How well do micro-economic factors explain obesity rates?

Assessing the influence of income and cost of diet on dietary intake and body mass index in a representative UK sample

Katherine Anne Timmins

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

The University of Leeds

School of Food Science and Nutrition

School of Medicine

February 2014

Declaration

The candidate confirms that the work submitted is her own, except where work which has formed part of jointly authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

This thesis features secondary analyses of established data sets. The candidate was not involved in survey design, data collection or primary data processing of these studies (the National Diet and Nutrition Survey (NDNS), the DANTE food cost database, Supermarket Nutrition Information Project (SNIP), and a UK Women's Cohort (UKWCS) sub-study). Credit for these data is detailed in the Acknowledgements. The candidate's contributions included data cleaning, linking data sets, data manipulation, analysis and interpretation.

Chapter 5 includes work which was featured in the jointly authored publication: Timmins, K., Morris, M., Edwards, K., Clarke, G. & Cade, J. 2013. Comparability of methods assigning monetary costs to diets: derivation from household till receipts versus cost database estimation from four-day food diaries. *EJCN*, 67(10):1072-1076. Michelle Morris and Katherine Timmins contributed equally in carrying out the research and analysis for this article, and in the manuscript preparation. The other co-authors provided supervisory guidance and contributed to the article drafts.

An article related to the work of Chapter 6 was also published as Timmins, K., Hulme, C. & Cade, J. 2013. The monetary value of diets consumed by British adults: an exploration into sociodemographic differences in individual-level diet costs. *Public Health Nutrition*, Oct 29:1-9 (epub). The candidate was solely responsible for carrying out the analysis and writing the first draft of this paper. J.E.C. and C.H., as PhD supervisors, provided guidance and commented on the article drafts.

Acknowledgements

It is fair to say that this thesis would not have emerged without the foresight of my supervisors, Professors Janet Cade and Claire Hulme, who secured the funding for this research topic, as well providing their unfailing guidance, helpful comments, shared wisdom and cheerful encouragement throughout. It has been a privilege.

I must also thank the Economic and Social Research Council and the Medical Research Council (ESRC/MRC) for funding this Joint Interdisciplinary studentship.

This research made extensive use of National Diet and Nutrition Survey (NDNS) data held by the UK Data Archive. I would like to acknowledge the survey creators, depositors and funders, including the National Centre for Social Research, the Northern Ireland Statistics and Research Agency, the Medical Research Council, University College London Medical School, the Food Standards Agency, the Department of Health and the UK Data Archive. The original data creators, depositors and copyright holders of the NDNS and the UK Data Archive bear no responsibility for their further analysis or interpretation. Crown copyright for the NDNS is held jointly with the National Centre for Social Research, and material is reproduced with the permission of the Controller of HMSO and the Queen's Printer for Scotland.

Acknowledgement for the creation of the DANTE food cost database is owed to Kevin Tarbutt and Edmund Parks. In addition, Chapter 5 draws on data from the following studies: the Supermarket Nutrition Information Project (SNIP), and a sub-study of the UK Women's Cohort Study (UKWCS). The work of Claire Oyston and Dr. Joan Ransley, who created the SNIP till receipt dataset, needs to be acknowledged, as does the contribution of Andrea Smyth in collecting till receipt data from the UKWCS.

Many people have provided guidance on various aspects of this research. Particular thanks are owed to my colleague and friend, Michelle Morris, who has been with me every step of the way and a pleasure to collaborate with. I am also grateful for the advice of my Assessors Drs Charlotte Evans and Sandy Tubeuf, and to have benefitted from the statistical expertise of Dr. Darren Greenwood. My thanks also go to the Nutrition Epidemiology Group, for some lively and thought-provoking discussions, and to colleagues in the Academic Unit of Health Economics.

Finally, never-ending thanks are owed to my parents, for always believing I could achieve anything; and to Jamie, for keeping me sane and so much more.

Abstract

Rates of obesity are predicted to increase, which is worrying given the association with adverse health outcomes. Cost of food or diet is one proposed contributor to an ‘obesogenic environment’. The “food price-obesity hypothesis” supposes that, with limited purchasing power, consumers may purchase energy-dense foods to obtain the maximum calories, resulting in excess energy intake.

This thesis attempts to gauge whether obesity may be attributed to food prices. Firstly, the published literature was synthesised. Secondly, the study examined how income and cost of diet are implicated in excess energy intake, as implied by the body mass index (BMI) and dietary energy density (DED), of adults in the National Diet and Nutrition Survey (NDNS).

The literature review revealed a heterogeneous body of studies that was generally supportive of the food price-obesity theory, but not conclusive. Studies of diet costs and DED overwhelmingly report a negative association. A limited number of studies investigating diet costs and BMI reported contradictory findings. The evidence linking income and DED was not strong.

In the NDNS sample, income was found to be negatively associated with DED, BMI, and overweight/obesity. In addition, a negative association was observed between diet costs and DED. There was no association between whole diet costs and BMI. In contrast, using proportional food group costs revealed some significant associations. This suggests that measuring how people apportion their food budget, rather than how much the whole diet is worth, may be insightful.

The thesis also addresses some methodological issues. Firstly, analyses demonstrate how equivalizing household income to take into account household composition can impact on findings. Secondly, a comparison of diet costing methods is presented.

Despite methodological challenges, the findings presented in the thesis suggest there is merit in pursuing research into diet costs, with many unexplored opportunities in this emerging field.

Table of Contents

List of tables

List of Figures

List of abbreviations

ACCRA	American Chamber of Commerce Research Association
AGHLS	Amsterdam Growth & Health Longitudinal Study
ANOVA	Analysis of variance
BA	Bland Altman
BMI	body mass index
BRFSS	Behavioral Risk Factor Surveillance System
C2ER	Council for Community & Economic Research
CAPI	computer-assisted personal interview
CARDIA	Coronary Artery Risk Development In Young Adults
CDC	Centers for Disease Control and Prevention
CDSR	Cochrane Database of Systematic Reviews
CI	confidence intervals
CRD	Centre for Reviews and Dissemination
CPI	Consumer Price Index
CSFII	Continuing Survey of Food Intake for Individuals
CSI	Coping Strategies Index
DAG	directed acyclic graph
DANTE	Diet and Nutrition Tool for Evaluation
DARE	Database of Abstracts of Reviews of Effects
DED	dietary energy density
DEFRA	Department for Environment, Food and Rural Affairs
DH	Department of Health
DHQ	diet history questionnaire
DINO	Diet In Nutrients Out
DLW	doubly-labelled water
DONALD	Dortmund Nutritional and Anthropometric Longitudinally Designed Study
EAR	Estimated Average Requirement
ECLS-K	Early Childhood Longitudinal Study – Kindergarten
ED	energy density
EFS	Expenditure & Food Survey
EI	energy intake
ESDS	Economic and Social Data Service
ESRC	Economic and Social Research Council
EU	European Union
F&V	fruit and vegetables

FAFH	food away from home
FAO	Food and Agriculture Organization
FFQ	food frequency questionnaire
FFPI	fast food price index
FPI	Food Price Index
FSA	Food Standards Agency
FVPI	fruit and vegetable price index
HBAI	Households Below Average Income
HDI	health development index (Ch3)
HDI	Healthy Diet Index (Ch 9)
HSCIC	Health and Social Care Information Centre
HSE	Health Survey for England
IMD	Index of multiple deprivation
INSEE	French National Institute of Statistics
IQR	interquartile range
LASA	Longitudinal Ageing Study Amsterdam
LIDNS	Low Income Diet and Nutrition Survey
MAR	mean adequacy ratio
MRC	Medical Research Council
MTF	Monitoring the Future survey
NASS	National Agricultural Statistics Service
NatCen	National Centre for Social Research
NDNS	National Diet and Nutrition Survey
NHANES	National Health and Nutrition Examination Surveys
NHS	National Health Service
NICE	National Institute for Health and Clinical Excellence
NISRA	Northern Ireland Statistics and Research Agency
NLSY97	National Longitudinal Survey of Youth
NOO	National Obesity Observatory
NS-SEC	National Statistics Socio-economic Classification
OECD	Organisation for Economic Co-operation and Development
OLS	ordinary least squares
ONS	Office for National Statistics
OR	odds ratio
PAF	Postcode Address File
PARADE	Partners of All Ages Reading About Diet and Exercise
PIR	poverty income ratio

PSID	Panel Study of Income Dynamics
PSU	primary sampling unit
PUFA	polyunsaturated fatty acid
RCT	randomized control trial
RMLS	Russia Longitudinal Monitoring Survey
RPI	Retail Price Index
SACN	Scientific Advisory Committee on Nutrition
SD	standard deviation
SE	standard error
SES	socioeconomic status
SHARE	Survey of Health, Ageing and Retirement Europe
SIMD	Scottish Index of Multiple Deprivation
SMD	standardized mean difference
SNIP	Supermarket Nutrition Information Project
SOS	Seattle Obesity Study
SUN	Suguimiento Universidad de Navarra
UCL	University College London
UKWCS	UK Women's Cohort Study
USDA	US Department of Agriculture
WCRF	World Cancer Research Fund
WFP	World Food Programme
WHO	World Health Organization
WISP	Weighed Intake Software Program

Chapter 1 Introduction

1.1 Foreword

Overweight and obesity have been recognised as the major health challenge of the 21st century (WHO, 2007, Butland et al., 2007). Defined as having a body mass index (BMI) of 25kg/m^2 and over or 30kg/m^2 and over, respectively, people classified as overweight or obese are at greater risk of a number of health problems, from cardiovascular disease and stroke to diabetes and osteoporosis (WHO, 2007). Results of a recent meta-analysis associate obesity in particular with higher all-cause mortality (Flegal et al., 2013). Given these adverse outcomes, predicted trends in the rates of overweight and obesity in many nations are worrying. Recent figures from England (HSCIC, 2013) indicate that the proportion of adults classified as overweight and obese has risen from 58% (males) and 49% (females) in 1993 to 65% and 58% respectively in 2011. If trends are to continue, it is predicted that by 2050 60% of British men, 50% of women and 25% of children will be classified as obese, with an estimated £9.7 billion in associated health costs (Butland et al., 2007). Slowing or even reversing such trends is undoubtedly in the interest of society.

1.2 The aetiology of obesity

In order to devise and implement effective interventions, it is first necessary to understand the aetiology of excess weight gain in the population. However, this has proved far from straightforward. At its simplest level, obesity can be explained as the result of positive energy balance, with an accumulation of excess energy. Whilst some authors emphasise the role of sedentarisation in Western society (Church et al., 2011), others propose that, in fact, average energy outputs have not changed appreciably in recent decades, and increased energy consumption is more likely to be the underlying problem (Scarborough et al., 2011).

However, to begin to understand the reasons for increased energy consumption, many researchers have emphasised the need to establish wider determinants. In other words, we need to identify the causes of positive energy balance, or even the causes of causes (Marmot and Bell, 2012). From this perspective, the factors contributing to obesity are acknowledged to be numerous and diverse – the Obesity System Map published in the 'Foresight report' (Butland et al., 2007) offers a well-recognised illustration of the complexity of the issue.

1.3 The obesogenic environment

Amongst the ‘causes of causes’ is the frequently cited ‘obesogenic environment’. The obesogenic environment refers to

“*the sum of influences that the surroundings, opportunities, or conditions of life have on promoting obesity in individuals or populations*” (p. 564, Swinburn et al., 1999).

The term encapsulates any influence on energy balance, including those which impact on physical activity. As one aspect of the obesogenic environment, the food environment refers to both the sources of food available and the factors that influence the purchase, preparation or consumption of that food (Holsten 2008). Cost of food or diet is one important factor contributing to the food environment.

Studying the proposed determinants of obesity in their entirety is beyond the scope of this thesis. Instead, the thesis will focus upon the role cost of food plays in the obesogenic environment.

1.4 The economics of obesity

This thesis aims to contribute to our understanding of the aetiology of obesity by investigating the potential role of micro-economic factors in food choice. ‘Micro-economics’ refers to the “branch of economics that studies individual units” (p.4, Sloman, 1999). This contrasts with the wider systems-focussed study of macro-economics, which is beyond the scope of this thesis. Instead, the ensuing chapters attempt to reconcile the overconsumption of energy by individuals with the aid of consumer choice theory. According to economic theory, such choices reflect decisions on how to allocate limited resources.

Dietary consumption is a collection of behaviours which are performed as a consequence of individuals’ decision-making. Faced with a number of options – what to eat, where, when and how much– behavioural economics suggests that people will choose combinations that best maximise their utility (‘utility’ in this sense refers to the “satisfaction a consumer gets from the consumption of all the units of a good consumed within a given time period” (p. 93, Sloman, 1999)). Maximum utility, however, is always constrained by scarcity: of time, of money, of social norms, preferences or health concerns. The ‘rational decision-maker’ of microeconomics is presented with a number of considerations that they must weigh up to arrive at the decision with the best utility.

The utility function of dietary consumption can be expressed as:

$$u_t = U(C_t, Z_t; W_t)$$

Equation 1.1

where an individual's utility over time (u_t) is a function of the consumption of both food (C_t) and other goods (Z_t), conditioned by their bodyweight (W_t) (Boizot-Szantai and Etile, 2005). This can be extended to observe that weight is a function of energy balance:

$$W_t = f(EI_t, EE_t)$$

Equation 1.2

where EE refers to energy expenditure and EI refers to energy intake over time. Furthermore, factors determining food consumption could be expressed as:

$$C_t = f(PP_t, FA_t, S_t, A_t, P_t)$$

Equation 1.3

where PP is purchasing power, FA refers to food availability, S is social influences, A is attitudes to food and P refers to taste preferences. Finally, purchasing power could be described as being a function of income (I), food prices (FP) and the consumption of other goods:

$$PP = f(FP, I, Z).$$

Equation 1.4

The amount of money available to an individual will influence their decisions on how to allocate that money to food. At the same time, the prices of foods will shape how that money can be allocated. Taken together, these aspects define the purchasing power of an individual.

The purpose of the above economic description is to illustrate the potential role of food prices in the obesogenic environment: they exert their influence via their part in determining purchasing power, which, along with availability and other environmental factors, will affect food consumption. If food prices over time encourage the consumption of certain foods which promote excess energy intake, they may be at least partly responsible for weight gain in the long-term. Increasingly, food prices and the global food system have been blamed for recent obesity trends (Drewnowski, 2007, Swinburn et al., 2011).

1.5 The "Food price-obesity" hypothesis

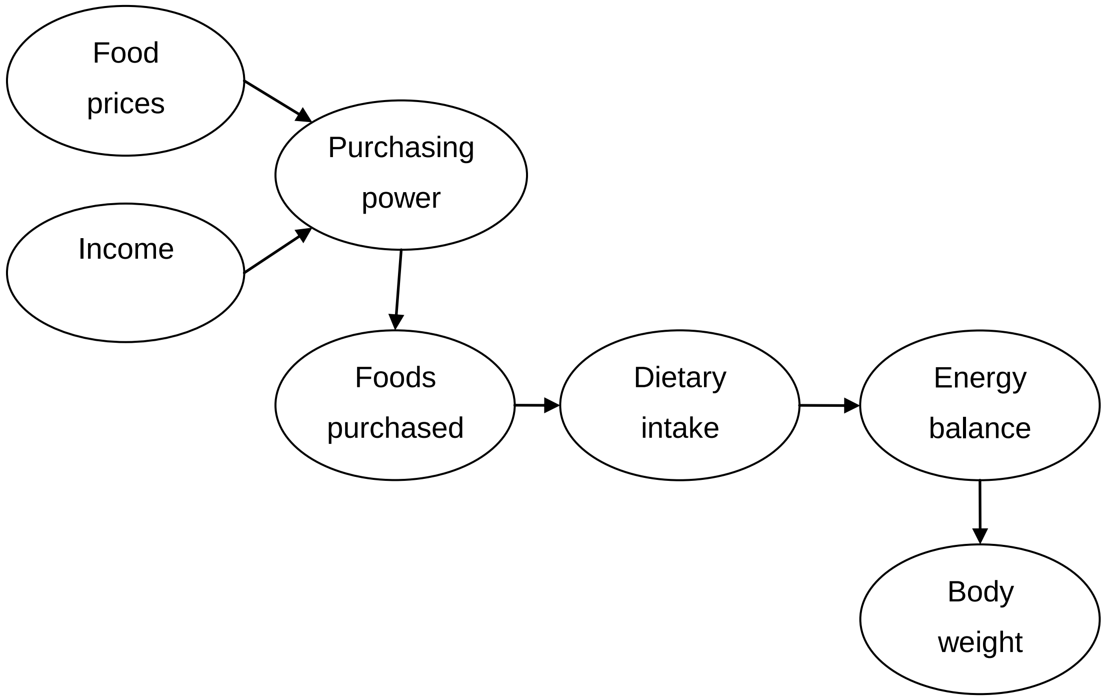
Affordability of nutrition has been a key topic for public health nutritionists for many years. Being unable to afford an adequate dietary intake clearly has repercussions in terms of malnutrition, both in terms of macro- and micro-nutrient intakes. In contrast, the hypothesis linking food prices and obesity implicates food prices in overconsumption, rather than inadequacy. This contrast implicitly recognises two categories of food insecurity: with hunger, or without hunger. The hypothesis argues that food insecurity *without* hunger could lead to an over-consumption of the cheaper calories found in more energy-dense foods (Adams et al., 2003, Dinour et al., 2007).

In recent years (2005-2010), the price of food has been acknowledged to have outpaced inflation considerably, across the globe (Brinkman et al., 2009). However, prior to this, the real price of food had been declining since the 1970s, in the United States at least (Drewnowski 2007; Cohen 2008). This trend in prices from the 1970s to the early 21st century reflects the increasing availability of foods, particularly following improvements in cultivation, storage and transport in the latter half of the 20th century (Cohen 2008; Drewnowski 2007). The change in availability appears to have been more marked for foods that require processing (Drewnowski 2007). These changes in manufacture and availability may be termed 'supply-side determinants' of food prices. Some researchers have found that price trends have been less favourable for energy-dilute foods, such as fruits and vegetables (Drewnowski and Darmon, 2005).

As described in Section 1.4, a decline in real food prices will have increased consumers' food purchasing power (assuming the prices and consumption of other goods, Z, remains stable). If the more energy-dense foods have become relatively cheaper, this has potentially favoured their purchase over that of less energy-dense foods, perhaps leading to an overconsumption of energy.

Figure 1.1 illustrates that there are several steps to the proposed causal pathway linking food prices and body weight. The food price-obesity hypothesis supposes that, with limited purchasing power, consumers will be encouraged to purchase more energy-dense foods in order to obtain as many calories as possible (maximising their utility). Meanwhile, dietary energy density has been linked to a tendency to positive energy balance in experimental settings (Prentice and Jebb, 2003) as well as to adiposity in observational studies (Perez-Escamilla et al., 2012). If people consume the energy-dense foods they purchase, they may be at risk of consuming excess energy and, in the absence of compensatory energy expenditure, there is a likelihood of weight gain, eventually leading to overweight and obesity over the long-term.

Figure 1.1 The proposed causal pathway between food prices and obesity



Food prices undoubtedly influence food purchases. The relationship between the price of a good and the quantity purchased of that good is exemplified in economists' 'demand curves' (Sloman, 1999), and food on the whole is no exception to this. There is a large breadth of economics literature which investigates the relationship between price and the quantity purchased for a wide variety of foods in many different regions, usually expressed as price elasticities (how 'elastic' a good is gives an indication of how much the quantity purchased responds to a change in price – a perfectly inelastic good will not see a change in demand following a price change, whereas an elastic good is responsive). Reviews of published food price elasticities (Andreyeva et al., 2010, Hawkes, 2009) indicate that food purchasing is influenced by food prices at the population level. At the individual level, too, experimental studies have observed food purchasing responsiveness to food price changes (see, for example, French, 2003, Block et al., 2010).

It should be noted that the above depiction offers a simplified account of the relationship between prices and purchasing. Whereas price elasticities imply a unidirectional relationship, it is acknowledged that the quantity of a good demanded (i.e. purchased) will in turn affect prices. This bidirectional influence makes the interpretation of food price studies complicated. For the purposes of explaining the economic aspect of obesity aetiology, a simplified unidirectional pathway was felt adequate to illustrate the reasoning behind the food price-obesity hypothesis. Nevertheless, it is an important issue to acknowledge, and the implications of this two-way relationship will be returned to in the Discussion (Chapter 9).

Whilst evidence abounds relating to the purchases of specific foods in relation to food prices, less is known about the overall effects on dietary intakes. Still, the food price-obesity theory has been gaining traction in recent years, with a growing number of researchers, practitioners and advocacy groups endorsing a fiscal approach to obesity prevention (Academy of Medical Royal Colleges, 2013, Sustain, 2013, Brownwell et al., 2009). There have been debates in several countries regarding this approach and targeted food or beverages taxes have already been implemented by some governments (see Chapter 9). However, the extent to which such proposals and policies reflect empirical evidence is unclear.

1.6 Obesity & inequality

If the above hypothesis holds true, this would have important implications in terms of health inequalities. Lower socioeconomic groups in the UK are recognised to suffer a greater incidence of poor health conditions, as well as an increased risk of all-cause mortality and decreased life expectancy (Acheson, 1998, Marmot and Bell, 2012). The role of diet in creating or reinforcing these socioeconomic differences in health has been asserted since at least the 1990s (James et al., 1997). If poor socioeconomic status causes unhealthy dietary choices, there are important ramifications for policy.

Darmon and Drewnowski (2008) describe a convincing body of literature investigating social class and diet in their non-systematic review. Socioeconomic disparities have been reported for the consumption of certain foods (such as whole grain) (James et al., 1997), the consumption of food groups (especially fruit and vegetables) (De Irala-Estevez et al., 2000), or in healthy diet scores (Kant and Graubard, 2007). In addition, data from at least one expenditure study suggests socioeconomic differences in food purchasing behaviour as well (Turrell and Kavanagh, 2005).

Much of the research in this area employs an aggregate index of socioeconomic position. These indices often incorporate proxy measures for economic status, social position, social environment, or social capital in an attempt to quantify the relative status of individuals or groups (Public Health England, 2013). Whilst useful in describing a phenomenon, aggregate indices may misrepresent the core causal relationship where separate aspects of the index are related to the health outcome in different ways (Benzeval et al., 2001, Macintyre et al., 2003, Carr-Hill and Chalmers-Dixon, 2005). Obesity is one outcome for which observed patterns of inequality differ according to the type of socioeconomic indicator chosen (Public Health England, 2013). Using data from the Health Survey for England (HSE), the National Obesity Observatory (NOO, 2010) found that, in women, lower socioeconomic status was associated with higher rates of obesity, but the pattern of obesity prevalence amongst men varied according to the socioeconomic indicator used.

Income is a component of socioeconomic status that could explain diet differences via an independent causal mechanism to other socioeconomic constituents. As explained above, income is an important contributor to purchasing power. As summarised by Marmot and Bell, “having insufficient money to lead a healthy life is a highly significant cause of health inequalities” (p.28, Marmot and Bell, 2012). According to the pathway illustrated in Figure 1.1, an increase in purchasing power will likely result in an increase in the purchase of food. This concept is encapsulated in Engel’s Law, which states that the quantity of food purchased will increase as income increases. Engel’s Law also stipulates that, whilst the absolute quantity of food rises, the proportion of income spent on food actually decreases as income increases (Zimmerman, 1932).

Engel’s Law has particular relevance when considering inequalities in access to a healthy diet: for instance, those on lower incomes, for whom food purchases take up a greater proportion of their income, will find it most difficult to adapt to food price increases. According to the FAO’s Coping Strategies Index (CARE/WFP, 2003), the typical first step for households facing food insecurity is to alter the diet by substituting cheaper foods, before compromising on quantity of energy intake. In other words, people turn to cheaper sources of calories (Drewnowski and Specter, 2004). This reiterates the concept of food insecurity without hunger introduced in Section 1.5. A discussion of income differences in dietary energy density and in obesity prevalence is presented in Chapter 4.

In summary, if purchasing power is as influential on dietary choices as is suggested above, then it is possible that income and food prices are driving some of the observed inequalities in health. Implicating these micro-economic factors in the obesity pathway could therefore offer support to public health policies which address the affordability of diets.

1.7 Thesis aim & objectives

This thesis will draw together the disciplines of nutrition, economics and epidemiology. The aim of the thesis is to determine the extent to which income and cost of diet are implicated in excess energy intake, as implied by current body mass index (BMI) and dietary energy density (DED) in a nationally representative sample.

The National Diet and Nutrition Survey (NDNS) was used to address this aim, as it contains investigator-measured anthropometric data as well as detailed dietary intake data. The outcome variables of DED and BMI were selected due to their key roles in the food price-obesity hypothesis. Income and cost of diet were chosen to represent the demand- and supply-side factors which help to shape purchasing power (Figure 1.1). Data on food price trends were not appropriate for the cross-sectional survey used, nor were data available on the food expenditure of NDNS participants, therefore the estimated cost of diets as consumed was used to denote the food price aspect of the hypothesis (see Chapter 5).

1.7.1 Objectives

To meet the primary research aim, the following objectives were identified:

1. To synthesise the published evidence linking food prices or diet costs with dietary energy density or weight status
2. To examine the relationship between income and BMI or overweight/obesity amongst NDNS adults
3. To assess whether income is related to DED amongst NDNS adults
4. To investigate the appropriateness of diet cost estimations, including the costing of food groups
5. To estimate and describe the diet costs of NDNS adults
6. To explore patterns in NDNS diet costs according to sociodemographic characteristics
7. To determine whether an association exists between diet costs and BMI or overweight/obesity amongst NDNS adults
8. To establish whether an association exists between diet costs and DED amongst NDNS adults
9. To discuss how evidence from the NDNS fits in with the food price-obesity hypothesis

The findings of this research could help to elucidate the micro-economic aspects of obesity aetiology, which in turn could guide public health interventions. Explorations in the socio-demographic differences in cost-related diet patterns may also contribute to the literature on inequalities in health, potentially identifying populations who may be at risk of adverse dietary changes in the face of future food prices.

1.8 Thesis structure

The thesis is organised into nine chapters. Table 1.1 outlines each chapter, and indicates how each relates to the objectives described above. A brief description of the content of each chapter is presented below.

Table 1.1 The thesis structure

	Chapter	Objective(s) to be met
Chapter 2:	Literature review	1
Chapter 3:	NDNS sample description	
Chapter 4:	Income, diet and BMI in the NDNS	2, 3
Chapter 5:	The DANTE food cost database	4
Chapter 6:	Estimating the diet costs of NDNS adults	5, 6
Chapter 7:	Diet costs, diet and BMI in the NDNS	7, 8
Chapter 8:	Food group costs & BMI in the NDNS	4, 7, 8
Chapter 9:	Conclusions	9

Chapter 2: Literature review

This chapter presents the results of a systematic search of the literature with a narrative synthesis of published findings from studies investigating the role of income, food prices or cost of diet in encouraging excess energy intake.

The review is organised in two sections, to reflect the two indicators of excess energy intake that form the focus of this thesis: firstly, dietary energy density and, secondly, body weight or mass. For each of these outcomes, literature will be considered in which the impact of the following three factors are investigated:

1. Food prices;
2. Dietary expenditure or diet cost; and
3. Income.

The synthesis identifies important gaps in knowledge and the methodological challenges faced by researchers in this area, to set the context for the analyses of later chapters.

Chapter 3: Sample description

The main analyses of this thesis use data from the National Diet and Nutrition Survey (NDNS). The NDNS is a national dietary assessment survey, designed to represent the general UK population. This chapter introduces the NDNS: its purpose and design, sampling techniques and data collection protocol. In addition, the chapter presents a description of the analytical sample, outlining some of the chief characteristics.

Chapter 4: Income in the NDNS

This chapter introduces the first empirical analyses of this thesis. In considering the micro-economic determinants of obesity, the primary focus of the chapter is on income, as an important factor in purchasing power. The methods used to measure income in the NDNS are outlined, and descriptive statistics are presented to show the income distribution in the sample. The chapter explores the relationship between income and diet – specifically energy density (kJ/g) – and the relationship between income and body mass index (BMI) amongst adults in the NDNS. In addition, the concept of equivalization is introduced – a variable seldom employed in nutrition epidemiology. A discussion around the suitability of equivalized versus household income will be incorporated, using results of analyses to illustrate the impact.

Chapter 5: The DANTE food cost database

A key supply-side determinant of food purchasing is the price of food. Direct data regarding the food prices encountered by NDNS participants is not available, however. Therefore, an estimation of the monetary cost of diets using an in-house database of national food prices will be used as a proxy for food prices. This chapter will describe the tool used, the DANTE (Diet and Nutrition Tool for Evaluation) food cost database, and how it is applied to estimate costs from diet records.

Despite the widespread employment of food price databases in diet cost research, no attempts at validating the method have been reported in the literature. The chapter will also present the results from a reanalysis of two previously conducted (unpublished) studies at the University of Leeds in which food purchase receipts and diet diary records were concurrently collected, allowing two methods of diet cost estimation to be compared.

Chapter 6: Estimating the diet costs of NDNS adults

Dietary costs have not previously been estimated for the NDNS sample. This chapter will describe the methods used to derive costs from the dietary data of the sample using the DANTE cost database, and then present descriptive results. In addition, comparisons between sociodemographic groups and by lifestyle variables will be explored.

Chapter 7: Diet costs, diet and BMI in the NDNS

If food prices influence dietary intake and energy balance, it may be the case that the inherent monetary value of diets is associated with dietary energy density or the body weight of people consuming those diets. This chapter takes the estimated diet costs presented in Chapter 6, and investigates how they relate to dietary energy density and body mass in the NDNS. Body mass will be considered both as a continuous variable (BMI) and, due to the clinical appropriateness of categories, as overweight/obesity incidence in a logistic regression.

Chapter 8: Food group costs & BMI in the NDNS

As an emerging research area, the best available method for investigating monetary aspects of diet is yet to be established. Whole diet costs are strongly related to energy intake, whereas energy-adjusted diet costs are closely associated with dietary energy density. As such, it can be problematic to disentangle the influence of energy intake or energy density in analyses using either construct. This chapter sets out a fresh approach to quantifying diet costs by examining the proportions of whole diet cost attributed to constituent food groups. Proportional costs give an indication of how people apportion their budget, as well as how these proportions change as budgets vary.

Methods and descriptive results will introduce the concept of constituent food group costs. In order to characterise these new variables, analyses will be included exploring the relationships within food group costs, between food group costs and whole diet costs, and in relation to proportional energy intake by food group. The chapter will then go on to examine the relationship between food group costs and BMI in regression analyses, and discuss if the new approach adds value to a traditional whole diet cost approach.

Chapter 9: Discussion & conclusion

The final chapter of the thesis will draw together the findings from the previous chapters, relating them to each other, and discussing how they fit with the food price-obesity hypothesis. Results of previous chapters will be interpreted collectively to develop overall conclusions. The implications for public health research and policy will be identified, and recommendations for future research suggested.

Chapter 2 Literature review

2.1 Summary

This chapter presents the results of a systematic search of the literature with a narrative synthesis of published findings from studies investigating the role of income, food prices or cost of diet in encouraging excess energy intake. There were six key relationships under investigation. The literature search was carried out on several databases in two separate phases (2011 and 2013), using a pre-established protocol.

A total of 44 articles were found to fit the inclusion criteria and were included in the review:

- No studies identified investigated dietary energy density in relation to food prices
- Nine studies investigated diet costs and dietary energy density
- Five studies investigated income and dietary energy density
- Twenty four articles investigated food prices and body weight (13 using adult samples, 10 focussing on children or adolescents; and one which used data from both)
- Seven studies were found to investigate body weight in relation to diet costs or expenditure
- A scoping search revealed four reviews regarding income and BMI or obesity (three of these studies were systematic), and 13 reviews related to income and energy density, or diet costs and energy density or BMI.

The findings relating to dietary energy density are largely in keeping with the prevailing hypothesis that economic factors influence the selection of energy-dense foods. The overall conclusion of this review is that the evidence – amongst adults, but not children – linking income or diet costs with dietary energy density is supportive of the theory. However, the review has identified that certain methodological issues limit our confidence in these results.

Heterogeneity amongst the literature makes it difficult to draw conclusions regarding micro-economic determinants of body weight. There are interesting results reported for many of the studies, reinforcing that this topic is a worthwhile area of investigation, but findings are largely mixed. Some results suggest that various subgroups – males or females, the near poor, or those with children – may elicit differing findings.

This synthesis of the literature helps to identify important gaps in knowledge and methodological challenges faced by researchers in this area. This sets the context for the analyses of later chapters.

2.2 Introduction

The overarching aim of this thesis is to determine the extent to which income and cost of diet are implicated in excess energy intake. Before examining data, however, it is first necessary to consider existing evidence around income, diet cost, diet and body weight.

This research area spans several different academic fields: nutrition, epidemiology, economics, politics, marketing, psychology and social geography, to name but a few. As such, there is a need to bring evidence together from these different disciplines. An interdisciplinary stance should best help to build a comprehensive picture of how micro-economic factors impact on diet and weight.

Synthesising the existing evidence was the first objective of the thesis identified in Chapter 1. This chapter presents the results of a systematic search of the literature with a narrative synthesis of published findings relevant to the research questions. Due to the cross-disciplinary nature of the research question, a broad variety of investigative approaches was anticipated, and, as such, a narrative synthesis was planned as opposed to a meta-analysis.

The results are organised in two sections, to reflect the two indicators of excess energy intake that form the focus of this thesis: firstly, dietary energy density and, secondly, body weight or mass. For each of these outcomes, literature will be considered in which the impact of the following three factors are investigated:

1. Food prices;
2. Dietary expenditure or diet cost; and
3. Income.

Of these three factors, income is expected to be the most widely researched, due to its acknowledged contribution to health inequalities (McDowell et al., 1997) and its frequent inclusion in socioeconomic indicators. Food prices and diet costs, on the other hand, form a more recent area of academic interest. Therefore, the review of literature around food prices and diet costs will be conducted in a systematic manner, to provide a comprehensive summary of the relevant literature. Literature investigating income and dietary energy density will similarly be searched and synthesised systematically. However, a comprehensive review of the literature surrounding income and body weight was judged beyond the scope of this chapter, both due to the extent and breadth of existing publications, and due to the identification of existing systematic reviews in a scoping search.

To summarise, the objectives of this review chapter are to synthesise any published evidence of associations between the following:

- Food prices and dietary energy density;
- Dietary expenditure or estimated diet cost and dietary energy density;
- Income and dietary energy density;
- Food prices and body weight or fatness;
- Dietary expenditure or estimated diet cost and body weight or fatness;
- Income and body weight or fatness.

The findings will help to identify important gaps in knowledge surrounding these relationships, and set the context for the analyses of later chapters.

2.3 Methods

The work for this literature review was conducted in two separate phases: the initial phase which focussed on food prices and dietary expenditure or cost, with a search conducted in January 2011; and a second phase, in which the initial search was updated and the literature on income was searched, in 2013. The search strategies and criteria were therefore separate for these two phases.

Before conducting the searches, a pre-established protocol was developed, in accordance with the Centre for Reviews and Dissemination (CRD, 2008). The protocol detailed the criteria, search strategy, literature sources and methods for data extraction and synthesis. Searching the literature entailed: firstly, identifying existing reviews; secondly, searching selected databases; and thirdly, citation searching.

Reviews were identified from the following catalogues: Cochrane Database of Systematic Reviews (CDSR); Database of Abstracts of Reviews of Effects (DARE); Evidence for Policy and Practice Information and Co-ordinating Centre (EPPICentre); NHS Economic Evaluation Database (NHS EED); and the National Institute for Health and Clinical Excellence (NICE). Reviews were also identified from the databases using the search strategy below.

The literature search was carried out in the following databases: CAB Abstracts, EMBASE, Food Science and Technology Abstracts, HMIC Health Management Information Consortium, Ovid MEDLINE and PsycINFO. The search strategies were developed for the Ovid MEDLINE interface, and adapted to suit other databases where necessary. The MEDLINE strategies can be found in Appendix A.

2.3.1 Criteria for inclusion

Criteria for inclusion and exclusion were pre-specified. However, during the second phase of the literature review work, in 2013, criteria were tightened, and articles found in the initial search were re-screened to reflect the new focus. This was due to the publication of a relevant systematic review in the interim (Lee et al., 2011). The main change was that dietary energy density was specified as the dietary outcome, rather than including all dietary outcomes.

In addition, it was decided during the second phase to exclude simulation studies. Again, this was due to the publication of a relevant systematic review since 2011 (Eyles et al., 2012), but the decision also reflected the fact that many of the simulation studies which predict anticipated effects of price changes – for example, as a consequence of taxation – utilise elasticities derived from purchasing data to model the effects on diet and health. This review was primarily concerned with dietary consumption, which may not be captured by purchasing data (see Section 2.3.1.1 below).

The search was not limited by date or country of origin. However, due to the resources available, it was judged pragmatic to include only English language articles. The decision was taken to also exclude grey literature (unpublished articles, theses and dissertations, non-peer reviewed articles), and include only those papers that reported findings from original research.

2.3.1.1 Literature on food prices or dietary expenditure/cost

For the purposes of this review, 'consumption' was taken to mean dietary intake, and not consumption in the traditional economic sense of purchasing. This was largely because the purchasing of food does not necessarily equate to the dietary consumption of that food (see, for example, Defra, 2010), and there is therefore additional potential for measurement error. (More information about the evidence on food prices and purchasing can be found in two systematic reviews of food price elasticities: Andreyeva et al. (2010) and Green et al. (2013)).

Other pre-specified criteria were:

Population

- Humans
- Healthy or non-diseased populations with risk factors
- Adults, and/or children, and/or adolescents, and/or elderly
- Males and females
- Any socio-economic grouping
- Populations not in a state of emergency or crisis, such as drought or other environmental disaster

Exposure

- Observations or manipulations in food prices (including beverages); food group prices; fast food prices or fruit & vegetable prices
- Observations or manipulations in calorie cost; fat cost; energy cost; or other macronutrient cost, derived from foods and/or beverages
- Observations or manipulations in food and/or beverage expenditure; or dietary expenditure – whole diet, or specific foods, beverages or food groups, but not relating to special diets for medical reasons
- Observations or manipulations of any of the above in the context of regional fiscal or taxation policy
- Observations or manipulations in food or beverage promotions, defined as the act of encouraging a sale by means of financial incentive such as price discounting, quantity discounting, or extra-product price promotions, and not promotion via non-financial means such as advertising

- Observations or interventions which do not relate to emergency relief or aid
- Data no earlier than 1900
- Studies not restricted to alcohol beverages in isolation

Outcomes

- Estimated energy density of diet
- Weight status as measured by body mass index (BMI – kg/m^2) or other markers of body composition; weight change

Study design

- RCTs and other intervention trials
- Cohort studies
- Case-control studies
- Cross-sectional studies

2.3.1.2 Literature on income

The 2013 phase of the literature review additionally incorporated a new search for income-related studies. The inclusion/exclusion criteria for this search were identical with regards to population and study design to the previous search (see Section 2.3.1.1). The criteria for exposure or outcome were as follows:

- Estimated energy density of diet
- Income – household, family or individual; annual or otherwise; equivalized or not; gross or net; but not a composite of socioeconomic status

The search relating to income and body weight was restricted to published reviews.

2.3.2 Study selection procedure

Citations and abstracts of all hits elicited by the above searches were exported to EndNote X5 (EndNote X4 during phase 1) and de-duplicated. Abstracts were then screened in EndNote.

Selection of relevant literature followed a two-step procedure: firstly, a screening of titles and abstracts; and secondly, examination of full-text articles against the checklist of inclusion and exclusion criteria.

During the 2011 phase of the literature review, a 10% random sample of abstracts was also screened by a second reviewer, PhD supervisor Claire Hulme. Comparison of screening results revealed 86% agreement between the reviewers. Following discussion and clarification of criteria, agreement was 96%, with the remaining 4% requiring full-text examination to resolve discrepancy. The 2013 phase of the search involved only the first reviewer.

Full-text screening was performed by the first reviewer only, with queries referred to the second reviewer. Reason for exclusion at this point was recorded, using codes denoting specific exclusion criteria.

2.3.3 Data extraction

All included papers were sorted according to exposure (diet cost/expenditure, food price, income) and outcome (dietary energy density, body weight, or combination of both). Extraction forms were developed using Microsoft Access, and relevant data extracted. The following data were extracted for all included articles: bibliographic details; country; sample size and main characteristics; year(s) of data collection; length of and loss to follow-up, where appropriate; exposure definition and measurement; outcome definition and measurement; statistical treatment, including comparison groups and subgroup analyses; results and p values. Extracted data were organised into tables in Microsoft Word, and are presented alongside a narrative synthesis of the findings.

2.3.4 Data analysis/synthesis

Due to the heterogeneity of study designs and broad range of disciplines anticipated in the literature, it was decided a priori that a meta-analysis of results would be inappropriate. The narrative synthesis instead seeks to organise the findings of the studies in such a way as to describe patterns – for example, the existence, direction or size of an effect – and attempt to uncover explanatory factors for such patterns, if any. Recommendations published by the Economic and Social Research Council (ESRC) Methods Programme (Popay et al., 2006) were followed.

2.3.5 Quality appraisal

Given the anticipated heterogeneity of studies in the literature review, it was judged inappropriate to apply a quality checklist to included studies. It is still important, however, to assess the strength of the evidence given the quality of the literature found. Therefore, efforts will be made in the synthesis of results to appraise each study in terms of the potential for bias brought about by the study design. Following the guidelines published by the Centre for Reviews and Dissemination (CRD, 2008), the quality assessment will consider: appropriateness of design, the reliability and validity of the chosen outcome measure, risk of bias brought about through sampling, statistical issues (including power), the quality of reporting and generalisability.

2.4 Results

2.4.1 Search results

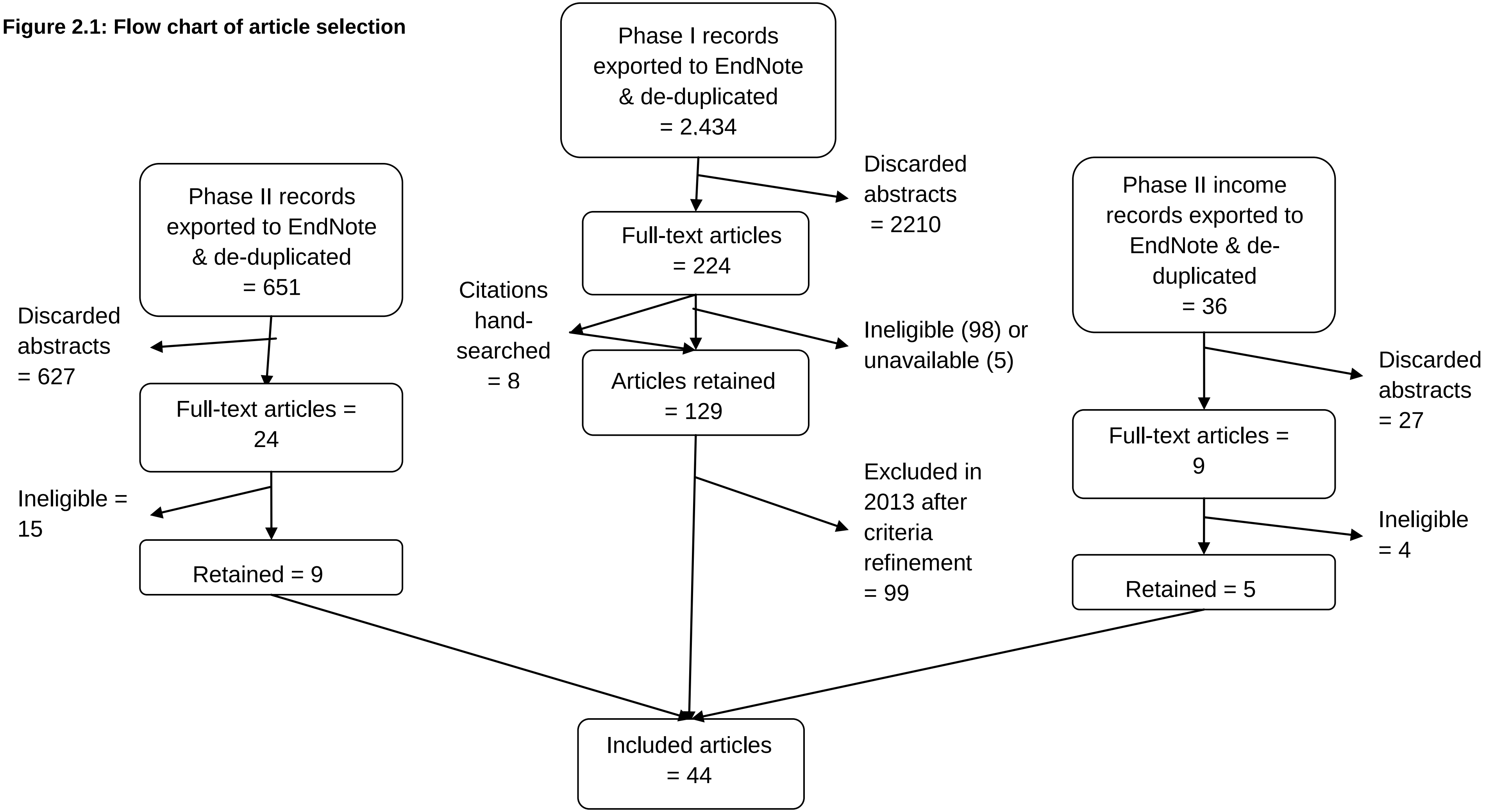
During phase 1 of the literature search, in 2011, 2,434 references were returned. Of these, 219 full-text articles were obtained and screened, from which 121 relevant articles were identified. Five articles were not obtainable: one was historical (published in 1947); one article was from a volume missing at the British Library; and the remainder were from geographically local publications which were not listed in the catalogue of the British Library. Additionally, seven review articles were downloaded and hand-searched for citations. From these, a further eight studies were identified, to bring the total of relevant articles to 129.

Following the adjustment to the inclusion criteria in 2013 (phase 2), a further 99 articles were excluded. Reasons for exclusion were: studies related to purchasing rather than dietary intake (n=63); studies were counterfactual and employed predictive, hypothetical models (n=9); or because studies investigated dietary intake or quality but not energy density (n=27).

The updated search in 2013 revealed a further 651 references returned by the phase 1 search strategy, that had been published between 2011 and 2013. In the abstract screening, 615 were identified as not fitting the criteria, and 12 were identified as reviews. Twenty four full-text articles were obtained, of which 9 met the criteria and were included in the review.

The Phase 2 search of income literature found 36 records, of which nine full-text versions were obtained. Of these, five were found to fit the criteria. This resulted in a total of 44 articles (excluding reviews) to be reviewed (Figure 2.1).

Figure 2.1: Flow chart of article selection



2.4.2 Existing reviews – food prices or diet costs

In addition to the 44 original articles, 13 systematic reviews were identified which included aspects of diet cost or food prices in relation to diet or overweight and obesity. Of these, six specifically focussed on economic aspects, whilst the focus of the remainder was on wider environmental exposures which incorporated food prices, diet cost or expenditure as just one aspect of the environment – at the national or local level (Harnack et al., Woodman et al., 2008, Holsten, 2008, Jaime and Lock, 2009, Steyn et al., 2009, Wilde et al., 2012, Conklin et al., 2013). This section will describe the reviews which share a similar research question to the current review, and not those concerned with wider environmental exposures. One of the six relevant reviews (An, 2012) was excluded during the second phase of this review, because the outcome of interest related to dietary consumption, and none of the included studies reported dietary energy density (reporting instead on other aspects of diet). Another review was excluded because it was not systematic (Goodman and Anise, 2006).

All of the relevant reviews were published in the past four years. Two were limited to US-based evidence (Powell and Chaloupka, 2009, Powell et al., 2013); two included all developed countries (Lee et al., 2011, Black et al., 2012). Only one of the reviews examined diet costs, as well as food prices (Lee et al., 2011); all others focussed on food prices only. The paper by Lee et al. (2011) was also the only review to include dietary energy density as an outcome.

Powell & Chaloupka, in 2009, reviewed US-based study literature to summarise the evidence surrounding food prices and BMI or obesity. Their synthesis indicated that the majority of studies reported statistically significant associations between food prices and BMI or the prevalence of obesity, and these were negative in direction for 'unhealthy foods' (energy-dense foods, fast food prices, sugar, whole milk) and positive for fruit and vegetables. Not all studies reviewed found statistically significant results for all exposures and outcomes; the reviewers suggest this may be due to differential elasticities for weight at different ends of the BMI distribution, as reported in three of the included studies. One study reviewed did not find any statistically significant results (Kim and Kawachi, 2006), but was described in the review as "weak statistical evidence", with a p value of 0.09. This study was the only one to examine state-level taxes as the exposure. The review identified important limitations in the literature reviewed, including comments on: inappropriate/unfeasible adjustment (for example, for income, or food outlet availability); a lack of longitudinal studies (only two in the review were not cross-sectional); the use of older data; and limited availability of price data (six of the nine reviewed studies employed a small database of non-representative food prices). The reviewers concluded that, whilst associations between food prices and BMI or obesity do exist, the effects are small, and fiscal interventions would therefore need to be non-trivial to produce measurable effects.

The aim of the systematic review by Lee et al. (2011) was to examine the effect of food costs on diet quality and disease risk. As such, the eligible literature encompassed evidence of a varied nature, including those studies which compared prices of healthy and unhealthy foods, those which considered whether a healthy diet was affordable, as well as those linking prices or costs to diets as consumed and to risk factors. Forty one articles were included in the review. Of these, 24 included dietary intake as an outcome and seven reported BMI or body weight as an outcome. The reviewers did not differentiate between food price data and diet cost (as estimated by applying food price data to dietary intake), and did not separate studies using these different approaches in their synthesis. The seven weight-related studies consisted of three studies employing food prices and four studies examining diet cost. Five of the studies found evidence of a negative relationship between food prices/diet cost and BMI or weight. The reviewers explained the null or contradictory findings of the remaining two studies to be due to methodological flaws, and concluded that the evidence of a relationship outweighed the evidence against.

Of the dietary intake articles included in the review by Lee et al. (2011), 11 investigated dietary energy density. Two of these studies reported findings from modelling studies. All of the studies indicated a negative relationship between costs and dietary energy density. The reviewers pointed out that the majority of these studies (n=7), examined energy cost as the exposure, which is methodologically problematic when linked to an outcome of energy density because of the creation of mathematical coupling (where energy is both the denominator in the exposure and the numerator in the outcome – see Lipsky (2009) for a discussion).

As in the previous review, Lee et al. (2011) highlight common flaws in the existing evidence, including: a majority of cross-sectional data; the validity of assumptions applied to food price data; and that many studies (all but two of the studies reporting dietary energy density) neglected to control for socio-economic variables such as education or income.

Black et al. (2012) focussed their review on subsidy programs amongst disadvantaged or low-income pregnant women and their children. Fourteen studies, published between 1980 and 2010 were identified, four of which reported maternal anthropometry. Many of the studies included reported outcomes relating to maternal dietary intake; however, none included dietary quality or dietary energy density per se. The reviewers found that evidence of an effect of food subsidies on maternal weight was inconclusive, although some positive significant effects were identified regarding fruit and vegetable intakes.

The final relevant review identified was that of Powell et al. (2013). This review included more recent evidence (published between 2007 and March 2012) of relationships between food prices, food consumption and body weight in the US. The search was limited to food prices of sugar-sweetened beverages, fast food or fruit and vegetables. Twenty studies were identified which related to BMI or weight, and 21 related to dietary consumption. The reviewers did not differentiate between consumption in the economic sense (i.e. purchasing) as opposed to the nutritional sense (dietary intake). Of the 21 consumption studies reviewed, 14 related to dietary intake, but none examined dietary energy density per se. The reviewers concluded that the published evidence suggested inverse relationships between food prices and food consumption; however, as already stated, this took into account both purchasing and intake studies. Examining the findings of intake studies alone (an approach not reported in the review), implies less strong evidence of a relationship, with significant negative relationships reported in just one of three studies of sugar-sweetened beverage prices, three of six studies of fast food prices, and three of five studies of fruit and vegetables prices.

The findings of studies investigating food prices and BMI or weight were mixed, and conflicting findings may have resulted from differing populations: there is the suggestion in the review that evidence differed depending on whether adults or children were studied, and that there were differential effects for low-income and higher income participants. The reviewers concluded that:

- Evidence of an effect of sweetened beverage taxes on weight outcomes was inconsistent, although one study found a significant association between beverage prices and children’s weight.
- There was fairly consistent evidence of a negative association between fast food prices and body weight, particularly amongst adolescents. Evidence was stronger for low- to middle-income participants.
- Findings linking fruit and vegetable prices to adult weight were mixed overall, but there was evidence of a positive association for women, and particularly those on low incomes.
- Amongst children and adolescents, all but four studies (out of 11), found significant evidence of a positive relationship of fruit and vegetable prices with body weight.

2.4.3 Dietary energy density

2.4.3.1 Food prices & DED

There were no studies identified which investigated dietary energy density in relation to food prices.

2.4.3.2 Diet costs & DED

Studies' designs and settings

Nine studies were found which investigated diet costs and dietary energy density, all published within the last 10 years (see Table). These were all cross-sectional. All studies were set in developed countries: three in France (Drewnowski et al., 2007, Maillot et al., 2007b, Andrieu et al., 2006), three in the US (Monsivais and Drewnowski, 2009, Townsend et al., 2009, Aggarwal et al., 2011) one in Japan (Murakami et al., 2007), one in the Netherlands (Waterlander et al., 2010) and one in Germany (Alexy et al., 2012). No studies from the UK fit the criteria.

Three studies used data from various nationally representative surveys of adults (Drewnowski et al., 2007, Maillot et al., 2007, Andrieu et al., 2006), one included data from two cohorts, one of which was comprised of elderly participants (Waterlander et al. 2010), and one made use of children and adolescent data (Alexy et al., 2012). Other studies used non-representative samples, often drawn from the research institution (for example, Murakami et al., 2007, Monsivais and Drewnowski, 2009). Townsend et al (2009) studied low-income women. Two of the studies used female-only samples (Townsend et al., 2009, Murakami et al., 2007). Sample sizes ranged from 112 (Townsend et al., 2009) to almost 4000 (Murakami et al., 2007).

Diet cost definition

All studies mapped food price information onto dietary intake data (see below). The majority of studies (n=8) expressed diet costs in relation to a standardized energy amount (2000kcal, 100kcal, 10MJ or 1000kJ). One study (Maillot et al., 2007) additionally reported daily costs. Of the two studies which did not use energy costs, Alexy et al (2012) used daily diet costs, as well as proportional food group costs, whilst Aggarwal et al (2011) employed the residuals of daily costs against energy intake in their analyses, in order to account for energy.

Price data

Several of the studies drew food price data from national statistics databases (Drewnowski et al., 2007, Maillot et al., 2007, Murakami et al., 2007), sometimes supplemented by additional sources. Many of the studies collected prices from local or national supermarket chains (Monsivais and Drewnowski, 2009, Waterlander et al., 2010, Alexy et al., 2012, Aggarwal et al., 2011) or local markets (Townsend et al., 2009). One study (Andrieu et al., 2006) obtained prices from a marketing research agency. The studies varied in the number of food items used to apply costs to dietary data: from 122 to 384 (where reported).

Alexy et al (2012) were the only investigators to examine food-group specific costs.

Assessment of diet

The three studies from France used 7-day records to assess dietary intake (Drewnowski et al., 2007, Maillot et al., 2007, Andrieu et al., 2006). The study by Murakami et al (2007) matched prices to diet history questionnaires, whilst that of Alexy et al (2012) used 3-day weighed records. The Dutch study (Waterlander et al., 2010) combined data from two cohorts, each of which used different diet assessment methods: one used interview to obtain intake information for the previous four-week period, whilst the second cohort used 24-hour recall. The remaining studies used food frequency questionnaires (FFQs) to assess diet.

Analytical approaches

The majority of studies (n=7) used regression techniques to analyse their data (see Table). These included least squares regression models and, in one study (Alexy et al., 2012), linear mixed effects models. ANOVA tests were also commonly employed. A few studies (Monsivais and Drewnowski, 2009, Waterlander et al. 2010) also reported correlation coefficients. All but three studies (Andrieu et al., 2006, Murakami et al., 2007, Aggarwal et al., 2011) used a combination of analytical techniques.

Analyses variously identified diet cost as the exposure variable and DED as the outcome (Andrieu et al., 2006, Murakami et al., 2007, Monsivais and Drewnowski, 2009, Townsend et al., 2009, Waterlander et al., 2010, Alexy et al., 2012, Aggarwal et al., 2011), or with DED as the exposure and diet cost as the outcome (Drewnowski et al., 2007, Maillot et al., 2007, Monsivais and Drewnowski, 2009, Waterlander et al., 2010, Alexy et al., 2012), and in some cases, analyses were included for both scenarios (Monsivais and Drewnowski, 2009, Waterlander et al., 2010, Alexy et al., 2012). Where diet cost was a predictor, five studies included the variable as categories (tertiles (Monsivais and Drewnowski, 2009, Townsend et al., 2009), quartiles (Andrieu et al., 2006) or quintiles (Murakami et al., 2007, Aggarwal et al., 2011)), and two included continuous cost variables (Townsend et al., 2009, Alexy et al., 2012). Of those studies in which diet cost was the outcome, this was always included as a continuous variable.

Models were, in most cases, adjusted for common covariates including age, sex and energy intake. Townsend et al (2009) adjusted for energy intake only. The analysis of Aggarwal et al (2011) included the most fully adjusted of the regression models, adjusting for ethnicity and household size in addition to the common covariates already listed. Alexy et al (20120) additionally incorporated interaction terms and non-linear terms in their mixed effects models. The analyses of Murakami et al (2007), Monsivais and Drewnowski (2009) and Waterlander et al (2010) were either unadjusted, or adjustments were not reported.

The multiple regression models of Drewnowski et al (2007) did not report individual coefficients, but only the p values and coefficient of determination.

Quality of studies

All of the studies were cross-sectional in design, and are therefore similarly at risk of the bias commonly associated with observational studies. Differences between the studies in terms of sampling, data collection and analysis, however, may have introduced different sources of bias.

The studies which relied upon nationally representative samples (Andrieu et al., 2006, Drewnowski et al., 2007, Maillot et al., 2007) are likely to have benefitted from these surveys' sampling designs which take into account selection bias when recruiting participants. In addition, findings from these samples are more likely to be generalisable (at least at the national level). Other studies – Murakami et al. (2007), Monsivais and Drewnowski (2009), Townsend et al. (2009) – may be considered weaker in quality in having to rely upon non-probability samples.

Study quality also differed in the reliability and validity of data collection methods. In terms of price data collection, established national economic surveys (as utilised by Andrieu et al., 2006, Drewnowski et al., 2007, Maillot et al., 2007 and Murakami et al., 2007) are likely to have developed price collection methods that attempt to minimise bias and are more likely to reflect the national distribution of prices. Conversely, collecting price information from a limited source (as performed by Monsivais and Drewnowski, 2009, Townsend et al., 2009, Waterlander et al., 2010, Aggarwal et al., 2011, and Alexy et al., 2012) risks introducing bias, if the participants in the study could purchase from a wider range of sources.

Dietary assessment methods have been widely investigated in terms of bias. All of the methods used in these studies relied upon self-reported dietary intake, which is subject to biases in reporting – most commonly under-reporting. All of the studies used established dietary assessment techniques (diaries, 24-hour recall, FFQs, DHQ). Little is known about the differences between these methods in terms of the further possible bias introduced when matched to food price data; however it could be conjectured that methods which quantify food consumption would be more appropriate for estimating diet costs. FFQs, which would necessitate the selection of representative foods for each item, risk introducing an additional level of bias with the assumptions inherent in these calculations. The studies of Monsivais and Drewnowski (2009) and Townsend et al. (2009) used FFQs and could be considered weaker in quality than those which, in particular, collected seven days of diet records (Andrieu et al., 2006, Drewnowski et al., 2007, Maillot et al., 2007).

In this area of research, perhaps the most important determinant of study quality is in the analytical approach. It has been remarked upon in the literature that including energy as both a numerator and denominator in the independent and dependent variables will result in potentially false positive findings, because the variables will be mathematically related and therefore automatically associated (Lipsky, 2009). Only three of the nine included studies attempted to address this: Aggarwal et al. (2001), Maillot et al, (2007) and Alexy et al. (2012), although Maillot et al. (2007) did not report the results of this analysis.

Findings

Of the analyses in which p values were reported, all but one test revealed a significant (in most cases, highly significant) negative relationship. Findings included significant differences in diet cost by categories of energy density (Drewnowski et al., 2007, Monsivais and Drewnowski, 2009, Maillot et al., 2007, Waterlander et al., 2010), differences in energy density by categories of diet cost (Andrieu et al., 2006, Murakami et al., 2007, Monsivais and Drewnowski, 2009, Townsend et al., 2009, Aggarwal et al., 2011) as well as significant negative trends and associations (Drewnowski et al., 2007, Maillot et al., 2007, Townsend et al., 2009, Waterlander et al., 2010, Alexy et al., 2012). Results were in a negative direction regardless of study quality: whether analyses were adjusted or not, how energy was accounted for, how diet costs were defined, and whether beverages were included in energy density estimates. Conclusions were similar across analytical approaches, samples and countries, and for both men and women. However, some findings indicated a stronger association for women (Drewnowski et al., 2007, Monsivais and Drewnowski, 2009).

The only non-significant result to be reported was that of Drewnowski et al (2009), who found that weekly diet costs did not differ significantly in unadjusted analyses between quintiles of dietary energy density amongst men in a nationally representative French sample.

Table 2.1 Study characteristics: studies investigating dietary expenditure/cost and dietary energy density

Ref	Country	Sample size	Sample characteristics	Exposure	Exposure details	Year(s) of exposure data collection	Outcome	Outcome assessment details	Year(s) of outcome data collection	Length to follow-up	Loss to follow-up
Andrieu et al. (2006)	France	1474	Nationally representative dietary survey; adults; 46% male	Dietary energy cost (€/10MJ)	Mean national prices (marketing research) x760 applied to 7d records.	1998 (diet), 1997 (prices)	Energy density; micronutrient intake	7d records matched to national nutrient database (895 items)	1998	N/A	N/A
Drewnowski et al. (2007)	France	1,985	National Survey on Individual Food Consumption: 15-92yrs, nationally representative	Dietary energy density (kcal/g)	7-day diary. EI divided by edible weight. Excluded water, diet beverages, tea and coffee	1999	Diet cost ($/7d or $/2000kcal)	Mean national retail prices taken from French National Institute of Statistics & supermarket sites. Adjusted for preparation & waste as per USDA. Collected in €, reported in $	1997	N/A	Excluded 511 under-reporters
Maillot et al. (2007b)	France	1,332	French National Agency for Food Safety survey: nationally representative	Dietary energy density (MJ/kg); Mean adequacy ratio (MAR)	7d food records. Excluded beverages in energy density calculations. MAR based on % of recommended intakes of 23 nutrients. Excluded alcohol, tea, coffee & drinking water	1999	Diet cost (€/10MJ, €/d)	Prices from marketing research (SECODIP), French National Institute of Statistics (INSEE) & supermarket websites. Adjusted for preparation & waste. Excluded alcohol, tea, coffee, drinking water.	1997		
Murakami et al. (2007)	Japan	3931	Female dietetic students, 54 institutions	Dietary energy cost (yen/1000kcal)	National Retail Price Survey (n=122) applied to DHQ (135 items)	2005 (diet), 2004 (prices)	Foods intake; nutrients intake	Intakes calculated from DHQ	2005	N/A	N/A
Monsivais and Drewnowski (2009)	USA, Pacific North-west	164	Staff of public university. Excluded those reporting FAFH >6/week	Dietary energy density (kcal/g)	152-item FFQ	2005-2006	Dietary energy cost ($/2000kcal)	Prices from 3 supermarket chains in Seattle region, for 384 component foods for each FFQ item, compiled using weighted means	2006	N/A	N/A

Table 2.1 (cont'd) Study characteristics: studies investigating dietary expenditure/cost and dietary energy density

Ref	Country	Sample size	Sample characteristics	Exposure	Exposure details	Year(s) of exposure data collection	Outcome	Outcome assessment details	Year(s) of outcome data collection	Length to follow-up	Loss to follow-up
Townsend et al. (2009)	California USA	112	Non-institutionalised low-income women, 20-55yrs; English-speaking; ethnically diverse	Dietary energy cost	Composite items assigned a mean food price ($/100g edible portion) from 8 markets	2006	Energy density; macronutrient intake	152-item FFQ (ref period previous 3mo)	Not reported		
Waterlander et al. (2010)	The Nether-lands	373 + 200	AGHLS: cohort recruited at 13yrs, mean age 36yrs. + LASA: 55-85yrs, community-dwelling elderly	Energy density	AGHLS: computer-assisted face-to-face interview: reference period of preceding 4 weeks. LASA: 2x 24hr recall. Beverages excluded (as well as fruit juices)	AGHLS: 2000. LASA: 2007	Diet costs (€/2000kcal)	Prices from 2 supermarket chains (44% market share)	2008		29 LASA & 40 AGHLS excluded as outliers
Aggarwal et al. (2011)	Washing-ton, USA	1903 (analytical 1266)	64% women; mean age 56yr, 57% college graduates	Diet cost (residual of $/d)	Retail prices from 3 local supermarkets applied to FFQ	2008-9 (date of price data collection not reported)	Energy density (kJ/g); Mean Adequacy Ratio (MAR)	ED from food only; MAR is index of % of daily recommendations for 11 nutrients (% adequacy/d)	2008-9	N/A	N/A
Alexy et al. (2012)	Germany	494	4-18yrs, DONALD study; 52% male	Diet cost (€/d), food group cost (% €/d)	Retail prices found for representative items from 8 food groups (n=356) applied to food group consumption (g/d) from 3d weighed records	2006-8 (diet), 2009 (prices)	Energy density	Excluding water & caloric beverages	2006-8	Not reported	Not reported

AGHLS Amsterdam Growth & Health Longitudinal Study; LASA Longitudinal Ageing Study Amsterdam; DONALD Dortmund Nutritional and Anthropometric Longitudinally Designed Study

Table 2.2 Results: studies investigating dietary expenditure/cost and dietary energy density

Ref	Exposure	Outcome(s)	Comparison/Subgroup	Statistical treatment	Adjustments	Results	p value	Summary of results
Andrieu et al. (2006)	Energy cost quartiles (€/10MJ)	Energy density		ANOVA	Age & sex	Q1 6.42 (6.30, 6.54), Q2 6.08 (5.96, 6.20), Q3 5.97 (5.85, 6.09), Q4 5.72 (5.60, 5.84)	0.0001	Energy density decreased with higher energy costs.
Drewnowski et al. (2007)	Dietary energy density (kcal/g)	Diet cost ($/week)	Women	Multiple regression	Energy intake, age	$R^2 = 0.38$	<0.001	Energy density and weekly diet costs were significantly positively associated.
	Dietary energy density (kcal/g)	Diet cost ($/week)	Men	Multiple regression	Energy intake, age	$R^2 = 0.44$	<0.001	
	Quintile of dietary energy density (kcal/g)	Diet cost ($/week)	Men	ANOVA	N/A	Q1 65.86 ± 22.49; Q2 68.06 ± 23.66; Q3 65.30 ± 24.96; Q4 60.97 ± 19.11; Q5 60.66 ± 19.50	0.023	Weekly diet costs significantly differed between quintiles of dietary energy density amongst men but not women, with lower diet costs in higher quintiles of ED.
	Quintile of dietary energy density (kcal/g)	Diet cost ($/week)	Women	ANOVA	N/A	Q1 51.35 ± 14.17; Q2 51.48 ± 15.60; Q3 51.74 ± 16.90; Q4 49.66 ± 16.90; Q5 47.81 ± 16.38	NS	
	Quintile of dietary energy density (kcal/g)	Diet cost ($/2000kcal)	Men	ANOVA	N/A	Q1 8.26 ± 2.61; Q2 7.95 ± 1.96; Q3 7.42 ± 2.07; Q4 6.62 ± 1.41; Q5 6.49 ± 1.41	0.001	Diet costs significantly differed between quintiles of energy density for both men and women, with a negative trend.
	Quintile of dietary energy density (kcal/g)	Diet cost ($/2000kcal)	Women	ANOVA	N/A	Q1 8.39 ± 1.96; Q2 7.78 ± 1.96; Q3 7.45 ± 1.74; Q4 7.07 ± 1.96; Q5 6.64 ± 1.68	0.001	

Table 2.2 (cont'd) Results: studies investigating dietary expenditure/cost and dietary energy density

Ref	Exposure	Outcome(s)	Comparison/Subgroup	Statistical treatment	Adjustments	Results	p value	Summary of results
Maillot et al. (2007b)	Energy density tertile	Diet cost (€/10MJ)		GLM	Age, energy intake	Individual figures not reported (bar chart)	<0.05	Energy density and energy costs were significantly negatively associated, whether energy density was included categorically or continuously.
	Energy density (MJ/kg)	Diet cost (€/10MJ)		Multivariate linear regression	Age, energy intake	β (SD): Men -0.235 (-.847); women -0.171 (0.897)	<0.0001	
Murakami et al. (2007)	Quintile of energy cost	Energy density (kcal/g)		Linear regression	Results for unadjusted model only shown		<0.0001	Energy density was significantly lower in increasing quintiles of diet cost.
Monsivais and Drewnowski (2009)	Dietary energy density (kcal/g) tertiles	Dietary energy cost ($/2000kcal)	Women	Bivariate methods with linear trend tests	None reported	Lowest 9.55 ± 1.82; middle 8.06 ± 1.25; highest 6.76 ± 0.87	<0.001	There was a significant negative trend in energy cost by energy density tertiles for both men and women.
	Dietary energy density (kcal/g) tertiles	Dietary energy cost ($/2000kcal)	Men	Bivariate methods with linear trend tests	None reported	Lowest 7.82 ± 1.28; middle 7.74 ± 1.27; highest 6.71 ± 1.15	0.006	
	Dietary energy density (kcal/g)	Dietary energy cost ($/2000kcal)		Least-squares regression	None reported	R^2 = 0.37	Not reported	Energy density and energy costs were weakly to modestly negatively associated (p value not reported). The correlation was stronger for women than men.
	Dietary energy density (kcal/g)	Dietary energy cost ($/2000kcal)	Men & women	Least-squares regression	None reported	Men R^2 = 0.09; women R^2 = 0.51	Not reported	
	Dietary energy cost ($/2000kcal) tertiles	Dietary energy density (kcal/g)	Women	Bivariate methods with linear trends tests	None reported	Lowest 1.60 ± 0.27; middle 1.33 ± 0.22; highest 1.12 ± 1.60	<0.001	There were significant negative trends in energy density by tertile of energy cost in both men and women (more strongly in women).
	Dietary energy cost ($/2000kcal) tertiles	Dietary energy density (kcal/g)	Men	Bivariate methods with linear trends tests	None reported	Lowest 1.58 ± 0.29; middle 1.51 ± 0.39; highest 1.35 ± 0.18	0.017	

Table 2.2 (cont'd) Results: studies investigating dietary expenditure/cost and dietary energy density

Ref	Exposure	Outcome(s)	Comparison/Subgroup	Statistical treatment	Adjustments	Results	p value	Summary of results
Townsend et al. (2009)	Dietary energy cost tertiles (excluding beverages) ($/2000kcal)	Dietary energy density (excluding beverages) (kcal/g)	Tertiles	ANOVA	None	Means: 1.77 ± 0.30; 1.55 ± 0.25; 1.31 ± 0.22	<0.001	Tertiles of energy cost significantly differed in dietary energy density, with the lowest cost tertile showing the highest density. This was true whether or not beverages were included in energy density estimates.
	Dietary energy cost tertiles (including beverages, except water) ($/2000kcal)	Dietary energy density (including beverages) (kcal/g)	Tertiles	ANOVA	None	Means: 1.02 ± 0.32; 1.01 ± 0.26; 0.80 ± 0.20	<0.001	
	Dietary energy cost (excluding beverages) ($/2000kcal)	Energy density (excluding beverages) (kcal/g)		Least-squares regression	Energy intake	R^2 = 0.40	<0.001	Energy costs and ED were significantly negatively associated after adjusting for energy.
Waterlander et al. (2010)	Diet costs (€/2000kcal)	Energy density (kJ/g)	AGHLS	Pearson's correlations		Men r=-0.505; women r= -0.413	<0.001	Energy costs and energy density were moderately negatively correlated for men and women.
	Diet costs (€/2000kcal)	Energy density (kJ/g)	LASA	Pearson's correlations		Men r=-0.559; women r= -0.562	<0.001	
	Energy density (kJ/g) quartiles	Diet costs (€/2000kcal)	LASA men	ANOVA		Q1 6.01 (SD 1.08); Q2 5.11 (0.87); Q3 5.18 (1.03); Q4 4.19 (0.61)	<0.001	In both men and women, for both samples, diet costs decreased with increasing quintiles of energy density.
	Energy density (kJ/g) quartiles	Diet costs (€/2000kcal)	LASA women	ANOVA		Q1 5.61 (SD 1.04); Q2 5.23 (0.97); Q3 4.76 (0.99); Q4 3.93 (1.03)	<0.001	
	Energy density (kJ/g) quartiles	Diet costs (€/2000kcal)	AGHLS men	ANOVA		Q1 5.09 (SD 0.80); Q2 4.86 (0.89); Q3 4.72 (0.62); Q4 4.01 (0.54)	<0.001	
	Energy density (kJ/g) quartiles	Diet costs (€/2000kcal)	AGHLS women	ANOVA		Q1 4.94 (SD 0.63; Q2 4.69 (0.54); Q3 4.56 (0.72); Q4 4.25 (0.67)	<0.001	

Table 2.2 (cont'd) Results: studies investigating dietary expenditure/cost and dietary energy density

Ref	Exposure	Outcome(s)	Comparison/Subgroup	Statistical treatment	Adjustments	Results	p value	Summary of results
Aggarwal et al. (2011)	Quintiles of energy-cost residuals (costs from food only)	Energy density (kJ/g)		Multivariable linear regression	Age, sex, ethnicity, household size, EI	B coefficient = -0.89	<0.0001	Every additional standard deviation of diet cost residual was associated with a significant reduction in energy density of 0.89kJ/g.
Alexy et al. (2012)	Energy density (kJ/d)	Diet cost (€/d)		Linear mixed effects model	ED*age, age*sex, age*age*sex	B coefficient = -0.20	<0.0007	Negative association (also for non-linear term kJ*kJ)
	Meat/sausage cost (% diet cost)	Energy density (kJ/g)		Linear mixed effects model	ED*ED, age	B coefficient = 5.5	<0.0001	Proportion of diet cost from meat/sausage, bread, confectionary, potatoes/rice/pasta all positively associated with energy density. Proportion of diet cost from dairy, vegetables and fruit all negatively associated with energy density. Proportional costs from convenience/fast foods not significantly associated.
	Dairy cost (% diet cost)	Energy density (kJ/g)		Linear mixed effects model	ED*ED, age	B coefficient = -14.3	<0.0001	
	Convenience/fast food cost (% diet cost)	Energy density (kJ/g)		Linear mixed effects model	ED*ED, age	B coefficient = 1.1	0.4916	
	Bread cost (% diet cost)	Energy density (kJ/g)		Linear mixed effects model	ED*ED, age	B coefficient = 2	0.0004	
	Vegetables cost (% diet cost)	Energy density (kJ/g)		Linear mixed effects model	ED*ED, age	B coefficient = -2.48	0.0003	
	Fruits cost (% diet cost)	Energy density (kJ/g)		Linear mixed effects model	ED*ED, age	B coefficient = -2.03	0.0008	
	Confectionary cost (% diet cost)	Energy density (kJ/g)		Linear mixed effects model	ED*ED, age	B coefficient = 1.38	0.0235	
	Potatoes/rice/pasta cost (% diet cost)	Energy density (kJ/g)		Linear mixed effects model	ED*ED, age	B coefficient = 1.28	0.0048	

2.4.3.3 Income & DED

Studies' designs and settings

Five studies were found to have investigated income and dietary energy density (Table), all of which were published recently (since 2006). Only one of these studies was based outside the US – that of Waterlander et al (2010), which used Dutch data. The Dutch study used data from the smallest sample sizes – 373 and 200 participants respectively from two cohorts. Three of the US studies used data from large, nationally representative surveys: two from the National Health and Nutrition Examination Surveys (NHANES) (Kant and Graubard, 2007, Kant and Graubard, 2013) and one from the Continuing Survey of Food Intake for Individuals (CSFII) (Mendoza et al., 2006). The final study (Aggarwal et al., 2011) used a moderate-sized (n=1318), regional-specific sample from Washington State. All of the studies were cross-sectional in design (even those which used data from longitudinal cohorts, such as Waterlander et al (2010), in which a single year of data collection was used). Most of the studies focused on adult samples, the exception being Mendoza et al (2006) and Kant and Graubard (2013) in which data from children and adolescents were analysed. One of the samples used by Waterlander et al was restricted to community-dwelling elderly.

Definition & measurement of income

Three of the studies expressed income in relation to the national poverty line: as a ratio (Kant and Graubard, 2007, 2013), or as a percentage (Mendoza et al., 2006). In each of these studies, the poverty line was year specific (where studies used more than one year of survey data collection) and specific to household composition. In these surveys, income was self-reported at interview for the family (NHANES) or household (CSFII) level. Income was then expressed as poverty categories: five (Kant and Graubard, 2007), four (Mendoza et al., 2006) or three (Kant and Graubard, 2013).

The other US-based study (Aggarwal et al., 2011) used self-reported household income, dichotomised into high and low categories (above or below the state median). In analyses, household size was used to adjust for differences in composition.

The study of Waterlander et al (2010) used data from two separate surveys, in which self-reported income was gathered differently: in the Amsterdam Growth and Health Longitudinal Study (AGHLS) sample, gross annual income was reported, whereas in the Longitudinal Ageing Study Amsterdam (LASA), net monthly household income was obtained. Participants were categorised into groups based upon the national median, in each case. Due to missing data, the authors were unable to adjust for household size or composition.

Assessment of dietary energy density

Dietary data was most commonly gathered through 24-hour recall (Kant and Graubard, 2007, 2013, Mendoza et al., 2006, the LASA sample of Waterlander et al., 2010). The exceptions to this were Aggarwal et al (2011), in which a FFQ was the tool used, and the AGHLS sample reported in Waterlander et al (2010), which employed face-to-face interviews for intakes in the preceding four-week reference period.

Three of the studies calculated dietary energy density from food only, excluding energy and mass values from beverages (Waterlander et al., 2010, Aggarwal et al.,2011, Kant and Graubard, 2013). Mendoza et al (2006) excluded water and human milk; whereas Kant and Graubard (2007) excluded beverages but not milk or 100% fruit juices in their calculations. The units used to express energy density also varied.

Analytical approaches

Table shows that multivariate regression was the most common approach (Kant and Graubard, 2007, Mendoza et al., 2006, Kant and Graubard, 2013). Different covariates were specified in the models of each of these studies: all appropriately adjusted for demographic variables, and Kant and Graubard (2007, 2013) adjusted for survey and data collection characteristics, as appropriate. Additional covariates chosen include food weight and total milk consumption (Mendoza et al., 2006), and Kant and Graubard (2013) additionally adjusted for household size and BMI. The analyses of Waterlander et al (2010) did not introduce covariates, using ANOVA and t tests

The study of Aggarwal et al (2011) differed in its approach, featuring energy density as the exposure and income category as the outcome. Logistic regression was used, adjusting for sociodemographic variables, household size and energy intake.

Quality of studies

All of the studies were cross-sectional in design (even where longitudinal data were available), and are thus subject to biases associated with observational studies. The studies using data from nationally representative surveys (Kant and Graubard, 2007, Mendoza et al., 2006, Kant and Graubard, 2013) were stronger in quality in terms of sampling design, but also employed more robust analytical techniques. In particular, these studies accounted for household composition, which is important when considering household income (see Chapter 4).

The studies of Aggarwal et al (2011) and Waterlander et al (2010) are additionally weakened by the employment of non- or semi-quantified dietary assessment techniques (FFQ and 4-week recall respectively), which could possibly introduce bias in the calculation of energy density (although the extent of this is unknown). In using unadjusted analytical techniques, Waterlander et al. are also unable to allow for confounding variables, which is an essential means of attempting to counteract bias in observational studies. Furthermore, the samples used in the study of Waterlander et al. (2010) were relatively small in size and, although the authors do not explicitly report power calculations in the report, they raise the issue of inadequate power in discussing the results.

In terms of children and adults, the studies involving children were of stronger quality. Of the studies involving adults, two of the three studies (Aggarwal et al., 2011 and Waterlander et al., 2010) were of poorer quality (for the reasons described above).

Findings

Amongst children, overall analyses failed to uncover significant differences in energy density between family poverty categories (Mendoza et al, 2006, Kant and Graubard, 2013). In subgroup analyses, Mendoza et al found a significant association between income and energy density amongst 0-to-4-year-old participants. Kant and Graubard (2013), however, did not find this in their age-stratified analyses.

Amongst adults, two studies reported significant findings: Kant and Graubard (2006) reported a significant negative relationship between poverty income ratio (PIR) and energy density; whilst Aggarwal et al (2011) found that the odds of having a higher income were significantly lower as DED increased. In unadjusted comparisons of Dutch income groups, Waterlander et al (2010) did not find any significant differences in DED by income groups in either sample.

