AN ECONOMIC ANALYSIS OF SOCIOECONOMIC VARIATION IN THE IMPACT OF OBESITY ON HEALTH AND HEALTH SERVICE USE

JONAS MINET KINGE

A thesis submitted for the degree of Doctor of Philosophy

UCL

Declaration

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

Publications and conference papers

At the time of submission of this thesis a paper consisting of substantial parts of Chapter 3 has been published in *Social Science & Medicine*. The complete reference is: Kinge, J.M., & Morris, S. (2010). Socioeconomic variation in the impact of obesity on health related quality of life. *Social Science & Medicine,* 71, 1864-1871.

Earlier versions of chapter 3, 4 and 5 have been presented at various conferences, detailed as follows:

Kinge, J.M., & Morris, S. The impact of socioeconomic status on the relationship between obesity and health care utilisation. Nordic Health Economists' Study Group meeting, Umeaa University, Sweden, 2010

Kinge, J.M., & Morris, S. Socioeconomic variation in the impact of obesity on life expectancy in the UK. Health Economists' Study Group meeting, University College Cork, Ireland, 2010

Kinge, J.M., & Morris, S. Socioeconomic status, obesity and health-related quality of life in England. Nordic Health Economists' Study Group meeting, University of Iceland, Iceland, 2009

Kinge, J.M., & Morris, S. The impact of socioeconomic status on the relationship between obesity and health related quality of life in England. Health Economists' Study Group meeting, University of Sheffield, England, 2009

Acknowledgements

This thesis is dedicated to my parents.

First and foremost I would like to thank my supervisor Stephen Morris. His comprehensive knowledge, dedication, and constructive approach have been of irreplaceable value to me. He has truly inspired and encouraged me and I could not imagine having a better advisor and mentor.

I would like to thank my second supervisor Emmanuel Stamatakis and my fellow PhD students, especially Laura Vallejo-Torres and Ed Kendall, for their practical advice and support in each stage of this process. Many thanks also to my other colleagues at the HCEG. Furthermore, I want to thank Joachim Lund, Holger Roth, and Torunn Erichsen for their encouragement.

Finally and very importantly, I would like to thank my parents, my brother and my sisters for supporting me in every possible way.

ABSTRACT

This thesis examines the relationship between obesity and health and health service use, and whether or not this relationship varies by socioeconomic status (SES). We start by generating hypotheses for this relationship based on a human capital model for health. We conduct a literature review on the topic and generate econometric models to be tested. We then provide an analysis of socioeconomic variation in the relationship between obesity and Health Related Quality of Life (HRQL). The results show that obesity is negatively associated with HRQL and that the negative association is more pronounced in lower SES individuals than higher SES individuals. We then conduct an analysis of socioeconomic variation in the association between obesity and HRQL in individuals with obesity-related comorbidities. The results show that the association between obesity and HRQL is more negative in individuals with these comorbidities than individuals without. Furthermore, it shows that obesity-related comorbidities are associated with greater reductions in HRQL in lower SES groups than in higher SES groups. We also examine SES variation in the impact of obesity on life expectancy and find that obesity increases mortality and reduces life expectancy. We find that the impact of obesity on mortality is more negative in lower SES women than in higher SES women. Mortality does not vary by SES groups in men. We then conduct an analysis of SES variation in the association between obesity and health service use. We use a range of health service indicators and find that obesity is associated with increased use. We also find that obesity will lead to greater use of some services in lower SES groups compared with higher SES groups. The main implication of this thesis is to illustrate that obesity can constitute different challenges across SES groups.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Introduction

In this chapter we start by explaining the background and motivation for the work in this thesis. We then go through the research aims and discuss the approaches we will use to reach these aims.

1.1 Background

In this thesis we examine the relationships between obesity and health, and obesity and health service use. Then we examine if these relationships vary by socioeconomic status (SES). We start by describing obesity and SES separately, before we discuss their relationship and how they might interact.

1.1.1 Obesity

Overweight and obesity are clinical terms used to describe excess body fat (Department of Health, 2011). Obesity is a state of being seriously overweight to the stage where it becomes a risk factor for poor health and the most common method of measuring obesity is Body Mass Index (BMI), measured as weight in kilogrammes divided by height in metres squared (kg/m^2). We use BMI and have grouped the BMI

categories according to World Health Organization guidelines (WHO, 1995), which is shown in Table 1.1.

Table 1.1: WHO BMI classification

Classification	**BMI (kg/m^2) cut points**
Underweight	<18.5
Normal weight	18.5-24.99
Overweight (pre-obese)	25-29.99
Obese Class I	30-34.99
Obese Class II	35-39.99
Obese Class III	≥40

Note
In the following analysis we pool some of the classes to ensure an adequate number of individuals in each class when we stratify by SES.

BMI has been criticised, e.g., because it does not incorporate body fat, which is an independent predictor of ill health (Burkhauser & Cawley, 2008). There are a range of other measures available: total body fat; percent body fat; waist circumference; and waist-to-hip ratio. Each of these measures has their strengths and weaknesses, and there is no consensus which measure of fatness is the best (Freedman & Perry, 2000; Burkhauser & Cawley, 2008). We have chosen to focus on BMI as it is the preferred measure by the WHO and is most widely used in social science research related to obesity (Burkhauser & Cawley, 2008). In our case this is relevant as we seek to compare our results to other studies.

We define obesity as a lifestyle related risk factor for poor health. By doing this we follow both public health traditions (Butland et al., 2007; Marmot, 2010; WHO, 2011) and health economics traditions (Contoyannis & Jones, 2004; Balia & Jones, 2008). Other harmful lifestyle related factors are smoking, excess drinking and drug use

(Marmot, 2010). However, the complicated nature of obesity distinguishes it from the other risk factors; it is a consequence of sedentary behaviour and excess energy consumption, it is not a description of the behaviour in itself. Not only is obesity a risk factor, it can be seen as an intermediate indicator of health. In the following we recognise that obesity has two effects on health. The first is the direct effect, which is, for example, reduced mobility. The second effect is the (indirect) risk factor effect. This is the increased risk of diabetes and other comorbidities caused by obesity, which then again harms health.

Rising obesity levels are a major issue in the UK. In England in 1993 15% of adults aged 16 years and over were obese (13% of men, 16% of women); by 2008 this figure had risen to 25% (24% of men, 25% of women) (National Centre for Social Research and Department of Epidemiology and Public Health, (UCL), 2009). The British government's Foresight programme shows that over half of the British adult population could be obese by 2050 (McPherson, Marsh, & Brown, 2007). This trend is worrying because obesity is, as mentioned, an important risk factor for a number of diseases including coronary heart disease, type II diabetes, hypertension and stroke (NHLBI, 1998). In England 7% of all deaths are attributable to obesity (House of Commons Health Committee, 2004). Obesity is also associated with decreased life expectancy. For instance, in the UK a 30-year old non-smoking man with a BMI of 35 kg/m^2 is projected to lose five years of life compared to a similar person with a BMI of 24 kg/m^2 (Mayhew, Richardson, & Rickayzen, 2009). The analogous result for women is a loss of two years. Results of a similar order of magnitude have been found in other countries (see, e.g., Peeters et al., 2003). As well as affecting premature mortality and life expectancy there

is also increasing evidence that obesity is associated with a loss in health related quality of life (HRQL) (see, e.g., Sach et al., 2007). Hence, obesity is a major public health problem presenting a direct and indirect burden on individuals and is also proven to have major implications for health service expenditures (National Audit Office, 2001). In 2002 the House of Commons Health Select Committee estimated that the total cost of obesity was close to £7 billion per year (including both direct and indirect costs) and that the direct health care costs of obesity constituted between 2.3% and 2.6% of the total annual NHS expenditure (House of Commons Health Select Committee, 2004). To generate sustainable responses to this problem we need evidence based, cost-effective solutions.

1.1.2 Socioeconomic status

The Acheson report (1998) highlights socioeconomic position as the most fundamental of the social determinants of health. In addition, socioeconomic disadvantage has been identified as a major reason for why other groups like some ethnic minorities have higher rates of poor health (Graham, 2004).

Monitoring and understanding socioeconomic conditions requires conceptual clarity about what socioeconomic parameters we are measuring, and why (NICE, 2008). A range of terms has been used (social class, social stratification, socio-economic position, socio-economic circumstances, deprivation), to describe variations in the living and working conditions of individuals, households or the areas. We base the definition of SES on a definition by Krieger (1997, p.345) of socioeconomic position:

"An aggregate concept that includes both resource-based and prestige-based measures, as linked to both childhood and adult social class position. Resource-based measures refer to material and social resources and assets, including income, wealth, educational credentials; terms used to describe inadequate resources include "poverty" and "deprivation". Prestige-based measures refer to individual's rank or status in a social hierarchy, typically evaluated with reference to people's access to and consumption of goods, services, and knowledge, as linked to their occupational prestige, income, and education level"

Strictly speaking this is a definition of socioeconomic position. The difference between socioeconomic status and socioeconomic position lies in the word "status". Some prefer not to use "status" as it refers to prestige and rank related characteristics, while "position" refers to actual resources (Krieger, Williams, & Moss, 1997). However, they are measuring the same thing and it will be referred to as socioeconomic status (SES) in the following text as this is more common in economic literature.

1.1.3 Obesity and socioeconomic status

Lifestyles that lead to obesity are becoming increasingly concentrated among lower SES groups (See, e.g., McLaren, 2007). Figure 1.1 and Figure 1.2 show the development of obesity across SES groups in England for men and women. We can see that obesity is steadily increasing across all social classes, however the share of obese individuals is higher in the lower social classes compared with the higher social classes. The social class gradient is more pronounced for women where professional women had an obesity prevalence of less than 15% in 2008 compared with almost 35% in the

unskilled manual classes. We can also see that the social gradient in the prevalence of obesity is consistent across time and appears to be increasing.

Figure 1.1: Prevalence of obesity in men (aged 16 and over) by social class, 1997-2008

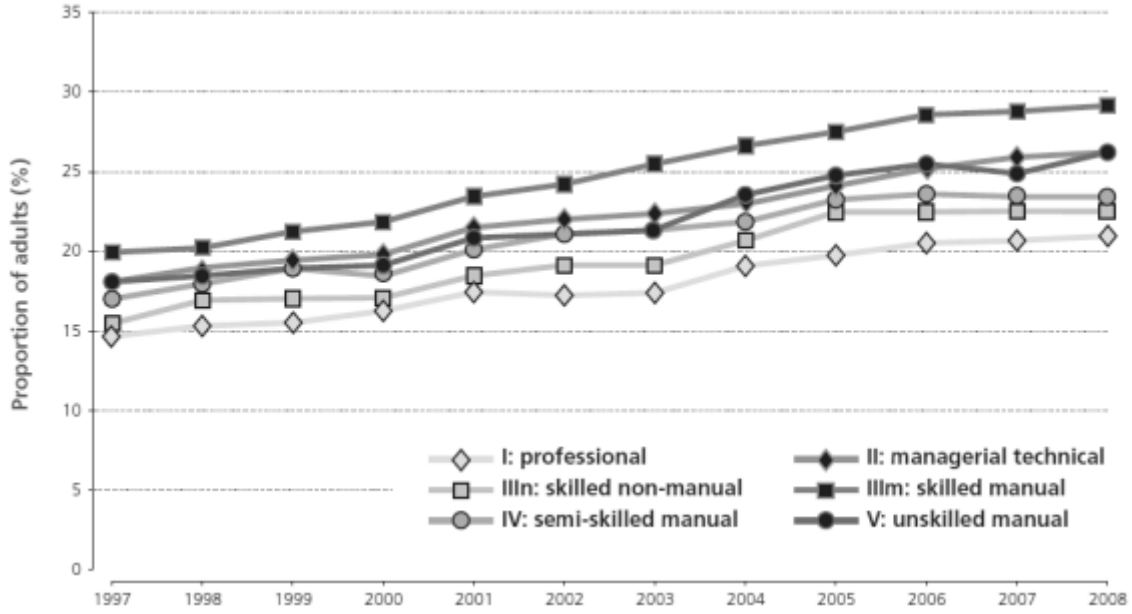
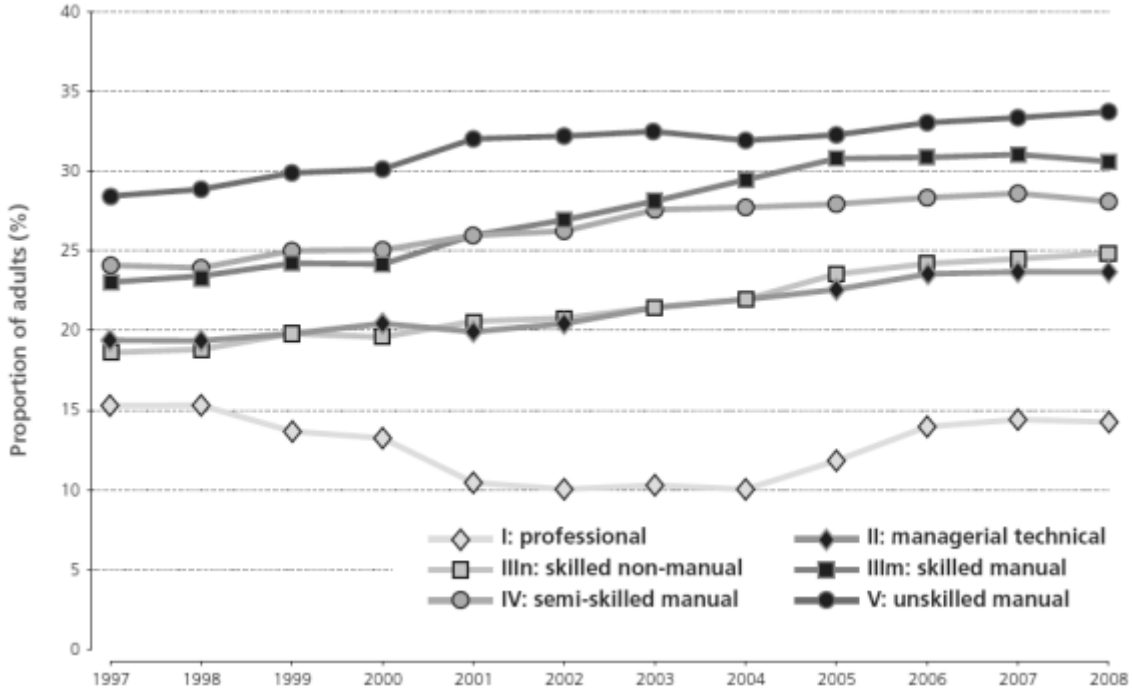


Figure 1.2: Prevalence of obesity in women (aged 16 and over) by social class, 1997-2008



Source: National Obesity Observatory, based on the Health Survey for England

Tackling inequalities in health is a primary aim in the English National Health Service (NHS). Hence, the socioeconomic pattern in the prevalence of obesity and other lifestyles has lead to a focus counteracting these lifestyles in lower SES groups. Many of the social policies implemented to address inequalities over recent years have been targeted at the most disadvantaged groups or areas (Marmot, 2010). For example, there has been a focus on increasing life expectancy in the so called spearhead group[1]. A similar concept to targeting is proportionate universalism (Marmot, 2010). This is universal interventions; however there is a social gradient in the distribution of resources, so that the resource use is proportionate to level of disadvantage. Furthermore, the Department of Health's White Paper (2010) suggests "A radical new approach" to increase the population health and decrease inequalities. In this approach they suggest that "the Government will consider different approaches for different groups of the population, taking account of the significant barriers that some people face" (Department of Health, p. 28, 2010). They continue to argue that local areas should develop tailored approaches that particularly target at risk groups and Directors of Public Health should be responsible for addressing inequalities in health by addressing the wider determinants of health through interventions aimed at smoking, alcohol consumption and obesity.

Based on the reasoning of the Department of Health (2010) and Marmot (2010), obesity may well be an area where it could be recommended to introduce

[1] Spearhead PCTs are the most health deprived areas of England. They are areas in the bottom fifth nationally for three or more indicators relating to life expectancy at birth, cancer and CVD mortality and the index of multiple deprivation.

interventions targeted at lower SES groups or spend proportionally more resources in lower SES groups. Furthermore, the Foresight report mentions as one of their "top five policy responses to obesity" was targeting health interventions for those at increased risk, such as low income groups (Butland et al., 2007).