Table 2.3 Study characteristics: studies investigating income and dietary energy density

Ref	Country	Sample size	Sample characteristics	Exposure	Exposure details	Year(s) of exposure data collection	Outcome	Outcome assessment details	Year(s) of outcome data collection	Length to follow-up	Loss to follow-up
Kant and Graubard (2007)	USA	36,600	NHANES I, II, III & 1999-2002. Nationally representative: adults aged 25-74yrs.	Poverty Income Ratio (PIR)	The ratio of total family income to the poverty threshold for each survey year for a family of given characteristics: <1 is below threshold.	1971-1975, 1976-1980, 1988-1994, 1999-2002	Dietary energy density	Kcal/g; food, milk & 100% fruit juices. Assessed by 24hr recall	1971-1975, 1976-1980, 1988-1994, 1999-2002	N/A	N/A
Mendoza et al. (2006)	USA	18,344	CSFII. Nationally representative, children & adults <20yrs. Mean age 9.3yr, 48.9% female.	Poverty category	Household income as a % of poverty level	1994-1996, 1998	Dietary energy density	Mean daily kcal/mean daily g; excluded water & human milk. Assessed by 2 nonconsecutive 24hr recalls (proxy interviews for children <6yrs)	1994-1996, 1998	N/A	Missing data left analytical sample of 11,284
Waterlander et al. (2010)	The Nether-lands	373 + 200	2 longitudinal cohorts: AGHLS: cohort recruited at 13yrs, mean age 36yrs LASA: 55-85yrs, community-dwelling elderly, mean age 69yrs,	Income category	AGHLS: 5 categories of gross annual income, recoded into 3 groups (below, at, or above Dutch modal income before tax). LASA: 11 categories net monthly household income, recoded into 2 groups (below or above modal Dutch income after tax).	AGHLS: 2000. LASA: 2007	Dietary energy density	kJ/g, calculated from ΣE/ΣW, beverages excluded. AGHLS: computer-assisted face-to-face interview: reference period of preceding 4 weeks. LASA: 2x 24hr recall.	AGHLS: 2000. LASA: 2007	N/A (cross-sectional data drawn from longitudinal cohorts)	
Aggarwal et al. (2011)	USA	1318 (analytical sample 1266)	SOS. Stratified sample with over-sampling of low income and ethnic minorities, adults, 64% female, mean age 56yr, 57% college graduates	Quintile of dietary energy density; SD dietary energy density	kJ/g, from food only, calculated from FFQ	2008-9	Household income (high vs low)	Self-report annual hhold income: 'high' defined as at or above state median ($50,000)	2008-9	N/A	69% response rate

Table 2.3 (cont'd) Study characteristics: studies investigating income and dietary energy density

Ref	Country	Sample size	Sample characteristics	Exposure	Exposure details	Year(s) of exposure data collection	Outcome	Outcome assessment details	Year(s) of outcome data collection	Length to follow-up	Loss to follow-up
Kant and Graubard (2013)	USA	39,822 (analytical)	NHANES I, II, III, 1999-2002 & 2003-2008. Nationally representative, children & adolescents aged 2-19yr	Poverty Income Ratio (PIR)	The ratio of total family income to the poverty threshold for each survey year for a family of given characteristics: <1 is below threshold.	1971-1974, 1976-1980, 1988-1994, 1999-2000, 2001-2002, 2003-2004, 2005-2006, 2007-2008	Dietary energy density	Kcal/g; food only; assessed by 1x 24hr recall	1971-1974, 1976-1980, 1988-1994, 1999-2000, 2001-2002, 2003-2004, 2005-2006, 2007-2008	N/A	N/A

NHANES - National Health and Nutrition Examination Surveys, CSFII – Continuing Survey of Food Intake for Individuals, AGHLS – Amsterdam Growth and Health Longitudinal Study, LASA – Longitudinal Ageing Study Amsterdam, SOS – Seattle Obesity Study

Table 2.4 Results: studies investigating income and dietary energy density

Ref	Exposure	Outcome(s)	Comparison/Subgroup	Statistical treatment	Adjustments	Results	p value	Summary of results
Kant and Graubard (2007)	Categories of PIR. category 1: <1.0; category 2: 1.0-1.99; category 3: 2.0-2.99; category 4: 3.0-3.99; category 5: ≥4.0	Dietary energy density (kcal/g)		Linear multiple regression	Sex, age, age^2, race/ethnicity, years of education, survey	Coefficients (SE): Category 1: 1.65 (0.02); 2: 1.68 (0.01); 3: 1.66 (0.01); 4: 1.65 (0.01); 5: 1.62 (0.01)	(Trend) 0.003	Higher PIR was associated with lower energy density.
Mendoza et al. (2006)	Poverty category (% of poverty line). category 1: <100%; category 2: 100-199%; category 3: 200-299%; category 4: ≥300%	Dietary energy density (kcal/g)		Linear regression	None	Means (95% CI): Category 1: 1.13 (1.11, 1.15); 2: 1.14 (1.12, 1.16); 3: 1.13 (1.11, 1.15); 4: 1.13 (1.12, 1.15)	ns	DED did not significantly differ between poverty categories.
			0-4 year-olds	Multivariate linear regression	Sex, age, age^2, education level of head of hhold, race/ethnicity, food weight, total milk	B coefficients (95% CI): Category 1: 0.03 (0.01, 0.05); 2: 0.03 (7.6×10^{-3}, 0.05); 3: 0.01 (-0.01, 0.04); 4: Reference	<0.05	There was a significant association amongst 0-4 year-old participants, with higher energy density associated with lower incomes. The association was not significant for older children.
			5-11 year-olds	Multivariate linear regression	As above	B coefficients (95% CI): Category 1: 8.0×10^{-3} (-0.03, 0.05); 2: -0.01 (-0.05, 0.03); 3: -0.01 (-0.04, 0.01); 4: Reference	ns	
			12-19 year-olds	Multivariate linear regression	As above	B coefficients (95% CI): Category 1: -0.04 (-0.08, 4.0×10^{-3}); 2: -0.02 (-0.06, 0.02); 3: -0.05 (-0.10, 5.6×10^{-3}); 4: Reference	ns	

Table 2.4 (cont'd) Results: studies investigating income and dietary energy density

Ref	Exposure	Outcome(s)	Comparison/Subgroup	Statistical treatment	Adjustments	Results	p value	Summary of results
Waterlander et al. (2010)	Income category. Category 1: below modal,; 2: modal; 3: above modal	Dietary energy density (kJ/g)	AGHLS men	ANOVA	N/A	Mean (SD): Category 1: 6.23 (0.92) 2: 6.49 (1.46) 3: 6.52 (1.34)	0.678	There were no significant differences in dietary energy density between any of the income categories.
	Income category: as above	As above	AGHLS women	ANOVA	N/A	Mean (SD): Category 1: 6.65 (1.13) 2: 6.40 (1.34) 3: 6.36 (1.09)	0.410	
	Income category. Category 1: below modal; category 2: above modal	As above	LASA men	T test	N/A	Mean (SD): Category 1: 5.82 (4.39) Category 2: 6.44 (1.55)	0.069	
	Income category: as above	As above	LASA women	T test	N/A	Mean (SD): Category 1: 6.78 (1.76) Category 2: 6.74 (1.26)	0.835	
Aggarwal et al. (2011)	Energy density quintile	Proportion classified as higher income		Multivariate logistic regression	Age, sex, race/ethnicity, household size, EI	% (95% CI): Q1 60.3 (45.6, 73.3) Q2 49.5 (35.0, 64.0) Q3 49.9 (35.5, 64.5) Q4 46.7 (32.6, 61.3) Q5 38.5 (25.8, 52.9)	<0.0001	Higher energy density was associated with a lower proportion of participants classified as high income.
	Energy density SD	Odds of being classified as higher income		Multivariate logistic regression	As above	β coefficient 0.77	<0.0001	
Kant and Graubard (2013)	Family PIR category. category 1: <130%; category 2: 130-349%; category 3: ≥350%	Dietary energy density (kcal/g)	Stratified into age groups: 2-5yr, 6-11yr, 12-19yr	Multivariable linear regression	Age, sex, race-ethnicity, survey cycle, month of measurement, weekday of recalled intake, education of household head, household size, and BMI-sex-age-percentile	Values not reported	All ns	Family PIR was unrelated to dietary energy density.

2.4.4 Body mass index or weight

2.4.4.1 Food prices & body weight

Studies' designs and settings

Twenty four articles investigating food prices and body weight were found to meet the study criteria: 13 using adult samples, 10 focussing on children or adolescents; and one which used data from both children and adults. Table shows details of the studies and Table 2.6 summarises their findings, organised into adult or children studies.

Studies used a combination of cross-sectional observations (n = 12), longitudinal or time series data (n = 11), and one study used a before-and-after comparison.

Children

Of the 11 studies which included data from children, all but two (Thomas et al., 1996, Black et al., 2013) were based in the US. The US studies drew from nationally representative samples: the Monitoring the Future (MTF) survey (Powell et al., 2007, Auld and Powell, 2009), the National Longitudinal Survey of Youth (NLSY97) (Powell, 2009, Powell and Bao, 2009), NHANES ((Fletcher et al., 2010b, Fletcher et al., 2009), and the Early Childhood Longitudinal Study – Kindergarten (ECLS-K) (Sturm and Datar, 2005, 2008, Sturm et al., 2010). The study based in Cote d'Ivoire (Thomas et al., 1996) drew from a random sample of households; whereas the remaining study, in Australia (Black et al., 2013), used a non-randomised sample of low-income Aboriginal children participating in a subsidy programme.

A varied selection of children's age ranges were apparent in the studies: four studies used a broad range (for example, from 2 to 17 years) (Powell and Bao, 2009, Black et al., 2013, Fletcher et al., 2009, 2010), whilst one study focussed on adolescents (Powell, 2009) and the remainder used data from younger children.

Adults

Studies using adult data were also predominantly conducted in the US: only four of the 15 were based elsewhere (Thomas et al., 1996, in Cote d'Ivoire; Asfaw et al., 2007, in Egypt; Staudigel, 2011, in Russia; and Lear et al., 2013, in Canada) and none were based in the UK. All but one of the American studies used data from large nationally representative surveys: the Behavioral Risk Factor Surveillance System (BRFSS) (Chou et al., 2004, Schroeter and Lusk, 2008, Cotti and Tefft, 2013, Kim and

Kawachi, 2006, Fletcher et al., 2010a), CSFII (Beydoun et al., 2008), the Panel Study of Income Dynamics (PSID) (Powell and Han., 2011), NLSY79 (Zhang et al., 2011) and the Coronary Artery Risk Development In Young Adults (CARDIA) study (Duffey et al., 2010).

Two of the non-US studies used data from nationally representative samples (Asfaw et al., 2007, in Egypt, and Staudigel, 2011, who used data from the Russia Longitudinal Monitoring Survey). The study in Cote d'Ivoire (Thomas et al., 1996) used a random sample of households. The survey by Lear et al (2013) used an opportunity sample of adults.

Two studies restricted their sample to women (Zhang et al., 2013, Asfaw, 2007).

Assessment of prices

There were three approaches to quantifying food prices in analyses: firstly, to use prices for a number of selected food items (the number of which ranged from four to 20); secondly, to create composite indices from food groups or types of food (for example, fast food, fruit and vegetables, or food eaten at home); or, thirdly, to compare regions or years according to taxes or subsidies.

Six studies looked at price data for individual food items (Thomas et al., 1996, Asfaw et al., 2007, Miljkovic et al., 2008, Duffey et al., 2010, Staudigel, 2011, Lear et al., 2013), the majority of which (n=4) were non-US based studies. The US-based studies took price data from national statistics (Miljkovic et al., 2008, Duffey et al., 2010), whereas the other studies used price data collected within the sample survey. Lear et al (2013) used the smallest sample of retailers to gauge prices, collecting data from just five supermarkets.

Twelve studies combined prices of individual items to give an index for a given food group or type. Most commonly, this was done for fast foods (n=6) or fruit and vegetables (n=6). Other indices reported were for: unhealthy foods (Zhang et al., 2011), food at home (Chou et al, 2004, Schroeter and Lusk, 2008, Powell, 2009), restaurant prices (Chou et al., 2004), or food groups (Sturm and Datar, 2005, 2008).

Five studies examined the effect of taxes on body weight. All of these used US data, and all focussed on soft drinks, whilst one study additionally analysed taxes on snack foods (Kim and Kawachi, 2006). Three of these studies compared taxes regionally, using state-level tax data (Fletcher et al., 2009, Sturm et al., 2010, Kim and Kawachi, 2006), whilst two examined changes in taxes temporally (Fletcher et al., 2010a, 2010b). One study examined the effect of a subsidy on fruit and vegetables

(Black et al., 2013) using a before-and-after analysis of participants in a state-funded programme.

One study (Asfaw et al., 2007) did not quantify food prices per se, but rather compared body weight before and after a general food price shock in Egypt.

Assessment of anthropometry

The majority of studies relied upon self-reported height and weight measurements (n=13), seven studies used professionally-measured anthropometry (Black et al., 2013, Fletcher et al., 2009, 2010a, 2010b, Sturm and Datar, 2005, 2008, Sturm et al., 2010, Duffey et al., 2010), and one study used a combination of self-report and professional measurements (Powell and Bao, 2009). Three studies did not report anthropometry measurement (Thomas et al., 1996, Asfaw et al., 2007, Staudigel, 2011).

The majority of studies (n=16) included continuous BMI (kg/m^2) (or z scores where appropriate) as the outcome. Other outcomes reported were: BMI categories (n=2), change in BMI (n=4), incidence/prevalence of overweight (n=4), incidence/prevalence of obesity (n=6) and body weight (n=3). Several studies (n=12) reported more than one body weight outcome.

Analytical approaches

All but one of the studies (Fletcher et al., 2010b) used multivariable regression techniques to test their hypotheses, adjusting for a wide range of confounders. As well as ordinary least squares (OLS) models (used in 11 studies), regression analyses employed a variety of model types, such as maximum likelihood probit, quadratic, fixed effects, random effects, lagged effects, two-stage least squares, logistic or multinomial.

Quality of studies

In adults, the included studies were generally found to be of acceptable quality. Although there were no randomized trials found to investigate food prices and body weight, many of the studies used longitudinal or time series data, with only three studies relying on cross-sectional designs (Thomas et al., 1996, Asfaw et al., 2007, Lear et al., 2013). Of these three studies, one (Asfaw et al., 2007) used a nationally representative sample in Egypt, one used a random sample of households (Thomas et al., 1996), but one used an opportunity sample (Lear et al, 2013). Probability sampling techniques will help protect against selection bias, and sound sampling approaches were reported in all but the opportunity sample reported by Lear et al (2013). A non-probability sample such as this is likely to introduce bias in the study. The study by

Lear et al. (2013) also reported the smallest sample size of the included studies, and the smallest range of price sources, with Thomas et al. (1996) also showing small numbers in these aspects. The small range of prices used may lead to values being used that do not reflect the distribution of prices in the study setting, and can thus be a source of measurement bias. Asfaw et al. (2007) used proxy prices in their analyses, although how these predicted prices were calculated is not clearly reported. This impacts on the judgement of quality of this study as it is difficult to assess how the methods could have introduced bias in the realisation of the independent variables.

In terms of analysis, the quality of the studies in adults was found to be good: all studies used appropriate and well-considered statistical analyses, with adjustment for important confounders as well as adjusting for design and longitudinal effects where necessary.

Despite strengths in sampling and analysis, a prominent shortcoming in the quality of the studies in adults was in measurement of the outcome. The majority of the studies – with the exceptions of Thomas et al. (1996), Asfaw et al. (2007) and Duffey et al. (2010) – relied on self-reported height and weight (one study (Staudigel, 2011) failed to report how BMI was measured). This is an important source of bias in BMI research, as participants tend to under-report weight and over-estimate height, although there is much variability in these tendencies (Gorber et al., 2007). As a result, the three studies using professionally measured anthropometry should be considered stronger in quality of outcome assessment. Taking all these sources of bias into account, the study of Duffey et al. (2010) was found to be particularly strong in terms of quality.

In children, the studies were of poorer quality than the adult studies in terms of reporting and in particular the reporting of analytical methods. On the other hand, more of the studies in children utilised objective measures of anthropometry, rather than self-reported height and weight.

The majority of the studies in children were of good quality in terms of samples used, with many of them using large nationally representative surveys – the only studies which did not use representative samples were Thomas et al. (1996) and Black et al. (2013). These two studies also suffered in terms of sample size, which will have resulted in a lower power to detect effects than in the other larger studies. The sample of Thomas et al. (1996), whilst modest in size and not nationally representative, was selected using sound, randomised methods to minimise selection bias. Black et al. (2013), using a before-and-after design in a subsidy programme, were unable to use a probability sample, making this study more open to sampling bias.

Several of the studies in children used longitudinal cohorts (Sturm and Datar, 2005, Sturm and Datar, 2008, Powell, 2009, Powell and Bao, 2009, Sturm, 2010), which may be considered of higher quality than cross-sectional samples.

The three studies of weaker quality in terms of BMI measurement (using self-reports) were Powell (2007), Auld and Powell (2009) and Powell (2009). Otherwise, the studies in children used researcher measured BMI, making them better quality in this respect than the majority of the adult studies.

The studies in children for the most part used sound sources of price data (or tax data). Thomas et al. (1996), however, used a narrow range of local prices, as already mentioned above, whilst Black et al. (2013) did not measure prices per se, relying on a before-and-after paradigm.

As well as employing a less robust design, the study of Black et al. (2013) may be criticised in its reporting of statistical analysis, with an unclear statement of treatment and whether analyses were adjusted for confounding variables, an important source of potential bias in non-experimental studies. Other studies which were unclear in their reporting of the statistical approaches used were: Auld and Powell (2009), in which significance was stated without supporting p values, and Fletcher (2010b), in which neither the sample size, year of data collection nor statistical treatment were reported. Otherwise, statistical approaches of the other studies were all appropriately selected and adjusted for confounding, perhaps with the exception of Fletcher (2009), in which only the year, quarter and state were adjusted for (omitting important confounding variables such as socioeconomic status, ethnicity, sex).

Taking all of the above into account, it seems the studies showing the best overall study quality include: one of the studies employing indices, Powell and Bao (2009); one of the studies investigating soft drink taxes, Sturm et al. (2010); and the two studies investigating food group prices, Sturm and Data (2005 and 2008).

Findings

Children

Studies reporting the effects of fast food prices on children's anthropometry reported mixed findings. Powell et al (2007) found a significant negative association between fast food prices and BMI or overweight; however, using a different modelling approach on the same sample (Auld and Powell, 2009), fast food prices were not found to be significantly associated with BMI (and the p value was not reported for the negative coefficient for overweight). Amongst adolescents in the NLSY97 (Powell, 2009), fast food prices were found to be negatively associated with BMI, but only in a longitudinal model and not in cross-sectional analysis of the data. Using data from younger children of the NLSY97 as well, Powell and Bao (2009) found no significant effect of fast food prices on BMI. The one study reporting a food-at-home price index found no association with BMI.

In terms of fruit and vegetables price indices, Auld and Powell (2009) and Powell and Bao (2009) found a significant positive association with BMI (but not overweight) in different samples. However, Powell et al (2007) failed to find a significant effect. The before-and-after observations of Black et al (2013) also failed to find an effect on children's body weight or fatness of a fruit and vegetable subsidy programme. In their two studies of younger children, Sturm and Datar (2005, 2008) found a highly significant positive association of a fruit and vegetable price index with BMI increase both at baseline and in the five-year follow-up.

Studies employing indices for other food groups (meat, dairy) found no significant associations with BMI change in children (Sturm and Datar, 2005, 2008).

Most of the studies of soft drink taxes amongst children did not find a significant association between tax presence or rate and BMI, overweight or change in BMI. However, one study (Sturm et al., 2010) did find a negative relationship between soda tax amount or indicator and change in BMI amongst those children who were at risk of overweight.

In Cote d'Ivoire, the prices of all foods tested were negatively associated with weight for height in children, both urban and rural.

Adults

Three studies (Beydoun et al., 2008, Powell and Han, 2011, Cotti and Tefft, 2013) found no significant association between fast food prices and BMI or obesity. In contrast, two studies found significant negative associations with BMI (Chou et al., 2004, Schroeter and Lusk, 2008) and obesity (Chou et al., 2004). One study (Zhang et al, 2011) found a significant negative association of unhealthy food prices with BMI and obesity, but only for some models (two-stage fixed effects), and only when the two wider definitions of unhealthy food were used.

Of the three studies which examined prices of food at home, one found a significant negative association (Chou et al., 2004), and one found no association (Cotti and Tefft, 2013) on BMI. One study reported significant coefficients in its models (Schroeter and Lusk, 2008), however the direction of the association differed according to whether a quadratic, log-linear or trans log model was specified. Chou et al (2004) also found a significant negative association between restaurant price indices and BMI and obesity.

Of the two studies reporting investigations in fruit and vegetable price indices and BMI or obesity in adults, neither found an overall significant association in their samples. However, in subgroup analyses, a significant positive association was identified by Powell and Han (2011) for poorer women, or women with children; whereas Beydoun et al (2008) found a significant negative association amongst those classified as ‘near poor’ according to the poverty income ratio (PIR).

The two studies investigating the effects of soft drinks taxes on adult BMI, overweight or obesity differ in their findings. Kim and Kawachi (2006) found no difference in the odds ratios for an increase in state obesity prevalence in states with no or a repealed soft drink tax compared to states with a 5% tax rate. In contrast, Fletcher et al (2010a) found that the soft drink tax rate was significantly and negatively associated with BMI and the proportion overweight (and additionally for the proportion obese if an incremental tax rate was used in the analysis). Associations were very small, but significant. In subgroup analyses, this finding held regardless of sex or education; however, no associations were apparent amongst Black participants or those aged 18-25yrs.

The studies investigating prices of individual food items reported mixed findings, and were based in several different settings. However, all of the studies reported significant associations with body weight for prices of at least some of the foods examined. Asfaw et al (2007) found negative associations with BMI for prices of baladi bread, sugar and rice, and positive associations with fruit, eggs and milk prices amongst Egyptian women. Miljkovic et al (2008) found negative associations between sugar beet and milk prices with overweight and obesity, and a positive association with the prices of potatoes in the US. Also in the US, Duffey et al (2010) found significant negative associations between soda prices and pizza prices on body weight, but no associations for whole milk or burger prices. In Russia, Staudigel (2011) found significant negative associations between BMI and onion, chicken or sausage prices (amongst the highest income tertile only), but positive associations for butter and beef prices (highest income only).

Finally, one study compared the BMI of respondents who shopped at more expensive or less expensive supermarkets (comparing the price of a standard basket of food) (Lear et al., 2013). They found a significant negative correlation between the basket price and BMI of shoppers. In adjusted analyses, shoppers at the two least expensive stores had a significantly higher BMI than that at the most expensive.

Table 2.5 Study characteristics: studies linking food prices and body weight

Ref	Country	Sample size	Sample characteristics	Exposure	Exposure details	Year(s) of exposure data collection	Outcome(s)	Outcome assessment details	Year(s) of outcome data collection	Length to follow-up	Loss to follow-up
Studies in children and adolescents											
Price indices for fast food, food at home, food away from home, fruit & vegetables											
Powell et al. (2007)	USA	72,854 observations	MTF Survey: nationally representative, 8th & 10th graders. Approx 50% male, majority (69%) white, mean age 14.7yrs	Fast food prices and fruit & veg prices	Two indices compiled: F&V using 7 items (potatoes, bananas, lettuce, sweet peas, tomatoes, peaches, frozen corn); fast food from 3 items (McDonald's 1/4-pounder with cheese, thin crust cheese pizza at Pizza Hut/Pizza Inn, fried chicken thigh & drum). Prices drawn from ACCRA Cost of Living Index reports. Deflated to 1982-1984. Matched to MTF by geocode (closest city)	1997-2003	BMI; overweight classification >= 95th percentile (2000 CDC Growth Chart)	Self-report height & weight	1997-2003	N/A	
Auld and Powell (2009)	USA	73,041	(MTF Survey (see above)	Fast food prices and fruit & veg prices	As above	1997-2003	BMI/over-weight status	Self-reported height & weight	1997-2003	N/A	N/A
Powell (2009)	USA	5,215	Drawn from NLSY97; 12- to 17-year-olds in 1997. 51.7% male, multi-ethnic	Fast food prices & food-at-home prices	Fast food price index as above. Data from ACCRA Cost of Living Index, matched to NLSY97 by geocode, deflated to 1982-4	1997, 1998, 1999, 2000	BMI	Self-reported anthropometry. Overweight classification: BMI>= 95th percentile (CDC growth chart)	1997, 1998, 1999, 2000	2, 4 and 6 yrs	Not reported

Table 2.5 (cont'd) Study characteristics: studies linking food prices and body weight

Ref	Country	Sample size	Sample characteristics	Exposure	Exposure details	Year(s) of exposure data collection	Outcome(s)	Outcome assessment details	Year(s) of outcome data collection	Length to follow-up	Loss to follow-up
Powell and Bao (2009)	USA	3,797	Drawn from NLSY97, mother-child pairs; 6-17-year-olds. 52% male, multi-ethnic	Fast food prices and fruit & veg prices	Fast food and F&V indices as above. Data from ACCRA Cost of Living Index, matched to NLSY97 by geocode, deflated to 1982-4	1998, 2000, 2002	BMI	Mixture of objective measurements & mothers' self-reports	1998, 2000, 2002	2 and 4 yrs	Not reported
Black et al. (2013)	Australia	174	Children, 2-17yrs, from low-income Aboriginal families	Subsidised fruit & vegetables	Families participated in community programme (x3), offering 88% subsidised boxes	2008-2010	% under/normal/ overweight or obese; body fat (%)	Health professional anthropometry; children centile charts; body fatness measured by UM030 monitor (n=22)	2008-2010	Median 370d	N=31 (18%)
Soft drinks											
Fletcher et al. (2010b)	USA	Not reported	NHANES III (1988-1994) and IV (1999-2006), 3-18yrs. Nationally representative	Soft drink taxes	States that have ever had a soft drink tax vs those without. Source not reported.	Not reported	BMI z-score, overweight or obesity incidence	Measured height & weight	1988-1994 & 1999-2006 (combined)		
Fletcher et al. (2009)	USA	34,000	NHANES III & NHANES1999, ages 3-18 yrs. 15% obese, 15% overweight	Changes in state soft drinks net tax rates	Information from web searches, LexisNexis database searches and Dept of Revenue websites & publications	1989-2006	BMI, %obese/over-weight	Anthropometry taken by trained health technicians	1989-1994, 1999-2006		21,040 (final sample)

Table 2.5 (cont'd) Study characteristics: studies linking food prices and body weight

Ref	Country	Sample size	Sample characteristics	Exposure	Exposure details	Year(s) of exposure data collection	Outcome(s)	Outcome assessment details	Year(s) of outcome data collection	Length to follow-up	Loss to follow-up
Sturm et al. (2010)	USA	6,866	ECLS-K. Nationally representative	Carbonated drinks tax rates	State-level tax data from Robert Wood Johnson Foundation. Separated into: (i) difference between taxes on carbonated drinks & that on foods; (ii) indicator of whether the carbonated drinks tax higher than food	January 2004	BMI change	Researcher-measured height & weight	1998 & 2004	6yrs	Not reported
Various food items											
Sturm and Datar (2005)	USA	6,918	ECLS-K, nationally representative	Real food price indices	Indices for meats, fruit & veg, dairy, and fast food derived from ACCRA food price information	Autumn 1999	Change in BMI	Professionally assessed anthropometry	Spring 1999-Spring 2002	Yearly	Original sample size 13,282
Sturm and Datar (2008)	USA	4,557	ECLS-K, nationally representative	Real food price indices	Indices for meats, and fruit & veg derived from ACCRA food price information	Autumn 1999	Change in BMI	Professionally assessed anthropometry	Spring 1999-Spring 2004	Bi-annually	2,361
Studies in both adults and children											
Various food items											
Thomas et al. (1996)	Cote d'Ivoire	160 households	Households randomly drawn from clusters. 50% urban, 50% rural. Children under 12yrs; adults 20-60 yrs	Food prices (real food price index rose 20% in 1988)	Local prices for: beef with bones; fresh fish; rice (imported); palm oil; eggs; sugar; plantain; manioc (unprocessed); purchased by enumerators. 3 prices for each commodity where possible	1989	Weight for height (children), BMI (adults)	Cote d'Ivoire Living Standards Survey (CILSS), 3rd wave.	1987/88	N/A	N/A

Table 2.5 (cont'd) Study characteristics: studies linking food prices and body weight

Ref	Country	Sample size	Sample characteristics	Exposure	Exposure details	Year(s) of exposure data collection	Outcome(s)	Outcome assessment details	Year(s) of outcome data collection	Length to follow-up	Loss to follow-up
Studies in adults											
Price indices for fast food, food at home, food away from home, unhealthy food, fruit & vegetables											
Chou et al. (2004)	USA	1,111,074	BRFSS. Nationally representative, 18yrs+	Restaurant prices and food-at-home prices	Full-service restaurant price taken from Census of Retail Trade; fast-food and food at home prices taken from ACCRA Cost of Living Index. Deflated by CPI	1984-1999	BMI and obesity incidence	Telephone interviews: self-reported height & weight (corrected for under-reporting)	1984-1999	N/A	N/A
Beydoun et al. (2008)	USA	7,331	USDA CSFII: nationally representative. 20-65yrs	Fast Food price index (FFPI); Fruit & veg price index (FVPI)	ACCRA Cost of Living Index; matched to CSFII by city & year. FFPI: 3 items; FVPI: 7 items	1994-1996	BMI; incident obesity	Self-report height & weight	1994-1996		original sample 16,103
Schroeter et al. (2005)	USA	202,323	Adults from BRFSS, nationally representative	Normalized (ie not real) fast food & food-at-home prices	CPI from US Dept of Labor Bureau & Labor Statistics (DOL/BLS)	2003	BMI & weight	Self-reported data	2003		Not reported
Powell and Han (2011)	USA	12,851 (analytical)	PSID panel; 47% men; original sample representative of US (low-income over-sampled)	Fast food price index; fruit and veg price index	ACCRA Cost of Living Index: 6 F&V items; 3 fast food items; matched by closest (straight-line) city to PSID	Unclear	BMI	Self-reported height & weight	1999, 2001, 2003, 2005	Not specified (only analytical sample size reported)	

Table 2.5 (cont'd) Study characteristics: studies linking food prices and body weight

Ref	Country	Sample size	Sample characteristics	Exposure	Exposure details	Year(s) of exposure data collection	Outcome(s)	Outcome assessment details	Year(s) of outcome data collection	Length to follow-up	Loss to follow-up
Zhang et al. (2011)	USA	6,622 (analytical)	Women from NLSY79, nationally representative	Unhealthy food prices	ACCRA price data (225 regions) for 21 foods used to create 3 indices: UFPI (sandwich, pizza, fried chicken); UFPII (UFPI + soft drink, beef, sausage, steak); UFPIII (UFPII + margarine, sugar, potatoes)	1985-2002	BMI; obesity	Self-reported height & weight	1985 (height) - 2002	Biennially	Not reported
Cotti and Tefft (2013) 6416	USA	711,081 (analytical; from ~4m)	BRFSS (US adults, non-representative)	Fast food price index; food-at-home price index	ACCRA Cost of Living Index: 2 fast food items; 13 grocery items, across 480 areas	1990-2008	BMI; obesity	Self-reported height & weight	1990-2008	N/A	
Soft drinks											
Kim and Kawachi (2006)	USA	Not reported	BRFSS	Taxes on soft drinks and snack foods	State-level presence, degree, absence and/or repeal of tax	1991-1998	Incidence of high rate of increase of obesity rate (>75th percentile)	Obesity rates calculated from self-reported height & weight	1991-1998		
Fletcher et al. (2010a)	USA	2,709,422	BRFSS adults, nationally representative, 57% overweight, 20% obese	Changes in state soft drinks tax rates	Both incremental (excl other taxes) and total taxes. Information from web searches, LexisNexis database searches and Dept of Revenue websites & publications	1990-2006	BMI, % obese, % overweight	Self-reported height & weight, adjusted using NHANES data (to correct self-report bias)	1990-2006		~10%
Various food items											
Asfaw (2007)	Egypt	>2,000 households	Mothers. Nationally representative	Food prices	Average price per 100kcal of 9 foods: baladi bread, sugar, oil, rice, fruits, vegetables, egg & milk, beef, pulses	1997	BMI	EIHS: 7d recall & anthropometry	1997		

Table 2.5 (cont'd) Study characteristics: studies linking food prices and body weight

Ref	Country	Sample size	Sample characteristics	Exposure	Exposure details	Year(s) of exposure data collection	Outcome(s)	Outcome assessment details	Year(s) of outcome data collection	Length to follow-up	Loss to follow-up
Miljkovic et al. (2008)	USA (California, Idaho, Texas, Minnesota, Michigan)	55,550 observations	Adults (mean age 46 yrs; 43% normal BMI, 40% overweight, 17% obese)	Past, current & future prices of sugar beet, potatoes and milk	State- and month-specific prices obtained from USDA National Agricultural Statistics Service (NASS); deflated to 1989	1990-1992, 1996-1998, 2001-2003	BMI category (normal, overweight, obese)	Self-reported height & weight.	1991, 1997, 2002	N/A	N/A
Duffey et al. (2010)	USA	11,972	CARDIA study, nationally representative, 18-30yrs	Food prices: soft drink, whole milk, hamburger, pizza	From Council for Community & Economic Research (C2ER) data; adjusted using CPI to 2006. Linked to cohort temporally & spatially	1985-1986, 1992-1993 & 2005-2006	Body weight (lb)	Measured by trained technician	1985-1986, 1992-1993 & 2005-2006	0, 7 and 20 years	19%, 28% (of original sample)
Staudigel (2011)	Russia	(full) 25,008 (analytical) 10,551	RLMS; adults; nationally representative	Food prices	Average prices (from high & low) for 20 common items, measured in RLMS	1994-2005	BMI	Measurement not reported	1994-2005	Mostly annually	6,307 only respond to 1 wave
Lear et al. (2013)	Canada	555	Opportunity samples from 5 supermarkets; adults	Food basket prices	Total basket cost (selecting cheapest goods) for: milk, bananas, tomatoes, eggs, rice, flour, sugar, bread.	Not specified	BMI	Self-report height & weight at time of survey	Not specified	N/A	

MTF Monitoring the Future Survey; NLSY97 National Longitudinal Survey of Youth; NHANES National Health Examination & Nutrition Survey; ECLS-K Early Childhood Longitudinal Study - Kindergarten cohort; BRFSS Behavioral Risk Factor Surveillance System; CSFII Continuing Survey of Food Intakes by Individuals; PSID Panel Study of Income Dynamics; CARDIA Coronary Artery Risk Development In Young Adults; RLMS Russia Longitudinal Monitoring Survey; EIHS Data from Egyptian Integrated Household Survey

Table 2.6 Results of studies linking food prices and body weight

Ref	Exposure	Outcome(s)	Comparison	Statistical treatment	Adjustments	Results	p value	Summary of results
Studies in children and adolescents								
Price indices for fast food, food at home, food away from home, fruit & vegetables								
Powell et al. (2007)	Price of fast food	BMI		OLS regression	Sex, grade, ethnicity, parental education, urbanicity, student income, student employment, maternal employment, physical activity, restaurant density	Coefficient: -0.3066 (SE 0.1397)	<0.05	The price of fast food is significantly negatively associated with BMI and % overweight.
	Price of fast food	Overweight (1=yes)		Maximum likelihood probit model	As above	Coefficient: -0.0224 (SE 0.0097)	<0.05	
	Price of fruit & veg	BMI		OLS regression	As above	Coefficient: 0.2688 (SE 0.2392)	ns	There was no significant association between fruit and vegetable prices and BMI or overweight.
	Price of fruit & veg	Overweight (1=yes)		Maximum likelihood probit model	As above	Coefficient: -0.0049 (SE 0.0153)	ns	
Auld and Powell (2009)	Price of fruit & veg	BMI		OLS model	Restaurant/supermarket density; poverty rate; per capita income; race; urbanicity; sex; mother employment; age; parental education	Coefficient: 0.6364 (t-ratio 2.72): males 0.374 (1.05); females 0.8640 (2.99)	Not reported	A positive and statistically significant effect was found, with a stronger association amongst females
		Overweight incidence		Probit model	As above	Coefficient: 0.0229 (t ratio 1.54): males 0.0402 (1.59); females 0.0104 (0.61)	Not reported	A positive association was found but this did not achieve statistical significance
	Price of fast food	BMI		OLS model	As above	Coefficient: -0.2555 (t ratio -1.90): males -0.2346 (-1.21); females -0.2583 (-1.50)	Not reported	There was a negative association between fast food price and BMI, but with only a marginal statistical significance
	Price of fast food	Overweight incidence		Probit model	As above	Coefficient: -0.0189 (t ratio -2.02): males -0.0205 (-1.43); females -0.0168 (-1.58)	Not reported	Each additional $1 is associated with a 2% decline in the prevalence of overweight

Table 2.6 (cont'd) Results of studies linking food prices and body weight

Ref	Exposure	Outcome(s)	Comparison	Statistical treatment	Adjustments	Results	p value	Summary of results
Powell (2009)	Price of fast food	BMI		Cross-sectional OLS model	# restaurants & food stores; ethnicity; living arrangement; parental income; adolescent income; maternal education & working hours; urbanicity	Coefficient: -0.7782 (SE 0.4281)	>0.05	An inverse relationship between fast food prices and BMI was evident in both models, however this only achieved statistical significance in the longitudinal analysis (where a $1 increase was estimated to reduce adolescent BMI by 0.646 units
	Price of fast food	BMI		Longitudinal individual-level fixed-effects model	As above	Coefficient: -0.6455 (SE 0.2979)	<0.05	
	Price of food at home	BMI		Cross-sectional OLS model	As above	Coefficient: -0.2187 (SE 0.7655)	>0.05	The negative relationship between food-at-home prices and BMI was not found to be statistically significant in either model
	Price of food at home	BMI		Longitudinal individual-level fixed-effects model	As above	Coefficient: -0.0807 (SE 0.7641)	>0.05	
Powell and Bao (2009)	Price of fruit & veg	BMI		Multivariate random effects model	# restaurants & stores; county-level income, ethnicity, gender, birthweight, breastfed, mother obesity, maternal marital status, maternal education, mother's work hrs, family income, urbanicity	Coefficient: 2.0143 (SE 0.7491)	<0.01	A significant positive association was found between the price of fruit and veg and children's BMI. Each $1 increase was estimated to increase BMI by 2 units (or 10% and 0.7% in % terms)
	Price of fast food	BMI		Multivariate random effects model	As above	Coefficient: -0.5068 (SE 0.3538)	ns	The negative association was not found to be statistically significant
Black et al. (2013)	88% subsidy on fruit & veg	Proportions of weight categories	Before & after subsidy	Stuart-Maxwell test	None	X^2 [3,125] = 1.33	0.721	The fruit and vegetable subsidy program was not associated with changes in body weight or fatness
	88% subsidy on fruit & veg	% body fat	Before & after subsidy	Paired t test or GLM regression	Unclear	22.5% vs 22.1%; test statistic not reported	ns	

Table 2.6 (cont'd) Results of studies linking food prices and body weight

Ref	Exposure	Outcome(s)	Comparison	Statistical treatment	Adjustments	Results	p value	Summary of results
Soft drinks								
Fletcher et al. ()	State soft drink tax	BMI z-score		Not stated	N/A	Mean BMI z-score: 0.427 vs 0.418	0.696	State soft drink taxes did not appear to be associated with children's BMI or proportions overweight or obese
	State soft drink tax	Obese		Not stated	N/A	Proportions: 0.148 vs 0.150	0.819	
	State soft drink tax	Overweight or obese		Not stated	N/A	Proportions: 0.297 vs 0.302	0.611	
Fletcher et al. (2009)	(Net) soft drink tax rate	Change in BMI z-score		OLS regression	Year, quarter, state	Coefficient: 0.015 (SE 0.016)	ns	State soft drink tax rates were not associated with changes in children's BMI or proportions overweight or obese
	(Net) soft drink tax rate	Change in % BMI categories		OLS regression	Year, quarter, state	Coefficients: Obese 0.009 (SE 0.006), overweight 0.002 (0.011), under-weight -0.002 (0.003)	All ns	
Sturm et al. (2010)	Higher soda tax amount	BMI change		OLS regression	Age; ethnicity; sex; family income; mother's education; physical activity; weekly TV; parent-child interaction; birth weight	Coefficient: -0.013	ns	The presence of a soda tax was associated with lower BMI, but the actual soda tax amount was not associated with BMI
	Higher soda tax indicator	BMI change		As above	As above	Coefficient: -0.085	<0.05	
	Higher soda tax amount	BMI change	At risk of overweight	As above	As above	Coefficient: -0.033	<0.05	Amongst children at risk of overweight, both the soda tax indicator and soda tax amount were associated with lower BMI
	Higher soda tax indicator	BMI change	At risk of overweight	As above	As above	Coefficient: -0.222	<0.05	
	Higher soda tax amount	BMI change	Family income <$25,000	As above	As above	Coefficient: -0.000	ns	Amongst low-income families, soda taxes were not significantly associated with children's BMI
	Higher soda tax indicator	BMI change	Family income <$25,000	As above	As above	Coefficient: -0.005	ns	
	Higher soda tax amount	BMI change	African American	As above	As above	Coefficient: 0.029	ns	Soda taxes were not associated with children's BMI amongst African Americans

Table 2.6 (cont'd) Results of studies linking food prices and body weight

Ref	Exposure	Outcome(s)	Comparison	Statistical treatment	Adjustments	Results	p value	Summary of results
Various food items								
Sturm and Datar (2005)	Fruit & veg price index (FVPI)	BMI change, KG-3rd grade		Two-level random effects model	Age, sex, family income, ethnicity, maternal education, physical activity, TV viewing, birthweight	Coefficient 0.114 (SE 0.033)	<0.001	Increasing the FVPI by 1SD was associated with a 0.11 increase in BMI unit, and highly significantly so
	Meats price index	BMI change, KG-3rd grade		Two-level random effects model	As above	Coefficient -0.025 (SE 0.031)	0.414	None of the other food group indices examined were significantly associated with BMI change
	Dairy price index	BMI change, KG-3rd grade		Two-level random effects model	As above	Coefficients etc not reported	ns	
	Fast food price index	BMI change, KG-3rd grade		Two-level random effects model	As above	Coefficients etc not reported	ns	
Sturm and Datar (2008)	Fruit & veg price index	BMI change, KG-5th grade		Two-level random effects model	Age, sex, family income, ethnicity, maternal education, physical activity, TV viewing, birthweight	Coefficient 0.182 (SE 0.045)	<0.001	The 5-year follow-up to (351) found similar results, with a 0.18 unit increase in BMI in response to a standard deviation rise in the price index of fruit and vegetables, but not meats
	Meats price index	BMI change, KG-5th grade		Two-level random effects model	As above	Coefficient 0.076 (SE 0.043)	0.078	
Studies in both adults and children								
Various food items								
Thomas et al. (1996)	Community price of all foods	Weight for height, children		Two-level regression	Age, urban/rural, health facilities, education, household composition	Wald statistics: urban 71.52, rural 111.70, all 90.48	All <0.01	A lower BMI in adults is associated with higher food prices, in general. Relationships are stronger in the rural subgroups, which exhibited statistically significant Wald statistics, in contrast to urban.
	Community price of all foods	log(BMI)		Two-level regression	As above	X2 Wald test: all 34.25, urban male 9.37, urban female 12.07, rural male 38.42, rural female 21.55	<0.01, 0.31, 0.15, <0.01, 0.01	

Table 2.6 (cont'd) Results of studies linking food prices and body weight

Ref	Exposure	Outcome(s)	Comparison	Statistical treatment	Adjustments	Results	p value	Summary of results
Studies in adults								
Price indices for fast food, food at home, food away from home, unhealthy food, fruit & vegetables								
Chou et al. (2004)	Fast food restaurant price	BMI		Multivariate regression	Ethnicity; sex; education; marital status; hhold income; age; cigarette & alcohol price; restaurant density	-1.216 (t ratio -1.67)	Not reported	A negative and statistically significant association was evident for each relationship for both outcomes, with the largest estimates reported for food-at-home prices
	Fast food restaurant price	Incident obesity		Logistic regression	As above	-0.034 (t ratio -0.58)	Not reported	
	Full-service restaurant price	BMI		Multivariate regression	As above	-0.687 (t ratio -4.28)	Not reported	
	Full-service restaurant price	Incident obesity		Logistic regression	As above	-0.047 (t ratio -3.83)	Not reported	
	Food at home price	BMI		Multivariate regression	As above	-6.462 (t ratio -3.37)	Not reported	
	Food at home price	Incident obesity		Logistic regression	As above	-0.530 (t ratio -4.28)	Not reported	
Beydoun et al. (2008)	FFPI	BMI		Multivariate linear regression	Age, gender, ethnicity, education, urbanicity, survey year, smoking, physical activity, self-rated health	Coefficient (SEE): 0.6 (1.0)	ns	The fast food price index was not significantly associated with BMI or obesity. Every additional $1 on the fruit and vegetable price index was associated with 3.9kg/m^2 lower BMI. When PIR tertiles were examined separately, the negative association was significant only amongst the near poor. Prices of fruit and vegetables were also associated with lower odds of being obese, but only amongst the near poor.
	FVPI	BMI		As above	As above	Coefficient (SEE): -3.9 (1.8)	<0.05	
	FFPI	BMI	Poverty income ratio (PIR) tertiles	As above	As above	Coefficients (SEE) (poor; near poor; non-poor): 3.6 (1.7), 0.4 (1.8), -0.3 (1.1)	all ns	
	FVPI	BMI	Poverty income ratio (PIR) tertiles	As above	As above	Coefficients (SEE) (poor; near poor; non-poor): -9.8 (5.7), -6.8 (2.8), -0.8 (2.0)	ns, <0.05, ns	
	FFPI (z-score)	Incident obesity		Logistic regression	As above	OR (95% CI): 1.07 (0.88, 1.31)	ns	
	FVPI (z-score)	Incident obesity		As above	As above	OR (95% CI): 0.88 (0.76, 1.04)	ns	

Table 2.6 (cont'd) Results of studies linking food prices and body weight

Ref	Exposure	Outcome(s)	Comparison	Statistical treatment	Adjustments	Results	p value	Summary of results
Beydoun et al. (2008) cont'd	FFPI (z-score)	Incident obesity	Poverty income ratio (PIR) tertiles	As above	As above	OR (95% CI) (poor; near poor; non-poor): 1.18 (0.87, 1.59), 1.04 (0.77, 1.40), 1.04 (.080, 1.36)	all ns	
	FVPI (z-score)	Incident obesity	Poverty income ratio (PIR) tertiles	As above	As above	OR (95% CI) (poor; near poor; non-poor): 0.77 (0.51, 1.17), 0.82 (0.67, 0.99), 0.95 (0.74, 1.22)	ns, <0.05, ns	
Schroeter et al. (2005)	Fast food restaurant price	BMI		Quadratic equation	Ethnicity, gender, education, marital status, income, age, alcohol price	Estimate -2.455 (t-value -7.33)	<0.01	Fast food restaurant prices were inversely related to BMI/weight, significantly so in two of the models, but not the log-linear model
	ln(food away from home price)	ln(weight)		Log-linear model	Ethnicity, gender, education, marital status, income, age, alcohol price, physical activity, F&V consumption, region	Estimate -0.044 (t-value -1.86)	ns	
	ln(food away from home price)	ln(weight)		Trans log model	As above	Estimate -66.644 (t-value -7.36)	<0.01	
	Food at home price	BMI		Quadratic equation	Ethnicity, gender, education, marital status, income, age, alcohol price	Estimate -3.860 (t-value -6.92)	<0.01	A statistically significant relationship was found between food-at-home prices and BMI or weight in all models, however the direction of effect differed by model
	ln (Food at home price)	ln(weight)		Log-linear model	Ethnicity, gender, education, marital status, income, age, alcohol price, physical activity, F&V consumption, region	Estimate 0.114 (t-value 4.90)	<0.01	
	ln (Food at home price)	ln(weight)		Trans log model	As above	Estimate -14.592 (t value -7.73)	<0.01	
Powell and Han (2011)	Fast food price	BMI		OLS regression	Race, age, age^2, zip code, number of children, price match quality, urbanization, median area-level household income, education, year	Coefficients: Men -0.2090 (SE 0.3309); Women -0.1612 (0.4180)	ns	Fast food prices were not associated with BMI in any of the models. Fruit and vegetable prices were not associated with BMI in the OLS regression. In longitudinal fixed effects models, fruit and vegetable prices were
	Fruit & veg price	BMI		OLS regression	As above	Coefficients: Men 0.1938 (SE) 4909); Women 0.7623 (0.5622)	ns	
	Fast food price	BMI		Longitudinal individual fixed effects model	Number of children, price match quality, urbanization, median area-level household income, education, year	Coefficients: Men 0.0724 (SE 0.1693); Women 0.2622 (0.2216)	ns	

Table 2.6 (cont'd) Results of studies linking food prices and body weight

Ref	Exposure	Outcome(s)	Comparison	Statistical treatment	Adjustments	Results	p value	Summary of results
Powell and Han (2011) cont'd	Fast food price	BMI	Subgroups: Below or above 130% poverty line	Longitudinal individual fixed effects model	As above	Coefficients: Poor men -0.2981 (SE0.9621); non-poor men 0.1307 (0.1711); poor women -0.159 (0.839); non-poor women 0.161 (0.232)	All ns	significantly positively associated with BMI only amongst women – in particular poor women, or those with children.
	Fast food price	BMI	Subgroups: with children or none	Longitudinal individual fixed effects model	As above	Coefficients: men without children 0.0156 (SE 0.2663); men with children 0.1309 (0.2488); women without children -0.0311 (0.3536); women with children 0.4053 (0.3126)	All ns	
	Fruit & veg price	BMI		Longitudinal individual fixed effects model	As above	Men 0.2744 (SE 0.2738); Women 0.6173 (0.3083)	Ns; <0.05	
	Fruit & veg price	BMI	Subgroups: Poor vs non-poor (130% federal poverty line)	Longitudinal individual fixed effects model	As above	Coefficients: Poor men -1.0617 (SE3.2861); non-poor men 0.3684 (0.2508); poor women 3.5553 (1.3703); non-poor women 0.3970 (0.3111)	Poor women <0.01; all other ns	
	Fruit & veg price	BMI	Subgroups: with children or none	Longitudinal individual fixed effects model	As above	Coefficients: men without children 0.1521 (SE0.3993); men with children 0.5454 (0.4177); women without children -0.1859 (0.5592); women with children 1.0950 (SE 0.4009)	Women with children <0.01; all other ns	

Table 2.6 (cont'd) Results of studies linking food prices and body weight

Ref	Exposure	Outcome(s)	Comparison	Statistical treatment	Adjustments	Results	p value	Summary of results
Zhang et al. (2011)	Unhealthy food price indices: UFP1, UFP2, UFP3	BMI		Fixed effects model	Age, family size, income, urbanization, region, marital status, food stamp participation	Coefficients: UFP1 0.06 (SE 0.05) UFP2 -0.01 (0.02) UFP3 -0.01 (0.02)	0.23, 0.92, 0.75	UFP1 (sandwich, pizza, fried chicken) was not significantly associated with BMI. UFP2 (UFP1 + soft drink, beef, steak, sausage) andUFP3 (EFP2 + margarine, sugar, potatoes) were significantly negatively associated with BMI, but only in the 2-stage random effects model.
	Unhealthy food price indices: UFP1, UFP2, UFP3	BMI		Random effects model	As above	Coefficients: UFP1 0.04 (SE 0.05) UFP2 -0.01 (0.02) UFP3 -0.01 (0.02)	0.36, 0.54, 0.43	
	Unhealthy food price indices: UFP1, UFP2, UFP3	BMI		2-stage fixed effects model	Age, income, urbanization, region, marital status, food stamp participation	Coefficients: UFP1 -0.03 (SE 0.03) UFP2 -0.05 (0.03) UFP3 -0.04 (0.03)	0.36, 0.09, 0.12	
	Unhealthy food price indices: UFP1, UFP2, UFP3	BMI		2-stage random effects model	As above	Coefficients: UFP1 -0.03 (SE 0.14) UFP2 -0.05 (0.01) UFP3 -0.05 (0.02)	0.78, <0.001, <0.001	
	Unhealthy food price indices: UFP1, UFP2, UFP3	Obesity		Fixed effects model	Age, family size, income, urbanization, region, marital status, food stamp participation	Coefficients: UFP1 0.89 (95% CI 0.60-1.32); UFP2 0.84 (0.76-0.94); UFP3 0.86 (0.79-0.95)	Not reported (see 95% CI)	UFP1 (sandwich, pizza, fried chicken) was not associated with odds of obesity. UFP2 (UFP1 + soft drink, beef, steak, sausage) andUFP3 (EFP2 + margarine, sugar, potatoes) were associated with a significantly reduced odds of obesity in all models.
	Unhealthy food price indices: UFP1, UFP2, UFP3	Obesity		Random effects model	As above	Coefficients: UFP1 0.83 (95% CI 0.59-1.16); UFP2 0.94 (0.89-1.00); UFP3 0.95 (0.90-0.99)	Not reported (see 95% CI)	
	Unhealthy food price indices: UFP1, UFP2, UFP3	Obesity		2-stage fixed effects model	Age, income, urbanization, region, marital status, food stamp participation	Coefficients: UFP1 0.81 (95% CI 0.31-2.08); UFP2 0.77 (0.68-0.88); UFP3 0.80 (0.77-0.84)	Not reported (see 95% CI)	
	Unhealthy food price indices: UFP1, UFP2, UFP3	Obesity		2-stage random effects model	As above	Coefficients: UFP1 0.87 (95% CI 0.70-1.08); UFP2 0.90 (0.85-0.96); UFP3 0.91 (0.89-0.95)	Not reported (see 95% CI)	

Table 2.6 (cont'd) Results of studies linking food prices and body weight

Ref	Exposure	Outcome(s)	Comparison	Statistical treatment	Adjustments	Results	p value	Summary of results
Cotti and Tefft (2013)	Fast food price index	BMI		OLS regression	Sex, age, race/ethnicity, income, employment, education, food retail outlet availability, state food stamp uptake, state food tax rate	Coefficient: -0.80	ns	There were no significant associations between BMI or obesity and fast food or food-at-home prices, regardless of the analytical approach used.
	Fast food price index	Obesity		OLS regression	As above	Coefficient: -0.006	ns	
	Fast food price index	BMI		OLS regression, with lagged variables	As above		ns	
	Fast food price index	Obesity		OLS regression, with lagged variables	As above		ns	
	Food-at-home price index	BMI		OLS regression	As above	Coefficient: -0.021	ns	
	Food-at-home price index	Obesity		OLS regression	As above	Coefficient: -0.001	ns	
	Fast food price index	BMI		2-stage least squares regression	Sex, age, race/ethnicity, income, employment, education, food retail outlet availability, state food stamp uptake, state food tax rate, indicator variables for county, year & quarter	Coefficient: 0.165	ns	
	Fast food price index	Obesity		2-stage least squares regression	As above	Coefficient: 0.002	ns	
Soft drinks								
Kim and Kawachi (2006)	Tax or absence/repeal of tax on soft drinks	Incidence of high (>75th percentile) rate of obesity rate increase	Reference group: states with a 5% tax	Multivariate-adjusted odds ratio	State median age, mean income, racial proportions, political party at 1992 elections	States without tax: OR 4.2 (CI 0.4-48.3); States with repealed tax: OR 13.3 (CI 0.7-262)	0.25; 0.09	States with no tax in place appeared four times as likely to have experienced a high rate of obesity rate increase; and those which had repealed a tax were reported 13 times as likely. Confidence intervals were wide, however, and neither findings achieved significance

Table 2.6 (cont'd) Results of studies linking food prices and body weight

Ref	Exposure	Outcome(s)	Comparison	Statistical treatment	Adjustments	Results	p value	Summary of results
Fletcher et al. (2010a)	Total soft drink tax rate	BMI		2-way fixed effects OLS framework	State, year, quarter, race, income, 1yr-lagged state unemployment, state cigarette tax	Coefficient: -0.0029	<0.01	The tax rate was associated with a significant but small decrease in BMI: a 1% increase was associated with a decrease of 0.003 units
	Total soft drink tax rate	% overweight, % obese		2-way fixed effects OLS framework	As above	Coefficients: obese -0.0001, overweight -0.0002	<0.1, <0.01	A 1% increase in total tax rate was associated with a decrease in obesity of 0.01% and in overweight of 0.02%. The latter relationship was statistically significant
	Incremental soft drink tax rate	BMI		2-way fixed effects OLS framework	As above	Coefficient: -0.0028	<0.01	The incremental tax rate was similarly associated to the outcomes as above; this time statistical significance was achieved in all cases
	Incremental soft drink tax rate	% overweight, % obese		2-way fixed effects OLS framework	As above	Coefficients: obese -0.0001, overweight -0.0002	<0.05, <0.01	
	Incremental soft drink tax rate	BMI	Income category	2-way fixed effects OLS framework	As above	Coefficients: <$10k -0.0153, $10-<$15k -0.0130, $15-<$20k -0.0099, $20-<$25k 0.0117, $25-<$35k 0.0032, $35-<$50k -0.0059, $50k+ -0.0081	<0.01, <0.01, <0.01, <0.01, <0.05, <0.01, <0.01	BMI was negatively associated with the soft drink tax rate at the tails of the income distribution (below $20k and above $35k), but positively so around the middle of the distribution. All results were statistically significant. A similar pattern was evident with the other outcomes, although degrees of significance varied
	Incremental soft drink tax rate	% obese	Income category	2-way fixed effects OLS framework	As above	Coefficients: <$10k -0.0008, $10-<$15k -0.0005, $15-<$20k -0.0008, $20-<$25k 0.0001, $25-<$35k 0.0002, $35-<$50k -0.0001, $50k+ -0.0005	<0.01, <0.01, <0.01, ns, <0.05, ns, <0.01	

Table 2.6 (cont'd) Results of studies linking food prices and body weight

Ref	Exposure	Outcome(s)	Comparison	Statistical treatment	Adjustments	Results	p value	Summary of results
Fletcher et al. (2010a) cont'd	Incremental soft drink tax rate	% overweight	Income category	2-way fixed effects OLS framework	As above	Coefficients: <$10k -0.0010, $10-<$15k -0.0005, $15-<$20k 0.0003, $20-<$25k 0.0002, $25-<$35k 0.0006, $35-<$50k -0.0005, $50k+ -0.0008	<0.01, <0.01, <0.01, <0.05, <0.01, <0.01, <0.01	
	Incremental soft drink tax rate	BMI	Sex	2-way fixed effects OLS framework	As above	Coefficients: Female -0.0040, male -0.0009	<0.01, <0.05	Small, but statistically significant coefficients were found for BMI in association with soft drink tax rate in both sexes
	Incremental soft drink tax rate	% obese	Sex	2-way fixed effects OLS framework	As above	Coefficients: Female 0.0000, male -0.0001	ns	No significant association between soft drink taxes and obesity prevalence was evident
	Incremental soft drink tax rate	% overweight	Sex	2-way fixed effects OLS framework	As above	Coefficients: Female -0.0005, male 0.0001	<0.01, <0.05	A modest and significant negative association was seen amongst females, whereas a small (although still significant) positive association was seen for males
	Incremental soft drink tax rate	BMI	Ethnicity	2-way fixed effects OLS framework	As above	Coefficients: Black -0.0012, white -0.0026, Hispanic -0.0164	ns, <0.01, <0.01	None of the outcomes were significantly associated with taxes amongst blacks; all were significant and negative for Hispanics; whilst amongst the white subgroup small significant relationships were apparent in considering BMI and overweight prevalence, but not obesity prevalence
	Incremental soft drink tax rate	% obese	Ethnicity	2-way fixed effects OLS framework	As above	Coefficients: Black -0.0001, white 0.0000, Hispanic -0.0021	ns, ns, <0.01	
	Incremental soft drink tax rate	% overweight	Ethnicity	2-way fixed effects OLS framework	As above	Coefficients: Black 0.0001, white -0.0002, Hispanic -0.0022	ns, <0.01, <0.01	
	Incremental soft drink tax rate	BMI	Education	2-way fixed effects OLS framework	As above	Coefficients: High school -0.0031, college -0.0076	<0.01, <0.01	Significant negative associations were apparent regardless of education classification; larger coefficients were observed for the college-educated subgroup
	Incremental soft drink tax rate	% obese	Education	2-way fixed effects OLS framework	As above	Coefficients: High school -0.0002, college -0.0004	<0.01, <0.01	

Table 2.6 (cont'd) Results of studies linking food prices and body weight

Ref	Exposure	Outcome(s)	Comparison	Statistical treatment	Adjustments	Results	p value	Summary of results
Fletcher et al. (2010a) cont'd	Incremental soft drink tax rate	% overweight	Education	2-way fixed effects OLS framework	As above	Coefficients: High school 0.0002, college -0.0004	<0.01, <0.01	
	Incremental soft drink tax rate	BMI	Age	2-way fixed effects OLS framework	As above	Coefficients: >65yrs -0.0038, 18-25yrs 0.0022, 25-40yrs -0.0032, 40-65yrs -0.0037	<0.01, ns, <0.01, <0.01	Significant negative associations were found for the age groups 25 years and older, but not for the younger age group (18 to 25)
	Incremental soft drink tax rate	% obese	Age	2-way fixed effects OLS framework	As above	Coefficients: >65yrs -0.0001, 18-25yrs 0.0000, 25-40yrs -0.0001, 40-65yrs 0.0000	All ns	No trend was obvious amongst any age group in terms of obesity prevalence
	Incremental soft drink tax rate	% overweight	Age	2-way fixed effects OLS framework	As above	Coefficients: >65yrs -0.0002, 18-25yrs 0.0001, 25-40yrs -0.0005, 40-65yrs -0.0001	ns, ns, <0.01, ns	The only significant association between soft drink taxes and overweight prevalence was found in the 25-40 age group
Various food items								
Asfaw (2007)	Average price per 100g of baladi bread	BMI	Primary sampling units	Modelling	Age, education, family size, urbanicity, expenditure, & for clustering (Huber-White sandwich estimators)	Coefficient: -0.119 (0.047)	p<0.05	A 1% increase in price of baladi bread is associated with a 0.12% reduction in BMI units
	Average price per 100g of sugar				As above	Coefficient: -0.112 (0.054)	p<0.05	A 1% increase in price of sugar is associated with a 0.11% reduction in BMI units
	Average price per 100g of oil				As above	Coefficient: -0.102 (0.062)	ns	There was a small but not statistically significant inverse relationship between the price of oil and BMI
	Average price per 100g of rice				As above	Coefficient: -0.203 (0.074)	p<0.01	A 1% increase in price of rice is associated with a 0.20% reduction in BMI units
	Average price per 100g of fruits				As above	Coefficient: 0.090 (0.037)	p<0.05	A 1% increase in price of fruits is associated with a 0.09% lower BMI

Table 2.6 (cont'd) Results of studies linking food prices and body weight

Ref	Exposure	Outcome(s)	Comparison	Statistical treatment	Adjustments	Results	p value	Summary of results
Asfaw (2007) cont'd	Average price per 100g of vegetables				As above	Coefficient: -0.004 (0.044)	ns	There was no significant relationship found between price of vegetables and BMI
	Average price per 100g of eggs & milk				As above	Coefficient: 0.137 (0.045)	p<0.01	A 1% decrease in price of eggs & milk is associated with a 0.14% reduction in BMI units
	Average price per 100g of beef				As above	Coefficient: 0.074 (0.101)	ns	There was no significant relationship found between price of beef and BMI
	Average price per 100g of pulses				As above	Coefficient: -0.001 (0.064)	ns	There was no significant relationship found between price of pulses and BMI
Miljkovic et al. (2008)	Price of sugar beet	Overweight or obese categories		Multinomial logit model	Age, income, education, sex, time, region, race, F&V consumption, historical & future prices (sugar beet, potatoes, milk)	Coefficients: overweight -0.23 (SE 0.01); obese -0.34 (SE0.01)	<0.01	An increase in the price of sugar beet significantly decreases the probability of being overweight or obese
	Price of potatoes	Overweight or obese categories		Multinomial logit model	As above	Coefficients: overweight 0.03 (SE 0.01); obese 0.06 (SE0.01)	<0.01	An increase in the price of potatoes significantly increases the probability of being overweight or obese
	Price of milk	Overweight or obese categories		Multinomial logit model	As above	Coefficients: overweight -0.06 (SE 0.01); obese -0.30 (SE0.02)	<0.01	An increase in the price of milk significantly decreases the probability of being overweight or obese
Duffey et al. (2010)	Soda price	Bodyweight (lb)		Pooled OLS regression	Study centre, age, race, sex, education, household income, family structure, time of data collection	Coefficient: -2.3 (SE 0.8)	<0.05	The prices of soda and pizza were negatively associated with bodyweight: every $1 increase was associated with 2.3lb and 1.3lb lower weight respectively. Whole milk and burger prices were not significantly associated with weight.
	Whole milk price	Bodyweight (lb)		As above	As above	Coefficient: -0.2 (SE 2.4)	ns	
	Burger price	Bodyweight (lb)		As above	As above	Coefficient: -0.4 (SE 1.9)	ns	
	Pizza price	Bodyweight (lb)		As above	As above	Coefficient: -1.3 (SE 1.9)	<0.05	

Table 2.6 (cont'd) Results of studies linking food prices and body weight

Ref	Exposure	Outcome(s)	Comparison	Statistical treatment	Adjustments	Results	p value	Summary of results
Staudigel (2011)*	Price of white bread	BMI		Fixed effects regression model	Age, marital status, work status, household size, pregnancy, education, year, community infrastructure, area median income	Coefficient: -0.0002 (SE 0.0014)	All ns	Prices of white bread, wheat flour, potatoes and cabbage were not found to be associated with BMI.
	Price of wheat flour	BMI		Fixed effects regression model	As above	Coefficient: 0.0010 (SE 0.0016)	All ns	
	Price of potatoes	BMI		Fixed effects regression model	As above	Coefficient: 0.0002 (SE 0.0014)	All ns	
	Price of cabbage	BMI		Fixed effects regression model	As above	Coefficient: -0.0013 (SE0.0017	All ns	
	Price of onions	BMI		Fixed effects regression model	As above	Coefficients: -0.0055 (0.0016), males -0.0030 (0.0027), females -0.0072 (0.0022), income tertile 1 -0.0021 (0.0023), tertile 2 -0.0082 (0.0017), tertile 3 -0.0062 (0.0025)	<0.01, ns, <0.01, ns, <0.01, <0.05	There was a significant negative association between the price of onions and BMI. This was not true of females only, nor the lowest income tertile.
	Price of oranges	BMI		Fixed effects regression model	As above	Coefficients: 0.0005 (SE 0.0028)	All ns	Prices of oranges and apples were not associated with BMI.
	Price of apples	BMI		Fixed effects regression model	As above	Coefficients: 0.0001 (SE 0.0011)	All ns	
	Price of beef	BMI		Fixed effects regression model	As above	Coefficients: 0.0014 (0.0034), males -0.0044 (0.0039), females 0.0053 (0.0044), income tertile 1 -0.0041 (0.0048), tertile 2 0.0028 (0.0040), tertile 3 0.0090 (0.0041)	Ns, ns, ns, ns, ns, <0.05	The price of beef was not associated with BMI, apart from amongst the highest income tertile, where there was a significant, small positive association.
	Price of pork	BMI		Fixed effects regression model	As above	Coefficients: -0.0045 (SE 0.0042)	All ns	Pork prices were not associated with BMI.

Table 2.6 (cont'd) Results of studies linking food prices and body weight

Ref	Exposure	Outcome(s)	Comparison	Statistical treatment	Adjustments	Results	p value	Summary of results
Staudigel (2011) cont'd	Price of chicken	BMI		Fixed effects regression model	As above	Coefficients: -0.0070 (0.0029), males -0.0079 (0.0040), females -0.0063 (0.0033), income tertile 1 -0.0049 (0.0032), tertile 2 -0.0074 (0.0040), tertile 3 -0.0096 (0.0047)	<0.05, ns, ns, ns, ns, <0.05	Chicken prices were negatively associated with BMI in the full sample, but in subgroup analyses were only significantly so amongst those with highest incomes.
	Price of sausages	BMI		Fixed effects regression model	As above	Coefficients: -0.0014 (0.0037), males 0.0023 (0.0050), females -0.0039 (0.0041), income tertile 1 0.0047 (0.0057), tertile 2 0.0022 (0.0060), tertile 3 -0.0117 (0.0052)	Ns, ns, ns, ns, ns, <0.05	No significant associated between sausage prices and BMI was apparent, except amongst the highest income tertile, where a negative association was apparent.
	Price of fresh milk	BMI		Fixed effects regression model	As above	Coefficients: -0.0032 (SE 0.0019)	All ns	The price of milk was not associated with BMI.
	Price of butter	BMI		Fixed effects regression model	As above	Coefficients: 0.0058 (0.0026), males 0.0110 (0.0033), females 0.0018 (0.0031), income tertile 1 0.0070 (0.0027), tertile 2 0.0014 (0.0044), tertile 3 0.0032 (0.0039)	<0.05, <0.01, ns, <0.05, ns, ns	BMI was significantly positively associated with the price of butter, but this was not the case for females or those in income tertiles 2 or 3.
	Price of cheese	BMI		Fixed effects regression model	As above	Coefficients: 0.0020 (SE 0.0016)	All ns	Prices of cheese, vegetable oil, sugar, cookies, fish and vodka were not associated with BMI.
	Price of vegetable oil	BMI		Fixed effects regression model	As above	Coefficients: 0.0012 (SE 0.0009)	All ns	
	Price of sugar	BMI		Fixed effects regression model	As above	Coefficients: 0.0028 (SE 0.0029)	All ns	
	Price of cookies	BMI		Fixed effects regression model	As above	Coefficients: -0.0007 (SE 0.0015)	All ns	

Table 2.6 (cont'd) Results of studies linking food prices and body weight

Ref	Exposure	Outcome(s)	Comparison	Statistical treatment	Adjustments	Results	p value	Summary of results
Staudigel (2011) cont'd	Price of fresh fish	BMI		Fixed effects regression model	As above	Coefficients: 0.0004 (SE 0.0011)	All ns	
	Price of vodka	BMI		Fixed effects regression model	As above	Coefficients: 0.0003 (SE (0.0012)	All ns	
	Price of white bread	Obesity		Logistic fixed effects regression	As above	Coefficients: 0.0000 (SE 0.0780)	All ns	Of all the analyses investigating food prices and obesity, a significant association was only found for the price of fresh milk amongst males. In all other cases, food prices did appear to be related to obesity.
	Price of wheat flour	Obesity		Logistic fixed effects regression	As above	Coefficients: 0.0012 (SE 0.0947)	All ns	
	Price of potatoes	Obesity		Logistic fixed effects regression	As above	Coefficients: -0.0345 (SE 0.0867)	All ns	
	Price of cabbage	Obesity		Logistic fixed effects regression	As above	Coefficients: 0.1213 (SE 0.0858)	All ns	
	Price of onions	Obesity		Logistic fixed effects regression	As above	Coefficients: -0.1467 (SE 0.0944)	All ns	
	Price of oranges	Obesity		Logistic fixed effects regression	As above	Coefficients: 0.0377 (SE 0.1620)	All ns	
	Price of apples	Obesity		Logistic fixed effects regression	As above	Coefficients: -0.0223 (SE 0.0827)	All ns	
	Price of beef	Obesity		Logistic fixed effects regression	As above	Coefficients: 0.1072 (SE 0.1936)	All ns	
	Price of pork	Obesity		Logistic fixed effects regression	As above	Coefficients: -0.2133 (SE 0.1846)	All ns	
	Price of chicken	Obesity		Logistic fixed effects regression	As above	Coefficients: 0.1227 (SE 0.1659)	All ns	
	Price of sausages	Obesity		Logistic fixed effects regression	As above	Coefficients: -0.0016 (SE 0.0014)	All ns	

Table 2.6 (cont'd) Results of studies linking food prices and body weight

Ref	Exposure	Outcome(s)	Comparison	Statistical treatment	Adjustments	Results	p value	Summary of results
Staudigel (2011) cont'd	Price of fresh milk	Obesity		Logistic fixed effects regression	As above	Coefficients: -0.1493 (SE 0.1002), males -0.4528 (0.1734), females -0.0039 (0.1187)	Ns, <0.01, ns	
	Price of butter	Obesity		Logistic fixed effects regression	As above	Coefficients: 0.2061 (SE 0.1277)	All ns	
	Price of cheese	Obesity		Logistic fixed effects regression	As above	Coefficients: 0.0155 (SE 0.1185)	All ns	
	Price of vegetable oil	Obesity		Logistic fixed effects regression	As above	Coefficients: -0.0385 (SE 0.0634)	All ns	
	Price of sugar	Obesity		Logistic fixed effects regression	As above	Coefficients: 0.1180 (SE 0.1890)	All ns	
	Price of cookies	Obesity		Logistic fixed effects regression	As above	Coefficients: 0.1179 (SE 0.0762)	All ns	
	Price of fresh fish	Obesity		Logistic fixed effects regression	As above	Coefficients: -0.0399 (SE 0.0581)	All ns	
	Price of vodka	Obesity		Logistic fixed effects regression	As above	Coefficients: 0.0245 (SE 0.0488)	All ns	
Lear et al. (2013)	Food basket price	BMI		Pearson correlation	None	R = -0.906	0.034	Supermarkets' food basket prices were negatively correlated with their shoppers' BMI.
	Food basket price	BMI		Multiple linear regression	Age, sex, car ownership, median income of residential area	Store 5 (most expensive basket) as comparator: Store 1: 3.66 (SE 0.94); Store 2: 3.73 (0.94); Store 3: 1.93 (0.88); Store 4: 1.52 (0.80)	<0.001; <0.001; 0.029; 0.057	The three supermarket with the least expensive baskets showed significantly higher BMI amongst their shoppers than the most expensive store. The two most expensive stores did not significantly differ.

* Subgroup coefficients presented only where significant associations observed.

2.4.4.2 Diet costs & body weight

Studies' designs and settings

Seven studies were found to investigate body weight in relation to diet costs or expenditure (Table 2.7). Four of these analysed data from cross-sectional surveys (Michaud et al., 2007, Murakami et al., 2007, Murakami et al., 2009, Lo et al., 2012), one was from an intervention study (Mushi-Brunt et al., 2007), and two (Rauber and Vitolo, 2009, Lopez et al., 2009) followed up longitudinal cohorts. Studies were based in a variety of countries – the US (Mushi-Brunt et al., 2007), Spain (Lopez et al., 2009), Brazil (Rauber and Vitolo, 2009), Taiwan (Lo et al., 2012), two from Japan (Murakami et al., 2007, 2009), and one cross-country study (Michaud et al., 2007) – and comprised a diverse range of populations, including children (two studies), elderly adults (two studies), and graduates/undergraduates (three studies). Sample sizes were generally large, and ranged from 354 to over 21,000.

Diet cost definition/assessment of expenditure

The majority of studies (n=5) estimated costs by matching national or supermarket prices to dietary data. Dietary assessment techniques included diet history questionnaire (DHQ) (Murakami et al., 2007, 2009), FFQ (Lopez et al., 2009) and 24-hour recall (Lo et al., 2012, Rauber and Vitolo, 2009). None of the studies matched prices to diet diary information. The number of food or beverage items priced ranged from 104 (Rauber and Vitolo, 2009) to 843 (Lo et al., 2012). Costs were expressed to a standardized energy amount (1000kcal or 1000kJ) in all but one of these studies (Lo et al., 2012) which utilised estimated daily costs for vegetables only.

Two of the studies used a measure of expenditure in the absence of dietary intake data. Expenditure was self-reported by participants either by telephone interview (Mushi-Brunt et al., 2007) or via questionnaire (Michaud et al., 2007). Mushi-Brunt et al employed household estimates of food expenditure; whereas Michaud et al calculated a measure of individual expenditure on food away from home relative to total reported expenditure.

Assessment of anthropometry

In one study, body weight was the exposure variable (Rauber and Vitolo, 2009); in all others it was the outcome. Four of the studies employed investigators or health professionals to measure anthropometry, whilst three relied upon participants' self-reports (Murakami et al., 2007, Lopez et al., 2009, Michaud et al., 2007). The latter study attempted to address self-report bias by adjusting their analyses to correct for bias.

Three studies included BMI (kg/m^2) as a continuous outcome in their analyses, one study investigated BMI categories only, and one study a binary outcome of obesity. The longitudinal study of Lopez et al also investigated change in body weight (kg), and Murakami et al (2009) additionally measured waist circumference. Both studies including children reported z-scores for BMI.

Analytical approaches

Given the variation in study design identified above, it is to be expected that the analytical approaches also vary. Table 2.8 details the analyses involved in each study. Four studies (Michaud et al., 2007, Lopez et al., 2009, Murakami et al., 2007, 2009) used multivariable regression techniques, adjusted for covariates. One study (Lo et al., 2012) reported only the results of a Chi^2 analysis, because BMI category was not the primary outcome of the study. The studies of Mushi-Brunt et al (2007) and Rauber and Vitolo (2009) used unadjusted comparisons (ANOVA and t test respectively).

Quality of studies

Unfortunately, the studies in this section of the literature review are considered to be of poorer quality on the whole than the studies in the other areas.

In working age adults, none of the studies (Murakami et al., 2007, Murakami et al., 2009, Lopez et al., 2009) used robust probability sampling methods. The sample of Lopez et al. (2009), however, was a longitudinal cohort, which has the advantage in terms of quality over the cross-sectional samples of Murakami et al. (2007) and Murakami et al. (2009). On the other hand, only Murakami et al. (2009) used objective measures of height and weight in adults, minimising the potential for self-report bias. The studies did not differ vastly in other aspects of study quality, such as statistical analyses, which were appropriately adjusted in all three studies, or price and dietary data collection methods, which, although different, are not yet established as differing in terms of quality.

In older adults, the studies (Michaud et al., 2007, Lo et al., 2012) were probably of better quality. Both used sophisticated sampling to create nationally representative samples of older adults, and were more than adequately powered with large samples. The study by Lo et al. (2012) was better quality in terms of data collection – using objectively measured anthropometry and applying national price data to 24-hour dietary recalls. On the other hand, the study of Michaud et al. (2007) used self-reported expenditure questionnaires and self-reported BMI. However, Michaud et al. (2007)

used more appropriate statistical analyses – multivariable regression (correcting for self-reported BMI) as opposed to Chi^2 analyses conducted by Lo et al. (2012).

The two studies in children (Mushi-Brunt et al., 2007, Rauber and Vitolo, 2009) again differed in terms of quality. Mushi-Brunt et al. (2007) used a non-probability sample, which allows the possibility of selection bias. Rauber and Vitolo (2009), on the other hand, used a longitudinal cohort, which is of stronger quality when trying to draw out causal evidence. However, this cohort was relatively small in size, which may result in an under-powered sample. The quality of both studies suffered in the analytical approaches taken: both using unadjusted comparison tests, which are unable to take into account confounding.

Findings

The studies using samples of female Japanese students (Murakami et al., 2007, Murakami et al., 2008b) reported small, but significant, negative associations between quintiles of diet cost and BMI or waist circumference.

The only longitudinal study to investigate diet costs and body weight (Lopez et al., 2009) found a relationship in the opposite direction to that of the studies by Murakami and colleagues (2007, 2009). Their results indicated that those with higher energy costs at baseline had significantly higher BMI at baseline, as well as significantly higher odds of weight gain over six years. However, the tendency towards higher odds of weight gain amongst those who had higher energy costs did not achieve statistical significance after adjusting for confounders.

Amongst the studies using elderly samples, findings were mixed. Lo et al. (2012) found that proportions in each BMI category differed by quintile of daily vegetable cost, with the lowest quintile containing the highest proportion of underweight and the lowest proportion of the most overweight category. In contrast, Michaud et al (2007) examined only the influence of food-away-from-home expenditure. After appropriate adjustments for confounding variables, the results indicated small negative coefficients in most subpopulations; however unadjusted correlations were positive, and p values were not reported in any scenario.

Neither of the studies involving children found a significant link between expenditure or diet cost and BMI percentile or risk of overweight.

Table 2.7 Study characteristics: studies investigating dietary expenditure/cost and body weight

Ref	Country	Sample size	Sample characteristics	Exposure	Exposure details	Year(s) of exposure data collection	Outcome	Outcome assessment details	Year(s) of outcome data collection	Length to follow-up	Loss to follow-up
Mushi-Brunt et al. (2007)	Missouri, USA	555 parent/ child dyads	Partners of All Ages Reading About Diet and Exercise (PARADE) intervention. Children aged 6-11yrs with a parent. 77% female; 65-71% African American	Household grocery expenditure	Telephone questionnaire	2000-2004	BMI/BMI percentile	Children: nurse-measured height & weight & CDC growth charts; Adults: Self-reported	2000-2004		Not reported
Michaud et al. (2007)	USA, Austria, Germany, Sweden, Nether-lands, Spain, Italy, France, Denmark, Greece	21,836	Drawn from Survey of Health, Ageing and Retirement Europe (SHARE) and Health & Retirement Study (US): nationally representative, 50yrs+	Expenditure on food away from home	Expenditure relative to total expenditure. Questionnaire-gathered self-reports.	2004	Incidence of obesity	Questionnaire-gathered self-reported height & weight. Corrected for self-report bias	2004		
Lopez et al. (2009)	Spain	19,057	Suguimiento Universidad de Navarra (SUN): prospective cohort of graduates, mean age 38.6yrs, 60% women	Daily food consumption costs (€/1000kcal)	Costs of foods derived from Ministry of Industry, Tourism & Commerce of Spain figures (n=136). Costs matched to baseline year. 18.3% prices taken from current supermarket websites. Matched to semi-quantitative FFQ	1999-2007	BMI; increase in bodyweight	Self-reported	1999-2007	2, 4 and 6 years	Retention rate 88%. After exclusions, 11,195
Rauber and Vitolo (2009)	Brazil	354	"Ten Steps in Action" (BRATSA I): children aged 3-4yrs, recruited 6-12mo at hospital	Risk of overweight (z score)	Professional-measured anthropometry	2005-2006	Mean expenditure (R$/1000kcal)	104 product prices taken from a large and a small retailer (means of 3 brands); corrected for waste/cooking. Estimated for 30 days. Matched to 2x 24hr recalls	2005-2006	4yrs	8

Table 2.7 (cont'd) Study characteristics: studies investigating dietary expenditure/cost and body weight

Ref	Country	Sample size	Sample characteristics	Exposure	Exposure details	Year(s) of exposure data collection	Outcome	Outcome assessment details	Year(s) of outcome data collection	Length to follow-up	Loss to follow-up
Murakami et al. (2007)	Japan	3931	Female dietetic students, 54 institutions	Dietary energy cost (yen/1000kcal)	National Retail Price Survey (n=122) applied to DHQ (135 items)	2005 (diet), 2004 (prices)	BMI	BMI from self-reports	2005	N/A	N/A
Murakami et al. (2008b)	Japan	1176	Female dietetic students, 15 institutions	Dietary energy cost (yen/1000kJ)	National Retail Price Survey (n=122) applied to DHQ (135 items)	2006/7 (diet), 2004 (prices)	BMI; waist circumference	Investigator measured	2006/7	N/A	N/A
Lo et al. (2012)	Taiwan	1911	50% male, adults 65yr+ from Elderly Nutrition and Health Survey	Daily cost of vegetables	Mean monthly prices obtained from national databases (n=628) + prices from supermarket (n=215) adjusted for inflation & applied to 24hr recall.	1999-2000 (diet), 1999-2000 & 2009 (prices)	BMI category	Physical examination	1999-2000	N/A	N/A

Table 2.8 Results: studies investigating dietary expenditure/cost and body weight

Ref	Exposure	Outcome(s)	Comparison/Subgroup	Statistical treatment	Adjustments	Results	p value	Summary of results
Mushi-Brunt et al. (2007)	Weekly household grocery spending	BMI/BMI percentile		ANOVA	None reported	F statistics not reported	ns	There was no significant difference in BMI between different household grocery spending levels.
Michaud et al. (2007)	Share of food expenditure on food away from home	Obesity incidence		Correlation	EI, % time eating out, time cooking, kcal/min eating	r=0.601 (excl US, r=-0.275)	Not reported	There was an apparent positive correlation between food expenditure away from home and obesity incidence. However, in Europe alone, the correlation appeared to be negative.
	Share of food expenditure on food away from home	Obesity incidence	Subgroup: US only	Logit regression	Age, income, hhold composition, marital status, smoking, education, wealth, physical activity	Point estimates: males: 0.241 (t = 1.93); females: -0.037 (-0.28)	Not reported	There appeared to be a negative relationship between food expenditure away from home and obesity in Europe and amongst American females, but positive amongst American males.
	Share of food expenditure on food away from home	Obesity incidence	Subgroup: Europe	Logit regression	Age, income, hhold composition, marital status, smoking, education, wealth, physical activity	Point estimates: males: -0.470 (t = -2.22); females: -0.566 (-3.15)	Not reported	
Lopez et al. (2009)	Daily food cost quintiles (€/1000kcal)	BMI		ANOVA	N/A	Q1 23 (SD 3.3); Q2 23.4 (3.4); Q3 23.6 (3.5); Q4 23.8 (3.5); Q5 24.2 (3.8)	<0.001	Those with higher daily food costs had a statistically significantly higher BMI at baseline.
	Daily food cost quintiles (€/1000kcal)	≥3kg weight gain within past 5yrs		Non-conditional logistic regression	Age, sex, EI, physical activity, smoking, snacking, alcohol, education, marital status, employment, dietary pattern scores	OR (95% CI): Q1 1 (ref); Q2 1.13 (1.02-1.26); Q3 1.06 (0.95, 1.19); Q4 1.14 (1.01, 1.29), Q5 1.13 (0.99, 1.29)	0.146	The tendency towards higher odds of weight gain amongst those who had higher food costs did not achieve statistical significance after adjusting for confounders.

Table 2.8 (cont'd) Results: studies investigating dietary expenditure/cost and body weight

Ref	Exposure	Outcome(s)	Comparison/Subgroup	Statistical treatment	Adjustments	Results	p value	Summary of results
	Daily food cost quintiles (€/1000kcal)	Average weight gain ≥0.6kg/yr		Non-conditional logistic regression	Age, sex, EI, physical activity, smoking, snacking, alcohol, education, marital status, employment, dietary pattern scores, baseline BMI	OR (95% CI): Q1 1 (ref); Q2 0.95 (0.83, 1.09), Q3 1.05 (0.92, 1.21), Q4 1.11 (0.96, 1.29), Q5 1.20 (1.02, 1.41)	0.007	Participants with the highest daily food costs had statistically significant higher odds of weight gain.
Rauber and Vitolo (2009)	Risk of overweight	Mean expenditure (R$/1000kcal)	No (≤1 z score) vs Yes (>1 z score)	t test	N/A	65.93 ± 14.55 vs 68.59 ± 20.17	0.208	No significant difference in expenditure per 1000kcal was found between children at risk of overweight and those not at risk.
Murakami et al. (2007)	Quintile of energy cost	BMI (kg/m^2)		Multivariable linear regression	PAL, Residence, Residential density, living status, smoking, alcohol, supplement, weight loss diet, rate of eating, EI.	Q1 21.1±0.1, Q2 21.1±0.1, Q3 20.9±0.1, Q4 21.0±0.1, Q5 20.8±0.1	0.0197	There was a significant p for trend between quintiles of energy cost, with a slight negative trend.
Murakami et al. (2008b)	Quintile of energy cost	BMI (kg/m^2)		Multivariable linear regression	Residence, residential density, living status, survey yr, smoking, weight loss, rate of eating, PAL	-0.38 (95% CI -0.60, -0.16)	0.0006	Every increase in energy cost quintile was associated with a lower BMI of 0.38kg/m^2 and a 1.46cm smaller waist circumference.
	Quintile of energy cost	Waist circumference (cm)		Multivariable linear regression	Residence, residential density, living status, survey yr, smoking, weight loss, rate of eating, PAL	-1.46 (95% CI -2.01, -0.90)	<0.0001	
Lo et al. (2012)	Quintile of daily cost of vegetables (NTD/d)	BMI category proportions	<18.5, 18.5-23.9, 24-26.9, ≥27	Chi^2	None (not primary outcome)	Q1 26.4%, 19.7, 16.4, 14.1; Q2 8.45%, 17.0, 17.7, 21.3; Q3 19.4%, 21.3, 19.1, 22.9; Q4 23.5%, 20.6, 22.8, 20.0; Q5 22.2%, 21.4, 24.0, 21.7	<0.001	Proportions of participants categorised in each BMI category differed between quintile of daily vegetable cost. The lowest quintile contained the highest proportion of underweight and the lowest proportion of the most overweight category.

2.4.4.3 Income & body weight - Existing reviews

The scoping search revealed four reviews regarding socioeconomic differences in BMI or obesity which reported on studies that investigated income separately. Three of these studies were systematic: the first of these (Sobal and Stunkard, 1989) was an extensive and comprehensive review of the literature to that date, which was updated in a comprehensive review in 2007 by McLaren (2007). The third systematic review, published most recently, was limited in scope to UK-based studies (El-Sayed et al., 2012). The remaining relevant review (Ball and Crawford, 2005) was semi-systematic in approach, and reported on the literature pertaining to weight *change* specifically.

The review of McLaren (2007) included 333 studies overall. Of these, 88 studies focussed on income and body weight in women, 78 studies reported on men, and 54 reported findings from men and women combined. The findings from these studies comprised results for 402 tested associations (which constituted 21% of all socioeconomic-body weight associations in the review). McLaren identified apparent differences in the results of studies depending upon the human development index (HDI) rating of the countries they were set in: in low- and mid-HDI settings, the majority of reported findings indicated a positive relationship between income and bodyweight. In high-HDI areas, for analyses of men and both sexes combined, the majority of findings were non-significant or curvilinear; in women, the majority of reported associations (49%) indicated a significant negative relationship, but 45% of associations were non-significant.

A predominantly negative finding in women between income and body weight agreed with the conclusions of Sobal et al in 1989, although the predominance was diminished in the more recent review of McLaren. Interestingly, McLaren found that, for some of the other socioeconomic indicators, such as education or occupation, there was in fact a predominance of negative associations. The author puts this down to the experience of a ‘transition’ in these countries, or possibly due to differential mechanistic influences of the alternative SES indicators. The inconsistent direction of findings for men apparent in McLaren’s review was also in keeping with the findings of Sobal et al (1989).

The review of El-Sayed et al (2012) identified just two UK-based studies to include individual-level income as a measure of SES in relation to obesity (6% of all included studies). The review, however, did not discuss income independently of overall SES. Referring to the original studies themselves, the findings indicated: higher odds (OR 1.36, 95% 1.21, 1.52) of being overweight with each higher income category (Lawlor et al., 2005); and a lower reported income in women who were obese compared to those who were not, but not when adjusted for confounding variables, nor in men (Viner, 2005).

Ball and Crawford (2005) identified nine studies that reported income in their review investigating SES and weight change. Studies investigating men separately tended to find no association between income and weight change: the one exception (Kahn and Williamson, 1991) was the finding that low-income men had higher odds of experiencing major weight gain. The study which did not stratify by sex also found no association between income and weight change. Amongst women, three of the studies found no significant associations. However, three studies found a significant association, with higher weight gain with lower incomes, two of which associations were negative (the direction of the association of the third was unidentified). Low-income women also experienced higher odds of weight loss in two studies. Income represented the most inconsistent of the SES indicators examined in the review, and the authors stress the differential associations with weight change of the different SES measures, as highlighted by previous review authors (see above).

Quality of reviews

None of these four reviews were without limitations. The oldest study in particular (Sobal and Stunkard, 1989) was lacking in a clear statement of the review methodology used, making it difficult to ascertain the possibility that bias may have been introduced in the process.

Review searches were all limited to English language articles (although Sobal and Stunkard, 1989 did not state search limits) and, where stated, to published non-grey literature only (explicitly stated in Ball and Crawford, 2005, and El-Sayed et al., 2012). None of the reviews reported taking steps such as double-screening or multiple researcher data extraction to minimise bias. Nor were quality appraisals explicitly performed for the included studies in any of the reviews, although El-Sayed et al. (2012) included a discussion of methodological limitations.

However, all of the studies provided clear statements of the inclusion criteria in their searches, including stated populations, exposure, comparisons and outcomes, and, with the exception of El-Sayed et al. (2012) in which only one database was searched and Sobal et al. (1989) in which the methods were not stated, the search strategies of the other two reviews were broad and comprehensive. Adequate details of the individual included studies were presented in the reviews of Ball and Crawford (2005) and El-Sayed et al. (2012), but Sobal and Stunkard (1989) and McLaren et al. (2007) reported only sample size, country and the presence and directions of significant associations. Although this may be considered a drawback in the quality of the reviews, it could be argued that, due to the extensive number of included studies in these two reviews, a summative approach such as taken was appropriate.

2.5 Discussion

This chapter sought to review the literature to date which has investigated the role of income, food prices and cost of diet in encouraging excess energy intake. The literature was searched in a semi-systematic manner, and data were extracted and organised into six sections to reflect the six key relationships under investigation (of two outcomes and three 'exposures').

Six reviews have recently been published with a focus on investigations into economic factors of food purchasing, dietary intake or body weight (see Section 2.4.2), which demonstrates an increase in interest surrounding these issues. However, each of the published reviews had important limits in its criteria – for example, restricting the search to US-based studies (Powell and Chaloupka, 2009, Powell et al., 2013), developed countries (Lee et al., 2011) or to subsidy effects (Black et al., 2012). The review of Lee et al (2011) was the closest in aim and criteria to the current chapter, but, despite being systematic, there are important shortcomings in the synthesis and conclusions of this review: firstly, there was a lack of differentiation between 'food price' and 'diet cost' data and methods, and secondly, the search strategy appears to have missed several important studies that were included in other (US-based) reviews. Additionally, none of these reviews incorporated income as an economic factor. Therefore, this chapter was necessary to draw together the existing evidence around income, food prices, diet cost, dietary energy density and body weight.

2.5.1 Economic factors and dietary energy density

The findings relating to dietary energy density are largely in keeping with the prevailing hypothesis that economic factors influence the selection of energy-dense foods.

Overwhelmingly, the studies linking diet costs and dietary energy density reported a negative association. However, as has been widely remarked in commentaries (Lipsky, 2009), in reviews (Lee et al., 2011), and in the commentary on study quality in Section 2.4.3.2, this observation may be the result of mathematical coupling, in which energy is included in calculating both the exposure and outcome. This is perhaps supported by the null result reported by Drewnowski et al. (2007) for female participants when daily diet costs, as opposed to energy costs, were analysed.

Conversely, this study was found to be of lower quality in comparison with the other studies, due to the statistical approach (unadjusted ANOVA) taken. A minority of the studies attempted to control for mathematical coupling by using residual values in their analyses (Aggarwal et al,. 2011, Maillot et al., 2007, who reported that they conducted these analyses, but the findings were the same and therefore were not reported separately, and Alexy et al., 2012), and these found a similar association. This shows that a negative association was reported regardless of study quality.

The evidence linking income and dietary energy density was less strong, but on the whole in agreement with the premise that economics influence diet selection. Two of the three studies amongst adult samples found evidence of lower energy densities amongst those with higher incomes. The one study which reported no significant findings, Waterlander et al (2010), was of poorer quality than the other two: the authors were not able to adjust for household size in their analyses comparing dietary energy density by income category, and they suggest themselves that the samples were underpowered to detect a significant difference. The poor quality of the study reporting no association perhaps indicates a false negative result.

Amongst children, a link between income and dietary energy density was not suggested by the evidence. Given the sound quality of the studies included, it may be concluded that no such association exists in children (albeit a conclusion from a small number of studies). Intuitively, a less strong link between household or family income and diet amongst children may be expected, given the varying degree of autonomy in food selection that children may have.

Interestingly, there were no studies identified in which food prices and dietary energy density was investigated. This is perhaps surprising, given the growing popularity of the food price-obesity argument (see below), and is a gap that needs to be filled if the purported causal pathway is to be substantiated.

Taking all this into consideration, the overall conclusion of this review is that the evidence – amongst adults, but not children – linking income or diet costs with dietary energy density is supportive of the theory that affordability is a determinant of dietary energy density. However, the review has identified that certain methodological issues need addressing and that the number of studies published to date is modest. Therefore, more research, particularly surrounding food prices, is needed to confer confidence to these conclusions.

The review of Lee et al (2011) was the only review identified to examine dietary energy density as an outcome. Although they did not consider studies which examine diet in relation to income, the authors included studies on diet costs, and similarly observed a consistently negative relationship with dietary energy density as was observed in this chapter. However, the authors concluded that the validity of the studies was questionable, given the mathematic relationship. As a result, they did not concede a meaningful relationship was apparent. In contrast, the review conducted here has identified that a handful of the studies avoided mathematical coupling by using residual values, and also found the negative relationship. Nevertheless, appropriately analysed studies constitute a minority and additional evidence would be useful.

2.5.2 Economic factors and body weight

The findings relating to weight are tantalising. Evidence predominantly came from studies investigating food prices, which varied widely in their focus and quality. This heterogeneity makes it difficult to draw conclusions. In particular, studies which looked at the prices or price trends of individual food items are problematic to synthesise, due to differing selections of foods and varying contexts used. However, all the studies taking such an approach found significant associations for at least some foods. This lends traction to the food price-obesity hypothesis.

Studies which included price indices for combinations of foods or food groups are arguably easier to compare. However, there is still much variation in the techniques used to derive and analyse such indices, and the results reported make it apparent that the choice of analytical model can be greatly influential on the outcome. This was particularly true of studies investigating fast food indices, in both adults and children. In children, the studies were found to be of mixed quality: the quality appraisal in Section 2.4.4.1 shows that many of the studies used self-reported measures of weight and height, and there were concerns identified regarding the reporting of statistical analyses. The best-quality study amongst children looking at fast food price indices, Powell and Bao (2009), found no association between fast food prices and BMI amongst young children. Self-reported data was again a feature of the studies investigating fast food prices with adults; however, studies in adults were judged to be of similar quality, but still presented mixed findings. The mixed findings and quality issues of studies around fast food prices and body weight make a confident conclusion unlikely.

In adults, findings for food-at-home indices were equally dependent upon analytical approach, though in children no associations were apparent (in the only study in children). The studies amongst adults were judged to be of similar quality, despite taking different analytical approaches, and all of these studies used self-reported anthropometric data. Food group price indices (not including fruit and vegetables – see below) did not reveal significant associations with children’s body weight; the two studies investigating this (Sturm and Datar, 2005 and Sturm and Datar, 2008) were of good quality, suggesting a true lack of association.

Evidence of a link between fruit and vegetable prices and weight was a little stronger: amongst children, three of the five studies found a significant positive association; and in adults, this significant positive relationship was echoed, although in both studies this was true only for certain subgroups (the ‘near poor’ (Beydoun et al., 2008) or poor women and women with children (Powell and Han, 2011)). The studies in adults were of a similar quality, both sharing the drawback of using self-reported height and weight, which risks introducing bias as a result of measurement error. In children, interestingly the three studies which reported a significant finding were all of better quality than the two which did not report the association: the latter two (Powell et al., 2007, and Black et al., 2013) can be criticised in terms of their quality (see Section 2.4.4.1). These findings suggest that, as the price of fruit and vegetable increases, these populations are less likely to purchase and consume fruit and vegetables. Given this observation, it could be conjectured that these people are instead purchasing more energy-dense foods and are more at risk of weight gain.

All of the studies investigating the effects of taxation focussed on soda or soft drinks. Only one of the three studies in children found a significant (negative) effect of soda taxes on body weight (Sturm et al., 2010), and one of two studies in adults (Fletcher et al., 2010a). Of the three studies in children, the two which failed to find an effect were judged to have quality issues particularly in terms of statistical analyses and the reporting of the analysis. In adults, the study of Fletcher et al differed in its approach in that the exposure variable was framed as a *change* in the tax rate, rather than just the presence, absence or degree of tax, as was used in the other studies. This, it could be argued, is an important differentiation to make, and could explain why other studies failed to find significant results – it is possible that the change in price as the result of a change in tax is more noticeable to consumers and therefore more influential on consumer behaviour and dietary consumption. The overall interpretation of these findings is that the evidence points towards an association or effect, but the evidence is limited to just a few studies.

Overall, studies investigating food prices and obesity imply that food prices have a role to play in obesity rates, yet the results highlight the difficulties in quantifying this relationship, making a consensus statement about the nature of this role unachievable – at least from the evidence to date. In children, there was just one study which identified prices of specific foods (Thomas et al., 1996). Although there were several potential shortcomings identified in the quality appraisal of this study, it nevertheless revealed many significant associations in a relatively small sample, using objective measures of anthropometry. Unfortunately, however, as this is the only study in children and the setting of this study was specific to an area of Cote d'Ivoire, it is unlikely to be generalisable to other settings or the wider population of children in general. Of the studies linking food item prices to bodyweight in adults, that of Duffey et al. (2010) was judged most favourably in the quality appraisal (Section 4.2.2.1), with a large, longitudinal, nationally representative sample, objective measurements of weight, and appropriate and clearly presented statistical analyses. Duffey et al. (2010) found a significant negative association between the prices of soda and pizza with bodyweight, but no association between prices of whole milk and burgers with bodyweight. Despite weaknesses in many of the study designs, and the heterogeneity of indicative foods selected, this avenue of investigation consistently has revealed significant relationships in the literature.

Studies of diet costs and body weight are mixed in their findings – in adults at least. Amongst children, associations between diet costs and BMI were not apparent. As stated above, children may be suspected to show a different relationship between economic factors and food choices (and therefore body weight), given the influence of parental mediation and varying degrees of autonomy in diet selection. Alternatively, both of these studies in children were identified as having flaws in the quality appraisal: one used a non-probability sample, the other relied upon a small sample size, and both studies used unadjusted analyses to test the relationship. Therefore, it cannot be determined whether a true lack of association exists in children, or whether the methodological limitations were responsible for the studies' findings. The studies using samples of older adults (Lo et al., 2012, Michaud et al., 2007) are difficult to interpret, in part due to the way in which results were reported (for example, Michaud et al did not report p values nor confidence intervals, and in the study of Lo et al, the relationship was not the primary outcome under investigation). Nevertheless, the findings suggest that lower diet costs attributable to vegetables are associated with more underweight participants, whilst estimated costs of food away from home could have a negative relationship with BMI. It is perhaps reasonable to suppose that economic factors may have a different role in diet selection amongst the elderly – a population which often has an over-representation of those on low incomes (potentially unable to afford adequate nutrition), as well as mobility issues which may have a significant bearing on body weight (either via reducing energy expenditure or impeding access to food). The results of these studies also could reflect a non-linear relationship between food prices and BMI – for example, vegetable costs being linked to underweight could be explained by these individuals being food insecure with hunger, but the relationship between costs and weight could be different as costs increase. Although it is not possible to ascertain this from the studies described, a non-linear relationship is considered in Chapter 7. In the end, given the differences in approach, the small number of studies and the quality issues of those studies, it is not possible to draw conclusions about the role of diet costs in determining BMI amongst older adults.

The potential age differences described above – in the proposed food price-body weight pathway – might be used to explain the mixed findings amongst studies of diet cost and body weight. Yet, the three studies amongst working-age adults also reported contradictory findings: in their studies of female Japanese students, Murakami and colleagues (2007, 2009) found significant negative (though small) associations between diet cost and anthropometric measurements; however, in a larger sample of Spanish graduates, Lopez et al (2009) found just the opposite at baseline, with higher BMI associated with higher diet costs. The reasons for this contrast are unclear. It may be due to the contrasting sample characteristics, or to the differences in assessment methods used, or the appropriateness of adjustments made in analyses. All three studies used non-probability sampling, so there is a possibility that the findings reflect different selection biases. Lopez et al. (2009) analysed a longitudinal sample, which may be considered of superior quality, but relied on self-reported anthropometry, unlike Murakami et al. (2009). With such a small pool of published studies available, and contrasting quality strengths, it may be too early to judge which findings are more convincing.

Finally, reviews suggest that, in developed countries, income is related to body weight amongst women but not men. This finding was repeated in most, but not all, of the studies included: as women reported higher incomes, they were more likely to report a lower BMI (or other anthropometric indicator). Of all the common indicators of socio-economic position, the reviews showed income as the least consistent predictor of BMI. This implies that affordability of a healthy diet is not as important as education or occupation in diet selection, although still clearly implicated. The quality appraisal of these reviews revealed shortcomings in terms of minimising bias in the screening and data extraction processes. However, in the majority of studies, the search strategies were transparent and comprehensive, which should increase confidence in the breadth of evidence represented.

In contrast to the conclusions of this review, Powell & Chaloupka (2009), in their review of US food price studies, concluded that the majority of studies indicated negative relationships between the prices of unhealthy foods and BMI and positive relationships for prices of fruit and vegetables. However, in their more recent update (Powell et al., 2013), conclusions were more mixed, and were similar to those described above. The review in this chapter adds to those of Powell and colleagues, in that the term ‘consumption’ is clearly differentiated, to identify only those studies that measured dietary intake, as opposed to purchases.

2.5.3 Methodologies

The studies described above demonstrate a variety of approaches, and this heterogeneity itself highlights how difficult it is to investigate these micro-economic factors. Given that almost every study was different in design, context, definition of diet cost and analytical approach, it is perhaps unsurprising to find inconsistent results.

Predictably, none of the relevant studies were randomised control trials (RCTs) – manipulations in income or food prices would be both practically and ethically problematic to implement. In the hierarchy of evidence (CRD, 2008), RCTs are often promoted as the ideal design. However, in the absence of such evidence, where it is inappropriate to conduct such studies, it is necessary, and indeed valuable, to consider the different types of evidence (Gortmaker et al., 2011). Aside from RCTs, the prospective cohort study is often regarded more favourably than a cross-sectional observational study. In this research field, cross-sectional data was dominant. Even some of the studies which drew data from longitudinal cohorts analysed them cross-sectionally. The exceptions to this were found amongst the food price-obesity literature, several of which employed time series data, and one study amongst the diet cost-obesity literature. Studies investigating dietary energy density were all cross-sectional in design. It is important to note that conclusions from the literature will be limited, as a result.

The majority of relevant studies were those which investigated effects of food prices. Even within this one approach to the research question, however, there was substantial variation: firstly, in how to incorporate food prices into a variable or variables (for example, using fast food indices); and secondly, in how to build analytical models. This makes it extremely challenging to compare or synthesise the results, with only a few studies sharing the same approach to measuring prices. Importantly, results depended upon the analytical method chosen.

As expected, no studies were identified which measured both expenditure and dietary intakes; the literature investigating diet costs therefore, without exception, used the same method, applying national (or local supermarket) prices to foods that were reported to be consumed. The comparability studies described in Chapter 5 indicate that diet costs estimated in this manner compare well to estimates from purchasing receipts (this work was also published in Timmins et al., 2013b). The studies did not use the same dietary assessment tool, however, and this may have had a bearing on the resulting diet cost estimates (see Murakami et al., 2008a, Monsivais et al., 2013 for comparisons). Estimating costs from reported dietary intake has limitations, one of the most important of which is that estimates will reflect the measurement error associated with the dietary assessment. The studies also varied in whether they chose mean prices, or lowest prices, to cost diets. This constitutes a major assumption regarding the prices actually encountered by participants. Chapters 5 and 6 discuss the strengths and weaknesses of this methodology in more detail.

There was also a lack of agreement amongst the studies reviewed which investigated income on how to realise the income variable, with some studies expressing it as a proportion of the government poverty line (which takes into account household size and composition), and others relying upon household measures.

Failing to take into account the composition or size of the household when using a household-level income variable could be misleading (Benzeval et al., 2001); therefore this area of research requires more attention.

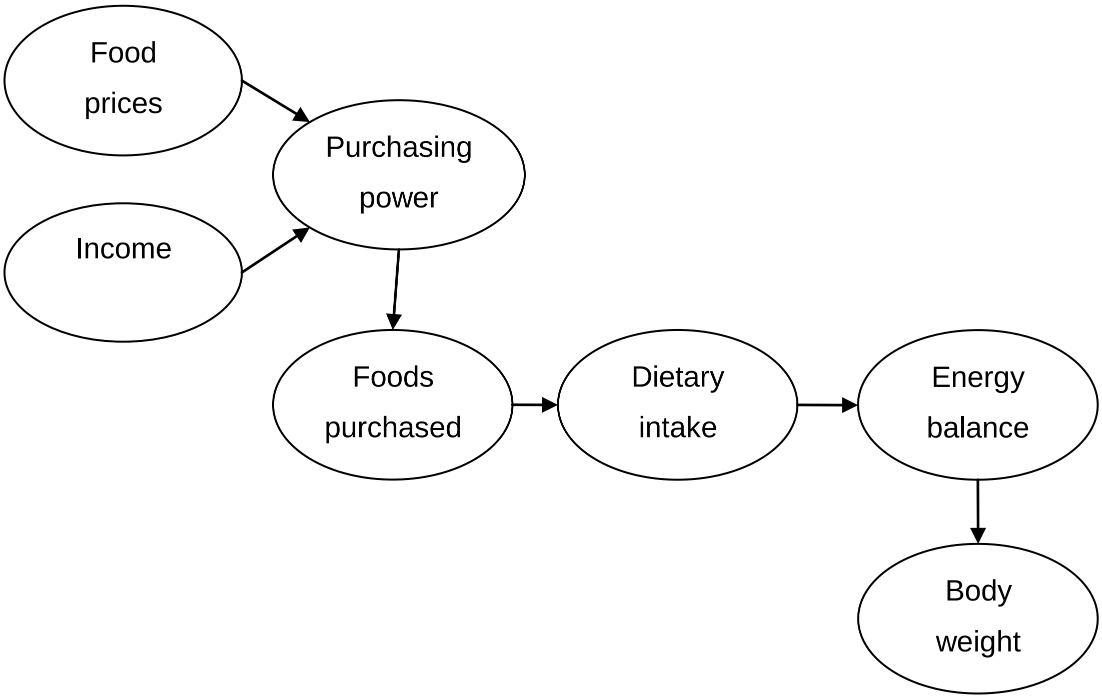
Despite a heterogeneous body of literature, this review found that several of the included studies shared sources of data – a few, for example, used data from the same sample (such as NHANES) and many of the US studies incorporated price data from the American Chamber of Commerce Research Association (ACCRA), acknowledged to be limited in its breadth and specificity of food items. This is an important limitation to the interpretation of these studies' findings, as well as to their generalisability.

Not all of the studies chose robust statistical techniques. Several employed unadjusted statistical comparisons (see, for example, Waterlander et al., 2010), or adjusted inappropriately or for too many covariates (for example, Murakami et al., 2007, Murakami et al., 2008b). Some studies were also weak in their reporting of results, for example not reporting actual statistical values (Mushi-Brunt et al., 2007) or p values (Michaud et al., 2007). The majority of the literature, however, employed well-considered analytical methods.

2.5.4 Implications for the "food price-obesity" hypothesis

Chapter 1 set out the conceptual framework which motivated this literature review – namely, the "food price-obesity hypothesis" (Section 1.5). The hypothesis, proposed and supported by several researchers and policy makers, suggests that food prices – via their impact on purchasing power – are responsible at least in part for recent obesity trends (see Figure 1.1 repeated below).

Figure 1.1 The proposed causal pathway between food prices and obesity



This literature review set out to establish evidence in support of, or in opposition to, the food price-obesity hypothesis. In the absence of studies which are able to measure all of the aspects of the proposed causal pathway, the focus of the review was separated into three main exposures – food prices, income and diet costs – as proxy measures of purchasing power.

The conclusions presented above support a link between income and dietary energy density, and between diet costs and dietary energy density. This can be represented conceptually by adapting the figure above, to show how these associations fit in the causal pathway (Figure 2.2 and Figure 2.3). It can be seen that, whilst income is directly in the causal pathway, diet costs offer a representation of foods purchased, and neither estimated diet costs nor purchasing data are direct measures of purchasing power. As mentioned above, there were no studies which assessed food prices and dietary energy density. Taking the evidence from both available angles (income and diet costs), it would seem the evidence supports a link between purchasing power and dietary energy density.

Figure 2.2 Income and dietary energy density within the food price-obesity framework

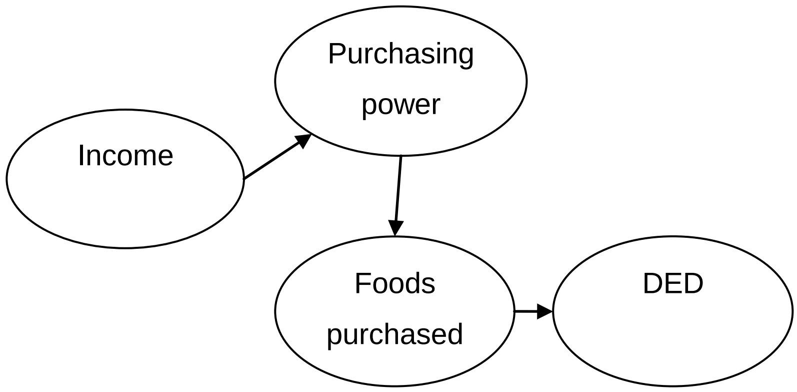
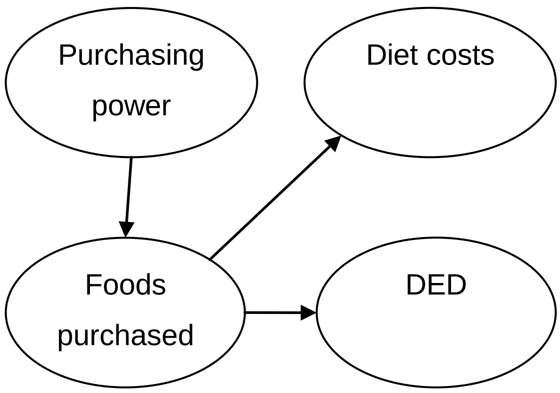


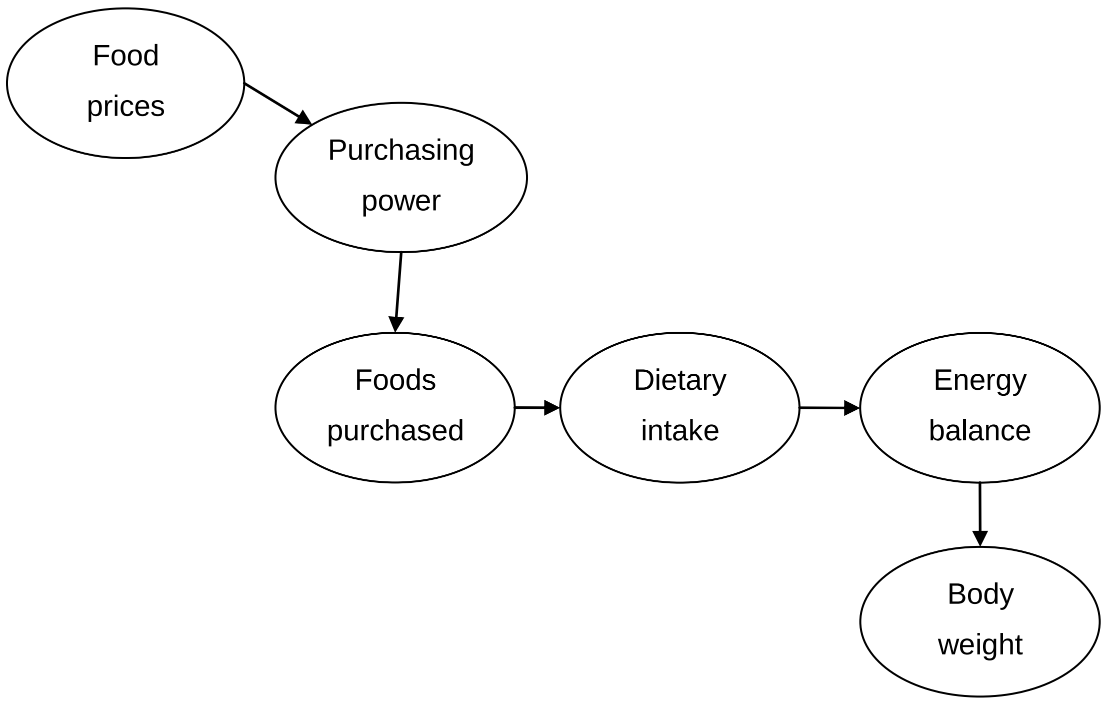
Figure 2.3 Diet costs and dietary energy density within the food price-obesity framework



The literature on obesity (or measures of bodyweight or composition) takes the hypothesized pathway to its endpoint; therefore, by necessity, these investigations are attempting to measure data further removed along the pathway. This point is worth stating, given the inconclusive findings of the literature review. It stands to reason that outcomes further along a causal pathway will be more difficult to ascertain, with more potential for confounding along the pathway. This is especially true of studies which are unable, by design (for example, if they are cross-sectional studies), to take into account the protracted duration of the proposed aetiology.

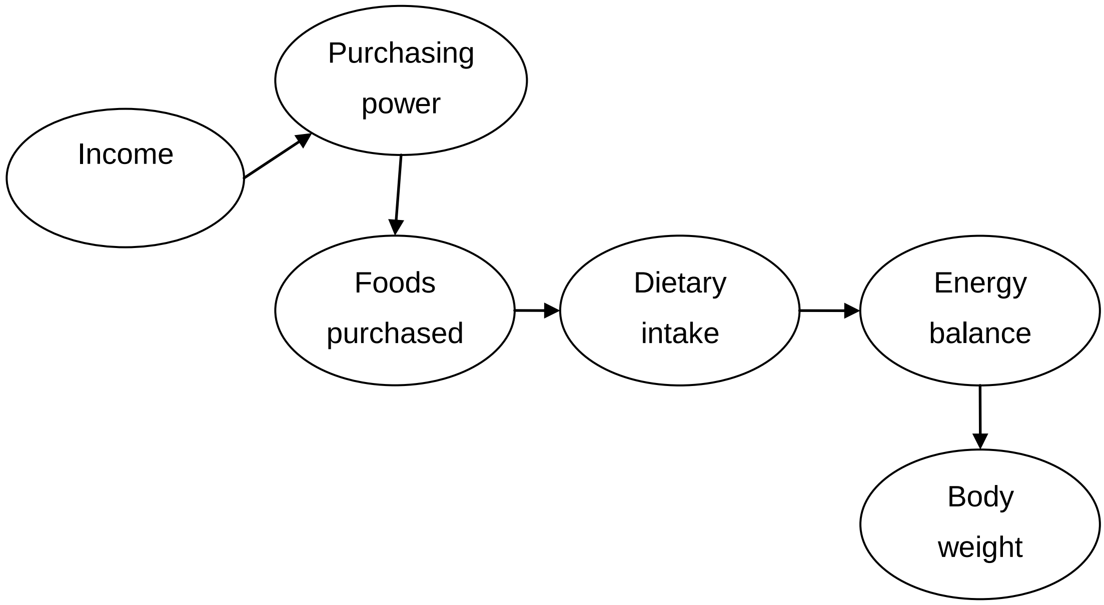
The literature review revealed a mixed and conflicting presentation of evidence. In contrast to the studies investigating dietary energy density as an outcome, the literature on bodyweight outcomes predominantly focussed on food prices as an exposure. This emphasis is perhaps unsurprising, given the widespread discussion of the food price-obesity hypothesis (see Figure 2.4). With such distal exposure and outcome variables, it would be expected that the chance of false negative results would be more likely in investigating the link between food prices and bodyweight. Nevertheless, a heterogeneous body of literature reported significant associations, implicating food prices – albeit defined in several different ways – in weight status.

Figure 2.4 Food prices and bodyweight within the food price-obesity framework



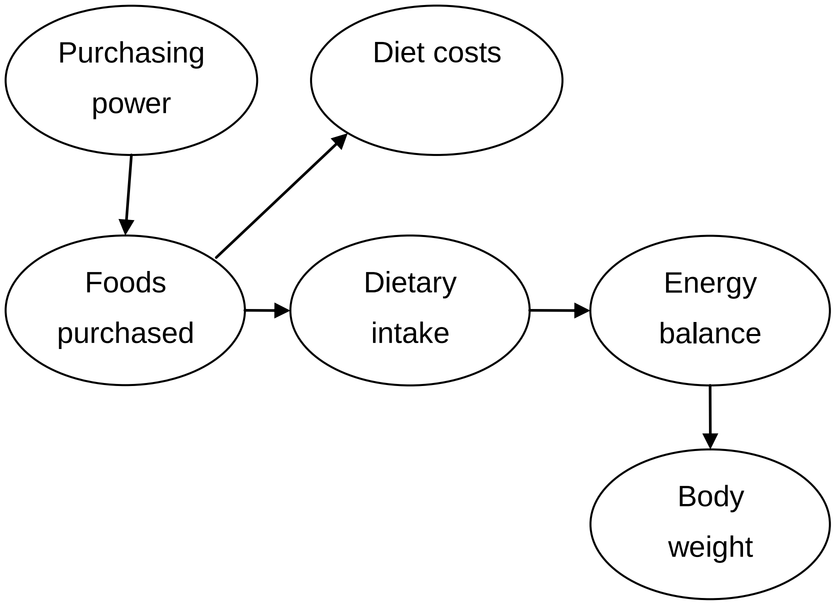
Studies focussing on the other chief determinant of purchasing power, income (see Figure 2.5), were similarly prevalent. The synthesis of evidence provided in the reviews suggests an association between income and bodyweight amongst women, though not men. Unfortunately, there were no studies apparent which assessed both aspects of purchasing power – income and food prices – together. This would be a valuable avenue of research to help ascertain whether the two variables exert their influence on bodyweight via their role in establishing purchasing power.

Figure 2.5 Income and bodyweight within the food price-obesity framework



The results of the review of literature examining diet costs and bodyweight were inconclusive. Suggested reasons for this were predominantly related to the small number of studies, the heterogeneity of study design and settings, and identified issues with study quality. There is also the possibility that estimates of diet costs are too far removed in the hypothesised food price-obesity pathway (as illustrated in Figure 2.6), therefore making it more difficult to identify associations. The analyses of later chapters in this thesis will add to the limited evidence base.

Figure 2.6 Diet costs and bodyweight within the food price-obesity framework



2.5.5 Wider implications

Due to the mixed findings uncovered, the key implications from this review relate to future directions for research, rather than implications for policy or interventions. Firstly, the results indicate some significant gaps in the literature: it is interesting to see that no studies have been published which investigate food prices and dietary energy density, despite the implication of energy density in the proposed mechanistic pathway between food prices and obesity. There was also a paucity of literature based upon UK data, and there appeared to be an over-reliance on cross-sectional designs, with a minority of longitudinal or time series analyses. There were also several other issues found with the quality of studies published in this field, particularly in an over-reliance on self-reported anthropometric measurements and poorly considered statistical analyses. There is much scope for improvement in future studies of this kind.

A consideration of the literature has also brought to light that there is some confusion with regards to definitions of terms – for example, some studies and existing reviews do not differentiate between consumption in the economics sense (in other words, purchasing) and dietary consumption or reported dietary intake. One review also was not clear on the difference between studies employing food prices as opposed to those estimating costs of diets. Clarification of the different terms and approaches is vital if a consistent message is to be communicated.

Finally, the synthesis above has indicated that the underlying relationships in this area are complicated. In particular, various subgroups – males or females, the near poor, those with children – may elicit differing findings, which suggests that the food price-obesity hypothesis may not be as straightforward as it is often portrayed. Future research must bear in mind such differential effects.

2.5.6 Limitations

Whilst this review has attempted to illustrate as comprehensively as possible what is known about food prices, diet costs, energy density and body weight, practical considerations have imposed certain important limitations. Firstly, the search was confined to literature published in English, peer-reviewed publications, and did not include studies which reported purchasing data or the modelling of hypothetical scenarios. The rationale for this latter exclusion was partly because of the drawbacks of assuming purchasing and consumption are equivalent; however, it is acknowledged that these studies could supplement the literature base and contribute to our understanding of the causal mechanisms.

Efforts were made to make the review process as systematic as possible. However, this review cannot be considered fully systematic, mainly because the majority of the review was conducted by a single reviewer. Furthermore, the pre-established protocol written in 2011 was adapted and refined in the second stage of the review. Despite the reasons for this being justified (due to the subsequent publication of new reviews), this could be considered a weakness.

Finally, the review was limited to a narrative synthesis of the findings, because the literature was too heterogeneous to conduct a satisfactory meta-analysis.

2.6 Conclusion

This literature review was necessary to draw together the different aspects of evidence relating to micro-economic factors of overweight and obesity. Previous reviews did not address these multiple aspects. The recentness of most publications indicates this is an area of increasing attention in the research community, and the disparate approaches and mixed understanding of terms mean that a synthesis of studies was timely.

Conclusions from the existing literature remain elusive. There are significant gaps in research, and existing studies are heterogeneous in design and setting and variable in quality. However, there are interesting results reported for many of the studies, reinforcing that this topic is a worthwhile area of investigation. The following chapters attempt to address some of the gaps identified – geographically (using UK data), and methodologically. However, the data available are not appropriate to address all of the gaps highlighted here – for example, longitudinal data are not available in the data set to be explored. An association between particular food prices and diet would have far-reaching consequences for public health initiatives, implying as it does that there may be fiscal means of counteracting the obesity ‘epidemic’.

Chapter 3 Sample description

3.1 Summary

The main analyses of this thesis will be conducted using data from the National Diet and Nutrition Survey (NDNS). The NDNS is a national dietary assessment survey, designed to represent the general UK population. This chapter will introduce the NDNS: its purpose and design, sampling techniques and data collection protocol. In addition, the chapter presents a description of the analytical sample, outlining some of the chief characteristics.

A brief discussion of the survey limitations is included. In particular, a description of energy intakes in the sample is presented, and the possible presence and potential influence of under-reporting considered.

This chapter will not cover the methods used in the derivation of new outcome variables from the sample data – for example, equivalized income, or diet costs. These will be explained in the chapters in which they feature (Chapters 4, 7 and 9).

Further details about the NDNS are available from the survey reports – for example, Bates et al. (2011).

3.2 Introduction

The NDNS is a national dietary monitoring programme, funded by the Food Standards Agency (FSA) and the Department of Health (DH). Previously, the survey comprised a series of one-off cross-sectional studies. Since 2008, however, a rolling programme was introduced, sampling around 1000 new participants each year. The purpose of the survey is to track national trends in dietary intake in relation to targets and recommendations, and to assess the nutritional status of different population groups. Therefore, every effort has been made to capture a nationally representative UK sample of individuals, aged 18 months and over. Sample recruitment methods are described in Section 3.3.

The survey is carried out by NatCen (the National Centre for Social Research), MRC HNR (Medical Research Council, Human Nutrition Research), the joint surveys team at the Department of Epidemiology and Public Health UCL (University College London), and NISRA (The Northern Ireland Statistics and Research Agency). Data sets are deposited with the UK Data Archive (NatCen et al., 2012). The original data creators, depositors and copyright holders of the NDNS and the UK Data Archive bear no responsibility for their further analysis or interpretation.

The analyses in Chapters 4, 6, 7 and 8 use data from the first two waves of the programme, 2008-2009 and 2009-2010. The original sample was comprised of both children and adults; however only adult data (≥19 years; n=1031) were included in the analyses of this thesis.

This chapter sets outs details about the NDNS recruitment and characteristics that are relevant for the interpretation of later results. The objectives are to:

1. Outline the survey design and sample recruitment;
2. Describe the sample characteristics;
3. Present descriptive results of pertinent dietary and anthropometric measurements;
4. Explain the derivation of new variables for this thesis; and
5. Discuss how the methodology and characteristics of the sample may be relevant to the interpretation of the analyses in subsequent chapters.

The descriptive results presented below relate to the analytical sample used in this thesis, and therefore may differ, albeit slightly, to the survey report. A discussion of sample weighting is also included.

3.3 Sample recruitment

In each year of data collection, a nationally representative sample of individuals is selected from private residences drawn from the Postcode Address File (PAF). Participants from private residences only are included.

A clustered sampling design was adopted to facilitate data collection: 27 addresses were randomly selected from each of 120 Primary Sampling Units (PSUs), themselves randomly selected from across the UK. Where there was more than one household at an address, one household was randomly selected. The interviewer then randomly selected up to one adult and one child from each household. Participants who were currently pregnant or breastfeeding were excluded from the survey. Eighteen of the 27 addresses were selected as ‘child booster’ addresses, so that at least two thirds of each PSU contained individuals aged 18 years and under. Booster samples were recruited from Scotland, Wales and Northern Ireland to enable cross-country comparisons.

In Years 1 and 2, 10% of the eligible addresses declined to take part before household selection. After selection, there was an overall response rate of 64% of households, presenting data from 2126 ‘fully productive’ individuals. This thesis is concerned with adult data only, of which there were 1031 ‘fully productive’ participants. Follow-up data (see Section 3.5) was missing for 24% of these individuals.

Seventeen participants were discovered to have incomplete dietary data (completing only three days of the four-day diet diary – see Section 3.5.2). This not only affects the daily diet cost estimates (which were calculated by assuming a total of four days’ dietary information), but also diminishes the level of confidence that can be attributed to the assumption that the dietary data reflect habitual intake. These participants were therefore excluded, leaving an analytical sample of 1014.

3.3.1 Sample weighting

Although designed to be a representative sample of the UK population, it is inevitable in survey sampling that non-response, clustering and other methodological factors result in a sample that deviates from the national demographic profile. For this reason, a weighting scheme for the NDNS has been calculated in an attempt to counteract any bias in selection probability or non-response. Weighting the sample in this manner allows the survey team to publish results that can be said to represent the dietary intakes of the UK's population.

There are two types of survey weights employed in the NDNS: selection weights and non-response weights. Selection weights are employed due to the sampling procedure employed by the survey: because the selection is made at an address level, it is possible that multiple dwelling units within a single address, or multiple catering units within a single dwelling unit, will be under-sampled. Selection weights are added to these units so that these dwelling units and catering units are not under-represented.

The method of applying non-response weights essentially involves replicating the responses of participants from a subgroup which experienced a higher rate of non-response. There is an underlying assumption that non-responding members of the subgroup would have responded similarly to responding members of the same subgroup in all aspects of the survey. Non-response weights in the NDNS are calculated with calibration methods using age, sex and Government Office Region.

The extent to which the NDNS sample differs to national estimates is apparent in the weighting procedure adopted in analysis of the survey: the proportions of the sample falling under different demographic categories before and after the application of sample weights is specified in Appendix B of the NDNS report (Bates et al., 2011).

The prime advantage of using an unweighted sample is the avoidance of relying on the assumption of within-group similarity in response. In addition, if the information used to create the sample weights is also included in a regression model, using sample weights will result in an inefficient model (Bloom and Idson, 1991). As described in Appendix B of the survey report, the demographic differences between the weighted and unweighted sample are minor. For these reasons, the analyses presented in this thesis use unweighted data, as has been the approach of other authors in this area (Chou et al., 2004). As a result, the investigations of subsequent chapters will be concerned with associations in the survey sample and cannot be considered representative of the population[1]. It should also be borne in mind that dwelling units in multiple-unit addresses and catering units in multi-occupied dwelling units are likely to be under-represented in the analytical sample.

---

[1] Readers may be interested to compare the survey results presented in Chapter 6 of this thesis to the population estimates published in TIMMINS, K., HULME, C. & CADE, J. 2013a. The monetary value of diets consumed by British adults: an exploration into sociodemographic differences in individual-level diet costs. *Public Health Nutrition* [Online]. Available: dx.doi.org/10.1017/S1368980013002905 [Accessed 28 Nov 2013].

3.4 Ethical considerations

Ethical approval for the NDNS was sought and obtained at the outset from the Oxfordshire A Research Ethics Committee, as well as Local Research Ethics Committees in the areas in which data were collected. Details of ethical approval can be found in the survey report (Bates et al., 2011). This ethical approval applies to secondary analyses of the available anonymised data, such as those conducted in this thesis.

3.5 Methods

Data were collected in two phases: firstly, a face-to-face interview ascertained participant characteristics, measurements of weight and height and administration of the four-day diary; 'fully productive' participants were then visited by a nurse for physical measurements (demi-span, waist and hip circumference, infant length, blood pressure), and blood and urine samples.

3.5.1 Participant characteristics

Characteristics relating to each participant were gathered using interviewer-administered CAPI (Computer Assisted Personal Interview) or, in the case of smoking and drinking behaviours, self-completion questionnaires.

A summary of the NDNS weighted sample characteristics is included in the survey report (Bates et al., 2011). However, a further description of the specific variables involved in these analyses, including that of newly derived categories, was deemed necessary for the analytical sample to be used in this thesis. Summary statistics for key sociodemographic variables are therefore presented below.

Categories for household income, employment (NS-SEC 8) and qualifications have been collapsed to facilitate analysis. Household income in the NDNS was assessed using 13 categories: <£5,000, £5,000 to £9,999, £10,000 to £14,999, £15,000 to £19,999, £20,000 to £24,999, £25,000 to £29,999, £30,000 to £34,999, £25,000 to £39,999, £40,000 to £44,999, £45,000 to £49,999, £50,000 to £74,999, £75,000 to £99,999, and £100,000 or more. These were collapsed to five bands: <£15,000, £15,000 to £24,999, £25,000 to £34,999, £35,000 to £49,999 and £50,000 or more.

Qualifications were collapsed from eight categories (degree or equivalent, higher education below degree, GCE A-level or equivalent, GCSE grades A-C, GCSE grades D-G or commercial qualifications, foreign or other, none, and still in full-time education) to four (degree or equivalent and higher education, GCE A-level or equivalent and foreign or other, GCSEs or commercial qualifications or currently still in full-time education, and none).

The NDNS uses the NS-SEC 8 categories to describe occupational class: higher managerial and professional, lower managerial and professional, intermediate occupations, small employers and own account workers, lower technical and supervisory, semi-routine occupations, routine occupations, never worked, and other. These were collapsed to four categories: managerial and professional (higher and lower), intermediate occupations and small employers and own account workers and lower technical and supervisory, routine and semi-routine occupations, and never worked with ‘other’.

3.5.2 Dietary data

Dietary consumption is measured in the NDNS by consecutive four-day un-weighed food diaries. Estimated (or un-weighed) food diaries are commonly used in dietary studies, due to their relatively low participant burden (compared to weighed intake), ease of administration and flexibility. In addition, they compare favourably to other assessment methods – for example, one comparison study found that the nutrient and food intakes calculated from un-weighed diaries did not significantly differ from those obtained by weighed diaries collected over 16 days (Bingham et al., 1994).

The diaries were provided to the participants on the first interview visit. The interviewer also contacted participants on the second or third day of the recording period, both to check recording and encourage completion. The selection of the start day for the diary recording period differed between Years 1 and 2, the main result of which was an over-sampling of weekend days in Year 1, and under-sampling of weekend days in Year 2. Details about day selection are described in Appendix A of Bates et al. (2011).

Portion size photographs were included for 15 commonly consumed foods, but all other portions had to be estimated using household measures or package weights. Diary data were coded and recorded using the DINO (Diet In Nutrients Out) software, which incorporates UK food composition data (FSA, 2002).

As well as recording foods and drink consumed, participants were asked to provide details for each eating occasion as to where it took place, with whom and whether it was at a table or whilst watching television. For each day they indicated whether the quantity they consumed was typical for them.

Section 3.6.2 below provides a summary of the dietary intake of the analytical sample used in this thesis. This includes:

- energy intake

- fruit and vegetable consumption (‘5 a day’)
- special diets
- unusual quantities of consumption, as reported
- food away from home (FAFH)
- alcohol consumption.

In addition, a summary is reported of responses to the interview question on the main place of household grocery purchasing. A full description of dietary intake is included in the survey report (Bates et al., 2011), and will not be repeated here.

Three of the variables listed above – alcohol consumption, unusual quantities of food consumed, and FAFH – were newly derived from existing NDNS data to aid analysis. These are described in turn below.

The calculation of ‘5 a day’ was performed by the NDNS team for the survey report and included as a binary variable (‘yes’ or ‘no’) in the data set. Achievement of the UK’s ‘5 a day’ recommendation was calculated from the dietary data, including composite dishes. The ‘5 a day’ criteria stipulate five portions, of 80g each, of fruit and vegetables, including dried fruit (30g for a portion) and up to one portion (150ml) of fruit juice, daily.

3.5.2.1 Alcohol consumption

Alcohol consumption was recorded in a number of ways in the NDNS, including by questionnaire (number of units in the previous week), as well as calculating average daily consumption from the diet diaries (by volume, in grams, and by per cent of total energy). Although both the self-report and diary methods are not without reporting bias, it was decided that the diary alcohol data were more appropriate given that it is the diary data which is included in the diet cost estimations.

Due to the highly skewed distribution of alcohol consumption (in grams), a categorical variable was derived. The cut-off points for four categories were specified, based upon Department of Health recommendations for drinking (NHS, 2012): higher risk (more than eight units per day for men, six for women); increasing risk (more than four units but less than or equal to eight units per day for men, and more than three but less than six for women); lower risk (four units or less per day for men, three units or less per day for women); and abstainers (participants who did not report drinking any alcohol during the four diary data collection days).

Grams of alcohol were converted to units using a conversion rate of one UK unit being equivalent to eight grams of alcohol (NHS, 2013).

3.5.2.2 Unusual daily intake reported

For each day of diet diary recording, participants were asked to indicate if this was typical of the amount they eat and drink in a day, more than usual or less than usual. Two variables were created in which the number of days were summed – firstly the number of days more than usual; and secondly the number of days less than usual.

3.5.2.3 Food away from home (FAFH)

The place where food was eaten was recorded in the diary alongside each food item – for example, “at home – kitchen” or “fast food outlet”. This information was used to identify when food was consumed away from home and to generate a variable summing the number of days, if any, on which participants consumed FAFH during the diet diary collection.

3.5.3 Anthropometric data

Height and weight measurements were taken by the trained interviewers on the first visit. Portable stadiometers and weighing scales measured to the nearest 0.1mm and 0.1kg respectively. Details of equipment and protocols followed are described in the survey report (Bates et al., 2011). BMI (body mass index) was calculated using the standard formula:

$$BMI = kg/m^2.$$

Equation 3.1

Categories of BMI follow the World Health Organization (WHO) definitions (WHO, 2006), shown in Table 3.1. Both BMI and BMI classifications exist within the NDNS data sets as ready-derived variables.

Table 3.1 WHO BMI classifications

BMI category	Category boundaries
Underweight	Less than 18.5 kg/m^2
Normal weight	18.5 to 24.99 kg/m^2
Overweight	25 to 29.99 kg/m^2
Obese	30 kg/m^2 or over
Morbidly obese	40 kg/m^2 or over

3.5.4 Statistical analyses

Summary statistics were generated for all variables. In addition, a number of univariate analyses were conducted to explore patterns in dietary intakes, including:

- Energy intakes between alcohol consumption groups, those who achieved or did not achieve ‘5 a day’, the frequency of FAFH reported, the frequency of unusual amounts of food consumed, and between BMI categories (bivariate linear regression, or p for trend).
- Sociodemographic characteristics according to alcohol consumption, achievement of ‘5 a day’ and BMI category (chi^2 comparisons).
- Alcohol consumption and achievement of ‘5 a day’ according to BMI category (chi^2 comparisons).

All tests between BMI categories excluded underweight participants (BMI <18.5kg/m^2; n=13) from the analysis. This was because of the small subgroup size, as well as the fact that the underweight were excluded from the main analyses of this thesis (see Section 4.3.4.3).

3.6 Descriptive results

3.6.1 Sociodemographic characteristics

Table 3.2 describes the sample in terms of sociodemographic categories. In total, 57% of respondents were female. Adults in the sample ranged from 19 to 94 years of age, with the mean age being 49.3 (SD 17.5). In terms of ethnicity, the sample was of a white majority (93%, n=940).

The mean number of people in the household was 2.5 (SD 1.3), with a range in size of one to nine persons. Almost a third of the sample (33%) lived in two-person households. In terms of main wage earner occupation, the majority of the sample (42%) fell under the “managerial and professional” description. A quarter of the sample reported having no qualifications (see Table 3.2). The distribution of reported household incomes is also shown in Table 3.2.

There were 226 participants who reported being a current regular cigarette smoker (22%), 247 reported ex-regular cigarette smokers (24%) and 541 who reported never having smoked cigarettes (53%).

Thirty eight per cent (n=389) reported a limiting long-standing illness, disability or infirmity. Of these, 209 (21%) stated their illness limited everyday activities. Twelve per cent (n=126) reported having an illness in the two weeks prior to interview which restricted their usual activity.

Table 3.2 Summary of sociodemographic characteristics of adults in years 1 and 2 of the NDNS rolling programme, combined (n=1014)

Variable	Proportion of sample (%)	Number of sample (n)
Sex		
Male	43%	434
Female	57%	580
Country		
England	81%	817
Wales	5%	53
Scotland	7%	70
Northern Ireland	3%	30
Run in*	4%	44
Ethnic group		
White	93%	940
Mixed ethnic	1%	9
Black or black British	3%	28
Asian or Asian British	2%	24
Any other group	1%	13
Employment classification		
Managerial & professional	42%	421
Intermediate, small employers & lower supervisory	30%	302
Routine & semi-routine	25%	250
Never worked & 'other'	4%	41
Qualifications		
Don't know/Not applicable	1%	8
Degree or higher education	33%	338
GCE A- level or equivalent (inc foreign qualifications)	17%	172
GCSEs or currently in full-time education	24%	245
No qualifications	25%	251
Household income		
No answer/refused	8%	76
Don't know	6%	63
Under £14,999	17%	174
£15,000 - £24,999	23%	237
£25,000 - £34,999	16%	165
£35,000 - £49,999	13%	130
£50,000 or more	17%	169
Age group		
19-29 years	14%	145
30-39 years	20%	202
40-49 years	18%	179
50-59 years	18%	184
60-69 years	15%	147
70 years and over	15%	157

*The 'Run-in' refers to the pilot sample of the NDNS, collected prior to the main survey, but able to be combined with the main survey results because field procedures remained identical.
NB – Percentages have been rounded to the nearest whole number

3.6.2 Dietary characteristics

3.6.2.1 Energy intake

Energy intakes of the analytical sample followed a normal distribution. Mean daily energy intake was 7699 kJ (95% CI 7544, 7854); mean daily food energy intake was 7242 kJ (95% CI 7103, 7381).

3.6.2.2 Special diets

There were 19 self-reported vegetarians, constituting less than 2% of the sample. In addition, 10% of the sample reported being on a special diet (n=103) (not including vegetarianism and veganism). Of these, more than half indicated that their diet had been recommended or prescribed by a medical professional. The majority of these special diets were weight-reducing diets (n=35).

Excluding all special dieters increased the mean energy intakes of the sample to 7790kJ/d total energy (95% CI 7627, 7953) and 7322kJ/d food energy (95% CI 7177, 7467).

3.6.2.3 Fruit & vegetable consumption

In terms of fruit and vegetable intake, 334 participants (33%) were found to achieve their ‘5 a day’. A greater proportion of ‘achievers’ were in managerial and professional occupations, and had a degree-level qualification. There was a lower proportion of achievers in the lowest income category, and vice versa. A greater proportion of those who did not consume ‘5 a day’ were in the youngest age group, and fell under the current smoker description.

3.6.2.4 Alcohol consumption

Forty per cent of the sample (32% males; 46% females) consumed no alcohol during the diary recording period, whilst 60% (68% males; 54% females) consumed at least some alcohol. Eight per cent (n=34) of men and 2% of women (n=13) had a mean daily alcohol unit consumption above the UK recommended limits (8 units and 6 units for men and women respectively).

3.6.2.5 Food away from home

A minority of the sample prepared or consumed all their food at home over the four days (n=213, 21%). Sixteen per cent (n=163) of participants ate out on all four days, 213 (21%) on three, 212 (21%) on two and 213 (21%) on just one day.

3.6.2.6 Place of purchase

The majority of the sample (74%, n=753) reported that they did their main grocery shop at a large supermarket.

3.6.2.7 Unusual quantity of consumption reported

Participants also indicated whether each day's food intake was the usual amount they tend to eat, less than usual or more than usual. Almost half of the sample (46%, n=464) reported that the amount they recorded was a typical day's consumption for them on all four days. The responses of the remainder are summarised in Table 3.3.

Table 3.3 Number of participants reporting an atypical quantity of food consumed, and number of days on which the atypical amount was reported

'More than usual amount'	'Less than usual amount'					
	0 days	1 day	2 days	3 days	4 days	Total
0 days	464	153	56	20	15	708
1 day	158	45	7	1	0	211
2 days	60	12	3	0	0	75
3 days	12	2	0	0	0	14
4 days	6	0	0	0	0	6
Total	700	212	66	21	15	1014

3.6.3 Anthropometric characteristics

A valid BMI measurement was missing for 76 individuals (8%). Of those for whom the index was available, the median BMI was 26.4kg/m^2 (IQR 22.9 to 30.0kg/m^2). BMI values showed a positive skew (skewness = 0.8). Twenty seven per cent of the sample were classified as obese (n=257), with 65% (n=607) either overweight or obese. The proportions of men and women in the sample classified in each BMI category can be seen in Figure 3.1 and Figure 3.2, and in Table 3.4.

Figure 3.1 Proportions of men in the NDNS sample within each BMI classification (n=434)

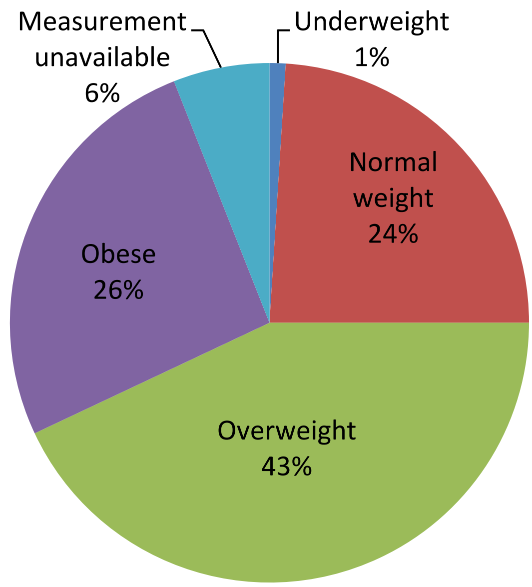


Figure 3.2 Proportions of women in the NDNS sample within each BMI classification (n=580)

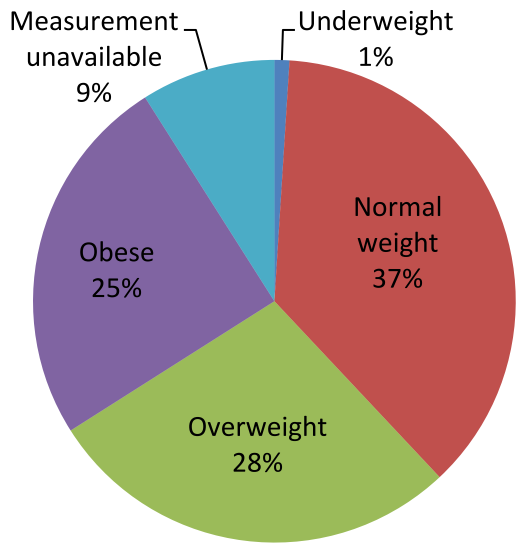


Table 3.4 Proportion of men (n=434) and women (n=580) in the NDNS sample (n=1014) within each BMI classification

BMI classification	Men (n)	Women (n)	Total (n)
Underweight	1% (5)	1% (8)	1% (13)
Normal weight	24% (106)	37% (212)	31% (318)
Overweight	43% (185)	28% (165)	35% (350)
Obese	26% (113)	25% (144)	25% (257)
Measurement unavailable	6% (25)	9% (51)	8% (76)

The characteristics of those without a valid BMI measurement are summarised in Table 3.5. Almost a third (32%) of those with missing BMI data were in the oldest age category (compared to 15% of the full sample). Compared to the full sample, a higher proportion of the participants with missing BMI were female (67% versus 57%), had no qualifications (30% versus 17%) and had missing income data (17% versus 6% did not know their income and 17% versus 8% gave no answer regarding income). The participants without a valid BMI had lower energy intakes than the whole sample (7024kJ total energy (SD 2128kJ); 6643kJ food energy (SD 2010kJ)). A higher proportion had never regularly smoked (58%), consumed no alcohol during data collection (46%) and achieved their ‘5 a day’ (39%) than compared to the whole sample.

Table 3.5 Characteristics of participants without a valid BMI measurement (n=76)

Variable		Proportion of sample (%)	Number of sample (n)
Sex			
	Male	33%	25
	Female	67%	51
Country			
	England	80%	61
	Wales	9%	7
	Scotland	6%	4
	Northern Ireland	1%	1
	Run in*	4%	3
Ethnic group			
	White	93%	71
	Mixed ethnic	0%	0
	Black or black British	7%	5
	Asian or Asian British	0%	0
	Any other group	0%	0
Employment classification			
	Managerial & professional	36%	27
	Intermediate, small employers & lower supervisory	33%	25
	Routine & semi-routine	26%	20
	Never worked & 'other'	5%	4
Qualifications			
	Don't know/Not applicable	3%	2
	Degree or higher education	33%	25
	GCE A- level or equivalent (inc foreign qualifications)	8%	6
	GCSEs or currently in full-time education	26%	20
	No qualifications	30%	23
Household income			
	No answer/refused	17%	13
	Don't know	17%	13
	Under £14,999	8%	6
	£15,000 - £24,999	24%	18
	£25,000 - £34,999	11%	8
	£35,000 - £49,999	8%	6
	£50,000 or more	16%	12
Age group			
	19-29 years	7%	5
	30-39 years	14%	11
	40-49 years	21%	16
	50-59 years	13%	10
	60-69 years	13%	10
	70 years and over	32%	24

*The 'Run-in' refers to the pilot sample of the NDNS, collected prior to the main survey, but able to be combined with the main survey results because field procedures remained identical.
NB - Percentages have been rounded to the nearest whole number

3.7 Univariate analyses

3.7.1 Fruit & vegetable consumption

Participants who achieved their ‘5 a day’ had a higher mean energy intake: 8316kJ (95% CI 8036, 8596) vs 7396kJ (95% CI 7214, 7578) total energy (t=-5.55 (df 1012), $p<0.0001$); and 7827kJ (95% CI 7572, 8082) vs 6955kJ (95% CI 6794, 7116) food energy (t=-5.89 (df 1012), $p<0.0001$).

3.7.2 Alcohol consumption

Mean energy intakes showed a linear increase across rising alcohol consumption categories (as measured by the diet diary), both including and excluding energy from alcohol ($p <0.001$, in both instances; see Table 3.6).

Table 3.6 Energy intakes by alcohol consumption category

	Alcohol consumption category				
	None (n=410)	**Lower risk (n=425)**	**Increasing risk (n=132)**	**Higher risk (n=47)**	**All consumers (n= 604)**
Mean daily energy, kJ (95% CI)	6826 (6619, 7033)	7733 (7514, 7952)	8922 (8537, 9307)	11578 (10660, 12496)	8292 (8086, 12496)
Mean daily food energy kJ (95% CI)	6788 (6582, 6994)	7360 (7144, 7576)	7738 (7370, 8106)	8756 (8018, 9494)	7551 (7368, 7734)

Alcohol consumption differed significantly by most of the sociodemographic variables, using chi^2 comparisons. Men were more likely to consume alcohol than women. The middle age groups (from 30 to 59 years) had the highest proportions of alcohol consumers. There was a greater proportion of alcohol consumers amongst the managerial and professional occupation group, those with a degree qualification, and those who were married or divorced. The proportion of alcohol consumers appeared to increase with household income category. There were smaller proportions of alcohol consumers amongst those who had ‘never worked’, single-person households or households with five or more people, and those who had been widowed. Ex-regular cigarette smokers and those who achieved their ‘5 a day’ were more likely to consume alcohol. All of these were statistically significant.

3.7.3 FAFH

Mean energy intakes (total energy and from food only) increased with the number of days on which food was consumed away from home ($p<0.001$ for both using unadjusted regression) (Table 3.7).

Table 3.7 Mean energy intakes (kJ) according to the number of days on which food away from home (FAFH) was consumed

	Number of days FAFH				
	0 (n=213)	**1 (n=213)**	**2 (n=212)**	**3 (n=213)**	**4 (n= 163)**
Mean daily energy intake, kJ (95% CI)	6853 (6540, 7166)	7298 (6998, 7598)	7937 (7595, 8279)	7856 (7550, 8162)	8815 (8420, 9210)
Mean daily food energy, kJ (95% CI)	6607 (6319, 6896)	6915 (6635, 7195)	7469 (7168, 7770)	7352 (7073, 7631)	8063 (7715, 8411)

3.7.4 Unusual quantity of consumption reported

Energy intakes according to the number of days atypical quantities were reported are shown in Table 3.8. Mean energy intakes generally increased with the increasing number of days on which participants stated intake was more than usual; however, this was not always apparent when participants also reported days on which consumption was less than usual.

Table 3.8 Mean energy intakes (kJ, standard deviations in brackets) according to the number of days on which atypical amount was reported

'More than usual amount'	'Less than usual amount'				
	0 days	1 day	2 days	3 days	4 days
0 days	7587 (2372)	7740 (2609)	7667 (2835)	6930 (3510)	4510 (1852)
1 day	8161 (2741)	7057 (1877)	7947 (2315)	5073 (-)	-
2 days	8445 (2052)	7986 (1849)	6881 (521)	-	-
3 days	8409 (1860)	11535 (65)	-	-	-
4 days	8649 (2900)	-	-	-	-

3.7.5 BMI

3.7.5.1 Sociodemographic differences in BMI

In chi^2 comparisons, proportions of normal weight, overweight and obese participants were found to significantly vary by age group, qualifications, marital status and cigarette-smoking status (Table 3.9). No significant differences were found for any of the other tested sociodemographic variables (survey year, country, ethnicity, household size, employment) or lifestyle variables (alcohol consumption, achievement of '5 a day', FAFH).

Table 3.9 Sociodemographic and lifestyle characteristics of normal weight, overweight and obese NDNS adults (n=925)

Variable	Normal weight (n=318)	Overweight (n=350)	Obese (n=257)	p (chi^2)
Age group				
19-29 years	22% (70)	11% (39)	9% (24)	
30-39 years	24% (76)	19% (66)	19% (49)	
40-49 years	17% (55)	17% (60)	18% (45)	
50-59 years	14% (46)	19% (68)	23% (60)	
60-69 years	11% (36)	15% (51)	19% (49)	
70 years and over	11% (35)	19% (66)	12% (30)	
				<0.01
Employment classification				
Managerial & professional	45% (142)	43% (152)	38% (97)	
Intermediate, small employers & lower supervisory	28% (88)	30% (105)	32% (82)	
Routine & semi-routine	24% (75)	23% (80)	27% (69)	
Never worked & 'other'	4% (13)	4% (13)	4% (9)	
				0.697
Qualifications				
Degree or higher education	39% (125)	34% (119)	27% (69)	
GCE A- level or equivalent (inc foreign qualifications)	21% (67)	16% (57)	15% (39)	
GCSEs or current full-time education	21% (68)	23% (80)	28% (71)	
No qualifications	18% (57)	26% (91)	30% (76)	0.004
Marital status				
Single, never married	36% (114)	25% (88)	22% (57)	
Married and living with partner	40% (126)	48% (167)	54% (140)	
Married but separated	3% (8)	4% (13)	2% (6)	
Divorced	14% (43)	11% (40)	14% (37)	
Widowed	8% (27)	12% (42)	7% (17)	
				0.003

Table 3.9 (cont'd)

Variable	Normal weight (n=318)	Overweight (n=350)	Obese (n=257)	p (chi²)
Cigarette-smoking status				
Never regularly smoked	55% (175)	53% (184)	52% (134)	
Ex-regular smoker	18% (58)	28% (99)	28% (72)	
Current regular smoker	27% (85)	19% (67)	20% (51)	
				0.01
Achieve '5 a day'				
Yes	35% (111)	32% (111)	31% (79)	
No	65% (207)	68% (239)	69% (178)	
				0.522
Alcohol consumption				
None	37% (118)	40% (140)	43% (111)	
Lower risk	47% (148)	41% (144)	39% (100)	
Increasing risk	11% (36)	14% (50)	13% (34)	
Higher risk	5% (16)	5% (16)	5% (12)	
				0.582
FAFH				
None	19% (59)	19% (67)	22% (56)	
1 day	19% (61)	21% (73)	22% (56)	
2 days	23% (72)	23% (80)	18% (46)	
3 days	25% (80)	19% (68)	21% (54)	
4 days	14% (46)	18% (62)	18% (45)	
				0.508
Less than usual quantity of food				
No days	67% (212)	72% (251)	68% (174)	
1 day	21% (68)	20% (69)	22% (57)	
2 days	8% (24)	6% (20)	6% (16)	
3 days	3% (10)	2% (7)	1% (3)	
4 days	1% (4)	1% (3)	3% (7)	
				N/A
More than usual quantity of food				
No days	66% (210)	72% (253)	69% (177)	
1 day	23% (73)	20% (70)	20% (52)	
2 days	9% (28)	6% (20)	9% (23)	
3 days	2% (6)	1% (4)	2% (4)	
4 days	<1% (1)	1% (3)	<1% (1)	
				N/A

Underweight excluded

3.7.5.2 Dietary differences by BMI category

Energy intakes by BMI classification, both including and excluding energy from alcohol, can be seen in Table 3.10. Differences between the normal weight, overweight and obese categories were found to be significant in both cases. The highest mean daily energy intake was found for the normal weight category.

Table 3.10 Daily energy intakes by BMI classification (n=1014)

	BMI category					P*
	Unavail-able (n=76)	Under-weight (n=13)	Normal weight (n=318)	Over-weight (n=350)	Obese (n= 257)	
Mean energy intake, kJ (95% CI)	7024 (6538, 7510)	7443 (6059, 8827)	7901 (7615, 8187)	7866 (7603, 8129)	7434 (7125, 7743)	0.03
Mean food energy, kJ (95% CI)	6643 (6183, 7103)	6832 (5742, 7922)	7440 (7196, 7683)	7386 (7147, 7625)	7000 (6716, 7284)	0.03

* Test excludes underweight

The proportion achieving their ‘5 a day’ did not differ by BMI classification, nor did the proportions in each alcohol consumption category. BMI categories were not found to differ according to the number of days on which an unusual quantity of food was consumed or the number of days food was eaten outside the home (Table 3.9).

3.8 Discussion

This chapter described the characteristics of the unweighted NDNS sample, and summarised the dietary intakes and habits of participants. In addition, methods for deriving new variables (a measure of alcohol consumption, the number of days on which FAFH was consumed, and whether unusual quantities were reported) were outlined. A discussion of how these characteristics may influence the interpretation of later analyses is presented below. The appropriateness of using unweighted estimates is also considered.

3.8.1 Representativeness of sample

The basic sociodemographic differences between the weighted and unweighted samples have already been published in Appendix B of the NDNS report (Bates et al., 2011). Comparisons of further characteristics of the unweighted sample in relation to national statistics are considered below.

3.8.1.1 Dietary characteristics

The mean energy intake of the analytical sample was 7.7 MJ per day, or 1831 kcal (7.2 MJ or 1730 kcal per day excluding alcohol). In the NDNS headline results report, energy intakes are reported separately for those aged 19 to 64 years (7.7MJ/d total and 7.3MJ/d food energy) and those aged 65 and over (6.9MJ/d total and 6.7MJ/d food energy).

A third of the analytical sample was found to achieve five portions of fruit and vegetables per day; this is slightly more than the 30% published in the report.

The proportion of vegetarianism reported in the survey, at less than 2%, is slightly lower than recent national estimates: the 2009 FSA survey on Public Attitudes to Food Issues (GfK Social Research, 2009), for example, found 3% of a nationally representative sample reported being vegetarian.

3.8.1.2 Cigarette smoking and alcohol consumption

Reported cigarette smoking in this NDNS sample was similar to that reported in 2011 by the Health Survey for England (HSE) (Craig and Mindell, 2011), which indicated that 23% of men and 19% of women were current regular smokers, 28% of men and 22% of women used to regularly smoke cigarettes, and 50% of men and 59%

of women had never smoked. This compares to 22%, 24% and 53% respectively in this sample (men and women combined).

The distribution of alcohol consumption varies widely between national surveys. This is most likely due to the variation in the methods used to measure alcohol consumption. Using computer-assisted interview, the HSE found that 31% of men and 46% of women reported consuming no alcohol in the previous week (Craig and Mindell, 2011). This is very similar to the proportions of the NDNS found to consume no alcohol during the four-day diet recording period: 32% of men and 46% women. Amongst those who reported consuming alcohol in the HSE, a 7-day drinking diary indicated that 56% of males and 52% females exceeded the recommended daily limits. This compares to a much lower estimate of NDNS adults who exceeded recommendations: 8% of men and 2% women. It should be borne in mind, however, that the estimates for this sample are derived from an average daily alcohol consumption, rather than the number of participants exceeding recommendations on any one day.

3.8.1.3 Income

The £15,000 - £24,999 household income category was the largest in this sample. This category sits slightly below the national median salary for the 2008-09 tax year of £25,800 (ONS, 2009). The household income reported here, however, does not take into account household composition. This topic is addressed in detail in Chapter 4.

3.8.1.4 BMI

The proportion of the sample classified as obese (26% of men and 25% of women) was slightly higher than that found by the Health Survey for England 2009 (Craig and Hirani, 2010) (22% men and 24% women). However, the percentages that were underweight (1% men and women), normal weight (24% men and 37% women) and overweight (43% men and 28% women) in the NDNS were found to be lower than HSE estimates, which reported 2% of men and women underweight, 32% of men and 41% of women normal weight, and 44% of men and 33% women overweight.

These differences represent minor discrepancies between the samples. As such, and because the HSE does not reflect data from Scotland, Wales and Northern Ireland, the NDNS sample can be judged an adequate representation of the BMI distribution found across the UK.

3.8.1.5 Summary

Many of the descriptive results for the analytical sample presented in this chapter are similar to the findings of other national statistics, as described above. The unweighted sample was therefore judged to adequately reflect national characteristics and deemed appropriate for the purposes of the research aims of this thesis.

3.8.2 Limitations

The assumptions surrounding dietary research form the basis of extensive discussion in the literature, and the associated limitations are well-documented. A summary of the main limitations, together with the potential implication for this research, are presented here.

Firstly, as in most dietary research, the accuracy of the data is dependent on the quality of the self-reported intake, and it is possible that diary entries contain errors and omissions, whether deliberate or unconscious. The suggestion of under-reporting, in this sample as in many others, indicates either a level of inaccuracy in recording or a behaviour modification in response to the measurement itself.

Energy intakes of this NDNS sample are lower than national recommendations for both men and women, which are 10.9MJ/d and 8.7MJ/d respectively (SACN, 2011). Low energy intakes in the NDNS have been commented upon in previous reports, where they have been attributed to under-reporting (SACN, 2011).

Furthermore, energy intakes were found to differ significantly between the BMI categories, with the obese exhibiting a lower mean intake than the other groups. Under-reporting of food intake by obese participants – for whom greater energy requirements would normally be expected – during dietary data collection has been observed in a number of other studies (for example, Rennie et al., 2007). Explanations for the phenomenon include altered diet recording due to social desirability motives, or the possibility that participants are reporting lower intakes due to following a weight-loss diet. In this sample, participants indicated whether or not they were adhering to a special diet; therefore, it will be possible to exclude these dieters in future sensitivity analyses. However, the possibility of a systematic bias in the measurement of diet amongst the obese could severely hamper the ability to draw conclusions in analyses concerning BMI categories. It could mask potential relationships in the data, and will be important to consider when interpreting later analyses. This issue will be further discussed in relation to such analyses, in Chapters 7 and 8.

Unfortunately, the presence of under-reporting is difficult to establish without measurements of energy expenditure or data on physical activity. Data of this type are not available for the full NDNS sample. A subset of the sample has been included in a doubly-labelled water (DLW) substudy (Bates et al., 2011), but at the time of writing, the results of this have yet to be published.

The coding of dietary data is also subject to limitations, including variability between coders, and the accuracy of the food and nutrient information assigned to the diary data. Food composition tables are restricted by the numbers and types of foods contained within them, and nearest alternatives may be substituted by coders where the actual food recorded is not available. As there is likely to be seasonal and manufacturer variation in the nutrient content of many foods, it is important to bear in mind that database-calculated intakes are estimates of actual intake.

In calculating population intakes, there is also an assumption that the behaviour recorded over four days is indicative of habitual intake. In fact, the optimum number of recording days needed to gauge habitual intakes differs according to the nutrient or micronutrient of interest (Willett, 1998). The choice of a four-day diary recording period was made for the rolling programme of the NDNS, as the seven-day diary collection of previous NDNS surveys was felt to be burdensome to participants and less appropriate for certain age groups. Comparison studies of alternative dietary assessment methods are cited in Appendix B of the survey report (Bates et al., 2011), alongside the rationale for tool selection. Whilst this selection indicates the NDNS investigators' confidence in the method chosen, it is unclear what the optimum data collection period would be for diet costs, and this should be borne in mind in later chapters.

3.8.3 Strengths

The NDNS makes use of sophisticated sampling and recruitment methods in order to best gather nationally representative data. It is currently the only representative national dietary survey in the UK, and as such is an important source of information about the population's diets. Comparisons of findings relating to specific variables to those of other national studies confirms the representativeness of the sample, even without employing survey weightings.

Another important advantage of the NDNS is that it collects professionally-measured anthropometry, rather than self-reports. This should help to minimise self-report bias that may be problematic in other studies. Although participants with missing or unavailable valid BMI measurements exhibited some differences in characteristics to the whole sample, the differences are likely to reflect the age profile of these participants, a third of which were in the oldest age category. The unavailability of BMI measurement amongst this age group is not unsurprising, given the difficulties associated with anthropometric measurements in older adults (Hirani and Mindell, 2008).

An additional strength of the NDNS worth mentioning is the collection of information in the diet diaries indicating where each meal was consumed. This is an underused variable in the NDNS data, and to date has only featured in one previously published work (Mak et al., 2012). However, it is of particular importance in the analyses of this thesis, in that it can be used to ascertain which foods were consumed away from home. Food away from home is often more expensive than that consumed within the household (Wrieden and Barton, 2011). The newly derived FAFH variable enables estimation of how much of the diet is eaten, and possibly purchased, away from home. The majority of the sample (79%) reported that they consumed food away from home on at least one day during data collection. The proportion of the sample which ate any food away from home is too large to exclude in later analysis. However, the newly derived variable identifying the number of days on which FAFH was consumed can be used as a covariate in subsequent regression analyses.

3.8.4 Other key points

In addition to the standard limitations identified above, which are widely recognised in dietary research, it is necessary to take into account a few more factors which could be expected to influence dietary intake or food budget during the period of data collection, and as such could possibly confer additional measurement error if habitual behaviour is assumed. These include: periods of sickness, food consumption away from home and special diets. Data regarding each of these potential factors were collected with the survey.

The number of participants reporting a period of illness during data collection, at 12% of the total sample, could skew estimates of energy intake, due to either reduced intake as a result of symptoms, or increased intake to aid recovery. However, because data on the typicality of the quantity consumed were collected for each day of diary recording, it should be possible to identify where illness has caused reduced or increased intake. Therefore, the indication of an unusual quantity consumed will be employed in sensitivity analyses, as opposed to reported illness.

A small proportion (10%) of the sample reported being on a special diet. As the restrictions imposed by the diet could over-ride other considerations in dietary choices, this could impact on the food budget, as well as influencing dietary intake. The influence of special diets will be examined in sensitivity analyses.

Place of purchase could influence dietary expenditure; however, it is not possible to account for this using the food cost database to estimate diet cost. However, given that almost three quarters of the sample listed supermarkets as the location of their primary household shopping, the measurement error associated with alternative places of purchase is likely to impinge on only a minority of the sample. Nevertheless, it is a factor to be borne in mind when interpreting the diet cost estimates (see Chapter 5).

Alcohol consumption is a behaviour of relevance to dietary research, even where energy from alcohol is excluded from analyses. In the NDNS sample, there was a positive relationship between energy intake – both including and excluding alcohol – and alcohol consumption category. This potentially indicates an increase in food consumption alongside the drinking of alcohol, or identifies a common underlying personality characteristic whereby those who are disposed to consume alcohol are also inclined to eat more.

Finally, it is important to note that, because the survey was designed to be nationally representative, older adults make up a large proportion of the sample, with almost a third (30%) aged 60 years and over. There are three key points arising from this demographic profile. Firstly, with no upper age limit in the NDNS recruitment, the oldest age category, '70 years and over', encompasses a broad range of ages (up to 94 years). The heterogeneity of this population group in terms of health, mobility and nutritional status is widely recognised (Keller, 2007). Secondly, the physiological and lifestyle changes associated with ageing are likely to impact on food selection, energy intake, and body composition (Gariballa and Sinclair, 2005). This may mean that the food price-obesity hypothesis is not as applicable to this age group. Thirdly, many of the adults in these top two age bands are likely to be retired. Not only does this imply this population is on lower incomes than younger age groups, but it may be that, due to the physical and social changes associated with ageing and retirement, the income they do report is not equivalent to that of working-age adults. In other words, the demand for other goods in later life will affect the proportion of income available for food. For example, spending more time in the home could increase demand for household heating – the consequence of which could be reduced purchasing power for food. (Indeed, seasonal variability in food insecurity has been observed amongst low-income elderly populations in the US (Nord and Kantor, 2006).)

In order to preserve the sample size for analyses in later chapters, and to describe diet costs for the UK population, the older adults will be included in the analytical sample. However, in many cases they will be excluded as part of sensitivity analyses.

3.9 Conclusions

The description of the sample and its dietary characteristics sets the scene for the interpretation of analyses in subsequent chapters. Data support that the analytical sample is appropriate in its description as a nationally representative sample. However, there are important considerations regarding methodological limitations that need to be taken into account. In particular, the NDNS, like all dietary surveys, suffers the potential for bias brought about by mis-reporting.

Despite this, the NDNS offers some data on important variables that may be crucial to the interpretation of later analyses. The consumption of food away from home, place of main grocery purchase and reporting of atypical quantities of food consumed will all be useful in clarifying the relationships examined in subsequent chapters.

BMI categorisation was found to significantly vary by age group, qualifications, marital status and cigarette-smoking status. This hints at the potential roles that economic and socio-demographic factors may play in the establishment of weight status. These potential relationships are explored in Chapter 4.

Chapter 4 Income in the NDNS

4.1 Summary

This chapter introduces the first empirical analyses of this thesis. In considering the micro-economic determinants of obesity, the primary focus of the chapter is on income, as an important factor in purchasing power.

As a determinant of food budget, income might be expected to influence food purchases, thereby affecting dietary consumption. People with a tighter budgetary constraint may need to adopt coping strategies in order to meet energy requirements, potentially selecting more energy-dense foods as a result. Energy-dense diets are linked to higher energy intakes and could therefore promote positive energy balance.

This chapter explores the relationship between income and energy density (kJ/g) and between income and body mass index (BMI) amongst adults in the NDNS. Findings reported in the literature are inconsistent on these topics, and this chapter includes a discussion about a possible explanation for this: much of the literature uses a measure of household income without accounting for household composition.

The concept of equivalization is introduced in the chapter, and the income distribution – both household and equivalized – of NDNS participants is described. Equivalization involves weighting household income to account for differences in household composition, and is seldom employed in nutritional epidemiology, although commonly used in economic studies and national statistics. The impact of equivalizing income on the results of the analyses is considered.

The results suggest a negative association between equivalized income category and energy density (excluding non-milk beverages). However, the trend was only statistically significant in linear models when those who reported their intake as unusual were excluded from analyses. The association was not evident in linear models when the non-equivalized income was used.

A significant negative association was also evident between income category and BMI, but only for the equivalized income variable. Furthermore, the odds of being overweight or obese were significantly lower with increasing categories of equivalized income but no significant result was obtained using household income categories.

The results illustrate the importance of equivalizing income to account for household composition appropriately. In this sample, equivalizing income had the effect of reclassifying 42% of participants. This potential misclassification could have important repercussions in research investigating income and health.

4.2 Introduction

Chapter 1 established overweight and obesity as the major health challenge of the 21st century (WHO, 2007, Butland et al., 2007). The aim of this thesis is to examine micro-economic factors in excess energy intake, considering both demand-side and supply-side factors in purchasing power. Income is a key demand-side factor due to its role as a determinant of food budgets (see Chapter 1). Income is also an available variable in the NDNS data set, therefore making it appropriate for examination in this chapter.

Socio-economic disparities have been widely documented in terms of both diet (James et al., 1997) and health (Marmot and Bell, 2012); however, much of the literature features an aggregate measure of socioeconomic status, making it difficult to tease out the independent influence of income (see Chapters 1 and 2). Aggregate measures are useful in defining inequalities, for which a single variable may be inadequate for capturing a complex phenomenon. On the other hand, it is recognised that the components of socioeconomic measures – most commonly, income, education and occupation – could vary in their predictive value because they are assumed to reflect different underlying mechanisms (Winkleby et al., 1992, Macintyre et al., 2003).

Income as an independent predictive factor is of importance in this thesis because of its role as a limiting factor in food purchasing. People with a tighter budgetary constraint may need to adopt coping strategies in order to meet energy requirements, potentially selecting more energy-dense foods as a result (CARE/WFP, 2003). Energy-dense diets have been linked to higher energy intakes (Prentice and Jebb, 2003) and could therefore promote positive energy balance and weight gain. Changes to diet in response to income ‘shocks’ have been documented (for example, von Hinke Kessler Scholder and Leckie, 2013, in Russia), highlighting the role of income in diet selection.