By using interventions targeted at vulnerable groups to tackle inequalities in health one makes assumptions about a homogenous condition across heterogeneous groups. If the prevalence of a problem is concentrated to a defined group, and we aim to target this group, the evidence required to inform decision makers should be from this group. Traditionally, it has been assumed that determinants of health (such as obesity, smoking and SES) are separate entities in their contribution to health and health service use. Limited evidence exists on if, and if so how, these determinants are combined and interact with each other. However, there is little reason to believe that unhealthy lifestyles are equally associated with health in high and low SES groups, and that this would have equal consequences for health service use. For example, low SES individuals may be less harmed from unhealthy lifestyles because, given the general health risk produced by limited resources, they have less to lose from damaging lifestyles. Conversely, low SES individuals might experience greater harm from poor lifestyles because poor material circumstances increase the health problems experienced from poor lifestyles. In other words there might be an interaction between poor lifestyle and SES in their relationship with health/health service use[2].

[2] This will be discussed further in Chapter 2, where we will present theories and discuss previous literature.

While the impacts of obesity on health and health service use have been investigated both in the UK (See, e.g., Macran, 2004; Sach et al., 2007; and, the Counterweight Project Team, 2005b) and in other countries (Jia & Lubetkin, 2005; and, Reidpath, 2002), little attention has been given to variations in these impacts across socioeconomic groups. Although we have information about prevalence of obesity across SES groups, see for example Figure 1.1 and 1.2, there is a need for information of the consequences of obesity across SES groups. Even if an obesity intervention was equally effective in terms of the change in obesity prevalence across SES groups the outcomes in terms of health might differ. Hence, previous studies not taking account of these interactions may fail to reflect the underlying reality of the relationship between various determinants and health/health service use when stratified by SES. This is important as it might have implications for how resources are distributed to improve efficiency of the health care system and tackle socioeconomic equity in health.

The *Acheson Report* (1998) highlighted as one of its key recommendations "that as part of health impact assessment, all policies likely to have a direct or indirect effect on health should be evaluated in terms of their impact on health inequalities, and should be formulated in such a way that by favouring the less well-off they will, wherever possible, reduce such inequalities". Assessing the impact of interventions (both targeted and universal) on health inequalities depends on clarity about how determinants link up to influence overall health and its unequal distribution. Public health interventions to reduce harmful lifestyles might be associated with increasing the overall health of the population, but have unknown effect on the overall

distribution of health across SES groups within the population. Furthermore, in economic evaluation of obesity interventions we might base our assumptions on biased measures of costs and effects, which could compromise economic efficiency in the health care system.

In summary obesity is a major concern in the UK, which continues to increase. Furthermore, obesity has a socioeconomic gradient and is identified as a determinant of health inequalities (Department of Health, 2010; Marmot, 2010). In counteracting health inequalities current government reports (e.g., the White Paper and Marmot review) recommend focusing on the lower SES groups either through targeting or spending proportionally more on lower SES groups. Furthermore, obesity specific reports recommend SES targeted interventions as means of tackling obesity in general. Little attention has been given to possible interactions. If one assume independence between health determinants like obesity and SES one may fail to reflect the underlying reality of production of health in populations, hence misinforming policy makers and misallocating resources (Birch et al., 2000). However, as will be shown in a literature review, in the following Chapter 2, little is known about heterogeneities between the determinants of health. In an attempt to add to this topic we analyse SES variation in the impact of obesity on health and health service use.

1.2 Aims and overview of thesis

1.2.1 Aims

In this thesis we want to investigate the consequences of obesity, in terms of health and health serviced use, and whether or not these consequences vary by SES groups. More specifically there are two general aims of this thesis:

1. Examine the relationships between obesity and health outcome; and obesity and health service use in the UK

2. Investigate whether or not there are socioeconomic variations in these relationships

In the thesis we focus on obesity and socioeconomic status. We have chosen to focus on adults (age$\geq$16) throughout this thesis. For individuals under the age of 16 we would need to apply a different set of socioeconomic measures (e.g., parent's education and income) and obesity measures (percentage above 95th BMI percentiles in the specific age group) which are less frequently available. Furthermore, we do not have measures of HRQL (since only individuals over the age of 16 are requested such information) and life expectancy (limited data where the follow up period is long enough) for this group.

1.2.2 Overview of thesis

Before a thorough empirical analysis, we construct a framework for examining the impact of obesity on health and health service use and how this varies by SES. We present the Commission on Social Determinants of Health (CSDH) framework for determinants of health and present a human capital model for health (the Grossman model). We use these models to analyse the impact of lifestyle related risk factors (like obesity) on health, and how they might vary by SES. We conduct a literature review on SES variation in the association between lifestyle related risk factors and health. The review shows that few studies have explored the manner in which SES and behavioural risk factors influence health and are characterised by inconsistent findings. In addition, limited research has been done on obesity and there is a complete absence of UK research investigating heterogeneities in the impact of obesity on health by SES groups. We then present a model for the relationship between obesity and health service use and how this relationship might vary by SES. Based on the models we develop econometric models for testing.

The preferred measure of health effects in the NICE public health reference case is Quality Adjusted Life Years (QALYs) (NICE, 2008). QALYs are essentially a composite measure consisting of HRQL and life expectancy; hence it takes into account both the quality and the quantity of life across health states. We have chosen to use HRQL (EQ-5D) and life expectancy as our measures of health in this thesis. Hence, we want to analyse the relationship between obesity and HRQL and how this varies by SES. We then also want to analyse the relationship between obesity and life expectancy and how this varies by SES.

Chapter 3, 4, 5 and 6 provide the empirical analysis of SES variation in the consequences of obesity. Chapter 3 investigates the relationship between HRQL and obesity, and whether or not this relationship varies by SES. Data is taken from four rounds of the *Health Survey for England* (2003–2006) for persons aged 16 and above. Banded total annual household income is regressed against a comprehensive set of SES indicators using interval regression. We use the equivalised predicted values from this model, categorised into quartiles, as our measure of SES. We regress HRQL (EQ-5D scores) against interactions between BMI and SES categories controlling for a range of covariates and generate predictions of mean EQ-5D scores. Obesity is negatively correlated with HRQL. The negative association between obesity and HRQL is greater in people from lower SES groups. Overweight and obese people in lower SES groups have lower HRQL than those of normal weight in the same SES group, and have lower HRQL than those in higher SES groups of the same weight. This trend is also observed after controlling for individual and household characteristics, although the statistical significance and magnitude of effects is diminished.

In Chapter 4 we investigate how obesity-related comorbidities modify the association between obesity and HRQL, and then how socioeconomic status influences this. We use the same dataset and SES groups as in Chapter 3. We regress HRQL (EQ-5D score) against interactions between BMI groups and obesity-related comorbidities, controlling for a range of covariates. We then include and interact with SES groups. We generate predictions of mean EQ-5D scores for each BMI/comorbidity/SES combination. The findings from this chapter suggest that the association between

obesity and HRQL is more pronounced in individuals suffering from obesity-related comorbidities. This chapter also finds that the relationship between obesity-related comorbidities and HRQL is more pronounced in lower SES groups.

In Chapter 5 we analyse the relationship between obesity and life expectancy. We use data from the *British Health and Lifestyle Survey* (1984-1985) and the longitudinal follow-up in June 2009 to investigate the impact of obesity on life expectancy for persons aged 40 and above. Our measure of SES is based on a prediction of income as in previous studies. We use both parametric and semi-parametric duration models and account for unobservable heterogeneity. We use a range of statistical tests to identify the parametric survival distribution that best fits our data. We look at both mortality and predicted life expectancy. Obesity increase mortality and is negatively correlated with life expectancy in men and younger women. In older women obesity is associated with increased life expectancy. The negative impact of obesity on mortality varies by SES groups in women, but not in men. The impact of obesity on life expectancy is more negative in lower SES women than in higher. This trend is also observed after controlling for longstanding illnesses, although the statistical significance and magnitude of effects is diminished.

Although we do not focus on a specific intervention and cannot calculate the cost of an intervention to the NHS we can calculate the general association between obesity and services provided by the NHS. This is done by analysing the association between obesity and health service use and how this varies by SES. We use data from ten rounds of the *Health Survey for England* (HSE: 1999-2008) and two rounds of the

British Household Panel Survey (BHPS: wave 14 (2004 - 2005) and wave 16 (2006 - 2007)) to investigate the association between obesity and a range of health services including medication categories based on the British National Formula (BNF) in Great Britain and how this varies by SES. Total annual household income is regressed against a comprehensive set of SES indicators in both datasets using panel and cross sectional regression methods. As in earlier chapters, we use the equivalised predicted values from these models, categorised into quartiles, as our measure of SES. We use a range of statistical methods including interval regression, nonlinear binary choice and count models to investigate the relationship between obesity and use. We make predictions of use based on the models. The methods and predictions are adjusted for data being either panel or cross sectional. We find that obesity is associated with increased use of a range of health services and medications. Furthermore, obesity in lower SES groups is associated with a greater increase in use than obesity in higher SES groups. However, the SES gradient was not identified for all health services. The findings are consistent across the datasets.

Chapter 7 concludes by pulling together the findings of the first six chapters. We discuss implications for policy and research. Furthermore, we discuss limitations and offer some suggestions for further research.

1.3 Summary

Obesity is a major public health problem that needs evidence based, cost-effective responses. Obesity is more prevalent in lower social classes and this social gradient has

persisted over time. This has led to a focus on counteracting obesity in lower SES groups. We argue that there is no reason to believe that obesity affects individuals across SES groups in the same way. The consequences of obesity might vary by SES groups and therefore so too might the outputs in terms of health and health service use. Furthermore, we summarise the methods we will use in meeting the general aims of this thesis. The following chapter will go into detail on earlier literature, theories and models mapping socioeconomic variation in the relationship between obesity and health/health service use.

CHAPTER 2

Modelling socioeconomic variation in the impact of obesity on health and health service use

2.1 Introduction

Generally health economic research and modelling assumes a production relationship between health and the determinants of health, i.e. an improvement in the determinants of health is seen to produce better health. Determinants of health can be lifestyle related risk factors, like smoking, diet and exercise, and it is also increasingly recognised that social, cultural and economic factors are associated with health status (see, e.g., Marmot, 2010). Traditionally little attention has been given to interactions between these determinants and it is assumed that health is produced from a number of determinants that all have a separate impact on health. However, the impact of one determinant might depend on other determinants.

The aim of the following chapter is to present a conceptual framework for the following analysis using a theoretical model. In addition, to construct hypotheses for SES variation in the impact of obesity on health and health service use based on this framework. This framework will form the basis for empirical models for testing.

To illustrate how SES and obesity affect health we build on earlier literature from both health economics and public health traditions. We start by explaining a conceptual framework for social determinants of health. We then build a theoretical model to illustrate socioeconomic variation in the impact of obesity on health. The theoretical model will be based on the Grossman model, which has been (and still is) the most influential economic model for understanding how variables such as age and lifestyle influence the demand for health. Further, we show a framework first presented by Birch et al. (1997) for socioeconomic variation in the impact of risk factors on health. We then generate hypotheses for how SES and lifestyle related risk factors interact and its consequences for health and discuss the hypotheses in light of a literature review. Last, we generate a model to illustrate socioeconomic variation on the impact of obesity on health service use and generate hypotheses for this relationship.

2.2 A conceptual framework for the social determinants of health

The Commission for Social Determinants of Health (CSDH) was a commission lead by Prof. Sir Michael Marmot set up by the WHO and was tasked to collect, collate, and synthesize evidence on the social determinants of health and their impact on health inequity, and to make recommendations for action to address that inequity (CSDH, 2008). The CSDH has developed a framework for the pathways and mechanism through which social determinants of health influences health (Solar and Irvin, 2007). This framework has three main elements consisting of different factors and is shown in Figure 2.1.

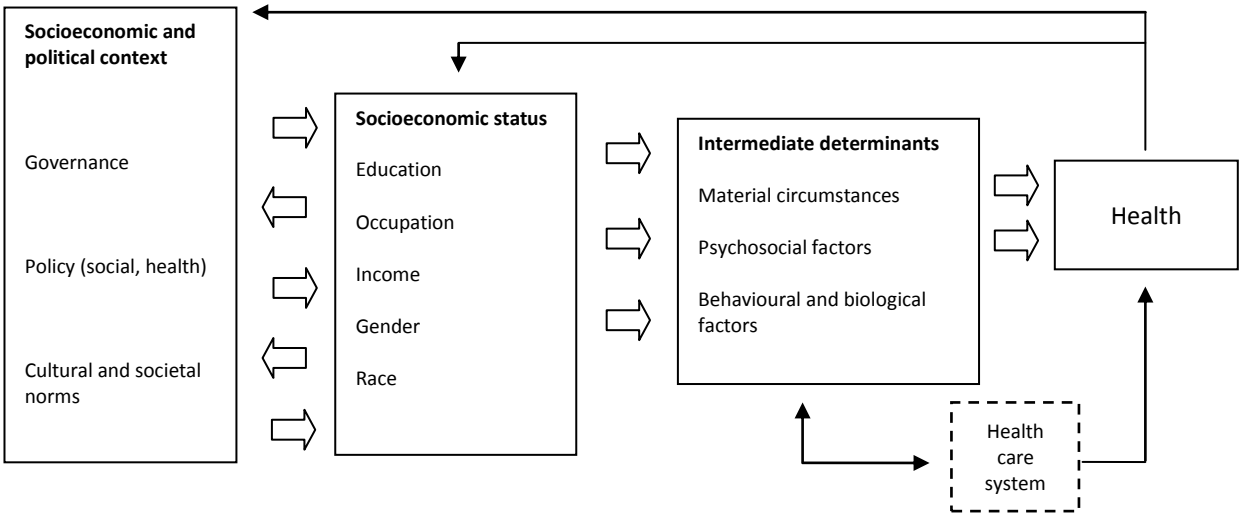
The first element *is socio-economic and political context*. These are factors that cannot be directly measured at the individual level. The factors are related to the function and the structure of the social system that influence people's health opportunities and the social gradient in health. Like a country's governance and social policy.

The second element of the framework consists of *structural determinants and socioeconomic status*[3]. Structural determinants are those who generate stratification and social class diversion in the society. They identify the most important structural stratifiers as: income, education, occupation, social class, gender and ethnicity.

The third element is the *intermediate determinants*. The main categories are material circumstances (e.g., housing quality and consumption potential), psychosocial circumstances (e.g., stressful living and social support), behavioural and biological factors (e.g., obesity and smoking).

[3] They use the term socioeconomic position, however we refer to this as socioeconomic status for consistency

Figure 2.1: The Commission on Social Determinants of Health conceptual framework



Source: Amended from Solar & Irwin, 2007

Figure 2.1 illustrates the main process in the CSDH framework. Moving from the left to right we see that the social and political context generates the socioeconomic positions and social classes. The socioeconomic position will then affect intermediate determinants which then affects health. The diagram also illustrates the reversed effects through which illness might affect socioeconomic position and widely prevalent diseases might affect social and political institutions.

The health care system is modelled as a separate element. The reason for this is that the health care system can influence health directly and it can alter the prevalence of intermediate determinants (for example, the GP might help a patient to quit smoking).

This framework is interesting as it illustrates that it is not SES in itself that influences health but it is behaviour and possibilities resulting from SES. For instance, income

increase material circumstances and education improves choice of diet, while it is the material circumstances and the diet which has an impact on health. In addition, it emphasises the importance of country specific data by including the political context in addition to the health care system. In the following we will use the Grossman model to focus on a part of this framework, which is the relationship between *socioeconomic status*, *intermediate determinants* and health. By using the Grossman model we will use economic theory to explain in detail how the determinants link up and heterogeneities between them.

2.3 A human capital model for health

We have now briefly discussed a framework that provides a basis for how health can be influenced by its separate determinants. In economic theory the Grossman model (Grossman, 2000) has been especially influential in understanding the impact various determinants on general health have. Not only has this model provided considerable insight into the allocation of resources (like time and money) into activities that generate health. It has also crossed academic fields and it is backed by empirical results, see Grossman (2000) for a discussion of empirical evidence. We will start by explaining the Grossman model framework. We will then use this framework to analyse how obesity and SES might influence health.