Despite this, the modest literature published on income and diet offers a mixed picture. Whilst several studies have documented dietary differences between higher-income and lower-income consumers, with better quality diets (variously defined) observed amongst those on higher incomes (Cassady et al., 2007, Cade et al., 1999, Darmon and Drewnowski, 2008, Hiza et al., 2013), this pattern has not always emerged (Waterlander et al., 2010, Du et al., 2004). Within the UK, the Low Income Diet and Nutrition Survey (LIDNS) suggested that dietary differences between income groups do not appear to be clear-cut: compared to the general UK population, LIDNS participants reported a lower consumption of wholemeal bread and vegetables and a higher consumption of soft drinks, processed meats, whole milk and sugar. On the other hand, the consumption of most foods and calculated nutrient intakes were broadly similar to the UK as a whole, as were the proportions of overweight and obese in the LIDNS sample (Nelson et al., 2007).

A small number of studies in adult populations have investigated dietary energy density as a marker of diet quality, in relation to income: three of these (Kant and Graubard, 2013, Waterlander et al., 2010, Aggarwal et al., 2011) are discussed in the systematic review presented in Chapter 2. A further study (Monsivais and Drewnowski, 2009) was not included in the review because the article did not present the results of formal comparison tests, and another two studies (Ricciuto and Tarasuk, 2007, Wrieden and Barton, 2011) were not included because they estimated dietary energy density from expenditure data, rather than measuring diet itself. Ricciuto and Tarasuk (2007) found a strongly significant negative relationship between income and energy density, excluding beverages, calculated from a Canadian national expenditure survey. Wrieden and Barton (2011), using Scottish expenditure data, also found significant deprivation group differences (using the Scottish Index of Multiple Deprivation) in both the energy density from food only, and that calculated from food and milk. In contrast, findings of studies based on dietary data were mixed. Monsivais and Drewnowski (2009) found no significant differences between categories of income in dietary energy density of food only; however there were some differences that reached statistical significance when beverages (except water) were included in energy density estimates. From Chapter 2 it can be seen that only two studies (Kant and Graubard, 2013, Aggarwal et al., 2011) clearly demonstrated a negative relationship between income and dietary energy density, whilst Waterlander et al. (2010) failed to find any income group differences in their study.

Contrasting findings may be due to the contexts of these studies – the findings of Waterlander et al. (2010), for example, included a sample of elderly Dutch adults. The other possibility is that there is disagreement on the effects of income on diet because income has been variously defined or measured in the different studies. Precise measurement of income is recognised to be difficult in survey design, and often broad categories of household income are used. Surveys also differ in whether they ask participants to include sources of income other than salary, and which sources they include. A key drawback of household income is that it is not equivalent across different household compositions: a household of two adults with an income of £30,000, for example, is likely to access a different standard of living than a family of six on the same income.

Equivalization is a method of weighting household income to take into account the size and composition of the household (simply put, the number of adults and children). Several equivalization scales have been developed, varying in their complexity – a summary of the most commonly employed has been compiled by Chanfreau and Burchardt (2008). Equivalization is a technique widely employed in economic studies and national statistics, but is seldom employed in nutrition research. Only one of the above studies investigating income and energy density (Kant and Graubard, 2013), for example, appears to have used equivalized income – one was not even able to account for household size (Waterlander et al., 2010).

In obesity research, national statistics, such as those from the Health Survey for England (HSE), do often employ an equivalized income variable. Data from the HSE (NOO, 2010) suggest a linear decrease in the prevalence of obesity with increasing quintiles of equivalized income for females, though the pattern for men is less clear. However, this trend was not formally tested for significance in the HSE. Obesity studies which include formal analyses using equivalized income have tended to come from outside the UK. For example, in Germany, the odds of being obese have been found to be higher amongst the lower and middle tertiles of equivalized income for both adult men and women (Schumann et al., 2011) and in the US the lowest tertiles using the Poverty Income Ratio (PIR) had higher odds of obesity than the highest tertile amongst participants in the National Health and Nutrition Examination Survey (NHANES) (Ali et al., 2011). However, as far as the author is aware, the relationships between equivalized income and BMI or the odds of being obese are yet to be tested formally in a UK sample.

The aim of this chapter is to explore the relationship between income and dietary energy density and the relationship between income and BMI amongst adults in the National Diet and Nutrition Survey (NDNS, see Chapter 3). In addition, the possibility that inconsistent findings in the literature could be due to a reliance on family income without accounting for household composition is considered. This is achieved by repeating analyses using either crude household income, household income adjusted for household size, or equivalized household income, and comparing the results. A similar approach is adopted by Benzeval et al. (2001) to identify the impact of equivalization of income in their study of long-term illness and self-reported health.

The following objectives will be addressed in this chapter:

1. To derive an equivalized income variable for adult participants in the NDNS;
2. To calculate and describe the dietary energy density of NDNS adults;
3. To assess whether income is related to dietary energy density amongst NDNS adults;
4. To examine the relationship between income and BMI or overweight and obesity amongst NDNS adults; and
5. To compare the influence of using equivalized versus crude household income or household income adjusted for size in testing the above relationships.

4.3 Methods

4.3.1 Sample

NDNS adult data from 2008-2010 were used. Further details about the survey design, sample recruitment and characteristics can be found in Chapter 3, as well as in the survey report (Bates et al., 2011). This section will briefly repeat the data collection methods used for the key exposure (income) and outcome variables (energy density and BMI). Following this, there will be a more detailed description of the methods used to derive new variables from the original data.

Household income was assessed by interview, with respondents placing themselves in one of 13 income brackets. Income data were given by 889 participants (the remaining 14% of the sample responded 'No answer/refused' or 'Don't know'). Categories were collapsed to five groups to facilitate analysis.

BMI was calculated from height and weight measurements taken by the interviewer during the first survey visit. Valid BMI measurements were unavailable for 81 participants. BMI classifications are based on WHO categories – overweight defined as a BMI between 25 and 29.9kg/m^2; obesity as 30kg/m^2 or over (see Chapter 3).

Dietary data was collected using four-day diet diaries (see Chapter 3). Diaries are coded by the NDNS data creators using DINO (Diets In Nutrients Out) software, which estimates energy intakes using nutrient information from the UK food composition tables (FSA, 2002). Mean daily energy intake is a readily derived variable in the NDNS data sets, both in terms of total energy, and separately for food energy and alcohol energy. Kilojoules was the standard measurement adopted in these analyses. Some adults in the survey completed only three days' worth of dietary data (n=17).

Only adults with complete diary data (four days) were included in the analytical sample. In addition, those without valid anthropometric measurements had to be excluded from analyses involving BMI. The analytical samples therefore comprised 875 and 814 respectively, from a possible 1014 adults. Unweighted sample data were used (see Section 3.3.1 for a discussion about the NDNS sample weighting scheme).

4.3.2 Derivation of equivalized income

There are a number of equivalence scales in use (see Chanfreau and Burchardt, 2008). The choice of scale used in these analyses was based upon the selection used for national figures. Until 2005/6, UK government statistics applied the

McClements scale (for example, in the Households Below Average Income, HBAI, report (DWP, 2013)). From 2006 onwards, the modified OECD scale was adopted to bring statistics in line with that of other departments and other members of the EU (Anyaegbu, 2010). The main difference between the scales is that the modified OECD scale assigns a single value to all children aged 14 years and under, and is therefore simpler than the McClements scale which includes several age-dependent values for children. The McClements scale uses a reference category (assigned a value of ‘1’) of a two-adult household. In this chapter, a rescaled modified OECD equivalence scale has been used to similarly assign a value of ‘1’ to a reference category of a two-adult household, as was practiced by Anyaegbu (2010).

To derive equivalized income, the midpoint of each household income category was used (with the exception of the extreme highest category, “£100,000 and over”, for which a value of £100,000 was chosen). The following formula was used to assign each participant with an equivalence index based upon the rescaled modified OECD equivalence scale:

$$\text{equivalence index} = (\#\text{Children}*0.2)+(((\#\text{Adults}-1)*0.33)+0.67).$$

Equation 4.1

Using this index, equivalized income could then be derived in the following manner:

$$\text{equivalized income} = \frac{\text{household income}}{\text{equivalence index}}.$$

Equation 4.2

The scale assumes a reference category of a two-adult household: the first adult in a household is allocated a value of 0.67, with each additional adult contributing a value of 0.33, and each additional child a value of 0.2. This takes into account economies of scale. The continuous equivalized income variable was then categorised to match the original NDNS household income classifications. Categories were collapsed to five groups to facilitate analysis.

4.3.3 Calculation of energy density

Energy density is typically defined as the average amount of energy consumed per gram of food in the diet. Several calculation methods have been employed in studies of energy density. These differ in their treatment of beverages in the calculation

– because liquids can contribute a disproportionate quantity of mass to the diet without adding much (if any) energy, investigators have variously excluded different types of beverages from energy density calculations. However, some beverages can contribute significant amounts of energy to the diet, and some beverages (such as milk) are consumed as foods as well as beverages, which has led to debate about which beverages, if any, should be excluded. Ledikwe et al. (2005) identified eight calculation methods:

- All food and beverages
- Food and energy-containing beverages
- Food, juice and milk
- Food and juice
- Food and milk
- Food and alcohol
- Food and liquid meal beverages
- Food only.

A comparison of these calculation methods in a US sample (Ledikwe et al., 2005) found substantial variation in energy density values. The authors concluded that the choice of method should reflect the purposes of the study. For instance, including liquid meals could be important amongst populations which consume a large amount of liquid meals, but is unlikely to make a significant impact in a more varied population.

In the comparison study, Ledikwe et al cautioned against including all beverages except water in calculating energy density. This they argued was because people who consume mainly water, at the expense of other beverages, would be assigned a higher dietary energy density value than those who consume, in particular, low energy beverages.

Dietary energy density was also the subject of a recent study in Scotland (Wrieden and Barton, 2011), in which values obtained through different calculation methods were compared for the Scottish population. Like Ledikwe et al above, the authors found substantial differences in energy density estimates depending on the method: including all beverages for example was found to halve the mean energy density of the sample. Based on their results, and in keeping with the WCRF recommendations, the authors advocated obtaining energy density estimates from food and milk, excluding non-milk beverages.

Energy density in this chapter was therefore calculated using the energy and mass totals from food and milk, excluding all non-milk beverages. The NDNS data sets were imported into an Access database (to enable the diet cost calculations of later chapters – see Section 6.3.1). Total intakes for energy (kJ) and mass (g) of food and milk were summed in Access. The totals were exported into the Stata data set, in which the new variable of energy density was generated by dividing the summed energy by the summed mass (kJ/g).

4.3.4 Analytical methods

The analyses in this chapter were designed to address the final three objectives of the chapter:

3. To assess whether income is related to dietary energy density amongst NDNS adults;
4. To examine the relationship between income and BMI or overweight and obesity amongst NDNS adults; and
5. To compare the influence of using equivalized versus crude household income or household income adjusted for size in testing the above relationships.

The sections below detail the analytical approaches taken for the key relationships under investigation.

All analyses were performed using Stata IC 12 (StataCorp, 2011).

4.3.4.1 Descriptive analyses

Descriptive analyses were run for the following variables: BMI, overweight and obesity, dietary energy density, household income and equivalized income. Sample means and standard deviations (or medians and interquartile ranges (IQR) where appropriate) are presented.

In addition, mean dietary energy density was calculated for sociodemographic subgroups, and by BMI classification. Univariate tests (ANOVA) were used to identify differences in energy density between these sociodemographic or BMI categories.
To gauge the effect of equivalization on participants’ categorisation into income bands, a contingency table was produced. The contingency table was produced by running a cross-tabulation of the frequency count in each income band using household income against equivalized income. Agreement between the income variables is denoted by the numbers falling into the diagonal cells: in other words, the diagonal cells (bold in **Figure 4.1**) show the number of participants that would be assigned to, for example, category A, regardless of the method used to define income. The contingency table can also be used to determine where the differences in categorisation lie.

Figure 4.1 A contingency table

Category	A	B	C	D
A	**XXXX**	xxx	xxx	xxx
B	xxx	**XXXX**	xxx	xxx
C	xxx	xxx	**XXXX**	xxx
D	xxx	xxx	xxx	**XXXX**

4.3.4.2 Income & dietary energy density

Mean energy density (with standard deviations) were calculated for each income category – using the crude household income variable and the equivalized income variable. Means were compared in univariate analyses (ANOVA).

To account for household size (but not composition), means adjusted for household size were also computed for the crude household income categories, and compared using ANOVA.

The relationship between income category and dietary energy density was compared using multivariable linear regression, adjusting for age, sex and occupation. The process of covariate selection is described in Section 4.3.4.6 below. Three models were run: these are detailed in Table 4.1.

Table 4.1 Summary of the variables included in the regression models investigating income and dietary energy density

Model	Exposure	Outcome	Covariates
1a	Crude household income	Dietary energy density	Age Sex Occupation
1b	Household income	Dietary energy density	Age Sex Occupation **Household size**
1c	Equivalized household income	Dietary energy density	Age Sex Occupation

In Models 1a, 1b and 1c above, the income variables were entered as dummy variables using the Stata 'i.' prefix. Due to the ordinal nature of these categories, it may be more appropriate to treat these variables as linear categories, rather than dummy variables, in the regression model. Therefore, a further three models (denoted 2a, 2b and 2c) were run in which the income variable was entered without the ‘i.’ prefix. These models were adjusted for the same covariates as identified for models 1a, 1b and 1c.

Sensitivity analyses were planned for the models described above. Firstly, it may be the case that participants who have consumed an unusual amount (reported for each day in the diet diary as ‘more than usual’ or ‘less than usual’) also consumed unusual or atypical types of foods. Secondly, those who are adhering to a special diet (see Section 3.6.2.2) are likely to make different dietary choices than they would do normally. Participants indicated whether they were following a special diet during the diary data collection period. In both of these instances, the atypical dietary choices have the potential to interfere with the hypothesised relationship between income and diet. Therefore, each of the models was run with the following sensitivity analyses:

- Excluding those who reported consuming an unusual amount (more or less than usual on any one day)
- Excluding those who reported that they were following a special diet.

4.3.4.3 Income & BMI

Analyses for the continuous BMI outcome were similar to those carried out for dietary energy density (above). Firstly, univariate analyses tested for differences in BMI between income groups – Kruskal Wallis ANOVA was used due to the skewed distribution of BMI.

The relationship between income and BMI was then further tested using multivariable linear regression analyses, in order to account for confounders. (Although the outcome, BMI, was skewed – see Section 3.6.3 – the linear regression model met the assumptions, and residuals were normally distributed.) As above, a model was run for each income variable, with a further model adjusting for household size (Table 4.2).

Table 4.2 Summary of the variables included in the regression models investigating income and BMI

Model	Exposure	Outcome	Covariates
3a	Crude household income	BMI	Age Sex Occupation
3b	Household income	BMI	Age Sex Occupation **Household size**
3c	Equivalized household income	BMI	Age Sex Occupation

In models 3a, 3b and 3c, the income variables were included in each model as dummy variables. A further three models – 4a, 4b and 4c – were run in which the income variables were treated as linear, rather than being entered using the ‘i.’ Stata prefix (see Section 4.3.4.2).

In addition to excluding those participants with incomplete diary data and those without a valid BMI measurement, underweight participants (BMI <18.5kg/m^2; n=13) were also excluded from the univariate and the regression analyses. As well as representing a small subgroup size, it was felt these participants – who it is assumed have experienced negative energy balance resulting in underweight – differ in their experience to those who have seen positive energy balance resulting in overweight or obesity. As overweight and obesity form a key focus in the aims of this thesis, the mechanisms underlying underweight were felt to be beyond the scope of this chapter.

No sensitivity analyses were judged practical for these analyses. In contrast to the analyses planned above (Section 4.3.4.2), excluding those who were on a special diet or reported unusual quantities of food would be expected to have little effect, because dietary data do not feature in the income-BMI models.

4.3.4.4 Income & overweight + obesity

Using a continuous variable, where available, will provide more information for a regression analysis than would categories of that variable (Naggara et al., 2011). Nevertheless, classifications of BMI are useful clinically in estimating risk of disease (WHO, 2006). For this reason, logistic regression models were also built to examine the income-BMI relationship. Due to lower participant numbers in the obese category, overweight and obese categories were combined to facilitate analyses. Logistic regression therefore investigated the odds of being classified as overweight or obese as opposed to being normal weight.

To enable this analysis, a binary outcome variable was generated where ‘0’ denotes normal weight (BMI between 18.5kg/m^2 and 24.9kg/m^2) and ‘1’ denotes overweight or obese (BMI of 25kg/m^2 or over). Descriptive analyses described the proportions of overweight and obese in each income category (crude household income and equivalized income). Chi^2 analyses were run to indicate differences between income bands. Confounding variables were then included in multivariate analyses of the effect of income on the odds of being overweight or obese (Table 4.3).

As described in the previous two sections (4.3.4.2 and 4.3.4.3), three models were run – one for each definition of income. Adjusted odds ratios are presented alongside 95% CI. The selection of covariates to include is described in Section 4.3.4.6. In all models, underweight participants were excluded (see Section 4.3.4.3).

Table 4.3 Summary of the variables included in the logistic regression models investigating income and overweight or obesity

Model	Exposure	Outcome	Covariates
5a	Crude household income	Overweight or obese	Age Sex Occupation
5b	Household income	Overweight or obese	Age Sex Occupation **Household size**
5c	Equivalized household income	Overweight or obese	Age Sex Occupation

Similarly to the previous regression models described (Sections 4.3.4.2 and 4.3.4.3), these logistic regression models (5a, 5b and 5c) were also run with the income variables treated linearly (models 6a, 6b and 6c), as opposed to being entered as dummy variables.

As above, no sensitivity analyses were deemed necessary.

4.3.4.5 Statistical power

The analyses described above are secondary analyses of an existing data set. The NDNS will have been powered to detect the survey's primary aims, and not necessarily the outcomes identified in this chapter. Therefore it is important to consider the power of the data to investigate the aims of this study, even though the sample size has already been dictated.

With an established sample size, the statistical power of the study to detect a desired effect size can be estimated. The hypothesized effect size is based on judgement – often based on the findings of previous literature or on clinical relevance. More power is necessary to detect a smaller effect size and vice versa, so the choice of a desirable effect value has important ramifications (Whitley and Ball, 2002).

In the interest of parsimony, a recommended approach is to frame the effect size on the expected difference in the outcome between two groups of the sample, dichotomising on the predictor variable (Greenwood, 2011). There are three key relationships that form the focus of this chapter: income and dietary energy density; income and BMI; and income and overweight or obesity. A calculation of power is necessary for each of these, so a hypothesized effect size will need to be chosen in each case, based upon the difference in the outcome that would be expected between high-income and low-income NDNS participants (see below).

From the hypothesized effect size, the standardized mean difference (SMD) can be calculated (the effect size divided by the expected standard deviation of the sample mean). This, combined with the desired α (typically 0.05 or 0.01) and the known sample size, forms the basis for sample size calculations using a nomogram. The nomogram is a method for determining power or sample size graphically, first proposed by Altman (1991) and described by Whitley and Ball (2002).

4.3.4.5.1 Income & energy density

Four studies report the outcome of dietary energy density according to income group. Three of these (Wrieden and Barton, 2011, Monsivais and Drewnowski, 2009, Ricciuto and Tarasuk, 2007) report energy density in kcal, and have been converted to kJ to allow comparison (using a conversion of 1kcal=4.18kJ). The studies varied in the number of groups compared: Ricciuto and Tarasuk examined deciles; Monsivais and Drewnowski compared four categories; Wrieden and Barton included quintiles; and Waterlander et al (2010) used tertiles for one sample and binary groups for the second. These studies also used varying methods of calculating energy density. Comparing the mean energy density of the extreme categories results in differences ranging from 0.04kJ/g to 0.92kJ/g.

Table 4.4 Summary of effect sizes from the literature investigating income and energy density

Study	Number of income categories	Difference between extreme categories (kJ/g)	Energy density calculation
Ricciuto and Tarasuk, 2007	10	-0.92 (fitted regression line)	Food only
Monsivais and Drewnowski, 2009	4	-0.54	Food only
		-0.46	Food + beverages, excluding water
Waterlander et al, 2010	3 (AGHLS sample)	Men 0.29 Women -0.29	Food only
	2 (LASA sample)	Men 0.62 Women -0.04	Food only
Wrieden and Barton, 2011	5*	-0.79	Food only
		-0.38	Food + milk

*categories of Scottish Index of Multiple Deprivation

Of the above, the paper by Wrieden and Barton is the only one to calculate energy density incorporating milk and food, as is the method used in this chapter. Using the difference, 0.38kJ/g, as the anticipated effect value, with a standard deviation of 1.4kJ/g, gives an SMD of 0.271. The analytical sample available for this analysis was 875. Using this information to trace the line through the nomogram, at α 0.05, estimates the power to detect this difference between groups to be 0.97, or 97%. This anticipated effect size 0.38kJ/g, was derived from a study which differs methodologically to the current analyses, in that purchasing and not dietary data were used to calculate energy density. Therefore a number of other potential effect sizes and significance scenarios are also presented in Table 4.5.

Table 4.5 Estimated power of the NDNS sample to detect hypothesized effect sizes (differences in dietary energy density) at significance cut-offs of 0.05 and 0.01

Anticipated effect size (difference, kJ/g)	SMD	Significance level (α)	power
0.38	0.271	0.05	0.97
		0.01	0.91
0.6	0.429	0.05	>0.995
		0.01	>0.995
0.2	0.143	0.05	0.80
		0.01	0.60

4.3.4.5.2 Income & BMI

The hypothesized difference in mean BMI between high- and low-income groups of the NDNS could be extrapolated from the most recent Health Survey for England (HSE) (Hirani, 2011), which reports the following age-standardized BMI means for equivalized income categories:

Table 4.6 Age-standardized mean BMI (kg/m^2) by quintiles of equivalized* income

Equivalized income quintile	Men	Women
Lowest	27.2	28.2
	27.4	28.0
	27.6	27.4
	27.8	26.7
Highest	27.5	25.8

*Equivalized using the McClements equivalence scale

Although there is no discernible trend amongst men, the data above suggest a negative trend between income and BMI amongst women, with a difference of 2.4kg/m^2 between the extreme income groups. A more modest difference would be expected if dichotomised groups, rather than extreme quintiles, were compared, especially if males and females are combined in analyses. Therefore, Table 4.7 lists the power calculations for a few hypothesized (dichotomous) effect sizes, using the nomogram as above. The SMD in this instance was computed by dividing the effect size by the standard deviation in the NDNS adults' BMI distribution, which is 5.3kg/m^2 (see Chapter 3). The sample size is reduced due to missing BMI measurements, allowing just over 800 participants. Table 4.7 indicates that the sample is adequately powered to detect a difference of 1.5kg/m^2 or more between groups.

Table 4.7 Estimated power of the NDNS sample to detect hypothesized effect sizes (differences in BMI) at significance cut-offs of 0.05 and 0.01

Anticipated effect size (difference, kg/m^2)	SMD	Significance level (α)	power
2.4	0.453	0.05	>0.995
		0.01	>0.995
1.5	0.283	0.05	0.97
		0.01	0.96
1	0.189	0.05	0.75
		0.01	0.50

4.3.4.5.3 Income & overweight or obesity

The analyses investigating overweight and obesity involve logistic regression models; therefore the hypothesized effect will be a difference in proportions of a dichotomised sample. The SMD for proportions are calculated with the following formula:

$$\frac{(p^1 - p^2)}{\sqrt{[\bar{p}(1 - \bar{p})]}}$$

Equation 4.3

where $p^1 - p^2$ is the difference and $\bar{p}$ is the mean of the two proportions (Whitley and Ball, 2002).

Data from the HSE (Hirani, 2011) are available giving prevalences of overweight and obesity combined, by equivalized income quintile. These indicate a difference of 8% amongst men and 13% amongst women between extreme quintiles

(Table 4.8). However, the direction of the trend is opposite in men to that of women, making it difficult to select a hypothesized effect size for analyses of combined men and women. Table 4.9 therefore shows a number of possible effect sizes and the power of the NDNS sample to detect each, as calculated using the nomogram. The calculations indicate that the study is not adequately powered to detect a difference in proportions less than 10%.

Table 4.8 Proportions of men and women classified as overweight or obese in the HSE, by equivalized* income quintile (%)

Equivalized income quintile	Men	Women
Lowest	63	62
	65	63
	67	60
	73	57
Highest	71	49

*Equivalized using the McClements equivalence scale

Table 4.9 Estimated power of the NDNS sample to detect different effect sizes (differences in proportions of overweight and obese)

Anticipated proportion overweight/obese (high income, low income)	Anticipated effect size (difference)	SMD	Significance level (α)	power
60%, 70%	10%	0.210	0.05	0.86
			0.01	0.70
63%, 68%	5%	0.105	0.05	0.30
			0.01	0.14
64%, 66%	2%	0.042	0.06	0.08
			<0.05	-

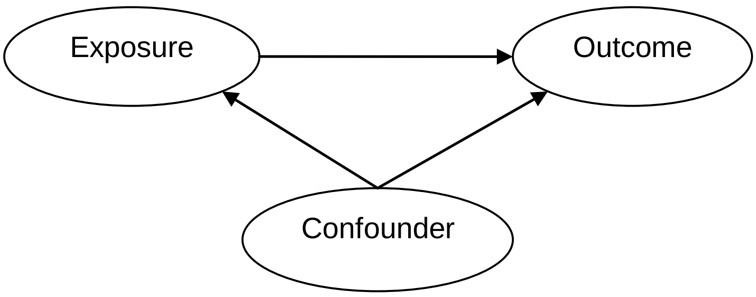
4.3.4.6 Selection of covariates

Regression goes beyond simply establishing an associative relationship: there is an underlying hypothesis regarding the *causal* relationship between exposure and outcome. In observational studies, especially those investigating chronic disease, there are likely to be other factors besides the exposure that also affect, or are causally related to, the outcome variable. Some of these may independently influence the outcome. On the other hand, where the aetiology is acknowledged to be multifactorial (as it is for obesity), many factors may also be related to each other, causally or by association. These inter-relationships mean that an observed exposure-outcome association could in fact be a reflection of other underlying relationships. In order to make a statement about the role of the exposure variable in the aetiology of the outcome, it is necessary to isolate its influence.

In an experimental setting, the influence of a causal variable can be isolated by controlling for other influences in the experimental design, and randomising participants to an exposed or control group. In epidemiology, controlling for the influence of other variables must be achieved by adjusting the regression model to include these influences. In this manner, the effect found of the exposure on the outcome takes into account the effects of the additional variables.

If a variable independently affects the outcome, controlling for its effect will not alter the observed effect of the exposure of interest. If the variable influences the exposure, with no direct effect on the outcome, it merely represents a step higher up the causal pathway. It is only where a variable has an effect both on the exposure and on the outcome (as in Figure 4.2) that the exposure-outcome relationship is confounded, and, if uncontrolled for, it cannot be determined whether the observed relationship is a true relationship between exposure and outcome or actually reflects the influence of the confounder.

Figure 4.2 Illustration of a confounding variable's relationship with exposure and outcome



Causal diagrams can help to establish the relationships between proposed determinants of the outcome. A directed acyclic graph (DAG) is a form of causal diagram that incorporates a priori assumptions about causal relationships in order to identify appropriate confounding variables (Greenland et al., 1999, Glymour, 2006). The premise of the DAG is that each variable is connected by arrows which demonstrate the direction of influence from one variable to another. Arrows are unidirectional, cementing the process of cause and effect. This unidirectionality also prevents cyclical relationships within the graph (hence the term 'acyclic').

The process of creating a DAG provides a rigorous method for working through whether or not variables confound the relationship under investigation. A DAG was created for both of the relationships forming the focus of this chapter: income and dietary energy density (Figure 4.3); and income and BMI (Figure 4.4). In each of these DAGs, the variables included in the regression models are in colour: green for the exposure variable, purple for the outcome and orange for the confounding variables.

From each DAG, it is possible to work backwards along the causal routes, or trace the ‘open backdoor pathways’ (Greenland et al., 1999) to find common causes of both exposure and outcome. Any variable along the ‘backdoor pathway’ can be adjusted for in the regression; however, it is recommended that adjustments are made at the minimum number possible in order to maximise the robustness and efficiency of the model (Bowers, 2008) and reduce the potential for collinearity.

In Figure 4.3 and Figure 4.4, the ‘backdoor pathways’ are depicted by the coloured arrows. Along each pathway, one variable has been selected (shown in orange) as a suitable adjustment. Both DAGs identified the same confounding variables: age, sex and employment (or occupation). Clearly it is reasonable to expect a link between occupation and income. Age is also connected to income, as increasing experience can be expected to attract higher pay grades. National statistics also identify a discrepancy in incomes between males and females in the UK (ONS, 2009).

These three variables were also linked to the outcome of dietary energy density in Figure 4.3. For age and sex, this putative causal pathway could be traced through the influence of age and sex on alcohol consumption, where alcoholic beverages will influence dietary energy density due to their liquid property. Although alcoholic beverages were excluded in the calculation of energy density in this sample (see Section 4.3.3), the literature suggests that food choices – including that of more energy dense foods – are associated with alcohol consumption (see, for example, Breslow et al., 2006, Breslow et al., 2013). Employment can be seen to exert its influence on energy density via another route: potentially encouraging the increased consumption of food away from home – found to be disproportionately energy-dense (Prentice and Jebb, 2003) – due to time constraints imposed by work commitments.

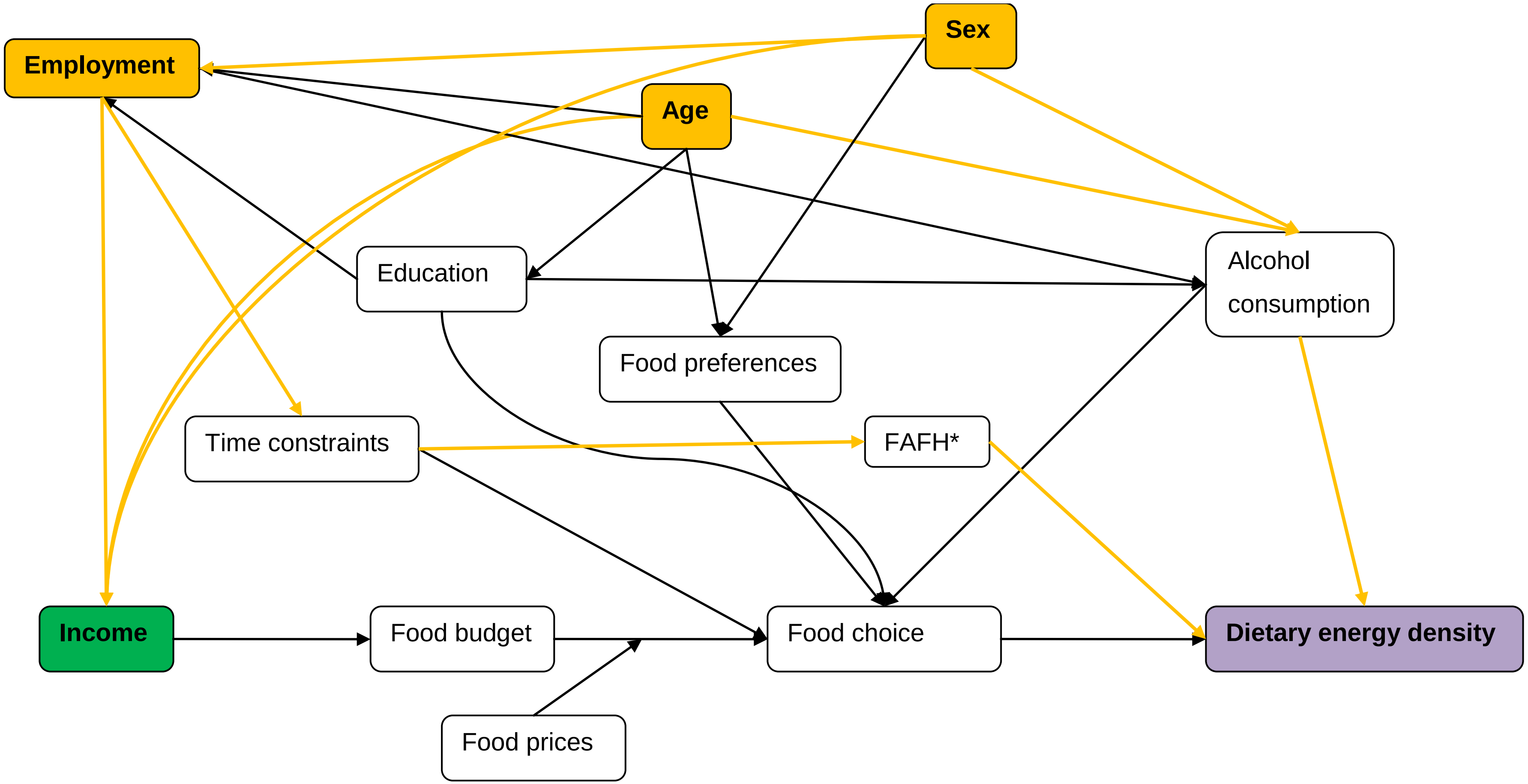
Commonly in dietary energy density research, energy intake is included as a covariate in regression analyses. However, in the current investigation, energy intake cannot properly be conceived of as a confounding variable: although it usually is correlated with dietary energy density, it would not be thought to influence income. For this reason, energy intake has not been included as a variable in the regression models, but in order to allow comparison to other research, sensitivity analyses will be performed in which energy intake is adjusted for.

Figure 4.4 illustrates the confounding influence of the same three variables on BMI. Age and sex are commonly recognised determinants of lean mass (Willett, 1998).

Employment may be less obviously linked to BMI, through its potential to influence daily physical activity (Proper and Hildebrandt, 2006). Of course, employment is not the sole determinant of physical activity, and it may therefore be preferable to adjust for physical activity itself. Unfortunately, this is not a variable available in the NDNS data. Smoking is another variable that is thought to be associated with BMI, and is often adjusted for in analyses investigating BMI as an outcome. However, its causal influence on income, and therefore its confounding influence, is less obvious. Rather than include smoking as a covariate in the main model, sensitivity analyses will be run in which smoking is also adjusted for.

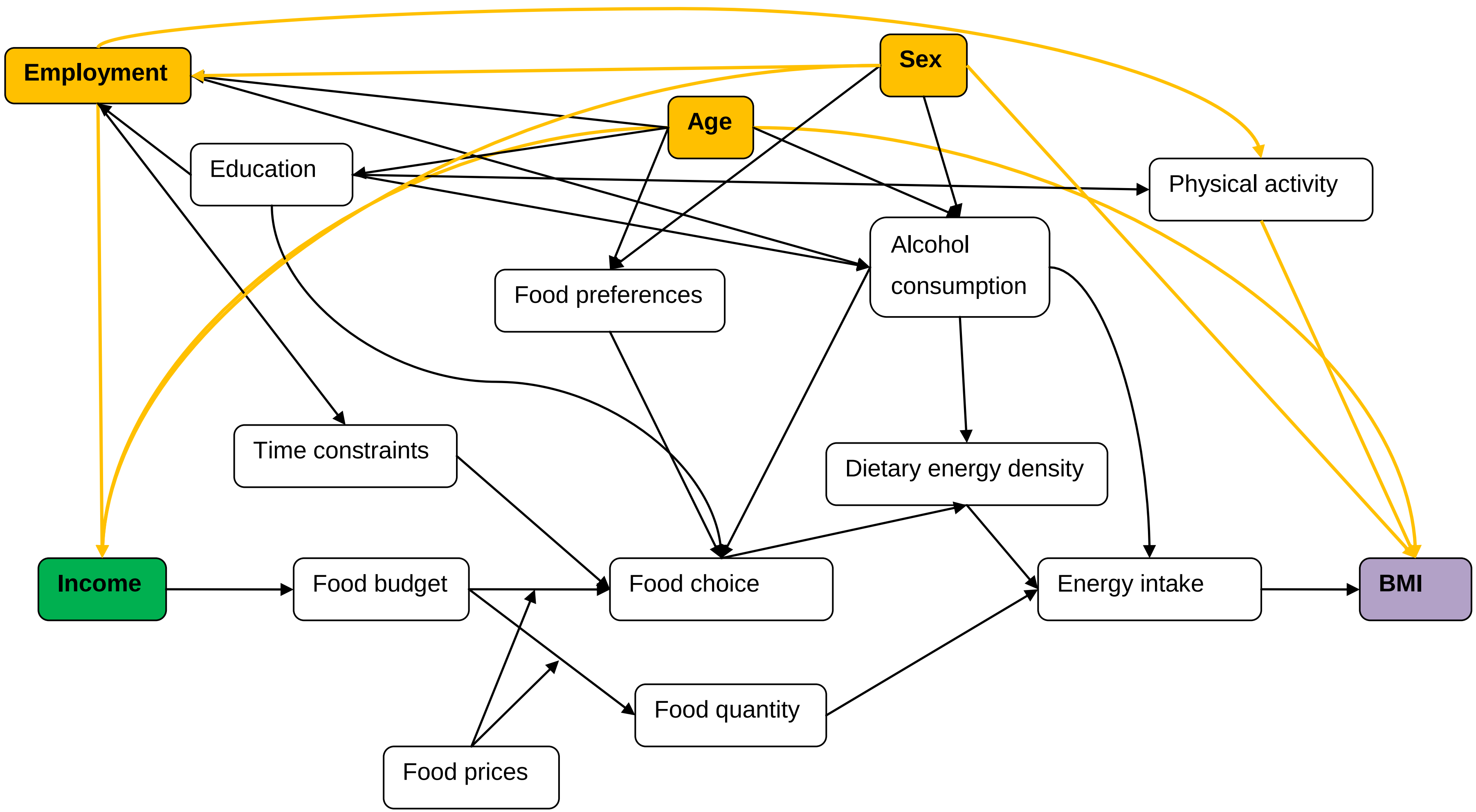
In addition to the confounding variables identified in the graphs, household size will be entered as a covariate in models 1b-6b. This is an approach often employed in nutrition epidemiology as an attempt to allow for differences in household composition. The results of this usual approach can then be compared to the results using an equivalized variable to determine which may be more useful in obesity research.

Figure 4.3 Directed Acyclic Graph (DAG) showing factors associated with income and dietary energy density



*Food away from home

Figure 4.4 Directed Acyclic Graph (DAG) showing factors associated with income and BMI



4.4 Results

4.4.1 Descriptive results

4.4.1.1 BMI

Of the participants with a valid BMI measurement, mean BMI was 27.5kg/m^2 (SD 5.3, n=938). The distribution of BMI values was positively skewed, however, with a median of 26.8kg/m^2 (IQR 23.8kg/m^2, 30.4kg/m^2). Twenty seven per cent (n=257) of the sample were obese, 65% (n=607) were either overweight or obese. The BMI of the sample is described in more detail in Chapter 3.

4.4.1.2 Dietary energy density

Mean dietary energy density (excluding non-milk beverages) was 6.38kJ/g, SD 1.42kJ/g (152kcal/100g, SD 34kcal/100g) (see Table 4.10).

Men had a higher mean dietary energy density than women (6.68kJ/g vs. 6.16kJ/g; p<0.01). Dietary energy density also differed significantly by age group, appearing to decrease with age (see Table 4.10). Other statistically significant differences were observed between categories of employment, household size and marital status, but not by qualification. Current regular smokers had a higher mean energy density than ex-smokers and non-smokers (6.94kJ/g compared to 6.17kJ/g and 6.25kJ/g respectively; p<0.01), and those who achieved ‘5 a day’ had a lower energy density, at 5.50kJ/g, than those who did not (6.82kJ/g; p<0.01).

Table 4.10 Mean dietary energy density* by sociodemographic groupings, kJ/g

Category (n)	Mean (kJ/g)	SD	p value (ANOVA)
Full sample (1014)	6.38	1.42	
Sex			
Female (580)	6.16	1.44	
Male (434)	6.68	1.34	
			<0.01
Age group			
19-29 years (145)	7.17	1.55	
30-39 years (202)	6.71	1.32	
40-49 years (179)	6.41	1.33	
50-59 years (184)	6.17	1.52	
60-69 years (147)	5.85	1.20	
70 years and over (157)	5.96	1.16	
			<0.01
Employment classification			
Managerial & professional (421)	6.30	1.44	
Intermediate, small employers, lower supervisory (302)	6.22	1.26	
Routine & semi-routine (250)	6.70	1.51	
Never worked & 'other' (41)	6.48	1.43	
			<0.01
Qualifications**			
Degree or higher education (338)	6.40	1.42	
GCE A-level or equivalent, foreign qualification (172)	6.43	1.35	
GCSEs/still in full-time education (245)	6.45	1.57	
No qualifications (251)	6.29	1.32	
			0.63
Household size			
1 person (268)	6.16	1.42	
2 people (336)	6.27	1.40	
3 or 4 people (327)	6.60	1.40	
5 or more people (83)	6.69	1.42	
			<0.01
Marital status			
Single, never married (289)	6.66	1.38	
Married and living with partner (467)	6.36	1.38	
Married but separated (30)	6.45	1.36	
Divorced (127)	6.23	1.52	
Widowed (101)	5.88	1.16	
			<0.01
Achieve '5 a Day'			
Yes (334)	5.50	1.05	
No (680)	6.82	1.38	
			<0.01
Cigarette-smoking status			
Never regularly smoked (541)	6.25	1.46	
Ex-regular smoker (247)	6.17	1.21	
Current regular smoker (226)	6.94	1.41	
			<0.01

*Energy density calculated from food and milk; **Missing qualifications data for n=8 participants

Mean dietary energy density within each category of BMI is presented in Table 4.11. An analysis of dietary energy density by normal weight, overweight and obese BMI categories did not reveal any statistically significant differences between these groups ($F(2, 922) = 1.44$; $p=0.24$).

Table 4.11 Mean dietary energy density in the NDNS sample for each BMI classification (n=1014)

BMI classification (n)	Mean (kJ/g)	SD	p value (ANOVA)
Not applicable (76)	6.00	1.61	
Underweight (13)	7.02	1.43	
Normal weight (318)	6.51	1.39	
Overweight (350)	6.35	1.45	
Obese (257)	6.38	1.52	
			0.24*

* tested only between normal weight, overweight and obese groups

4.4.1.3 Income

As described in Chapter 3, the most commonly reported crude annual household income category was between £15,000 and £24,999 (23%). Seventeen per cent reported an income below this, 16% reported £25,000 to £34,999, 13% £35,000 to £49,999 and 17% £50,000 or more. The remaining 14% (n=139) of the sample either did not know their annual household income or declined to answer.

Using equivalized income categories, participants were evenly split amongst the bottom three income categories, with 20% (n=198) having an equivalized income below £15,000, 20% (n=202) in the category of £15,000 to £24,999, 19% (n=197) in the £25,000 to £34,999 category, 14% (n=142) in the £35,000 to £49,999 category and 13% (n=136) in the highest income category.

Following equivalization, 42% of the sample (n=371) were reclassified into different income brackets (see Table 4.12): 163 moved up to a higher category and 208 moved down.

Table 4.12 Cross-tabulation of household income and equivalized household income: number in each category

Household income	Equivalized income: £14,999 or less	£15,000 to £24,999	£25,000 to £34,999	£35,000 to £49,999	£50,000 or more	Total
£14,999 or less	**143**	31	0	0	0	174
£15,000 to £24,999	52	**106**	79	0	0	237
£25,000 to £34,999	3	52	**73**	37	0	165
£35,000 to £49,999	0	13	39	**62**	16	130
£50,000 or more	0	0	6	43	**120**	169
Total	198	202	197	142	136	875

Participants who did not change income category following equivalization are shown in bold.

4.4.2 Income & energy density

Missing income data for 139 participants left an analytical sample of 875. Mean dietary energy density according to categories of income are presented in Table 4.13. Mean energy density differed significantly between categories of equivalized income (p=0.04) but not between crude household income categories (p=0.08). A linear trend by income category was not obvious.

Table 4.13 Mean energy density by income category, using reported household or equivalized income (n=875)

	Equivalized household income				Crude household income			
Category	n	Mean DED (kJ/g)	SD	p value (ANOVA)	n	Mean DED (kJ/g)	SD	p value (ANOVA)
£14,999 or less	198	6.62	1.55		174	6.47	1.47	
£15,000 to £24,999	202	6.46	1.29		237	6.34	1.45	
£25,000 to £34,999	197	6.18	1.37		165	6.30	1.32	
£35,000 to £49,999	142	6.45	1.55		130	6.72	1.53	
£50,000 or more	136	6.34	1.26		169	6.35	1.29	
				0.04				0.08

Household size-adjusted means are presented in Table 4.14. Differences in mean energy density between household income categories were not statistically significant when adjusting for household size.

Table 4.14 Mean energy density by reported household income category, adjusted for household size (n=875)

Category	n	Mean (kJ/g)	95% CI	p value (ANCOVA)
	Household income, adjusted for size			
£14,999 or less	174	6.33	6.09, 6.57	
£15,000 to £24,999	237	6.14	5.91, 6.37	
£25,000 to £34,999	165	6.05	5.77, 6.32	
£35,000 to £49,999	130	6.41	6.10, 6.72	
£50,000 or more	169	6.03	5.73, 6.32	
				0.06

In the multivariable regression models (adjusted for age, sex and occupation), equivalized income was not found to be associated with dietary energy density (see Table 4.15). On the other hand, crude household income was significantly associated with energy density, both with and without adjustment for household size (overall p=0.03 and p=0.04 respectively). As an example, the coefficient for the highest household income category indicates that this category was associated with a lower dietary energy density of 0.27kJ/g compared to the lowest category (adjusting for household size). Overall findings were similar when energy intake was included in the model in the sensitivity analyses.

Table 4.15 Multivariable linear regression of income categories (a separate model for each income variable definition) on dietary energy density (excluding non-milk beverages), adjusted for age, sex and occupation (n=875)

	Model 1a Crude household income		Model 1b Household income adjusted for size		Model 1c Equivalized income	
Category	Coefficient (95% CI)	p	Coefficient (95% CI)	p	Coefficient (95% CI)	p
£14,999 or less (reference)	-	-	-	-	-	-
£15,000 to £24,999	-0.07 (-0.33, 0.19)	0.61	-0.09 (-0.35, 0.18)	0.51	-0.13 (-0.40, 0.13)	0.33
£25,000 to £34,999	-0.32 (-0.61, -0.02)	0.04	-0.35 (-0.65, -0.05)	0.02	-0.33 (-0.60, -0.05)	0.02
£35,000 to £49,999	0.12 (-0.21, 0.45)	0.48	0.07 (-0.26, 0.41)	0.67	-0.18 (-0.49, 0.13)	0.25
£50,000 or more	-0.21 (-0.53, 0.10)	0.19	-0.27 (-0.59, 0.06)	0.10	-0.30 (-0.62, 0.02)	0.07
Overall		0.04		0.03		0.17

Table 4.16 Sensitivity analysis of income regressed on dietary energy density (excluding non-milk beverages), adjusted for age, sex, occupation and food energy (n=875)

Category	Model 1a Crude household income Coefficient (95% CI)	p	Model 1b Household income adjusted for size Coefficient (95% CI)	p	Model 1c Equivalized income Coefficient (95% CI)	p
£14,999 or less (reference)	-	-	-	-	-	-
£15,000 to £24,999	-0.14 (-0.40, 0.11)	0.29	-0.16 (-0.41, 0.10)	0.23	-0.12 (-0.38, 0.13)	0.34
£25,000 to £34,999	-0.37 (-0.65, -0.08)	0.01	-0.40 (-0.69, -0.11)	0.01	-0.31 (-0.57, -0.05)	0.02
£35,000 to £49,999	0.03 (-0.29, 0.34)	0.87	-0.02 (-0.34, 0.30)	0.91	-0.22 (-0.52, 0.07)	0.14
£50,000 or more	-0.26 (-0.56, 0.04)	0.09	-0.31 (-0.62, -0.01)	0.05	-0.30 (-0.61, 0.00)	0.05
Overall		0.04		0.02		0.15

When income was included in the multivariable linear regression models as a linear variable (and not a series of dummy variables, as above), in each case, there was no significant association. The adjusted R^2 for each model was similar.

Table 4.17 Linear regression of income on energy density, with income categories treated as continuous, adjusting for age, sex and occupation (n=875)

	Model 2a Crude household income	Model 2b Household income adjusted for size	Model 2c Equivalized income
Adjusted R^2	0.109	0.110	0.112
Coefficient	-0.03	-0.04	-0.07
(95% CI)	(-0.10, 0.04)	(-0.12, 0.03)	(-0.14, 0.00)
P value	0.42	0.25	0.06

Excluding those who reported an unusual amount of food consumed (n=514) did not alter the results in the models where income was entered as dummy variables. However, when these participants were excluded in the models in which income was entered as a linear variable, the coefficient for equivalized income achieved statistical significance (-0.14; 95% CI -0.25, -0.03; p=0.01). This suggests a lower energy density of 0.14kJ/g with each increasing equivalized income band.

Excluding participants who were following a special diet (n=91) resulted in no significant associations for any of the income variables treated as dummy variables. Where the income variables were treated linearly, however, and those on special diets were excluded, equivalized income was found to be significantly associated with energy density: with each progression up through the income bands, a lower energy density of 0.09kJ/g was predicted (95% CI -0.16, -0.01; p=0.02).

4.4.3 Income & BMI

The sample was further reduced to 825 due to missing BMI data (n=50). BMI differed between equivalized (p=0.04) but not crude household (p=0.08; not adjusted for household size) income categories (Table 4.18).

Table 4.18 Median BMI (kg/m^2) by income category, using reported household or equivalized income (n=825)

	Equivalized household income				Crude household income			
Category	n	Median BMI	IQR	P*	n	Median BMI	IQR	P*
£14,999 or less	190	26.9	23.5, 31.8		168	26.8	23.2, 31.8	
£15,000 to £24,999	190	27.5	23.9, 31.6		219	27.4	24.2, 31.0	
£25,000 to £34,999	184	27.1	24.3, 30.1		157	27.7	24.1, 31.2	
£35,000 to £49,999	134	26.5	24.0, 30.2		124	26.4	23.8, 30.1	
£50,000 or more	127	25.5	23.2, 28.9		157	26.0	23.3, 29.4	
				0.04				0.08

* p for Kruskal-Wallis ANOVA

Despite skewness of the outcome variable (see Section 3.6.3), residuals for the regression analysis were found to be normally distributed and the assumption of constant variance was also met. Regression of dummy income categories on BMI, adjusted for age, sex and occupation, revealed no significant association using either income variable (Table 4.19).

Table 4.19 Multivariable linear regression of income categories (a separate model for each income variable definition) on BMI, adjusted for age, sex and occupation (n=814)

Category	Model 3a Crude household income Coefficient (95% CI)	p	Model 3b Household income adjusted for size Coefficient (95% CI)	p	Model 3c Equivalized income Coefficient (95% CI)	p
£14,999 or less (reference)	-	-	-	-	-	-
£15,000 to £24,999	-0.30 (-1.39, 0.79)	0.59	-0.376 (-1.47, 0.72)	0.50	-0.13 (-1.23, 0.96)	0.81
£25,000 to £34,999	0.26 (-0.97, 1.48)	0.68	0.12 (-1.12, 1.35)	0.85	-0.72 (-1.85, 0.40)	0.21
£35,000 to £49,999	-0.47 (-1.82, 0.87)	0.49	-0.68 (-2.05, 0.69)	0.33	-0.74 (-2.01, 0.53)	0.25
£50,000 or more	-0.74 (-2.06, 0.57)	0.27	-0.99 (-2.33, 0.36)	0.15	-1.31 (-2.66, 0.03)	0.06
Overall		0.55		0.42		0.31

Underweight (n=11) excluded

The inclusion of smoking status in the models, in the sensitivity analysis, altered coefficients slightly, but resulted in similar overall p values for the income variables.

Table 4.20 Multivariable linear regression of income categories (a model for each income variable definition) on BMI, adjusted for age, sex, occupation and smoking (n=814)

Category	Model 3a* Crude household income Coefficient (95% CI)	p	Model 3b* Household income adjusted for size Coefficient (95% CI)	p	Model 3c* Equivalized income Coefficient (95% CI)	p
£14,999 or less (reference)	-	-	-	-	-	-
£15,000 to £24,999	-0.37 (-1.46, 0.73)	0.51	-0.44 (-1.53, 0.66)	0.44	-0.30 (-1.39, 0.80)	0.60
£25,000 to £34,999	0.13 (-1.09, 1.35)	0.84	-0.00 (-1.24, 1.23)	>0.99	-0.78 (-1.91, 0.34)	0.17
£35,000 to £49,999	-0.56 (-1.90, 0.79)	0.42	-0.75 (-2.12, 0.62)	0.28	-0.89 (-2.16, 0.38)	0.17
£50,000 or more	-0.87 (-2.19, 0.45)	0.20	-1.10 (-2.45, 0.25)	0.11	-1.46 (-2.81, -0.12)	0.03
Overall		0.53		0.40		0.25

Underweight (n=11) excluded

*Sensitivity analysis: also adjusting for smoking status.

However, treating the income variables as linear in the models (Table 4.19) revealed a significant negative relationship with equivalized income: each increasing equivalized income category was associated with 0.33kg/m^2 lower BMI. No significant relationship was evident for household income, whether crude or adjusted for household size.

Table 4.21 Linear regression of income on BMI, with income categories treated as continuous, adjusting for age, sex and occupation (n=814)

	Model 4a Crude household income	Model 4b Household income adjusted for size	Model 4c Equivalized income
Adjusted R^2	0.0028	0.0058	0.0078
Coefficient	-0.17	-0.23	-0.33
(95% CI)	(-0.47, 0.14)	(-0.54, 0.09)	(-0.63, -0.02)
P value	0.29	0.16	0.04

Underweight (n=11) excluded

4.4.4 Income & overweight + obesity

The proportion of adults classified as overweight or obese differed according to equivalized but not crude household income in the univariate analyses (see Table 4.22). The lowest proportion was amongst those in the highest equivalized income category (53%).

Table 4.22 Proportion of adults classified as overweight or obese in each income category (n=814)

	Equivalized household income			Crude household income		
Category	n	% overweight + obese	p value*	n	% overweight + obese	p value*
£14,999 or less	183	67		162	65	
£15,000 to £24,999	188	68		215	70	
£25,000 to £34,999	183	70		157	69	
£35,000 to £49,999	133	63		123	63	
£50,000 or more	127	53		157	57	
			0.02			0.09

* chi^2 comparison; underweight (n=11) excluded

Using dummy income variables, logistic regression found no overall association between income and the odds of being classified as overweight or obese: this was true whichever definition of income was used (see Table 4.23). However, the odds of being overweight or obese were significantly lower in the highest equivalized category compared to the lowest (OR 0.54, 95% CI 0.31, 0.92, p=0.03), indicating 46% lower odds. Odds ratios did not differ by household income category, whether or not the model was adjusted for household size.

Table 4.23 Results of logistic regression models investigating income and the odds of being overweight or obese (n=814)

Category	Model 5a Crude household income OR (95% CI)	p	Model 5b Household income adjusted for size OR (95% CI)	p	Model 5c Equivalized income OR (95% CI)	p
£14,999 or less (reference)	1.00	-	1.00	-	1.00	-
£15,000 to £24,999	1.20 (0.76, 1.89)	0.44	1.15 (0.73, 1.82)	0.55	0.94 (0.60, 1.49)	0.80
£25,000 to £34,999	1.20 (0.72, 2.00)	0.48	1.11 (0.67, 1.87)	0.68	0.96 (0.60, 1.55)	0.88
£35,000 to £49,999	1.01 (0.58, 1.74)	0.98	0.90 (0.51, 1.57)	0.70	0.78 (0.46, 1.31)	0.35
£50,000 or more	0.83 (0.49, 1.41)	0.49	0.73 (0.42, 1.25)	0.25	0.54 (0.31, 0.92)	0.03
Overall		0.54		0.39		0.15

Underweight (n=11) excluded

Treating the income variables as linear (Table 4.24) revealed significantly lower odds of being overweight or obese with increasing equivalized household income category. Odds ratios did not achieve statistical significance for crude household income nor household income adjusted for household size.

Table 4.24 Results of logistic regression models investigating income and the odds of being overweight or obese (n=814): income entered as linear variables

	Model 6a Crude household income	Model 6b Household income adjusted for size	Model 6c Equivalized income
Adjusted R^2	0.0414	0.0458	0.0450
OR	0.94	0.91	0.87
(95% CI)	(0.83, 1.07)	(0.80, 1.04)	(0.77, 0.99)
P value	0.34	0.15	0.03

4.4.5 Equivalized income as a continuous variable

The median equivalized income of the sample, prior to categorisation into income bands, was £26,119 (IQR £16,917 to £41,045). Median equivalized income by BMI categorisation is shown in Table 4.25. Equivalized income was found to significantly differ by weight category (chi^2 statistic = 7.525, p=0.02), with the lowest median equivalized income, £23,326pa, found amongst the obese.

Table 4.25 Median equivalized income by normal weight, overweight and obese categories (n=814)

	Normal weight (n=284)	Overweight (n=300)	Obese (n=230)
Median equivalized income, £pa (IQR)	£26,453 (£17,155 to £47,500)	£26,119 (£17,500 to £39,583)	£23,326 (£16,071 to £35, 714)

4.5 Discussion

This chapter presents a description of incomes in the NDNS, with an aim to explore whether the way in which the income variable is defined has any bearing on whether it is found to be associated with diet or BMI. The results suggest that accounting for household composition (through equivalization) will affect results and therefore our interpretation of the income-diet and income-BMI relationships. The three outcomes investigated – dietary energy density, BMI and overweight and obesity – were all found to be associated with equivalized income. This implicates monetary considerations in the diet selection and consequent weight status of British adults.

4.5.1 Dietary energy density in the NDNS

The mean dietary energy density of the sample, at 6.38kJ/g, is greater than the WCRF recommended goal of 5.23kJ/g (stated in kcal, 125kcal/g) (WCRF, 2007). However, the estimates from this NDNS sample were not as high as recent estimates of energy density from food and milk in Scotland – 7.23kJ/g (Wrieden and Barton, 2011). Other population means using this method have ranged between 5.69kJ/g and 7.61kJ/g (see Ledikwe et al. (2005) for a review).

The findings here are in agreement with previous studies (Ledikwe et al., 2005, Marti-Henneberg et al., 1999) in finding statistically significant differences in dietary energy density between males and females, and by age group. In addition, differences in energy density were found between categories of several other key variables in this sample, including employment and household size. These latter variables, of course, could be related to energy density through their relationship with income – the primary variable under investigation in this chapter. It can be difficult to tease apart the relative influence of these closely related variables. Other studies (Wrieden and Barton, 2011) have found differences in dietary energy density according to aggregate measures of socioeconomic status. The results of this chapter support the literature in indicating that nutritional inequalities exist in the UK (James et al., 1997).

The one sociodemographic variable for which differences in energy density were not apparent was qualification. This is perhaps surprising, given the presumed close relationship between qualifications, occupation and income. This finding also contrasts with observations in other samples (Monsivais and Drewnowski, 2009, Kant and Graubard, 2013) which indicated decreasing energy density with increasing levels of education. However, it may be that the method of classification used in this study, and the categories specified, were inadequate for detecting differences. For example, some of the original survey qualification categories had to be collapsed due to small numbers of participants. It is possible that dichotomising the sample into those with no qualifications and those with any qualifications and comparing the two groups may have shown significant differences in energy density – the means presented in Table 4.10 indicate a lower DED amongst those with no qualifications than the other qualification categories, although this was not formally tested.

It is possible that much of the disparity in dietary energy density is due to differences in the consumption of fruit and vegetables. Fruit and vegetables tend to have a higher water content than many other types of food, and are probably therefore an important influential factor in the energy density of the diet. In support of this, dietary energy density was found to significantly differ between those who did and those who did not achieve their '5 a day' in this sample. Differences in achievement of '5 a day' are reported in Chapter 3: many of these differences lay between the same categories as did differences in energy density – for example by age and by occupation; however differences were not evident between males and females or by household size, in contrast to the results for energy density comparisons. This suggests that fruit and vegetable intake is only part of the explanation for energy density variation.

Interestingly, dietary energy density was found to differ according to cigarette smoking status. This is potentially indicative of behaviour clustering – in other words, those who engage in smoking may be more likely to also engage in other harmful behaviours, such as consuming a poor (energy-dense) diet (Schuit et al., 2002, Poortinga, 2007). An alternative explanation is that if, as the results here suggest, income is related to dietary energy density, cigarette smoking may influence dietary selection through its impact on the available budget for food. However, it is not possible from the NDNS data to determine which of these hypotheses is most likely.

4.5.2 Income and energy density

The univariate comparisons revealed significant differences in dietary energy density between equivalized income bands, but not between household income categories. There is a clear implication here that equivalizing has created a more appropriate income variable, allowing differences to be detected. The results of the multivariate models, however, are more complicated to interpret.

Treating the income variable as four dummy variables in the model (with the 'i.' prefix) results in a better model fit when household income is used (either adjusted for household size or not), but not when equivalized income is the predictor. On examining the coefficients for each category, it appears that the household income category of £25,000-£34,999 is the only category for which confidence intervals do not span zero.

This could lead to the conclusion that it is this category that differs most from the baseline comparator (the lowest income category) – perhaps indicative of a non-linear relationship. A non-linear relationship to income is credible, given previous reports of a non-linear relationship between income and health (Benzeval et al., 2001).

Regardless of the shape of the relationship, the initial conclusion from the multivariate regression analyses might be that household income is sufficient, or even more appropriate than the equivalized variable, for investigating income-diet hypotheses. However, the sensitivity analyses invite a different interpretation.

Excluding participants who were following a special diet in Models 1a, 1b and 1c (in which income categories were entered as dummy variables using the 'i.' prefix) resulted in no overall significant p values for income – whereas including these participants suggested a significant association for household but not equivalized income. When special dieters were excluded from Models 2a, 2b and 2c (where the 'i.' prefix was not used), a significant association with dietary energy density was evident for equivalized but not household income. Furthermore, the significant association between equivalized income and energy density similarly emerged when Models 2a, 2b and 2c were run excluding those who had reported an unusual amount of food consumed. This is perhaps even more interesting, because the number of participants excluded in this latter sensitivity analyses constituted more than half of the analytical sample. A reduction in sample size of this degree would be expected to decrease the power of the regression analyses substantially, therefore making a statistically significant result less, not more, likely.

The results of both sensitivity analyses – especially given the decreased power of the sample sizes – suggest that there is indeed a relationship between income and dietary energy density, but the consumption of atypical diet during the diary data collection period is masking this relationship when the full sample is analysed.

4.5.3 Income and BMI

Treating the income variables as ordinal in the models, rather than entering them as dummy variables, resulted in a better model fit for both the linear and logistic regression analyses. The findings indicated that equivalized income was the more useful income variable in testing the income-BMI relationship. Every higher equivalized income category was associated with $0.33kg/m^2$ less in BMI, or 13% lower odds of being classified as overweight or obese.

These findings support the HSE observations (NOO, 2010; see Section 1.6) amongst UK women of an inverse linear relationship between income and the prevalence of obesity. The analyses here go further in adjusting for confounding variables and formally testing the relationship, to find evidence of this association in both men and women. The results support international data in showing lower odds of being obese amongst those with higher incomes (Schumann et al., 2011, Ali et al., 2011).

Previously, authors have commented on the apparent non-linearity of the association between income and BMI – even describing it as an expected shape given that there is both the demand for food and the demand for an ideal body weight which may compete with or offset each other (Lakdawalla and Philipson, 2009). These analyses, in contrast, suggest that there is a linear relationship amongst British adults. One possible explanation for this is that the underweight were excluded from the regression analyses. As described in Section 4.3.4.3, this exclusion was made because it was felt that the mechanisms underlying negative energy balance would be different to those underlying positive energy balance and should therefore be considered separately. Considering only positive energy balance may have enabled the linear association to emerge.

Furthermore, analyses using the continuous variable of equivalized income (before categorization) indicated that the categories of normal weight, overweight and obese were found to significantly differ in their median equivalized income estimates, with the lowest mean, £23,326, amongst the obese. The median equivalized income estimates of the normal weight and overweight participants, at around £26,000, were more in line with the national UK median salary (see below).

Taken together, the results of this chapter indicate that income is significantly negatively associated with both dietary energy density and with overweight and obesity. If the hypothesis is correct – that increasing income allows for the purchase of more expensive, less energy-dense diets and therefore a decreased likelihood of weight gain – then a relationship between energy density and BMI would be anticipated. However, energy density was not found to significantly differ by BMI in this sample. This is in contrast to other findings in the literature – such as those of Cox and Mela (2000) – and perhaps reflects the fact that the analysis in this chapter was not adjusted for other variables, or that sensitivity analyses were not performed because this relationship was not the primary purpose of this chapter.

4.5.4 Equivalizing income

The median equivalized income was estimated at £26,100. It is not possible to compare this with the median income before equivalization, because household income is a categorical variable in the NDNS. However, £26,100 is in line with the national median UK salary of the 2008-09 tax year, £25,800 (ONS, 2009). This is unsurprising given the design of the NDNS as a nationally representative survey, and lends credence to the appropriateness of the sample for these analyses.

The discrepancy in proportions across income bands between equivalized and the crude income variable indicates the extent to which relying upon a non-equivalized income variable could misclassify participants. Table 4.12 shows that 43% of the sample may have been misclassified if household size and composition were not taken into account. Furthermore, this misclassification occurred in both directions – in other words, participants could have been reclassified into either higher or lower income bands. This would have important ramifications for analyses involving income by affecting participants' ranking, which could obscure relationships, especially those that are linear.

The results of the analyses – the regression analyses in particular – demonstrate the impact that misclassification can have on interpretations: using the equivalized income variable revealed a significant association between income and BMI, and between income and the odds of being overweight or obese, where crude household income did not. The advantage of equivalized income was also displayed in the income-energy density investigations, although the significant association only became apparent on excluding certain participants in the sensitivity analyses.

The issues surrounding income measurement are more complex and numerous than the simple adjustment for household composition implies: there are numerous arguments documented around the best method for gauging income – whether to use wage only, adjusting for tax benefits, accounting for indicators of wealth and so on. These concerns are too numerous to cover in detail in this thesis. Nevertheless, the findings of this chapter suggest that a simple adjustment of already collected household income data can be useful, and will enhance comparability across different household sizes and compositions.

There have been few studies published which have set out to examine both equivalized and non-equivalized household income variables in relation to health, and none in relation to diet, as far as the author is aware. Benzeval et al (2001) compared odds ratios for self-reported health and limiting or long-standing illness, between quintiles of family income, net individual income or equivalized family income. Their results indicated that the equivalized income variable gave the best statistical fit. The findings of this chapter support the conclusions of Benzeval et al that equivalizing income is the most appropriate method, and extends this conclusion to investigations involving BMI or diet.

4.5.5 Limitations

The conclusions of this chapter are limited insofar as the NDNS provides only cross-sectional data. The broad hypothesis underlying the rationale for these analyses – relating the effect of income onto diet selection and subsequent weight – cannot be tested with cross-sectional data, which is inappropriate for statements about causality.

Measuring income accurately is not straightforward. Most studies, such as the NDNS, must rely upon self-reported income. Participants may mis-estimate total household income, or they may purposefully over- or under-estimate due to social desirability pressures (Hebert et al., 1995). It is not uncommon in such surveys for a large number to not report income at all. In this survey, income data were missing for 8% of the sample (n=139). This may have created a form of self-selection bias in which the analytical sample consisted of only those participants who were able or willing to divulge income information, and should be borne in mind when interpreting findings.

In addition, it has previously been noted that income at a single time-point may not give an accurate representation of economic status, as factors such as prior income, savings, income shocks or other life course-specific situations may be influential. For this reason, other authors (Benzeval et al., 2001, for example) have recommended a life course approach to quantifying income, using longitudinal data to gain a clearer picture. No such data are available with the NDNS. Nevertheless, the findings of this chapter suggest that a cross-sectional measurement of income, adjusted for household composition, can be enough to reveal interesting patterns.

The analyses involving dietary energy density in this chapter rely upon self-reported data. The drawbacks of this – particularly in terms of mis-reporting – are discussed in Chapter 3. If participants in the NDNS have not accurately reported the types of foods consumed as well as the quantity consumed, the estimates for dietary energy density could be biased as a result. Without physical activity data, it is difficult to determine if under- or over-reporting of energy has taken place (see Chapter 3), and as such the potential for biased energy density estimates is hard to assess. Having said that, the mean energy density estimates for this sample are in line with those reported in other studies (see Section 4.5.1), which suggests that these estimates are plausible.

Whilst energy density may be considered as one indicator of diet quality, it remains a crude measure of quality. As noted earlier in this chapter (Section 4.3.3), estimates of energy density depend upon the method chosen to calculate them – including or excluding beverages, for example, can make a considerable difference to the estimate. Ideally, more refined measures of quality – such as dietary pattern analysis – would be informative in assessing dietary differences between income groups. However, this was beyond the scope of this thesis, and would have to be the subject of future research. Although crude, energy density provides an easily constructed variable that could differentiate types of dietary choice at a basic level. Furthermore, it has relevance to the rationale for this chapter – that lower food budgets may encourage the selection of energy-dense foods in order to maximise the energy obtained for a given amount of money.

4.5.6 Strengths

This study benefits from using data from a survey designed to be nationally representative. This could give insight into how income may be related to BMI or dietary energy density in the UK. A further strength is the use of professionally measured anthropometry, as opposed to self-reported height and weight, which will have helped to minimise the bias associated with the BMI variables.

Furthermore, the analysis is one of only a few to directly compare results using different income variables, and the first to do so in relation to dietary data, as far as the author is aware. Here, the advantages of using an equivalization index to correct household incomes are plainly demonstrated in the context of dietary research.

4.6 Conclusions

This chapter introduced an important demand-side factor in the micro-economics of diet selection – income. Whilst other demand-side factors are also presumed to be involved in the processes of food purchasing, income is a defining variable in the affordability of diet, and data on income were available in the NDNS data set, making it an ideal focus of study.

The results of the sensitivity analyses presented above agree with some of the literature in finding a negative relationship between income and energy density – those on the lowest incomes reported the most energy-dense diets. This agrees with the theory underlying the rationale for these analyses, implying that those on lower incomes could be motivated to consume energy-dense foods which provide more energy per serving.

The analyses also indicated a negative relationship between income and BMI. This has not always been evident in UK statistics. However, this is the first time that a formal analysis of BMI and overweight/obesity prevalence has been performed in a representative UK sample using an appropriately equivalized income variable.

The inclusion of models employing either crude household income or equivalized income clearly illustrates the necessity of accounting for household composition in an appropriate fashion: examination of the results demonstrates how the conclusions of the analyses would have been very different had equivalization not been performed. In most cases, crude household income, even when adjusted for household size in the model, was not associated with the outcomes tested in this chapter. Equivalized income is a variable seldom employed in nutritional epidemiology, and the findings of this chapter highlight the potential detriment of this oversight.

The fact that income was related to both dietary energy density and BMI in this chapter could imply that these three variables are causally related – the hypothesis being that restricted income encourages the consumption of energy-dense diets leading to a propensity for excess energy intake and thus higher BMI. The caveat with the present analysis is that causality cannot be determined from cross-sectional data. A further caution in interpreting the results in this way is that BMI and energy density were not significantly associated in this sample.

Nevertheless, being able to link income to diet and diet-related health will have important repercussions for public health. Attempts to intervene in diet or BMI may be hindered by neglecting to take into account underlying socioeconomic influences. As it was summed up by the 'Marmot review': "Having insufficient money to lead a healthy life is a highly significant cause of health inequalities." (Marmot and Bell, 2012).

What was known previously:

- Socioeconomic disparities in diet and health are present in the UK.
- Energy-dense diets are linked to higher energy intakes.
- Income is a defining factor in the affordability of diets.
- There is conflicting evidence of an association between income and dietary energy density.
- There is limited evidence of a linear relationship between income and BMI and between income and overweight or obesity prevalence in adults.
- Household income should be framed with reference to household composition, but is seldom equivalized in nutrition epidemiology.

What this chapter adds:

- Income is negatively and linearly associated with dietary energy density in the NDNS.
- Income is negatively and linearly associated with BMI amongst NDNS adults.
- The odds of being overweight or obese are significantly lower with increasing income bands.
- Obese adults in the NDNS have a lower median equivalized income than those who are normal or overweight.
- The use of crude household income can result in different findings and interpretations compared to when equivalized household income is used.

Chapter 5 The DANTE food cost database

5.1 Summary

One of the primary aims of this thesis is to describe and examine the monetary costs of adults' diets in the National Diet and Nutrition Survey (NDNS). Direct data regarding the food prices encountered by NDNS participants, however, is not available. Therefore, a means of estimating the monetary cost of diets is necessary. This chapter will introduce the tool that will be used to estimate NDNS diet costs, the DANTE (Diet and Nutrition Tool for Evaluation) food cost database, and describe the methods employed in its construction.

The food cost database houses information on national food prices and is integrated within the in-house nutritional database, DANTE, which is used to store and analyse the nutritional aspects of dietary data. This enables a price to be applied to the quantity of each food reported in a diet diary, FFQ or other assessment tool, alongside traditional nutrient analyses. From this it can be estimated how much an individual's diet may have cost had they purchased their food at average prices.

Given the degree of inference associated with this approach, its validity may be questioned. Unfortunately, however, there is no gold standard against which to validate the DANTE cost database. Instead, this chapter presents results from comparability studies using data from two previously conducted studies, in which diet costs estimated by the DANTE cost database were compared to calculations from household till receipts.

Testing for agreement using Bland Altman plots, the comparability studies revealed mean differences between the methods as low as £0.02, with 95% limits of agreement between £3.22 and -£3.08. This suggests that the DANTE cost database is useful in estimating diet costs of larger samples. At the individual level, however, the differences in estimates between the methods are potentially substantial, as indicated by the wide limits of agreement.

Understanding how methods differ in their estimates of diet cost is important for interpreting the results of diet cost research – such as those presented in the subsequent chapters of this thesis.

5.1.1 Acknowledgements

This chapter builds upon the work of previous research at the University of Leeds. I would like to acknowledge in particular the efforts of: Kevin Tarbutt (funded via a Rank Prize Fund) and Edmund Parks who created and updated the DANTE cost database; Claire Oyston and Joan Ransley who created the SNIP till receipt dataset; and Andrea Smyth for conducting the UKWCS subsample study.

The current analysis of the comparability studies was undertaken in collaboration with another PhD student within the University of Leeds' Nutritional Epidemiology Group, Michelle Morris. This work also forms the basis of a publication in the European Journal of Clinical Nutrition (Timmins et al., 2013b), again produced in collaboration with Michelle Morris.

5.2 Introduction

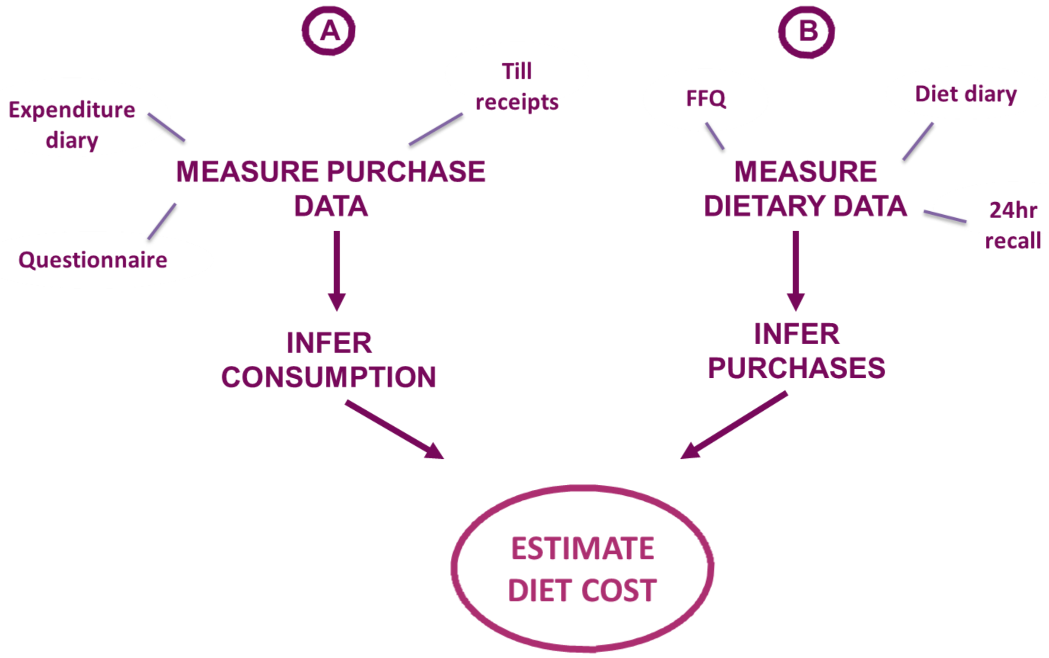
A key supply-side determinant of food purchasing is the price of food (Chapter 1). It has often been described as an important contributor to the obesogenic environment (Drewnowski and Darmon, 2005, Monsivais et al., 2010, Chaloupka and Powell, 2009), and behavioural studies have repeatedly found that manipulation of food prices affects both purchasing behaviour (for example, French, 2003, Ni Mhurchu et al., 2009), as well as being consistently reported by participants as key influences on their purchasing and consumption decisions (Steptoe et al., 1995, Nelson et al., 2007). Experimental observations, however, lack external validity, and findings may not be applicable outside the experimental setting. Identifying the actual effect of food prices on diet and health in a real-world setting is necessary, but challenging.

Chapter 2 synthesised the methodological approaches employed in the field. At a population level, there are a number of methods that have been used to measure food prices, which can then be matched to sales data, or population-level data on diet and health. However, there is a need to measure cost at the individual level, in order to link food prices to health outcomes (Murakami et al., 2008a). Yet calculating the financial cost of a person's dietary intake is far from straightforward. Measurement of diet cost is made difficult by the fact that people do not purchase foods in the exact quantities that they eat. Nor do they necessarily purchase at the time of consumption. Factors which make diet cost assessment problematic include: free food, shared food, foraged or homegrown food, food away from home (FAFH), food waste, promotional discounts, bulk buying, food from the storecupboard or freezer, seasonal fluctuations in prices, and variation in prices according to retail outlet. It is therefore necessary to estimate, rather than measure, diet costs.

Methods for estimating diet costs can be broadly categorised into two procedures (Figure 5.1): firstly, purchase data can be measured, from which dietary consumption is inferred, or, vice versa, costs can be inferred from dietary assessment data.

Methods falling into category A attempt to measure individuals' purchases of food and drinks. This can be achieved through: an expenditure diary, in which participants record their purchases for a set period; the collection of till receipts for all household purchases during a given period; or using an expenditure questionnaire with a single time-point of administration. Once food and drink expenditure has been calculated from these data, assumptions are made about how much of the purchased goods were consumed by the individual. These assumptions may attempt to take into account household composition as well as anticipated food waste.

Figure 5.1 Methodology routes for estimating diet costs: A) from purchase data or B) from dietary data



Methods in category B, on the other hand, have the advantage of using best available methods for dietary assessment. Using dietary data, it is then possible to apply prices to the foods consumed – commonly these are housed in a database of national prices. The key assumptions of this method are: firstly, that foods consumed are priced around the national average value; and secondly, that participants have purchased all the foods consumed. Neither can be said to be true in every case, and therefore the estimated costs represent the inherent monetary value of the diet, rather than actual expenditure.

Each of the cost estimation methods described above has its advantages and disadvantages. A summary of the methods, with a brief appraisal of their strengths and weaknesses, is outlined in Table 5.1. None of these methods are able to account for all of the factors proposed to influence actual expenditure, as described above.

The analyses in Chapters 6, 7 and 8 examine diet costs of a nationally representative adult sample, the NDNS. As the NDNS is a dietary survey, with no information on expenditure, the diary data will be matched to a database of food prices, as described in route B of Figure 5.1. The food prices to be used are held within an in-house database, referred to as the ‘DANTE food cost database’. Due to the element of approximation inherent in this costing method, it was felt important to gauge how the diet costs estimated using the DANTE food prices and dietary intake compare to other methods of assigning costs to diets. This chapter introduces the DANTE food cost database, and details two comparability studies using data from prior research projects at the University of Leeds, which had collected costs of diet from measured purchases. A re-analysis of these data will be presented.

In summary, the purposes of this chapter are to describe:

1. The food cost database used;
2. The method of linking this database to dietary data; and
3. How estimates using this method compare to alternative methods.

Table 5.1 Methods used in the literature for estimating individual-level diet costs

Method	Application	Example	Advantages	Disadvantages
Till receipt collection, expenditure diary	Commonly used to estimate national Consumer Price Index Also used to estimate food or nutrient availability	The Living Costs and Food Survey (Defra, 2009)	Suitable for large population samples; easy to administer	Burdensome for respondents Limited to period of diary/receipt keeping Does not assess dietary consumption – if used, consumption is estimated from expenditure, with a correction factor for waste Cannot account for storecupboard patterns; free food; or shared food
Retrospective expenditure questionnaire	To gather reported habitual food expenditure or budgets	Turrell & Kavanagh (2005)	Single time point of administration Low burden on participants	Retrospective, therefore a probability of recall bias Reliance on self-report data Estimates usually ask for aggregate food level, so information may be lacking for specific food items
Estimation using published price databases	To estimate dietary expenditure where dietary information but not expenditure information is available	Ryden & Hagfors (2011)	Can be applied to typical dietary surveys in the absence of expenditure data	Actual expenditure is not measured Sources of price information may differ to chief sources of groceries amongst the population Estimates of expenditure rely on averaged price data National-level price data may not be matched at the regional level Consumption must be back-transformed to purchase quantities to calculate prices after adjusting for waste or water retention/loss Cannot account for variations in expenditure caused by, for example, homegrown or free food, promotions, or FAFH

5.3 The DANTE food cost database

5.3.1 Introduction

National food price data do exist in the UK (Defra, 2009). However, household food items are coded into just 250 aggregated groups (and 250 further categories for eating out purchases). Therefore, in 2004, a more detailed catalogue of UK food prices was compiled at the University of Leeds. This database incorporated low, medium and high prices for over 3,000 food and drink items. In addition, the prices were integrated into the Diet and Nutrition Tool for Evaluation (DANTE), dietary analysis software which utilises nutrient information from McCance and Widdowson's Composition of Foods (Holland et al., 1991). The database, which will be referred to as the 'DANTE food cost database', offers a unique tool for estimating diet costs alongside traditional dietary intake data collection and nutrient analysis.

5.3.2 Population of the DANTE cost database

Price information was collected from a variety of sources, but chiefly Tesco online (www.tesco.com). Price information for items not available from this source, such as niche products, was located from other outlets' websites - including Sainsbury's (www.sainsburys.co.uk) – or specialist stores. The lowest, highest, and mean prices in pence were calculated per edible 100g for each item (or 100ml where appropriate). Where weight information was unavailable (for example, for fruit pie or cake slices), 100g was estimated from standard food portion sizes (MAFF, 1994). The price for 100g was mapped onto each DANTE food item code. On occasion, no price data were available for an item; in such instances (n=398), the price was based on an appropriate equivalent, judged on product type and nutritional content. Promotional offers affecting unit price were disregarded as anomalous data.

Following the initial data collection, food price information was found to be missing for 346 items. These were added in May-June 2008, in the same manner. To allow for inflation, the consumer price index (CPI) was used to adjust the prices in line with those collected in 2004. After this expansion, the database numbered 3,192 items.

The food cost database was populated in 2004 by a placement medical student, and expanded in 2008 by another postgraduate student.

5.3.3 Using the DANTE cost database to assign costs to diets

The prices per 100g (or 100ml) are housed in the database as additional vectors: in other words, they are listed for each food as are nutrients. From this, it is possible to multiply the cost (high, mean or low) by the quantity consumed to estimate the cost of the food eaten. Figure 5.2 shows an example of the nutrient analysis output, showing a calculated cost (in pence) for the quantity of each food consumed. The costs of all food and drinks consumed by a participant can then be summed to provide a total cost, and divided by the number of days to give a daily diet cost estimate.

Figure 5.2 Snapshot of the DANTE cost database food item estimates

FoodNumbe	FoodName	MainFoodGı	Total_Grams	Calculatedcost
10040	FAT SPREAD (62-72% FAT) NOT POLYUNSATURATED	21	1[illegible]	3.708
608	MILK SEMI-SKIMMED PASTEURISED SUMMER	11	106.[illegible]	5.4735
126	BREAD WHITE TOASTED	2	65.8[illegible]	5.02335
10072	PRAWN COCKTAIL SNACKS, MAIZE / RICE FLOUR CORN SNACKS,	42	1[illegible]	13.94
2251	BOILED SWEETS BARLEY SUGAR BUTTERSCOTCH GLACIER MINTS	43	42	20.7[illegible]

5.3.4 Strengths & limitations

The DANTE cost database boasts a key advantage over household expenditure data: using dietary assessment methods provides data at the individual level which is important when investigating the economic determinants of diet and health. It is important to clarify that the cost estimates given by the database, however, reflect the estimated inherent value of the diet, rather than being a measurement of expenditure. The value of a person’s diet may not reflect the prices they encountered in purchasing their food.

The creation of the database relied heavily upon a single source, the Tesco website. This means that the price ranges collected may not reflect that found nationally. Furthermore, because the database creation was carried out historically, it is unclear if there was a protocol for systematically selecting alternative sources where items were not listed on the Tesco website, nor is it documented for which items this was necessary. It is also difficult retrospectively to assess whether the indices used to adjust for inflation in the expansion of the database were adequate.

The database houses three levels of cost for many of the foods it contains. This provides options for the researcher, but in reality it may be difficult to gauge which level of pricing is most appropriate for each participant or sample. Geographical variations, as well as retailer availability and access (Wrigley et al., 2002, Jiao et al., 2012), could affect costs encountered. Using mean prices could result in an overestimation of diet costs for groups which consistently purchase foods at lower than average prices (or vice versa). In addition, promotions or price discounts, due to their transient nature, could not be incorporated into the DANTE cost database, and therefore, where these are used by individuals to stretch their budget (Beatty, 2010), this cannot be taken into account.

This estimation method will not be able to account for food purchased and eaten away from home – restaurant or takeaway meals, for example. Foods and drinks consumed outside the home are likely to be higher in cost than would be estimated by the DANTE cost database. Free, shared, or foraged food will similarly be treated as purchased and consumed within the household.

A final point about the DANTE cost database is that such databases will reproduce any biases incurred through dietary misreporting. Dietary assessment is recognised to be prone to measurement error (Freedman et al., 2011). This error will be reproduced in the cost estimates, where it exists. Under-reporting of food consumption, for example, will result in an underestimation of diet cost.

The DANTE cost database offers a method that is easy to apply to existing dietary survey data, and has advantages in its level of detail and in its ability to provide individual-level estimates. However, as identified above, the method is associated with several limitations. It would be valuable, therefore, to ascertain how this method compares to other methods for estimating diet costs. The following section describes two studies that carried out such a comparison.

5.4 Comparability with other costing methods

5.4.1 Introduction

This section describes two comparability studies conducted in existing data sets using the DANTE food cost database. Little is known about the validity of price databases in estimating costs from dietary assessment. This is largely because there is no gold standard against which the method can be validated, as all diet costing approaches involve a degree of inference. It therefore maybe worthwhile to assess the extent to which methods assigning prices to dietary assessment instruments agree with measures of expenditure, and the need for this has been documented (Murakami et al., 2008a).

Comparability of some diet cost methods has been investigated in the literature (Murakami et al., 2008a, Aaron et al., 2013, Monsivais et al., 2013). The first of these, conducted in a Japanese population, compared cost estimates using a price database applied to weighed dietary records against estimates of the same price database applied to a diet history questionnaire. The means across four time points of administration were correlated by 0.64 in women and 0.69 in men (Pearson's product moment). However, both methods in this study inferred purchases from consumption, the comparison being between cost estimates of different dietary assessment tools.

The comparability study of Aaron et al. (2013) examined estimates from store prices applied to a FFQ against estimates derived from till receipts along with 24-hour recalls in a sample of low-income women in California. Collecting dietary data alongside till receipts allowed the investigators to judge the quantity consumed by individuals and to account for free, non-purchased food. Bland Altman plots revealed a mean difference in the daily diet cost estimates of the two methods of $0.14, with 95% limits of agreement of -$7.76 and $7.48. This means that, in 95% of cases, individual diet cost estimates are likely to underestimate by $7.76 or overestimate by $7.48. This is a fairly wide interval considering the mean daily cost estimate of each method was found to be around $6.00.

The most recently published study, Monsivais et al. (2013), compared three methods of diet cost estimation, again using a US sample. The first method, like Aaron et al. (2013), concurrently collected till receipts with dietary assessment, but employed food diaries as opposed to 24-hour recalls. The second method estimated costs from the food diaries using a database of supermarket prices. The third method also used supermarket prices, but applied them to FFQs. The results indicated that the FFQ method estimated lower diet costs than the other two methods; however the mean difference between the FFQ method and food diaries combined with supermarket prices was small ($0.62, compared to a daily diet cost average of between $8 and $10). The mean difference between receipt cost estimates and the food diary estimates from supermarket prices was $-1.76.

There are no previously published studies in the UK comparing diet cost estimates from different methods. As this is the intended method for Chapters 6, 7 and 8, comparing the diet diary method to expenditure records will help in interpreting the findings. Given the discrepancies reported in nutrient values between FFQ and diet diary methods (see, for example, Bingham et al., 1994), which could be assumed to also apply to cost values, these further comparisons are necessary in order to add to the comparisons already presented in the literature.

Two prior research projects within the University of Leeds Nutritional Epidemiology Group independently collected food purchase receipts alongside diet diary records. Each data collection allows examination of the usefulness of the DANTE costing tool for populations of differing characteristics: one was carried out in the same year as the cost database creation, 2004, using a subsample of single-living females drawn from the UK Women's Cohort Study (UKWCS); and the other sample was taken from the Supermarket Nutrition Information Project (SNIP) which took place in 1998-99. Analyses on these prior studies had been carried out with the same objective – to attempt to validate the DANTE food cost database as a means of diet cost estimation.

Abstracts relating to these data have been presented at the Nutrition Society Meeting, 2005 (Oyston et al., 2005, Smyth et al., 2005). However, it was identified that there were drawbacks to the analytical methods used: the UKWCS study did not apply a correction for waste to the till receipts, nor did it report mean difference or limits of agreement; whilst analysis of the SNIP data did not make use of all available data, and applied the consumer price index (CPI) to adjust for inflation. The analysis carried out for this chapter employs new methods.

The following objective was identified at the outset:

- To check the level of agreement between till receipt records of food bought and the cost estimate produced by DANTE for food consumed at home, using robust statistical methods.

5.4.2 Samples & data collection

This section (5.4.2) describes the samples and data collection methods for the two studies: the UKWCS and SNIP. Table 5.2 summarises the characteristics of each sample. The work described in this section (study design, recruitment, collection and inputting of data) were all performed prior to this PhD project, by other investigators. This previous work is acknowledged in Section 5.1.1.

5.4.2.1 UKWCS subsample

In 2004, 200 single-living women, randomly selected from the UKWCS cohort[2], were approached to participate in a food cost study, with fifty women agreeing to take part. The purpose of the study was to compare diet cost estimates from dietary assessment against those from till receipts. Participants were asked to complete a four-day food diary (two weekdays and two weekend days), and collect all till receipts for a two-week period. Participants indicated in the diaries if foods were homegrown or bought outside the usual household purchases (for example, at a work canteen). Diaries were entered and coded using DANTE. Homegrown food or FAFH (either assumed or as indicated) were not included in the diary coding.

Complete data were returned for 36 of the women (72% response rate, from those who agreed to participate). Participants were aged between 52 and 81 years, were of a majority professional occupation class, and 89% white (the remaining 11% of the sample did not report ethnicity).

5.4.2.2 SNIP sample

The SNIP study's main aim was to assess the validity of using supermarket purchase information to estimate nutrient intake (Ransley et al., 2003). As such, the sample of households (n = 284) was recruited from the Tesco Clubcard database held at the Roundhay store in Leeds. The study was conducted in 1998-1999.

Households were instructed to collect till receipts of all purchases of food for human consumption made over a 28-day period. In addition, a weighed intake diet diary was completed for every member of the household over four days (three weekdays and one weekend day). (Other dietary assessment methods were employed in the SNIP; however the diaries only were considered for use in the validation of the cost database.) Diet diaries were coded using the Weighed Intake Software Program (WISP), for Windows v1.2. WISP is a nutrient analysis package with a similar premise

---

[2] More information about the UKWCS can be found in study reports (for example, Cade et al. 2004, The UK Women's Cohort Study: comparison of vegetarians, fish-eaters and meat-eaters. *Public Health Nutrition* 7(7): 871-878).

to DANTE. Foods and drinks are included as individual items, with nutrient information per 100g assigned to each. In both WISP and DANTE, much of the nutrient information is taken from the UK's nutrient reference tables (Holland et al., 1991).

The completion rate of the SNIP study was 75%, with data available for 214 households, comprised of 522 individuals. The sample was reduced to 326 individuals from 161 households after excluding individuals with missing household composition data. The final sample had a mean household size of two (ranging from one to five), and included adults (n=256, 79%) and children (n=69, 21%). White ethnicity comprised the majority (94%), and 53% were female. A more detailed description of the sample can be found in Ransley (2002).

Table 5.2 Characteristics of the samples

Descriptor	SNIP sample (1998-99)	UKWCS sample (2004)
Individuals, n	326	36
Households, n	161	36
Mean household size (range)	2 (1-5)	1
% White	94	89*
% Female	53	100
Age range, years	1-87	52-81
% Adult	79	100
Social class of the majority	Intermediate and junior non-manual (50%)	Professional (39%)
BMI adults, kg/m^2 (95% CI)	25.01 (24.45 to 25.57)	25.06 (22.90 to 27.22)
Mean[1] daily energy intake[1], MJ (95% CI)	7.15 (6.88 to 7.43)	7.89 (7.16 to 8.62)

*The remaining 11% of the sample did not report ethnicity

[1] Energy intakes as calculated from diet diaries

5.4.3 Data cleaning

On examination, it was felt the data would benefit from further cleaning and re-analysis. This was undertaken with the help of fellow PhD candidate Michelle Morris. Quality Assurance (QA) checks were carried out for each sample, in which raw data for a random subsample (5% of the SNIP sample and 10% of the UKWCS sample) were checked against the data recorded in the databases. Details of the results of these QA checks can be found in Appendix B.

5.4.3.1 Till receipt data

Recorded totals for food expenditure from the till receipts were compared to those calculated from raw data. Only minor discrepancies were apparent in all the QA checks.

On re-examining the UKWCS till receipts, the originally recorded data was mostly identical to the raw data, with a discrepancy found for only one participant.

Recalculated raw till receipt totals for the SNIP sample were within 1% of the originally calculated totals for almost half the QA sample, and totals for just two participants were found to differ by more than 5%. As a result, the general level of accuracy of receipt calculations was deemed satisfactory for the purposes of this study.

5.4.3.2 Dietary data

For the dietary data, diaries were re-entered, and energy intake totals compared to the originally coded data.

The originally coded UKWCS data was again found to have a satisfactory level of agreement with the estimated energy totals of the QA analysis – showing a difference in energy intake of less than 5% for all participants examined.

In the QA check of the originally coded diary data of the SNIP sample, on the other hand, energy intakes appeared to vary widely. On examination of the data, it became apparent that a large number of foods were missing in the originally coded data. It was discovered that this was due to a mismatch between food item codes from WISP and those in DANTE. Although both programs use codes from McCance & Widdowson's nutrient tables (Holland et al., 1991), some codes have been updated in subsequent editions or supplements.

A total of 868 food item codes were missing from DANTE. The food codes were updated manually as a result, to match the SNIP data to the DANTE codes. One hundred and nineteen codes could not be replaced in this fashion, however, either because they did not appear in any edition of McCance & Widdowson (for example, diet lemonade) or because they were unique recipes. The most commonly occurring of these were hand-searched in the original diaries so an equivalent DANTE code could be assigned to each. Following replacement, 169 individuals still had missing data; and were excluded from further analyses.

Following the correction of food codes, as described above, a second QA check was attempted on the diary data of the SNIP sample. This time, energy intakes of re-entered data were found to be within 1% of the original energy intakes for the majority

(42%) of the QA sample. Results for five participants of the sample, however, showed a difference in estimated energy intakes of greater than 10%. Nevertheless, the general level of accuracy was deemed satisfactory for the purposes of the validation study.

5.4.4 Estimation of diet costs

UKWCS subsample

Diet diary data was assigned a cost using the DANTE food cost database as described in Section 5.3.3. A daily mean cost was calculated from the total cost recorded.

Till receipts were summed, following exclusion of non-food items. Totals were divided by the number of days (14) to give a daily estimate. To account for waste resulting from spoilage, inedible parts or discarding, a correction factor of -15% was applied to the till receipt figures (as recommended by the Department for Environment Food and Rural Affairs (Defra 2010)).

SNIP sample

Total diet costs were generated from the diet diary information using the mean values in the DANTE cost database, and an average taken across the days.

The original 2004 comparability study (see Section 5.4.1) used the Consumer Price Index (CPI) to adjust the DANTE food cost database for inflation from 1998/99 (when the SNIP data were collected) to 2004 (the year the DANTE food cost database was populated). However, the CPI contains an inflation estimate averaged across a range of consumption goods, not limited to food. As such, it was considered a crude tool for adjusting the price information. Instead, for the present study, data from the Office for National Statistics (ONS, 2011) were used to calculate an inflation index for each of the 27 food groups for which there are data. These will reflect the different rates of inflation experienced by each food group. The food groups and indices are listed in Appendix C. These were applied manually to the 1998/99 DANTE costs to bring them in line with 2004 prices.

The total household expenditure on food was divided by the household size to give a per capita diet cost. A correction factor of -15% was again applied to account for waste and spoilage (see above). The corrected total was then divided by the number of days of data collection (28) to express as a daily average.

5.4.5 Analytical methods

Summary statistics were generated for each cost estimation method (DANTE cost database with diet diaries and till receipt calculations) within each sample. Pearson product-moment (for normally distributed data) and Spearman's rank (for non-normal data) correlations were conducted for each sample.

The daily diet costs calculated from the till receipts were tested for agreement to the costs estimated by DANTE using Bland Altman (BA) difference plots. BA plots assess the agreement of two methods by plotting – for each participant – the mean of the two methods against the difference between the two methods (Bland and Altman, 1986). In this manner, it is possible to ascertain if one method biases measurements (showing the mean difference of the whole sample) as well as gauging limits within which we would expect to find individual-level differences in the measurements.

In the SNIP sample, sensitivity analyses were undertaken, excluding the top 5% of estimates in each collection method. In addition, subgroup analyses were also performed in the SNIP sample, with separate BA plots for males and females, and for adults and children. Sensitivity analyses and subgroup analyses were not possible in the UKWCS subset, due to the small sample size.

Statistical analyses were performed using Stata IC 11 (StataCorp, 2011).

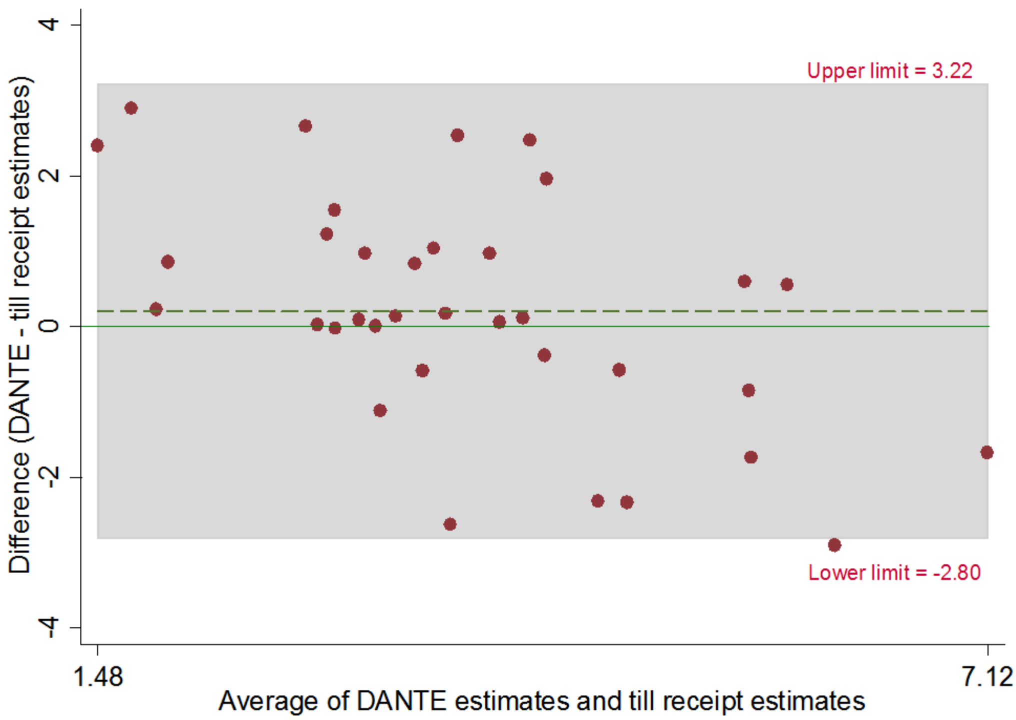
5.4.6 Results

5.4.6.1 UKWCS subsample

The data were normally distributed for both cost estimation methods. The mean daily cost given by the till receipts (adjusted for waste) was £3.75 (SD £1.83); for DANTE it was £3.96 (SD £1.08). The estimates of the two methods were moderately and significantly correlated (r = 0.547; 95% CI 0.261, 0.745; p<0.001).

Plots of the differences between the means indicated normal distribution[3]. In plotting the differences, there was one outlier evident, which was subsequently excluded. A Bland Altman plot of the differences can be seen in Figure 5.3. The mean difference between the methods was £0.21 (range: -£2.90 to £2.90), with 95% limits of agreement (±2σ) of -£2.80 and £3.22. No noteworthy bias toward over- or underestimation was evident (indicated by dashed green line on Figure 5.3).

Figure 5.3 Bland Altman plot of the difference between DANTE daily estimated cost and till receipt daily estimated cost (adjusted for waste), for the UKWCS subsample (n=35)



[3] In assessing level of agreement, the assumption is that the differences between the variables are normally distributed, rather than the variables themselves. This is because the limits of agreement are based upon the standard deviation (σ) of the differences. See Bland & Altman (1999) for a further discussion.

5.4.6.2 SNIP

Diet cost estimates for the SNIP data were found to deviate from a normal distribution. There was also an outlier evident in plotting the differences. On investigation, it was found that the outlier was due to large volumes of alcohol and bottled water consumed by one individual, which was not reflected in the till receipt data that had been averaged across the household. The outlier was dropped from subsequent analyses. Following removal of the outlier, the median daily cost estimated by DANTE was £2.88 (IQR £2.01, £3.72); the median daily cost calculated from till receipts (adjusting for waste) was £2.71 (IQR £2.16 to £3.73). The estimates of the two methods were found to be significantly, though not strongly, correlated (Spearman's ρ = 0.384; 95% CI 0.287, 0.473; p<0.0001).

The mean difference between the estimates of the two tools was £0.10. Differences ranged from -£4.29 and £5.91, and the distribution of differences appeared normal. The Bland Altman plot of the differences can be seen in Figure 5.4, which shows 95% limits of agreement (±2σ) of £2.88 (upper) and -£3.08 (lower). Degree of bias was minimal. However, it is apparent on the plot that the spread of scatter points widens as the mean difference between the methods increases. This was confirmed by the fitting of a regression trend where the 95% confidence limits were seen to widen along the x axis Figure 5.5. This demonstrates reduced agreement at higher costs.

Figure 5.4 Bland Altman plot of the difference between DANTE daily estimated cost and till receipt daily estimated cost (adjusted for waste), for the SNIP study (n=325)

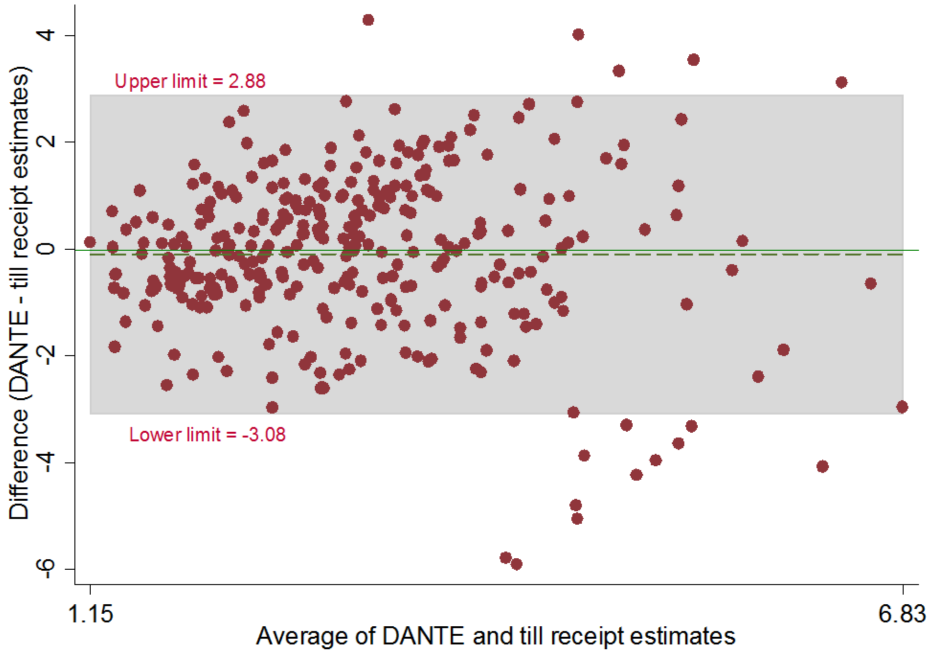
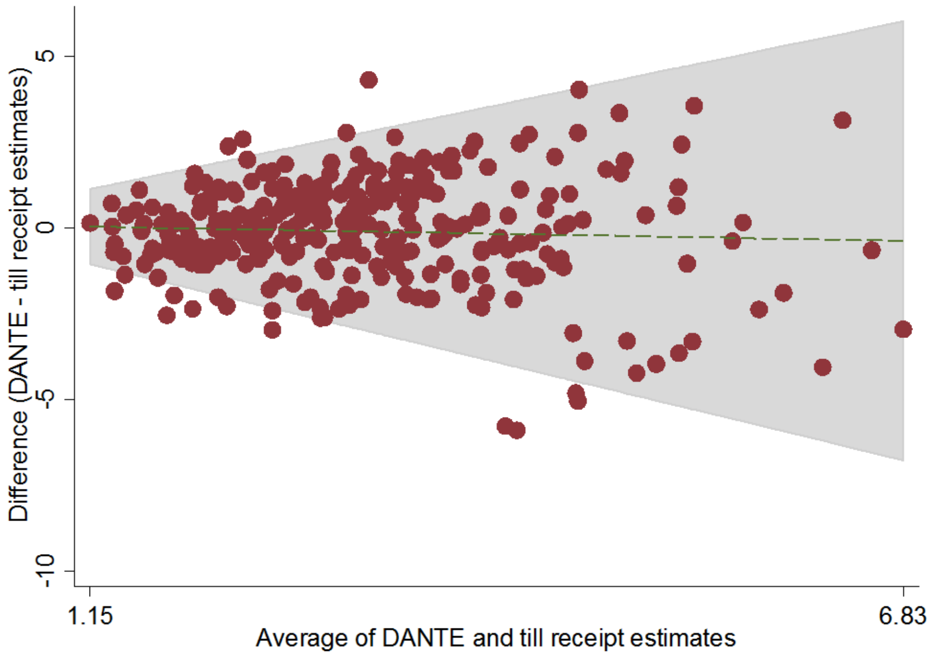


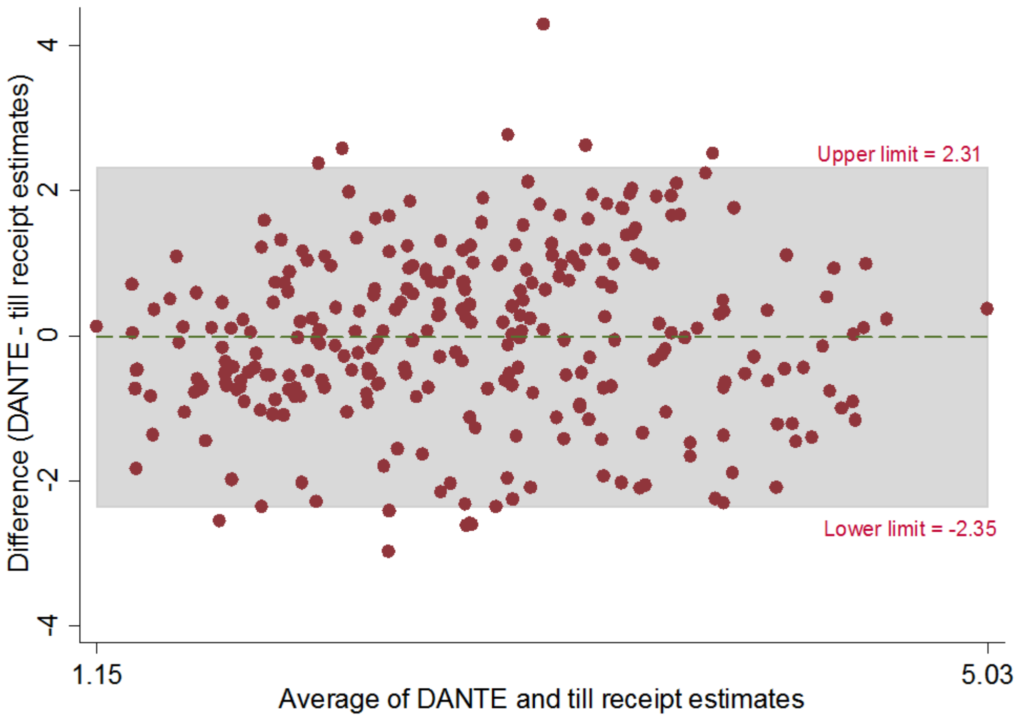
Figure 5.5 Bland Altman plot showing the differences between DANTE and till receipt estimates for the SNIP study, regression trend fitted (n=325)



Sensitivity analysis

A sensitivity analysis was performed by excluding the top 5% of values in each collection method, giving a sample size of 292. This resulted in lower estimated daily costs: median £2.75 using DANTE (IQR 1.88 to 3.55); and £2.58 from the till receipts (IQR £2.09 to £3.45). The BA plot (Figure 5.6) showed narrower bias (mean difference = £0.02) and limits of agreement (£2.31, -£2.35). In addition, there was no obvious fanning evident in the plot.

Figure 5.6 Bland Altman plot of the difference between DANTE daily estimated cost and till receipt daily estimated cost (adjusted for waste), with top 5% values excluded (n=292)



Subgroup analyses

The median daily cost estimated by DANTE for males (n=152) was £3.07 (IQR £2.15, £3.89); for females (n=172) it was £2.63 (IQR £1.78, £3.51). Correlation coefficients (Spearman's rho) were similar for both males and females (Table 5.4).

Children displayed lower estimated costs compared to adults, especially when using DANTE estimated costs (Table 5.3). Analyses revealed cost estimates to be less strongly correlated when adults and children were tested separately (Table 5.4): adults' cost estimates from till receipts and those from DANTE were significantly correlated (ρ = 0.354, 95% CI 0.242, 0.457; p<0.0001); however cost estimates for children were not significantly correlated (ρ = 0.197, 95% CI -0.045, 0.418; p = 0.354).

Table 5.3 Median (IQR) estimated daily dietary costs (£) of sample subgroups

Subgroup	DANTE cost database (£)	Till receipts (£)
Males (n=152)	3.07 (2.15 to 3.89)	2.76 (2.15 to 3.75)
Females (n=172)	2.63 (1.78 to 3.51)	2.69 (2.16 to 3.72)
Adults (n=256)	3.06 (2.32 to 3.10)	2.77 (2.26 to 3.81)
Children (n=67)	1.83 (1.39 to 2.51)	2.31 (1.96 to 2.96)

Table 5.4 Correlations between till receipt and DANTE cost database estimations

	Spearman's rho	95% CI	p value
Males (n=152)	0.375	0.229, 0.504	<0.0001
Females (n=172)	0.401	0.268, 0.520	<0.0001
Adults (n=256)	0.354	0.242, 0.457	<0.0001
Children (n=67)	0.197	-0.045, 0.418	0.354

Bland Altman plots were created separately for each subgroup – males, females, adults and children – and sensitivity analyses excluding the top 5% were performed in each case. Mean differences and 95% limits of agreement are presented in Table 5.5; the plots can be seen in Figure 5.7 and Figure 5.8. All subgroup plots showed widening limits of agreement, indicating reduced agreement at higher costs.

Males exhibited a similar pattern in agreement to the whole sample, both with and without the top 5%. On exclusion of the top 5%, females showed a reduction in the widening limits of agreement, but not to the extent of the whole sample, or of males.

On excluding children, the mean difference was as small as £0.01, although limits of agreement remained similar to the whole sample estimates. Although the limits of agreement narrowed on excluding the top 5%, the mean difference between the methods increased when adults were analysed alone.

Table 5.5 Summary of Bland Altman subgroup results, with or without the top 5% (£)

				Excluding top 5%		
	Mean difference[1] (bias, £)	95% limits of agreement (£)		Mean difference[1] (bias, £)	95% limits of agreement (£)	
		Lower	Upper		Lower	Upper
Full sample	-0.10	-3.08	2.88	-0.02	-2.35	2.31
Males (n=152)	0.07	-2.95	3.09	0.16	-2.21	2.52
Females (n=172)	-0.27	-3.16	2.63	-0.19	-2.42	2.04
Adults (n=256)	0.01	-3.08	3.09	0.11	-2.18	2.41
Children (n=67)	-0.55	-2.86	1.75	-0.50	-2.67	1.67

[1] Mean of DANTE cost database minus till receipt estimates

Figure 5.7 Bland Altman plots for males and females, including and excluding the top 5%

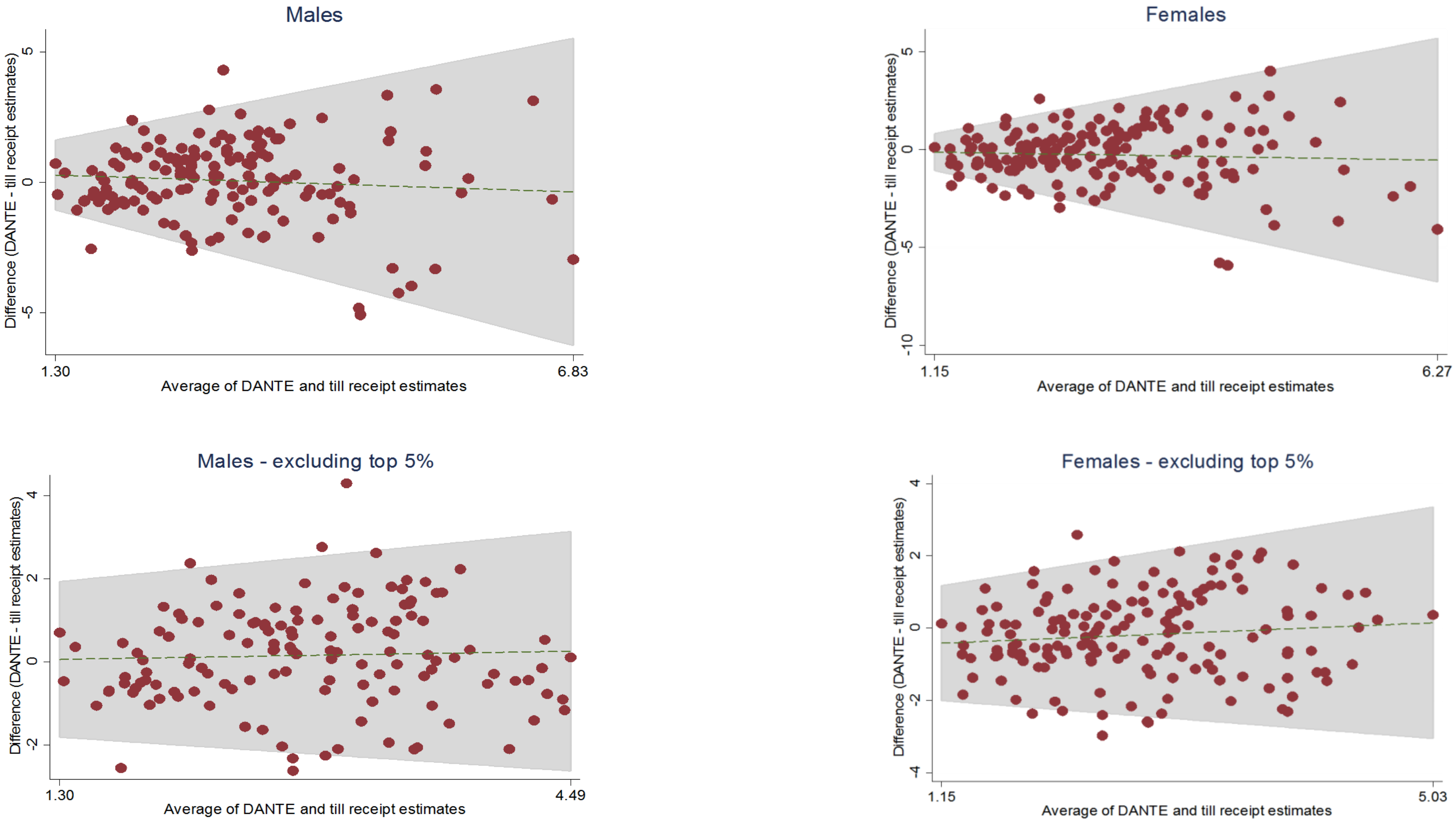
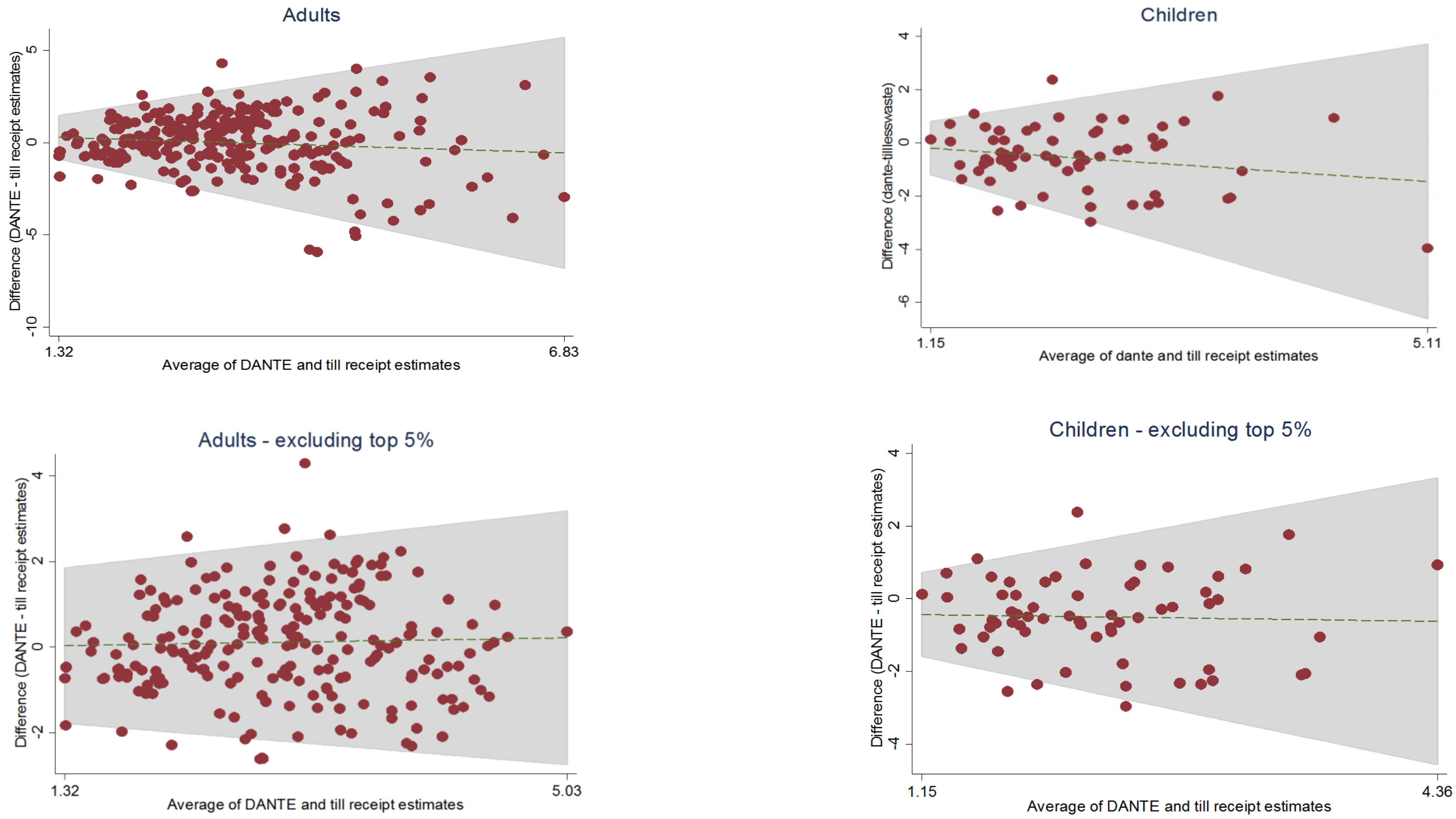


Figure 5.8 Bland Altman plots for adults and children, including and excluding the top 5%



5.5 Discussion

Previous comparability studies have examined estimates of a food cost database applied to different dietary assessment tools (Murakami et al., 2008a), or the estimates of a price database applied to a FFQ against till receipt estimates (Aaron et al., 2013). One study (Monsivais et al., 2013) took a similar approach to the analyses in this chapter, in comparing receipts against diet diary estimates. This chapter adds to the work of Monsivais et al. (2013), to enable better interpretation of the analyses in subsequent chapters of this thesis.

Analyses of the UKWCS subsample and the SNIP study produced similar results: the mean difference between the cost estimates of the two methods was modest in both cases, and both BA plots displayed comparable limits of agreements. The results suggest that the DANTE cost database could overestimate or underestimate the daily diet cost for an individual by roughly £3.00. With a mean daily cost of around £3.00, this constitutes a potentially substantial difference. However, the full-sample mean differences were as little as £0.10, which suggests that the two methods agree relatively well in estimating dietary expenditure at a group level.

These findings are not dissimilar to the study of Aaron et al. (2013), in which a small mean difference ($0.14) was apparent between estimates from till receipts and a cost database applied to FFQ, whilst the limits of agreement were fairly wide. In the latter study, the limits of agreement exceeded the average daily cost estimate, being around ±$7.50 compared to a mean cost estimate of around $6.00. The wider limits of agreement may have been due to the dietary assessment method used, or it may be a result of the different sample and setting used. Monsivais et al. (2013) found a larger mean difference between till receipt estimates and costs calculated from diet diaries using market prices, but a small mean difference between FFQ and diary estimates. Again, the slightly different findings could be due to sample and context differences.

In contrast to the study of Aaron et al. (2013), examination of the SNIP sample showed evidence of widening limits of agreement with increasing estimated diet costs. When the more expensive diets in the sample were excluded, both the mean difference between the two methods and the limits of agreement were reduced. This implies that the database and till receipt estimates agree best for the 95% of the sample spending less on their diets.

In the subgroup analyses, the between-group differences of both methods were in the same direction. These were greater when using DANTE to estimate costs, rather than till receipts. There was variation in the methods' agreement between males and females, and between adults and children. In particular, the DANTE cost database estimates for children varied noticeably from the till receipt values, on average exhibiting lower costs. This most likely reflects a drawback in the till receipt method, which assumed an equal consumption across members of the household. In actuality, both the quantity and composition of diet is likely to differ across the family unit (Bates et al., 2011), patterns which are more likely to be captured using dietary assessment. The results of this study support this, showing decreased agreement in the subgroups likely to consume a smaller quantity of food.

The fact that the SNIP study took place in a year different to that of the cost data collection is of particular usefulness, in that it permits the assessment of employing inflation correction factors in such a database. Correction factors can be derived from the annual food price indices compiled by food group by the Office for National Statistics (Defra, 2009). Adjustments of prices (by food type) made according to national price indices have been previously found to yield similar estimates to real-time estimates for some, though not all, of a small sample of food items (Friel et al., 2001). Finding a way to apply the cost data to different time periods will augment its usability, and uniquely allow comparisons in trends across time. It would be informative to formally test the consequences of separating the CPI into food group-specific indices on estimate accuracy as compared to simply applying the CPI.

Estimating dietary expenditure will always have its limitations, but the food cost database remains both a pragmatic method for large-scale dietary research, and the one most likely to deliver the clearest picture of individual-level diet costs. Further explorations might investigate whether the accuracy of DANTE estimations differs according to various demographic or household characteristics or dietary patterns.

5.5.1 Limitations

As a comparison study, it is difficult to draw conclusions about the validity of either method used. However, understanding more about how the best available methods relate to each other could help to enhance the comparability of findings across the literature, whilst a more precise measure of the actual cost of daily intake is still lacking.

Many of the limitations associated with the DANTE database diet cost estimates have already been discussed in Section 5.3.4. In brief, it should be remembered that the method of applying a cost database to dietary data will inevitably echo any biases or measurement error associated with the dietary assessment tool used. Secondly, cost databases tend not to be able to account for food away from home (FAFH), free foods, promotions, or price variability. It is possible that the limits of agreement seen in this study reflect a variation in product prices: only the mean costs of each food item from the DANTE cost database were employed, whereas lower-than-average or higher-than-average costs may have been represented in the till receipts. In future applications of the DANTE cost database, there is the potential to use the low and high values within the database.

Within the DANTE cost database, some infrequently consumed foods lack cost information - for example, some exotic fruits (rambutan) and offal (trotters and tails). None of these foods occurred in the UKWCS food diary data, but six uncosted items were reported in the diaries of four participants from the SNIP sample. This may have resulted in an underestimation of expenditure for these participants. It is unlikely that the small amounts involved will have skewed the results.

5.5.2 Strengths

As mentioned in Section 5.3.4, the DANTE cost database boasts a key advantage over household expenditure data in that using dietary assessment methods provides data at the individual level. This is of particular relevance when investigating the economic determinants of diet and health.

In addition, this study is valuable in that the samples examined exhibit different characteristics: only single women, who shop for one person, were included in the UKWCS sample; whereas complete households were recruited in the SNIP study. The samples also differed in size and in the year in which the data was collected. This variation adds strength to the conclusions, with similar findings for both samples.

5.6 Conclusions

Cost of diet is likely to warrant an increasingly important role in public health research. The increasing economic pressures of recent years have elicited growing concern about the affordability of a healthy diet, and establishing whether diet costs contribute to inequalities in health could have far-reaching policy implications.

This chapter has introduced the DANTE food cost database, the tool which is to be used in the following chapters to explore diet costs in the NDNS. A description of the main limitations of this approach was included, and the extent of these limitations was assessed by comparing the DANTE food cost database to the alternative method of estimating diet costs, using household till receipts.

The DANTE food cost database linked to a dietary assessment tool agrees well with estimates from household expenditure at a sample level, for two contrasting samples. This suggests that calculating the cost of food using dietary assessment data is useful in estimating the monetary value of a population's diets. At the individual level, diet cost estimates showed less agreement. In the SNIP study, agreement was stronger for the 95% of the population spending less on their diets, and for adults.

This comparison of methodologies was critical for the interpretation of diet cost research. The results suggest that using a cost database linked to food composition tables is a pragmatic method for large-scale dietary research. This should help improve confidence in the findings of Chapters 6, 7 and 8.

What was known previously:

- There is a need to measure diet cost at the individual level, in order to link food prices to health outcomes.
- Identifying the actual effect of food prices on diet and health in a real-world setting is challenging, and investigators often rely upon estimates of diet costs.
- The DANTE food cost database holds national UK price information for over 3,000 food and beverage items which can be linked to dietary data to estimate diet costs.
- Due to the element of approximation inherent in costing diets, it is important to gauge how diet costing methods compare.
- There are no previously published studies in the UK comparing diet cost estimates from different methods.

What this chapter adds:

- The DANTE food cost database linked to a dietary assessment tool agrees well with estimates from household expenditure at a sample level, for two contrasting samples.
- At the individual level, estimates were found to differ by as much as £3.00 per day.
- Agreement was stronger for the 95% of the population with lower diet costs and for adults.
- This chapter adds to the work of previous authors, in populations outside the UK, to enable better interpretation of diet cost analyses.

Chapter 6 Estimating the diet costs of NDNS adults

6.1 Summary

According to the food price-obesity hypothesis, varying prices of foods may determine the selection of some foods over others, potentially encouraging the purchase of energy-dense foods. Whilst the role of income in energy balance is discussed in an earlier chapter (Chapter 4), the next few chapters concentrate on a key supply-side determinant of dietary purchases – food prices. The previous chapter (Chapter 5) introduced a tool for inferring the costs of diets as recorded in dietary surveys, the DANTE cost database. This chapter describes the costs of British adults' diets, as estimated by applying the DANTE cost database to the dietary data of the National Diet and Nutrition Study (NDNS).

A cost was assigned to each food and beverage (excluding water) recorded in the adult diet diaries from the first two years of the NDNS (2008/09-2009/10; n=1014). Daily diet costs were calculated, both including and excluding costs of alcoholic beverages, and costs per 10MJ were also calculated in order to improve comparability across individuals with differing energy requirements. The chapter presents descriptive results of these estimated diet costs, including descriptive statistics by sociodemographic groups and other lifestyle variables.

The median daily diet cost of the sample was £2.84 (IQR £2.27, £3.64). Energy intake and daily diet cost were strongly correlated. The median energy-adjusted cost was £4.05 per 10MJ (£3.45, £4.82). Univariate analyses indicated that diet costs differed significantly between categories of many of the sociodemographic variables. Observed differences were, for the most part, as anticipated.

Multivariable regression assessed the effects of each variable on diet costs after adjustment, indicating that: food energy intake, income and fruit and vegetable intake were associated with daily diet costs; whilst sex, BMI category, income and fruit and vegetable intake were associated with diet costs per 10MJ.

This is the first time monetary costs have been applied to the diets of NDNS adults. The findings suggest that certain sociodemographic groups in this sample consume diets of lower monetary value. The potential influence of inflation was also considered by comparing unadjusted diet costs with those estimated after the application of food group-specific inflation indices. The results set the context for the investigations into diet costs, dietary energy density and BMI in the NDNS which are the subject of the following chapters.

Some of the analyses in this chapter form the basis of a publication in the journal Public Health Nutrition (Timmins et al., 2013a). The results presented in the article differ to those included in this chapter, however, in that survey weights were applied to the analyses. This was because the emphasis of the article was on describing the estimated diet costs of British adults, whereas this chapter is intended as a precursor to the regression analyses of Chapter 7 and 8 (see Section 3.3.1).

6.2 Introduction

The purchasing power of an individual is determined by their income, the prices of foods, and the person's consumption of other goods (see Chapter 1). Consumers must reconcile their purchases within their food budget; therefore, varying prices may determine the selection of some foods over others. The price of food is reported as a dominant aspect of conscious decision-making in many samples (Steptoe et al., 1995, Connors et al., 2001, Shepherd et al., 2006). In the UK, almost a third of respondents in the Low Income Diet and Nutrition Survey (LIDNS) identified price, value or budget to be the most important influence on their dietary choices (Nelson et al., 2007).

Whilst national-level food price and consumption data enable the monitoring of trends, they cannot portray the effects prices may have on individuals' dietary behaviour. Price elasticities can reveal changes in demand in response to price changes of a specific food (see Andreyeva et al., 2010 for a review); however, it is difficult from this type of data to elucidate changes to the whole diet and to dietary intake that might occur as a result. Allocating prices to all the foods consumed by an individual, on the other hand, can give an indication of the cost of their whole diet.

Although not an accurate reflection of individuals' own food expenditure (see Chapter 5), prices applied to foods consumed could indicate the value of diet that they can afford (or choose to afford). It may be possible from this estimation of diet costs to speculate the extent to which price considerations have guided food selection.

Previous publications have used national or local food price databases to apply a monetary value to the diets of American (Monsivais and Drewnowski, 2009, Rehm et al., 2011), French (Darmon et al., 2004, Maillot et al., 2007a), Dutch (Waterlander et al., 2010), Spanish (Schroder et al., 2006, Lopez et al., 2009) or Japanese (Murakami et al., 2007) populations (see Chapter 2). To date there have been no such studies in a representative UK sample, however, and dietary costs have never been estimated for the NDNS.

This chapter describes for the first time the monetary value of adults' diets in the NDNS. The method of costing diets – using the DANTE cost database – is outlined in Chapter 5. As newly derived variables, a thorough exploration of descriptive statistics by sociodemographic and other subgroups are included, along with some univariate tests for comparison and correlation. In addition, the chapter will explore the appropriateness of applying different inflation indices to diet cost estimates. Diet costs were estimated both including and excluding alcohol.

The research presented in this chapter satisfies the following objectives:

1. Estimate and describe the diet costs of NDNS adults
2. Explore patterns in NDNS diet costs according to sociodemographic characteristics
3. To investigate the appropriateness of diet cost estimations

Elucidating patterns in diet costs could have implications for the targeting of public health nutrition messages. In addition, individual-level data allow the exploration of relationships between diet costs, dietary quality and health outcomes. Such investigations are the focus of Chapters 7 and 8; therefore this chapter sets the scene for these later chapters.

6.3 Methods

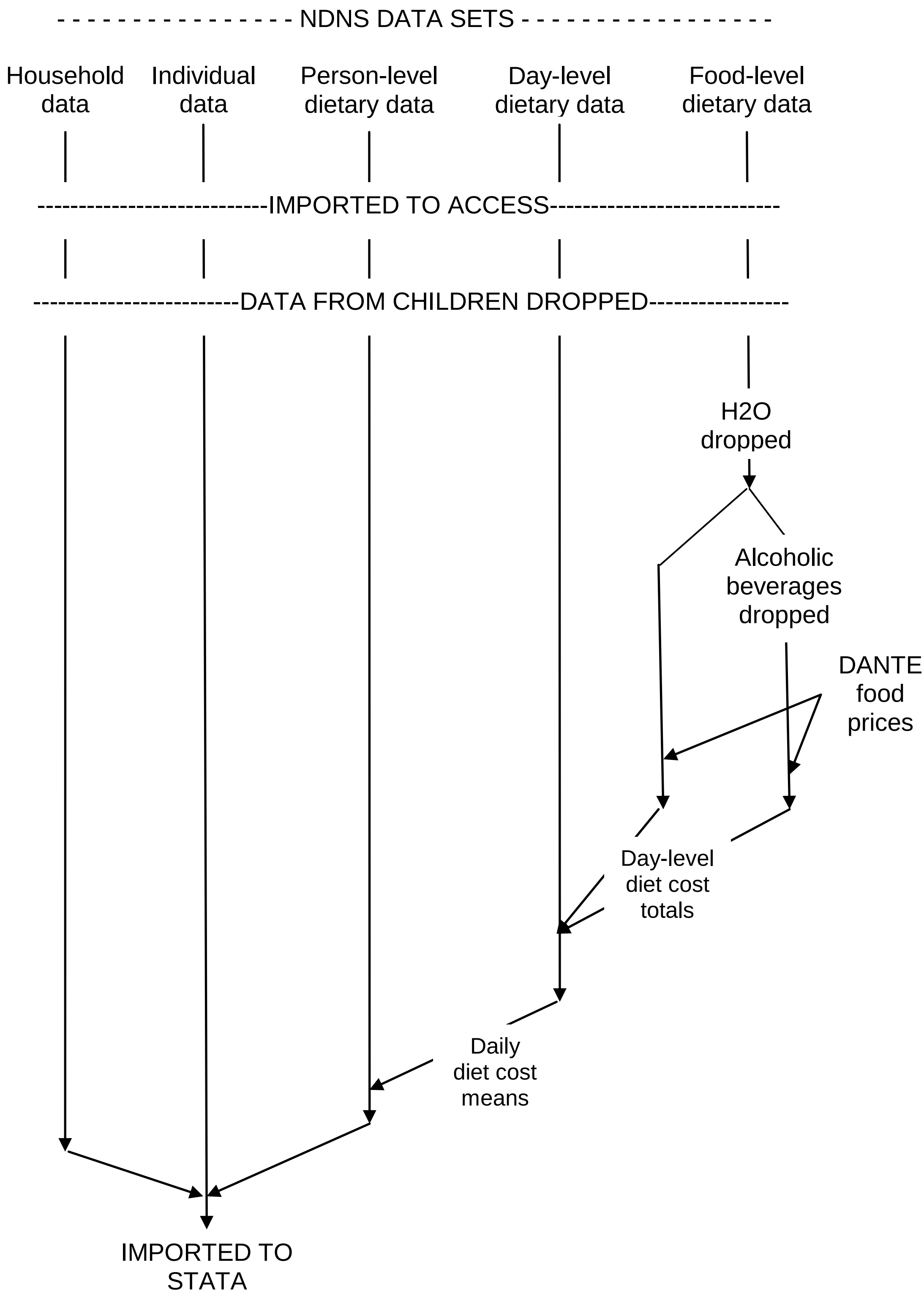
The data used in this chapter were previously collected and compiled by other investigators. The methods employed to furnish these data sets are summarised in Chapters 3 and 5. Details on the sample and data collection can be found in these chapters. Information regarding ethical approval is also contained within these chapters. The sections below describe the linking of the data sets to derive diet costs, along with the statistical methods adopted.

6.3.1 Linking the data sets

The NDNS data set was downloaded from the Economic and Social Data Service (ESDS) repository in December 2011, as a Stata (StataCorp, 2011) data file. Appropriate variables were selected (see Chapter 3) and exported to a database in Microsoft Access 2007. This was to allow the data to be linked to the DANTE cost database, also housed in Access.

To assign a cost to individuals' diets in the NDNS, it was necessary to allocate a price to each food or beverage consumed. This was achieved by linking the NDNS data to the DANTE cost database. However, the food codes employed by DINO, the tool used to code the NDNS, differ to those of DANTE. Therefore, it was necessary to first match the food item descriptions of the two databases, then to add a vector to the DANTE food table containing the DINO codes. In this manner, the tables could be linked via the DINO codes. Figure 6.1 summarises the process involved in linking the databases.

Figure 6.1 Flowchart depicting the process of linking data sets to calculate diet costs



A list of unique food items occurring in the NDNS adult food data was generated (a total of 3416 items); these were then manually matched to the DANTE food items, and the appropriate DINO code was entered in the DANTE cost database.

Both databases incorporate data from the UK food composition tables (FSA, 2002), and, as such, it was possible to match many of the food items exactly (30% of foods). Of the remainder, 232 items (7%) were not appropriate for inclusion in the analyses (these comprised the supplements food group, medicines and sundries such as sugar-free chewing gum and sugar replacements), and 40% (n=1369) were assigned to duplicate entries in the DANTE database. Duplicate entries were necessary where the DINO database included several versions of a food to allow for more detailed dietary assessment: for example, 'fried egg' has five separate DINO codes to reflect different cooking oils (blended oil, butter, lard, margarine, PUFA), whereas in the DANTE cost database there is only one option for fried egg (in vegetable oil). Matches described as a 'close alternative' mostly reflected minor discrepancies in the item description – for example, "peas boiled in salted water" could be matched to "peas boiled in water".

The DANTE cost database does not contain prices for a number of food items that were unavailable at the time of the database creation (see Chapter 5) – for example, some game items (pheasant, partridge) and ethnic foods (plantain, enchilada). Twenty of the foods in the NDNS data set, consumed by 62 adults, were found to have missing costs in the DANTE cost database. The problem was irresolvable at this stage in the cost database's development, due to the time that had lapsed, and these missing costs remain a limitation in the diet cost estimates.

6.3.2 Assigning costs to diets

The diet cost variables that needed to be calculated from the linked data sets were:

- Daily diet cost estimates, including alcohol;
- Daily diet cost estimates, excluding alcohol;
- A calculation of diet cost including alcohol in relation to total energy intake; and
- A calculation of diet cost excluding alcohol in relation to food energy.

Before the DANTE costs were applied to the dietary data, a subset table was created in which water was excluded. Uncarbonated water was excluded from the diet cost calculations because it was not possible to distinguish from the data whether the water consumed was free tap water or purchased bottled water.

Once estimated, the values for each individual were added to the Stata NDNS data set as a new variable (for each cost parameter). Whilst the variables themselves were estimated in pence, for the purposes of clarity in interpretation, figures are reported as GBP£.

6.3.2.1 Daily diet cost

Mean daily diet costs were derived for each individual by multiplying the food price in the DANTE cost database by the quantity of food consumed, summing for each day, and calculating the average across the four days[4] of dietary intake data collection:

$$\text{Daily diet cost } (\pounds\ \text{day}^{-1}) = \frac{\underline{\sum(\text{ DANTE price (p/g)*quantity food consumed (g)})} \div 100}{\text{Number of days (4).}}$$

Equation 6.1

Costs excluding alcohol were derived using the same formula applied to the Access subset in which alcoholic beverages were excluded.

6.3.2.2 Correcting for inflation

The database was populated using 2004 prices, whereas the NDNS data was collected between 2008 and 2010. During that time, prices will be expected to have increased as a result of inflation. One way to correct for this and bring the 2004 prices in line with 2008-2010 prices would be to inflate the 2004 prices using the national index of inflation, the Consumer Price Index (CPI). The CPI is an inflation index averaged across a range of goods (see Section 5.4.4). Amongst these goods, the Office for National Statistics (ONS) includes indices for 27 different food groups, which averaged together make up the Food Price Index (FPI). An inspection of the food group data reveals varying patterns between the inflation rates of different food groups. Therefore, it cannot be assumed that applying costs from a different year will not modify the patterns of costs observed. On the other hand, the differences in inflation rates between food groups may be so slight as to make little difference.

[4] Participants with less than four days dietary data were excluded from analyses in this thesis – see Chapter 3.

To assess the possibility that food-group specific indices might more sensitively reflect the changes in prices faced by consumers, dietary costs were estimated under three alternative scenarios:

- Unadjusted, using the 2004 prices of the DANTE cost database;
- Adjusted using 27 food group-specific indices;
- Adjusted using a flat rate of inflation, the Food Price Index (FPI).

The price indices were derived from ONS data (ONS, 2011), from the detailed Consumer Price Index (CPI) reference tables. These tables use the reference year 1987 (1987=100). For the purposes of this study, new indices were derived with a reference year of 2004. This was achieved using the following formulae:

$$2008/09 \text{ index} = (((\text{Index08\_09} - \text{Index04})/\text{Index04}) * 100) + 100; \text{ or}$$

Equation 6.2

$$2009/10 \text{ index} = (((\text{Index09\_10} - \text{Index04})/\text{Index04}) * 100) + 100$$

Equation 6.3

where Index08_09 refers to the average of all months' indices from February 2008 to March 2009 (the NDNS Year 1 data collection period); Index09_10 refers to the average of all months' indices from April 2009 to March 2010 (Year 2 data collection period); and Index04 refers to the index at June 2004.

Two new vectors were added to the DANTE cost database: one containing the price of each food item adjusted using the FPI formula; and one with prices adjusted using the food group indices (see Appendix C for a full list of the indices). The DANTE food group codes were matched to the ONS food groups manually, before the correct food group index could be applied. The new vectors were populated using the following formula:

$$\text{Index-adjusted price} = \text{DANTE price} * \text{index}/100.$$

Equation 6.4

The new DANTE prices could then be used to create two new variables in the NDNS data set, containing estimated costs after adjusting for inflation. In order to apply the correct indices for each of the two years of data collection, the NDNS sample was split into each wave before the index-adjusted prices were assigned, then the sample was merged again.

As the FPI-adjusted costs were created using a single index, the relative differences in costs within the sample, and the proportions of cost attributed by each food group will be the same as the unadjusted diet costs. The food group-specific indices, however, may have created some differences in these relative and proportional costs, because each participant is likely to have consumed different quantities of each food group. Therefore, FPI-unadjusted costs and food group-adjusted costs were calculated by sociodemographic variables, and presented side-by-side to allow comparison.

6.3.2.3 Diet costs per 10MJ

Energy-adjusted costs were also calculated to control for the varying energy requirements associated with differing demographic groups (such as age). As with most nutrients (Willett, 1998), diet costs were predicted to be correlated with energy intake. Adjusting for energy should enable the identification of factors associated with diet costs independently of energy intake, making subgroup comparisons easier to interpret.

Daily costs were adjusted to 10MJ, selected as a midpoint between estimated average requirements (EARs) for males and females (SACN recommends EARs of 10.9MJ for men and 8.7MJ for women (adults aged 19+) (SACN, 2011)). The energy-adjusted daily diet cost was calculated using the following formula:

energy-adjusted cost = (mean daily diet cost (£)/mean daily energy intake (MJ)) x 10.

Equation 6.5

6.3.3 Analytical methods

Statistical analyses were performed using Stata IC 12 (StataCorp, 2011). After linking the databases and derivation of new variables, the data were imported into a new Stata data set. This data set combined individual-, household- and day-level data at the individual level, and also contained the diet cost variables.

Outliers for both diet cost variables were identified. To rule out the possibility of implausibly extreme diet cost estimates, the coded diaries of the participants in the top and bottom 1% of diet cost were examined. The foods and drinks consumed were judged to be plausible, and there were therefore no exclusions on this basis. Higher diet cost estimates appeared to be largely attributable to costs from alcoholic beverages or takeaway coffees.

Median daily diet costs (£ day^{-1}) and median energy-adjusted costs (£ $10MJ^{-1}$) were calculated for the whole sample and for each category of the following variables:

- Age group
- Sex
- Employment
- Qualifications
- Equivalized household income
- Household size
- Marital status
- BMI classification
- Cigarette smoking status
- Alcohol consumption category
- ‘5 a day’ achievement.

These variables were the sociodemographic indicators available in the NDNS data set which had adequate numbers of participants within each category (see Chapter 3). All cost variables were positively skewed; therefore median and interquartile ranges (IQR) are presented.

Sociodemographic differences in daily and energy-adjusted diet costs were tested using Kruskal-Wallis ANOVA. Where appropriate (with ordinal variables) a test for trend was used. A significance level of 5% was set.

Multivariable regression models were built to assess the strength of each variable’s relationship with diet costs (one model for daily diet costs, and one model for energy-adjusted diet costs), adjusting for the other variables. Due to probable collinearity, not all variables were included in the regression models. However, because these models are intended as exploratory rather than explanatory, unlike the regression analyses that have featured previously, there is not a single ‘exposure’ variable around which to build a direct acyclic graph (DAG). The variables selected for inclusion in the model were therefore chosen on the basis of anticipated sociodemographic differences. A minimum number of variables were included, to avoid including those variables which are highly correlated. Those selected a priori were: age group, sex, equivalized household income, BMI category, smoking status and ‘5-a-day’ achievement. In addition, energy intake from food was included in the model with daily diet costs, but not in the model for energy-adjusted diet costs, because energy was used in the derivation of the latter variable. With the exception of energy intake, all covariates were categorical. However, only cigarette-smoking status was entered in the model as dummy variables (using the Stata ‘i.’ prefix), because all other variables were made up of ordered categories. Regression models were built only for diet costs calculated without costs of alcoholic beverages.

Despite non-normal distribution of the diet cost variables, the residuals of each regression model were found to follow a normal distribution, and the dependent variables were found to have constant variance, therefore meeting the assumptions for linear regression.

6.4 Results

6.4.1 Whole sample

The median daily diet cost and energy-adjusted diet cost of the full sample, both including and excluding alcohol, can be seen in Table 6.1.

Table 6.1 Median values and interquartile ranges for average daily diet costs (£ day^{-1}) and costs adjusted to 10MJ for the whole sample (n=1014)

		Median	IQR
Daily diet cost (£ day^{-1})	Including alcohol	3.47	2.57, 4.83
	Excluding alcohol	2.84	2.27, 3.64
Energy adjusted diet cost (£ 10MJ^{-1})	Including alcohol	4.73	3.83, 6.00
	Excluding alcohol	4.05	3.45, 4.82

6.4.2 Diet costs & energy intake

The mean daily energy intake of the sample was 7699kJ (SD 2515kJ). Mean energy from food was 7242kJ (SD 2250kJ). The relationship between daily diet cost and energy intake was strongly positively correlated, both including alcohol (correlated with total energy: Spearman's rho = 0.68; 95% CI 0.65, 0.72) and excluding alcohol (correlated with food energy: Spearman's rho = 0.66; 95% CI 0.63 to 0.69). See Figure 6.2 and Figure 6.3.

Figure 6.2 Daily diet costs (p d^{-1}) including costs from alcohol plotted against total energy intake (kJ), NDNS adults (n=1014)

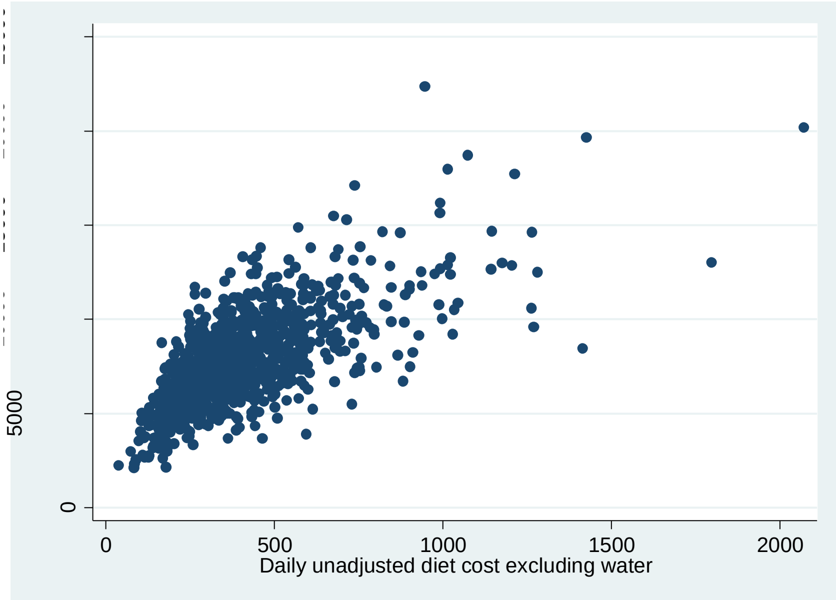
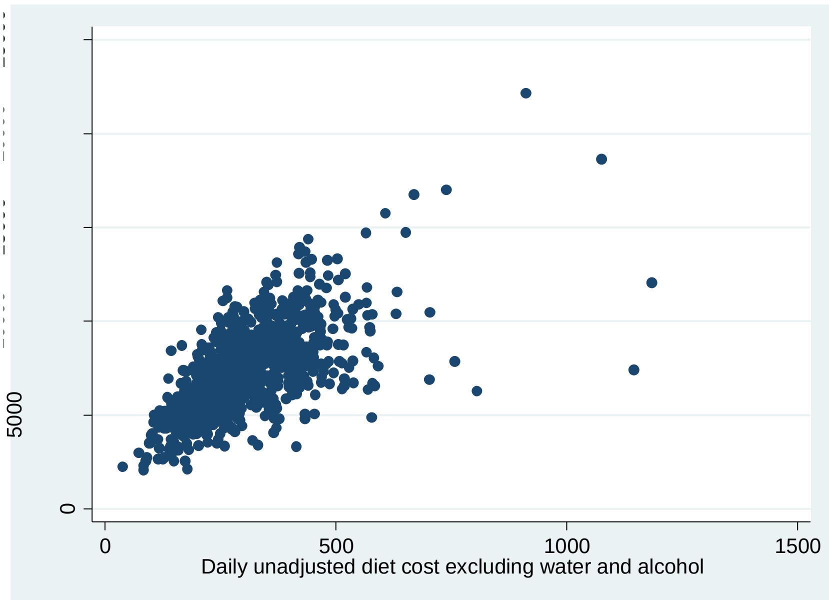


Figure 6.3 Daily diet costs (p d^{-1}) excluding costs from alcohol plotted against food energy intake (kJ), NDNS adults (n=1014)



6.4.3 Diet costs by sociodemographic characteristics

Table 6.2 shows the median diet costs by sociodemographic group, excluding alcohol. Kruskal-Wallis ANOVA revealed significant differences in daily diet costs by several sociodemographic variables when costs of alcoholic beverages were excluded. These were: sex, employment, marital status, qualifications and income. In addition, costs were found to differ by the lifestyle indicators of cigarette smoking, alcohol consumption and achievement of ‘5 a day’. All of these differences persisted regardless of whether daily diet costs or costs per 10MJ were compared.

Similar differences were apparent when Kruskal-Wallis analyses compared diet costs including alcohol (Table 6.3): daily diet costs were found to differ by sex, employment, marital status, qualifications, income, cigarette smoking, alcohol consumption and achievement of ‘5 a day’. In addition, daily diet costs differed significantly between age groups. However, some of these contrasts were not statistically significant when costs per 10MJ were tested – namely, sex, age group and smoking.

Diet costs were not found to differ significantly between the categories of household size or BMI category.

Table 6.2 Median values and interquartile ranges for daily diet costs (£ day^{-1}) and costs adjusted to 10MJ for sample subgroups. Alcohol excluded (p values for Kruskal-Wallis ANOVA) (n=1014)

			Daily diet cost (£ d^{-1})			Energy-adjusted diet cost (£ $10MJ^{-1}$)		
Variable		**n**	**Median**	**IQR**	**p**	**Median**	**IQR**	**p**
Sex	Male	434	3.14	2.43, 4.02	<0.01	3.80	3.24, 4.47	<0.01
	Female	580	2.69	2.20, 3.30		4.28	3.67, 4.99	
Age group	19-29 years	145	2.74	2.09, 3.41	0.18*	3.78	3.34, 4.56	0.05*
	30-39 years	202	2.96	2.35, 3.71		4.07	3.41, 4.82	
	40-49 years	179	2.90	2.21, 3.74		4.12	3.48, 4.66	
	50-59 years	184	2.94	2.44, 3.75		4.17	3.62, 5.07	
	60-69 years	147	2.84	2.25, 3.77		4.32	3.75, 5.12	
	70 years and over	157	2.59	2.14, 3.20		3.90	3.27, 4.65	
Employment	Managerial & professional	421	3.10	2.52, 3.93	<0.01	4.27	3.69, 5.01	<0.01
	Intermediate, lower supervisory & small employers	302	2.90	2.27, 3.56		4.00	3.53, 4.92	
	Routine & semi-routine	250	2.52	2.01, 3.01		3.75	3.14, 4.37	
	Never worked & other	41	2.56	1.99, 3.28		3.87	3.17, 4.93	
Marital status	Single, never married	289	2.78	2.20, 3.70	<0.01	3.87	3.36, 4.69	0.02
	Married	467	2.96	2.36, 3.70		4.09	3.54, 4.87	
	Married but separated	30	2.96	2.15, 4.00		4.23	3.31, 5.50	
	Divorced	127	2.74	2.26, 3.40		4.29	3.47, 4.92	
	Widowed	101	2.47	2.02, 3.10		4.16	3.30, 4.76	
Qualifications (n=1006)	Degree or higher education	338	3.13	2.52, 3.99	<0.01	4.27	3.67, 5.07	<0.01
	GCA A-level or equivalent, foreign qualification	172	2.89	2.39, 3.65		4.05	3.52, 4.79	
	GCSEs/still in full-time education	245	2.81	2.24, 3.72		4.01	3.35, 4.94	
	No qualifications	251	2.48	1.99, 3.03		3.78	3.21, 4.45	

Table 6.2 (cont'd) Median values and interquartile ranges for daily diet costs (£ day^{-1}) and costs adjusted to 10MJ for sample subgroups.

Variable		n	Daily diet cost (£ d^{-1}) Median	IQR	p	Energy-adjusted diet cost (£ 10MJ^{-1}) Median	IQR	p
Equivalized household income (n=875)	Under £14,999	198	2.59	2.07, 3.11	<0.01 *	3.73	3.25, 4.43	<0.01*
	£15,000-£24,999	202	2.69	2.26, 3.33		3.97	3.37, 4.58	
	£25,000-£34,999	197	2.89	2.28, 3.65		4.00	3.49, 4.84	
	£35,000-£49,999	142	3.17	2.52, 4.15		4.19	3.67, 5.04	
	£50,00 and over	136	3.31	2.66, 4.19		4.55	3.80, 5.35	
Household size	1 person	268	2.77	2.19, 3.45	0.58*	4.11	3.43, 4.79	0.07*
	2 people	336	2.95	2.34, 3.74		4.15	3.56, 5.01	
	3 or 4 people	327	2.81	2.28, 3.56		3.96	3.38, 4.60	
	5 or more people	83	2.89	2.21, 3.61		3.91	3.39, 4.35	
BMI category	NA/Missing	76	2.66	2.09, 3.54	0.09†	4.25	3.30, 5.25	0.26 †
	Underweight (<18.5kg/m^2)	13	1.98	1.75, 2.47		3.41	2.80, 4.15	
	Normal weight (18.5 – 24.9kg/m^2)	318	2.89	2.36, 3.58		4.01	3.45, 4.83	
	Overweight (25.0 – 29.9kg/m^2)	350	2.93	2.34, 3.72		4.07	3.46, 4.78	
	Obese (30kg/m^2 and over)	257	2.78	2.18, 3.50		4.15	3.57, 4.76	
Smoking	Never smoked	541	2.92	2.31, 3.70	<0.01	4.12	3.54, 4.89	<0.01
	Ex-smoker	247	2.91	2.37, 3.67		4.11	3.55, 4.99	
	Current smoker	226	2.55	2.09, 3.22		3.82	3.20, 4.56	
Alcohol consumption	None	410	2.59	2.10, 3.23	<0.01*	3.90	3.39, 4.68	0.01*
	Low risk	425	2.93	2.39, 3.84		4.15	3.54, 4.95	
	Increasing risk	132	3.17	2.54, 3.71		4.13	3.50, 4.92	
	High risk	47	3.18	2.54, 4.46		3.89	3.18, 4.59	
Achieve '5 a day'	Yes	334	3.41	2.81, 4.21	<0.01	4.52	3.84, 5.40	<0.01
	No	680	2.60	2.10, 3.23		3.86	3.31, 4.57	

* test for trend on ordered categories; † test for trend, excluding NA/Missing and Underweight

Table 6.3 Median values and interquartile ranges for daily diet costs (£ day^{-1}) and costs adjusted to 10MJ for sample subgroups. Including alcohol (p values for Kruskal-Wallis ANOVA) (n=1014)

			Daily diet cost (£ d^{-1})			Energy-adjusted diet cost (£ $10MJ^{-1}$)		
Variable		**n**	**Median**	**IQR**	**p**	**Median**	**IQR**	**p**
Sex	Male	434	4.21	2.95, 5.75	<0.01	4.68	3.73, 6.07	0.73
	Female	580	3.11	2.36, 4.12		4.76	3.88, 5.96	
Age group	19-29 years	145	3.10	2.27, 4.60	0.03*	4.31	3.55, 5.42	0.83*
	30-39 years	202	3.75	2.71, 4.94		4.84	3.81, 5.94	
	40-49 years	179	3.86	2.70, 5.39		4.91	4.03, 6.46	
	50-59 years	184	3.75	2.90, 4.90		5.14	4.06, 6.45	
	60-69 years	147	3.48	2.56, 4.92		4.80	3.99, 6.37	
	70 years and over	157	2.94	2.31, 3.70		4.25	3.60, 5.36	
Employment	Managerial & professional	421	3.92	2.96, 5.37	<0.01	5.12	4.19, 6.39	<0.01
	Intermediate, lower supervisory & small employers	302	3.31	2.60, 4.63		4.57	3.78, 6.04	
	Routine & semi-routine	250	2.93	2.23, 4.06		4.28	3.55, 5.37	
	Never worked & other	41	2.90	1.99, 4.02		4.23	3.26, 5.79	
Marital status	Single, never married	289	3.42	2.39, 4.90	<0.01	4.56	3.71, 5.60	<0.01
	Married	467	3.60	2.80, 4.96		4.88	3.89, 6.18	
	Married but separated	30	3.48	2.17, 4.69		4.98	3.85, 6.48	
	Divorced	127	3.46	2.69, 4.64		4.86	4.10, 6.54	
	Widowed	101	2.80	2.15, 3.82		4.45	3.62, 5.64	
Qualifications	Degree or higher education	338	4.03	2.96, 5.48	<0.01	5.11	4.17, 6.39	<0.01
(n=1006)	GCA A-level or equivalent, foreign qualification	172	3.52	2.75, 4.84		4.87	3.88, 5.82	
	GCSEs/still in full-time education	245	3.53	2.60, 4.68		4.86	3.80, 6.34	
	No qualifications	251	2.83	2.14, 3.85		4.24	3.56, 5.08	

Table 6.3 (cont'd) Median values and interquartile ranges for daily diet costs (£ day^{-1}) and costs adjusted to 10MJ for sample subgroups, including alcohol

			Daily diet cost (£ d^{-1})			Energy-adjusted diet cost (£ 10MJ^{-1})		
Variable		**n**	**Median**	**IQR**	**p**	**Median**	**IQR**	**p**
Equivalized household income (n=875)	Under £14,999	198	2.88	2.24, 3.91	<0.01*	4.25	3.52, 5.34	<0.01*
	£15,000-£24,999	202	3.31	2.60, 4.28		4.53	3.79, 5.48	
	£25,000-£34,999	197	3.61	2.68, 4.84		4.73	3.87, 6.10	
	£35,000-£49,999	142	3.94	3.11, 5.35		5.12	4.32, 6.29	
	£50,00 and over	136	4.54	3.16, 6.14		5.85	4.55, 7.17	
Household size	1 person	268	3.33	2.47, 4.45	0.37*	4.65	3.84, 5.95	0.46*
	2 people	336	3.63	2.65, 5.10		4.85	3.97, 6.28	
	3 or 4 people	327	3.47	2.60, 4.73		4.69	3.72, 5.89	
	5 or more people	83	3.47	2.50, 5.17		4.60	3.72, 5.68	
BMI category	NA/Missing	76	3.29	2.19, 4.22	0.10†	4.64	3.55, 6.24	0.23†
	Underweight (<18.5kg/m^2)	13	2.37	1.94, 4.20		4.31	3.26 ,5.42	
	Normal weight (18.5 – 24.9kg/m^2)	318	3.43	2.67, 4.75		4.65	3.79, 5.79	
	Overweight (25.0 – 29.9kg/m^2)	350	3.63	2.67, 5.07		4.80	3.86, 6.07	
	Obese (30kg/m^2 and over)	257	3.29	2.49, 4.71		4.73	3.89, 6.07	
Smoking	Never smoked	541	3.47	2.60, 4.86	0.04	4.73	3.80, 5.99	0.13
	Ex-smoker	247	3.61	2.79, 4.84		4.89	3.93, 6.13	
	Current smoker	226	3.29	2.35, 4.68		4.65	3.78, 5.78	
Alcohol consumption	None	410	2.60	2.10, 3.23	<0.01	3.89	3.39, 4.65	<0.01
	Low risk	425	3.75	3.01, 4.72		5.00	4.17, 5.97	
	Increasing risk	132	5.80	4.87, 6.79		6.59	5.75, 7.73	
	High risk	47	9.28	7.21, 10.46		7.88	7.05, 9.39	
Achieve '5 a day'	Yes	334	4.21	3.11, 5.43	<0.01	5.23	4.35, 6.49	<0.01
	No	680	3.17	2.34, 4.40		4.49	3.70, 5.67	

* test for trend on ordered categories; † test for trend, excluding NA/Missing and Underweight

6.4.4 Regression analyses

Multivariable regression indicated that daily diet costs increased significantly as energy intake increased, after adjusting for the other variables (see Table 6.4): it can be seen that each additional 100kJ was associated with an additional 3 pence in diet cost (95% CI £0.03, £0.03). There was a significant increase – of 44 pence (95% CI £0.31, £0.56) – in daily diet costs for those who achieved '5 a day' compared to those who did not. There was also a significant overall effect of household income category on diet costs, with an increase of 15p associated with each progression up through the categories

Table 6.5 presents the results of the adjusted regression diet costs per 10MJ. This model revealed significant effects of household income and achieving '5 a day', as did the daily diet cost model. In this second model, moving into a higher income category was associated with an increase of 19p per 10MJ, and those who achieved '5 a day' had an energy-adjusted cost of 56p more than those who did not. In contrast to the first model, however, a significant effect was observed for sex, with females showing costs of 46p per 10MJ higher than males, and for BMI, with an additional 2p associated with each progression up through BMI categories.

After adjustment, no significant effects were apparent in either model for age group or cigarette-smoking status.

Table 6.4 Regression of sociodemographic and lifestyle variables on estimates of daily diet cost (n=814)

Variable	Coefficient (difference in diet cost, pence)	95% CI	Overall p value
Sex*	12.47	-0.24, 25.18	0.054
Age group	-0.42	-4.05, 3.20	0.818
Food energy (100kJ)	3.13	2.84, 3.43	<0.001
BMI category†	0.87	-0.21, 1.95	0.115
Cigarette smoking status‡			0.475
Current regular smoker	-9.02	-23.72, 5.69	
Ex-regular smoker	-3.77	-17.74, 10.19	
Achieve 5 a day	43.95	31.45, 56.46	<0.001
Equivalized household income§	15.21	11.01, 19.40	<0.001

* Reference category = males
† Underweight participants (BMI<18.5kg/m^2) excluded
‡ Compared with participants who have never regularly smoked (reference category)
§ Household income categories: under £14,999; £15,000-£24,999; £25,000-£34,999; £35,000-£49,999; £50,00 and over

Table 6.5 Regression of sociodemographic and lifestyle variables on estimates of diet cost per 10MJ (n=814)

Variable	Coefficient (difference in diet cost, pence)	95% CI	Overall p value
Sex*	45.58	28.72, 62.45	<0.001
Age group	1.43	-3.88, 6.75	0.596
BMI category†	1.77	0.17, 3.37	0.030
Cigarette smoking status‡			0.890
Current regular smoker	-5.13	-26.83, 16.57	
Ex-regular smoker	-2.67	-23.37, 18.02	
Achieve 5 a day	56.39	38.10, 74.67	<0.001
Equivalized household income	19.17	12.95, 25.39	<0.001

* Reference category = males

† Underweight participants (BMI<18.5kg/m²) excluded

‡ Compared with participants who have never regularly smoked (reference category)

§ Household income categories: under £14,999; £15,000-£24,999; £25,000-£34,999; £35,000-£49,999; £50,00 and over

6.4.5 Inflation index comparisons

Table 6.6 shows the diet cost estimates using each of the three different indices to correct for inflation. In comparing the food group-adjusted and FPI-adjusted costs (which both account for inflation), the median difference between the estimates was £0.03 (IQR -£0.04, £0.08). As a percentage of diet costs, differences between the estimates of the two indices ranged from -4% to 9% (median 0.6%, IQR -1%, 2%). Excluding alcohol, the median difference was £0.06 (IQR £0.02, £0.11) or 2% (range -4%, 9%; IQR 0.5%, 3%).

Table 6.6 Median estimated daily diet costs (£ day^{-1}) for the whole sample (n=1014), by method of adjustment

Diet cost estimation method	Including alcohol		Excluding alcohol	
	Median	IQR	Median	IQR
Prices unadjusted	3.47	2.57, 4.83	2.84	2.27, 3.64
Prices adjusted by food group	4.22	3.16, 5.78	3.48	2.81, 4.42
Prices adjusted by FPI	4.18	3.11, 5.79	3.42	2.75, 4.35

Table 6.7. presents the p values of Kruskal-Wallis ANOVA (or tests for trend where appropriate) comparing the estimated diet costs of sociodemographic and lifestyle categories when diet costs are estimated using costs adjusted for inflation

using 27 food group indices, or using a flat rate of inflation (FPI); Table 6.8 shows the estimated diet costs by sociodemographic and lifestyle categories. Estimated costs exclude costs of alcoholic beverages. Table 6.7 indicates that, in the majority of cases, results of comparison tests are similar and conclusions would be the same regardless of the inflation index applied. The one exception to this is the comparison by age group, where diet costs including alcohol were found to significantly differ when a flat rate of inflation (the FPI) was applied, but not where the food group indices were used. Where alcohol was excluded from diet costs, however, age groups were not found to significantly differ.