The Grossman model is a model derived from human capital theory where individuals increase their health capital by investments in health. The main contribution of the Grossman model is that the demand for health care is derived from its ability to

produce health capital and not for the direct utility. This means that one does not gain utility directly from antibiotics. One will gain utility from the increased health capital (reduced duration/impact of illness) resulting from taking this drug. In this way this model introduces a production relationship between health capital and multiple influences on health. Hence, health capital is both demanded and produced by individuals. We start with a simple representation of a utility function:

$$U = U(H, O) \tag{2.1}$$

Equation 2.1 illustrates that people gain utility from two goods: health capital (*H*), which in this case is measured as the sum of the number of healthy days over a lifetime, and consumption of all other goods (*O*). The number of healthy days is derived from the individuals health stock (*HS*) (*H=f(HS)*). The health stock at time t is determined by health stock in the previous period (HS_{t-1}), the depreciation in the health stock (d_t) and investments in health (I_t).

$$HS_t = HS_{t-1} - d_t + I_t \tag{2.2}$$

Health stock depreciates over time because of effects like aging and poor lifestyles. Depreciation can be seen as "wear and tear" of health and can be offset by investments in health through, for example, eating more fruit or consuming health care. The larger the depreciation is, the larger investment is needed to offset this depreciation. Both I_t and O_t are produced in the household and have a similar pattern:

$$I_t = I(M_t, T_{Ht}) \quad (2.3)$$

$$O_t = O(X_t, T_{Ot}) \quad (2.4)$$

Production of I_t uses medical care (M_t) and O_t uses other market goods (X_t). Both I_t and O_t are affected by time spent on their production T_{Ht} and T_{Ot}.

Individuals maximise their utility within two constraints, time budget and monetary budget. The time budget is the total time available (365 days per year) and it can be distributed on four time consuming activities. The time spent on producing health (T_{Ht}) and other goods (T_{Ot}), the time spent working (T_{Wt}) and the time spent being sick (T_{St}).

$$T_t = T_{Ht} + T_{Ot} + T_{St} + T_{Wt} \quad (2.5)$$

The monetary budget depends on the income from work, the wage rate (W), the market goods and their prices (P_m, P_X)[4].

$$P_M * M + P_X * X = T_W * W \quad (2.6)$$

This equation represents the present value of each of its components. Hence, it is discounted with the discount rate (r).

[4] In England the NHS is free of charge hence in this case P_M=0. However, M_t does not only represent health services provided by the NHS but also other goods that contribute to increased health.

Maximising utility gained from health means that the individual sets the marginal benefit of health capital equal to its marginal costs. Marginal benefit of health capital consists of two parts. The first is what Grossman calls the investment benefit, which is based on the fact that healthy days can be used to work and this generates income. The marginal benefit of this monetary benefit is denoted MB_{Mt}. The investment benefit consists of the marginal product of health (the number of healthy days generated from one unit increase in health stock) denoted as MP_H and wage.

$$MB_{Mt} = MP_{Ht} * W_t \tag{2.7}$$

The second benefit is what Grossman has called the consumption benefit. This is the utility gained directly from health and is the value of the marginal product of health multiplied by the value of a healthy day. The marginal product of healthy days is measured as the ratio of the undiscounted marginal utility of healthy days (MU_H) to the marginal utility of wealth (MU_W).

$$MB_{Ht} = MP_{Ht} * (1+r)^t * (MU_{Ht} / MU_{Wt}) \tag{2.8}$$

As mentioned we maximise utility when marginal benefit is equal to marginal costs. The marginal benefit consists of the investment benefit and the consumption benefit ($MB_{Mt} + MB_{Ht}$). The marginal costs consists of the marginal cost of investment (*MCI*) multiplied with the opportunity cost of investment (the interest rate (*r*) minus the change over time in *MCI* (ΔMCI)) and the depreciation rate.

$$MB_{Mt} + MB_{Ht} = MCI_{t-1}[(r - \Delta MCI_{t-1}) + d_t] \quad (2.9)$$

This equation fully determines the optimal quantity of health at any time. Hence, people will invest in health up until the point where the marginal benefit from an investment (the consumption benefit and the investment benefit) is equal to the marginal costs incurred by the investment. If we divide both sides of equation 2.9 on *MCI* we get:

$$\frac{MB_{Mt} + MB_{Ht}}{MCI_{t-1}} = [(r - \Delta MCI_{t-1}) + d_t] \quad (2.10)$$

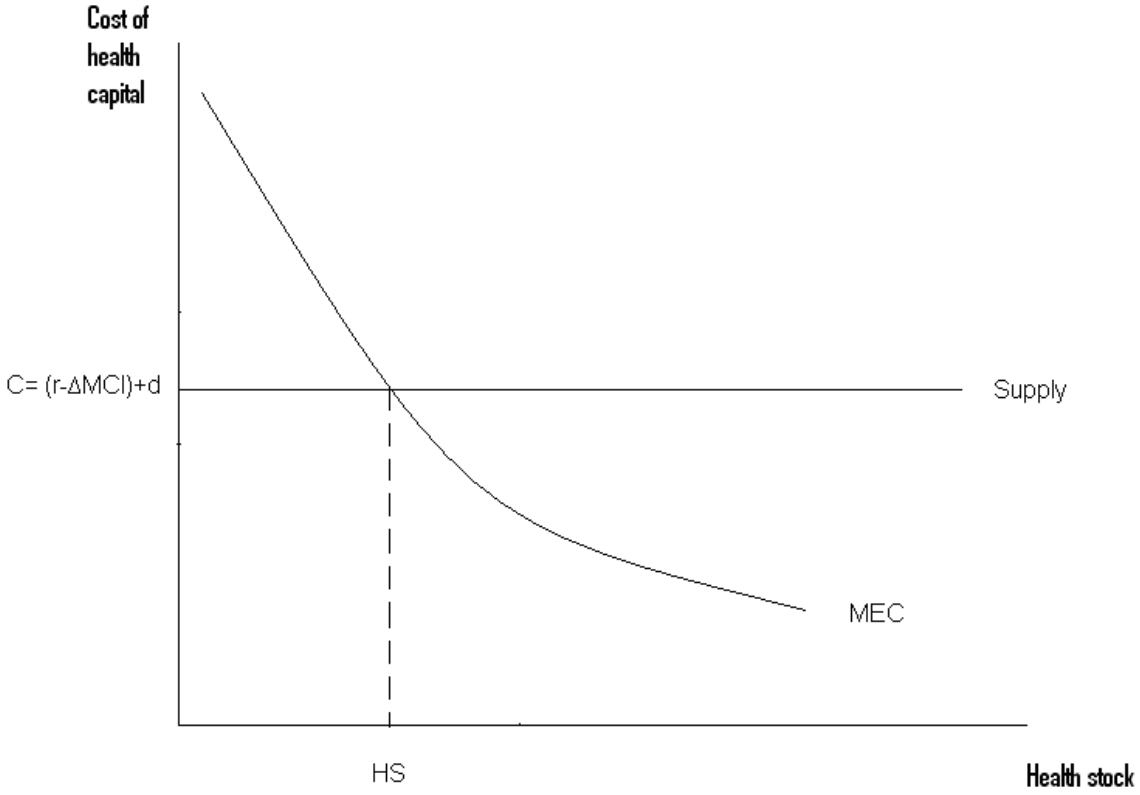
On the left hand side we have what we call the total rate of return on an investment in health and the term of the right hand side is from now on referred to as *the user cost of health capital*. In equilibrium the total rate of return on an investment in health must equal the user cost of health capital.

We will now demonstrate the equilibrium health in a figure which later will be used to make predictions. Due to the complicated relationship between the consumption benefit and the investment benefit we cannot estimate both in one model. To demonstrate some of the predictions of the model we focus on the investment model. The advantage of using this model compared with the consumption model is that it allows more easy interpretation in form of graphs. Although some predictions differ in the investment and the consumption model, it will not have consequences for the

predictions of how obesity and SES we influence health. For example, we will later illustrate how obesity reduces health in the investment model, but the same would have happened if we had used the consumption model, although a different mechanism would have worked.

So when we use the investment model the utility gained from health is in terms monetary units. This is illustrated in Figure 2.2 using the concept MEC (marginal efficiency of capital) which shows the relationship between the stock of health (HS) and the rate of return on an investment in health. It is assumed a diminishing marginal productivity of health as the health stock increases (Grossman, 2000), hence the convex shape of the MEC curve. The horizontal line (supply) shows the relationship between the stock of health and the cost of health capital ($C= (r-\Delta MCI)+d$). Since, the cost of capital is independent of the stock; the supply curve is a straight line. We can now see the equilibrium health stock (*HS*) which is where the MEC curve crosses the supply curve. We will later use this result to illustrate predictions about changes in equilibrium health based on changes in the determinants.

Figure 2.2: Equilibrium health



The Grossman model has been widely used as a basis for empirical testing and there is a range of evidence supporting the predictions made by the model. However, some of its assumptions have been heavily criticised. One of the main critiques it receives is that the model assumes that individuals have perfect information and make rational well informed choices about the time of death. In addition, it does not take into account uncertainty (such as the effect of health care on health). An extended model has taken account of this criticism by making the model probabilistic and assuming that individuals face a probability distribution of depreciation rates (Grossman, 2000). In our case this implies that lifestyle related *risk* factors increase the risk of a negative health outcome. However, as we discussed in the introduction, obesity works both by a direct effect and by an indirect (risk factor) effect. We discuss this further below.

We have now presented the Grossman model framework. In the following part we will go into detail on how obesity and SES influence health in this framework.

2.3.1 Lifestyles, SES and demand for health

We start by explaining how lifestyle related risk factors influence health capital. We then explain how the model can be modified to take into account SES variation in the impact of risk factors on health capital.

2.3.1.1 Lifestyle related risk factors in the Grossman model

When we want to include lifestyle related risk factors in the model we have two options. Either we can include lifestyle related risk factors as part of the investment in health or we can include them as part of the depreciation rate. Traditionally lifestyle related risk factors have been included as a part of the investment in health alongside medical care. However, it has been argued that this is not satisfactory as it extends the vector of medical care to a range of variables with widely different interpretations. Furthermore, it is assumed that investments in health is non-negative (Grossman, 2000), which makes it difficult to handle disinvestments in health. Muurinen (1982) suggests, however, that some behaviours such as smoking might cause use-related deterioration of health and that these should be modelled as a part of the health depreciation rate. Hence, in modified versions of the Grossman model the depreciation rate has also been a function of other factors like lifestyle related risk factors in addition to age (Muurinen, 1982; Reid, 1996; Grossman, 2000; Bolin et al., 2006). This means that the depreciation rate is defined as a function of both “the age

of the stock" and "the intensity of the use of the stock". To model this we denote the depreciation rate as a product of t (age), lifestyle (L_t) and all other variables (X_t), like environmental factors (such as pollution).

$$d_t = d(t, L_t, X_t) \tag{2.11}$$

Following Grossmans (2000) extension of the model we consider the depreciation rate effect to be probabilistic. In the introduction we explained that we have two effects of obesity, the direct and indirect effect (risk factor effect). The direct effect increases depreciation in the same way as the direct effect of age. While age reduces physical strength obesity may reduce mobility, for example. Hence, the probability that an individual experiences a negative health impact from the direct effect is close to or equal to 1. The indirect effects increase the probability of longstanding illnesses and health shocks. Hence, there is uncertainty surrounding whether or not individuals experience this effect and which of these effects the individual experiences. Hence, the risk factor effect has a range of probabilities <1. The depreciation rate is therefore the expected depreciation rate based on age, lifestyle and other variables.

2.3.1.2 SES in the Grossman model

In the original Grossman model (1972) there were two SES factors included: wage and education. These work to alter health in two ways. Higher wages increase the opportunity to purchase health care, i.e. high income individuals can purchase more expensive health care goods (this is more relevant in countries with a private health

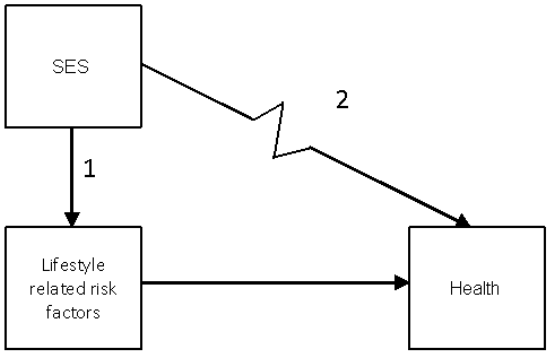
care system). Education increases the efficiency of an investment in health care, which raises the optimal stock of health. For example, individuals with a university degree will gain more in terms of health from a GP visit than individuals with no education. Based on both these factors the Grossman model predicts that high SES individuals will have a higher health stock on average across the population.

It has later been argued that this approach does not deal with SES in a satisfactory way and in extended versions of the Grossman model it has been argued that SES factors like education should be a part of the depreciation rate (Muurinen, 1982; van Doorslaer, 1987; and, McGuire, Henderson & Mooney, 1994). The results of doing this conform to the original Grossmans model, but the justification is different. In the latter SES is positively associated with health because it increases the general health productivity of an individual, and therefore reduces the marginal cost of investment in health. However, by including SES as a part of the depreciation rate it is seen as operating by redirecting the choice of different production processes in a way which leads to a lifestyle which is less use intensive in terms of health (Muurinen, 1982). For example, low SES individuals are more likely to make lifestyle choices that lead to a poorer health outcome. This is consistent with the CSDH framework, which also predicts that high SES individuals have a higher average health "level", as they have lower prevalence of negative health determinants.

We illustrate this in the Figure 2.3 below. SES works indirectly in two ways. SES has an impact on the prevalence of risk factors, illustrated by arrow 1. In addition, SES has an indirect impact on health through other *intermediate determinants* from the CSDH

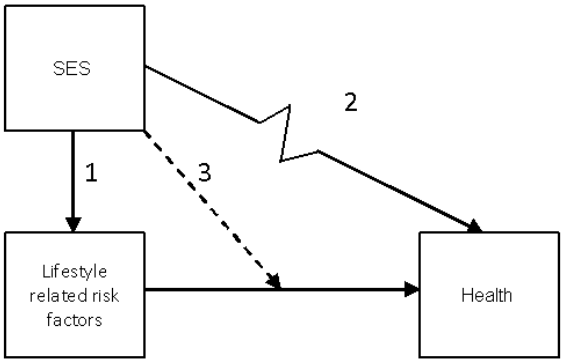
framework (see Figure 2.1), which is illustrated by arrow 2. These other intermediate determinants could be material circumstances (e.g., housing quality and consumption potential), psychosocial circumstances (e.g., stressful living and social support) (CSDH, 2008).

Figure 2.3: Impact of lifestyle and SES on health



We also introduce a third way that SES might influence health. This is by modifying the impact of lifestyle related risk factors on health. We illustrate this in Figure 2.4 below. The third way will work so that it modifies the negative impact of risk factors on health, which is showed by arrow 3.

Figure 2.4: SES modified the impact of lifestyle on health



If SES modifies the effect of lifestyle related risk factors we assume that SES influences the impact of lifestyle related risk factors on the health depreciation rate[5]. In equation 2.12 the depreciation rate has been altered to take into account this SES effect:

$$d_t = d(t, L_t, X_t; SES_t) \tag{2.12}$$

The semicolon (;) in front of SES_t illustrates that SES modifies the impact of the other variables (like lifestyle related risk factors) on the depreciation rate. We substitute 2.12 into equation 2.9:

$$MB_{Mt} + MB_{Ht} = MCI_{t-1}[(r - \Delta MCI_{t-1}) + d(t, L_t, X_t; SES_t)] \tag{2.13}$$

Equation 2.13 shows that the level of health capital depends on lifestyle related risk factors however; a homogenous risk factor will have unequal consequences for the depreciation rate depending on SES.

2.3.2 Predictions of the model

We will now use Grossman's investment model to illustrate how a homogenous lifestyle related risk factor can have heterogeneous consequences for equilibrium health. Let's assume that we have two equal individuals who are both normal weight. They do both have the same health stock *HS* and the cost of health capital *C*. This is

[5] We will later present and discuss a range of theories for why, and if so how, health SES might modify the health depreciation rate.

shown in Figure 2.5. Now assume that they both become obese, but that individual 2 has more health problems related to obesity. This could be a result of, for example, individual 1 has more knowledge and resources available to tackle either the direct or indirect effect (or both) of obesity. Then obesity will result in a higher depreciation rate in individual 2 compared to individual 1 (although the obesity status is the same). That would mean that the user cost of health capital is higher for individual 2 (*C**= (r-ΔMCI)+d***) than individual 1 (*C*= (r-ΔMCI)+d**). Hence, equilibrium health will decrease more for individual 2 (*HS***) than for individual 1 (*HS**) when they become obese.

Figure 2.5: Variation in equilibrium health due to shift in health depreciation

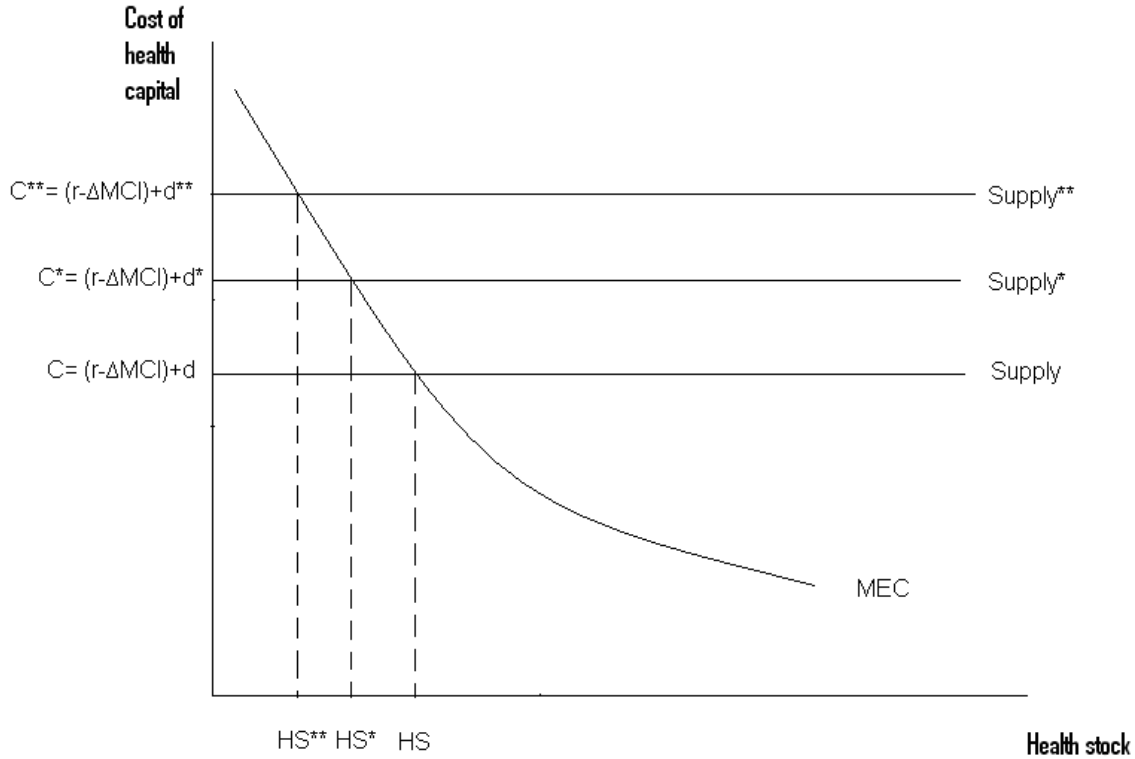
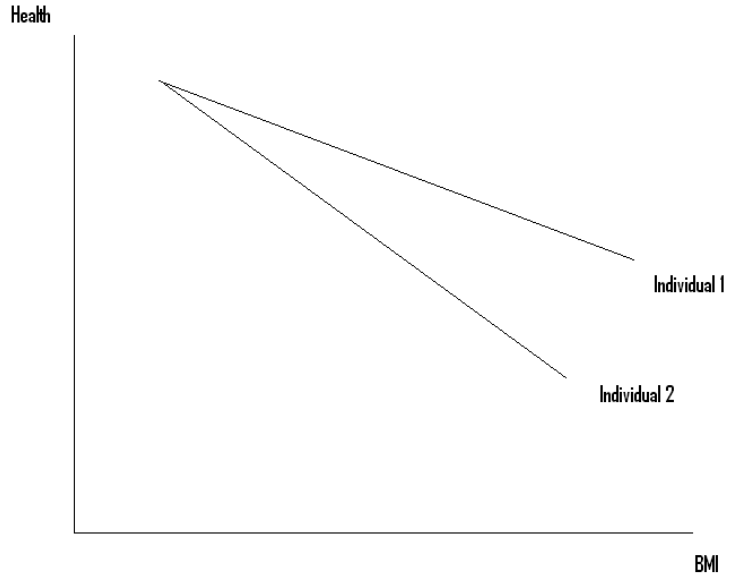


Figure 2.6 below shows individual 1 and 2 in a framwork where we have health, which is produced by health stock (*H=f(HS)*), on the vertical axis and BMI on the horizontal

axis. Both start out with the same health, however individual 2 has a higher depreciation rate when BMI increases. For the sake of the presentation we present a linear relationship between BMI and health. As illustrated in Figure 2.5 a higher depreciation rate when BMI increases will lead to more negative impact on equilibrium health stock, which then again reduces health. As a result we can see the disparities in health between individual 1 and 2 increase when BMI increase. So BMI excerts a more negative effect in individual 2 than in individual 1.

Figure 2.6: Relationship between equilibrium health and BMI



Using the Grossman model we have shown how a separate impact of lifestyle related risk factors on the health depreciation rate will generate disparities in health when individuals have a 'poor' lifestyle compared to having a 'good' lifestyle. We will later in this chapter discuss why there might be a heterogeneous impact of obesity on health.

2.3.3 Empirical testing of heterogeneities in the impact of lifestyle on health

We have presented a detailed outline of the structure for SES variation in the impact of obesity on health. This model has provided hypotheses for how our relevant determinants link up. Based on this we will now develop an econometric model suitable for empirical testing.

We illustrated that equilibrium level of health depends on the marginal investment and consumption benefits ($MB_{Mt} + MB_{Ht}$), the marginal costs of an investment in heath (MCI) and the user cost of health capital ($C= (r-\Delta MCI)+d$). Based on Grossman (2000) we include the depreciation rate as a separate element and we write the empirical formulation of the demand function for health capital (H_t) as:

$$H_t = \varepsilon_0 + \varepsilon_1(MB_{Mt} + MB_{Ht}) + (1-\varepsilon_2)(MCI) + (1-\varepsilon_3)d_t + u_t \quad (2.14)$$

The ε's are parameters to be estimated. The depreciation rate depends indirectly on SES (based on literature by, for instance, CSDH (2008) we know that SES is positively associated with health through material circumstances). In addition, we let the depreciation vary by lifestyle (L_t) and other variables (Z_t) like age:

$$d_t = a_0 + a_1 L_t + a_2 SES_t + a_3 Z_t + u_t \quad (2.15)$$

The a's are parameters to be estimated. We swap 2.15 into 2.14:

$$H_t = \varepsilon_0 + \varepsilon_1(MB_{Mt} + MB_{Ht}) + (1-\varepsilon_2)(MCI) + (1-\varepsilon_3)(a_0 + a_1 L_t + a_2 SES_t + a_3 Z_t + u_t) + u_t$$

(2.16)

We want to look at the impact of a change in lifestyle related risk factors on health. In our model lifestyle and SES influences health stock (H_t) through the health depreciation rate. We have included the health depreciation rate as a separate element in the function for health capital. Based on reasoning by, for example, Muurinen, we let lifestyle and SES work only through the depreciation rate. Hence, we keep both marginal benefit and marginal costs of an investment constant when risk factors change. So to investigate the impact of behavioural risk factors on health, we model the impact of the depreciation rate on health capital:

$$H_t = (\varepsilon_0 + (1-\varepsilon_3)a_0) + (1-\varepsilon_3)a_1 L_t + (1-\varepsilon_3)a_2 SES_t + (1-\varepsilon_3)a_3 Z_t + ((1-\varepsilon_3)u_t + u_t)$$

$$H_t = b_0 + b_1 L_t + b_2 SES_t + b_3 Z_t + e_t$$

(2.17)

To let the impact of risk factors vary by SES we include an interaction between risk factors and SES.

$$H_t = b_0 + b_1 L_t + b_2 SES_t + b_3 L * SES + b_4 Z_t + e_t \quad (2.18)$$

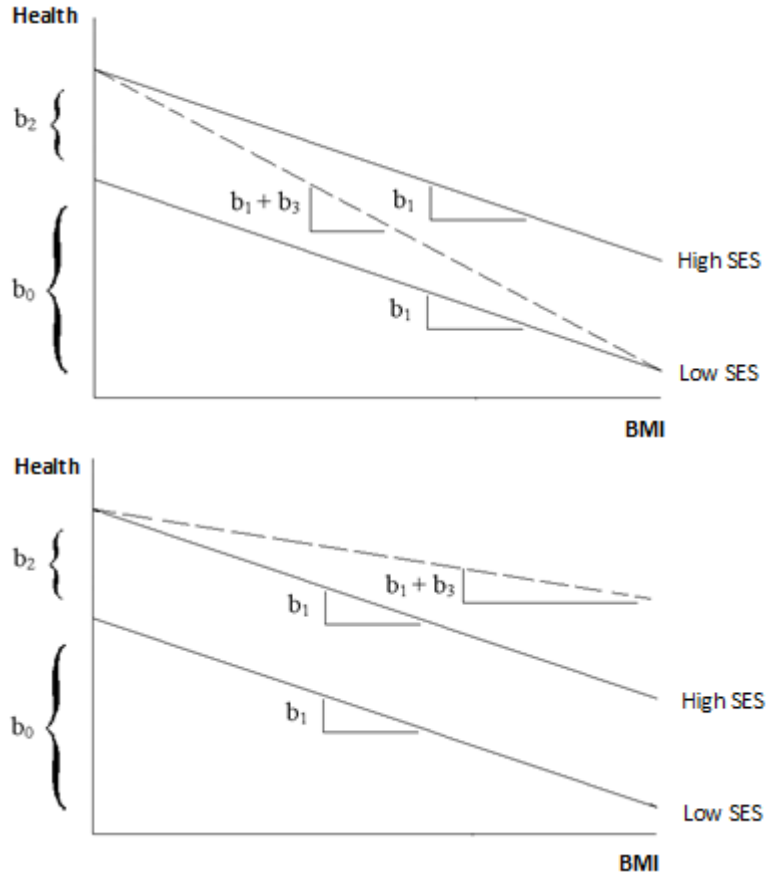
To test for a SES variation in the impact of lifestyle we can test the hypothesis H0: $b_3=0$. If $b_3 \neq 0$ it means that H0 is rejected and the impact of lifestyle on health is not equal across SES groups.

In Figure 2.7 we draw the relationship between BMI and health based on equation 2.18, with and without an interaction. For the sake of presentation we show a linear relationship[6] between BMI and health[7]. We now see the implications of rejecting H0. The solid lines present an equal relationship between BMI across SES groups (i.e. $b_3=0$). The dotted lines in the two graphs represents relationships where $b_3 \neq 0$. In the first graph $b_3 < 0$ and in the second graph $b_3 > 0$.

[6] It has been showed that this relationship is concave as both underweight and overweight/obesity has a negative impact on health.

[7] For the sake of presentation we have drawn the figures so that the high SES group always lies above the low SES group, however in the real relationship they might cross each other at some point.

Figure 2.7: SES variation in the impact of BMI on health



We have now discussed the mechanism behind a heterogeneous impact of obesity on health and how we can test for it. However, we have not discussed why there might be an interaction. To begin to answer this we have conducted a literature review to analyse theories and empirical evidence on heterogeneities in the relationship between lifestyle related risk factors and health. However, before we discuss these theories we present an earlier model developed by Birch et al. (1997) to illustrate SES variation in the impact of lifestyles on health. This is a useful supplement to our model because it takes a different perspective. While our framework looks at individual level variations in health, their framework looks at the consequences for the whole community in terms of a community's resource expenditures.

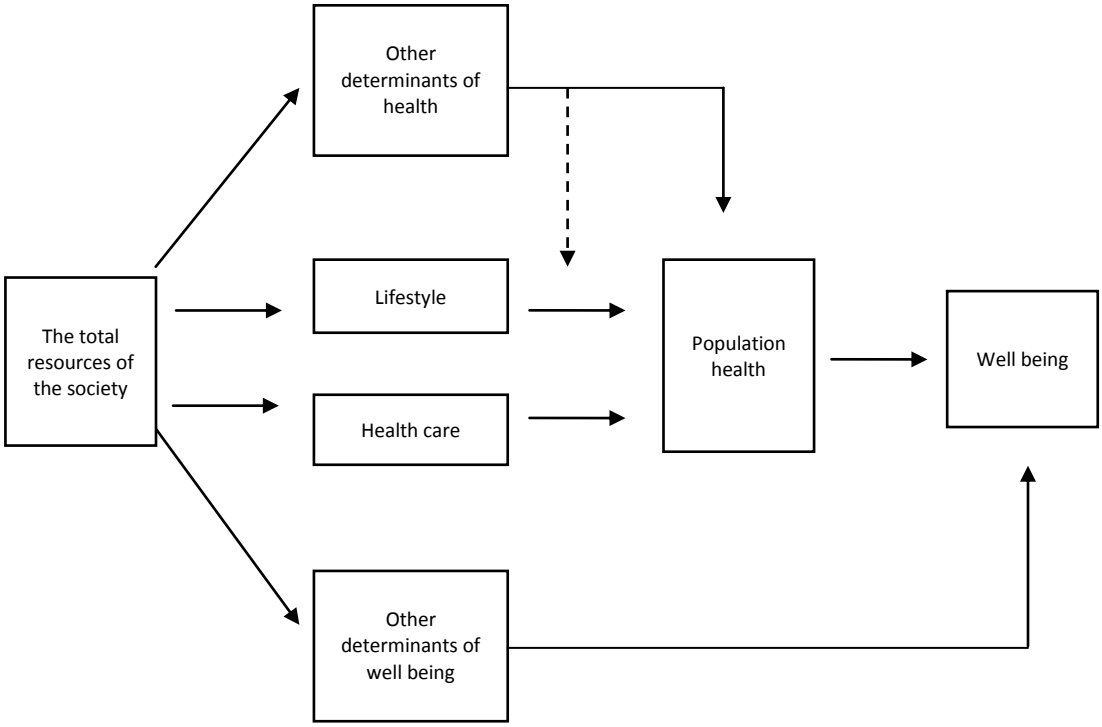
2.4 Overview of pathways to health in the population

In this chapter we presented an individual level model based on the Grossman model to explain heterogeneities in the impact of obesity on health. Birch, Jerret and Eyles (1997 & 2000) have presented a similar framework that takes account of socioeconomic variation in the association between lifestyle related risk factors and health. However, the main distinction between the previous and the "Birch framework" is that the "Birch framework" takes the perspective of the society, while the Grossman model is an individual level model. In this way the "Birch model" combines the CSDH framework presented earlier and our modified Grossman model. The aim is to illustrate that the demand for health improvements compete for the resources of the society with other utility generating activities. In addition, demand for health care competes with other health improving activities. It then illustrates how heterogeneities in the pathways to health can influence how the resources should be distributed to maximise the well being of the population.

In Figure 2.8 we see the framework presented by Birch, Jerret and Eyles (1997 & 2000). This framework illustrates that society can distribute it resources to maximise its well being. Well being consists of population health and other determinants of well being. Health is determined by lifestyle, health care and other determinants of health. Other determinants of health are considered to be environmental and socioeconomic factors. This framework shows that other determinants of health can influence health in two ways. Firstly, there is the direct effect in terms of the impact of the adverse effect of poor socioeconomic conditions on health. Second, socioeconomic factors can influence health indirectly as a moderator of other health determinants such as

lifestyle on health (represented by the dashed line). For example, the health risk of smoking or obesity might be related to socioeconomic status.

Figure 2.8: SES variation in the impact of lifestyle related risk factors on health



As we can see from this model the determinants of well being compete for limited resources and one must chose the interventions that maximise the well being of the society. Hence, if resources are devoted to counteracting lifestyle related risk factors, we need to know how these influence health. Birch, Jerret and Eyles (1997 & 2000) argue that reducing the prevalence of lifestyle related risk factors are likely to have systematically different health effects among social groups, and resource allocation based on average population values could lead to misallocation of resources. If one does not take into account the interaction between health determinants one may fail to reflect the reality of production of health in populations.

2.5 Hypothesis on pathways to health

We have discussed various frameworks for socioeconomic variation in the impact of lifestyle on health. However, we have not discussed why we might reject H0, and if so the direction of the relationship. We assumed that lifestyle related risk factors would affect health through the health depreciation rate. We also suggested including SES as a part of the depreciation rate. In Figure 2.7 we illustrated three hypotheses on the relationship which are; (a) a more pronounced impact in the high SES groups ($b_3 < 0$), (b) more pronounced impact in the lowest SES groups ($b_3 > 0$) and (c) SES does not modify the relationship ($b_3 = 0$) (similar effect).