Table 6.7 P values from Kruskal-Wallis ANOVA for differences in daily diet costs between categories of sociodemographic variables

	Including alcohol		Excluding alcohol	
Variable	FPI-adjusted p value	Food-group index adjusted p value	FPI-adjusted p value	Food-group index adjusted p value
Sex	<0.01	<0.01	<0.01	<0.01
Age group*	0.03	0.06	0.19	0.40
Employment	<0.01	<0.01	<0.01	<0.01
Marital status	<0.01	<0.01	<0.01	<0.01
Qualifications	<0.01	<0.01	<0.01	<0.01
Equivalized income*	<0.01	<0.01	<0.01	<0.01
Household size*	0.36	0.41	0.56	0.62
BMI classification*	0.45	0.46	0.77	0.72
Smoking status	0.04	0.03	<0.01	<0.01
Alcohol	<0.01	<0.01	<0.01	<0.01
Achieve ‘5 a day’	<0.01	<0.01	<0.01	<0.01

*test for trend on ordered categories

Table 6.8 Median daily diet costs (£ day^{-1}) and interquartile ranges (IQR) by sociodemographic and lifestyle variables, excluding alcohol (n=1014)

			Daily diet cost (£ d^{-1}), FPI-adjusted		Daily diet cost (£ d^{-1}), food group index-adjusted	
Variable		**n**	**Median**	**IQR**	**Median**	**IQR**
Sex	Male	434	3.81	2.91, 4.80	3.86	3.00, 4.86
	Female	580	3.26	2.64, 3.93	3.31	2.70, 4.03
Age group	19-29 years	145	3.30	2.53, 4.15	3.32	2.52, 4.17
	30-39 years	202	3.53	2.86, 4.56	3.60	2.90, 4.55
	40-49 years	179	3.51	2.62, 4.43	3.57	2.76, 4.48
	50-59 years	184	3.53	2.91, 4.59	3.63	3.05, 4.64
	60-69 years	147	3.43	2.76, 4.52	3.53	2.84, 4.57
	70 years and over	157	3.16	2.54, 3.81	3.25	2.59, 3.89
Employment	Managerial & professional	421	3.73	3.01, 4.72	3.80	3.07, 4.79
	Intermediate, lower supervisory & small employers	302	3.46	2.76, 4.29	3.53	2.81, 4.38
	Routine & semi-routine	250	3.03	2.47, 3.68	3.04	2.49, 3.70
	Never worked & other	41	3.05	2.42, 3.88	3.05	2.47, 3.99
Marital status	Single, never married	289	3.34	2.62, 4.43	3.37	2.66, 4.50
	Married	467	3.57	2.87, 4.43	3.65	2.93, 4.57
	Married but separated	30	3.51	2.54, 4.92	3.59	2.64, 4.92
	Divorced	127	3.32	2.73, 4.15	3.34	2.76, 4.21
	Widowed	101	3.04	2.43, 3.76	3.08	2.43, 3.84
Qualifications (n=1006)	Degree or higher education	338	3.76	3.08, 4.78	3.84	3.10, 4.92
	GCA A-level or equivalent, foreign qualification	172	3.51	2.90, 4.33	3.61	2.94, 4.42
	GCSEs/still in full-time education	245	3.36	2.72, 4.53	3.39	2.74, 4.54
	No qualifications	251	3.02	2.38, 3.71	3.03	2.43, 3.72

Table 6.8 (cont'd) Median daily diet costs (£ day^{-1}) and interquartile ranges (IQR) by sociodemographic and lifestyle variables, excluding alcohol

			Daily diet cost (£ d^{-1}), FPI-adjusted		Daily diet cost (£ d^{-1}), food group index-adjusted	
Variable		**n**	**Median**	**IQR**	**Median**	**IQR**
Equivalized household income (n=875)	Under £14,999	198	3.08	2.50, 3.73	3.12	2.53, 3.80
	£15,000-£24,999	202	3.24	2.77, 4.01	3.33	2.82, 4.03
	£25,000-£34,999	197	3.47	2.76, 4.32	3.49	2.83, 4.45
	£35,000-£49,999	142	3.81	3.06, 4.97	3.87	3.15, 5.08
	£50,00 and over	136	3.98	3.21, 5.01	4.08	3.31, 5.13
Household size	1 person	268	3.30	2.65, 4.19	3.36	2.72, 4.29
	2 people	336	3.55	2.81, 4.59	3.60	2.88, 4.61
	3 or 4 people	327	3.42	2.75, 4.27	3.48	2.82, 4.35
	5 or more people	83	3.49	2.66, 4.45	3.41	2.79, 4.44
BMI category	NA/Missing	76	3.18	2.50, 4.27	3.26	2.53, 4.34
	Underweight (<18.5kg/m^2)	13	2.44	2.12, 3.04	2.43	2.13, 3.24
	Normal weight (18.5 – 24.9kg/m^2)	318	3.47	2.82, 4.27	3.56	2.91, 4.35
	Overweight (25.0 – 29.9kg/m^2)	350	3.54	2.84, 4.53	3.63	2.89, 4.54
	Obese (30kg/m^2 and over)	257	3.38	2.61, 4.27	3.43	2.67, 4.35
Smoking	Never smoked	541	3.51	2.79, 4.45	3.61	2.87, 4.52
	Ex-smoker	247	3.52	2.85, 4.38	3.58	2.90, 4.52
	Current smoker	226	3.06	2.50, 3.88	3.13	2.52, 3.95
Alcohol consumption	None	410	3.13	2.54, 3.86	3.19	2.59, 3.94
	Low risk	425	3.52	2.87, 4.63	3.63	2.93, 4.72
	Increasing risk	132	3.78	3.07, 4.50	3.81	3.06, 4.55
	High risk	47	3.88	3.01, 5.28	3.97	3.05, 5.16
Achieve '5 a day'	Yes	334	4.14	3.40, 5.10	4.26	3.45, 5.23
	No	680	3.14	2.53, 3.87	3.19	2.59, 3.94

6.5 Discussion

This chapter set out to estimate and describe the diet costs of NDNS adults, then explore the patterns in these costs according to a number of sociodemographic and lifestyle characteristics.

This is the first time a monetary value has been applied to individuals' diets in the NDNS. These costs are estimates of the inherent monetary value of diets, as opposed to actual expenditure. Despite this difference, the daily estimated diet costs of this sample are similar to national expenditure estimates, when excluding costs from alcohol: the inflation-adjusted median estimates in the NDNS were £3.42 (FPI) and £3.48 (food group indices), compared to £3.50 per person per day reported in Family Food 2010 (£24.50 per week) (Defra, 2012). When costs from alcohol are included, the NDNS daily estimates, at £4.18 (FPI) or £4.22 (food group indices), are slightly higher than expenditure data suggest, at £27.57 per week, or £3.94 per day. It could be conjectured that this discrepancy could be due to cheaper sources of alcohol being purchased (whereas the DANTE cost database assumes a mean cost), or it could be due to measurement error associated with dietary consumption data (although over-reporting of alcohol consumption does not typically feature in dietary surveys).

The estimated monetary value of diets was closely correlated with energy intake in the NDNS, indicating that those with higher energy requirements face higher diet costs. Due to this relationship, adjusting diet costs to 10MJ should allow a more fair comparison between groups of individuals who are likely to have different energy requirements (for example, between men and women).

6.5.1 Diet costs of sociodemographic groups

Univariate comparisons highlighted some interesting differences between subgroups in this sample, even after adjusting diet costs to 10MJ. Men were estimated to have higher daily diet costs than women in this sample, but lower diet costs per 10MJ. This is a pattern similarly reported in a French (Maillot et al., 2007b) and a US (Monsivais and Drewnowski, 2009) sample, although not apparent in all studies of this type (Rehm et al., 2011). The pattern likely reflects the higher energy intakes that tend to be observed in males, with diet costs and energy intakes being strongly correlated in this sample. After adjusting for energy, males exhibited lower costs, probably as a result of having more energy-dense diets, a sex difference similarly reported in US (Ledikwe et al., 2005) and Mediterranean (Marti-Henneberg et al., 1999) samples. In the multivariable analysis, however, sex no longer had a significant effect on daily diet costs, although a difference was still apparent when diet cost per 10MJ was the dependent variable. This suggests that the difference between males and females reflects sex differences other than energy intake – for example, in fruit and vegetable consumption.

Those in managerial and professional positions showed higher diet costs than other occupations; as did those with higher compared to lower educational qualifications. Although other studies have not investigated occupation group differences in diet costs, differences in education have previously been described in other countries (Monsivais et al., 2010, Rehm et al., 2011, Monsivais and Drewnowski, 2009), as in this study. The influence of education and occupation on diet costs could be indirect, through probable links between these socioeconomic variables and income, which could determine food budgets. Alternatively, diet selection may be influenced by education independently, with occupation and income being consequential rather than causal. Although both education and occupation are frequently used markers of socioeconomic status, education appears to be more strongly associated with dietary habits (De Irala-Estevez et al., 2000, Giskes et al., 2010), perhaps reinforcing the latter interpretation. On the other hand, one study examining fruit and vegetable consumption by strata of education reported increasing consumption as incomes increased within each stratum (Lallukka et al., 2010). The authors additionally found that participants with the highest reported education level but low incomes did not consume more fruit and vegetables than the lowest educated.

Both diet cost variables were found to increase monotonically with income categories in this sample. The effect of income on diet costs was still significant after adjusting for other variables in the regression models. This is in keeping with Engel's observation that expenditure on food will increase as income increases (Zimmerman, 1932; see Chapter 1). The increase in cost per 10MJ with rising income categories is particularly interesting: because the food price database uses mean values and does not distinguish between different types of the same product, it implies that the additional costs incurred by the higher income categories are a result of the selection of different foods, rather than merely 'trading up' to higher quality, more expensive versions of the same items. In reality, higher income participants may also have 'traded up' in addition to choosing different foods than did lower income subjects, which would augment the observed diet cost differences. Similar income effects have been observed in some (Monsivais and Drewnowski, 2009, Rehm et al., 2011), though not all (Waterlander et al., 2010), comparable studies. (The authors of the latter study suggest the lack of significance may be attributed to a lack of statistical power in the sample, or inappropriate income measurement.)

The findings suggest that those who are or have been married tend to have higher diet costs, whilst the widowed show the lowest costs. One interpretation is that this is due to an over-representation of the elderly amongst the widowed, who may be more likely to be on lower incomes. This is the first time marital status has been included in a study of dietary costs, although Murakami et al (2007) reported significant differences according to 'living status', with or without family. In contrast, there were no significant differences in diet costs by household size in the NDNS sample.

6.5.2 Diet costs and lifestyle variables

Interestingly, diet costs (either per day, or per 10MJ) were not found to differ between BMI categories, yet a significant positive association (p=0.03) was apparent between BMI category and diet cost per 10MJ when adjusting for other variables (Table 6.5). In light of the food price-obesity hypothesis described in Chapter 1, this finding is particularly interesting. The relationship between diet costs and BMI is investigated more thoroughly in Chapter 7.

Differences in diet costs per 10MJ were also evident between smokers and non-smokers in this study, with current regular cigarette smokers showing the lowest diet costs. It could be speculated from this relationship that the monetary costs of smoking impinge upon the food budget. Conversely, the findings may reflect a clustering of behaviours (smoking and poor diet). The latter interpretation is supported by the observation that cigarette smoking status was not found to be significantly related to daily diet costs or diet costs per 10MJ after adjusting for other variables. In other populations, comparisons between smokers and non-smokers have resulted in mixed findings (Murakami et al., 2007, Lopez et al., 2009); although the same studies found similar trends for alcohol consumption.

In this sample, the observation of increasing daily diet costs with increasing alcohol consumption could also be attributed to the concomitant increasing intakes of food energy (not presented). However, those who consumed no alcohol exhibited a similar median cost to the highest alcohol consumers when adjusted to 10MJ, suggesting that the observed differences are not solely due to the energy differences between the consumption groups, and again supporting a behaviour-cluster interpretation. A previous study (Breslow et al., 2006) has identified a significant pattern of lower diet quality with increasing alcohol consumption, but only a few have reported increasing food energy intakes (Kesse et al., 2001), and one validation study suggests a tendency to over-report food intake amongst higher risk alcohol consumers (Zhang et al., 2001). On the other hand, it is also possible that drinking behaviours are linked to disposable incomes and thereby affect food budgets.

Diets containing five portions (400g) or more of fruit and vegetables per day were found to be of higher monetary value than those that featured fewer. This supports the findings of previous research suggesting that people who score more favourably on healthy diet indicators (Schroder et al., 2006, Maillot et al., 2007b, Cade et al., 1999, Ryden and Hagfors, 2011), as well as those who consume more fruit and vegetables in particular (Rehm et al., 2011), tend to spend more money on food or consume higher value diets. In addition, the findings presented here go further than many of the other studies in showing that the relationship between fruit and vegetable consumption and diet costs remains even after adjusting for other economic and demographic factors. Whilst some studies report that a diet adhering to national guidelines is theoretically achievable on low incomes (for example, Cassady et al., 2007 in the US), others have found that modelling diets to be both palatable and nutritionally adequate does increase costs (Darmon et al., 2006). One study in Ireland predicted that the cost of adhering to proposed guidelines, whilst achievable in theory, would take up to 100% of the income from welfare for an adolescent male (Flynn et al., 2011).

The current study did not investigate costs according to wider measures of diet quality nor adherence to guidelines other than fruit and vegetables. Nevertheless, the results imply that the better quality diets, as signified by the consumption of fruit and vegetables, were of higher intrinsic monetary worth. It cannot be determined from this study design whether diet costs were influential in participants' food selection; nevertheless, the relationships evident between diet costs and socioeconomic markers are interesting, with potential policy implications.

6.5.3 Inflation indices

Whilst several investigators have matched food price databases to nutrition survey data with a different year of data collection, there does not appear to have been an investigation into the possible influence of inflation. The results above present for the first time a comparison between inflation adjustment methods.

Reassuringly, diet costs estimated using a flat rate of inflation (the FPI) appeared to be similar to those adjusted by the different food group inflation indices.

For the whole sample, the median difference between the two diet costs was just £0.02 a day excluding alcohol. However, a key aim of the inflation investigations was to determine if the different adjustments would have consequences in terms of interpreting between-group differences in estimated diet costs. Looking at different subgroups (Table 6.8), the inflated diet costs are on the whole similar whether the FPI or food group indices were used (with differences ranging up to +£0.12, or +3% of diet cost). It is perhaps interesting, though, that almost all of the subgroup estimates were higher when food group indices had been applied, as compared to when the FPI was applied. Although differences observed in this sample were modest, it is possible that greater time lags in the collection of price data and dietary data in other studies (compared to the 4-5 year difference between the NDNS and the DANTE food cost database) would result in larger differences, potentially leading to an increasing bias towards underestimation of diet costs if the FPI is used.

As a cautionary note, it is possible that food group indices imply a spurious level of accuracy. Whilst it is intuitive to account for differential rates of inflation between food groups, it should be noted that there still remains a large degree of assumption-making in the compilation of these indices – for example, the assumption that the reference food items used to calculate each index, along with their weights, give an accurate indication of the whole food group's price changes over time. Full details of the assumptions inherent in the methods are described in the Office of National Statistics supporting documents (ONS, 2011).

In terms of univariate analyses, few differences in p values were evident when food group indices or the FPI were used. This, coupled with only minor differences in each index's effect on the cost estimates of each category, implies that there is little to distinguish the two when used in this sample. The exception was the effect of adjustment on age group comparisons. The different p values shown in Table 6.7 suggest that the age groups have been unevenly affected by the price changes of certain food groups. From these results, it is difficult to identify which food groups may be culpable and how, but this is an interesting area for future investigation.

The implication of these explorations is that researchers need to consider carefully the different approaches to handling data collected in differing years. Applying a flat rate of inflation could, on the face of it, offer a simple route to estimating diet costs that appear meaningful to another year's experience of pricing. However, ignoring the relative influences of different food groups on diet cost inflation risks losing an important level of detail. If possible, a comparison of different approaches may be advisable, as was performed here. In the end, pragmatic considerations may influence the approach adopted. In the NDNS, different years of data collection are combined (2008-9 and 2009-10; see Chapter 3). In order to achieve an adequate sample size for later analyses (Chapters 7 and 8), it is advantageous to combine these years of data collection. This would not be feasible if different inflation indices had been applied to each year, making the years incomparable. For this reason, as well as the fact that these investigations revealed little difference in the effect of indices in this sample, unadjusted costs were adopted for the analyses of the ensuing chapters.

6.5.4 Limitations

As demonstrated above, assigning costs to dietary data using a food price database is a potentially insightful methodology. It is not without limitations, however. A full discussion of these limitations can be found in Chapter 5, but consideration needs to be given to some key points that are relevant for the interpretation of the results in this chapter and this is offered below.

Firstly, it should be noted that these diet cost estimates will inevitably echo any measurement error associated with the dietary assessment tool from which they are extrapolated. Under-reporting of food consumption, for example, will result in an underestimation of diet cost. Where under-reporting may be more prevalent amongst certain subgroups, as it has been suggested to be for those classified as obese for example (Rennie et al., 2007), the resulting bias could influence the results of subgroup comparisons. In this sample, energy intake was found to vary significantly between BMI categories, with the lowest energy intake reported in the obese. This perhaps suggests that such bias exists within the sample. The relationship between diet costs and BMI is explored more thoroughly in Chapter 7. Chapter 3 contains a more in-depth discussion of limitations in dietary assessment.

This method of costing has limits in establishing the role of diet costs in food selection. Firstly because the results imply that the diets of certain subgroups are worth more, not necessarily that these populations spend more on their diets. The value of a person’s diet may not reflect the prices they encountered in purchasing their food: although 74% of this sample indicated that the majority of their household grocery purchases were made in large supermarkets (see Chapter 3), prices are known to vary by area and according to retailer type (Cummins and Macintyre, 2002). In addition, the food cost database does not account for restaurant or takeaway meals, which are likely to be higher than those estimated, and thought to account for 31% of all food and drink purchases in England (Defra, 2009). Food away from home (FAFH) has been demonstrated in the UK to be roughly three times that of equivalent foods eaten in the home (Wrieden and Barton, 2011). It can be assumed that accounting for these costs would result in higher estimated diet costs for those who consumed FAFH during the data collection, which are not reflected in the estimates presented above. In addition, the DANTE cost database does not identify free, shared or foraged food. Secondly, as a cross-sectional study, it is impossible to gauge whether diets of a lower monetary value are selected as a result of budgetary considerations, or whether the value of a diet merely reflects a preference for cheaper foods driven by other factors.

6.5.5 Strengths

These findings add to the literature on social inequalities in diet and health. Many of the patterns revealed here appear to substantiate speculated differences in diet costs, which should impart confidence to the costing method.

The existence of diet cost differences between certain groups of people could have implications in the consideration of proposed fiscal interventions to combat public health issues such as obesity (as suggested in one recent report, (Sustain, 2013)), that may differentially affect socioeconomic groups. Modelling studies have indicated that this would be the case, and taxation measures are likely to be economically regressive (Nnoaham et al., 2009). This is concerning, given that the differences between sociodemographic groups observed here are likely to be conservative (Section 6.5.1).

Individual-level diet costs allow the investigation of diet costs in relation to health outcomes. Chapters 7 and 8 present such investigations, where the associations between diet costs and energy density, and diet costs and BMI are examined.

6.6 Conclusion

This study is the first attempt to quantify individual diet costs for a representative UK sample. Diets of adults in the NDNS were matched to a food cost database to derive an estimated daily diet cost and a cost per 10MJ for each participant. The findings suggest that certain subgroups in the UK consume diets of lower monetary value. Observed differences were, for the most part, in the directions anticipated. Costing diets in this manner is constrained by the measurement error associated with dietary assessment. Nevertheless, the derivation of these cost variables paves the way for the investigations into the links between diet costs, diet quality and health which are the subject of the following chapters.

What was known previously:

- There is a need to measure diet cost at the individual level, in order to link food prices to health outcomes.
- Previous studies have used national food price databases to apply a monetary value to the diets of dietary surveys, but to date there have been no such studies in a representative UK sample.
- In other populations, people who score more favourably on healthy diet indicators and those who consume more fruit and vegetables tend to spend more money on food or consume higher value diets.

What this chapter adds:

- This is the first time monetary costs have been applied to the diets of NDNS adults.
- Many of the patterns revealed substantiate speculated differences in diet costs, which should impart confidence to the costing method.
- Diet costs were not found to differ significantly between the categories of household size or BMI category.
- Better quality diets, as signified by the consumption of fruit and vegetables, were of higher intrinsic monetary worth, even after adjusting for other economic and demographic factors.
- Income and fruit and vegetable intake appear to be key drivers of both daily diet costs and costs per 10MJ.
- On the whole, there was little difference in using a flat rate of inflation compared to the food group-adjusted indices, although comparisons suggests that age groups were unevenly affected by the price changes of certain food groups.
- The existence of diet cost differences between certain groups of people could have implications in the consideration of proposed fiscal interventions that may differentially affect socioeconomic groups.

Chapter 7 Diet costs, diet and BMI in the NDNS

7.1 Summary

If food prices influence dietary intake and energy balance, it may be the case that the inherent monetary value of diets is associated with dietary energy density (DED) or the body weight of people consuming those diets. The previous chapter (Chapter 6) presented diet costs for a nationally representative dietary survey, the National Diet and Nutrition Survey (NDNS). This chapter investigates how these estimated diet costs relate to DED and body mass in the sample.

The relationship between diet costs and dietary energy density (excluding non-milk beverages) was assessed using quintile comparisons, and with multivariable regression using the residuals method. Multivariable regression tested for a linear association between diet costs per 10MJ and BMI; polynomial models tested for non-linear relationships. Finally, logistic regression was used to gauge the effect of diet costs per 10MJ on the proportion of the sample overweight and obese.

The results indicated a strong negative association between the monetary cost and the energy density of diets. On the other hand, the data did not support an association of diet costs with BMI or classifications of overweight and obese. The possibility of a non-linear relationship between diet costs and BMI was also rejected. Interestingly, energy intake increased with increasing energy density, suggesting that an over-consumption of calories with increasing energy density is credible. The lack of association between diet costs and body mass may be due to the study design and potential self-reporting bias.

Whether the approach taken here is capable of implicating monetary factors in obesity remains to be seen. More prospective investigations would be ideal, given the protracted nature of obesity aetiology. In the meantime, there is still scope to explore this emerging field of study using cross-sectional data. The following chapter (Chapter 8) explores a new approach to the research question, in which diet costs are characterized in terms of the constituent food groups.

7.2 Introduction

The primary aim of this thesis is to determine the role of micro-economic factors in excess energy intake. A common theory in the recent literature is that rising obesity rates may be attributed to trends in food prices (the 'food price-obesity hypothesis' – see Chapter 1). Research into this hypothesis, however, is at an early stage, with few studies able to confirm or refute such a link (the results of a comprehensive literature search on the topic are presented in Chapter 2).

Diet costs have been linked, positively and significantly, to a variety of measures of dietary quality, such as nutrient density (Monsivais et al., 2010), indices of healthfulness (Cade et al., 1999, Bernstein et al., 2010) and dietary patterns (Ryden et al., 2008). However, fewer studies have attempted to address the outcome of weight or BMI, and none have done so in the UK (see Chapter 2).

From the literature review (Table 2.7), it can be seen that there have been just three studies of working-age adults published which investigated weight in relation to diet costs as estimated from dietary data. The findings published from these studies in Japan and Spain are mixed. Murakami et al (2007), for example, found a small but significant negative relationship between diet costs and BMI in a sample of Japanese female students; the same group, however, failed to repeat this finding in a subsequent study using laboratory-measured weight and height rather than self-reports (Murakami et al., 2008b). In a prospective Spanish cohort study, Lopez et al (2009) found a significant increase in BMI at follow-up with increasing quintiles of diet cost per 1000kcal at baseline. Whilst the odds ratio for a total weight gain of 3kg or more during the study was not significant (once adjusted for confounders), the highest quintile of diet cost was significantly associated with a 20% increase in the odds of gaining an average of at least 0.6kg per year.

BMI may give an indication of positive energy balance, but the protracted nature of weight gain makes it difficult to investigate putative links using cross-sectional data. An alternative approach is to examine aspects of the diet which may give an indication of excess energy intake. Increasing energy density, for example, has been suggested to encourage excess energy consumption, and has been linked to adiposity (see Chapter 1). A broader range of literature has been published relating dietary costs to energy density compared to the literature relating costs to BMI. All of these studies reported a strong negative relationship between diet costs and energy density, in France (Darmon et al., 2004, Maillot et al., 2007b), the Netherlands (Waterlander et al., 2010), the USA (Monsivais and Drewnowski, 2009, Townsend et al., 2009), Scotland (Wrieden and Barton, 2011) and Sweden (Ryden and Hagfors, 2011). In all of these studies, the relationship held regardless of the energy density calculation method used (including all beverages except water, excluding all beverages, or excluding non-calorific beverages), the method of dietary data collection (recall, diary, FFQ or from national expenditure data), and the analytical method employed (quintile comparisons, correlation or adjusted regression).

This chapter presents analyses which examine the relationships between diet costs and the outcomes of energy density and BMI, using the estimated dietary costs of the NDNS as a representative UK sample. The estimated costs of this sample have already been presented in Chapter 6. The following objectives will be addressed in this chapter:

1. To determine whether an association exists between diet costs and BMI or overweight amongst NDNS adults; and
2. To establish whether an association exists between diet costs and dietary energy density amongst NDNS adults.

The three key relationships under investigation are: daily diet costs and dietary energy density; costs per 10MJ and BMI; and costs per 10MJ and overweight and obesity. The selection of which diet cost variable to involve in each relationship was based upon discussion in previous publications: whilst costs adjusted to 10MJ improve comparability across populations with differing energy requirements, other investigators have cautioned against linking energy-adjusted costs with an outcome also derived using energy values – such as energy density. This is due to the resultant mathematical coupling, in which the same variable appears in the numerator of one variable and the denominator of the other (Lipsky, 2009). In this case, kJ is the numerator in the energy density calculation and the denominator in energy-adjusted diet cost. Observed relationships between two such variables could then reflect their algebraic relationship, as opposed to the hypothesized causal association. However, it is still necessary to control for energy intake, given its close association with diet costs (demonstrated in Chapter 6).

Adjustment for energy intake is a challenge that is not new to nutritional research, and various approaches have been considered in depth (Willett, 1998). One proposed alternative to a straightforward nutrient density approach is the ‘residuals method’, in which the residual values of a model, with energy intake as the independent variable and the nutrient in question as the dependent variable, are used to represent diet costs in the final model. Although not a nutrient, it is suspected that diet cost is subject to similar considerations with respect to total energy intake. With this in mind, the residuals method described by Willett will be adopted for the regression model which features dietary energy density as the outcome. This approach has similarly been taken in recent research in the field (for example, Aggarwal et al., 2011, Maillot et al., 2007b).

In contrast, the variables of energy-adjusted diet cost (£ per 10MJ) and BMI do not share this problem of mathematical coupling. In this instance, using energy-adjusted values will assist in the interpretability of the model coefficients. This approach is described by Willett as the ‘multivariate nutrient density method’ (Willett, 1998), in which a nutrient density variable is entered into a model alongside total energy intake. This method will be adopted in this chapter for the regression models investigating the outcomes of BMI or overweight and obesity.

The relationships identified above have never been formally investigated in a national sample in the UK before. The findings of the following analyses are expected to make an important contribution to the evidence base.

7.3 Methods

7.3.1 Sample

The analyses in this chapter, like the preceding chapters, are based upon data from the NDNS. The survey design, sample recruitment and characteristics are described in Chapter 3. Only adults with complete diary data (four days) and valid anthropometric measurements were included in the analytical sample. In addition, the decision was taken to exclude those participants with a BMI of less than 18.5kg/m^2 (n=13), as it was assumed these participants were (or had been) in negative energy balance resulting in underweight, and negative energy balance was felt to be subject to influences different to that of positive energy balance and beyond the scope of this thesis. These exclusions resulted in a sample of 925, from a possible 1031, adults.

7.3.2 Estimation of dietary costs

The exposure variables (dietary costs) were estimated by linking the NDNS data with the DANTE food cost database using Microsoft Access. This database is described in Chapter 5, whilst the method for linking it to the NDNS data is outlined in Chapter 6. Costs were estimated as a daily mean for each participant, as well as per 10MJ of energy intake. Descriptive results for both estimations can be found in Chapter 6.

The diet costs used below exclude costs and energy from alcohol. This is because alcohol, as a relatively expensive commodity, has a skewing effect on the diet costs of those who consume it, with the potential to skew results. Furthermore, alcohol may not be considered part of the food budget by individuals, and is therefore separate to the hypothesised causal relationships under investigation in this chapter.

7.3.3 Calculation of energy density

Dietary energy density was a newly created variable for this sample, derived by dividing the total energy intake (kJ) by total mass of food consumed (g). The methods of calculation can be found in Chapter 4, along with summary statistics. Energy density is expressed as kJ/g.

7.3.4 Analytical methods

This chapter aims to establish two key relationships: between diet costs and energy density, and between diet costs and BMI. In addition to considering BMI as a continuous variable, it is useful to investigate proportions classified as overweight or obese due to the clinical relevance of these classifications. There are therefore three relationships under investigation. The exposures and outcomes for each of these are summarised in Table 7.5. The methods employed to assess each relationship are detailed in the sections below.

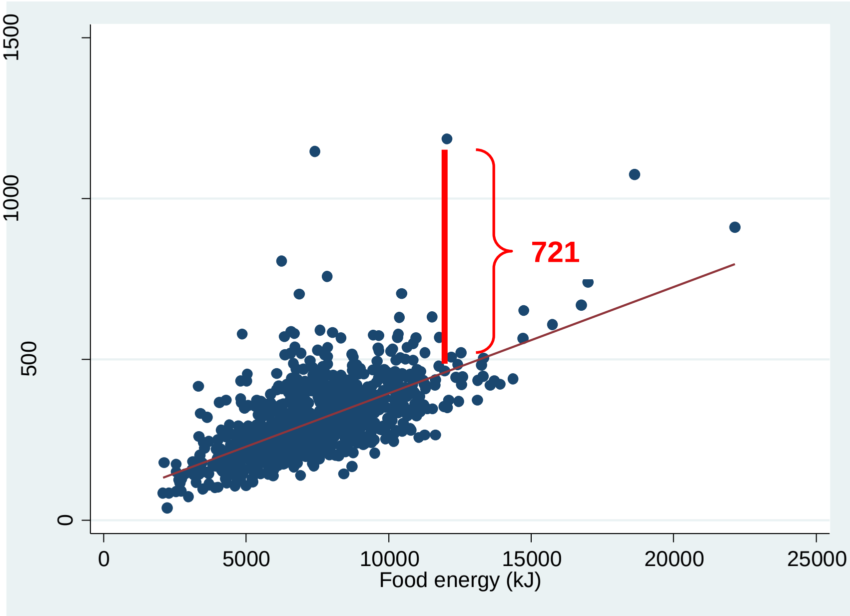
After derivation of all relevant variables, the data were analysed in Stata 12 (StataCorp, 2011). Descriptive analyses for the key variables have been performed in previous chapters (Chapters 3, 4 and 6), but means and SDs (or medians and IQR) are reiterated in the Results section below.

7.3.4.1 Diet costs & energy density

The sample was split into quintiles of dietary energy density to enable an initial exploration of patterns of daily diet costs, energy-adjusted diet costs, energy intake and BMI according to dietary energy density. Means and/or medians for each quintile are presented. In addition, adjusted means were estimated for each quintile, after adjusting for age and sex (and also energy intake when estimating mean daily diet cost). The 95% CI for the adjusted means are included.

To explore the association between diet costs and energy density, the residuals method was adopted (see Section 7.2). To enable this, a new vector was added to the Stata data set, containing the residual values from the regression of daily diet cost excluding alcohol (p d^{-1}) on daily food energy intake (kJ). These residuals provide a measure of diet cost that is uncorrelated with energy intake. Figure 7.1 illustrates how the residual value is calculated for an example individual. To aid in the interpretation of the regression coefficients, studentised residuals were generated in Stata, arrived at by dividing each residual by an estimate of its standard deviation. These studentised residuals represent the exposure variable of diet cost.

Figure 7.1 Scatter plot of daily diet costs (pence per day) against food energy (kJ), showing line of best fit and example residual value (n=1014)



On satisfying the assumptions for regression (homoscedasticity and normally distributed residuals), a linear multivariable regression was run with the studentised residuals of diet cost as the primary predictor variable and dietary energy density (kJ/g) as the dependent variable. Covariates to include in the model were selected following a priori consideration of confounding in the relationship. This is detailed in Section 7.3.4.5 below.

In addition, a sensitivity analysis was performed to establish the possibility of undue influence by the older participants in the sample. This was performed by repeating the adjusted model excluding participants aged 70 years and over (n=157). This subgroup was shown to have the lowest diet costs (see Chapter 6) and it has been suggested that older adults may be subject to different factors in dietary selection (Gariballa and Sinclair, 2005).

Regression results are presented in the form of coefficients and 95% confidence intervals (CI), for both unadjusted and adjusted analyses.

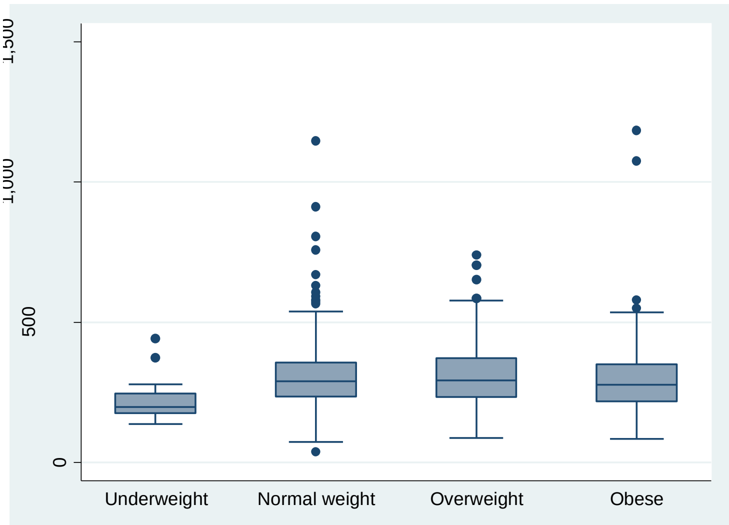
7.3.4.2 Diet costs & BMI (continuous)

The relationship between diet costs per 10MJ and continuous BMI was estimated using multivariable linear regression. Whilst the residual approach to energy adjustment was employed in the investigation of dietary energy density (see Section 7.3.4.1 above), for the outcome of BMI, which is not mathematically coupled to energy-adjusted costs, Willett's 'multivariate nutrient density method' (Willett, 1998) was selected. This was chosen as the best option due to the improved interpretability of regression coefficients, and given the potentially erroneous assumption of the residual approach that there is an equivalent effect at all levels of energy intake. The adjusted model also controlled for age, sex, employment and smoking status; the process of covariate selection is detailed in Section 7.3.4.5 below.

Sensitivity analyses were carried out to establish the possibility of undue influence by certain participants. These were performed by repeating the adjusted model excluding the participants identified each time. Firstly, participants who indicated consuming an atypical quantity of food during the data collection period were excluded. Secondly, those who reported being on a special diet (not counting vegetarianism or veganism) were excluded. Both of these exclusions were made because of the underlying assumption that the dietary data provide an indication of usual diet, whereas those on a special diet or consuming an atypical amount will not, by definition, be recording their usual diet. Finally, an analysis was run excluding those aged 70 years and over, for the same reasons as identified for the dietary energy density model (Section 7.3.4.1).

Given the observed patterns in Chapter 6 (see also Figure 7.2), it was suspected that the exposure and outcome variables may not in fact exhibit a linear relationship. Nonetheless, there may be a relationship between the variables, albeit a non-linear one. If this was the case, the standard linear regression model would be unable to detect the relationship.

Figure 7.2 Box plot displaying the means and distributions of daily diet cost excluding alcohol for each BMI category (n=938)



Polynomial regression is a method for detecting non-linear relationships. The technique involves incorporating higher order effects alongside the main effects in the model: in other words, diet costs would be entered as a predictor variable, along with a variable containing squared values of diet costs, a variable of cubed values and so on (Equation 7.1 shows an example of a model including squared and cubic values).

$$\text{BMI} = \beta_0 + \beta_0 \,\text{diet costs} + \beta_1 \text{ diet costs}^2 + \beta_2 \text{ diet costs}^3 + e$$

Equation 7.1

The regression equation above, if appropriate for the data, would fit a cubic regression line, as shown in Figure 7.3. Figure 7.4 illustrates the fitted curve of a quadratic relationship.

Figure 7.3 Illustration of a cubic regression line

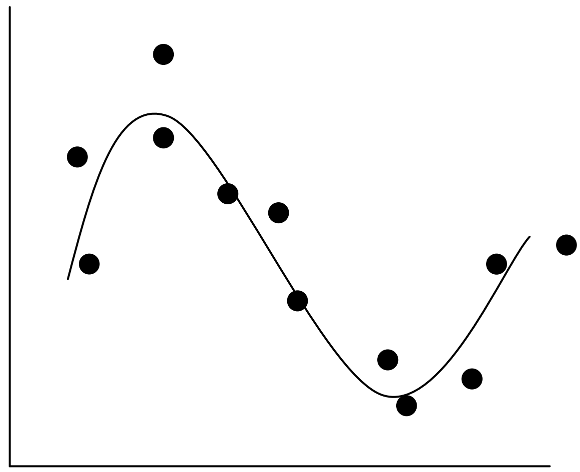
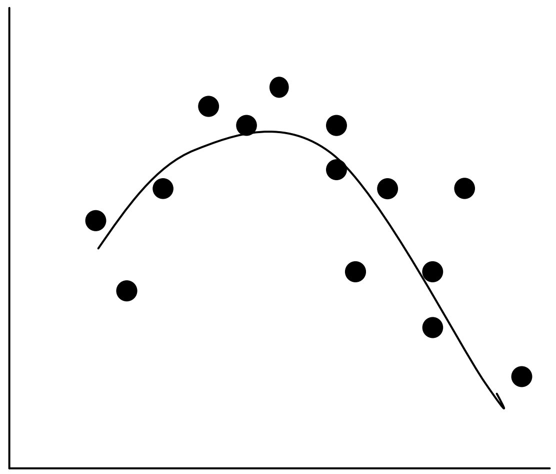


Figure 7.4 Illustration of a quadratic regression line



Alternatives to the straightforward polynomial model described above include the use of fractional polynomials (for instance β_1 diet costs$^{1/2}$) or employing non-parametric (or local influence) models that incorporate splines or lowess smoothing. The danger of non-parametric methods, which use values 'local' to the regression line to influence the shape of the line, is that there is a potential for over-fitting (Royston, 2005). Locally fitted lines also make it difficult to interpret findings or compare results to other studies (Royston, 2005).

In a polynomial regression, the number of higher order effects added to the regression model will depend upon the hypothesized shape of the relationship between the exposure and outcome variables. Some researchers advocate a stepwise approach to selecting the number of higher orders, comparing model fit after each subsequent addition of a higher order effect (Royston, 2005, McDonald, 2009). Conversely, it has been argued that, in epidemiology, it should be possible to anticipate the shape of the curve, and from this choose the polynomial to include (Greenland, 1995). In biological sciences it has been observed that any order higher than cubic is unlikely (McDonald, 2009).

For the reasons outlined above, polynomial regression was identified as the most suitable approach to detect a potentially non-linear relationship between diet costs and BMI. Due to the relative novelty of the exposure-outcome relationship, it was decided to run and compare models of both the quadratic and cubic orders, in addition to the standard linear regression.

7.3.4.3 Diet costs & overweight+obesity

As well as being expressed on a continuous scale, BMI is commonly grouped into categories of risk (WHO, 2006). Whilst information may be lost by grouping a variable in this manner (Naggara et al., 2011), the BMI risk categories are of clinical and public health significance. For this reason, a logistic regression was performed to assess the relationship between diet costs and a binary outcome variable of normal weight ('0') versus overweight and obese ('1').

The BMI categories of overweight and obese were combined for this analysis due to small participant numbers in the obese category. Unadjusted and adjusted odds ratios are presented alongside 95% CI. Adjustments were made for the same confounding variables as identified in the linear regression (see Section 7.3.4.5). The same sensitivity analyses were planned as for the diet cost-BMI regression (see Section 7.3.4.2 above).

7.3.4.4 Statistical power

As discussed previously (Section 4.3.4.5), the methods above relate to secondary analyses of an existing data set and therefore the sample size is already predetermined. A consideration of the power of the sample size to detect an effect in these particular analyses is necessary, however.

The desired effect size on which to base power calculations is a matter of judgement. For the analyses in this chapter, a similar approach to choosing the desired effect size will be adopted as was described in Section 4.3.4.5 – dichotomizing the predictor variable (diet costs) and estimating the expected difference in the outcome between those with high diet costs and those with lower costs. The nomogram method (see Whitley and Ball, 2002) (described in Section 4.3.4.5) will be used to gauge the expected power, knowing the available sample size, significance level and standardized mean difference (SMD; calculated from the expected effect size). A power calculation will need to be performed for each of the key relationships under investigation in this chapter: diet cost and dietary energy density; diet cost and BMI; and diet cost and overweight and obesity. These are described in turn below.

7.3.4.4.1 Diet costs & energy density

Chapter 2 describes the literature investigating diet costs and dietary energy density. From Section 2.4.3.2, it can be seen that four studies reported energy density by quantiles of diet cost (Aggarwal et al., 2011, Andrieu et al., 2006, Monsivais and Drewnowski, 2009, Townsend et al., 2009). Table 7.1 shows the difference in mean energy density between extreme quantiles of each study. Two of these studies (Townsend et al., 2009, Monsivais and Drewnowski, 2009) reported energy density in kcal, so were converted to kilojoules to allow comparison (1kcal = 4.186kJ).

Table 7.1 Summary of effect sizes from the literature investigating diet costs and energy density

Study	Number of income categories	Difference between extreme categories (kJ/g)	Energy density calculation
Andrieu et al, 2006	5	0.7	Food + caloric beverages
Monsivais and Drewnowski, 2009	4	Men 2.0 Women 1.0	Food only
Townsend et al, 2009	3	1.9	Food only
Aggarwal et al, 2011	5	2.7	Food only

Using the most conservative effect size from the literature – of 0.7kJ/g – and the standard deviation of energy density values in the NDNS (1.42kJ/g – see Chapter 4), gives a standardized mean difference of 0.493. With the available sample size of 507 participants in each group (dichotomizing a full sample of 1014), the nomogram indicates that the NDNS sample is highly powered to detect a difference of 0.7kJ/g (Table 7.2).

Table 7.2 Estimated power of the NDNS sample to detect hypothesized effect sizes (difference in dietary energy density)

Anticipated effect size (difference, kJ/g)	SMD	Significance level (α)	power
0.7	0.493	0.05	>0.995
		0.01	>0.995

7.3.4.4.2 Diet costs & BMI

Three studies investigated diet costs (estimated from dietary assessment) and BMI in working-age adults (see Chapter 2). These studies found a difference between extreme quintiles of diet cost of 0.9kg/m^2 (Murakami et al., 2007), 0.2kg/m^2 (Murakami et al., 2008b) and 1.2kg/m^2 (Lopez et al., 2009).

The NDNS sample has a mean BMI of 27.5kg/m^2 with a standard deviation of 5.3kg/m^2. Given a sample size of just over 800 participants (smaller than above due to missing or invalid BMI measurements), Table 7.3 shows the power estimated from the nomogram for each of the hypothesized effect sizes taken from the literature. It can be seen that the NDNS sample is inadequately powered to detect a small difference in BMI, but has around 85% power to detect a difference similar to that observed by Lopez et al, of 1.2kg/m^2.

Table 7.3 Estimated power of the NDNS sample to detect hypothesized effect sizes (difference in BMI)

Anticipated effect size (difference, kg/m^2)	SMD	Significance level (α)	power
0.2	0.038	0.05	0.10
		0.01	<0.05
0.9	0.170	0.05	0.65
		0.01	0.45
1.2	0.226	0.05	0.85
		0.01	0.74

7.3.4.5.3 Diet costs & overweight and obesity

The investigations involving diet costs and overweight and obesity use logistic regression. Therefore it is necessary to hypothesize a difference in proportions between the dichotomized groups, rather than a difference in means. The calculation for the SMD of proportions can be seen in Equation 4.3, Section 4.3.4.5.3.

Only one of the studies found in the literature review (Section 2.4.4.2) reported BMI category proportions (Lo et al., 2012), although this was between daily expenditure on vegetables only, and amongst a sample of elderly Taiwanese. In this study, the lowest quintile of vegetable expenditure was 30.5% overweight and obese, whilst the highest was 45.7%. Using Equation 4.3, this gives an SMD of: 0.152/0.486 = 0.313. Again using the nomogram, with an available sample size to allow two groups of about 400 participants each, the power calculated to such a difference in proportions is 98% at the 5% significance level (Table 7.4).

Table 7.4 Estimated power of the NDNS sample to detect the hypothesized effect size (difference in proportions overweight and obese)

Anticipated effect size (difference, %)	SMD	Significance level (α)	power
0.152	0.313	0.05	0.98
		0.01	0.94

7.3.4.5 Selection of covariates

For each of the relationships under investigation in this chapter – diet costs and dietary energy density, and diet costs and BMI – a directed acyclic graph (DAG) was created to identify appropriate confounding variables. Graphically linking the variables in this manner provides a rigorous method for confounder selection, as described in Chapter 4. The DAGs for the analyses in this chapter are shown in Figure 7.5 and Figure 7.6.

Figure 7.5 Directed Acyclic Graph (DAG) showing factors associated with dietary energy density and diet costs

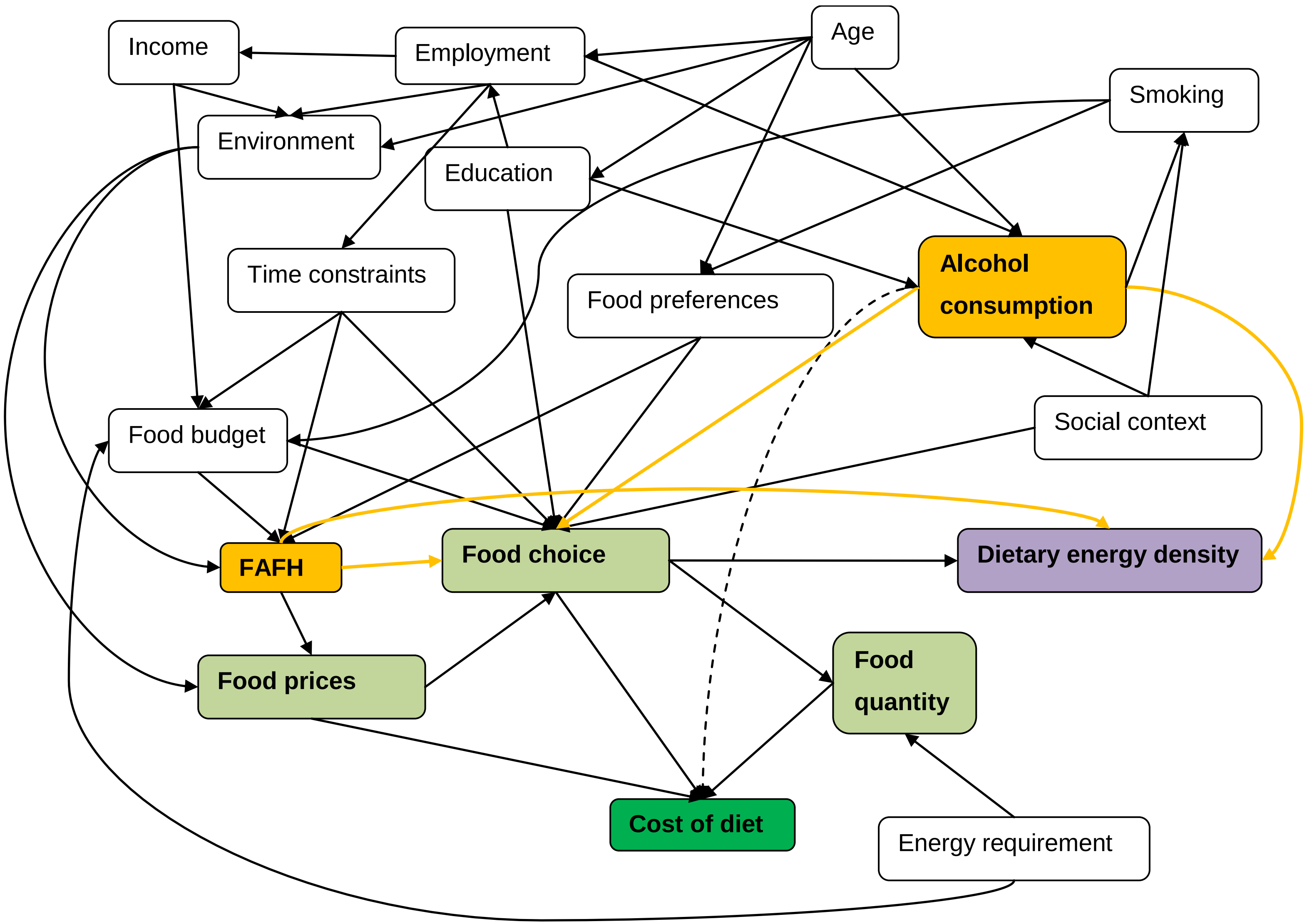
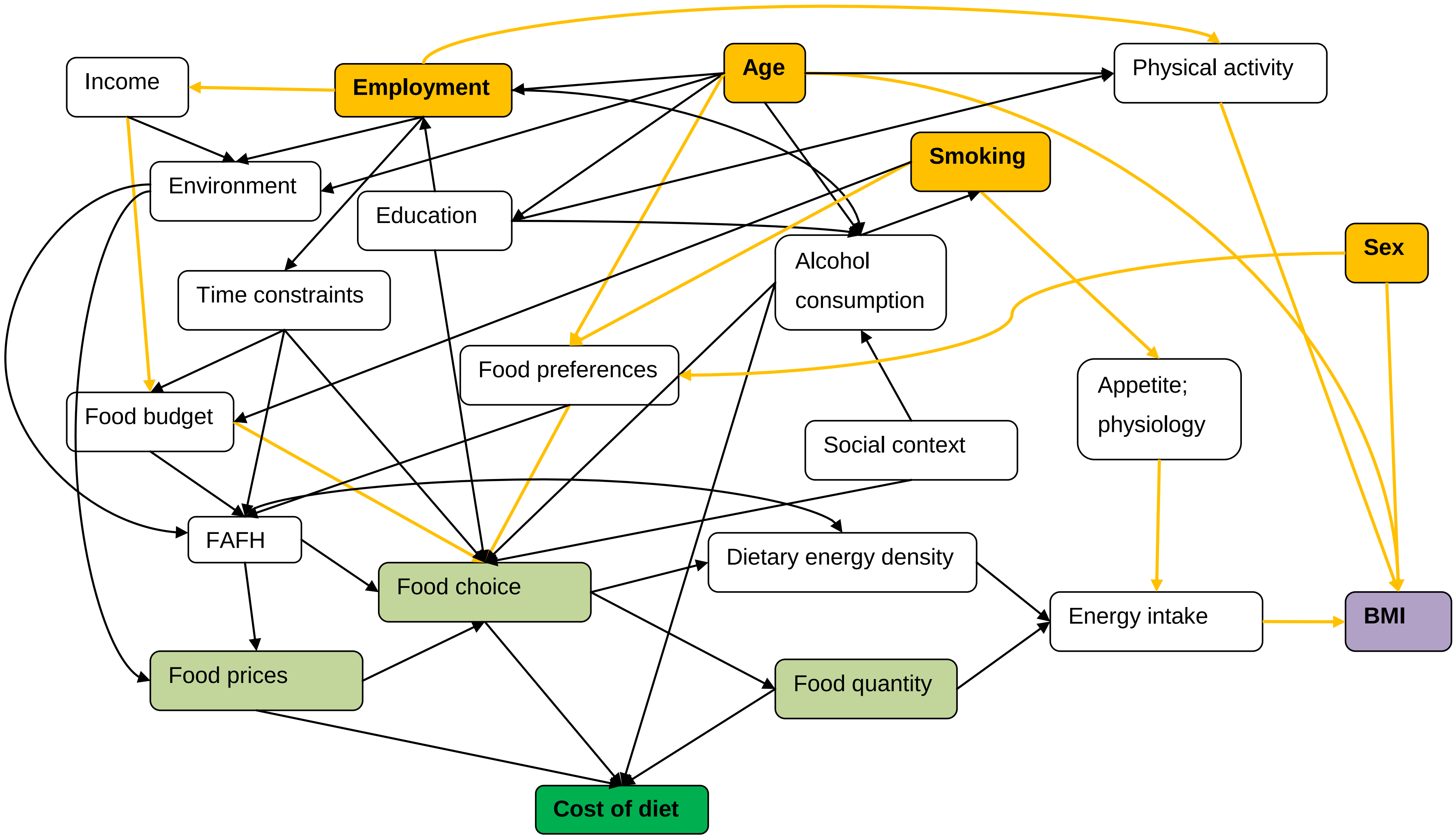


Figure 7.6 Directed Acyclic Graph (DAG) showing factors associated with body mass index and diet costs



In each DAG shown above, the exposure variable is depicted in dark green, whilst the outcome is shown in purple. It should be noted that there are no direct causal arrows linking the exposure and outcome in either graph. This is because diet cost in itself is not the cause of dietary quality or BMI. Rather, diet costs can be regarded as the representation of the interplay between food prices, food choice and food quantity (shown in pale green in the DAGs).

Tracing the ‘open backdoor pathways’ in these DAGs (see Section 4.3.4.6) reveals common causes of both exposure and outcome. In Figure 7.5 and Figure 7.6, the ‘backdoor pathways’ from the outcome were linked to the variables contributing to cost of diet (food prices, food choice and food quantity). These are depicted by the coloured arrows. Along each pathway, one variable has been selected (shown in orange) as a suitable adjustment. These were judged to be the minimum number of covariates able to capture the confounding pathways (the minimum being desirable in order to enhance the efficiency and robustness of the model (Bowers, 2008)).

For the diet cost-energy density relationship, two confounding variables were identified from the DAG (Figure 7.5): food away from home (FAFH) and alcohol consumption. Food away from home was identified as a cause of diet costs, because food ready to consume is usually of a higher price than that prepared in the home (Wrieden and Barton, 2011). At the same time, FAFH has been documented to be disproportionately energy-dense, compared to food at home (Prentice and Jebb, 2003), thus it can be said to confound the relationship between diet costs and dietary energy density. Alcohol consumption is also shown to be influential on diet costs, due to its relative expense, and on energy density, as would be expected for a liquid. Although beverages were excluded in the calculation of energy density, and diet costs are expressed excluding costs attributed to alcohol, the results presented in Chapter 6 hint at different dietary habits according to alcohol consumption group, even where costs or energy from alcohol are excluded. It was therefore retained as a confounding variable.

From Figure 7.6 it can be seen that the confounders identified in the hypothesised diet cost-BMI relationship are age, sex, employment and smoking status. Each of these have been linked in the literature to BMI: age and sex exert their influence via their roles as determinants of lean mass (Willett, 1998); smoking has been linked to weight status (Canoy et al., 2005), possibly due to an influence on appetite and subsequent eating behaviour; and employment can be said to influence daily physical activity (Proper and Hildebrandt, 2006), consequently impacting on energy balance.

Age, sex and smoking can also be linked to food choice, through their influence on food preferences. Differences in dietary choices according to these variables have been documented in a number of studies (Cade and Margetts, 1991, Breslow et al., 2006, Herman and Polivy, 2010, Renner et al., 2012). On the other hand, employment is connected to the exposure variable in the DAG via another route, impacting on food budgets, which depend upon the income received for employment. The fact that diet costs were shown to increase with income category in the NDNS (Chapter 6) supports this assumption.

A summary of the confounding variables chosen for each model is presented in Table 7.5. In addition to the confounders identified using the DAGs, each model was also adjusted for energy intake. This was due to the close relationship observed between energy intake and diet costs (see Figure 6.3). Issues relating to adjustment for energy in nutrition research are discussed further in Sections 7.2, 7.3.4.1 and 7.3.4.2 above.

Table 7.5 Variables to be included in each of the adjusted models

Exposure	Outcome	Adjustments
Residual of diet cost against energy intake	Dietary energy density	Alcohol consumption Food away from home Energy intake
Energy-adjusted diet cost	BMI	Energy intake Age Sex Smoking Employment
Energy-adjusted diet cost	Overweight+obesity (logistic)	Energy intake Age Sex Smoking Employment

7.4 Results

The median energy-adjusted diet cost of the sample was £4.05 per 10MJ (IQR £3.45, £4.82), excluding costs and energy from alcohol. Mean dietary energy density (excluding non-milk beverages) was 6.4kJ/g (SD 1.4kJ/g). The median BMI was 26.4kg/m^2 (IQR 22.9, 30.0; n=938). Sixty five per cent of the sample (n=607) were classified as either overweight or obese. More details on descriptive analyses can be found in Chapters 3 (BMI), 4 (energy density) and 6 (diet costs).

7.4.1 Diet costs & energy density

Values for average energy density, food energy, diet costs and BMI for each quintile of dietary energy density are presented in Table 7.6. Mean energy intake can be seen to increase with increasing quintiles of energy density. All other variables show no obvious trend by quintile, with the exception of diet costs per 10MJ when costs and energy from alcohol are excluded, where a decrease in cost is observed with increasing energy density quintiles.

Table 7.6 Mean and median values for each quintile of dietary energy density (1=lowest) (n=1014)

	Quintile of dietary energy density				
	1	**2**	**3**	**4**	**5**
Mean energy density, kJ/g (95% CI)	4.5 (4.4, 4.6)	5.6 (5.6, 5.6)	6.3 (6.3, 6.3)	7.1 (7.1, 7.1)	8.4 (8.3, 8.5)
Mean food energy, kJ (95% CI)	6012 (5753, 6271)	6776 (6510, 7042)	7424 (7135, 7713)	7728 (7435, 8021)	8278 (7930, 8626)
Median daily diet cost, £ d^{-1} (IQR)	3.34 (2.50, 4.69)	3.36 (2.55, 4.78)	3.37 (2.65, 4.62)	3.72 (2.75, 5.43)	3.54 (2.52, 4.86)
Median daily diet cost excluding alcohol, £ d^{-1} (IQR)	2.87 (2.27, 3.71)	2.74 (2.23, 3.42)	2.89 (2.31, 3.67)	2.91 (2.30, 3.66)	2.82 (2.12, 3.58)
Energy-adjusted diet cost, £ 10MJ^{-1} (IQR)	5.44 (4.50, 6.88)	4.89 (4.09, 6.13)	4.47 (3.72, 5.75)	4.67 (3.76, 5.83)	4.11 (3.33, 5.16)
Energy-adjusted diet cost excluding alcohol, £ 10MJ^{-1} (IQR)	4.93 (4.29, 5.71)	4.28 (3.68, 4.90)	3.87 (3.52, 4.57)	3.85 (3.36, 4.52)	3.42 (2.96, 4.07)
Mean BMI, kg/m^2 (95% CI)	28.7 (27.9, 29.5)	27.2 (26.5, 27.9)	28.0 (27.3, 28.7)	27.0 (26.3, 27.7)	26.8 26.0, 27.6)

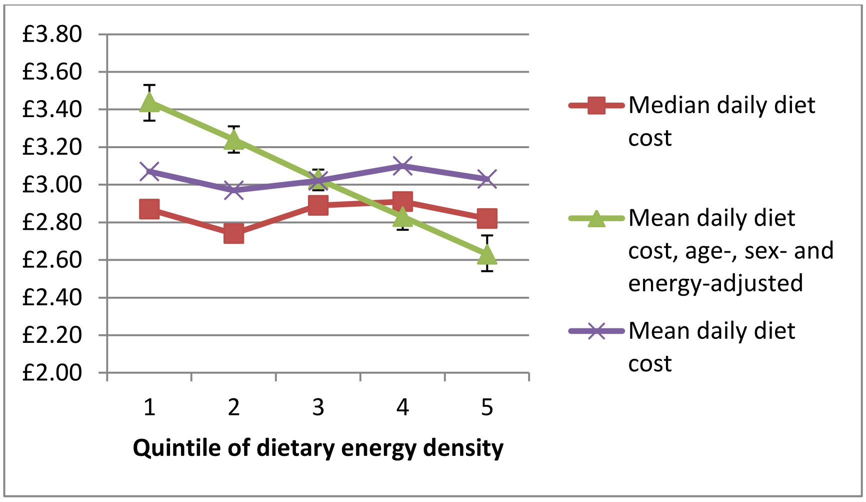
Mean quintile values for the diet cost variables are shown in Table 7.7, after adjusting for age and sex (and, in the case of daily diet costs, energy). Figure 7.7 and Figure 7.8 illustrate how these adjusted means compare to the unadjusted values. It can be seen that a negative relationship between diet costs and energy density is only apparent when energy is taken into account.

Table 7.7 Adjusted mean diet costs by quintile of dietary energy density (1=lowest), excluding alcohol (n=1014)

	Quintile of dietary energy density				
	1	**2**	**3**	**4**	**5**
Mean daily diet cost, £ d^{-1} (95% CI)	4.46 (4.26, 4.66)	4.22 (4.08, 4.36)	3.98 (3.87, 4.09)	3.74 (3.60, 3.88)	3.50 (3.29, 3.70)
Mean daily diet cost excluding alcohol, £ d^{-1} (95% CI)	3.44 (3.35, 3.54)	3.24 (3.17, 3.31)	3.03 (2.98, 3.09)	2.83 (2.77, 2.90)	2.63 (2.53, 2.72)
Mean cost per 10MJ, £ $10MJ^{-1}$ (95% CI)	5.20 (4.29, 6.11)	4.62 (3.71, 5.53)	4.18 (3.28, 5.08)	4.26 (3.38, 5.15)	3.75 (2.88, 4.61)
Mean cost per 10MJ excluding alcohol, £ $10MJ^{-1}$ (95% CI)	4.48 (3.86, 5.10)	3.73 (3.11, 4.35)	3.39 (2.78, 4.00)	3.36 (2.75, 3.96)	3.03 (2.44, 3.62)

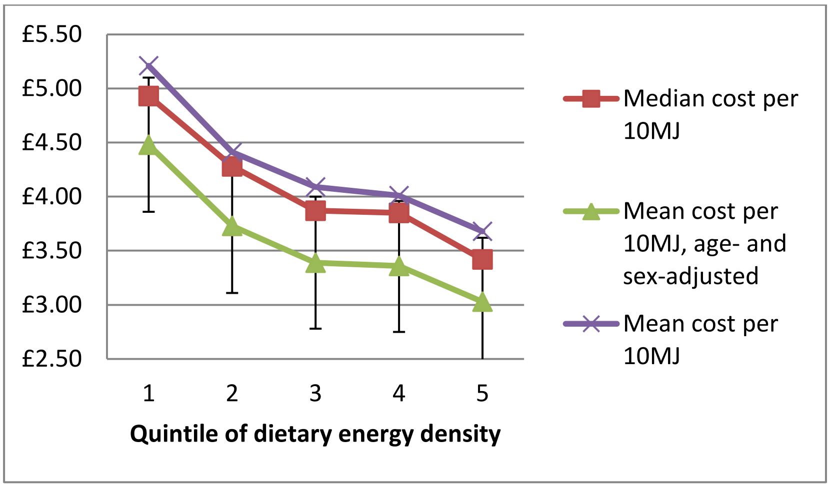
All values adjusted for age and sex. Daily diet costs also adjusted for food energy intake

Figure 7.7 Average daily diet costs for each quintile of dietary energy density (1=lowest), both with and without adjustments (n=1014)



Error bars show 95% CI for adjusted means

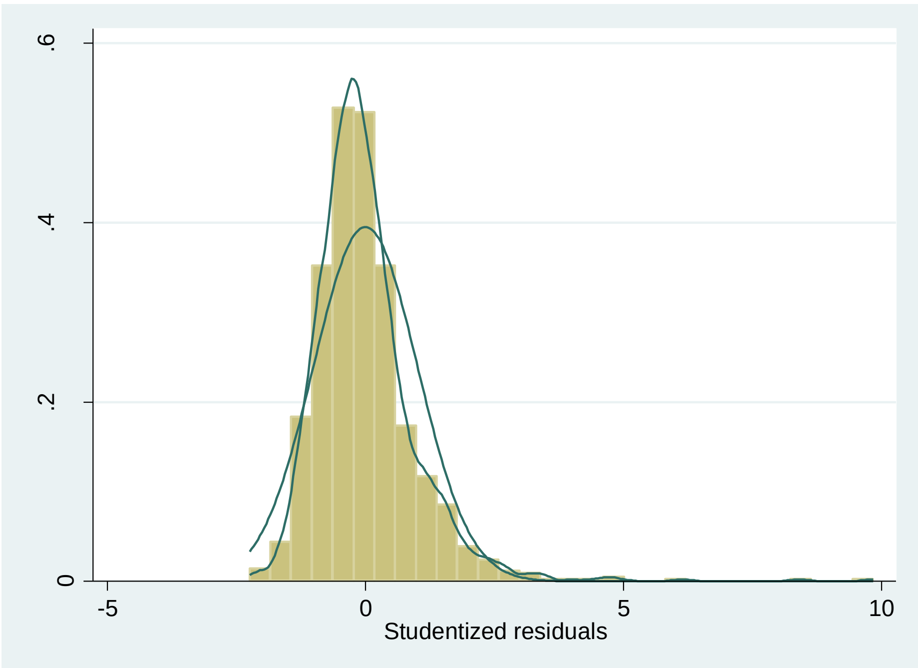
Figure 7.8 Average diet costs per 10MJ for each quintile of dietary energy density (1=lowest), both with and without adjustments (n=1014)



Error bars show 95% CI for adjusted means

The distribution of the studentised residual values from a diet cost-energy intake regression can be seen in Figure 7.9. Examination of the residuals of the full model revealed deviation from normality (skewness = 1.03, kurtosis = 6.72). Regression should be robust enough to handle this degree of non-normality (Bowers, 2008). A plot of the residuals against fitted values indicated constant variance. The assumptions for the regression analysis were judged to be met.

Figure 7.9 Histogram of studentised residuals from daily diet cost excluding alcohol (£ d^{-1}) plotted against food energy intake (n=1014)



Unadjusted linear regression revealed a significant negative association between diet costs (independent variable) and dietary energy density (dependent variable) (Table 7.8). After adjusting for alcohol consumption and FAFH, a stronger slope was indicated by the coefficient. As the residuals have been standardized, this is interpretable as a decrease in energy density of 0.46kJ/g for each additional standard deviation above the diet cost that would be expected for a given energy intake.

Excluding participants aged 70 years and over resulted in no appreciable differences to the estimates or model fit.

Table 7.8 Regression of diet cost on dietary energy density (residual method)

	n	Coefficient for diet cost residual	95% CI	p value	r^2
Unadjusted model	1014	-0.433	-0.516, -0.351	<0.001	0.095
Adjusted model*	1014	-0.455	-0.533, -0.377	<0.001	0.217
Adjusted model, excluding participants ≥70 years	857	-0.445	-0.530, -0.360	<0.001	0.212

Costs and energy from alcohol excluded
* Adjusting for energy intake, alcohol consumption and food away from home (FAFH)

7.4.2 Diet costs & BMI

Table 7.9 presents the results for the regression of energy-adjusted diet costs on BMI. Due to missing BMI values (n=76) and exclusion of underweight (n=13) the sample size was reduced to 925. There was no apparent effect of energy-adjusted diet cost on BMI, either unadjusted, or after adjusting for age, sex, employment classification and smoking. Sensitivity analyses did not improve model fit.

Table 7.9 Multivariable regression of diet costs (pence per 10MJ) on BMI (kg/m^2)

	n	Coefficient for Diet cost (p 10MJ-1)	95% CI	p value	r^2
Unadjusted model	925	<0.001	-0.001, 0.003	0.503	0.012
Adjusted model*	925	0.001	-0.002, 0.004	0.444	0.039
Adjusted model, excluding participants reporting atypical amounts	419	<0.001	-0.004, 0.004	0.870	0.031
Adjusted model, excluding special dieters	833	<0.001	-0.002, 0.003	0.814	0.046
Adjusted model, excluding participants ≥70 years	794	<0.001	-0.003, 0.003	0.950	0.057

Costs and energy from alcohol excluded. Adjusted for energy intake, alcohol consumption, smoking status and food away from home (FAFH)

A linear relationship was not apparent in the plot of the exposure and outcome (Figure 7.10). Polynomial regression did not find a significant non-linear effect of energy-adjusted diet costs on BMI. Figure 7.11 shows the quadratic line of best fit. Applying a Lowess smoothing function fitted a line closer to horizontal (not shown).

Figure 7.10 Energy-adjusted diet cost (pence per 10MJ) against BMI (kg/m^2) (n=925)

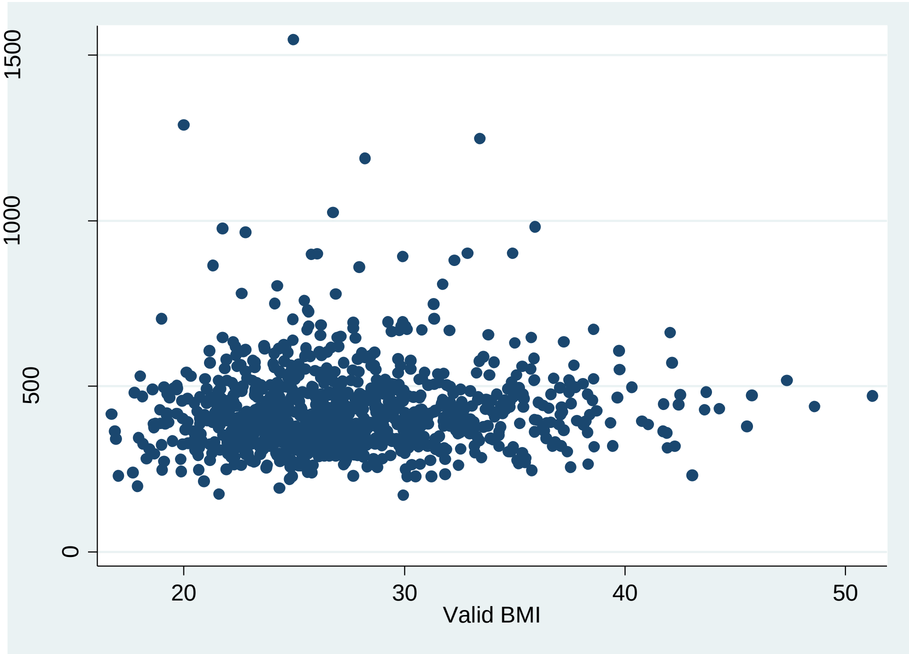
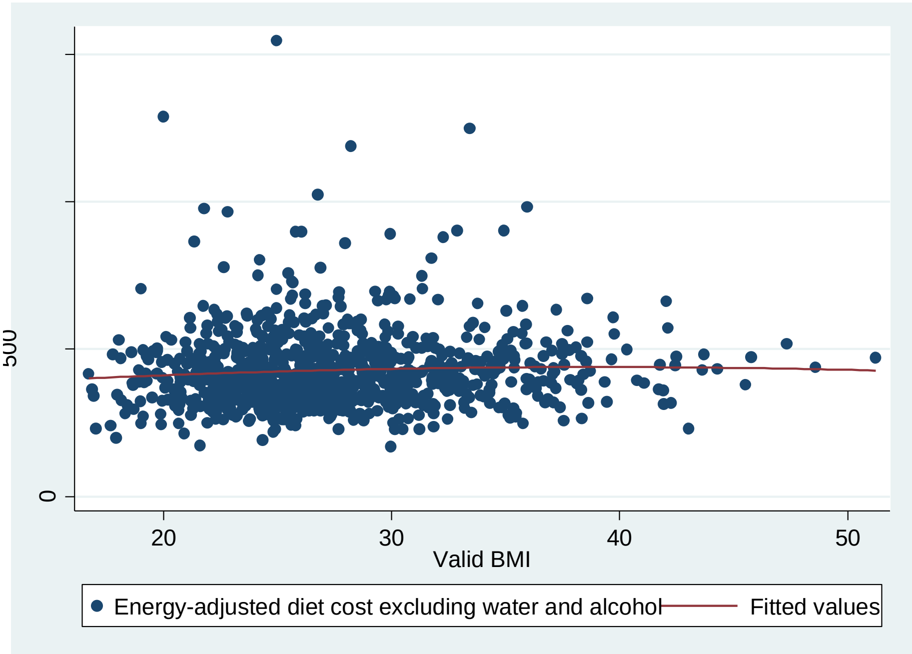


Figure 7.11 Energy-adjusted diet costs (pence per 10MJ) against BMI (kg/m^2), with 'curve' showing quadratic fitted values (n=925)



7.4.3 Diet costs & overweight+obesity

Logistic regression identified no difference in the odds of having a BMI of 25kg/m^2 or greater dependent on diet cost per 10MJ, either when adjusted for confounding variables, or in an unadjusted model (Table 7.10). Sensitivity analyses did not improve model fit.

Table 7.10 Logistic regression of energy-adjusted diet cost (pence per 10MJ) on the odds of being classified overweight or obese

	n	Odds ratio for diet cost (p 10MJ-1)	95% CI	p value	pseudo r^2
Unadjusted model	925	1.000	0.999, 1.001	0.660	0.002
Adjusted model*	925	1.000	0.999, 1.002	0.557	
Adjusted model, excluding participants reporting atypical amounts	419	1.000	0.999, 1.002	0.627	0.040
Adjusted model, excluding special dieters	833	1.000	0.999, 1.001	0.934	0.062
Adjusted model, excluding participants ≥70 years	794	1.000	0.999, 1.001	0.645	0.061

Underweight participants (BMI<18.5kg/m^2) excluded; costs and energy from alcohol excluded

* Adjusted models include as covariates: energy intake, age, sex, smoking status, and employment category

7.5 Discussion

The aim of this chapter was to identify whether there was evidence of a relationship between diet costs and dietary energy density or between diet costs and BMI amongst adults in the NDNS. This was enabled by the estimation of individual-level diet costs for this sample, as described in Chapter 6. While there was evidence to indicate a diet cost-energy density link, the data did not support an association with BMI or classifications of overweight and obese. This is the first study of this kind in a UK-wide population.

7.5.1 Diet costs & energy density

A strong inverse relationship was evident between the monetary cost and the energy density of diets. This confirms the findings of previous studies which similarly reported a negative association (Darmon et al., 2004, Maillot et al., 2007b, Monsivais and Drewnowski, 2009, Townsend et al., 2009, Waterlander et al., 2010, Ryden and Hagfors, 2011, Wrieden and Barton, 2011). The majority of previously published studies, however, did not make allowance for the mathematical coupling inherent in researching these variables (see Section 7.2; Lipsky, 2009). The three exceptions were: Darmon et al (2004), who counteracted the coupling by using an interaction term for energy intake and diet weight in their regression model; Maillot et al (2007b) who conducted both a multivariate density model and a model using the residual method; and Aggarwal et al. (2011) who used residuals. This chapter, like Maillot et al and Aggarwal et al, revealed evidence of a strong relationship when residuals are entered into a model to represent diet costs. This suggests that the observed relationship is not wholly due to mathematical artefact, as argued by some (Lipsky, 2009).

This link between the costs of diets and their energy density extends the observations in individual foods: that energy-dense foods tend to cost more than energy-dilute foods (Drewnowski et al., 2004, Waterlander et al., 2010). Establishing this pattern for the whole diet, rather than just for some constituent foods, enhances the relevance of such an observation in relating it to actual consumption.

Interestingly, each quintile of dietary energy density exhibited a similar median daily diet cost, especially when costs from alcohol were included (Table 7.6). This implies that, whilst it is possible that consumers in the highest quintile may be motivated to maximise the perceived return of their money (in purchasing more energy for the same daily cost), being restricted by a food budget might not be a primary motivator in the selection of more energy-dense foods. This finding is in contrast to that of Drewnowski et al (2007), however, in which weekly costs were apparently lower in the highest two quintiles of energy density. It is also contrary to the results of a linear programming study (Darmon et al., 2003), in which a forced cost constraint resulted in the selection of a more energy-dense diet.

Fruit and vegetable intake is another factor that has been consistently linked to the energy density of diets (see, for example, Waterlander et al., 2010): increasing the consumption of fruit and vegetables appears to have a diluting effect on energy density, as might be expected given the water content associated with fruits and vegetables. In this sample, dietary energy density did differ between those who did and those who did not achieve their ‘5 a day’, as did estimated diet costs (see Chapter 6). The degree to which this food group alone influences these outcomes could be helpful in interpreting the findings. Chapter 8 explores the role of food groups in diet costs, dietary energy density and consumption by BMI categories in the NDNS.

7.5.2 Diet costs & BMI

The second objective of the chapter was to explore the relationship between diet costs and BMI. In this study, there was no such association evident. Other published studies have presented evidence of both positive and negative associations between diet costs and BMI, as well as the absence of a relationship (see Section 7.2).

Possibly, these diverse findings are due to the heterogeneity of study design and samples. Cross-sectional surveys, in particular, may be criticised for attempting to extrapolate prior behaviour from current dietary practices. Obesity is acknowledged to have a protracted development, and current weight is a consequence of past, rather than current, behaviours. Therefore, measuring BMI and dietary behaviour at the same time point might not be meaningful. However, the study by Lopez et al. (2009) reported mixed findings using a prospective cohort design.

A second possible interpretation for the mixed findings is that the relationship proposed is more complex than a simple linear association. However, the polynomial investigations of this chapter failed to find evidence of a non-linear relationship.

The simplest conclusion, of course, is that there is no link between weight and diet costs - in this population or in general. Whilst this possibility cannot be ruled out, the limitations of this study, and the inconclusive results of previous investigations, caution against the full acceptance of this conclusion. Furthermore, there is evidence of a strong association between diet costs and dietary energy density, showing a link between current diet costs and current diet quality.

In this sample, as in other papers (Drewnowski et al., 2007, Waterlander et al., 2010), energy intake was positively associated with energy density. Given this observation, it follows that the consumption of an energy-dense diet, of lower inherent monetary value, will lead over time to a greater intake of energy than would an energy-dilute, more expensive diet. The potential for 'passive over-consumption' was previously described in a review of dietary energy density and the regulation of food intake in relation to fast foods (Prentice and Jebb, 2003). Several observational studies appear to corroborate this theory, reporting significant differences in dietary energy density according to BMI (Cox and Mela, 2000, Hartline-Grafton et al., 2009). In contrast, there were no statistically significant differences between the energy density estimates for the BMI categories of the NDNS (see Table 4.11, Section 4.4.1.2). Whilst results vary markedly depending upon the inclusion or exclusion of different (or all) beverages (Cox and Mela, 2000, Hartline-Grafton et al., 2009), the findings described above were not replicated in the NDNS even where the same specification was used in the calculation of energy density.

Whilst diet costs can be linked to energy density, and, in most cases (see Section 1.5), energy density linked to BMI (although not in the NDNS), there is a lack of evidence explicitly making the connection between all three variables along the proposed pathway. As far as the author is aware, this is the first investigation to examine the links between all three. There has been one other study which indirectly assesses this tripartite relationship. A pan-European and American analysis demonstrated a link between expenditure on FAFH (as a proportion of total food expenditure) and the relative risk of obesity in older adults, but for men only (Michaud et al., 2007). However, whilst FAFH is widely acknowledged to be more energy dense (Prentice and Jebb, 2003), the investigators in this study did not explicitly measure dietary energy density. The NDNS findings presented here failed to implicate either diet costs or energy density in the prevalence of obesity. This is disappointing, given the patterns observed in the literature, and it may be alleged that the explanation for this lies in the potential that the sample is biased, with systematic under-reporting suspected among the obese (see Section 7.5.3 below for a more detailed discussion).

7.5.3 Limitations

Readers of this thesis will already be familiar with the limitations associated with the NDNS data collection and with the DANTE cost database, which are described in

Chapters 3, 5 and 6. The discussion below relates some of the key points already described to the findings of this chapter.

The dietary data in particular are prone to measurement error, and this error has the potential to bias results. The suspected presence of mis-reporting is likely to result in ‘classical measurement error’, or, in the case of systematic mis-reports – for example, the obese consistently under-reporting energy intake – there may be ‘differential measurement error’ (Gannon, 2009). Under-reporting amongst obese participants is a phenomenon described in the literature (for example, Rennie et al., 2007), and, as discussed in Chapter 6, the pattern of energy intakes by BMI category in this NDNS sample is suggestive of this.

This suspected error, along with the challenges of cross-sectional design in researching BMI, makes it increasingly unlikely that the hypothesised associations would be detectable. If there is under-reporting as described, the data are unlikely to reveal a relationship even where there is one. Instead, the conclusions of this chapter must be predominantly based upon the analyses with dietary energy density as the outcome. Even this, however, despite resulting in strong associations with diet costs, will have been influenced by measurement error. Mis-reporting might explain the lack of detectable differences in energy density found between the BMI categories (Section 4.4.1.2), in contrast to the findings of other published studies. One way to limit the influence of such a bias would be to exclude those participants suspected of under-reporting. To do so, however, would require information on the physical activity of the participants, which is not available. (Physical activity data has been collected in a substudy of the NDNS but not for the full sample – results of the substudy had not been published at the time of writing – see Chapter 3).

The power calculations presented in Section 7.3.4.4 indicate only modest power of this sample to detect the proposed differences in BMI, the highest estimate being 64%. It is possible that the lack of significant association with diet costs is due to inadequate sample size. However, the power calculations indicated a stronger power of the sample to detect a difference in proportions of overweight and obese, at 88%, and logistic regression also failed to find a significant association. One limitation of these power calculations, it must be noted, is that they are based upon data from the only available studies, which involve populations different to the NDNS. These populations – US low-income women, elderly Taiwanese adults, female Japanese students, and a Spanish student cohort – are assumed to differ from the UK population (the latter two samples have a lower mean BMI than the NDNS, for example).

Finally, it needs reiterating that the diet costs analysed in this chapter represent the *inherent value* of the diet, as opposed to a measure of food expenditure amongst the participants. The costs reported are expressed in 2004 prices (the year the DANTE cost database was populated) and are not directly comparable to UK expenditure data from 2008-10, when the NDNS data were collected. In addition, the DANTE cost database is unable to account for foods eaten away from home. Despite this limitation, however, diet costs were still found to be associated with energy density after controlling for FAFH in this sample.

7.5.4 Strengths

Despite the drawbacks identified above, this chapter makes a useful contribution to the literature. In particular, it applies these methods for the first time to a national UK sample. As well as being geographically wider in sampling than the only other similar UK study (Wrieden and Barton, 2011), the current analyses estimate dietary intakes and costs from individual-level consumption as opposed to the data used by Wrieden & Barton (2011) from the Expenditure and Food Survey (EFS).

Much of the published literature on energy-adjusted diet costs and energy density has been criticised for failing to account for the mathematical coupling involved. The analyses in this chapter employ the residual method of Willett (1998) to address this limitation, confirming the existence of a true relationship. Furthermore, analytical methods new to this area of research are employed to determine – and reject – the possibility of a non-linear trend between diet costs and BMI, for the first time.

Finally, this study is the first to combine an exploration of diet costs with that of energy density and BMI. Despite the limitations in dietary data collection, a key strength of the NDNS is the use of professionally-measured anthropometry. This at least reduces the measurement error potential for this variable.

7.6 Conclusion

This is the first study of this kind in a UK-wide population. The analyses of this chapter took the individual-level diet costs estimated for the NDNS sample in Chapter 6, and linked them to both a dietary outcome and professionally-measured anthropometric data. The analyses confirm a diet cost-energy density link that is not due to mathematical artefact. The UK has no current set guidelines regarding dietary energy density. However, in relation to the WCRF recommendation of 5.23kJ/g (stated in kcal, 1.25kcal/g (WCRF, 2007)), all but the lowest quintile of energy density in this sample exceeded the goal. The implication of the regression results is that progression towards such a goal would be accompanied by a resultant increase in dietary costs for the majority of the NDNS sample.

On the other hand, the data did not support an association with BMI or classifications of overweight and obese. Whether the approach taken here is capable of implicating monetary factors in obesity remains to be seen. Prospective investigations, which include an assessment of energy expenditure to enable the identification of under-reporting, would be recommended for further investigations.

The following chapter takes a slightly different approach to the overall question, given the problems identified in this chapter regarding measurement error in energy intakes. The contribution of food groups are examined individually, and assessed in terms of explanatory power and usefulness in this field of research.

What was known previously:

- Chapter 6 identified significant sociodemographic differences in the diet costs of British adults.
- Diet costs have been linked, positively and significantly, to a variety of measures of dietary quality, including dietary energy density.
- Much of the published literature on energy-adjusted diet costs and energy density has been criticised for failing to account for mathematical coupling.
- Evidence of associations between diet costs and BMI is mixed, but no studies have investigated this in a UK sample.

What this chapter adds:

- This is the first study of this kind in a UK-wide population.
- A strong negative association between diet costs and energy density was evident, and the evidence confirms this is not due to mathematical artefact.
- The data did not support an association with BMI or overweight/obesity.
- The possibility of a non-linear trend between diet costs and BMI was tested, for the first time, and rejected.
- Mis-reporting might explain the lack of detectable differences in energy density found between the BMI categories.

Chapter 8 Food group costs & BMI in the NDNS

8.1 Summary

Chapters 6 and 7 describe analyses in which food prices were applied to the diets reported in the NDNS. This chapter extends these investigations using a fresh approach, in which diet costs for each of eight constituent food groups are analysed.

The rationale behind this chapter was that food group costs might offer a more detailed representation of diet costs than whole diet costs. Examining the costs of food groups is a little researched area. Analyses in this chapter explored: firstly, the relationships between the food groups costs and whole diet costs; secondly, how food group costs differed according to sociodemographic and other characteristics; and finally whether food group costs were associated with BMI or overweight.

Overall, foods in the meat, fish, eggs and beans category were found to be responsible for the greatest proportions of diet costs. However, alcoholic and non-alcoholic beverages were found to be the strongest determinants of whole diet costs in a multiple linear regression model. Comparisons revealed differences in at least one of the proportional food group costs between categories of almost all of the sociodemographic variables. The food group which differed the most according to socioeconomic variables was fruit and vegetables.

The linear regression and logistic regression models of food group costs on BMI or overweight/obesity revealed some significant associations, in contrast to the analyses using whole diet costs (see Chapter 7). A negative association was apparent between BMI and proportional costs of high-fat and high-sugar foods, suggesting a protective effect of this food group cost. In the logistic regression, the significant effect of high-fat and -sugar food group costs was no longer evident; instead associations were found for fruit and vegetables (negative) and the meat food group (positive).

These findings suggest that normal weight, overweight and obese individuals apportion their food budget differently. In contrast, whole diet costs do not differ by BMI category. This implies that it is not the food budget per se that encourages positive energy balance, but rather how people apportion their budget, and suggests that costing diets in this manner could have some use in future research into diet costs. The differences observed for high-fat and –sugar foods possibly reflect some bias from under-reporting. These methodological challenges make it difficult to ascertain the role of food group costs in excess weight. Nevertheless, sociodemographic observations could have implications for policy.