There are some theories that support the first hypothesis (a), which is that high SES worsens and low SES limits the negative influence of lifestyle related risk factors on health. Blaxter (1990) argued that lifestyle factors were important determinants of health among those who were well off and not among those who were less well off in terms of occupation status. Blaxter argues that individuals in low SES groups have so many other factors around them that worsen health, so improving lifestyle will not make much difference. High SES individuals on the other hand, who have few factors around them that worsen health, will be able to achieve the full benefits from improving their lifestyle. Also some economists that use a rationale choice framework support this hypothesis either directly or indirectly. They assume that individuals with long lives have more to lose by increasing the probability of death at any point in time (Adda & Lechene, 2001; Carbone, Kverndokk & Røgeberg, 2003). As individuals with a

low SES have a shorter life expectancy on average they experience less harm from poor lifestyles and therefore are more likely to have a poor life style (like smoking and being obese). Individuals in high SES groups have a longer life expectancy; hence they have more incentives to invest in healthy behaviours. In this argument lies the assumption that high SES individuals have a higher impact on their absolute life expectancy than low SES individuals.

The second hypothesis (b) is that high SES limits and low SES worsens the negative influence of lifestyle related risk factors. Birch, Jerret and Eyles (2000) suggest that the effect is opposite of Blaxter (1990). The idea is that given the economic resources and favourable social circumstances, high SES individuals are more capable of tackling the harm of unhealthy lifestyles. While low SES groups are already weak and are more likely to suffer the comorbidities of unhealthy lifestyles. They also suggest that the underlying cause of the poor lifestyle (in their case smoking) might differ between SES groups and that the reason for the lifestyle might interact with the lifestyle to increase the health consequences. This can also be thought of as a "ceiling effect" where high SES groups have reached their potential health and improvements in lifestyles makes little difference. Low SES groups have not reached their full potential health and have more room for improvements (Pampel & Rogers, 2004). Higher SES groups might develop a better resistance to comorbidities since they have better medical care and more comfortable living conditions.

Finally, the third hypothesis (c), which is a similar relationship between BMI and health, means that there is no SES variation in the impact of lifestyle related risk

factors on health. This suggests that SES and risk factors contribute independently to health such that there is no interaction between them.

We have now presented three hypotheses on the association between lifestyles and health. In the following section we will present a literature review on this topic to investigate the empirical evidence.

2.5.1 Literature review on SES variation in the association between lifestyle related risk factors and health

The aim of this review is to identify studies that investigate SES variation in the relationship between lifestyle related risk factors and health.

2.5.1.1 Search strategy

We used three search strategies. First, we searched the databases MEDLINE and EMBASE using controlled vocabulary terms. In this review we choose to focus on the four main lifestyle related risk factors identified in Marmot (2010), which were obesity, smoking, alcohol and drug use[8]. Search terms and number of identified studies in each stage can be found in Appendix 2.1.

We included all papers that could be related to smoking, obesity, alcohol or drug consumption either directly or indirectly. We included only papers examining *variation* in the impact/association of the lifestyle related risk factors on health by variables

[8] These have been chosen as they were identified as the most important lifestyle related determinants of socioeconomic inequalities in health by Marmot (2010).

related to our definition of socioeconomic status (see Chapter 1). Hence, we do not include papers looking at the association between lifestyle related risk factors and health controlling for SES variables. In addition, we focus on SES variation hence we excluded papers looking at heterogeneities based on other groups, e.g., gender and ethnicity. We restricted the search to papers written in English. We did not have any restrictions on years, geographical area or population.

We screened titles and abstracts and identified a total of 7 papers in the first stage. In the second stage we searched the references of the papers identified in the first stage. In the third stage we used the citation search "referenced by" in Google Scholar to identify papers that had referenced our identified papers. We identified another 7 papers based on the reference searches. Hence, we found a total of 14 papers. A summary of the country, year, methods and results of identified papers can be found in Table 2.1.

Table 2.1: Summary of identified studies on SES variation in the association between lifestyle related risk factors and health

Author, country, year	Dataset	Method	Results
Marang-van de Mheen, Davey Smith, & Hart, Scotland, 1999	Renfrew and Paisley study (1972-1976) and the NHS central register, Men: n=7045 women: n=8348, age 20+	Used Cox models to investigate the impact of smoking on mortality interacting smoking and SES (measured by social class and area deprivation). They fitted three models depending on the covariates (the most advanced model controlled for age, cholesterol, diastolic BP, BMI, adjusted FEVI, deprivation, angina, ECG ischemia and bronchitis.	Did not find significant interactions, hence they concluded that the impact is similar across SES groups.
Thrane, Norway, 2006	National Screening program (1998-1999), n=19741, age 40-43	Used logistic regression techniques to investigate the impact of lifestyle (BMI, smoking, physical activity and diet) interacted with education on self-rated health controlling for gender, employment status, welfare support and loneliness	Find significant interactions between different lifestyles and education
Schafer, Ferraro, & Williams, USA, 2011	National Social Life, Health and Aging Project (2005-2006), n=3005, age 57-85	Fit models with interactions between BMI and low SES on C-reactive protein levels	Find significant interactions between BMI and low SES and conclude that BMI has a more negative effect in lower SES groups
Makela & Paljarvi, Finland, 2008	Finish Drinking Habit Survey (1969, 1976, 1984) , n=8199, age 15-69	Used Cox models to investigate the impact of alcohol and SES (based on occupation) and their interactions on mortality or hospitalisation controlling for gender, period, age, and marital status.	They did not find significant interactions between SES and alcohol consumption.
Pampel & Rogers, USA, 2004	National Health Interview Survey (1990), n= 41,104, age 18+	Use tobit, logit and Cox models to investigate the impact of smoking and SES (measured by education and occupation) and its interactions on morbidity, self-rated ill health and mortality. Controlling for age, gender, ethnicity, marital status and family income.	Find significant effect of the interactions between SES and smoking for men and women on morbidity and self-rated ill health but not on mortality.
Krueger & Chang, USA, 2008	National Health Interview Survey (1990) linked to mortality data in 1997, n=43335, age 18+	Used Gompertz models to investigate the impact of stress and lifestyle (smoking, alcohol, physical activity) on mortality. Fitted separate models for three SES (based on family income and education) groups. Controlled for age, gender, race/ethnicity, and marital status.	High baseline levels of former smoking and physical inactivity increased the impact of stress on mortality in the general population as well as among those of low socioeconomic status (SES), but not middle or high SES
Birch, Jerret, Wilson, Law, Elliot, & Eyles, Canada, 2005	Telephone survey of four neighbourhoods where 300 households were selected in each neighbourhood. (1996)	Used logistic regression to investigate the impact of smoking and SES (neighbourhoods) and their interactions on self-rated health. Controlling for age, gender and marital status.	They find a significant interaction between smoking and the deprived downtown neighbourhood. Smoking had less impact.
Birch, Jerret, & Eyles, Canada, 2000	Sante Quebec Survey (1992-1993), n=23,564, age 15+	Used logistic regressions to investigate the impact of smoking and SES (income, education and occupation) and their interactions on self-rated poor health. Controlling for age, gender and marital status.	The difference between smokers and non-smokers was significantly greater among groups of the population with low incomes and without employment, but significantly less among sub-groups with lower levels of education.
Davey Smith & Shipley, England, 1991	Whitehall Study (1967-1969), Men: n=16930, age 40-64. Note: only civil servants	Calculated mortality percentage for different smoking and occupation categories.	Found that mortality was higher for low SES smokers than high SES smoker and that the increase from non smokers to smokers was higher.
Singh-Manoux, Britton, Kivimaki,	Whitehall II (1985),	Fitted separate regressions by SES groups (measured by employment) and used MANCOVA	There was only a significant negative impact of IMT on

Gueguen, Halcox & Marmot, England, 2008	n=10308, age 35-55. Note: only civil servants	with interactions between SES and IMT controlling for age gender. The dependent variable was cognitive functioning.	cognitive functioning in the lowest SES group.
Laaksonen, Sarlio-Lahteenkorva, Leino-Arjas, Martikainen, & Lahelma, Finland, 2005	Finish cross sectional surveys (2000, 2001, 2002), men: n=1799 women: n=7148, age 40-60	Use linear regression methods to investigate the impact of BMI and SES (working conditions and occupation) and their interactions on the physical and mental scale of SF-36. Control for Smoking, alcohol use, marital status, and menopausal status (for women).	Socioeconomic position did not modify the association between BMI and health. However, there were some significant interactions between working conditions and BMI.
Kooiker, Christiansen, The Netherlands and Denmark, 1995	Two datasets. Dutch National Survey of General Practice (1987/88), n=11038, age. Danish Health Survey (1983), n=3149, age 25-64	Use MANOVA to test for significant interactions between lifestyle (BMI>27, smoking, drinking and no sports) and SES (social class, urbanisation, tenure status, health insurance, living arrangements, and daily activities) on five dependent variables (subjective health, number of health complaints, days with complaints, presence of one or more chronic diseases, GHQ score). Controlling of age and gender.	Neither the Dutch nor the Danish study finds interactions between lifestyles and SES on any of the measures of health.
Marmot, Rose, Shipley & Hamilton, England, 1978	Whitehall Study (1967-1969), Men: n=18403, age 40-64. Note: only civil servants	Used logistic models to investigate the impact of smoking on coronary heart disease mortality. Control for age. Fitted separate models by occupation groups.	Find a higher prevalence of smoking related diseases in the lower SES groups. These diseases appear to be not only related to more smokers, but also a higher impact of smoking in the lower SES groups.
Williamson, Canada, 2000	National Population Health Survey (1994-1995), n=11941, age 20-64	Used OLS to investigate the interaction between smoking status and five income groups on self rated health and number chronic conditions. Controlling for age, education, gender and marital status.	Did not find that the effect of smoking and physical activity depends on income

2.5.1.2 Results

We identified 14 studies, which were all from Europe or North America. Of the three UK studies identified (Marmot, Shipley & Hamilton, 1978; Davey Smith & Shipley, 1991; Singh-Manoux et al., 2008) none had investigated SES variation in the association between obesity and health.

Smoking is the risk factor that has received the most attention. Blaxter (1990) hypothesised that harmful habits like smoking has a greater impact on health in non-manual than in manual social classes. Their findings support this hypothesis when using a composite measure of physiological indices as the outcome measure. However,

this study has later been criticised for not providing confidence intervals or any formal statistical tests for differences between the groups (Marang-van de Mheen et al., 1999). In addition, it has been replicated using mortality as an outcome measure (Marang-van de Mheen et al., 1999) and the authors do not find significant differences in mortality and conclude that the association is equal across SES groups. On the contrary to Blaxter (1990), Birch, Jerret & Eyles (2000) argue that the association between smoking and health decrease in higher SES groups compared with lower. They find support for this by modelling interactions between SES variables and smoking on self-rated health. Pampel & Rogers (2004) find a significant effect of the interactions between SES and smoking for men and women on morbidity and self rated ill health but not on mortality. Their analysis supports the results by Birch, Jerret & Eyles (2000) by showing a stronger association in the low SES groups. Similar findings were found by Marmot, Shipley & Hamilton (1978) and Davey Smith & Shipley (1991) who find a more pronounced association between smoking and smoking related diseases in the lower SES groups. However, there are also some studies that show no interactions between smoking and SES and its association with subjective health measures (Kooiker & Christiansen, 1995; Williamson, 2000; Thrane, 2006).

Some studies have investigated SES variation in the association between health and BMI. Kooiker & Christiansen (1995) interacted a dummy for BMI>27 with a composite SES measure and tested if this interaction was significant using MANOVA tests. The dependent variables were a range of subjective and physiological health measures. They do not find significant interactions. This is similar to the findings of Thrane (2006) who interact BMI as a linear term with education. Laaksonen, Sarlio-Lahteenkorva,

Leino-Arjas, Martikainen, & Lahelma (2005) investigated the association between BMI (in three groups) interacted with working conditions and occupational status and SF-36 score. They do not find significant interactions for occupational class however; they do find significant interactions for working conditions.

Thrane (2006) also investigate interactions between SES and physical activity and diet and find some indications of an interaction with physical activity. This is contradicted by Williamson (2000) who does not find an interaction between SES and physical activity on self-rated health. Kooiker & Christiansen (1995) have looked at interactions between SES and drinking and exercise. They do no find significant interactions. Similarly, Makela & Paljarvi (2008) interacted alcohol use and occupation and looked at it association with mortality. They also do not find significant interactions.

We can see that the direction of this relationship might depend on the type of risk factor that is studied. For instance, the relationship might differ for smoking and obesity. In addition, it might be that this relationship differs depending on the type of health measure used. For example it looks like there is more support for an impact on health related quality of life and less support for an impact on mortality. In addition, the results might be sensitive to the country of origin.

The results are also sensitive to the underlying regression model. Most of these studies are based on nonlinear models like logit and survival models. These models look at multiplicative interactions and they do not quantify the implications of the

interactions. There is a general lack of clarity regarding what these studies actually investigates as they do not quantify the consequences.

To conclude, there is little evidence to support the hypothesis of a more negative impact of lifestyle related risk factors on health in the highest SES groups as originally suggested and tested by Blaxter (1990). However, numerous studies provide support of the hypothesis of a more negative impact lifestyle related risk factors on health in the lower SES groups. In addition, there is support of the hypothesis of an equal impact of lifestyle on health across SES groups. Although most of the evidence is on smoking, some studies investigate BMI/obesity and provide inconsistent results. In general, more research is needed to establish if and how socioeconomic variation in the impact of lifestyle related risk factors on health transpires.

2.6 Health service use

We have illustrated, using the Grossman model, how lifestyle related risk factors increase the health depreciation rate, which is negatively correlated with equilibrium health capital. We also included the possibility of SES variation in the impact of risk factors on the health depreciation rate. A further interest in our research is the consequences this has for health service use.

When health depreciation increases, equilibrium health will decrease, this means that a lower health capital is demanded. However, a lower demand for *health capital* does

not mean a lower demand for *health care*. As health depreciates individuals invest more in health care to augment their declining health stock.

When analysing demand for health care it is common to take into the consideration the concept of *need*. Need is a set of factors that is believed to determine use. We do take into consideration need in the following model of health service use, which is based on an economic model by Sutton et al. (2002). Need is an unobserved variable and the standard assumption is that need depends on current health status.

Health service use is however not a direct measure of need and there are other non-need variables affecting use (Sutton et al., 2002). Non-need variables are variables like ethnicity, gender and income. Although, Sutton et al. (2002) include SES as a part of need, we consider SES to be a non-need variable. For example, health service use may also depend on propensity to use across socioeconomic groups, but propensity to use should not increase actual need. Hence, it is only the current health status that decides how much one needs health care.

We would also expect supply variables like waiting times and number of GP's in an area to have an effect on use. Such supply variables are affected by past decisions of allocation of resources (which are affected by need) and failure to include these could lead to endogeneity issues (omitted variable bias).

We model utilisation of health services (U) as a function of need (N), non-need (NN) and supply factors (S).

$$U_t = f(N_t, NN_t, S_t) \tag{2.19}$$

Need is a function of health (H_t). We consider SES to be a non-need variable in addition to other non-need variables (X_t).

$$\begin{aligned} N_t &= f^2(H_t) \\ NN_t &= f^3(SES_t, X_t) \end{aligned} \tag{2.20}$$

Health is affected by lifestyle and SES. And as we see below, SES influences need indirectly by its indirect impact on health through intermediate determinants. So you do not need more health care because of low SES in itself. But you need more health care because of the reduced health you have as a result of reduced material circumstances, for example.

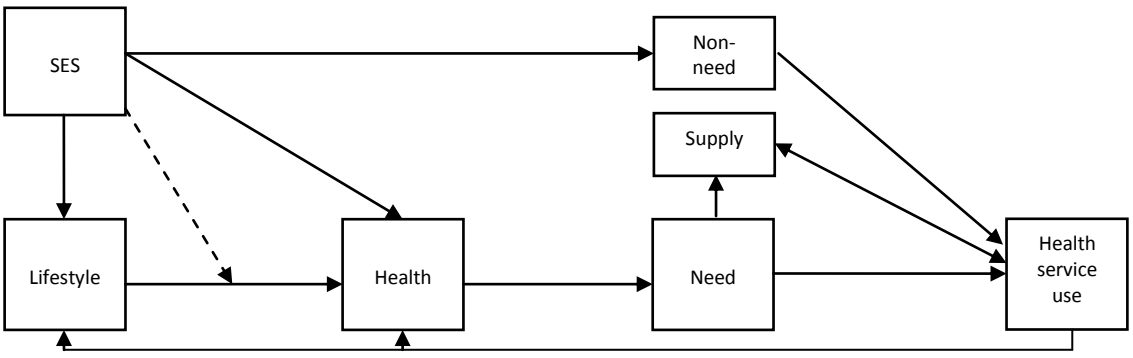
$$\begin{aligned} H_t &= f^4(L_t; SES_t) \\ N_t &= f^2(H_t) = f^5(L_t; SES_t) \end{aligned} \tag{2.21}$$

Hence, we believe lifestyle, SES, supply, and other non-need variables to affect health service use.

$$U_t = f^5(L_t, SES_t, S_t, X_t) \tag{2.22}$$

The model is illustrated in Figure 2.9. As we can see this model is an extended version of the model presented for the impact of lifestyle on health in Figure 2.4. So health service use is determined by need, non-need and supply factors. Need depends on health, which again is determined by SES, lifestyle and its interaction. So by controlling for supply and other non-need variables we can determine the effect of lifestyles on health and then again on health service use. We can also investigate SES variation in the impact of lifestyle on use.

Figure 2.9: Impact of lifestyle and SES on health service use



In figure 2.9 we also illustrate that health service use can affect lifestyle. For instance, the GP might help an individual to stop smoking or improve diet. Hence, there are potential endogeneity issues, if this is not accounted for in the analysis.

2.6.1 Empirical testing of heterogeneities in the impact of lifestyle on health service use

Based on the Grossman model we hypothesise that obesity increases the health depreciation rate, which will reduce the equilibrium health. We used a model to illustrate that health affects need for health services. We also showed that use might

not be determined only by need but also by other factors, non-need variables and supply. Based on this model we present the following econometric specification for health service use (m_t).

$$m_t = \varepsilon_0 + \varepsilon_1 H_t + \varepsilon_2 SES_t + \varepsilon_3 S_t + \varepsilon_4 X_t + u_t \quad (2.23)$$

Where the amount of health service used by an individual depends on the SES of the individual, health (H_t), the supply (S_t) of health care and other non-need variables (X_t). The ε's are parameters to be estimated. Earlier we modelled health at time t a function of obesity (L_t), SES and other covariates like age (Z_t).

$$H_t = b_0 + b_1 L_t + b_2 SES_t + b_3 L * SES + b_4 Z_t + e_t \quad (2.24)$$

The b's are parameters to be estimated. As we can see we have included an interaction between lifestyle and SES. Hence, if the impact of lifestyle on health differs by SES, we would also expect the impact of lifestyle on health service use to differ by SES. We swap equation 2.24 into equation 2.23 to illustrate how lifestyle will affect health service use:

$$\begin{aligned} m_t &= \varepsilon_0 + \varepsilon_1 (b_0 + b_1 L_t + b_2 SES_t + b_3 L * SES_t + b_4 X_t + e_t) + (1 - \varepsilon_2) SES_t + \varepsilon_3 S_t + u_t \\ m_t &= (\varepsilon_0 + \varepsilon_1 b_0) + \varepsilon_1 b_1 L_t + (\varepsilon_1 b_2 + (1 - \varepsilon_2)) SES_t + \varepsilon_1 b_3 L * SES_t + \varepsilon_1 b_4 X_t + \varepsilon_3 S_t + (u_t + \varepsilon_1 e_t) \\ m_t &= a_0 + a_1 L_t + a_2 SES_t + a_3 L * SES_t + a_4 X_t + a_5 S_t + e_t \end{aligned} \quad (2.25)$$

To test for a SES variation in the impact of lifestyle on health service use we can test the hypothesis H0: $a_3=0$. If $a_3 \neq 0$ it means that H0 is rejected and the impact of lifestyle on health service use is not equal across SES groups.

We have posed three hypotheses for how lifestyle might affect health status. Similar hypotheses can be illustrated for health service use. Hence hypothesis (a) illustrates a more negative impact of lifestyle related risk factors on health in the higher SES groups. Hence, high SES individuals will have a more pronounced impact of increased BMI on primary care use, which would mean that $a_3 > 0$.

Hypothesis (b) is when low SES individuals have a more negative impact of risk factors on health. Hence, this means that when BMI increases the slope is steeper for the low SES groups and $a_3 > 0$. Therefore, increased BMI leads to larger disparities in health service use between lower and higher SES groups.

The third hypothesis is when we do not reject H0 and $a_3 = 0$, which means a similar effect of lifestyles on depreciation across SES groups. Hence, in scenario (c) BMI has an equal impact across SES groups on health service use.

There might be variations in the sign of a_2, i.e. whether or not the high or the low SES groups use most of the service at baseline. There are indications of lower SES groups using more primary care services than higher SES individuals both with and without controlling for need (Morris, Sutton, & Gravelle, 2005). For other services there might be a different relationship at baseline and it might be that for some services higher SES

individuals have more use at baseline (normal weight). There are evidence to suggest that high SES individuals have a higher use of specialist/secondary care services when need is controlled for (Morris, Sutton, & Gravelle, 2005). However, the implications of the interactions are similar.

We do not have a literature review of SES variation in the impact of obesity on use in this chapter. However, we do have a review in Chapter 6 looking at obesity and health service use.

2.7 Summary

In this chapter we have presented a framework for analysis in the following chapters in this thesis.

We showed the CSDH framework for how the pathways and mechanism through which social determinants of health influences health. The idea is that socioeconomic status has an impact on intermediate determinants (e.g., material circumstances and lifestyle related risk factors). This again will have an impact on health.

We presented a human capital model for health (the Grossman model) and use this model to analyse how the impact of lifestyle related risk factors might vary by SES. In this model risk factors like obesity affect the health depreciation rate. An increased user cost of health capital will decrease the consumption of health capital and individuals who are obese will have a lower equilibrium health. We modified the model

to take into account SES variation in the impact of obesity and its consequences for equilibrium health. We then show how this framework has been modelled earlier by Birch et al. (1997) in a "macro" perspective and how this has consequences for distributing the resources in the society to maximise well being.

Three hypotheses on the relationship between risk factors and health are presented and discussed. These are: (a) high SES worsens and low SES limits the negative influence of lifestyle related risk factors on health; (b) high SES limits and low SES worsens the negative influence of lifestyle related risk factors on health; and, (c) the impact of lifestyle related risk factors on health is equal across SES groups (similar effect).

We presented a review of the literature investigating SES variation in the impact of lifestyle related risk factors on health. We find that there is limited evidence in support of hypotheses "(a) high SES worsens and low SES limits the negative influence of lifestyle related risk factors on health", which originally was found by Blaxter. However, more support of the hypotheses (b) and (c), meaning that the current debate is between these two hypotheses. Although most of the evidence is on smoking, some studies investigate BMI/obesity and provide inconsistent results.

We further analyse how SES variation in the impact of lifestyle related risk factors on health depreciation might have consequences for health service use. We use a framework where health influences need, which then again influences health service

use. We show that supply could influence each of these parts. Based on this framework we present three hypotheses for health service use.

In the following four empirical chapters we will test these three hypotheses. To find which of these has support using British data.

CHAPTER 3

Socioeconomic variation in the association between obesity and health related quality of life

3.1 Introduction

In Chapter 2 we built an economic model for the relationship between lifestyle and health and how this might vary by socioeconomic status. In this chapter we want to test this hypothesis. We presented an econometric specification for empirical testing:

$$H_t = b_0 + b_1 L_t + b_2 SES_t + b_3 L * SES_t + b_4 X_t + e_t \quad (3.1)$$

Where health status (H_t) is determined by lifestyle (L_t), SES and other variables (X_t), which are variables like age and gender. The lifestyle variable in our case is obesity. To test for a SES variation in the impact of obesity we can test the hypothesis H0: $b_3=0$. If $b_3 \neq 0$ it means that H0 is rejected and the association between obesity and health is not equal across SES groups.

In this chapter we measure health status with Health Related Quality of Life (HRQL). HRQL aims to capture the multidimensional concept of health and both clinicians and policymakers are recognising the importance of this to inform patient management

and policy decisions. For example, it is used in assessing Patient Reported Outcome Measures (PROMs). PROMs are measures of health gain and the collection is co-ordinated by the department of health (The Health and Social Care Information Centre, 2011). They are used for assessing the clinical quality of providers and supporting the reduction of health inequalities. In addition, the National Institute for Health and Clinical Excellence in England use cost-effectiveness analysis in economic evaluation of public health interventions and the preferred measure of benefit is quality-adjusted life years based on HRQL.

The aims of this chapter are to investigate the relationship between HRQL and obesity, and to investigate whether or not this relationship varies by SES. We start with a review of the literature of the relationship between obesity and HRQL. We then undertake the analysis using data from a large individual-level health survey, which includes interviewer-measured rather than self-reported height and weight and, a comprehensive set of individual and household characteristics that allows us to control for factors that affect the relationship between obesity and HRQL.

3.2 Literature review of the association between obesity and HRQL

The objective of the review was to identify studies that investigate the relationship between obesity and HRQL.

3.2.1 Search strategy

The literature was searched in three stages[9]: First, we searched the journals: *Obesity Reviews; Social Science & Medicine; Health Economics; and, Journal of Health Economics*. *Obesity Reviews* is the highest ranked obesity journal by impact factor (Thompson ISI, 2010). *Health Economics* and *Journal of Health Economics* were chosen as they are the two top ranked journals in the field of health economics in the *Thomson Reuters Journal Citation Reports*. *Social Science & Medicine* is the world's most cited social science journal (Thomson ISI, 2010) and it has a focus on health economics. The search term used was:

(obesity OR obese OR overweight OR weight loss) AND (quality of life OR health status OR health outcome OR rated health OR health index)

From this search we identified 5 papers. We then searched the references of the papers identified. In the third stage we used the "cited by" option in Google scholar to identify papers that had referenced the papers identified in the two first stages. We ended up with a total of 21 studies.

We included all papers that investigated the associations between a measure of obesity/overweight and HRQL (e.g., EQ-5D, self-rated health, SF-36 and HUI). We excluded papers that looked at physical status in other ways (e.g., blood pressure or diabetes status). We exclude all studies not in the English language.

[9] We initially used the search term ((("Obesity"[Mesh]) OR "Body Mass Index"[Mesh]) AND "Health Status Indicators"[Mesh]) OR "Quality of Life"[Mesh] in PubMed and got 93 754 results.

From each paper we extracted the year, country, methods and results, which can be found in Table 3.1.

Table 3.1: Literature review of studies describing the relationship between obesity and HRQL

Author, country, year	Dataset	Methodology	Key results
Wiczinski, Doring, John, & von Lengerke, Germany, 2009	General population survey in the Augsburg region, n=2732, age 35-74	Used linear regression techniques to investigate the impact of 4 BMI groups (excluding underweight) on the mental and physical scale of SF-12 controlling for age, education, family status, kind of health insurance, and place of residence	Obesity was associated with the physical but not the mental scale of SF-12
Brown, Dobson, & Mishra, Australia, 1998	Australian Longitudinal Survey for Women's Health, Women: n=13431, age 45-49	Calculated mean SF-36 score for 5 BMI groups adjusted for area of residence, education, smoking, exercise and menopausal status for women.	Increased BMI was associated with a significant reduction in the score of the physical component of the SF-36 but not the mental component
Han, Tijhuis, Lean, & Seidell, The Netherlands, 1998	Monitoring Risk Factors in Health in The Netherlands (1995), men: n=1885 women: n=2156, age 20-59	Used logistic regression to estimate the risk of poor health (a score below 66.7% of the standardised score on SF-36). Adjusted for age, smoking, alcohol consumption, physical activity, education, marital status, employment, household composition, intimate contact, and parity.	Higher BMI increased the probability of having poor health
Lean, Han, & Seidell, The Netherlands, 1998	Monitoring Risk Factors in Health in The Netherlands (1993-1995), men: n=5887 women: n=7018, age 20-59	Used logistic regression analysis to investigate the odds for poor quality of life (a score below 66.7% of the standardised score on SF-36), by waist circumference. Adjusted for age, smoking, education, alcohol consumption, physical activity, employment status, household composition, discussion of personal matters with other people, and parity	Higher waist circumference was related to higher probability of having poor health
Le Pen, Levy, Loos, Banzet, & Basdevant, France, 1998	SOFRES, n=1000 (500 obese; 500 non-obese), age 18+	Compares the association between obesity (BMI>30) and SF-36 and OSQOL. Using a cross-sectional survey with a control group matched based on sex, age, and employment status	The quality of life of patients with severe obesity is impaired, but it mainly affects the physical consequences of the disease. Comparison of the specific OSQOL with the generic SF-36 shows clearly that the two kinds of scales correspond to different objectives and could be considered as complementary.
Fine et al., USA, 1999	Nurses Health Study (1992-1996), women: n=40098, age 46-71	Used OLS to investigate the impact of weight change (in three groups: losers, maintainers, gainers) on SF-36 score controlling for age, cigarette smoking, levels of physical activity, alcohol consumption, and self reported comorbid	Weight gain was associated with decreased SF-36 scores while the association with weight loss was opposite.

		conditions	
Lean, Han, & Seidell, The Netherlands, 1999	Monitoring Risk Factors in Health in The Netherlands (1993-1995), men: n=5887 women: n=7018, age 20-59	Used logistic regression methods to measure the impact of BMI in three categories on the probability of low SF-36 score (less than 66.7% of the standardised score). Adjusted for age, smoking, education, alcohol consumption, physical activity, employment status, household composition, discussion of personal matters with other people, and parity	Individuals with higher BMI had a higher probability of having a low SF-36 score
Doll, Petersen, & Stewart-Brown, UK, 2000	Postal survey within hte old Oxford Regional Health Authority (1997), n=8889, age 18-64	ANOVA models looking at the impact of BMI (measured as 5 BMI groups) on SF-36 score, controlling for age, gender and frequency of health service utilisation	Overweight and obesity are associated with poor levels of subjective health status, particularly in terms of physical well-being
Ford, Moriarty, Zack, Mokdad, & Chapman, USA, 2001	Behavioural Risk Factor Surveillance System (1996), n=109076, age 18+	Used logistic regression to investigate the impact of 6 BMI groups on binary quality of life measures, adjusting for age, gender, race or ethnicity, educational attainment, employment status, smoking status, and physical activity status	Participants with a self-reported BMI>18.5 and participants with a self-reported BMI>30 reported impaired quality of life
Larsson, Karlsson, & Sullivan, Sweden, 2002	Swedish cross sectional survey, n=5633, age 16-64	Used regressions to look at the impact of BMI on SF-36 scores controlling for age, sex, physical activity and disability. Used ANOVA test to control for the impact of three BMI groups on SF-36 scores.	Find a relationship between BMI and physical but not mental health
Heo, Allison, Faith, Zhu, & Fontaine, USA, 2003	Behavioural Risk Factor Surveillance System (1999), n=155989, age 18+	Created a binary self-rated health variable and use logistic regression methods to look at the impact of BMI (as continuous and as six categories) controlling for age, sex, marital status, educational attainment, annual income, smoking status, and employment status. They also control for comorbidities	Associations between BMI and HRQL indices were J-shaped. Joint pain and comorbidities may mediate BMI-HRQL associations
Groessl, Kaplan, Barrett-Connor, & Ganiats, USA, 2004	Longitudinal cohort study (1992-1995), n=1326, mean age 72	Used ANOVA and ANCOVA methods to investigate if there was a significant association between BMI (in four groups) on quality of well being scale in individuals with a mean age of 72. Controlling for age, gender, smoking history, and exercise	Obese older adults tend to have lower HRQOL than those who are overweight or of normal BMI. Overweight did not have a significant impact.
Yan et al., USA, 2004	Cross sectional survey in Chicago (1996), men: n=3981 women: n=3099, age 65+	Used generalised linear models and logistic regressions to investigate the impact of four BMI groups on HSQ-12 domains controlling for age, race, education, smoking, and alcohol intake. Separate analysis by gender in a population aged 65 and older	Obesity was associated with lower health perception and poorer physical and social functioning (women only) but not impaired mental health
Macran, UK, 2004	Health Survey for England (1996), n=11783, 18+	Used OLS and MANOVA to investigate the impact of BMI (three and five groups) on EQ-5D controlling for age gender and comorbidities.	Most of the apparent relationship between BMI and HRQL could be accounted for by age and the presence of long-standing illness. However women's HRQL did appear to be sensitive to their weight.
Jia & Lubetkin, USA, 2005	Medical Expenditure Panel Survey (2000), n=15438, age 18+	Used linear and spline regression methods to investigate the impact of obesity (five groups) on SF-12 and EQ-5D scores controlling for smoking, physical activity, clinical conditions, age, race/ethnicity, sex, and income	Persons with obesity had significantly lower HRQL than those who were normal weight and such lower scores were seen even for persons without