8.2 Introduction

Chapters 6 and 7 presented a description of the inherent values of adults' diets in the NDNS, in terms of daily costs and per 10MJ. Whilst sociodemographic patterns in diet costs were evident, and diet costs were found to be associated with dietary energy density, analyses failed to uncover a relationship between costs of the whole diet and measures of body mass and obesity. This adds to the already conflicting findings from other studies reported in the literature (see Chapter 2).

As an emerging research area, the best available method for investigating the monetary aspects of diet is yet to be established. As stated in Chapter 6, estimated daily diet costs are strongly associated with, and largely determined by, the quantity of food consumed. To be confident that analyses involving diet costs are not merely reflecting the quantity of food consumed, it is necessary to control for energy. Expressing costs per unit of energy (or a standard amount, such as 10MJ) goes some way to addressing this, as does the residual method in a regression model (Chapter 7). However, both of these methods have their drawbacks.

The primary disadvantage of relating dietary costs to a standardized energy amount is that information about the experience of the individual risks being lost. The costs estimated by a food cost database signify the inherent monetary value of diets, rather than actual expenditure. Nevertheless, they may offer insight into how people reconcile their food purchase decisions within a given food budget. Adjusting the costs to 10MJ gives an indication of energy cost, but it has been argued that energy cost, as a construct which is unavailable to the consumer at the point of purchase, is unlikely to guide food purchasing decisions (Lipsky, 2010). Therefore, estimating the inherent value of dietary energy is of limited use if investigators are interested in making statements about dietary choices. The residual method can similarly be considered a representation of energy cost, and is thus subject to the same limitation.

Although consumers may not explicitly base purchasing decisions on a calculation of energy cost, cost of food has been extensively reported as an important determinant of dietary decision-making (Steptoe et al., 1995, Shepherd et al., 2006, Nelson et al., 2007). If cost is not considered by the consumer in terms relative to energy, the challenge is to find a measure able to capture the influence of cost.

Lower-budget consumers, for whom food costs are perhaps a more salient aspect of food purchasing, may apportion their food budget differently to those with a more generous budget. Examining the contributions of constituent food groups to whole diet costs gives an indication of how people may apportion their budget, as well as how these proportions change as budgets vary.

A handful of previously published studies have reported proportional food group costs or the energy costs of food groups (Cade et al., 1999, Murakami et al., 2007, Ryden et al., 2008, Ryden and Hagfors, 2011, Alexy et al., 2012). These investigations either reported the relationships between each food group cost and whole diet costs (Murakami et al., 2007, Ryden and Hagfors, 2011), or examined proportional food group costs according to dietary pattern (Ryden et al., 2008), healthfulness score (Cade et al., 1999), or dietary energy density (Alexy et al., 2012).

There is nothing in the literature comparing proportional food group costs by sociodemographic variables. However, describing these costs for population subgroups could be informative, particularly from a policy perspective where reducing health inequalities is a priority. Nor have food group costs been described by BMI category before, although food prices – and of energy-dense foods in particular – have been hypothesized to be culpable in the aetiology of obesity trends (Drewnowski and Darmon, 2005) and researchers have previously used whole diet costs to explore this theory (Murakami et al., 2007, Murakami et al., 2008b, Lopez et al., 2009).

The aim of this chapter is to explore the proportional food group costs of NDNS adults' diets, in relation to BMI. The analyses will satisfy the following thesis objectives:

1. To investigate the appropriateness of diet cost estimations, including the costing of food groups;
2. To estimate and describe the diet costs of NDNS adults;
3. To explore patterns in NDNS diet costs according to sociodemographic characteristics; and
4. To determine whether an association exists between diet costs and BMI or overweight amongst NDNS adults.

These objectives are also addressed in Chapters 6 and 7 with respect to whole diet costs. This chapter expands on the previous chapters' investigations by analysing the costs of constituent food groups. The costs of the constituent food groups will be newly derived for this chapter. As such, the analyses include explorations to help characterise these new variables: examining the relationships within food group costs, between food group costs and whole diet costs, and in relation to proportional energy intake by food group.

In contrast to the previous chapters, this chapter will not examine dietary energy density as an outcome. This is due to the disproportionate influence of a few food groups on dietary energy density – for example, fruit, vegetables and dairy products are associated with a higher water content and therefore a lower energy density than other food groups (Darmon et al., 2004). Instead the focus will be on BMI and obesity prevalence. Given that there was little trend evident in whole diet costs by BMI, the difference in food group costs could be informative.

8.3 Methods

8.3.1 Sample

This chapter again makes use of 2008-2010 NDNS data (NatCen et al., 2012). Information about the survey – design, ethical approval, recruitment, response rate and sample characteristics – is given in Chapter 3. Dietary consumption is measured in the NDNS by consecutive four-day un-weighed food diaries. Respondent characteristics were ascertained during a face-to-face interview, and anthropometric measurements (including height and weight) were measured by health professionals.

The analytical sample included only adults with complete diary data (n=1014). In addition, participants missing a valid BMI measurement and those who were classified as underweight (BMI<18.5kg/m^2) were excluded from the regression analyses, leaving a sample of 925 participants, from a possible 1031. The rationale for these exclusion criteria are described elsewhere (Chapter 4).

8.3.2 Calculation of food group costs

Costs for each food group were calculated both in absolute terms and as a percentage of the whole diet cost. Daily costs for each food group will help describe the range of costs experienced for that food group, as well as indicating the median monetary value of each food group as consumed by this sample. On the other hand, expressing food group costs as a proportion of whole diet cost will help illustrate how the cost of each food group contributes to the total diet cost, and in this way could indicate how food budgets are composed.

In contrast to the previous chapter, it was considered inappropriate to standardize food group costs to a common energy amount (such as 10MJ). The main reason for this is because of the focus on percentage costs in regression analyses (see below), which, being proportional, would be equivalent at all energy amounts (expressing costs as a percentage in itself could be considered a form of standardization). For the descriptive analyses, there are other reasons to avoid standardizing food group costs to a common energy amount: there are no recommendations for energy intakes from food groups, nor are they commonly reported, making the selection of a standard amount arbitrary, and most likely unhelpful for interpretation.

8.3.2.1 Absolute costs

Costs were estimated using the DANTE food cost database. A price was applied to each food item and beverage consumed in the NDNS by linking the data sets in Microsoft Access. A full description of this method is presented in Chapter 6. Cost calculations excluded water, and involved prices unadjusted for inflation. No outliers were excluded.

Food items in the NDNS data each contain a main food group code, of which there are 60 defined in the dietary analysis software used for the survey (Diets In Nutrients Out, or DINO). A summed cost was calculated for each of these 60 food groups for each participant using the Microsoft Access database. In a similar method used for calculating whole diet costs, daily costs (C_i^k) for each food group (k) were calculated by summing the prices (p) of the foods consumed in the quantities (q) consumed by each individual (i), and dividing by the number of days (d) to give a daily average:

$$C_i^k = \frac{\sum(p_i^k q_i^k)}{d}$$

Equation 8.1

For the purposes of analysis, it was necessary to collapse the 60 food groups into a smaller number of categories. In keeping with current UK guidelines (the eatwell plate, DH, 2011), eight food groups were chosen:

- meat, fish, eggs and beans
- fruit and vegetables
- starchy foods
- milk and dairy
- foods high in fat and/or sugar
- non-alcoholic beverages
- alcoholic beverages
- miscellaneous foods.

More detail for these food groupings can be found in Appendix D. A look-up file was created manually to match each of the 60 DINO food group codes to the appropriate food group listed above. Daily costs (£ d^{-1}) were then derived for each of the eight food groups by summing the appropriate DINO food group totals. Prices were not corrected for inflation (see Chapter 6 for a discussion of inflation adjustments).

8.3.2.2 Proportional costs

The daily food group costs were merged into the Stata (StataCorp, 2011) data set (see Chapter 6) as eight new variables. For each participant, the food group costs were then divided by the daily diet cost to give a percentage, or proportional cost. This was performed both including and excluding alcohol from calculations:

a) costs of each of the eight food groups were divided by the whole diet cost including alcohol, to give eight new variables;
b) and costs of the seven food groups (alcoholic beverages comprising the excluded food group) were divided by the whole diet cost that excluded costs from alcohol, resulting in seven further variables.

8.3.2.3 Other food group variables

In addition to the proportion of whole diet cost that each food group contributed, proportional values were also calculated for the contribution to energy intake (kJ) of each food group, and the proportion of total diet mass (g).

8.3.3 Analytical methods

8.3.3.1 Descriptive statistics

Stata IC 12 (StataCorp, 2011) was used for all statistical analyses. Summary statistics of each food group cost variable (of which there are 23) were calculated for the whole sample. Distributions were positively skewed in every case; medians and interquartile ranges (IQR) are presented. Descriptive statistics for proportions of energy intake and of mass for each food group are also shown.

Spearman correlations explored how absolute food group costs (£ d^{-1}) related to each other and to the whole diet cost (both £ d^{-1} and £ $10MJ^{-1}$). A multivariable regression assessed the strength of each food group's cost in predicting whole diet costs, after adjusting for the other food groups.

8.3.3.2 Sociodemographic comparisons of proportional costs

Subgroup comparisons were made using proportional, rather than absolute, food group costs. As described in the Introduction (Section 8.2), this was because proportional costs were felt to potentially reflect the reconciliation of the food budget.

Kruskal-Wallis ANOVAs compared proportional food group costs (% whole diet cost) between categories of sex and age. Age- and sex-adjusted geometric means were calculated for categories of each of the following variables:

- employment,
- equivalized income,
- qualifications,
- household size,
- marital status,
- cigarette-smoking status,
- ‘5-a-day’ achievement and
- alcohol consumption category.

A linear regression analysis was performed for each logged proportional food group cost variable, adjusting for age and sex, to identify between-group differences (112 models in total). Due to the number of tests, a significance level of 1% was set.

8.3.3.3 Food group costs and BMI

As the primary outcome of interest in this chapter, summary statistics for BMI categories are presented separately. Age- and sex-adjusted differences between BMI categories were assessed using linear regression analysis with each logged proportional food group costs. In contrast to the other between-group comparisons above (which investigated only proportional costs), absolute costs between BMI categories were also compared.

The relationship between each food group cost and BMI (kg/m^2) was investigated using multivariable linear regression, adjusting for age, sex, employment and energy intake. (The selection of covariates is described in Section 7.3.4.5.)

Proportional food group costs (% whole diet cost) were used to assess these potential relationships. This was in keeping with the rationale outlined in the Introduction to this chapter (Section 8.2): with the idea that how people apportion their food budget may be more informative than the diet cost per se.

Proportions are similarly employed in analyses investigating energy from macronutrients – for example, per cent energy from fat – and the treatment and interpretations are analogous. As an illustrative example, in order to isolate the influence of energy from fat from the influence of total energy, it would be necessary to hold total energy intake constant (this is equivalent to having isoenergetic treatment arms in experimental studies). If absolute values of energy intake were used in a regression analysis, an increase in energy from fat would also in effect be an increase in total energy, if energy from carbohydrate and from protein are included in the model.

By extension, including all absolute food group costs in a regression model would lead to similarly problematic interpretations: for example, a coefficient for fruit and vegetable costs would be the expected effect of increasing fruit and vegetable costs while holding all other food group costs constant, but this would equate to an increase in total diet cost, therefore not isolating an effect of fruit and vegetable costs.

This problem in interpretation holds true for proportional values, in that it is not feasible that one constituent could vary whilst the other constituents are held constant – if all proportional values are included, they should add up to 100%, whereas an increase in one proportion but not the others would in theory see the whole adding up to more than 100%. In macronutrient studies, a common solution to this is to exclude one macronutrient in the regression model – chosen as a referent (Willett, 1998). As an example, we could select per cent energy from carbohydrate as the referent. The coefficient of the macronutrient of interest – for example, fat – would then be interpreted as the effect of both an increase in per cent energy from fat along with a corresponding decrease in per cent energy from carbohydrate so that total energy remains constant.

The choice of referent macronutrient would have an important bearing on conclusions regarding the role of energy from fat in the example above – a substitution effect is implied, and the substitution of fat for protein may well have different results. In the case of food group costs, it is unclear how a food group would be chosen as a referent – there is no evidence as yet upon which to base such a decision – but this would have an important impact upon the interpretation of results. In addition, the effect of all, and not just one, of the food groups is of interest in this chapter, and omitting a food group from the multivariable model would not allow an assessment of that food group's influence on the outcome.

An alternative solution is to run a separate regression analysis for each food group. By including each food group in a separate model, there is no statement made about where the substitution is taking place, or which other food groups experience a corresponding change in proportional costs as a result of a change in proportional cost from the included food group. By not including other food groups in the model, it is implied that the corresponding change is shared across all the excluded food groups.

Separate models for each food group were therefore judged the most straightforward method for interpretation, and, because this chapter is concerned with the effect of each food group and not just one food group, the number of p values of interest would be the same whether they were in separate or combined models. There was therefore judged to be no additional risk of false positive results using this approach. For these reasons, fifteen separate models were run: eight models (one for each food group) including costs from alcoholic beverages; and a further seven models excluding costs from alcohol.

The models were repeated with the following sensitivity analyses: excluding those who reported consuming an unusual amount ('less than usual' or 'more than usual' – see Chapter 3); excluding those who reported adhering to a special diet; and excluding those aged 70 years and over. The rationale for these exclusions is explained in Section 7.3.1. A significance level of 1% was set.

8.3.3.4 Food group costs and overweight and obesity

Similarly to Chapter 7, logistic regression was used to investigate the relationship between food group costs and the binary outcome of normal ('0') or overweight and obese ('1'). As described previously, this was due to the clinical significance of BMI classifications. Overweight and obese categories were combined.

As for the linear regression described above, logistic regression models were run separately for each food group, both including and excluding costs from alcoholic beverages, giving a total of 15 analyses. Models were adjusted for age, sex, employment and energy intake. Odds ratios (OR) and 95% CI are presented. A significance level of 1% was again adopted due to multiple tests.

The same sensitivity analyses were conducted as above (Section 8.3.3.3).

8.3.3.5 Statistical power

The analyses described above are secondary analyses of already-collected data. As such, they are constrained by the available sample size. It is still worthwhile, however, to estimate the power given by the sample size for these new analyses in order to judge whether the power is sufficient to detect the alternative hypothesis.

Many of the analyses of this chapter are exploratory or descriptive in nature (the sociodemographic comparisons, for example). This section will concentrate on estimating the statistical power for the hypothesis-driven analyses, investigating BMI or overweight/obesity as outcomes.

Knowing the sample size a priori, it is possible to calculate the statistical power of the study that is needed to detect the effect size that is expected (see Section 4.3.4.5). Unfortunately, no published studies are available from which to hypothesize an expected effect size. Therefore, for the linear regression analyses, a desirable effect size of 1kg/m^2 was selected arbitrarily. This would be the desired effect size if the sample were dichotomised based upon the predictor variable – an approach for regression power calculations advocated by Greenwood (2011).

Using the nomogram method first put forward by Altman (1991) and described by Whitley and Ball (2002) and in Chapter 4, the power can be estimated using the known sample size, the desired α (in this case, 0.01), and the standardized mean difference (SMD). A desired difference between group means of 1kg/m^2 and a standard deviation for BMI in the sample of 5.19kg/m^2 gives an SMD for this sample of 0.193. Drawing a line through these points on the nomogram indicates a study power of around 0.50, or 50%. Power calculations for alternative effect sizes and significance levels are presented in Table 8.1.

Table 8.1 Estimated power of the NDNS sample to detect hypothesized effect sizes

Anticipated effect size (difference, kg/m^2)	SMD	Significance level (α)	power
1	0.193	0.05	0.76
		0.01	0.50
1.5	0.289	0.05	0.97
		0.01	0.92
2	0.385	0.05	>0.995
		0.01	0.994

For the logistic regression analyses (with overweight and obese as the outcome), it is necessary to calculate the SMD from the estimated or desired difference in proportions. Again, there is no established literature which provides data on which to base this estimate; therefore a range of desired differences were chosen arbitrarily, and the power calculations for these given in Table 8.2 below.

Table 8.2 Estimated power of the NDNS sample to detect hypothesized effect sizes

Anticipated proportion overweight/obese (high food group cost, low food group cost)	Anticipated effect size (difference)	SMD	Significance level (α)	power
60%, 65%	5%	0.103	0.05	0.31
			0.01	0.15
60%, 70%	10%	0.210	0.05	0.82
			0.01	0.67
58%, 73%	15%	0.316	0.05	0.99
			0.01	0.97

8.4 Results

8.4.1 Descriptive results

Median and mean cost values for each food group are presented in Table 8.4 alongside the proportion each food group contributes to the total daily diet cost.

Costs attributed to meat, fish, eggs and beans were the highest of the food groups, and constituted the greatest proportion of whole diet costs. The next largest contributor to whole diet costs was the fruit and vegetable food group, which contributed approximately half as much as meat, fish, eggs and beans.

The cost of each food group was found to be significantly and positively correlated with food energy intake (Spearman's rank coefficients all $p<0.001$).

Table 8.3 shows the proportions of whole diet cost contributed by each food group, alongside the proportions of food energy intake (EI) and proportions of total daily food weight (g).

Table 8.3 Median proportions of food energy intake (EI), mass (g) and daily diet cost (£) contributed by each food group (costs of alcohol excluded)

Food group	Median proportion of EI (%)	Median proportion of g (%)	Median proportion of whole diet cost (%)
Starchy foods	25	8	10
Fruit & vegetables	9	11	15
Meat, fish, eggs & beans	22	7	29
Milk and dairy	11	7	8
Foods high in sugar and/or fat	23	4	10
Non-alcoholic beverages	1	49	13
Miscellaneous foods	3	1	4

8.4.2 Relationships between food group and whole diet costs

Between the food groups, the strongest correlation was observed between the dairy and fruit and vegetable groups ($r = 0.29$); all other pairwise correlations were less than 0.2 (Figure 8.1).

Table 8.4 Median and mean diet costs of NDNS adults by food group (n=1014)

Food group	Mean cost, £ d^{-1} (95% CI)	Median cost, £ d^{-1} (IQR)	Mean % of whole diet cost (95% CI)	Median % of whole diet cost (IQR)	Mean % of whole diet cost excluding alcohol (95% CI)	Median % of whole diet cost excluding alcohol (IQR)
Starchy foods	0.36 (0.34, 0.38)	0.28 (0.19, 0.45)	10 (10, 10)	8 (5, 13)	12 (12, 12)	10 (7, 15)
Fruit & vegetables	0.53 (0.50, 0.56)	0.44 (0.23, 0.72)	14 (13, 15)	12 (7, 20)	17 (16, 18)	15 (9, 23)
Meat, fish, eggs & beans	0.92 (0.88, 0.96)	0.82 (0.56, 1.17)	25 (24, 26)	24 (16, 32)	30 (29, 31)	29 (21, 38)
Milk and dairy	0.27 (0.26, 0.28)	0.22 (0.12, 0.36)	8 (8, 8)	6 (3, 10)	9 (9, 9)	8 (5, 12)
Foods high in sugar and/or fat	0.32 (0.31, 0.33)	0.27 (0.15, 0.44)	9 (9, 9)	8 (4, 12)	11 (11, 11)	10 (6, 15)
Alcoholic beverages	0.94 (0.84, 1.04)	0.30 (0.00, 1.24)	17 (16, 18)	9 (0, 30)	-	-
Non-alcoholic beverages	0.49 (0.45, 0.53)	0.36 (0.22, 0.55)	13 (12, 14)	10 (6, 16)	15 (14, 16)	13 (8, 19)
Miscellaneous foods	0.17 (0.16, 0.18)	0.12 (0.05, 0.23)	5 (5, 5)	3 (1, 6)	6 (6, 6)	4 (2, 8)

Figure 8.1 Pairwise scatter plots of food group costs, line of best fit shown

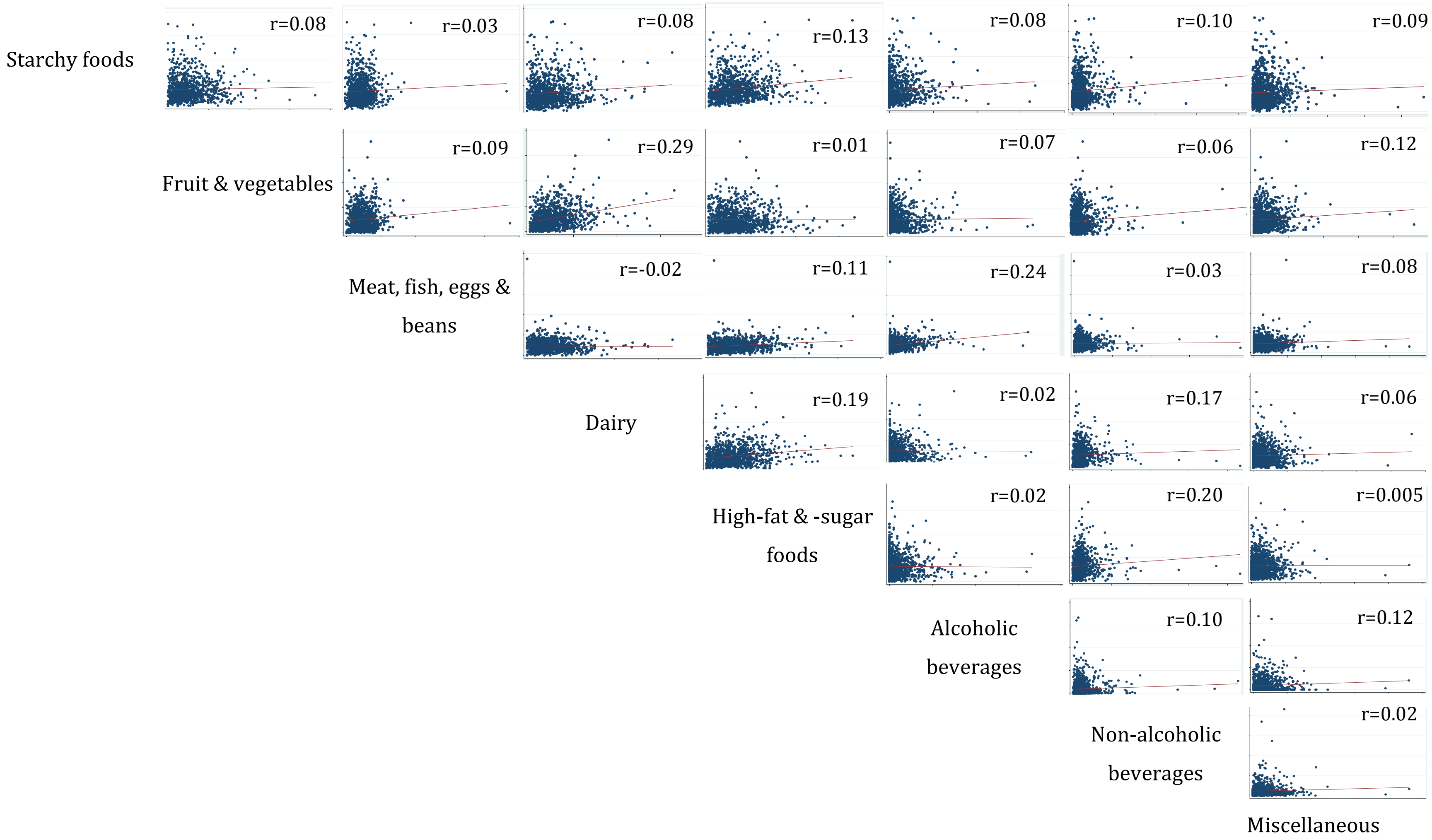


Table 8.5 and Table 8.6 display the results of the multiple linear regression models. The coefficients represent the difference in diet costs (£ d^{-1} or £ $10MJ^{-1}$) that is associated with each additional £0.01 from the food group variable. Alcoholic beverages were found to have the strongest effect on whole daily diet costs and on diet costs per 10MJ (Table 8.5), with an associated increase of £0.75 in daily diet cost or £0.58 in cost per 10MJ for every additional penny spent on alcohol. Excluding costs from alcohol (Table 8.6), non-alcoholic beverages and the meat, fish, eggs and beans category were found to have the largest effect on daily diet costs, whilst non-alcoholic beverages and fruit and vegetables were the strongest predictors for diet costs per 10MJ, after controlling for other food groups.

The dairy food group was the only group not to be significantly associated with all the whole diet cost variables: it was not a significant predictor of diet costs per 10MJ when alcohol was excluded (p=0.12).

Table 8.5 Multiple regression of food group costs on whole daily diet cost (n=1014; r^2 = 1.00) and on diet costs per 10MJ (r^2=0.678)

Variable	Daily diet cost (£ d^{-1})	p	Diet cost per10MJ (£)	p
Starchy foods	0.12	<0.001	-0.08	<0.001
Fruit & vegetables	0.19	<0.001	0.24	<0.001
Meat, fish, eggs & beans	0.27	<0.001	0.17	<0.001
Milk and dairy	0.09	<0.001	-0.06	<0.01
Foods high in sugar and/or fat	0.11	<0.001	-0.17	<0.001
Alcoholic beverages	0.75	<0.001	0.58	<0.001
Non-alcoholic beverages	0.27	<0.001	0.37	<0.001
Miscellaneous foods	0.08	<0.001	0.08	<0.001

Table 8.6 Multiple regression of food group costs on whole daily diet cost (n=1014; r^2 = 1.00) and on diet costs per 10MJ (r^2=0.556), excluding costs from alcohol

Variable	Daily diet cost (£ d^{-1})	p	Diet cost per10MJ (£)	p
Starchy foods	0.22	<0.001	-0.07	<0.001
Fruit & vegetables	0.35	<0.001	0.36	<0.001
Meat, fish, eggs & beans	0.49	<0.001	0.25	<0.001
Milk and dairy	0.17	<0.001	-0.03	0.12
Foods high in sugar and/or fat	0.20	<0.001	-0.17	<0.001
Non-alcoholic beverages	0.50	<0.001	0.53	<0.001
Miscellaneous foods	0.15	<0.001	0.11	<0.001

8.4.3 Comparisons by sociodemographic and other variables

8.4.3.1 Proportions of diet cost from alcohol

The percentage of daily diet cost attributable to alcohol differed significantly between the sexes and between age groups: see Table 8.7. Once adjusted for age and sex, the proportional costs of alcohol for all other demographic groups were not found to differ.

In terms of lifestyle variables, costs attributed to alcohol were higher amongst current smokers (24%) than ex-regular smokers (21%) or those who had never smoked (18%; p=0.004). Those who did not achieve their '5 a day' fruit and vegetables also had higher proportions of diet cost attributed to alcoholic beverages (23% versus 17%, p<0.001).

As the inclusion or exclusion of alcohol affects the proportional cost estimates and due to the differences identified above, the estimates presented below show results both including and excluding alcohol.

8.4.3.2 Proportional food group costs by age and sex

Table 8.7 shows the median proportions of diet costs contributed by each food group for males and females and by age strata. Proportions of cost from fruit and vegetables, dairy, non-alcoholic beverages and miscellaneous foods were significantly higher amongst females (all p<0.001), whilst the proportion of cost given to alcoholic beverages was significantly higher amongst males (15% versus 3%, p<0.001). Age groups were also found to differ significantly in their proportional costs for all food groups except non-alcoholic beverages and miscellaneous foods.

When costs from alcoholic beverages were excluded (Table 8.8), median proportional costs differed between females and males for fruit and vegetables, which constituted a greater proportion of females' diet costs (17% versus 14%), as did miscellaneous foods (5% vs 4%), whereas lower proportions of cost were exhibited amongst females for meat, fish, eggs and beans (29% vs 32%). Age groups differed in the proportional costs for starchy foods, fruit and vegetables, dairy, and foods high in sugar and/or fat.

As a result of these differences, proportional costs in the following sections are adjusted for age and sex.

Table 8.7 Age and sex differences in median proportional food group costs, including alcohol (p from Kruskal-Wallis ANOVA) (n=1014)

Food group	Sex			Age group						
	Male	Female	p value	19-29 years	30-39 years	40-49 years	50-59 years	60-69 years	70 years and over	p value*
Starchy foods	8%	8%	0.56	12%	11%	8%	7%	7%	7%	<0.001
Fruit & vegetables	10%	14%	<0.001	9%	12%	11%	13%	14%	15%	<0.001
Meat, fish, eggs & beans	23%	24%	0.75	25%	24%	22%	21%	25%	26%	<0.001
Milk and dairy	6%	7%	<0.001	5%	6%	6%	7%	7%	9%	<0.001
Foods high in sugar and/or fat	8%	8%	0.16	9%	8%	8%	6%	8%	10%	<0.001
Alcoholic beverages	15%	3%	<0.001	0%	10%	16%	15%	5%	0%	<0.001
Non-alcoholic beverages	9%	11%	<0.001	11%	10%	9%	10%	12%	10%	0.60
Miscellaneous foods	3%	4%	<0.001	3%	3%	4%	3%	3%	4%	0.18

Table 8.8 Age and sex differences in median proportional food group costs, excluding alcoholic beverages (p from Kruskal-Wallis ANOVA) (n=1014)

Food group	Sex			Age group						
	Male	Female	p value	19-29 years	30-39 years	40-49 years	50-59 years	60-69 years	70 years and over	p value*
Starchy foods	11%	10%	0.02	14%	12%	11%	9%	9%	8%	<0.001
Fruit & vegetables	14%	17%	<0.001	11%	15%	14%	17%	19%	17%	<0.001
Meat, fish, eggs & beans	32%	29%	<0.001	30%	28%	30%	29%	30%	30%	0.58
Milk and dairy	7%	8%	0.03	5%	7%	8%	8%	9%	11%	<0.001
Foods high in sugar and/or fat	11%	9%	0.19	11%	10%	10%	8%	9%	11%	<0.001
Non-alcoholic beverages	12%	13%	0.12	12%	12%	12%	13%	13%	12%	0.55
Miscellaneous foods	4%	5%	<0.01	4%	4%	5%	4%	4%	5%	0.19

* test for trend

8.4.3.3 Proportional costs by other sociodemographic variables

Table 8.9 and Table 8.10 show the age- and sex-adjusted proportional costs of each food group (geometric means) according to sociodemographic and lifestyle variables. There were no significant differences according to household size. For all other comparisons, significant differences ($p<0.01$) were found in at least one food group.

The proportion of diet cost from fruit and vegetables was found to differ between categories of all sociodemographic variables (employment, qualifications, equivalized income and marital status), regardless of whether alcohol was included: those with a higher income, higher qualifications, in managerial and professional positions, and those who were single had greater proportional costs for fruit and vegetables.

The greatest proportions for starchy foods were found amongst those who had the lowest incomes and those who had never worked, both when alcohol was included or excluded.

The proportion of cost attributed to meat, fish, eggs and beans was greater amongst those with no qualifications. When alcohol was included, it was also found to differ by employment category, with those in professional and managerial roles having lower proportional costs for this food group.

The dairy, high-fat and high-sugar and miscellaneous food groups did not show differences after costs from alcohol were discounted. However, when alcohol was included, dairy costs differed by marital status (with the lowest proportions amongst separated, divorced and widowed categories), miscellaneous costs differed by qualification (those with no qualifications having the highest proportion), and the high-fat/high-sugar costs differed by qualification (with lower proportions seen in those with A-levels and above) and equivalized income (where proportions decreased as income category increased).

Non-alcoholic beverages were not found to differ between any of the sociodemographic categories.

8.4.3.4 Proportional cost comparisons by lifestyle variables

The lifestyle variables compared were: achievement of ‘5 a day’, cigarette-smoking status and (when costs from alcohol excluded) alcohol consumption category.

Cigarette smoking was associated with lower proportional costs for fruit and vegetables, and for dairy foods. When costs from alcohol were included, those not currently smokers had higher proportional costs for starchy foods. When alcohol costs were excluded, this was no longer statistically significant, but non-smokers had lower proportional costs for meat, fish, eggs and beans.

Participants who achieved five portions of fruit and vegetables per day had almost double the proportion of diet cost attributed to fruit and vegetables. In addition, they were found to have lower proportions of cost attributed to meat, fish, eggs and beans, high-fat and high-sugar foods, and non-alcoholic beverages. When alcohol was excluded, there was also a difference between achievers and non-achievers in the proportion of cost contributed by starchy foods.

Finally, participants in the higher alcohol consumption categories were found to have higher proportional costs for the meat, fish, eggs and beans category, after discounting the costs from alcohol.

Table 8.9 Age- and sex-adjusted geometric mean proportional food group costs, including alcoholic beverages (p values from regression*) (n=1014)

Category	n	Starch (%)	p	F&V (%)	p	Meat, etc (%)	p	Dairy (%)	p	Hi-sugar & fat (%)	p	Non-alc beverage (%)	p	Misc (%)	p
Employment classification			<0.01		<0.01		<0.01		0.44		0.03		0.07		0.98
Managerial & professional	421	7		12		21		6		6		9		3	
Intermediate	302	8		11		22		6		7		10		3	
Routine & semi-routine	250	9		9		23		6		7		10		3	
Never worked & 'other'	41	10		8		25		5		8		11		3	
Equivalized income**			<0.01		<0.01		0.03		0.11		<0.01		0.90		0.03
Under £14,999	174	9		9		23		6		7		10		3	
£15,000 - £24,999	237	8		10		22		6		7		10		3	
£25,000 - £34,999	165	8		11		22		6		7		9		3	
£35,000 - £49,999	130	7		12		21		5		6		9		3	
£50,000 or more	169	7		13		20		5		6		9		3	
Qualifications***			0.02		<0.01		<0.01		0.99		<0.01		0.78		<0.01
Degree or higher ed	338	7		14		20		6		6		9		3	
GCE A- level or equivalent	172	8		12		21		6		6		9		3	
GCSEs or FT education	245	8		10		23		6		7		10		3	
No qualifications	251	9		8		24		6		7		10		4	
Marital status			0.24		<0.01		0.54		<0.01		0.25		0.89		0.05
Single, never married	289	8		12		22		6		7		10		3	
Married	467	8		11		22		6		7		9		3	
Married but separated	30	8		10		22		5		6		9		3	
Divorced	127	8		10		22		5		6		9		3	
Widowed	101	7		9		22		5		6		9		3	

Table 8.9 (cont'd) Age- and sex-adjusted geometric mean proportional food group costs, including alcoholic beverages (p values from regression*)

Category	n	Starch (%)	p	F&V (%)	p	Meat, etc (%)	p	Dairy (%)	p	Hi-sugar & fat (%)	p	Non-alc beverage (%)	p	Misc (%)	p
Household size			0.20		0.79		0.54		0.09		0.78		0.39		0.32
1 person	268	8		11		22		5		7		9		3	
2 people	336	8		11		22		6		7		9		3	
3 or 4 people	327	8		11		22		6		7		10		3	
5 or more people	83	8		11		21		6		6		10		3	
Cigarette-smoking status			0.01		<0.01		0.01		<0.01		0.81		0.59		0.87
Never regularly smoked	541	8		13		21		6		7		9		3	
Ex-regular smoker	247	8		10		22		6		7		10		3	
Current regular smoker	226	7		8		24		5		7		10		3	
Achieve '5 a Day'			0.01		<0.01		<0.01		0.03		<0.01		<0.01		0.78
Yes	334	7		19		19		6		5		8		3	
No	680	8		8		23		6		7		10		3	

*Adjusted in each case for age and sex. P values for overall effect of categorical variables. **Data missing for n=139. ***Data missing for n=8.

Table 8.10 Age- and sex-adjusted geometric mean proportional food group costs, excluding costs from alcoholic beverages (p values from regression*) (n=1014)

Category	n	Starch (%)	p	F&V (%)	p	Meat, etc (%)	p	Dairy (%)	p	Hi-sugar & fat (%)	p	Non-alc beverage (%)	p	Misc (%)	p
Employment classification			<0.01		<0.01		0.05		0.12		0.11		0.20		0.65
Managerial & professional	421	9		16		26		7		8		11		4	
Intermediate	302	10		13		27		7		8		12		4	
Routine & semi-routine	250	11		12		28		7		9		12		4	
Never worked & ‘other’	41	11		10		29		7		9		13		4	
Equivalized income**			<0.01		<0.01		0.67		0.61		0.11		0.24		0.16
Under £14,999	174	11		11		27		7		9		11		4	
£15,000 - £24,999	237	10		12		27		7		9		12		4	
£25,000 - £34,999	165	10		14		27		7		8		12		4	
£35,000 - £49,999	130	9		15		27		7		8		12		4	
£50,000 or more	169	9		17		27		7		8		13		3	
Qualifications***			0.09		<0.01		<0.01		0.46		0.01		0.69		0.02
Degree or higher ed	338	9		17		26		7		8		12		3	
GCE A- level or equivalent	172	10		15		27		7		8		12		4	
GCSEs or FT education	245	10		12		28		7		9		12		4	
No qualifications	251	10		10		29		7		9		12		4	
Marital status			0.58		<0.01		0.11		0.02		0.55		0.59		0.09
Single, never married	289	10		15		26		8		8		12		4	
Married	467	10		14		27		7		8		12		4	
Married but separated	30	10		13		28		7		8		12		4	
Divorced	127	10		12		28		7		8		12		3	
Widowed	101	10		12		29		6		8		12		3	

Table 8.10 (cont'd) Age- and sex-adjusted geometric mean proportional food group costs, excluding costs from alcoholic beverages (p from regression*)

Category	n	Starch (%)	p	F&V (%)	p	Meat, etc (%)	p	Dairy (%)	p	Hi-sugar & fat (%)	p	Non-alc beverage (%)	p	Misc (%)	p
Household size			0.17		0.78		0.51		0.07		0.76		0.31		0.28
1 person	268	9		13		28		7		8		11		4	
2 people	336	10		14		27		7		8		12		4	
3 or 4 people	327	10		14		27		7		8		12		4	
5 or more people	83	10		14		27		8		8		12		4	
Cigarette-smoking status			0.01		<0.01		<0.01		<0.01		0.48		0.11		0.60
Never regularly smoked	541	10		15		26		8		8		11		4	
Ex-regular smoker	247	10		13		28		7		8		12		4	
Current regular smoker	226	9		11		31		6		9		13		4	
Alcohol consumption			0.03		0.81		<0.01		0.02		0.05		0.57		0.02
Abstainers	410	10		14		26		8		9		12		3	
Lower risk	425	10		14		28		7		8		12		4	
Increasing risk	132	9		13		29		7		8		12		4	
Higher risk	47	9		13		31		6		7		12		5	
Achieve '5 a Day'			<0.01		<0.01		<0.01		0.06		<0.01		<0.01		0.59
Yes	334	9		24		24		8		7		10		4	
No	680	10		10		29		7		9		13		4	

*Adjusted in each case for age and sex. P values for overall effect of categorical variables. **Data missing for n=139. ***Data missing for n=8.

8.4.3.5 Food group costs by BMI category

In terms of absolute costs, significant differences were found between the normal weight, overweight and obese categories for starchy foods, and high-fat and high-sugar foods (Table 8.11). This indicates that the overweight and obese spent less on starchy foods and less on high-fat and –sugar foods.

Table 8.11 Median food group costs in the NDNS, by BMI category (unadjusted) (n=938)

BMI category	n	Starchy food	Fruit & veg	Meat, etc	Dairy foods	High-fat/ sugar foods	Non-alcoholic beverages	Misc
Underweight	13	£0.16	£0.22	£0.73	£0.14	£0.31	£0.25	£0.11
Normal weight	318	£0.32	£0.47	£0.78	£0.21	£0.30	£0.36	£0.13
Overweight	350	£0.28	£0.42	£0.89	£0.23	£0.28	£0.37	£0.12
Obese	257	£0.28	£0.42	£0.87	£0.21	£0.22	£0.34	£0.13
P for trend*		<0.01	0.034	0.041	0.466	<0.01	0.430	0.749

*Underweight excluded from analyses

Proportional costs differed between BMI categories only for the high-fat and high-sugar food group, adjusting for age and sex. This was true regardless of whether or not alcohol was included (Table 8.12 and Table 8.13), with the obese having less of their diet costs attributable to this food group.

Table 8.12 Food group costs in the NDNS as a proportion of daily diet cost (including costs from alcohol), by BMI category (n=938)

BMI category	Starchy food	Fruit & veg	Meat, etc	Dairy foods	High-fat/ sugar foods	Non-alcoholic beverages	Misc	Alcohol
Underweight	8%	12%	20%	6%	10%	9%	3%	18%
Normal weight	8%	11%	21%	6%	8%	9%	3%	19%
Overweight	8%	11%	22%	6%	7%	10%	3%	21%
Obese	8%	10%	23%	5%	5%	10%	3%	22%
P for trend*	0.310	0.099	0.053	0.371	<0.001	0.497	0.315	0.152

*Underweight excluded from analyses

Table 8.13 Food group costs in the NDNS as a proportion of daily diet cost (excluding costs from alcohol), by BMI category (n=938)

BMI category	Starchy food	Fruit & veg	Meat, etc	Dairy foods	High-fat/ sugar foods	Non-alcoholic beverages	Misc
Underweight	10%	15%	25%	8%	12%	11%	4%
Normal weight	10%	14%	26%	7%	10%	11%	4%
Overweight	10%	13%	27%	7%	8%	12%	4%
Obese	10%	13%	29%	7%	7%	12%	4%
P for trend*	0.308	0.091	0.031	0.358	<0.001	0.414	0.287

*Underweight excluded from analyses

8.4.4 Food group costs & BMI

Table 8.14 and Table 8.15 display the results of the linear regression analyses of food group costs and BMI. Whether costs from alcohol were excluded or not, there was a significant effect apparent from the proportion of diet cost attributed to foods high in sugar and fat: every additional percentage of cost from this food group was associated with a lower BMI of just over 9kg/m^2 (including alcohol 95% CI -14.73, -4.03, p<0.01; excluding alcohol 95% CI -14.20, -4.01, p<0.01).

Sensitivity analyses for the models including alcohol found similar results in all cases, although the significant coefficient for the high-fat and –sugar foods was smaller when excluding those on a special diet (n=833; b=-7.14; 95% CI -12.55, -1.72) or those over the age of 70 (n=794; b= -7.10; 95% CI -12.99, -1.21). When alcohol was not included, excluding these groups had a similar effect on the high-fat and –sugar food group’s coefficient (b=-7.06 (95% CI -12.22, -1.90) and -6.82 (95% CI -12.40, -1.12) respectively). All other estimates were found to be similar in the sensitivity analyses.

Table 8.14 Linear regression of proportional food group costs (including alcohol) on BMI (n=925)

	Coefficient	95% CI	p value	R^2
Model 1 Starchy foods	-0.17	-6.30, 2.97	0.48	0.028
Model 2 Fruit & vegetables	-0.89	-4.32, 2.54	0.61	0.027
Model 3 Meat, fish, eggs & beans	2.21	-0.60, 5.03	0.12	0.030
Model 4 Milk and dairy	-5.65	-11.80, 0.51	0.07	0.031
Model 5 Foods high in sugar and/or fat	-9.38	-14.73, -4.03	<0.01	0.039
Model 6 Non-alcoholic beverages	0.49	-2.86, 3.85	0.77	0.027
Model 7 Miscellaneous foods	-0.04	-0.35, 0.28	0.83	0.029
Model 8 Alcoholic beverages	0.69	-1.13, 2.50	0.46	0.028

Adjusted for age, sex, employment and energy intake

Table 8.15 Linear regression of proportional food group costs (excluding alcohol) on BMI (n=925)

	Coefficient	95% CI	p value	R^2
Model 1a Starchy foods	-0.98	-5.26, 3.31	0.67	0.027
Model 2a Fruit & vegetables	-0.69	-3.93, 2.56	0.68	0.027
Model 3a Meat, fish, eggs & beans	2.94	0.39, 5.48	0.02	0.032
Model 4a Milk and dairy	-5.07	-10.75, 0.62	0.08	0.030
Model 5a Foods high in sugar and/or fat	-9.10	-14.20, -4.01	<0.01	0.040
Model 6a Non-alcoholic beverages	0.51	-2.52, 3.54	0.74	0.027
Model 7a Miscellaneous foods	3.86	-1.89, 9.61	0.19	0.029

Adjusted for age, sex, employment and energy intake

8.4.5 Food group costs & overweight and obesity

The adjusted logistic regression models (including alcohol) identified significant effects on the odds of being overweight or obese for the fruit and vegetable food group (OR 0.09, 95% CI 0.02, 0.38, p<0.01) (see Table 8.16). This implies a 91% reduction in the odds of being overweight or obese for every additional 1% of diet cost attributed to fruit and vegetables. A similar odds ratio was evident in the model excluding alcohol

(OR 0.11, 95% CI 0.03, 0.45, p<0.01) (Table 8.17). Additionally, in the analyses excluding alcohol, there was evidence of a significant effect of the meat, fish, eggs and beans food group (OR 5.59, 95% CI 1.85, 16.89, p<0.01), suggesting a five times increase in the odds of being overweight or obese for every additional per cent of cost. No other food group costs were found to have a significant impact on the odds.

Table 8.16 Adjusted logistic regression of food group costs (including alcohol) on the odds of being classified overweight or obese (n=925)

Food group cost (studentised residuals)	Odds ratio	95% CI	p value
Model 1 Starchy foods	0.29	0.04, 1.93	0.20
Model 2 Fruit & vegetables	0.09	0.02, 0.38	<0.01
Model 3 Meat, fish, eggs & beans	3.67	1.09, 12.37	0.04
Model 4 Milk and dairy	0.32	0.02, 4.21	0.39
Model 5 Foods high in sugar and/or fat	0.07	0.01, 0.64	0.02
Model 6 Non-alcoholic beverages	1.37	0.32, 5.80	0.67
Model 7 Miscellaneous foods	4.28	0.23, 78.67	0.33
Model 8 Alcoholic beverages	1.82	0.83, 3.98	0.14

Adjusted for age, sex, employment and energy intake

Table 8.17 Adjusted logistic regression of food group costs (excluding alcohol) on the odds of being classified overweight or obese (n=925)

Food group cost (studentised residuals)	Odds ratio	95% CI	p value
Model 1a Starchy foods	0.41	0.07, 2.41	0.33
Model 2a Fruit & vegetables	0.11	0.03, 0.45	<0.01
Model 3a Meat, fish, eggs & beans	5.59	1.85, 16.89	<0.01
Model 4a Milk and dairy	0.45	0.04, 4.87	0.51
Model 5a Foods high in sugar and/or fat	0.10	0.01, 0.84	0.03
Model 6a Non-alcoholic beverages	1.75	0.47, 6.49	0.41
Model 7a Miscellaneous foods	7.39	0.56, 98.24	0.13

Adjusted for age, sex, employment and energy intake

Excluding those who reported consuming an unusual amount resulted in odds ratios for fruit and vegetables that were no longer significant (including alcohol) (n=419; OR 0.12, 95% CI 0.02, 1.02, p=0.05). The other food groups had similar ORs and probability values, with the exception of the miscellaneous food group: OR 344.19, 95% CI 1.70, 69850.28, p=0.03.

When costs from alcohol were not included, each of the sensitivity analyses revealed a similar pattern to when alcohol was included.

8.5 Discussion

This chapter builds upon the findings presented in Chapters 6 and 7, describing the diet costs of NDNS adults in terms of constituent food groups. The primary aim was to investigate how proportional food group costs relate to participants' BMI, and the chapter also explored the newly derived variables' relationships to each other and to diet costs as a whole, as well as sociodemographic differences in these costs.

There are no directly comparable national food group expenditure data with which to compare the figures estimated here for the NDNS. However, the Family Food report (Defra, 2012) describes expenditure in pence per person for a number of food categories. Combining foods to best match the food groups used here, and calculating the expenditure as a proportion, gives the estimates presented in Table 8.18. The only two categories which offer a direct match in description to this chapter's food groups are fruit and vegetables, and alcoholic beverages. With the exception of non-alcoholic beverages and foods high in sugar and/or fat, expenditure figures estimate higher proportional costs than the dietary data estimated costs. This may reflect the waste associated with these other food groups, which would result in lower quantities actually being consumed. This remains conjecture, however, in the absence of a more controlled comparison. Interestingly, the proportional estimates given to meat, fish and eggs are similar, despite the differences in the categorisation.

Table 8.18 Proportional food group costs estimated from NDNS diets compared to national expenditure data*

NDNS food group	% diet cost excl alcohol	% diet cost incl alcohol	Defra equivalent food groups	% of total food & non-alcoholic drink expenditure	% of total food & drink expenditure
Starchy foods	10	8	Potatoes & cereals	23	21
Fruit & vegetables	15	12	Fruit & vegetables	18	16
Meat, fish, eggs & beans	29	24	Carcase meat, non-carcase meat & meat products, fish & eggs	29	26
Milk and dairy	8	6	Milk & cream, & cheese	11	10
Foods high in sugar and/or fat	10	8	Fats & oils, sugar & preserves, & confectionary	7	6
Alcoholic beverages	-	9	Alcoholic beverages	-	11
Non-alcoholic beverages	13	10	Beverages & soft drinks	6	5
Miscellaneous	4	3	No equivalent	-	-

*Figures derived from Family Food 2010 data (Defra, 2012)

Although not necessarily a reflection of actual food expenditure, the estimates presented in this chapter deliver insight into the food spending patterns of the participants. Proportional costs could give an indication of the share of their food budget people are willing to apportion to different types of food.

These results are the first to examine the food group costs from dietary data of a nationally representative sample. All of the other similar studies have published food group costs from non-representative samples (Cade et al., 1999, Murakami et al., 2007, Ryden et al., 2008, Ryden and Hagfors, 2011, Alexy et al., 2012), most of which contrasted results from different dietary patterns. Identifying patterns in diet costs from a representative sample is advantageous in terms of judging the implications of public health interventions.

8.5.1 Food group costs in the NDNS

The biggest driver of whole diet costs in this sample was costs attributed to alcoholic beverages (with a sample median of 9% proportional costs). The decision to include or exclude alcohol from proportion calculations modified the results of comparisons; this suggests that this decision could have an important impact in the interpretation of results, which is something that research in this area should take into account.

The impact of alcohol on the results also highlights its role in the dietary expenditure of British adults. From these results, it could be inferred that people allocate potentially significant amounts of their budgets to purchase alcoholic beverages. This is something that would need to be addressed in any budget-focussed intervention. It would be interesting to see the effects of, for example, a minimum unit pricing policy, if introduced, on the proportional cost for alcohol.

After alcohol, non-alcoholic beverages, meat, fish, eggs and beans and fruit and vegetables formed the largest contributors to diet costs. This agrees relatively well with the only other previously published food group cost estimates for British adults (Cade et al., 1999), in which the costs of both the healthiest diets (according to the Healthy Diet Index) and the least healthy diets in the UK Women's Cohort Study (UKWCS) were predominantly made up from these three food groups. The full sample estimates were not presented in the paper; however the proportional food group costs for the least healthy participants map well onto the NDNS patterns observed here, with the exception that NDNS adults had lower costs attributed to fruit and vegetables (15% compared to 29% in the UKWCS lowest HDI group) and higher costs from alcohol (9%

versus 7%). These differences were more marked when comparing the results of the healthiest UKWCS diet group. This may be explained by sample differences: the NDNS is a nationally representative sample, whereas the UKWCS drew from a comparatively healthy, older age group of women, over-sampled for vegetarians.

In other cultures, too, parallels can be drawn for proportional food group costs. Murakami et al (2007), for example, found the largest contributors to diet cost in a sample of female Japanese students to be meat, fish and shellfish (32%), followed by vegetables (16%) and confectionaries (12%). This compares to the NDNS estimates of 29% for meat, fish, eggs and beans, 15% for fruit and vegetables, and 10% for high-fat and high-sugar foods. Those in the highest quintile of diet cost in Murakami et al's study had almost four times as much of the cost attributed to fish and shellfish, four times as much for vegetables, and three times as much for fruit, suggesting that these groups may have been important drivers of total diet cost in this sample.

In a slightly different approach, Ryden et al (2008) compared food group costs of two trial arms: the control group, and those who had received the intervention of modifying their diet to a Mediterranean diet. It was found that, proportionally, the Mediterranean diet group had greater costs attributed to fish, followed by vegetables and fruit. The control group, on the other hand, showed higher proportions for meat, then dairy foods, then beverages.

Finally, in a sample of German children, Alexy et al. (2012) found the greatest proportional costs to be in the meat/sausage category (16%) followed by dairy (15.8%) and convenience/fast food (11%). Proportional costs were presented separately for fruit and vegetables in this study, with a mean 6.5% and 7.8% of diet costs attributed to these food groups respectively. Proportional costs for confectionary were fairly low in this sample, at just over 6% on average.

There are several points to be made from these cross-cultural comparisons. Firstly, the use of different food groups makes it difficult to directly compare samples: the studies outlined above described a greater number of food groups than used for the NDNS analyses. It is possible that collapsing foods to eight groups, whilst useful in interpreting the costs in relation to UK recommendations, could have resulted in a loss of information. From the analyses presented in this chapter, it is unclear whether those with higher proportional costs for the meat, fish, eggs and beans group had higher costs for fish or higher costs for meat, for example. Despite this, it appears that there are commonalities across countries (specifically the UK, Sweden and Japan) in that the key contributors to diet costs appear to be meat, fish and shellfish, and vegetables.

The second point is that proportional costs vary according to dietary patterns. This was an aspect of food group costs that was not explored fully in this chapter, but could have important implications for dietary interventions. The healthiest diet group of the UKWCS, for example, had almost 50% of diet costs attributed to fruit and vegetables. Likewise, in the NDNS, those who achieved the ‘5 a day’ recommendation had significantly higher proportional costs for fruit and vegetables (although not as high as the UKWCS, at 24%) as well as differences in a number of other food group costs. This hints at a shift in the proportions of expenditure and not just an increase in fruit and vegetable costs in order to meet the recommendation.

The observations that food group costs differ by diet quality leads to a question of how proportional costs relate to overall diet costs. In this chapter, multivariate regression was used to identify the key drivers of whole diet costs. It may also have been interesting to have examined proportions of food group costs by strata of whole diet costs. As described above, the food group analyses indicated much higher proportional costs were attributed to fruit and vegetables amongst those who achieved 5 a day; in Chapter 6 it was described that those who achieved 5 a day also showed significantly higher whole diet costs. It could be the case that fruit and vegetable proportional costs are systematically related to whole diet costs. Examining food group costs according to, for example, quintiles of whole diet costs might illustrate key differences in the make-up of diet costs, according to their estimated worth.

It is also possible to compare the proportional food group costs estimated here with proportional costs found in national expenditure data. In the 2010 Family Food report (Defra, 2010), it is apparent that the largest proportion of household food expenditure (excluding ‘eating out’ expenditure) was meat and meat products, fish and eggs (combined), followed by cereals, and fruit and vegetables (these two groups showing a similar proportion), then alcoholic beverages. The differences seen between these expenditure data and the results of this chapter – in particular the differences between proportions of cost attributed to alcoholic beverages and non-alcoholic beverages (which are both much lower in the expenditure data) – could be explained in a couple of ways. Firstly, it may be the case that promotional and lower-than-median cost beverage items are purchased to a greater extent than can be accounted for using the DANTE cost database. Secondly, eating out purchases are reported separately in the national expenditure data, whereas they were treated as at-home purchases in the costing methods of this chapter, rather than being excluded altogether. Beverages comprise a food group that are commonly consumed outside the home (Defra, 2010).

The contrast between proportions of cost, energy and mass from each of the food groups has not been presented before. The differences between the three suggest that proportional food group costs do not simply reflect the quantities of each food group consumed, and can be seen to be a useful construct.

8.5.2 Between-group comparisons of food group costs

None of the previously published studies reported contrasts between sociodemographic categories, making these NDNS results the first to do so. All of the variables examined except household size showed statistically significant differences in proportional costs in at least one of the food groups. Comparisons of whole diet costs similarly found differences in these variables, although not for age groups (Chapter 6).

Proportional costs for fruit and vegetables showed the most number of significant between-group differences in the analyses by socioeconomic and other variables. This is in keeping with reported consumption differences of this food group from other sources – for example, by income (Defra, 2009).

Males and females were found to differ in their proportional costs for fruit and vegetables and meat, fish, eggs and beans (after excluding alcohol). This may help to explain the differences in whole diet costs described in Chapter 6, as a higher proportion of females' diet costs was attributed to fruit and vegetables, one of the less energy dense food groups.

Interestingly, proportional costs between equivalized income groups differed only for the starchy foods, fruit and vegetables and (when alcohol was included) high-fat and high-sugar food groups. However, comparisons of whole diet costs (Section 6.4.3) found a significant trend in costs with increasing income categories. One interpretation of this is that not only were the higher income groups spending more on fruits and vegetables, they were also spending proportionately more of their budget on this food group.

The food groups which showed significantly different proportional costs were not the same for each of the socioeconomic indicator variables (occupation, income and education). Significant trends in whole diet costs were apparent for all these variables (see Chapter 6), but the results presented here imply that each variable influences diet costs in different ways – for example, whole diet costs were found to differ significantly between employment categories (Table 6.2), but on comparing proportional food groups costs, only the proportions of cost attributed to starchy foods and to fruit and vegetables were found to differ, whilst proportions of all other food group costs did not differ by employment.

Where food group cost analyses contrasted most to those of whole diet costs, however, was in comparing age groups: age group differences were apparent between the proportional costs of most food groups. In examining whole diet costs, no significant trends were apparent. This suggests that the whole diet estimations are missing an important level of detail, and food group costs are more descriptive of diet costs.

Also in contrast to whole diet costs was the linear trend in fruit and vegetable proportional costs by marital status. The age- and sex-adjusted means suggest that the single category had higher costs attributed to fruit and vegetables, with lower proportions observed amongst the married, separated and divorced, and the lowest amongst the widowed. This is interesting in that it does not appear to match the patterns observed for whole diet costs, for which the married had the highest costs. The reasons for this are unclear, although it might be useful to examine whether socioeconomic status is unevenly distributed by marital status – for example, only 12% (n=10) of the widowed participants in the NDNS were in the highest two equivalized income categories (compared to 26-36% of the other marital categories).

Differences in proportional food group costs according to cigarette smoking status perhaps support the suggestion made in Chapter 7 of behaviour clustering. The differences in apportioning the diet costs imply that whole diet cost differences are not simply the result of a reduced food budget due to cigarette purchasing.

8.5.3 Food group costs & BMI

Whereas whole diet cost estimates revealed no significant differences between BMI categories, examination of proportional food group costs showed some significant trends. In terms of absolute costs, tests for trend indicated that the obese spent the least on starchy foods and on foods high in sugar and fat. As a proportion of whole diet costs, only the high-fat and high-sugar food group costs were found to differ significantly by BMI category. This was also the only significant predictor of BMI in the linear model.

The different findings for the absolute costs and proportional costs of the starchy food group illustrate the impact of expressing costs as a proportion of whole diet costs: whilst lower costs were apparent amongst the obese for starchy foods, the proportion of total diet cost was found to be the same across BMI categories. It is possible that the lower costs attributed to high-fat and high-sugar foods found in this chapter are indicative of under-reporting or measurement error. Unfortunately there is no other literature available with which to compare these proportional cost findings, as this is the first study of this kind. Nor are there physical activity data available for the NDNS to evaluate the presence of under-reporting (see Chapter 3).

There is no consensus in the literature as to whether certain food groups in particular are prone to be under-reported. Some studies (eg Krebs-Smith et al., 2000) suggest that under-reporting is uniform across dietary composition. Bailey et al (2007) similarly found lower intakes reported amongst identified under-reporters for a majority (18 of 24) of food groups. On the other hand, contrasting reports have identified food group-specific mis-reporting (Mendez et al., 2003): Lafay et al (2000), for example, found significantly less reported intakes in under-reporters (both in terms of frequency and portion sizes) specifically for foods high in fat and/or carbohydrate.

The sensitivity analyses found that excluding special dieters resulted in a reduced coefficient for high-fat and –sugar foods. It may be conjectured that this is as a result of the removal of some bias, if dieters are more likely to under-report items from this food group. Excluding the elderly had a similar effect, though the reasons for this are unclear. In both cases, the coefficients were still found to be statistically significant.

Whilst it can be conjectured that under-reporting has influenced the investigations of BMI, it is important at the same time to consider other interpretations of these findings. As well as issues potentially arising from dietary measurement error, there may be additional methodological issues in the approach taken to express food group costs – for example, in the expression of food groups as a percentage, the definition of the food groups, or the use of mean food item costs in characterising food budget allocation. Comparisons to national expenditure data (see above) show a number of discrepancies that are difficult to explain. All in all, it must be stressed that these early explorations into food group costs warrant further attention.

8.5.4 Food group costs & overweight and obesity

In the logistic regression analyses, there was no significant impact on the odds of being overweight or obese from the high-fat and high-sugar food group. Instead, significantly reduced odds were apparent with increasing proportions of cost from fruit and vegetables. When alcohol was excluded, there was also a significant effect from the meat, fish, eggs and beans food group, which showed more than five times the odds for every additional percentage of diet cost.

These findings suggest that assigning a greater proportion of the food budget to fruit and vegetables is protective against weight gain. Conversely, a greater proportion attributed to meat, fish, eggs and beans was associated with excess weight. The latter observation was not statistically significant when costs from alcoholic beverages were included in the calculations. This may be because of the relationship in this sample between alcohol consumption and proportional costs for the meat, fish, eggs and beans food group: those who consumed more units of alcohol had higher proportional costs without taking into account alcohol costs (Table 8.10), but proportions were lower when alcohol was included because increasing proportions of the whole diet cost were given to alcoholic beverages, consequently reducing proportions for other food groups.

The story implied by the logistic regression differs to that suggested by the linear regression models, with different food groups implicated. This could be because the overweight were combined with the obese participants, perhaps outweighing any under-reporting amongst the obese – specifically with respect to high-fat/sugar foods.

It is unclear why the sensitivity analysis in which those who reported consuming an unusual amount were excluded failed to achieve significance for fruit and vegetables and the meat food group. It is possible that this was a result of reduced statistical power following the removal of almost half the sample number.

A final caveat needs to be made with respect to the odds ratios given by these models. The models investigating fruit and vegetables reported ratios of 0.09 (including alcohol) and 0.114 (excluding alcohol). These imply a huge reduction (of 81% and 79%) in the odds of being overweight and obese for every additional percentage of diet costs spent on this food group. Excluding alcohol, the odds ratio for the meat food group was similarly extreme, showing more than five times an increase in the odds. These ratios are difficult to interpret, and it is advisable that the conclusions from these findings are tentatively made, reflecting exploratory investigations of a new approach.

Again, these anomalous findings might bring into question the appropriateness of this methodology. As mentioned in the discussion of the unexpected negative association between costs from high-fat and –sugar foods with BMI (see 8.5.3), it is unclear whether these findings are a reflection of measurement error or in fact reveal a methodological limitation in the way food group costs have been expressed. For example, given the different distributions of the proportional costs of each food group, a percentage point change might represent a disproportionately large change in a food group such as the miscellaneous foods, to which only a small proportion of diet costs were attributed, but signify a much smaller difference in a food group with more variability in proportional costs, such as fruit and vegetables. Investigating food group costs separately in the regression models also limits our interpretation, in that it is difficult to interpret how the percentage change in one food group corresponds to the necessary opposite change in the other food groups – the model might assume a corresponding change equally distributed across the other food groups, but whether this is a realistic scenario is questionable, especially given the varying distributions of the proportional food group costs.

8.5.5 Implications

These food group costs do not represent actual expenditure of NDNS adults. Nevertheless, they offer potential insight, albeit tentatively, into how people divide their food budgets.

This insight into budgeting could have implications for public health interventions. Firstly, it can be used to communicate how a healthy diet can be achievable without having to increase your budget: it is not how much you spend, rather how you assign your budget. For example, the results imply that achieving the '5 a day' recommendations entails reassigning the proportions usually spent on other food groups to fruit and vegetables. Studies suggest that perceptions of expense are a deterrent of dietary change (Mushi-Brunt et al., 2007), so emphasising dietary improvement without an increased food expenditure could be important. This could be both positive and negative from a public health perspective: whilst the message that a healthy diet is achievable within current budgets is positive, reapportioning the food budget may be a more complex message to convey.

Describing the proportional food group costs could also be useful in forecasting the impact of proposed fiscal interventions. With the help of price elasticities (both own- and cross-price), the effect of targeted subsidies or taxes, for example, could be predicted with regards to proportional costs as well as whole diet costs. Having some idea of how budget apportioning differs amongst socioeconomics groups could aid in identifying where interventions could have differential effects: this would be possible with elasticities specific to income groups which have been published (see Nnoaham et al., 2009, for example). The results presented here show age and sex differences in proportional food group costs, suggesting that fiscal policies would affect males and females and different age groups disproportionately.

Food choices are a product of several factors, including culture, lifestyle choices and taste preferences (Steptoe et al., 1995, Connors et al., 2001). Cost, as well as being a limiting factor, when expressed in relation to the food budget might give an indication of willingness to pay. The proportion of the food budget assigned to a particular food group could be seen to be a reflection of values – in other words, it could indicate whether certain foods are perceived to be worth their cost.

The above illustrates some of the potential implications of findings from food group cost investigations. It is important to note, however, that the findings of this chapter are chiefly to present a fresh methodology for diet cost investigations. As mentioned throughout this discussion (and in particular in Section 8.5.6 below), a number of anomalous findings in the results of this Chapter indicate that, whilst the approach has potential, the methodology requires further attention in order to confidently interpret investigations of this kind. For example, it would be interesting to perform a comparison study of DANTE-estimated food group costs against calculations of food group expenditure collected from purchasing data collected at the same time as the dietary data. Anthropometric data on the same sample would be additionally useful to test the associations of both expenditure and estimated diet costs with BMI.

8.5.6 Limitations

Whilst proportional food groups offer a new perspective on diet costs, the underlying methods are essentially the same as in Chapters 6 and 7. As such, the same limitations in methodology apply. These are discussed in previous chapters, but the main points are repeated below.

Firstly, the NDNS is cross-sectional, and any conclusions as to causality are restricted. Further to this, without actual expenditure data, the costs reported here are estimated, and refer to the inherent value of the diets, rather than actual expenditure.

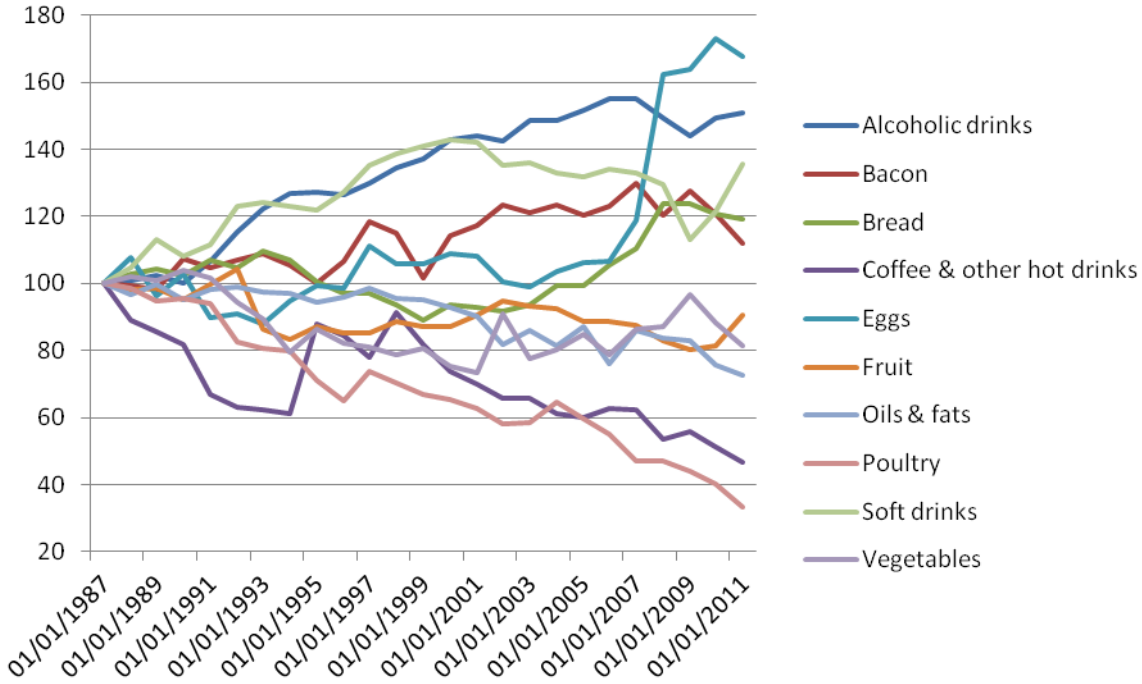
Secondly, the costing method uses only median prices from the database, whereas participants were most likely faced with a wide variety of prices which would influence budgeting decisions. The food cost database cannot take into account price discounts or promotions or the consumption of free food, nor can it estimate the costs of foods purchased and eaten away from home.

Thirdly, the costing method is likely to echo any measurement error associated with the dietary assessment itself. The NDNS relies upon the self-reported intakes of diet diaries, which are subject to mis-reporting (both conscious or non-conscious) as well as behaviour change in response to the assessment. The issue of bias as a result of suspected under-reporting in relation to this chapter's results is discussed above. Further research is recommended in which physical activity or metabolic measurements could be used to identify mis-reporters and determine how mis-reporting related to food group costs.

Another potential limitation of this investigation is the definition of the food groups. These were selected on the basis of UK recommendations (Section 8.3.2); however all other similar studies in the literature used a larger number of narrower food groups. It is possible that combining foods into eight groups resulted in some detail being lost. For example, costs for meat products and for fish products showed different patterns in the study by Ryden et al (2008), yet these were combined along with eggs and beans in the current analyses. At an even narrower level of detail, longitudinal studies suggest that there are even differential effects of meat consumption on BMI by type of meat (Gilsing et al., 2012).

Finally, the costs presented in this chapter were calculated using the original prices of the DANTE cost database, from 2004. Whilst proportional costs would remain the same after applying the FPI or RPI to account for inflation (which would be a flat rate applied to all food groups), it is possible that different results would be obtained with food group-specific inflation indices, of which the ONS publishes 27 (see Section 6.4.5). Figure 8.2 offers an illustration of how food group prices do not increase at an equal rate. This was not addressed in this chapter, and warrants further investigation.

Figure 8.2 Price inflation in the UK of 10 food groups between 1987 and 2011, relative to the Food Price Index (FPI=100) (data from ONS, 2011)



8.5.7 Strengths

Like the two chapters that precede it, this chapter makes an important contribution to the field in that it is the first attempt to describe at this level of detail the diet costs of a nationally representative British sample. It also shows a more thorough analysis of food group costs – both in relation to each other and in relation to whole diet costs – that has been absent in the literature. This new approach could have additional policy relevance compared to research into whole diet costs or food expenditure.

8.6 Conclusions

Estimating the costs of food groups as a proportion of whole diet costs is a little researched avenue of investigation, yet one that is potentially insightful. The results presented here add to a small international literature base, with the analyses constituting the first conducted with a nationally representative sample and the first to examine sociodemographic patterns.

One of the key objectives of the chapter was to assess the appropriateness of food group costs as a means of quantifying dietary costs – in particular whether this method adds value to a traditional whole diet cost approach. The findings contrast with those of Chapters 7 & 8, suggesting that food group costs confer additional information. However, the linear regression analyses found a negative association between BMI and proportional costs of high-fat and high-sugar foods, suggesting a protective effect of this food group cost. This finding perhaps supports the conjecture that there is evidence of mis-reporting amongst participants of higher BMI. In the logistic regression, where the overweight and obese categories were combined, the significant effect of high-fat and high-sugar food group costs was no longer apparent. Conversely, anomalous results might be reflection of methodological issues associated with the food group costing rather than dietary measurement error. These results highlight the need for physical activity or metabolic data in future dietary research in order to be able to account for mis-reporting.

The implication of these findings is that dietary change could be achieved by readdressing how food budgets are divided, rather than by incurring additional cost. This has the potential for a more acceptable public health message in addressing health inequalities. Food group costs could also provide a means of modelling the effects of targeted fiscal policies on different sociodemographic groups.

It would have been inappropriate to assess food group costs in relation to dietary energy density (see Section 8.2). However, it would be an interesting topic for further research to see how proportional food group costs varied in this sample according to some other indicator of dietary quality.

Despite the methodological limitations, this initial exploration into the proportional food group costs of NDNS adults has uncovered some interesting results. This suggests that costing diets in this manner could have some use in future research into diet costs.

What was known previously:

- In the NDNS, whole diet costs do not appear to be associated with BMI (Chapter 7). This adds to the already conflicting findings in the literature.
- As an emerging research area, the best available method for investigating the monetary aspects of diet is yet to be established.
- Proportional costs could give an indication of the share of a food budget people apportion to different types of food.
- A few studies have reported diet costs by food group, and suggested that proportional costs vary by dietary patterns.
- There is nothing in the literature comparing proportional food group costs by sociodemographic variables, nor by BMI category.

What this chapter adds:

- These results are the first to examine the food group costs of a nationally representative British sample.
- Presented is the most thorough analysis to date of food group costs – in relation to each other, in relation to whole diet costs, and according to sociodemographic status.
- Foods in the meat, fish, eggs and beans category were found to be responsible for the greatest proportions of diet costs, but alcoholic and non-alcoholic beverages were the strongest determinants of whole diet costs.
- All of the variables examined except household size showed statistically significant differences in proportional costs in at least one of the food groups. The proportional food group cost which showed the most differences was that of fruit and vegetables.
- In contrast to analyses using whole diet costs, significant associations with BMI and overweight/obesity were apparent, suggesting that normal weight, overweight and obese individuals apportion their food budget differently.
- A negative association was apparent between BMI and proportional costs of high-fat and high-sugar foods. This could constitute evidence of mis-reporting amongst participants of higher BMI.
- In logistic regression analyses, associations with overweight/obesity were found for fruit and vegetables (negative) and meat, fish, eggs and beans (positive), but the odds ratios are difficult to interpret, and conclusions are tentative.
- The implication is that it is not the food budget per se that encourages positive energy balance, but rather how people apportion their budget. This has relevance for public health messages.

Chapter 9 Discussion & conclusion

9.1 Introduction

The aim of this thesis was to examine whether income and cost of diet are implicated in excess energy intake, using data from the representative UK dietary survey, the National Diet and Nutrition Survey (NDNS). In this final chapter, the findings from the previous chapters are drawn together and collectively considered in light of this aim. The conclusions from the work will also be discussed in relation to the food price-obesity hypothesis (Section 9.3) which provided the motivation for these analyses.

Policy implications are particularly pertinent in this area of research, due to the role that fiscal interventions (such as taxation or subsidisation) could play in manipulating the pathway between purchasing power and food choice. However, given that the findings from this thesis are largely exploratory and a key theme that has emerged is the methodological difficulty of researching this area, the discussion below will focus primarily on the implications for researchers.

9.2 Summary of research findings

To meet the main aim (reiterated above), the work of this thesis was divided into meeting nine objectives. A summary of the findings that meet each objective are summarised in turn below.

1. To synthesise the published evidence linking food prices or diet costs with dietary energy density (DED) or weight status.

Chapter 2 presented the results of a semi-systematic review of the literature to meet this objective. The key findings included:

- Studies of diet costs and DED overwhelmingly reported a strong negative association.
- The evidence linking income and DED was less strong: two of the three studies amongst adults found evidence of lower DED with higher incomes. Amongst children, evidence of a link was not apparent.
- No studies were found to investigate food prices and DED.
- Published reviews suggest that, in highly developed countries, income is related to body weight amongst women but not men: as women report higher incomes, they are more likely to report a lower body mass.

- The limited number of studies investigating diet costs and BMI in adults reported contradictory findings. Quality of these studies varied. Associations between diet costs and BMI were not apparent amongst children.
- Studies investigating food prices and BMI varied widely in approach.
    - All of the studies testing prices of individual food items found significant associations for at least some foods, though these studies are problematic to synthesise.
    - There were mixed findings reported for fast food prices and body weight.
    - In adults, findings for food-at-home indices were dependent upon analytical approach; in children, no significant associations were apparent.
    - In terms of fruit and vegetable indices, three of five studies in children found a significant positive association; in adults, a significant positive association was found only for certain subgroups.
    - Only one of the three studies investigating the effect of soft drink taxes in children found a significant effect of soda taxes on body weight, and one of two studies in adults.

The overall conclusion of Chapter 2 was that the evidence – amongst adults, but not children – is generally supportive of the food price-obesity theory.

2. To examine the relationship between income and BMI or overweight/obesity amongst NDNS adults.

Chapter 4 addressed this second objective. The results indicated that:

- Income is negatively and linearly associated with BMI amongst NDNS adults, including both men and women.
- The odds of being overweight or obese are significantly lower with increasing income bands.
- Obese adults in the NDNS have a lower median equivalized income than those who are normal or overweight.
- The use of household income can result in different findings and interpretations compared to when equivalized household income is used.

3. To assess whether income is related to DED amongst NDNS adults.

Chapter 4 also presented findings for the third objective, concluding that equivalized household income is negatively and linearly associated with DED in the NDNS.

4. To investigate the appropriateness of diet cost estimations, including the costing of food groups.

This objective was inherent in Chapters 5, 6 and 8. From the findings, the following statements about the appropriateness of diet cost estimations can be made:

- The DANTE food cost database linked to a dietary assessment tool agrees well with estimates from household expenditure at a sample level – at the individual level or amongst higher spenders, diet cost estimates agreed less well (Chapter 5).
- In testing whole diet costs, there was little difference in using a flat rate of inflation compared to the food group-adjusted indices; although age groups may have been unevenly affected by the price changes of certain food groups (Chapter 6).
- Analyses using proportional food group costs produced different results to those using whole diet costs, suggesting that assessing how people apportion their food budget, rather than how much they spend on food, may be more useful (Chapter 8).

5. To estimate and describe the diet costs of NDNS adults.

Objective 5 was met by two key chapters: Chapter 6 described whole diet costs, expressed as daily (£ d^{-1}) or energy-adjusted (£ $10MJ^{-1}$) amounts; Chapter 8 described the costs of eight constituent food groups. The findings indicated:

- A median daily diet cost of £2.84 (IQR £2.27, £3.64) and a median energy-adjusted cost of £4.05 (£3.45, £4.82).
- Better quality diets, as signified by the consumption of fruit and vegetables, were of higher intrinsic monetary worth, even after adjusting for other economic and demographic factors.
- ➢ Foods in the meat, fish, eggs and beans category tended to account for the greatest proportion of whole diet costs.
- ➢ Beverages – alcoholic and non-alcoholic – were the strongest determinants of whole diet costs.

6. To explore patterns in NDNS diet costs according to sociodemographic characteristics.

The sixth objective of this thesis was addressed by both Chapters 6 and 8:

- ➢ Patterns in whole diet costs substantiated speculated sociodemographic differences, such as by income.
- ➢ In terms of food group costs, statistically significant differences were apparent for all of the sociodemographic variables (except household size) in at least one of the food groups. The proportional food group cost which showed the most differences was that of fruit and vegetables. Interpretation of the findings of this new approach is tentative.

7. To determine whether an association exists between diet costs and BMI or overweight/obesity amongst NDNS adults.

The findings from Chapters 7 and 8 suggest that there is no association – linear or non-linear – between whole diet costs and BMI in the NDNS. In contrast, analyses using proportional food group costs revealed some significant associations:

- A negative association was apparent between BMI and proportional costs of high-fat and high-sugar foods.
- A negative association was found between proportional costs of fruit and vegetables and overweight/obesity.
- A positive association was seen between proportional costs of meat, fish, eggs and beans and the odds of overweight-obesity, but the odds ratios are difficult to interpret.

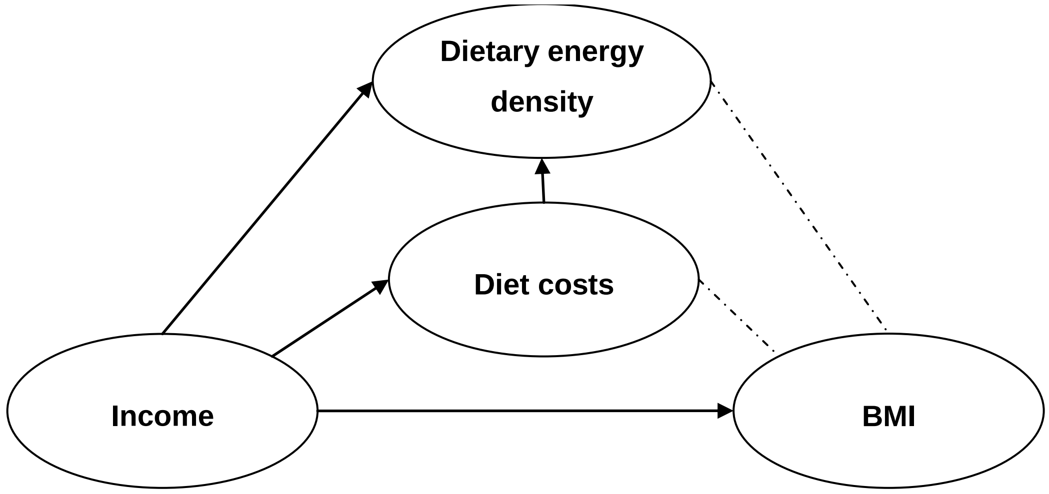
8. To establish whether an association exists between diet costs and DED amongst NDNS adults.

Chapter 7 examined this objective and the findings indicated a strong negative association between diet costs and DED that is not due to mathematical artefact.

9.3 Revisiting the 'food price-obesity hypothesis'

This section addresses the final objective outlined in Chapter 1. The findings from Chapters 4, 6 and 7 are broadly illustrated in Figure 9.1. The diagram indicates partial support for the food price-obesity hypothesis (see Figure 1.1, Section 1.5), in that income was found to be related: firstly, to diet costs, which is in keeping with the purported role of income in determining purchasing power; secondly to dietary energy density, implying that lower incomes encourage more energy dense diets; and thirdly to BMI, implicating this demand-side factor in obesity prevalence. In addition, diet costs – which theoretically reflect purchasing power – were negatively associated with DED.

Figure 9.1 Associations in the NDNS between key variables of the food price-obesity hypothesis

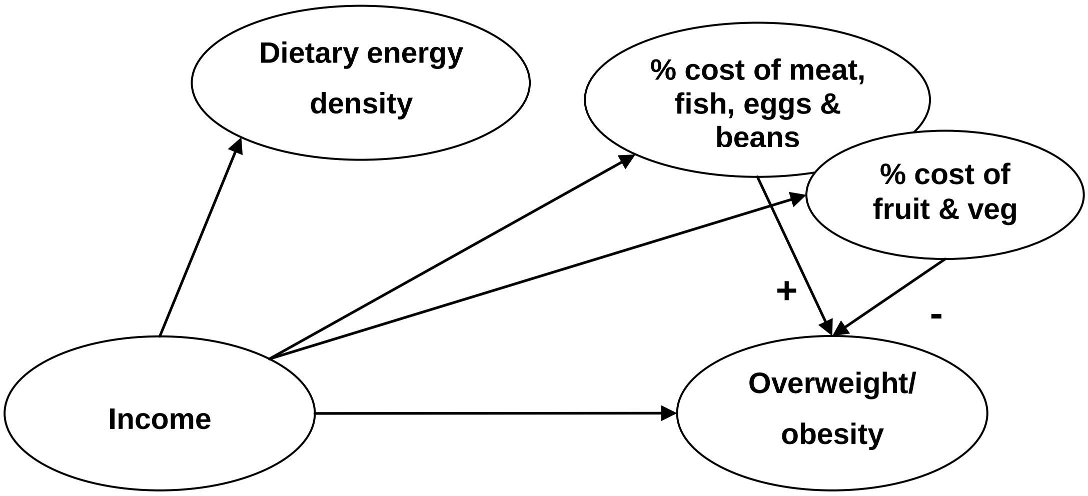


Solid lines represent significant associations; dashed lines show where associations would be expected according to the hypothesis, but were not apparent in the NDNS.

However, analyses in this sample failed to link diet costs or dietary energy density with BMI or overweight and obesity. A possible interpretation is that purchasing power does influence dietary choices, but that dietary energy density does not lead to positive energy balance; whilst income is related to BMI via a different mechanism (perhaps as a marker for socioeconomic status). Alternatively, explanations for these observations could be methodological – a lack of observed effect being due to insufficient sample size, for example, or due to self-reporting bias in the dietary assessment (see limitations below).

Chapter 8 probed further to determine whether the lack of associations, both in this research and in the literature, may be the result of inadequately capturing the diet cost variable. The results from this chapter could be seen to supplement Figure 9.1 as shown in Figure 9.2.