			chronic diseases known to be linked to obesity.
Kortt, & Clarke, Australia, 2005	Australian National Health Survey (1995), n=19301, age 18+	Separate regression analyses for males and females to model the association between utility (measured by SF-36), BMI (three groups, excluding underweight), controlling for obesity-related medical conditions, age, smoking status, age left school, and income	Results suggest that BMI is negatively associated with utility
Laaksonen, Sarlio-Lahteenkorva, Leino-Arjas, Martikainen & Lahelma, Finnland, 2005	Finish cross sectional surveys (2000, 2001, 2002), men: n=1799 women: n=7148, age 40-60	ANOVA was used to assess differences in SF-36 of three BMI groups controlling for age. The effect of SES and working conditions on the associations between body weight and SF-36 was examined with multiple regression analysis. Included interactions between BMI as a continuous variable and SES groups as categorical variables.	Body weight was associated with physical health only. Socioeconomic position did not modify the association between BMI and health. However, they found an effect of working conditions.
Sach et al., UK, 2007	Patients at one UK general practice (2004), n=1865, age 45+	Used linear regression methods to investigate the impact of six BMI groups on EQ-5D, EQ-VAS and SF-6D controlling for age, sex, smoking status, and presence of the 10 co-morbidities. In addition, they used logistic regression methods to investigate the impact of probabilities of reporting at being at level one in each of the domains in the respective HRQL measure.	The EQ-5D, EQ-VAS and SF-6D were in agreement that, relative to a normal BMI, obesity is associated with a lower HRQL, even after controlling for patient characteristics and comorbidity.
Wee et al., Singapore, 2008	Recruited Primary Care patients (), n=411, age 21+	Used linear and nonlinear regression methods to investigate the impact of four BMI groups on EQ-5D/EQ-VAS, Health Utilities Index (HUI2 & HUI3) and the SF-6D controlling for marital status, smoking status, work status, presence of chronic medical conditions and Family Functioning Measures scores	Consistent with available studies, obese subjects reported worse HRQL than normal-weight subjects.
Mond & Baune, Germany, 2009	German Health Interview and Examination Survey (1997-1999), n=4181, 18-65	Used linear regression to investigate the impact of BMI in three groups on SF-36 scores. Controlled for age, social status and a range of medical conditions.	Found relationship between physical functioning for mild obesity in women and moderate obesity in men. Found no association between weight status and psycho-social functioning, in women, whereas overweight was associated with better perceived psycho-social functioning in men.
Lee et al., Wales, 2005	Used hospital treated patients in Wales (2002-2003), n=30020, age 18+	Used linear regression techniques investigating the association between BMI as a linear variable on EQ-5D score controlling for age across individuals with and without diabetes.	Found that BMI is negatively correlated with EQ-5D score in patients with and without diabetes.

3.2.2 Findings

There is evidence that obesity is inversely associated with HRQL. Of the 21 identified studies that met the inclusion criteria and investigate the associations between obesity and HRQL, four were from the UK (Lee et al., 2005; Doll, Petersen, & Stewart-Brown, 2000; Macran, 2004; Sach et al., 2007), and they all provide evidence of a negative association between obesity and HRQL. Some studies have investigated the association between obesity and HRQL separately on a physical and a mental scale (Wiczinski et al., 2009; Brown, Dobson, & Mishra, 1998; Le Pen, Levy, Loos, Banzet, & Basdevant, 1998; Doll, Petersen, & Stewart-Brown, 2000; Larsson, Karlsson, & Sullivan, 2002; Yan et al., 2004; Laaksonen, Sarlio-Lahteenkorva, Leino-Arjas, Martikainen & Lahelma, 2005; Mond & Baune, 2009). These studies demonstrate an association between obesity and HRQL on the physical but not the mental QOL scale. Five of the studies have used EQ-5D (Lee et al., 2005; Macran, 2004; Jia & Lubetkin, 2005; Sach et al., 2007; Wee et al., 2008) and they all identify an inverse association between obesity and EQ-5D score.

The most common method used is the ANOVA method to test for significant variation, and then present summary statistics of the HRQL variable for different BMI groups. A gender difference is commonly observed as the studies that fit separate models for men and women tend to find a more pronounced effect for women.

In these studies there might be issues with omitted variables, which are other explanatory variables that have not been included in the model that are correlated with the dependent variable (HRQL) and the independent variables (BMI groups). One

such variable is longstanding illnesses. An illness (such as different cancers) might alter BMI status and HRQL. By controlling for illness variables one could avoid this potential bias. However, controlling for illness variables is also criticised. It has been argued that controlling for comorbidities is an overadjustment as these factors are part of the pathway between obesity and the output variable (See Zizza, 2004). Hence, it can be seen as controlling for the risk factor effect of obesity and by doing this one looks only at the direct health effects.

A number of studies have identified a negative correlation between obesity and HRQL when they control for comorbidities (Fine et al., 1999; Heo, Allison, Faith, Zhu, & Fontaine, 2003; Jia & Lubetkin, 2005; Kortt, & Clarke, 2005; Sach et al., 2007; Wee et al., 2008). However, studies that fit regressions with and without illness variables find that controlling for illnesses reduce the negative association between obesity and HRQL. Macran (2004) and Mond & Baune (2009) found that most of the association between obesity and HRQL can be explained by the presence of longstanding illnesses. Especially in men, where they do not identify a relationship when longstanding illnesses are controlled for.

We identified one study that investigates if the effect of obesity varies by SES measures. Laaksonen, Sarlio-Lahteenkorva, Leino-Arjas, Martikainen & Lahelma (2005) investigate if the association between BMI and HRQL varies by occupational status and working conditions (consisting of perceived job control and demand) using Finnish data. They show that the association between BMI and HRQL, measured using the physical function component of the SF-36, is modified by working conditions. This was

not found for the mental health scale of the SF-36. Although they found that working conditions modified the relationship between BMI and HRQL they did not find any evidence for occupational class (managers and professionals, semiprofessionals, routine non-manual workers, and manual workers). They used linear regression methods and interacted BMI (as a continuous variable) with occupational class and working conditions. Laaksonen et al. might have found a different result if they had allowed a nonlinear impact of BMI on health. Furthermore, this study looks only at public sector workers between the ages of 40 – 60 yrs. Hence, this sample does not cover some socioeconomic groups, such as the unemployed and entrepreneurs. In addition, the age range is restricted which further reduce the variation in the population. Furthermore, morbidly obese individuals have higher unemployment rates and will not be included in this sample. This study was not undertaken in the UK, and results might vary between countries, socioeconomic status might have different consequences across countries. In addition, they use self-rated height and weight, which has been criticised as individuals have a tendency to underreport weight and over report height (Gorber, S. C., Tremblay, Moher, & Gorber, B., 2007). The following study overcomes these limitations by applying BMI categories allowing for the nonlinear impact of BMI on HRQL, using a representative population of the English population, and using nurse measured height and weight.

Of the English studies only the study by Macran (2004) use a representative sample for the English population aged 18 and over. However, this study use a dataset from 1996, while there are similar data available for the years 2003-2004.

To conclude, there is evidence of a negative relationship between obesity on HRQL. This relationship persists after controlling for comorbidities. However, limited evidence exists on whether or not this effect varies across SES groups.

3.3 Methods

3.3.1 Data and variables

The analysis is based on data from four rounds (2003-2006) of the *Health Survey for England* (HSE) (National Centre for Social Research & Department of Epidemiology and Public Health (UCL)). The HSE is a repeated cross-sectional nationally representative survey which draws a different sample every year of individuals living in England. The sample is selected every year using multi-stage stratified probability sampling design with postcode sectors selected at the first stage and household addresses selected at the second stage. Stratification is based on geographical areas and not on individual characteristics. Respondents are interviewed on a range of topics including their health (including obesity and EQ-5D score), and their socioeconomic status. Only participants aged 16 or above completed the EQ-5D questionnaire and are included in this study.

Our measure of HRQL is the EQ-5D (The Euroqol Group, 1990), which is included in the HSE in 1996 and 2003–2006. This measure is commonly used in cost-effectiveness analyses to calculate quality-adjusted life years. The HSE includes other measures of HRQL like self-rated health; however EQ-5D is used in this chapter as it is the preferred measure of HRQL in adults in cost-effectiveness analyses considered by the National

Institute for Health and Clinical Excellence in England (NICE, 2008). The EQ-5D is a generic measure of HRQL. It is applicable to a wide range of health conditions and treatments, and provides a descriptive profile that is reducible to a single index value for health status. The EQ-5D descriptive system consists of five dimensions – mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each dimension is scored at one of three levels, depending on whether the respondent has no problems (score=1), some problems (=2) or serious problems (=3) with each dimension. This descriptive system defines 243 EQ-5D health states, plus 'unconscious' and 'dead', giving 245 states in total. Each EQ-5D health state can be converted to a single summary index by applying a formula that attaches weights to each of the levels in each dimension. A number of such formula, or value sets, are available for different countries, based on the valuation of EQ-5D health states from general population samples. The HSE uses weights obtained from the UK population by the Measurement and Valuation of Health Group (The MVH Group, 1995; Dolan, 1997). After applying these weights, an EQ-5D index score of one represents full health, a score of zero is equivalent to death, and negative scores represent health states worse than death.

Our obesity measure is based on body mass index (BMI), measured as weight in kilograms divided by height in metres squared (kg/m^2). BMI is computed from the height and weight measures obtained during the interviewer visit in the HSE; it is not based on self reported height and weight, which means that the likelihood of systematic measurement error is reduced. Obesity is measured as a categorical variable based on four BMI categories, defined according to World Health Organization guidelines (WHO Expert Committee on Physical Status, 1995): normal weight, BMI

18.5-24.9 kg/m^2; overweight, BMI 25-29.9 kg/m^2; class I obesity, BMI 30-34.9 kg/m^2; and, class II / III obesity BMI >35 kg/m^2. In our data there are 171 men and 491 women in obesity class III (BMI >40 kg/m^2); we combine obesity classes II and III to ensure an adequate number of observations in each cell when we interact SES and obesity category. Our data also has small number of respondents in the underweight group (BMI <18.5 kg/m^2; 155 men, 305 women). We do not combine underweight and normal weight because there is some evidence that the underweight have lower HRQL than the normal weight (Jia & Lubetkin, 2005). We therefore include underweight as a separate category (interacted with SES where appropriate). Based on the discussion in chapter 1, we do not report the results for this group in the following analysis.

Our SES variable is a composite measure, designed to capture multiple dimensions of SES (individual, household and area socioeconomic variables) rather than focusing on a single aspect (such as income). We use a regression-based approach to construct the SES measure and regress (using interval regression) total annual household income reported in 31 income bands (including an open-ended top category) against a set of socioeconomic variables. We calculate predicted values from this model, which are then equivalised using McClements household score provided in the HSE to account for household size and composition. The socioeconomic variables used as covariates in the interval regression are: education qualifications (measured in seven categories); social class of household reference person (HRP; six categories); cars owned by household (four categories); economic activity status for last week (11 categories); housing tenure (five categories); bedrooms in household (five categories); index of multiple deprivation quintile (SOA level) (five categories); and, whether or not the

person was an income support claimant (yes/no). The predicted SES measure is explained partly but not exclusively by the total annual household income variable. When we regress the predicted SES measure against the 30 income band indicators (excluding the lowest category) the model R^2 is 0.2647. The predicted SES measure is then divided into quartiles for use in the EQ-5D models.

In the EQ-5D models we also include covariates for: age (cubic function); ethnicity (nine categories); marital status (six categories); smoking status (four categories); Government Office Region of residence (nine categories); predicted equivalised total annual household income (continuous variable); and, survey year (four categories). We also run a model controlling for number (five categories) and type (fourteen categories) of longstanding illness; limiting longstanding illness (yes/no); and, acute ill health days the last two weeks (five categories).

3.3.2 Statistical analysis

We investigate differences in EQ-5D scores and individual and household characteristics by BMI category using one-way ANOVA models for continuous variables and χ^2 tests for categorical variables. We then regress EQ-5D score (h_i) against the BMI categories[10] using OLS, with normal weight as the omitted category:

$$h_i = \beta_0 + \beta_1 \sum_{a=1}^{5} BMI_{ai} + \beta_2 Z_i + u_i \qquad (3.2)$$

[10] As described above we include underweight and interact it with SES as appropriately. However, we do not report the results for this group and denote only the BMI groups of interest in the equations

Where Z_i is other variables like age and smoking status and *i* indexes the individuals. The β's are vectors of parameters to be estimated.

We then rerun this model including interaction terms between the BMI categories and the SES quartiles, with normal weight/SES quartile 1 (most deprived) as the omitted category.

$$h_i = \beta_0 + \beta_1 \sum_{b=1}^{4} SES_{bi} + \beta_2 \sum_{a=1}^{5} BMI_{ai} + \beta_3 \sum_{a=1}^{4} \sum_{b=1}^{5} SES_{bi} * BMI_{ai} + \beta_4 Z_i + u_i \qquad (3.3)$$

Three versions of each model were estimated: unadjusted models including the BMI categories or BMI/SES categories only; multivariate models that controlled for all the individual and household characteristics except for the longstanding and acute illness variables; and, multivariate models that controlled for all the individual and household characteristics including the longstanding and acute illness variables. We run separate models for men and women. After every regression we compute the predicted mean EQ-5D score for each BMI or BMI/SES category. In the multivariate models we do this by fixing the covariates at their whole-sample mean values. In this case, the variation in predicted mean EQ-5D scores is a function of the impact of obesity on EQ-5D scores and how this varies by SES, and is not affected by the values of the covariates and comorbidities. We also compute predicted mean EQ-5D scores for each BMI/SES category by setting the individual and household characteristics to the mean values of each SES category, and then to the mean values of each BMI category. In these cases the values of the covariates used in the predictions are more representative of the

individuals in each SES or obesity category, and variation in predicted mean EQ-5D scores reflects both differences in the impact of obesity on EQ-5D scores and how this varies by SES, but also differences in the values of the individual and household characteristics and between categories.

Standard errors for the predicted mean values were computed using the standard error calculation based on Korn & Graubard (1999), which accounts for the variation in the covariates that would arise from repeated sampling, and is appropriate to use with a representative sample of the population when making inferences about the underlying population.

We test for equality in predicted mean EQ-5D scores between BMI categories (across all SES groups combined and within SES groups) and between SES quartiles within BMI categories using Wald tests and F- tests.

We test for statistically significant interactions between obesity and SES (i.e., whether or not the association between obesity and EQ-5D score depends on SES, and vice versa) for each gender in the adjusted and unadjusted models using two-way ANOVA.

We apply survey individual-level selection weights reported in the HSE to each observation. This adjusts for the fact that different observations have different probabilities of selection and participation in the survey. It is also possible that, due to the sampling strategy used in the HSE, observations are independent across Primary Sampling Units (PSUs), but not within PSUs. If this is the case then if we use estimators

that assume independence within these clusters the standard errors on our regression coefficients will be too small and we will overestimate the statistical significance of the independent variables in our models. We therefore control for clustered sampling within PSUs using unique PSU/year identifiers that produce Huber/White/sandwich robust variance estimators that allow for within-group dependence (Kish & Frankel, 1974).

To maximise the sample size, we included individuals in the EQ-5D models with missing income data for whom we could compute a predicted SES value. In the EQ-5D models we included an indicator for missing income (yes/no), to control for the possibility that income may not be missing at random. To investigate the effect of imputing missing income values we repeated our analysis in individuals with non-missing income data only.

P values below the 5% level (t statistics higher than ≈1.9) are regarded as statistically significant. Values between 5% and 10% (t statistics between ≈1.6 and ≈1.9) are regarded as weakly significant. Analyses were undertaken using Stata SE version 11.

3.4 Results

The total number of respondents in the HSE in 2003-2006 was 61,603. Forty five thousand nine hundred and eighty five were aged 16 or above and 42,002 had EQ-5D data. Thirty three thousand seven hundred and sixteen observations were included in the income regression and predicted SES values were computed for 42,825

observations. A total of 33,105 observations (15,142 men, 17,963 women) had EQ-5D scores and predicted income data and were used in the EQ-5D models. The numbers of observations in each BMI/SES category are in the Table 3.2.