Figure 9.2 Associations in the NDNS, showing food group rather than whole diet costs



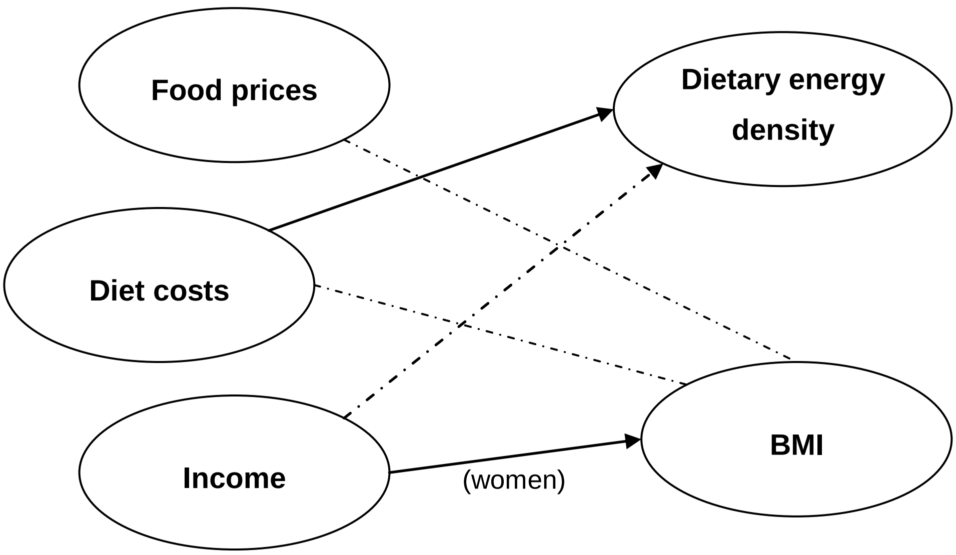
Incorporating costs from food groups modifies the relationships as shown above: diet costs, in the form of proportional costs for food groups, are now shown to be associated with BMI. Specifically, the evidence from Chapter 8 suggested links between costs from the meat, fish, eggs and beans food group, and those from fruit and vegetables, with BMI. Chapter 8 also implied a relationship between BMI and costs from high-fat and high-sugar foods; this has been omitted from Figure 9.2 on the assumption that this was a reflection of measurement error rather than an underlying relationship.

Although the link between proportional costs of food groups and DED was not tested in the analyses of this thesis, reports of associations between food group costs and the overall energy density of the diet can be seen in the literature (albeit using costs from differently categorised food groups) (Alexy et al., 2012). Taken together, a tentative inference is that the way in which people apportion their food budget, and not just the magnitude of that budget, affects energy balance.

The findings from the food group analyses are interesting and potentially insightful. However, a caveat to bear in mind is that the analyses were exploratory. Unusual odds ratios (see Section 8.4.5) are difficult to interpret, and may limit the confidence placed in the findings. Nevertheless, the chapter highlights a possible new avenue for diet cost research – further suggestions for research are given below.

The conclusions from the NDNS analyses can be compared to the findings from the literature review (Chapter 2; illustrated in Figure 9.3). Findings from the literature are altogether more mixed. However, the NDNS results confirm an association between diet costs and energy density found in other studies, as well as the link between income and BMI. The other relationships depicted in Figure 9.1 have been observed to be significant in at least some of the studies; however, the literature is not in agreement on these associations, and the links remain tentative.

Figure 9.3 Illustration of relationships implied by the literature review results (Chapter 2)



Solid arrows represent relationships with supporting evidence; dashed lines show where evidence is lacking or contradictory.

9.3.1 Limitations of the food price-obesity hypothesis

The main criticism of the food price-obesity hypothesis is its reductionist approach. The relationships portrayed in Figure 9.1, Figure 9.2 and Figure 9.3 and the pathway presented in Figure 1.1 all imply simple, linear connections between the variables. However, as mentioned previously (Chapter 1), at least some of these variables are likely to have a bidirectional relationship: for example, whilst food prices might influence the types and quantities of food purchased and consumed, the amount of food purchased (the demand) also plays a role in determining food prices.

The theoretically efficient representation of food prices as a single variable is also problematic when addressing the hypothesis with real data. This becomes obvious when examining the published evidence (Chapter 2), where there is a large degree of heterogeneity in studies' definitions and measurement of 'food prices'. In high-income countries especially, there is a wide variety of different foods and beverages on sale, each of which is likely to have a different degree of elasticity – for example, one review

(Andreyeva et al., 2010) identified eggs, sugars and sweets, cheese and fats and oils to be more inelastic than soft drinks or meat among other items. Further complicating the picture are cross-price elasticities, which indicate the effect of a price change of one food item on the consumption of another item: a price increase in some foods may encourage the consumption of replacement foods (supplements), or a decrease in consumption of both the more expensive food and another item that is commonly consumed alongside the first (complements). These economic phenomena will inevitably have whole-diet implications that may be difficult to predict.

Another detraction of the food price-obesity hypothesis are the assumptions made about the experience of the consumer in the decision-making process. Firstly, a degree of conscious reasoning is supposed – the consumer is assumed to take into account price information when purchasing foods. In actuality, it has been proposed that unconscious, as well as conscious, processes are involved in dietary decision making (Kremers et al., 2006). Unconscious processing is said to occur where environmental cues directly, or automatically, influence the choice behaviour. With hundreds of dietary decisions to be made every day, it has been suggested that consumers develop automatic heuristics to guide decisions (Scheibehenne et al., 2007) – habit may be considered one of these 'shortcuts', for example.

The degree to which consumers take into account food price information and correctly evaluate this in relation to relative prices (relative in time or to other foods) may be questioned. In the case of energy cost, information regarding costs per MJ are not readily available to the consumer, therefore requiring a great deal of conscious processing and numerical ability (Lipsky, 2009). Levels of numeracy are socially patterned in the UK (Bynner and Parsons, 1997), which would suggest that those individuals most capable of identifying the cheapest calories are in fact the individuals who are more likely to have higher socioeconomic status, including income, and therefore least likely to need to maximise the cost per calorie. Evidence from other samples (Turrell and Kavanagh, 2005, Ryden and Hagfors, 2011) do indicate that those exhibiting the higher diet costs tend both to earn more and to have had more years of education, suggesting that these groups are not motivated to achieve low diet costs.

Following this, if not all consumers are able to process the real price information, another line of enquiry worth pursuing relates to consumer *perceptions* of price. Studies have shown that a nutritionally adequate diet is possible to achieve within a strict budget constraint (Cassady et al., 2007, Maillot et al., 2008), yet this is not perceived to be the case by a majority of respondents (Cox et al., 1997). Cost is a frequently reported barrier to consuming healthier diets (Nelson et al., 2007). Qualitative evidence suggests that concern about food costs varies across individuals. The degree of 'food cost-concern' has been associated to food purchasing independently of income (Turrell & Kavanagh, 2005). Furthermore, price sensitivity has been linked to waist circumference (Gandal and Shabelansky, 2010), and attitudes to food prices linked to dietary energy density (Bowman, 2006).

One way in which food prices are brought to the attention of consumers, and may instil a perception of good value, is through promotions. The term 'promotion' encompasses a broad range of approaches, but here it is chiefly used to refer to financial incentive, such as price discounting, quantity discounting, or extra-product price promotions. In marketing, price is recognised as a conspicuous stimulator of consumption (Chandon & Wansink, 2011). Price promotions are often transient in nature, and data on their effects on purchasing tend to be short-term, though convincing (Hawkes, 2009, Chandon and Wansink, 2011). What is less clear is how this impacts on long-term behaviour and energy balance. Price promotions add further to the complexity of this field of study, making it difficult to evaluate their impact. However, such promotions may have a key role to play in the food price-obesity hypothesis which is yet to be addressed.

Another assumption inherent in the food price-obesity hypothesis is that income is related to obesity via its effect on purchasing power. In fact, as a marker of socioeconomic status, it is possible that an observed association between income and body weight may in fact be reflecting an association with socioeconomic status more generally, and not (solely) because income allows the purchase of more food. This is supported by the conclusions of systematic reviews (see Chapter 2) in which income was identified as the least consistent of socioeconomic indicators in predicting BMI. In addition, low income may be associated with more harsh environments, themselves linked to unhealthy dietary choices (Laran and Salerno, 2013). From another viewpoint, it has been suggested that income inequality itself, rather than absolute incomes, could be causally related to health (Wilkinson and Pickett, 2010).

Finally, the food price-obesity hypothesis is guilty of not recognising the influence of other probable determinants of weight gain and obesity. These include aspects of the obesogenic environment (Swinburn et al., 1999), as well as wider influences as detailed in the Foresight report (Butland et al., 2007). In particular, some of the key factors in dietary decision-making worth mentioning are availability, the retail environment, and time costs. 'Time cost' refers to the amount of time needed to purchase, transport and prepare foods. One study estimated that time considerations made up to as much as 49% of the ‘cost’ deliberation of an individual when purchasing food (Davis and You, 2010). As such, the potential for minimisation of monetary cost will be dependent upon how much a person values their time (Leung and et al., 1997). Availability in this sense primarily relates to the geographical location of food retail outlets, their accessibility, their density and the types of store that are accessible. It has been suggested that, in having limited access to cheaper stores and restricted travel options (which might disallow bulk purchasing), the poor are faced with higher costs for the same foods (Beatty, 2010). Observations of shoppers suggest that BMI varies by type of store (Chaix et al., 2012, Lear et al., 2013), and at least one before-and-after study has measured a change in dietary intakes following the arrival of a supermarket to a ‘food desert’ (Wrigley et al., 2002). These descriptions of wider determinants are not intended to be exhaustive; however they illustrate how the influence of food prices sits within a complex environment with many factors affecting dietary choices.

9.4 Limitations of this research

Many of the limitations of this research have been detailed throughout the preceding Chapters. This summary is intended to recapitulate the main drawbacks of the approaches taken in the analyses, which are important to take into account when interpreting the results.

A key limitation is in trying to establish a causal relationship using cross-sectional data. As mentioned previously, the development of obesity is usually assumed to take place across a protracted time period. Measuring body weight and diet concurrently may be misleading where diet has changed through time, and current diet is no longer a reflection of the dietary consumption which led to weight gain. Having said that, studies of year-to-year comparisons of dietary assessment have shown little within-subject variation in nutrient intakes (Willett, 1998). Still, whilst cross-sectional analyses may indicate interesting patterns that could be potentially meaningful, it is not possible to make firm conclusions about causality.

A second important drawback of the research in this thesis relates to the assumptions made in the diet costing method. An in-depth discussion of these was related in Chapter 5, but in brief, these include the assumption that mean national prices give an indication of food costs, whereas the foods consumed by individuals could have cost more or less than average, or even been without financial cost (such as free or foraged food). In addition, the DANTE cost database cannot take into account where foods have been purchased on promotion.

Costing diets in this manner is also likely to echo the measurement error for which dietary assessment is notorious. Self-reporting bias, whether conscious or unconscious, will result in biased estimates of diet costs as well as of dietary intake. Even without mis-reporting, diet assessment may not accurately reflect usual intakes if eating patterns are altered in response to the act itself of keeping diet records (Rebro et al., 1998). Measurement error may also have affected the dietary energy density estimates made in this thesis.

A third drawback is that the observations and conclusions made in the preceding chapters apply to a British adult population and therefore may not generalise to other populations. Food markets differ across the globe, and the interplay between price concerns and culture are also likely to vary (see, for example, a comparison of food away from home expenditure in Europe and in the US – Michaud et al. (2007)). Within the UK as well, observations and results may not extend to children and adolescents, who were not included in the analytical sample. The results of some studies in the literature (see Chapter 2) hint towards a lack of relationship between food prices and weight in children, and it would be interesting to see if (and how) economic factors play a role in the diets of British adolescents and children.

It is also worth pointing out that many of the analyses and findings presented in this thesis are the result of exploratory investigations. New approaches to investigating the food price-obesity hypothesis cross-sectionally – for example, in using food group specific proportional costs – are described as potential avenues for further research, but results presented here are difficult to interpret with confidence. Therefore, all conclusions stated in this Chapter and in preceding chapters' discussion are tentatively given.

Finally, as mentioned above, income and diet costs represent only a small part of a complex problem. Understanding economic determinants of diet and health could offer routes to useful interventions (see section 9.4 below), but are unlikely to be wholly responsible for the obesity epidemic. For a more complete consideration of obesity causes, the Obesity Systems Map gives a thorough representation (Butland et al., 2007).

9.5 Implications

As well as exploring the hypothesis that food prices are influential in the development of obesity, this thesis examined several methodological points in the research area. Therefore, implications of the findings can be considered both in terms of the wider implications – for example, what the results could mean for policy – as well as implications for the researcher in this field. These will be covered separately below.

9.5.1 Implications for researchers

The first methodological point made in this thesis related to the measurement of income. Although there are known issues with capturing information on income, and household income may be considered a crude variable (see Chapter 4 for a discussion), it remains a pragmatic variable to gather data on in large-scale surveys. The investigations of this thesis illustrated how crude household income can be enhanced as a variable by equivalizing for household size and composition. Despite being a common approach in many surveys, equivalization appears to be less common in nutritional epidemiology. The implications of Chapter 4 are that investigators in nutritional epidemiology would be well served to apply a well-established equivalization index to household income data, rather than the more common approach of adjusting for household size only.

The second major methodological point made in this thesis related to the validity of diet costing methods. Food purchasing and food consumption are often treated interchangeably in this field of research – as discovered in the review of the literature in Chapter 2. However, as discussed in Chapter 5, they are not necessarily equivalent: food purchased may not be consumed, and foods consumed may not have been purchased, for example. Chapter 5 went on to explore how diet costs estimated from consumption data relates to costs calculated from purchasing data. This was not a validity study as such, given that neither method can be considered a gold standard. However, it will help to interpret the estimates of studies using a food price database costing approach. The implication for the researcher is that the choice of method – purchasing versus estimating costs from consumption – will give different estimates of diet costs at an individual level, but will probably give similar mean values at a population level. The purpose of the research question should determine the importance to gauge individual-level costs, which will aid in the choice of estimation method.

Costing diets using databases of food prices is already a common approach in the literature, with several researchers applying food price data to dietary data collected in a different year. In Chapter 6 of this thesis, the role of inflation in estimating diet costs was explored. The findings suggested that a flat rate of inflation (the Food Price Index, FPI) might not reflect the prices faced by some people, if they consume proportionately more of a particular food group that may have seen a higher rate of inflation than that averaged across all foods. Looking at the individual food group indices, which are available publicly in the UK from the Office for National Statistics (ONS) (see Appendix C), it can be seen that some foods, such as vegetables, experienced a greater price increase than the overall FPI, whilst others showed less of a price increase, such as soft drinks or poultry. The recommendation from this finding is that researchers should, where possible and pragmatic to do so, apply more sensitive inflation indices – such as by food group – to food price databases that need correcting to another year of data collection. In this instance, it was not judged pragmatic to do so, given the advantage in terms of sample size of combining years of data collection.

The final methodological finding of this thesis relates to proportional food group costs. Initial findings from the exploratory investigations of Chapter 8 suggest that there may be value in characterising diet costs in more detail than is offered by whole diet costs. However, the results presented in Chapter 8 indicate that more work is needed in food group costing before the approach can be adopted with confidence. The implication for other researchers is that consideration needs to be given to alternative approaches to costing diets, but it is as yet too early to recommend the best methods. Ideas for further investigations in this area are presented in Section 9.6 below.

9.5.2 Implications for policy

Currently, efforts to stem the rise in overweight and obesity in the UK do not appear to have had an obvious impact, and many practitioners, researchers and advocacy groups have called for new solutions (Limb, 2013, Academy of Medical Royal Colleges, 2013). Taken on their own, the results of this thesis imply that a means to combat the obesity epidemic could lie in interventions which target individuals' purchasing power. This would entail increasing incomes (a policy sphere perhaps beyond the reach of public health) or manipulating food prices to encourage healthier dietary choices.

An obvious, and well-debated, means of food price manipulation would be to impose a food- or nutrient-based tax to increase the prices of 'unhealthy' items, or to introduce subsidies for foods considered healthier. Some governments already have instituted health-motivated taxes on foods or nutrients: for example, Hungary on 'junk food' (Holt, 2011), Denmark for a limited period on saturated fat (Jensen and Smed, 2013) and, most recently, Mexico on soft drinks and junk food (Boseley, 2013).

Evidence for the effects of such taxes is still thin on the ground as yet (Timmins, 2011). The majority of the evidence stems from modelling studies using estimated price elasticities, most of which imply significant effects on purchases, but some of which caution about unintended substitution effects (see Eyles et al., 2012, for a review). Other evidence can be found in experimental manipulations or from a few natural experiments (Mytton et al., 2012). Whilst these types of study similarly indicate significant negative effects of taxes on purchasing, evidence for an effect on body weight or obesity is not strong (see Chapter 2 for a discussion of soda tax studies in the US).

Paucity of evidence may not be the sole argument against taxation policies, with many opponents emphasising the potentially regressive nature of taxes (bringing into focus the ethics of these policies) as well as questioning the size of effect (Winkler, 2012). Ultimately, policies such as these will only be brought into effect if they are politically acceptable (Swinburn et al., 2011). Regardless of the evidence, acceptability may be present if the intervention is judged to be proportionate (Nuffield Council on Bioethics, 2007). In addition, fiscal policies may be attractive to policy makers in that they have the potential to be a cost-effective approach (Lehnert et al., 2012).

The results of the final analyses in this thesis (Chapter 8) may offer an additional, albeit tentative, perspective on the taxation debate. Food group cost analyses suggested that it is not the food budget itself that drives dietary selection, but how people are willing to apportion their budget. If this is the case, it points toward an educational approach, in which efforts could be made to educate about food budgeting. As described in Chapter 1, cost of food is frequently identified by consumers as a barrier to healthy eating. Communicating how to achieve a healthy diet within budget could be a welcome message to the British public.

A final implication arising from this thesis relates to monitoring and surveillance. Given the indications in the literature as well as the policy debates around fiscal measures, there may be an argument for the monitoring of dietary expenditure along with national dietary surveys. Data on both expenditure and food consumption would add considerably to the evidence base, and could prove useful in determining the appropriateness of fiscal interventions.

Of course, these results should not stand in isolation. Even if a fiscal policy is deemed necessary, it is unlikely to be sufficient. As stated above, the food price-obesity theory neglects the myriad other causes of obesity that have been proposed. Obesity is a multi-faceted, complex issue which will in all likelihood require a similarly multi-faceted approach with multiple interventions (Butland et al., 2007).

9.6 Possibilities for future work

This thesis explores relationships from an emerging research area, in which there is plenty of scope for future investigation. The literature review in Chapter 2, for instance, identifies some large gaps in research: firstly, there have been no published studies investigating food prices and dietary energy density; and secondly, longitudinal studies are scarce, particularly when considering diet costs. Studies to address either of these gaps would help to complete the picture of the food price-obesity hypothesis.

In terms of diet, the focus of this research was on dietary energy density, due to its association (observed and theoretical) with positive energy balance. However, economic factors of diet have the potential to impact on more than energy balance and obesity. In this sample, for example, patterns in fruit and vegetable intake were clearly different according to income and diet cost. This may imply inequalities in other macro- and micro-nutrient intakes which are important for public health. A recent review (Rao et al., 2013) has synthesised a number of studies which have investigated diet costs in relation to other indicators of dietary quality or patterns, and indicated broadly consistent findings that the healthier diets cost more than the least healthy. However, the number of studies was small, and heterogeneous, and results varied according to whether daily diet costs or energy-adjusted costs were used. Furthermore, there were no studies in the UK, and in particular, none in the NDNS. This would be an interesting area to pursue in the nationally representative sample.

The results of Chapter 8 show an initial foray into food group costs and how they relate to whole diet costs and to diet and health. As a fairly new approach, there remains much that could be explored in this representation of diet costs. For example, comparisons to expenditure data would be enlightening, as would investigations in which energy expenditure was assessed in order to account for the possible effect of under-reporting of particular food groups. It would also be an interesting avenue to further examine how proportional food groups costs relate to whole diet costs – do proportional costs in the lowest quintile of whole diet costs differ to those of the highest quintile, for example? Finally, further work is needed to determine the most appropriate groupings of foods: the eight categories derived in Chapter 8, although based upon UK recommendations, may have been too broadly defined. Chapter 8 also indicated that food budget apportioning may be a crucial step between food prices and diet, potentially providing an alternative policy approach (see above). As far as the author is aware, there have been no trials in which a food budget educational intervention is implemented and evaluated. This could be an important focus of future research.

A final suggestion would be to extend the investigations presented here by incorporating later waves of the NDNS sample. Sample size may have been an issue in the analyses investigating BMI, and, in addition, the sample size limited the design of this study in that a thorough investigation incorporating both income and diet costs together was not feasible.

9.7 Concluding remarks

Rates of obesity and overweight are of real concern to healthcare providers, and, as yet, attempts to stem the trends in the UK have not yielded great success. This thesis attempted to gauge whether obesity rates may be attributed to trends in food prices, by, first, synthesising the published evidence and, second, investigating the purchasing power (with equivalized income and diet costs as proxies) of NDNS adults along with dietary energy density and BMI. It is the first time monetary costs have been applied to the diets of NDNS adults, and the thesis also introduced the novel approach of linking proportional food group costs to a health outcome.

Findings from both the literature and the NDNS analyses confirm socioeconomic differences in diet costs, and indicate a negative relationship between diet costs and dietary energy density. Evidence linking food prices or diet costs to body weight, however, is less conclusive. A key output of this thesis has been to highlight the methodological (and theoretical) difficulties in researching this question: available data are abounding with assumptions, and it is challenging to draw out relationships in such a complex system. However, these initial results suggest that there is merit in pursuing this line of research.

A cheap, healthy diet is not an oxymoron. Nevertheless, cost of food may be a crucial contributor to the obesogenic environment, dominating food purchase decisions and perhaps encouraging unhealthy diets. With growing concerns about sustainability and the future of the food industry, diet cost research can only grow in its contribution to the knowledge base.

References

AARON, G. J., KEIM, N. L., DREWNOWSKI, A. & TOWNSEND, M. 2013. Estimating dietary costs of low-income women in California: a comparison of 2 approaches. *Am J Clin Nutr,* 97**,** 835-841.
ACADEMY OF MEDICAL ROYAL COLLEGES 2013. Measuring up: the medical profession's prescription for the nation's obesity crisis.
ACHESON, D. 1998. Inequalities in health: report of an independent inquiry. London: HMSO.
ADAMS, E. J., GRUMMER-STRAWN, L. & CHAVEZ, G. 2003. Food insecurity is associated with increased risk of obesity in California women. *J Nutr,* 133**,** 1070-1074.
AGGARWAL, A., MONSIVAIS, P., COOK, A. J. & DREWNOWSKI, A. 2011. Does diet cost mediate the relation between socioeconomic position and diet quality? *EUROPEAN JOURNAL OF CLINICAL NUTRITION* 65**,** 1059-1066.
ALEXY, U., BOLZENIUS, K., KOPPER, A., CLAUSEN, K. & KERSTING, M. 2012. Diet costs and energy density in the diet of German children and adolescents. *European Journal of Clinical Nutrition,* 66**,** 1362-3.
ALI, M. K., BARKER, L., MCKEEVER BULLARD, K., NARAYAN, K. M. V., BECKLES, G. L., IMPERATORE, G. & STEVENS, M. R. 2011. Disease risks in US children and young adults. Analyses from NHANES 1999-2008. *Diabetes Care,* 34**,** 1998-2004.
ALTMAN, D. G. 1991. *Practical Statistics for Medical Research,* London, UK, Chapman & Hall.
AN, R. P. 2012. Effectiveness of subsidies in promoting healthy food purchases and consumption: a review of field experiments. *Public Health Nutrition,* 16**,** 1215-1228.
ANDREYEVA, T., LONG, M. W. & BROWNELL, K. D. 2010. The impact of food prices on consumption: a systematic review of research on the price elasticity of demand for food. *American Journal of Public Health,* 100**,** 216-22.
ANDRIEU, E., DARMON, N. & DREWNOWSKI, A. 2006. Low-cost diets: more energy, fewer nutrients. *European Journal of Clinical Nutrition,* 60**,** 434-6.
ANYAEGBU, G. 2010. Using the OECD equivalence scale in taxes and benefits analysis. *Economic & Labour Market Review,* 4**,** 49-54.
ASFAW, A. 2007. Do Government Food Price Policies Affect the Prevalence of Obesity? Empirical Evidence from Egypt. *World Development,* 35**,** 687-701.
AULD, M. C. & POWELL, L. M. 2009. Economics of Food Energy Density and Adolescent Body Weight. *Economica,* 76**,** 719-40.
BAILEY, R. L., MITCHELL, D. C., MILLER, C. & SMICIKLAS-WRIGHT, H. 2007. Assessing the effect of underreporting energy intake on dietary patterns and weight status. *J Am Diet Assoc***,** 64-71.
BALL, K. & CRAWFORD, D. 2005. Socioeconomic status and weight change in adults: a review. *Social Science & Medicine,* 60**,** 1987-2010.
BATES, B., LENNOX, A., BATES, C. & SWAN, G. 2011. *National Diet and Nutrition Survey, Headline results from Years 1 and 2 (combined) of the Rolling Programme (2008/2009-2009/2010).* [Online]. Available: http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/documents/digitalasset/dh_128550.pdf.
BEATTY, T. K. M. 2010. Do the Poor Pay More for Food?: Evidence from the United Kingdom. *American Journal of Agricultural Economics,* 92**,** 608-621.
BENZEVAL, M., JUDGE, K. & SHOULS, S. 2001. Understanding the relationship between income and health: how much can be gleaned from cross-sectional data? *Social Policy & Administration,* 35**,** 376-396.

BERNSTEIN, A. M., BLOOM, D. E., ROSNER, B. A., FRANZ, M. & WILLETT, W. C. 2010. Relation of food cost to healthfulness of diet among US women. *American Journal of Clinical Nutrition,* 92**,** 1197-203.

BEYDOUN, M. A., POWELL, L. M. & WANG, Y. 2008. The association of fast food, fruit and vegetable prices with dietary intakes among US adults: Is there modification by family income? *Social Science & Medicine,* 66**,** 2218-2229.

BINGHAM, S., GILL, C., WELCH, A., DAY, K., CASSIDY, A., KHAW, K., SNEYD, M., KEY, T., ROE, L. & DAY, N. 1994. Comparison of dietary assessment methods in nutritional epidemiology : weighed records v. 24 h recalls, food-frequency questionnaires and estimated-diet records. *British Journal of Nutrition,* 72**,** 619-643.

BLACK, A. P., BRIMBLECOMBE, J., EYLES, H., MORRIS, P., VALLY, H. & O DEA, K. 2012. Food subsidy programs and the health and nutritional status of disadvantaged families in high income countries: a systematic review. *BMC Public Health,* 12**,** 1099.

BLACK, A. P., VALLY, H., MORRIS, P. S., DANIEL, M., ESTERMAN, A. J., SMITH, F. E. & O'DEA, K. 2013. Health outcomes of a subsidised fruit and vegetable program for Aboriginal children in northern New South Wales. *Medical Journal of Australia,* 199**,** 46-50.

BLAND, J. M. & ALTMAN, D. G. 1986. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet,* i**,** 307-310.

BLOCK, J. P., CHANDRA, A., MCMANUS, K. D. & WILLETT, W. C. 2010. Point-of-Purchase Price and Education Intervention to Reduce Consumption of Sugary Soft Drinks. *American Journal of Public Health,* 100**,** 1427-1433.

BLOOM, D. E. & IDSON, T. L. 1991. The practical importance of sample weights. *Proceedings of the Survey Research Methods Section, ASA***,** 620.

BOIZOT-SZANTAI, C. & ETILE, F. 2005. The food prices/ body mass index relationship: Theory and evidence from a sample of French adults. Ivry-sur-Seine.

BOSELEY, S. 2013. Mexico to tackle obesity with taxes on junk food and sugary drinks. *The Guardian*.

BOWERS, D. 2008. *Medical Statistics from Scratch: an introduction for health professionals,* Chichester, John Wiley & Sons Ltd.

BOWMAN, S. A. 2006. A comparison of the socioeconomic characteristics, dietary practices, and health status of women food shoppers with different food price attitudes. *Nutrition Research,* 26**,** 318-324.

BRESLOW, R. A., CHEN, C. M., GRAUBARD, B. I., JACOBOVITS, T. & KANT, A. K. 2013. Diets of drinkers on drinking and nondrinking days: NHANES 2003-2008. *Am J Clin Nutr,* 97**,** 1068-75.

BRESLOW, R. A., GUENTHER, P. M. & SMOTHERS, B. A. 2006. Alcohol Drinking Patterns and Diet Quality: The 1999–2000 National Health and Nutrition Examination Survey. *Am J Epidemiol,* 163**,** 359-366.

BRINKMAN, H. J., DE PEE, S., SANOGO, I., SUBRAN, L. & BLOEM, M. W. 2009. High Food Prices and the Global Financial Crisis Have Reduced Access to Nutritious Food and Worsened Nutritional Status and Health. *Journal of Nutrition,* 140**,** 153S-161S.

BROWNWELL, K. D., FARLEY, T., WILLETT, W. C., POPKIN, B. M., CHALOUPKA, F. J., THOMPSON, J. W. & LUDWIG, D. S. 2009. The public health and economic benefits of taxing sugar-sweetened beverages. *N Engl J Med.,* 360**,** 1599–1605.

BUTLAND, B., JEBB, S. A., KOPELMAN, P., MCPHERSON, K., THOMAS, S., MARDELL, J. & PARRY, V. 2007. Tackling Obesity: Future Choices - Project Report (2nd ed). Available: http://www.bis.gov.uk/foresight/our-work/projects/published-projects/tackling-obesities/reports-and-publications [Accessed 21 February 2011].

BYNNER, J. & PARSONS, S. 1997. *It Doesn't Get Any Better. The impact of poor basic skills on the lives of 37 year olds,* London, The Basic Skills Agency.

CADE, J. & MARGETTS, B. M. 1991. Relationship between diet and smoking - is the diet of smokers different? *J Epidemiol Community Health,* 45**,** 270-272.

CADE, J., UPMEIER, H., CALVERT, C. & GREENWOOD, D. 1999. Costs of a healthy diet: analysis from the UK Women's Cohort Study. *Public Health Nutrition,* 2**,** 505–512.

CANOY, D., WAREHAM, N., LUBEN, R., WELCH, A., BINGHAM, S., DAY, N. & KHAW, K. 2005. Cigarette smoking and fat distribution in 21,828 British men and women: a population-based study. *Obes Res,* 13**,** 1466-75.

CARE/WFP 2003. The Coping Strategies Index: Field Methods Manual. Nairobi: CARE and WFP.

CARR-HILL, R. & CHALMERS-DIXON, P. 2005. The Public Health Observatory Handbook of Health Inequalities Measurement. *In:* LIN, J. (ed.).

CASSADY, D., JETTER, K. & CULP, J. 2007. Is Price a Barrier to Eating More Fruits and Vegetables for Low-Income Families? *Journal of the American Dietetic Association,* 107**,** 1909-1915.

CHAIX, B., BEAN, K., DANIEL, M., ZENK, S. N., KESTENS, Y., CHARRIERE, H., LEAL, C., THOMAS, F., KARUSISI, N., WEBER, C., OPPERT, J.-M., SIMON, C., MERLO, J. & PANNIER, B. 2012. Associations of supermarket characteristics with weight status and body fat: a multilevel analysis of individuals within supermarkets (RECORD study). *PloS ONE,* 7**,** e32908.

CHALOUPKA, F. J. & POWELL, L. M. 2009. Price, availability, and youth obesity: Evidence from Bridging the Gap. *Prev Chronic Dis,* 6.

CHANDON, P. & WANSINK, B. 2011. Is food marketing making us fat? A multi-disciplinary review. INSEAD Working Paper collection.

CHANFREAU, J. & BURCHARDT, T. 2008. Equivalence scales: rationales, uses and assumptions. Available: http://www.scotland.gov.uk/Resource/Doc/933/0079961.pdf [Accessed September 24 2013].

CHOU, S.-Y., GROSSMAN, M. & SAFFER, H. 2004. An Economic Analysis of Adult Obesity: Results from the Behavioral Risk Factor Surveillance System. *National Bureau of Economic Research, Inc, NBER Working Papers*.

CHURCH, T., THOMAS, D., TUDOR-LOCKE, C., KATZMARZYK, P. & EARNEST, C. 2011. Trends over 5 Decades in U.S. Occupation-Related Physical Activity and Their Associations with Obesity. *PLoS ONE,* 6**,** e19657.

CONKLIN, A. I., MAGUIRE, E. R. & MONSIVAIS, P. 2013. Economic determinants of diet in older adults: systematic review. *Journal of Epidemiology and Community Health,* 67**,** 721-727.

CONNORS, M., BISOGNI, C. A., SOBAL, J. & DEVINE, C. M. 2001. Managing values in personal food systems. *Appetite,* 36**,** 189-200.

COTTI, C. & TEFFT, N. 2013. Fast food prices, obesity, and the minimum wage. *Economics and Human Biology,* 11**,** 134-147.

COX, D. N., ANDERSON, A. S., LEAN, M. E. J. & MELA, D. J. 1997. UK consumer attitudes, beliefs and barriers to increasing fruit and vegetable consumption. *Public Health Nutrition,* 1**,** 61-68.

COX, D. N. & MELA, D. J. 2000. Determination of energy density of freely selected diets: methodological issues and implications. *International Journal of Obesity,* 24**,** 49-54.

CRAIG, R. & HIRANI, V. 2010. Health Survey for England 2009. Volume 1: Health and lifestyles. Available: http://www.hscic.gov.uk/pubs/hse09report [Accessed July 24 2013].

CRAIG, R. & MINDELL, J. 2011. Health Survey for England 2011. London.

CRD 2008. Systematic Reviews: CRD's guidance for undertaking reviews in healthcare. York.
CUMMINS, S. & MACINTYRE, S. 2002. The price and availability of food in Glasgow: a systematic study of an urban foodscape. . *Urban Studies,* 39**,** 2115-2130.
DARMON, M., FERGUSON, E. L. & BRIEND, A. 2003. A cost constraint alone has adverse effects on food selection and nutrient density: An analysis of human diets by linear programming. *J. Nutr.,* 132**,** 3764–3771.
DARMON, N., BRIEND, A. & DREWNOWSKI, A. 2004. Energy-dense diets are associated with lower diet costs: a community study of French adults. *Public Health Nutrition,* 7.
DARMON, N. & DREWNOWSKI, A. 2008. Does social class predict diet quality? *Am J Clin Nutr,* 87**,** 1107–17.
DARMON, N., FERGUSON, E. L. & BRIEND, A. 2006. Impact of a cost constraint on nutritionally adequate food choices for French women: an analysis by linear programming. *J Nutr Educ Behav,* 38.
DAVIS, G. C. & YOU, W. 2010. The Time Cost of Food at Home: General and Food Stamp Participant Profiles. *Applied Economics,* 42**,** 2537-52.
DE IRALA-ESTEVEZ, J., GROTH, M., JOHANSSON, L., OLTERSDORF, U., PRATTALA, R. & MARTINEZ-GONZALEZ, M. A. 2000. A systematic review of socio-economic differences in food habits in Europe: consumption of fruit and vegetables. *European Journal of Clinical Nutrition,* 54**,** 706-714.
DEFRA 2009. Family Food: A report on the 2009 Family Food Module of the Living Costs and Food Survey. The Office for National Statistics: Defra.
DEFRA 2010. Household Food and Drink Waste linked to Food and Drink Purchases. Department for Environment Food and Rural Affairs.
DEFRA 2012. Family Food 2010 (revised).
DH. 2011. *The eatwell plate* [Online]. Available: http://webarchive.nationalarchives.gov.uk/+/www.dh.gov.uk/en/Publichealth/Nutrition/DH_126493.
DINOUR, L. M., BERGEN, D. & YEH, M.-C. 2007. The food insecurity-obesity paradox: a review of the literature and the role food stamps may play. *Journal of the American Dietetic Association,* 107**,** 1952-1961.
DREWNOWSKI, A. 2007. The real contribution of added sugars and fats to obesity. *Epidemiologic Reviews,* 29**,** 160-71.
DREWNOWSKI, A. & DARMON, N. 2005. Food choices and diet costs: an economic analysis. *J. Nutr.,* 135**,** 900-904.
DREWNOWSKI, A., DARMON, N. & BRIEND, A. 2004. Replacing fats and sweets with vegetables and fruits — A question of cost. *Am J Public Health,* 94**,** 1555–1559.
DREWNOWSKI, A., MONSIVAIS, P., MAILLOT, M. & DARMON, N. 2007. Low-energy-dense diets are associated with diet quality and higher diet costs in French adults. *J Am Diet Assoc,* 107**,** 1028-1032.
DREWNOWSKI, A. & SPECTER, S. E. 2004. Poverty and obesity: the role of energy density and energy costs. *American Journal of Clinical Nutrition,* 79**,** 6-16.
DU, S., MROZ, T. A., ZHAI, F. & POPKIN, B. M. 2004. Rapid income growth adversely affects diet quality in China - particularly for the poor! *Soc Sci Med,* 59**,** 1505-15.
DUFFEY, K. J., GORDON-LARSEN, P., SHIKANY, J. M., GUILKEY, D., JACOBS, D. R. & POPKIN, B. M. 2010. Food price and diet and health outcomes: 20 years of the CARDIA study. *Arch Intern Med.,* 170**,** 420-426.
DWP 2013. Households below average income (HBAI): 1994/95 to 2011/12.
EL-SAYED, A. M., SCARBOROUGH, P. & GALEA, S. 2012. Unevenly distributed: a systematic review of the health literature about socioeconomic inequalities in adult obesity in the United Kingdom. *BMC Public Health,* 12**,** 18.

EYLES, H., NI MHURCHU, C., NGHIEM, N. & BLAKELY, T. 2012. Food pricing strategies, population diets, and non-communicable disease: a systematic review of simulation studies. *PLoS medicine,* 9**,** e1001353.

FLEGAL, K. M., KIT, B. K., ORPANA, H. & GRAUBARD, B. I. 2013. Association of all-cause mortality with overweight and obesity using standard body mass index categories. *JAMA,* 309**,** 71-82.

FLETCHER, J., FRISVOLD, D. & TEFFT, N. 2009. The Effects of Soft Drink Taxes on Child and Adolescent Consumption and Weight Outcomes. *Department of Economics, Emory University*.

FLETCHER, J. M., FRISVOLD, D. & TEFFT, N. 2010a. Can Soft Drink Taxes Reduce Population Weight? *Contemporary Economic Policy,* 28**,** 23-35.

FLETCHER, J. M., FRISVOLD, D. & TEFFT, N. 2010b. Taxing soft drinks and restricting access to vending machines to curb child obesity.

FLYNN, M. A. T., O'BRIEN, C. M., ROSS, V., FLYNN, C. A. & BURKE, S. J. 2011. Revision of food-based dietary guidelines for Ireland, Phase 2: recommendations for healthy eating and affordability. *Public Health Nutrition,* 15**,** 527-537.

FREEDMAN, L. S., SCHATZKIN, A., MIDTHUNE, D. & KIPNIS, V. 2011. Dealing with dietary measurement error in nutritional cohort studies. *J Natl Cancer Inst,* 103**,** 1086-92.

FRENCH, S. A. 2003. Pricing Effects on Food Choices. *J. Nutr.,* 133**,** 841S–843S.

FRIEL, S., NELSON, M., MCCORMACK, K., KELLEHER, C. C. & THRISKOS, P. 2001. Methodological issues using Household Budget Survey expenditure data for individual food availability estimation: Irish experience in the DAFNE pan European project. *Journal of Public Health Nutrition,* 4**,** 1143-1148.

FSA 2002. *McCance and Widdowson's The Composition of Foods, Sixth summary edition.,* Cambridge, Royal Society of Chemistry.

GANDAL, N. & SHABELANSKY, A. 2010. Obesity and price sensitivity at the supermarket. *Forum for Health Economics & Policy.* Tel Aviv: Tel Aviv University.

GANNON, B. 2009. The influence of economic incentives on reported disability status. *Health Economics,* 18**,** 743-759.

GARIBALLA, S. & SINCLAIR, A. 2005. Ageing and older people. *In:* GEISSLER, C. & POWERS, H. (eds.) *Human Nutrition.* 11th ed. PA, USA: Elsevier Churchill Livingstone.

GFK SOCIAL RESEARCH 2009. Public Attitudes to Food Issues. Report for the Food Standards Agency. London.

GILSING, A. M. J., WEIJENBERG, M. P., HUGHES, L. A. E., AMBERGEN, T., DAGNIELE, P. C., DGOLDBOHM, A. & VAN DEN BRANDT, P. A. 2012. Longitudinal changes in BMI in older adults are associated with meat consumption differentially, by type of meat consumed. *J Nutr,* 142**,** 340-349.

GISKES, K., AVENDANO, M., BRUG, J. & KUNST, A. E. 2010. A systematic review of studies on socioeconomic inequalities in dietary intakes associated with weight gain and overweight/obesity conducted among European adults. *Obesity Reviews,* 11**,** 413-429.

GLYMOUR, M. M. 2006. Using causal diagrams to understand common problems in social epidemiology. *In:* OAKES, M. & KAUFMAN, J. (eds.) *Methods in Social Epidemiology.* Jossey-Bass.

GOODMAN, C. & ANISE, A. 2006. *What is known about the effectiveness of economic instruments to reduce consumption of foods high in saturated fats and other energy-dense foods for preventing and treating obesity?,* Geneva, World Health Organization.

GORBER, S. C., TREMBLAY, M., MOHER, D. & GORBER, B. 2007. A comparison of direct vs self-report measures for assessing height, weight and body mass index: a systematic review. *Obesity Reviews,* 8**,** 307-326.

GORTMAKER, S. L., SWINBURN, B., LEVY, D. E., CARTER, R., MABRY, P. L., FINEGOOD, D. T., HUANG, T., MARSH, T. & MOODIE, M. 2011. Changing the future of obesity: science, policy, and action. *The Lancet,* 378**,** 838-847.

GREEN, R., CORNELSEN, L., DANGOUR, A. D., TURNER, R., SHANKAR, B., MAZZOCCHI, M. & SMITH, R. D. 2013. The effect of rising food prices on food consumption: systematic review with meta-regression. *BMJ,* 346**,** f3703.

GREENLAND, S. 1995. Dose-Response and Trend Analysis in Epidemiology: Alternatives to Categorical Analysis. *Epidemiology,* 6**,** 356-365.

GREENLAND, S., PEARL, J. & ROBINS, J. 1999. Causal diagrams for epidemiologic research. *Epidemiology,* 10**,** 37-48.

GREENWOOD, D. Statistical modelling in health research: Sample size & post hoc tests. MEDR5150M, 16 June 2011 University of Leeds.

HARNACK, L., BLOCK, G. & LANE, S. 1997. Influence of selected environmental and personal factors on dietary behavior for chronic disease prevention: a review of the literature. *Journal of Nutrition Education,* 29**,** 306-312.

HARTLINE-GRAFTON, H. L., ROSE, D., JOHNSON, C. C., RICE, J. C. & WEBBER, L. S. 2009. Energy density of foods, but not beverages, is positively associated with body mass index in adult women. *EUROPEAN JOURNAL OF CLINICAL NUTRITION,* 63**,** 1411-1418.

HAWKES, C. 2009. Sales promotions and food consumption. *Nutrition Reviews,* 67**,** 333-342.

HEBERT, J. R., CLEMOW, L., PBERT, L., OCKENE, I. S. & OCKENE, J. K. 1995. Social desirability bias in dietary self-report may compromise the validity of dietary intake measures. *Int J Epidemiol,* 24**,** 389-398.

HERMAN, C. P. & POLIVY, J. 2010. Sex and gender differences in eating behavior. *In:* CHRISLER, J. C. & MCCREARY, D. R. (eds.) *Handbook of Gender Research in Psychology.* New York: Springer.

HIRANI, V. 2011. Chapter 10, Adult anthropometric measures, overweight and obesity. *In:* CRAIG, R. (ed.) *Health Survey for England - 2010, Respiratory health.* hscic.

HIRANI, V. & MINDELL, J. 2008. A Comparison of Measured Height and Demi-span Equivalent Height in the Assessment of Body Mass Index among People Aged 65 Years and Over in England. *Age Aging,* 37**,** 311-317.

HIZA, H. A., CASAVALE, K. O., GUENTHER, P. M. & DAVIS, C. A. 2013. Diet quality of Americans differs by age, sex, race/ethnicity, income and education level. *J Acad Nutr Diet,* 113**,** 297-306.

HOLLAND, B., WELCH, A. A., UNWIN, I. D., BUSS, D. H., PAUL, A. A. & SOUTHGATE, D. A. T. 1991. *McCance and Widdowson's The Composition of Foods,* Cambridge, Royal Society of Chemistry.

HOLSTEN, J. E. 2008. Obesity and the community food environment: a systematic review. *Public Health Nutrition,* 12**,** 397–405.

HOLT, E. 2011. Hungary to introduce broad range of fat taxes. *Lancet,* 378**,** 75.

HSCIC 2013. Statistics on Obesity, Physical Activity and Diet: England, 2013. NHS Information Centre.

JAIME, P. C. & LOCK, K. 2009. Do school based food and nutrition policies improve diet and reduce obesity? *Preventive Medicine,* 48**,** 45-53.

JAMES, W. P. T., NELSON, M., RALPH, A. & LEATHER, S. 1997. The contribution of nutrition to inequalities in health. *British Medical Journal,* 314**,** 1545-1549.

JENSEN, J. & SMED, S. 2013. The Danish tax on saturated fat - short run effects on consumption, substitution patterns and consumer prices of fats. *Food Policy,* 42**,** 18-31.

JIAO, J., MOUDON, A. V., ULMER, J., HURVITZ, P. M. & DREWNOWSKI, A. 2012. How to identify food deserts: measuring physical and economic access to supermarkets in King County, Washington. *Am J Public Health,* 102**,** e32-9.
KANT, A. K. & GRAUBARD, B. I. 2007. Secular trends in the association of socio-economic position with self-reported dietary attributes and biomarkers in the US population: National Health and Nutrition Examination Survey (NHANES) 1971-1975 to NHANES 1999-2002. *Public Health Nutrition,* 10**,** 158-167.
KANT, A. K. & GRAUBARD, B. I. 2013. Family income and education were related with 30-year time trends in dietary and meal behaviors of American children and adolescents. *Journal of Nutrition,* 143**,** 690-700.
KELLER, H. H. 2007. Promoting food intake in older adults living in the community: a review. *Appl Physiol Nutr Metab,* 32**,** 991-1000.
KESSE, E., CLAVEL-CHAPELON, F., SLIMANI, N., VAN LIERE, M. & GROUP, T. E. N. 2001. Do eating habits differ according to alcohol consumption? Results of a study of the French cohort of the European Prospective Investigation into Cancer and Nutrition (E3N-EPIC). *Am J Clin Nutr,* 74**,** 322-7.
KIM, D. & KAWACHI, I. 2006. Food taxation and pricing strategies to "thin out" the obesity epidemic. *American Journal of Preventive Medicine,* 30**,** 430-7.
KREBS-SMITH, S. M., GRAUBARD, B. I., KAHLE, L. L., SUBAR, A. F., CLEVELAND, L. E. & BALLARD-BARBASH, R. 2000. Low energy reporters vs others: a comparison of reported food intakes. *EUROPEAN JOURNAL OF CLINICAL NUTRITION,* 54**,** 281-287.
KREMERS, S. P. J., DE BRUIJN, G.-J., VISSCHER, T. L. S., VAN MECHELEN, W., DE VRIES, N. K. & BRUG, J. 2006. Environmental influences on energy balance-related behaviors: a dual-process view. *International Journal of Behavioral Nutrition and Physical Activity,* 3**,** 9.
LAFAY, L., MENNEN, L., BASDEVANT, A., CHARLES, M. A., BORYS, J. M., ESCHWEGE, E., ROMON, M. & GROUP, F. S. 2000. Does energy intake underreporting involve allkinds of food or only speci®c food items? Results from the Fleurbaix Laventie Ville Sante (FLVS) study. *International Journal of Obesity,* 24**,** 1500-1507.
LAKDAWALLA, D. & PHILIPSON, T. 2009. The growth of obesity and technological change. *Econ Hum Biol,* 7**,** 283-293.
LALLUKKA, T., PITKANIEMI, J., RAHKONEN, O., ROOS, E., LAAKSONEN, M. & LAHELMA, E. 2010. The association of income with fresh fruit and vegetable consumption at different levels of education. *European Journal of Clinical Nutrition,* 64**,** 324-327.
LARAN, J. & SALERNO, A. 2013. Life-history strategy, food choice, and caloric consumption. *Psychological Sciences,* 24**,** 167-173.
LAWLOR, D. A., BATTY, G. D., MORTON, S. M., CLARK, H., MACINTYRE, S. & LEON, D. A. 2005. Childhood socioecnomic position, educational attainment, and adult cardiovascular risk factors: the Aberdeen children of the 1950s cohort study. *Am J Public Health,* 95**,** 1245-1251.
LEAR, S. A., GASEVIC, D. & SCHUURMAN, N. 2013. Association of supermarket characteristics with the body mass index of their shoppers. *Nutrition Journal,* 12.
LEDIKWE, J. H., BLANCK, H. M., KHAN, L. K., SERDULA, M. K., SEYMOUR, J. D., TOHILL, B. C. & ROLLS, B. J. 2005. Dietary Energy Density Determined by Eight Calculation Methods in a Nationally Representative United States Population. *J Nutr.,* 135**,** 273-278.
LEE, J., RALSTON, R. A. & TRUBY, H. 2011. Influence of food cost on diet quality and risk factors for chronic disease: a systematic review. *Nutrition & Dietetics,* 68**,** 248-261.

LEHNERT, T., SONNTAG, D., KONNOPKA, A., RIEDEL-HELLER, S. & KONIG, H.-H. 2012. The long-term cost-effectiveness of obesity prevention interventions: systematic literature review. *Obesity Reviews,* 13.

LEUNG, P. & ET AL. 1997. Effect of Preparation Time on the Minimum-Cost Diet. *Journal of Consumer Affairs,* 31**,** 204-17.

LIMB, M. 2013. Pressure grows on government to act on obesity as figures show effect on health and the NHS. *BMJ,* 346**,** f1218.

LIPSKY, L. M. 2009. Are energy-dense foods really cheaper? Reexamining the relation between food price and energy density. *American Journal of Clinical Nutrition,* 90**,** 1397-1401.

LIPSKY, L. M. 2010. Reply to N Darmon and M Maillot. *AJCN,* 91**,** 1069-1070.

LO, Y.-T., CHANG, Y.-H., WAHLQVIST, M. L., HUANG, H.-B. & LEE, M.-S. 2012. Spending on vegetable and fruit consumption could reduce all-cause mortality among older adults. *Nutrition Journal,* 11**,** 113.

LOPEZ, C. N., MARTINEZ-GONZALEZ, M. A., SANCHEZ-VILLEGAS, A., ALONSO, A., PIMENTA, A. M. & BES-RASTROLLO, M. 2009. Costs of Mediterranean and western dietary patterns in a Spanish cohort and their relationship with prospective weight change. *Journal of Epidemiology & Community Health,* 63**,** 920-927.

MACINTYRE, S., MCKAY, L., DER, G. & HISCOCK, R. 2003. Socio-economic position and health: what you observe depends on how you measure it. *Journal of Public Health Medicine,* 25**,** 288-294.

MAFF (ed.) 1994. *Food Portion Sizes,* London: The Stationary Office.

MAILLOT, M., DARMON, N., DARMON, M., LAFAY, L. & DREWNOWSKI, A. 2007a. Nutrient-dense food groups have high energy costs: an econometric approach to nutrient profiling. *J. Nutr.,* 137**,** 1815–1820.

MAILLOT, M., DARMON, N., VIEUX, F. & DREWNOWSKI, A. 2007b. Low energy density and high nutritional quality are each associated with higher diet costs in French adults. *American Journal of Clinical Nutrition,* 86**,** 690-6.

MAILLOT, M., FERGUSON, E. L., DREWNOWSKI, A. & DARMON, N. 2008. Nutrient profiling can help identify foods of good nutritional quality for their price: a validation study with linear programming. *J. Nutr.,* 138**,** 1107–1113.

MAK, T. N., PRYNNE, C. J., COLE, D., FITT, E., ROBERTS, C., BATES, B. & STEPHEN, A. 2012. Assessing eating context and fruit and vegetable consumption in children: new methods using food diaries in the UK national diet and nutrition survey rolling programme. *Int J Behav Nutr Phys Act,* 9**,** 126.

MARMOT, M. & BELL, R. 2012. Fair society, health lives. *Public Health,* 126**,** S4-10.

MARTI-HENNEBERG, C., CAPDEVILA, F., ARIJA, V., PEREZ, S., CUCO, G., VIZMANOS, B. & FERNANDEZ-BALLART, J. 1999. Energy density of the diet, food volume and energy intake by age and sex in a healthy population. *European Journal of Clinical Nutrition,* 53**,** 421-428.

MCDONALD, J. H. 2009. *Handbook of Biological Statistics,* Baltimore, Sparky House Publishing.

MCDOWELL, D. R., ALLEN-SMITH, J. E. & MCLEAN-MEYINSSE, P. E. 1997. Food expenditures and socioeconomic characteristics: focus on income class. *American Journal of Agricultural Economics,* 79**,** 1444-1451.

MCLAREN, L. 2007. Socioeconomic status and obesity. *Epidemiologic Reviews,* 29**,** 29-48.

MENDEZ, M. A., WYNTER, S., WILKS, R. & FORRESTER, T. 2003. Under- and overreporting of energy is related to obesity, lifestyle factors and food group intakes in Jamaican adults. *Public Health Nutrition,* 7**,** 9-19.

MENDOZA, J. A., DREWNOWSKI, A., CHEADLE, A. & CHRISTAKIS, D. A. 2006. Dietary energy density is associated with selected predictors of obesity in U.S. Children. *Journal of Nutrition,* 136**,** 1318-22.

MICHAUD, P.-C., VAN SOEST, A. H. O. & ANDREYEVA, T. 2007. Cross-country variation in obesity patterns among older Americans and Europeans. *Forum for Health Economics & Policy,* 10.
MILJKOVIC, D., NGANJE, W. & CHASTENET, H. D. 2008. Economic factors affecting the increase in obesity in the United States: differential response to price. *Food Policy,* 33**,** 48-60.
MONSIVAIS, P. & DREWNOWSKI, A. 2009. Lower-energy-density diets are associated with higher monetary costs per kilocalorie and are consumed by women of higher socioeconomic status.[Erratum appears in J Am Diet Assoc. 2009 Jul;109(7):1296]. *Journal of the American Dietetic Association,* 109**,** 814-22.
MONSIVAIS, P., MCLAIN, J. & DREWNOWSKI, A. 2010. The rising disparity in the price of healthful foods: 2004–2008. *Food Policy,* 35**,** 514-520.
MONSIVAIS, P., PERRIGUE, M. M., ADAMS, S. L. & DREWNOWSKI, A. 2013. Measuring diet cost at the individual level: a comparison of three methods. *EUROPEAN JOURNAL OF CLINICAL NUTRITION,* 67**,** 1220-1225.
MURAKAMI, K., SASAKI, S., OKUBO, H., TAKAHASHI, Y., HOSOI, Y. & ITABASHI, M. 2007. Monetary costs of dietary energy reported by young Japanese women: association with food and nutrient intake and body mass index. *Public Health Nutrition,* 10**,** 1430-9.
MURAKAMI, K., SASAKI, S., TAKAHASHI, Y., OKUBO, H., HIROTA, N., NOTSU, A., FUKUI, M. & DATE, C. 2008a. Comparability of weighed dietary records and a self-administered diet history questionnaire for estimating monetary cost of dietary energy. *Environmental Health Insights,* 1**,** 35-43.
MURAKAMI, K., SASAKI, S., TAKAHASHI, Y., UENISHI, K. & GROUP, T. J. D. S. S. F. N. A. B. 2008b. Monetary cost of dietary energy is negatively associated with BMI and waist circumference, but not with other metabolic factors, in young Japanese women. *Public Health Nutrition,* 12**,** 1092-1098.
MUSHI-BRUNT, C., HAIRE-JOSHU, D. & ELLIOTT, M. 2007. Food spending behaviors and perceptions are associated with fruit and vegetable intake among parents and their preadolescent children. *J Nutr Educ Behav,* 39**,** 26-30.
MYTTON, O., CLARKE, D. & RAYNER, M. 2012. Taxing unhealthy food and drinks to improve health. *BMJ* [Online], 344. Available: http://dx.doi.org/10.1136/bmj.e2931 [Accessed 22 December 2013].
NAGGARA, O., RAYMOND, J., GUILBERT, F., ROY, D., WEILL, A. & ALTMAN, D. G. 2011. Analysis by Categorizing or Dichotomizing Continuous Variables Is Inadvisable: An Example from the Natural History of Unruptured Aneurysms *American Journal of Neuroradiology,* 32**,** 437-440.
NATCEN, MRC, RESOURCE CENTRE FOR HUMAN NUTRITION RESEARCH & SCHOOL, U. M. 2012. National Diet and Nutrition Survey, 2008-2010. 3rd ed. Colchester, Essex: UK Data Archive.
NELSON, M., ERENS, B., BATES, B., CHURCH, S. & BOSHIER, T. 2007. Low Income Diet and Nutrition Survey. London.
NHS. 2012. *The risks of drinking too much* [Online]. Available: http://www.nhs.uk/Livewell/alcohol/Pages/Effectsofalcohol.aspx [Accessed 6 July 2013.
NHS. 2013. *Alcohol units* [Online]. Available: http://www.nhs.uk/Livewell/alcohol/Pages/alcohol-units.aspx [Accessed 6 July 2013.
NI MHURCHU, C., BLAKELY, T., JIANG, Y., EYLES, H. C. & RODGERS, A. 2009. Effects of price discounts and tailored nutrition education on supermarket purchases: a randomized controlled trial. *American Journal of Clinical Nutrition,* 91**,** 736-747.

NNOAHAM, K. E., SACKS, G., RAYNER, M., MYTTON, O. & GRAY, A. 2009. Modelling income group differences in the health and economic impacts of targeted food taxes and subsidies. *International Journal of Epidemiology,* 38**,** 1324-1333.

NOO 2010. Adult obesity and socioeconomic status. *National Obesity Observatory data briefing.* NHS.

NORD, M. & KANTOR, L. S. 2006. Seasonal variation in food insecurity is associated with heating and cooling costs among low-income elderly Americans. *J Nutr,* 186**,** 2989-2944.

NUFFIELD COUNCIL ON BIOETHICS. 2007. *Public health: ethical issues* [Online]. Available: http://www.nuffieldbioethics.org/public-health [Accessed 22 December 2013.

ONS. 2009. Annual Survey of Hours and Earnings, 2009 Results. Available: http://www.ons.gov.uk/ons/rel/ashe/annual-survey-of-hours-and-earnings/2009-revised/index.html [Accessed July 24 2013].

ONS. 2011. *Consumer Price Indices* [Online]. Office for National Statistics. Available: http://www.ons.gov.uk/ons/rel/cpi/consumer-price-indices/index.html [Accessed July 19 2011].

OYSTON, C., RANSLEY, J. K., THOMAS, J. D. & CADE, J. E. 2005. Validation of the integral costing of DANTE – Diet and Nutrition a Tool for Evaluation. *Proceedings of the Nutrition Society,* 64**,** 102a.

PEREZ-ESCAMILLA, R., OBBAGY, J. E., ALTMAN, J. M., ESSERY, E. V., MCGRANE, M. M., WONG, Y. P., SPAHN, J. M. & WILLIAMS, C. L. 2012. Dietary Energy Density and Body Weight in Adults and Children: A Systematic Review. *Journal of the Academy of Nutrition and Dietetics,* 112**,** 671-684.

POORTINGA, W. 2007. The prevalence and clustering of four major lifestyle risk factors in an English adult population. *Prev Med,* 44**,** 124-128.

POPAY, J., ROBERTS, H., SOWDEN, A., PETTICREW, M., ARAI, L., RODGERS, M., BRITTEN, N., ROEN, K. & DUFFY, S. 2006. Guidance on the Conduct of Narrative Synthesis in Systematic Reviews. Economic and Social Research Council.

POWELL, L. M. 2009. Fast food costs and adolescent body mass index: Evidence from panel data☆. *Journal of Health Economics,* 28**,** 963-970.

POWELL, L. M., AULD, M. C., CHALOUPKA, F. J., O'MALLEY, P. M. & JOHNSTON, L. D. 2007. Access to fast food and food prices: relationship with fruit and vegetable consumption and overweight among adolescents. *Advances in Health Economics & Health Services Research,* 17**,** 23-48.

POWELL, L. M. & BAO, Y. 2009. Food prices, access to food outlets and child weight. *Economics & Human Biology,* 7**,** 64-72.

POWELL, L. M. & CHALOUPKA, F. J. 2009. Food prices and obesity: Evidence and policy implications for taxes and subsidies. *The Milbank Quarterly,* 87**,** 229–257.

POWELL, L. M., CHRIQUI, J. F., KHAN, T., WADA, R. & CHALOUPKA, F. J. 2013. Assessing the potential effectiveness of food and beverage taxes and subsidies for improving public health: a systematic review of prices, demand and body weight outcomes. *Obesity Reviews,* 14**,** 110-128.

POWELL, L. M. & HAN, E. 2011. Adult obesity and the price and availability of food in the United States. *American Journal of Agricultural Economics,* 93**,** 378-384.

PRENTICE, A. M. & JEBB, S. A. 2003. Fast foods, energy density and obesity: a possible mechanistic link. *Obesity Reviews,* 4**,** 187–194.

PROPER, K. I. & HILDEBRANDT, V. H. 2006. Physical activity among Dutch workers - differences between occupations. *Preventive Medicine,* 43**,** 42-45.

PUBLIC HEALTH ENGLAND 2013. Social and economic inequalities in diet and physical activity.

RANSLEY, J. K. 2002. *An Investigation into the Use of Supermarket Receipts as a New Method of Estimating Fat and Energy in the Household Diet.* PhD, University of Leeds.
RANSLEY, J. K., DONNELLY, J. K., BOTHAM, H., KHARA, T. N., GREENWOOD, D. & CADE, J. 2003. Use of supermarket receipts to estimate energy and fat content of food purchased by lean and overweight families. *Appetite,* 41**,** 141-148.
RAO, M., AFSHIN, A., SINGH, G. & MOZAFFARIAN, D. 2013. Do healthier foods and diet patterns cost more than less healthy options? A systematic review and meta-analysis. *BMJ Open* [Online], 3. Available: doi:10.1136/bmjopen-2013-004277 [Accessed 22 December 2013].
RAUBER, F. & VITOLO, M. R. 2009. Nutritional quality and food expenditure in preschool children. *Jornal de Pediatria,* 85**,** 536-40.
REBRO, S. M., PATTERSON, R. E., KRISTAL, A. R. & CHENEY, C. L. 1998. The effect of keeping food records on eating patterns. *Journal of the American Dietetic Association,* 98**,** 1163-1165.
REHM, C. D., MONSIVAIS, P. & DREWNOWSKI, A. 2011. The quality and monetary value of diets consumed by adults in the United States. *Am J Clin Nutr,* 94**,** 1333-9.
RENNER, B., SPROESSER, G., STROHBACH, S. & SCHUPP, H. T. 2012. Why we eat what we eat. The Eating Motivation Survey (TEMS). *Appetite,* 59**,** 117-128.
RENNIE, K. L., COWARD, A. & JEBB, S. A. 2007. Estimating under-reporting of energy intake in dietary surveys using an individualised method. *British Journal of Nutrition,* 97**,** 1169-1176.
RICCIUTO, L. E. & TARASUK, V. S. 2007. An examination of income-related disparities in the nutritional quality of food selections among Canadian households from 1986-2001. *Soc Sci Med,* 64**,** 186-198.
ROYSTON, P. 2005. Multivariable regression models with continuous covariates, with a practical emphasis on fractional polynomials and applications in clinical epidemiology. *Germand Stata Users' Group Meetings 2005.* Berlin.
RYDEN, P. & HAGFORS, L. 2011. Diet cost, diet quality and socio-economic position: how are they related and what contributes to differences in diet costs? *Public Health Nutrition,* 24**,** 1-13.
RYDEN, P., MATTSSON SYDNER, Y. & HAGFORS, L. 2008. Counting the cost of healthy eating: a Swedish comparison of Mediterranean-style and ordinary diets. *International Journal of Consumer Studies,* 32**,** 138-146.
SACN 2011. Dietary Reference Values for Energy. London: TSO.
SCARBOROUGH, P., BURG, M., FOSTER, C., SWINBURN, B., SACKS, G., RAYNER, M., WEBSTER, P. & ALLENDER, S. 2011. Increased energy intake entirely accounts for increase in body weight in women but not in men in the UK between 1986 and 2000. *Br J Nutr,* 105**,** 1399-1404.
SCHEIBEHENNE, B., MIESLER, L. & TODD, P. 2007. Fast and frugal food choices: Uncovering individual decision heuristics. *Appetite,* 49**,** 578-589.
SCHRODER, H., MARRUGAT, J. & COVAS, M. I. 2006. High monetary costs of dietary patterns associated with lower body mass index: a population-based study. *International Journal of Obesity,* 30**,** 1574-9.
SCHROETER, C., LUSK, J. & TYNER, W. 2005. Determining the impact of food price and income changes on obesity. *American Agricultural Economics Association Annual Meeting.* Providence, Rhode Island.
SCHUIT, A. J., VAN LOON, A. J., TIJHUIS, M. & OCKE, M. 2002. Clustering of lifestyle risk factors in a general adult population. *Prev Med,* 35**,** 219-224.
SCHUMANN, B., KLUTTIG, A., TILLER, D., WERDAN, K., HAERTING, J. & GREISER, K. H. 2011. Association of childhood and adult socioeceonomic indicators with

cardiovascular risk factors and its modification by age: the CARLA Study 2002-2006. *BMC Public Health,* 11**,** 289.
SHEPHERD, J., HARDEN, A., REES, R., BRUNTON, G., GARCIA, J., OLIVER, S. & OAKLEY, A. 2006. Young people and healthy eating: a systematic review of research on barriers and facilitators. *Health Education Research,* 21**,** 239-257.
SLOMAN, J. 1999. *Economics,* London, Financial Times/Prentice Hall.
SMYTH, A., CADE, J. E. & THOMAS, J. 2005. How much does what we eat cost? Validation of a new database approach to estimating food costs. *Proceedings of the Nutrition Society,* 64**,** 102a.
SOBAL, J. & STUNKARD, A. J. 1989. Socioeconomic status and obesity: a review of the literature. *Psychological Bulletin,* 105**,** 260-75.
STATACORP 2011. Stata Statistical Software: Release 11. College Station TSL.
STAUDIGEL, M. 2011. How (much) do food prices contribute to obesity in Russia? *Economics & Human Biology,* 9**,** 133-47.
STEPTOE, A., POLLARD, T. M. & WARDLE, J. 1995. Development of a measure of the motives underlying the selection of food: the food choice questionnaire. *Appetite,* 25**,** 267-84.
STEYN, N. P., PARKER, W., LAMBERT, E. V. & MCHIZA, Z. 2009. Nutrition interventions in the workplace: Evidence of best practice. *South African Journal of Clinical Nutrition,* 22 (3)**,** 111-117.
STURM, R. & DATAR, A. 2005. Body mass index in elementary school children, metropolitan area food prices and food outlet density. *Public Health,* 119**,** 1059-68.
STURM, R. & DATAR, A. 2008. Food prices and weight gain during elementary school: 5-year update. *Public Health,* 122**,** 1140-3.
STURM, R., POWELL, L. M., CHRIQUI, J. F. & CHALOUPKA, F. J. 2010. Soda taxes, soft drink consumption, and children's body mass index. *Health Affairs,* 29 (5)**,** 1052-1058.
SUSTAIN 2013. A Children's Future Fund: How food duties could provide the money to protect children's health and the world they grow up in. London.
SWINBURN, B., EGGER, G. & RAZA, F. 1999. Dissecting obesogenic environments: the development and application of a framework for identifying and prioritizing environmental interventions for obesity. *Preventive Medicine,* 29**,** 563-570.
SWINBURN, B., SACKS, G., HALL, K. D., MCPHERSON, K., FINEGOOD, D. T., MOODIE, M. & GORTMAKER, S. L. 2011. The global obesity pandemic: shaped by global drivers and local environments. *Lancet,* 378**,** 804-814.
THOMAS, D., LAVY, V. & STRAUSS, J. 1996. Public Policy and Anthropometric Outcomes in the Cote D'Ivoire. *Journal of Public Economics,* 61**,** 155-92.
TIMMINS, K. 2011. On with the pennies, off with the pounds? The use of taxation policies in obesity prevention. *Perspectives in Public Health,* 131**,** 155-157.
TIMMINS, K., HULME, C. & CADE, J. 2013a. The monetary value of diets consumed by British adults: an exploration into sociodemographic differences in individual-level diet costs. *Public Health Nutrition* [Online]. Available: dx.doi.org/10.1017/S1368980013002905 [Accessed 28 Nov 2013].
TIMMINS, K., MORRIS, M., EDWARDS, K., CLARKE, G. & CADE, J. 2013b. Comparability of methods assigning monetary costs to diets: derivation from household till receipts versus cost database estimation from four-day food diaries. *EUROPEAN JOURNAL OF CLINICAL NUTRITION,* 67**,** 1072-1076.
TOWNSEND, M. S., AARON, G. J., MONSIVAIS, P., KEIM, N. L. & DREWNOWSKI, A. 2009. Less-energy-dense diets of low-income women in California are associated with higher energy-adjusted diet costs. *Am J Clin Nutr,* 89.
TURRELL, G. & KAVANAGH, A. 2005. Socio-economic pathways to diet: modelling the association between socio-economic position and food purchasing behaviour. *Public Health Nutrition,* 9**,** 375-383.

VINER, R. M. T. J. 2005. Adult socioeconomic, educational, social, and psychological outcomes of childhood obesity: a national birth cohort study. *BMJ,* 330**,** 1354.
VON HINKE KESSLER SCHOLDER, S. & LECKIE, G. 2013. Income shocks, food expenditures, calorie intake and body weight: a multilevel structural equation modelling analysis. *Bulleting of the International Statistical Institute,* 68**,** 1-6.
WATERLANDER, W. E., DE HAAS, W. E., VAN AMSTEL, I., SCHUIT, A. J., TWISK, J. W. R., VISSER, M., SEIDELL, J. C. & STEENHUIS, I. H. M. 2010. Energy density, energy costs and income – how are they related? *Public Health Nutrition,* 13**,** 1599-1608.
WCRF 2007. Food, Nutrition, Physical Activity and the Prevention of Cancer: a Global Perspective. *Second Expert Report.*
WHITLEY, E. & BALL, J. 2002. Statistics review 4: Sample size calculations. *Critical Care,* 6**,** 335-341.
WHO. 2006. *BMI classification* [Online]. Available: http://apps.who.int/bmi/index.jsp?introPage=intro_3.html [Accessed 21 Mar 2013.
WHO 2007. The challenge of obesity in the WHO European Region and the strategies for response. *In:* BRANCA, F. N., H; LOBSTEIN; T (ed.). Geneva: The World Health Organization.
WILDE, P. E., LLOBRERA, J. & VALPIANI, N. 2012. Household Food Expenditures and Obesity Risk. *Current Obesity Reports,* 1**,** 123-133.
WILKINSON, R. & PICKETT, K. 2010. *The spirit level: why equality is better for everyone,* London, Penguin.
WILLETT, W. C. E. 1998. *Nutritional Epidemiology,* New York, Oxford University Press.
WINKLEBY, M. A., JATULIS, D. E., FRANK, E. & FORTMANN, S. P. 1992. Socioeconomic status and health: how education, income, and occupation contribute to risk factors for cardiovascular disease. *Am J Public Health,* 82**,** 816-820.
WINKLER, J. T. 2012. Why soft drink taxes will not work. *British Journal of Nutrition,* 108**,** 395-396.
WOODMAN, J., LORENC, T., HARDEN, A. & OAKLEY, A. 2008. Social and environmental interventions to reduce childhood obesity: a systematic map of reviews. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.
WRIEDEN, W. & BARTON, K. 2011. The Scottish Diet: Estimations of energy density and expenditure. Edinburgh.
WRIGLEY, N., WARM, D., MARGETTS, B. M. & WHELAN, A. 2002. Assessing the impact of improved retail access on diet in a 'food desert': a preliminary report. *Urban Studies,* 39**,** 2061-2082.
ZHANG, J., TEMME, E. H. M. & KESTELOOT, H. 2001. Alcohol Drinkers Overreport Their Energy Intake in the BIRNH Study: Evaluation by 24-Hour Urinary Excretion of Cations. *Journal of the American College of Nutrition,* 20**,** 510-519.
ZHANG, Q., CHEN, Z., DIAWARA, N. & WANG, Y. 2011. Prices of unhealthy foods, food stamp program participation, and body weight status among U.S. low-income women. *Journal of Family and Economic Issues,* 32**,** 245-256.
ZIMMERMAN, C. C. 1932. Ernst Engel's law of expenditures for food. *The Quarterly Journal of Economics,* 47**,** 78-101.

Appendix A

The following search strategies were used in the literature review in Chapter 2.

Search strategy 1: Literature on food prices or dietary expenditure/cost

1. exp Diet/ or Nutritional Requirements/
2. exp Diet, Fat-Restricted/ or Diet, Vegetarian/ or Diet, Mediterranean/
3. exp Diet Records/ or Diet Surveys/
4. exp Food Habits/ or Food Preferences/
5. ((soda or carbonated or sweet* or sugar* or soft) adj3 (beverage* or drink*) adj2 (purchase* or consumption or intake)).tw.
6. ((food* or diet*) adj2 (choice* or purchase* or consumption or selection or intake)).tw.
7. ((food* or eating or diet*) adj2 behavi*).tw.
8. ((energy or kcal or MJ or joule or calor* or fat) adj2 (intake or consum* or density)).tw.
9. (nutri* adj2 (intake or consum* or density)).tw.
10. or/1-9
11. Economics/ and (Food/ or Diet/ or Food Preferences/ or Food Habits/)
12. Food/ec [Economics]
13. diet/ec [Economics]
14. exp models, economic/ and (Food/ or Diet/ or Food Preferences/ or Food Habits/)
15. Programming, Linear/ and (Food/ or Diet/ or Food Preferences/ or Food Habits/)
16. (((financ* or monetary) adj2 cost*) and (food* or diet or kcal or MJ or joule or calor* or fast food or drink* or beverage* or fruit* or vegetable* or snack*)).tw.
17. Fees/ and Charges/ and (Food/ or Diet/ or Food Preferences/ or Food Habits/)
18. (food* adj3 (cost* or price* or pricing* or expenditure or spend* or budget*)).tw.
19. (fast food* adj3 (cost* or price* or pricing* or expenditure or spend* or budget*)).tw.
20. ((drink* or beverage*) adj3 (cost* or price* or pricing* or expenditure or spend* or budget*)).tw.
21. (snack* adj3 (cost* or price* or pricing* or expenditure or spend*)).tw.
22. (diet adj3 (cost* or price* or pricing* or expenditure or spend*)).tw.
23. ((energy or kcal or MJ or joule or calor*) adj3 (price* or pricing*)).tw.
24. (nutrient* adj3 (price* or pricing* or cost*)).tw.
25. ((food or fat or snack* or drink* or beverage*) adj3 tax*).tw.
26. ((food or fruit* or vegetable* or fat or drink* or beverage*) adj3 subsid*).tw.
27. ((food or fat or snack* or drink* or beverage*) adj3 (discount* or promotion*)).tw.
28. (food* adj3 (cost* or price* or pricing* or fiscal) adj2 (policy or policies)).tw.
29. ((shopping or market or supermarket or food or grocer*) adj2 basket* adj4 (cost* or price* or pricing* or expenditure or spend* or budget*)).tw.
30. ((price or demand or nutrient* or food*) adj2 elastic*).tw.

31. (((price or pricing) adj2 effect) and (food* or diet or kcal or MJ or joule or calor* or fast food or drink* or beverage* or fruit* or vegetable* or snack*)).tw.
32. (((price* or pricing) adj2 (change* or manipulation*)) and (food* or diet or kcal or MJ or joule or calor* or fast food or drink* or beverage* or fruit* or vegetable* or snack*)).tw.
33. or/11-32
34. exp obesity/
35. exp Body Weight Changes/
36. Overweight/
37. Nutritional Status/
38. exp Body Mass Index/ or Body Fat Distribution/ or Skinfold Thickness/ or Waist-Hip Ratio/
39. Obes*.tw.
40. Overweight.tw.
41. Overnutrition.tw.
42. (Overeat* or over-eat*).tw.
43. (weight adj3 (reduc* or maint* or control* or gain or loss or chang*)).tw.
44. (body adj3 (weight* or size or fat or mass)).tw.
45. (BMI or body mass index).tw.
46. (obes* adj3 (prevent* or control)).tw.
47. or/34-46
48. 10 or 47
49. 33 and 48
50. exp animals/ not (exp animals/ and exp humans/)
51. exp Veterinary Medicine/
52. exp Animal Experimentation/
53. exp Climatic Processes/
54. exp HIV Infections/
55. exp Drug Costs/
56. exp Food, Fortified/ec
57. Dietary Supplements/
58. Obesity/pp or Obesity/dt
59. Hypothalamic Diseases/
60. Weight loss/de or Eating/de
61. Food contamination/
62. or/50-61
63. 10 or 11 or 12 or 13 or 14 or 15
64. or/16-32
65. 63 or 47
66. 64 and 65

67. 66 not 62
68. limit 67 to english language

Search strategy 2: Literature on income

1. exp Diet/ or Nutritional Requirements/
2. exp Diet Records/ or Diet Surveys/
3. exp Food Habits/ or Food Preferences/
4. ((energy or kcal or MJ or joule or calor*) adj2 dens*).tw.
5. (nutri* adj2 dens*).tw.
6. or/1-5
7. Economics/ and (Food/ or Diet/ or Food Preferences/ or Food Habits/)
8. Food/ec [Economics]
9. diet/ec [Economics]
10. (income* or salar*).tw.
11. or/7-10
12. 6 and 11
13. exp animals/ not (exp animals/ and exp humans/)
14. exp Veterinary Medicine/
15. exp Animal Experimentation/
16. exp Climatic Processes/
17. exp HIV Infections/
18. exp Drug Costs/
19. exp Food, Fortified/ec
20. Dietary Supplements/
21. Obesity/pp or Obesity/dt
22. Hypothalamic Diseases/
23. Weight loss/de or Eating/de
24. Food contamination/
25. or/13-24
26. 12 not 25
27. limit 26 to english language
28. or/4-5
29. 11 and 28
30. 29 not 25
31. limit 30 to english language

Appendix B

Quality Assurance checks for comparability studies

UKWCS subsample

A Quality Assurance was carried out on a 10% random sample (n = 4) of the participants. Of the till receipts, entries for three of the sample were found to be 100% accurate, but there was a discrepancy of 7% in total expenditure for one participant. A further 10% random sample (n = 4) was then checked, showing agreement of 100%.

Diary entries of the first random sample above (n=4) were checked. Table A shows the discrepancies in energy intakes between the raw data and the electronically coded diaries (less than 5% of the total daily energy estimates).

Table A Difference in estimated daily energy intake (kcal) following Quality Assurance check of UKWCS diaries

ID	Total kcal d^{-1}	Error in kcal d^{-1}	% difference
75636	1486	17	1.14%
74221	1623	12	0.74%
53812	2426	-114	-4.7%
8850	1600	-23	-1.4%

SNIP sample

Results of the QA checks for till receipt totals from a 5% subsample of the SNIP study (n = 26; from 25 households) can be seen in Table B. All but two of the totals were within 5% of the original data entry.

Table B Difference in estimated daily expenditure per household member following Quality Assurance check of SNIP till receipts

Household ID	Cost per day per household member: original data entry (£)	Cost per day per household member: QA check (£)	Difference (£)	% difference
12	2.22	2.20	-0.02	-0.88%
15	2.72	2.96	0.25	9.08%
21	1.84	2.01	0.18	9.68%
25	4.44	4.48	0.04	0.94%
26	5.57	5.57	0.00	0.00%
30	3.16	3.15	-0.01	-0.33%

Household ID	Cost per day per household member: original data entry (£)	Cost per day per household member: QA check (£)	Difference (£)	% difference
32	3.52	3.52	0.00	0.10%
33	3.19	3.11	-0.07	-2.30%
35	3.52	3.49	-0.03	-0.97%
37	2.32	2.32	-0.01	-0.23%
42	3.48	3.48	0.00	0.00%
55	6.37	6.37	0.00	0.00%
57	4.23	4.12	-0.12	-2.77%
59	2.32	2.35	0.03	1.15%
67	3.32	3.23	-0.09	-2.83%
70	4.79	4.76	-0.04	-0.74%
75	2.85	2.84	-0.01	-0.43%
76	4.24	4.25	0.01	0.21%
136	3.54	3.54	0.00	0.03%
138	3.21	3.13	-0.09	-2.69%
139	4.36	4.36	0.00	0.04%
214	4.27	4.29	0.02	0.47%
218	5.50	5.42	-0.07	-1.35%
223	2.79	2.80	0.01	0.38%
225	2.51	2.53	0.03	1.00%

Table C shows the results from the QA check of diary entries for the 5% sample (n = 26) following data cleaning (see Chapter 5), with accuracy framed in terms of daily energy consumption estimates (kcal).

Table C Difference in daily energy intake (kcal) following QA check of SNIP diaries

ID	Total kcal d^{-1}	Error in kcal d^{-1}	% difference
00301011	2267	-103	-4.5%
00381011	1336	-181	-13.5%
00411011	1151	0	0%
00581011	984	0	0%
00691011	2950	-37.3	-1.3%
00701011	2444	-14	-0.01%
00801011	1552	-25	-2%
00861011	1867	-212	-11%
00871011	2202	-70	-3%
00931011	1563	-80	-5%

ID	Total kcal d^{-1}	Error in kcal d^{-1}	% difference
00981011	1469	-141	-9.6%
01111011	906	0	0%
01531011	1325	-345	-26%
01561011	1685	-184	-11%
01621011	988	0	0%
01831011	2192	9	0.4%
01911011	1444	531	37%
02021011	1365	-16	-1%
02051011	1701	-67	-4%
03541011	1717	-46	-3%
03551011	2770	0	0%
03571011	902	5	0.5%
03621011	1306	-95	-7%
05431011	788	-12.6	-1.6%
05521011	1084	-5	-0.5%
05631011	1165	3	0.3%

Appendix C

The DANTE food cost database was populated in 2004. The dietary data analysed in this thesis were collected in years different to the DANTE cost database population. To adjust for inflation, indices were applied to bring the food prices in line with those faced by the dietary survey participants. Table D shows the indices by food group that were applied to the DANTE cost database for the comparability study in Chapter 5 (1998/1999 index), and to the NDNS data (Chapter 6), to account for change in price over time. These were derived from national data (ONS, 2011), from indices used to calculate the Consumer Price Index (CPI). The food prices in the DANTE food cost database were multiplied by the appropriate food group index and divided by 100 to give the inflation-adjusted price.

Table D Food group-specific inflation indices (June 2004 = 100)

Food group	1998/1999* Index	2008/2009* Index	2009/2010* Index
Bacon	85.87	119.3	122.9
Beef	97.85	123.7	130.6
Biscuits & Cakes	95.30	119.8	122.8
Bread	86.78	133.4	135.4
Butter	96.62	133.6	133.4
Cereals	101.22	121.8	128.8
Cheese	93.78	123.6	126.8
Coffee	110.41	117.0	123.9
Eggs	93.88	154.1	161.2
Fish	95.08	122.8	129.1
Fresh Milk	88.42	134.9	141.2
Fruit	92.19	112.0	117.7
Fresh Fruit	92.04	113.3	118.3
Lamb	77.00	112.1	123.7
Milk Products	98.49	119.0	119.0
Oils & Fats	102.05	122.6	121.9
Other foods	99.23	108.3	113.2
Other meat	94.97	116.5	121.5
Pork	81.35	119.3	126.3
Potatoes	102.91	115.9	119.1
Poultry	95.85	112.6	113.3
Soft Drinks	101.38	110.4	114.5
Sugar & Preserves	95.34	117.2	129.9
Sweets & Chocolate	86.59	121.6	130.4
Tea	113.57	116.7	130.1
Vegetables	95.81	139.2	143.1
Alcohol	89.67	112.3	116.5
Total food (FPI)	93.91	118.3	123.1

* November to July averaged

Appendix D

The eatwell plate food groups

The eight food groups selected for use in Chapter 8 were based upon the food groups that make up the Department of Health's eatwell plate (DH, 2011). The eatwell plate is a graphical representation of UK dietary recommendations which indicates proportions of five basic food groups. These are:

- Bread, rice, potatoes, pasta and other starchy foods
    - Including breakfast cereals, oats, maize, cornmeal, polenta, millet, spelt, couscous, bulgur wheat, pearl barley, yams and plantains
- Fruit and vegetables
    - Including dried fruit, and fruit and vegetable juices
- Milk and dairy
    - Including milk, cheese, yoghurt, fromage frais, cottage cheese, cream cheese, Quark
- Meat, fish, eggs, beans and other non-dairy sources of protein
    - Including fresh, frozen or canned varieties of fish, eggs, nuts, beans and other pulses
- Foods and drinks high in fat and/or sugar
    - Including cakes, biscuits, chocolate, sweets, puddings, pastries, ice cream, jam, honey, crisps, butter, margarine and spreads, oil, cream, mayonnaise

For the analyses in Chapter 8, beverages were categorised into separate groupings to the above, to give two further food groups: non-alcoholic beverages and alcoholic beverages.

A final food group – 'miscellaneous' – was created for foods which did not fall into the basic groups described above. These were already coded as a 'miscellaneous' food group in the NDNS data set. Miscellaneous foods included: vinegar, Marmite, sauces and condiments, gravy thickener, soy sauce, herbs and spices, salt and pepper.