Table 3.2: Numbers of observations by BMI category, gender and SES quartile

	Normal weight	Overweight	Class I obesity	Class II/III obesity
Men				
All	4,535	6,795	2,854	803
SES quartile				
1 (most deprived)	1,080	1,417	714	217
2	1,127	1,582	701	207
3	1,107	1,786	721	211
4 (least deprived)	1,221	2,010	718	168
Women				
All	7,261	6,019	2,802	1576
SES quartile				
1 (most deprived)	1,545	1,645	923	607
2	1,777	1,559	764	400
3	1,908	1,492	626	330
4 (least deprived)	2,031	1,323	489	239

Source: HSE 2003-2006

Selected descriptive statistics across BMI categories are in Table 3.3, while full descriptive statistics are in Appendix 3.1. Thirty six percent of the sample were normal weight, 39% (the modal category) were overweight and 24% were obese (17% class I obesity, 7% class II/III obesity). The mean EQ-5D score across the whole sample was 0.871, falling from a mean value of 0.900 in the normal weight category to 0.782 in the class II/III obesity category ($P<0.01$). The mean predicted SES measure was highest in the normal weight category and lowest in the class II/III obesity category; the modal SES quartile was quartile 4 (least deprived) in the normal weight category and quartile 1 (most deprived) in the class I and II/III obesity categories. Forty six percent of the sample was men and the mean age was 48 years. There was significant variation between BMI categories with respect to gender, age, ethnicity, marital status, smoking status, education qualifications, social class of HRP, car ownership, economic activity,

housing tenure, bedrooms in household, area deprivation and claiming income support (all *P*<0.01).

Table 3.3: Selected descriptive statistics by BMI category

	Whole sample	Normal weight	Overweight	Class I obesity	Class II / III Obesity	*P* value
Observations						
Number	33,105	11,796	12,814	5,656	2379	
%	100	36	39	17	7	
EQ-5D score	0.871	0.900	0.876	0.840	0.782	<0.01
Predicted SES measure	29,652	30,914	30,159	27,831	25,821	<0.01
Missing income	14	14	14	14	15	0.46
SES quartile						
1 (most deprived)	25	22	24	29	35	
2	25	25	25	26	26	
3	25	26	26	24	23	
4 (least deprived)	25	28	26	21	17	<0.01
Gender						
Men	46	38	53	50	34	
Women	54	62	47	50	66	<0.01
Age (years)	48	43	50	52	50	<0.01

Source: HSE 2003-2006

Table 3.4 shows the results of the interval regression model of banded total annual household income against the individual, household and area socioeconomic variables. The coefficients show the predicted change in income if each of the dependent variables takes the value 1. Education, social class of HRP, number of cars and bedrooms in household are all positively correlated with income. Compared to being in paid employment all other types of economic activity are negatively associated with income. Claiming income support is negatively correlated with income, as is area deprivation. We used these coefficients to generate our predicted SES measure. Respondents were categorised into quartiles based on the predicted SES values. The

range of values for equivalised predicted income was -18,720[11] to 146,702; the cut-points used to set the quartiles were 17,542, 27,728 and 39,215.

[11] We convert the predicted income into quartiles and use these as our SES measure and do not use the actual income values in the further analysis. Hence, it is not a problem that we obtain negative predicted income values.

Table 3.4: Interval regression of total annual household income against SES indicators

	Coef.	z
Educational qualifications		
Degree or equivalent	Base category	
Higher education below degree	-9,799	-15.11
NVQ3/GCE A Level or equivalent	-7,685	-10.53
NVQ2/GCE O Level or equivalent	-10,231	-16.31
NVQ1/CSE other grade or equivalent	-11,363	-14.11
Foreign/other	-13,419	-17.25
No qualification	-10,779	-16.44
Social Class of HRP		
Professional	Base category	
Managerial technical	-4,809	-4.42
Skilled non-manual	-13,251	-11.51
Skilled manual	-13,268	-11.99
Semi-skilled manual	-14,341	-12.61
Unskilled manual	-14,652	-12.90
Cars owned by household		
Household has no car	Base category	
One	300	0.59
Two	7,561	10.79
Three or more	16,337	11.54
Economic activity status for last week		
In paid employment or self-employed	Base category	
Going to school or college full time	-10,331	-4.51
On a government training scheme	-15,454	-6.98
Doing unpaid work	-2,156	-0.27
Waiting to take up paid work already obtained	-3,446	-0.59
Looking for paid work or a government training scheme	-14,299	-14.23
Intending to look for work but prevented by temporary sickness	-14,720	-9.99
Permanently unable to work because of long-term sickness	-11,422	-15.50
Retired from paid work	-12,587	-23.01
Looking after home or family	-10,348	-9.87
Doing something else	-19,093	-10.11
Housing tenure		
Own outright	Base category	
Buying with help of a mortgage or loan	4,310	6.59
Pay part rent and part mortgage	-1,130	-0.62
Rent	-928	-1.32
Live rent free and/or squatting	-6,967	-4.74
Bedrooms in household		
One	Base category	
Two	683	1.12
Three	3,034	4.55
Four	10,627	10.63
Five or more	25,708	12.88
IMD quintile (SOA level)		
1 (least deprived)	Base category	
2	-2,937	-3.87
3	-3,946	-5.00
4	-5,558	-7.20
5 (most deprived)	-6,138	-7.96
Income support claimant		
No	Base category	
Yes	-2,041	-2.49
Constant	47,940	31.21
Observations	33,716	
Adjusted R^2	0.077	

Notes

HRP = household reference person. IMD = Index of Multiple Deprivation. SOA = super output area.

The unadjusted EQ-5D regressions are in Table 3.5. For all men combined and all women combined EQ-5D scores decreased significantly at higher levels of obesity (both $P<0.01$). The mean EQ-5D score among men of normal weight was 0.910, and for those in the overweight, class I obesity and class II / III obesity categories it was 0.894, 0.866, and 0.821, respectively. In women analogous figures were 0.904, 0.866, 0.827, and 0.775, respectively; the negative correlation of overweight and obesity with EQ-5D score were greater for women than for men.

We repeated the unadjusted analyses for all men combined and all women combined restricting the sample to individuals with non-missing income data only (28,374 observations; 13,016 men, 15,358 women). We then compared the predicted mean EQ-5D scores produced by both models for each of the BMI categories in both men and women using t-tests. There were no significant differences for every BMI category in both genders (all tests $P>0.23$). In the remainder of this chapter we focus on the results for all individuals in whom we were able to compute predicted SES values.

Table 3.5: Association between obesity and EQ-5D score (unadjusted)

	Normal weight				Overweight				Class I obesity				Class II/III obesity				Equal
	Coef.	**t**	**Mean**	**SE**	**Coef.**	**t**	**Mean**	**SE**	**Coef.**	**t**	**Mean**	**SE**	**Coef.**	**t**	**Mean**	**SE**	**means**
Men																	
All	Base category		0.910	0.002	-0.017	-4.76	0.894	0.002	-0.044	-9.38	0.866	0.003	-0.089	-9.00	0.821	0.008	*P*<0.01
SES quartile																	
1 (most deprived)	Base category		0.838	0.007	-0.038	-3.62	0.800	0.006	-0.072	-5.43	0.766	0.009	-0.160	-6.41	0.679	0.019	*P*<0.01
2	0.080	8.65	0.919	0.004	0.057	6.03	0.895	0.004	0.028	2.57	0.867	0.006	-0.019	-0.98	0.819	0.015	*P*<0.01
3	0.095	10.42	0.933	0.003	0.078	8.66	0.916	0.003	0.062	5.93	0.900	0.005	0.046	3.00	0.884	0.010	*P*<0.01
4 (least deprived)	0.107	12.25	0.946	0.003	0.097	11.25	0.936	0.002	0.087	9.07	0.925	0.004	0.079	5.74	0.918	0.009	*P*<0.01
Equal means	*P*<0.01				*P*<0.01				*P*<0.01				*P*<0.01				
Women																	
All	Base category		0.904	0.002	-0.038	-11.54	0.866	0.003	-0.078	-16.21	0.827	0.004	-0.130	-17.04	0.775	0.007	*P*<0.01
SES quartile																	
1 (most deprived)	Base category		0.841	0.006	-0.048	-5.48	0.792	0.006	-0.083	-7.54	0.758	0.009	-0.152	-10.26	0.689	0.013	*P*<0.01
2	0.072	9.90	0.913	0.003	0.033	4.07	0.873	0.005	-0.004	-0.39	0.836	0.008	-0.039	-2.61	0.802	0.013	*P*<0.01
3	0.079	10.81	0.919	0.003	0.056	7.03	0.896	0.005	0.030	3.06	0.871	0.007	0.001	0.10	0.842	0.012	*P*<0.01
4 (least deprived)	0.091	12.87	0.931	0.003	0.073	9.54	0.914	0.004	0.045	4.49	0.885	0.007	0.009	0.63	0.850	0.012	*P*<0.01
Equal means	*P*<0.01				*P*<0.01				*P*<0.01				*P*<0.01				

Notes

The dependent variable is EQ-5D score. The coefficients and t-statistics show the impact of the obesity variables on EQ-5D scores based on OLS. The mean and SE values are predicted mean EQ-5D scores, and their standard errors, for each BMI or BMI/SES category.

While the magnitude of the effect was reduced for all men combined and all women combined in the multivariate models, the negative correlation between obesity and EQ-5D scores remained significant for all men and all women after controlling for individual and household characteristics except for the longstanding and acute illness variables (Table 3.6; both $P<0.01$). Compared with normal weight, class II/III obesity was associated with a mean difference in EQ-5D score of -0.067 (t=-7.2, $P<0.01$) in men and -0.100 (t=-14.06, $P<0.01$) in women. For all men combined, while overall there was a significant negative correlation between obesity and EQ-5D score, there were no significant differences between overweight and normal weight.

Table 3.6: Association between obesity and EQ-5D score (controlling for individual and household characteristics)

	Normal weight				Overweight				Class I obesity				Class II/III obesity				Equal
	Coef.	**t**	**Mean**	**SE**	**Coef.**	**t**	**Mean**	**SE**	**Coef.**	**t**	**Mean**	**SE**	**Coef.**	**t**	**Mean**	**SE**	**means**
Men																	
All	Base category		0.898	0.002	-0.002	-0.45	0.897	0.002	-0.020	-4.25	0.878	0.003	-0.067	-7.20	0.831	0.007	*P*<0.01
SES quartile																	
1 (most deprived)	Base category		0.874	0.007	-0.012	-1.22	0.862	0.006	-0.044	-3.35	0.831	0.009	-0.139	-5.81	0.735	0.019	*P*<0.01
2	0.039	4.52	0.914	0.004	0.037	4.01	0.911	0.004	0.013	1.17	0.887	0.006	-0.036	-1.87	0.838	0.014	*P*<0.01
3	0.033	3.63	0.907	0.003	0.036	3.89	0.911	0.003	0.025	2.35	0.899	0.005	0.001	0.09	0.876	0.010	*P*=0.04
4 (least deprived)	0.021	2.00	0.895	0.004	0.023	2.13	0.897	0.004	0.017	1.48	0.891	0.005	0.004	0.27	0.879	0.010	*P*=0.35
Equal means	*P*<0.01				*P*<0.01				*P*<0.01				*P*<0.01				
Women																	
All	Base category		0.889	0.002	-0.014	-4.35	0.875	0.002	-0.044	-9.53	0.845	0.004	-0.100	-14.06	0.789	0.007	*P*<0.01
SES quartile																	
1 (most deprived)	Base category		0.882	0.006	-0.028	-3.31	0.854	0.007	-0.061	-5.81	0.821	0.009	-0.141	-9.99	0.740	0.013	*P*<0.01
2	0.026	3.72	0.908	0.004	0.006	0.76	0.888	0.005	-0.03	-2.91	0.852	0.008	-0.074	-5.14	0.808	0.012	*P*<0.01
3	0.011	1.42	0.893	0.004	0.004	0.43	0.886	0.005	-0.018	-1.81	0.863	0.007	-0.054	-3.87	0.828	0.012	*P*<0.01
4 (least deprived)	-0.007	-0.69	0.875	0.006	-0.011	-1.07	0.870	0.006	-0.033	-2.61	0.849	0.009	-0.064	-4.10	0.817	0.013	*P*<0.01
Equal means	*P*<0.01				*P*<0.01				*P*=0.01				*P*<0.01				

Notes

The dependent variable is EQ-5D score. The coefficients and t-statistics show the impact of the obesity variables on EQ-5D scores based on OLS. The mean and SE values are predicted mean EQ-5D scores, and their standard errors, for each BMI or BMI/SES category computed by setting the individual and household characteristics to their whole-sample mean values. The individual and household characteristics are age, age squared and age cubed, predicted equivalised total annual household income, ethnicity, marital status, smoking status, Government Office Region of residence and survey year.

Similar results, although at reduced levels, were found controlling for longstanding and acute illnesses as well (Table 3.7; both $P<0.01$). Compared with normal weight, class II and class III obesity was associated with a mean difference in EQ-5D score of -0.028 (t=-3.9, $P<0.01$) in men and -0.049 (t=-8.8, $P<0.01$) in women, after controlling for individual and household characteristics including the longstanding and acute illness variables.

Table 3.7: Associations between obesity and EQ-5D scores (controlling for individual and household characteristics including the longstanding and acute illness variables)

	Normal weight				Overweight				Class I obesity				Class II/III obesity				Equal
	Coef.	**t**	**Mean**	**SE**	**Coef.**	**t**	**Mean**	**SE**	**Coef.**	**t**	**Mean**	**SE**	**Coef.**	**t**	**Mean**	**SE**	**means**
Men																	
All	Base category		0.893	0.002	0.000	0.080	0.893	0.002	-0.006	-1.650	0.886	0.003	-0.028	-3.920	0.865	0.006	P<0.01
SES quartile																	
1 (most deprived)	Base category		0.881	0.006	-0.002	-0.260	0.878	0.006	-0.019	-1.910	0.861	0.009	-0.074	-4.050	0.807	0.017	P<0.01
2	0.019	2.700	0.899	0.004	0.017	2.370	0.898	0.004	0.007	0.790	0.887	0.006	-0.020	-1.250	0.861	0.014	P=0.03
3	0.016	2.220	0.897	0.004	0.020	2.640	0.900	0.003	0.020	2.260	0.900	0.005	0.019	1.550	0.900	0.010	P=0.9
4 (least deprived)	0.013	1.530	0.894	0.004	0.014	1.540	0.894	0.004	0.014	1.500	0.895	0.006	0.010	0.780	0.891	0.011	P=0.99
Equal means	*P=0.05*				*P<0.01*				*P<0.01*				*P<0.01*				
Women																	
All	Base category		0.879	0.002	-0.008	-3.220	0.871	0.002	-0.022	-5.710	0.857	0.003	-0.049	-8.780	0.830	0.005	P<0.01
SES quartile																	
1 (most deprived)	Base category		0.877	0.005	-0.021	-3.170	0.856	0.006	-0.029	-3.460	0.848	0.008	-0.078	-6.910	0.799	0.011	P<0.01
2	0.010	1.730	0.886	0.003	0.002	0.420	0.879	0.004	-0.022	-2.660	0.855	0.007	-0.029	-2.590	0.848	0.010	P<0.01
3	0.004	0.660	0.881	0.003	-0.002	-0.320	0.874	0.004	-0.007	-0.840	0.870	0.006	-0.031	-2.730	0.846	0.010	P<0.01
4 (least deprived)	-0.002	-0.290	0.874	0.005	-0.003	-0.380	0.873	0.005	-0.017	-1.670	0.859	0.008	-0.027	-2.140	0.849	0.011	P=0.02
Equal means	*P=0.08*				*P<0.01*				*P=0.14*				*P<0.01*				

Notes

The dependent variable is EQ-5D score. The coefficients and t-statistics show the impact of the obesity variables on EQ-5D scores based on OLS. The 'Mean' values are predicted mean EQ-5D scores for each BMI or BMI/SES category computed by setting the age variables, the individual characteristics and the longstanding illnesses to their sample mean values. The individual and household characteristics are age, age squared and age cubed, predicted equivalised total annual household income, ethnicity, marital status, smoking status, Government Office Region of residence, survey year, the longstanding illness variables (the number and type of longstanding illnesses, limiting longstanding illness) and acute illness.

The analysis by SES quartile reveals socioeconomic variation in the relationship between obesity and HRQL. The relationship between obesity and EQ-5D score depends on SES: the interaction between obesity and SES was statistically significant for both genders in the unadjusted models, the models controlling for individual and household characteristics except for the longstanding and acute illness variables, and in the model controlling for longstanding and acute illnesses as well (all $P<0.01$).

In the unadjusted models, in every SES quartile and in both men and women, mean EQ-5D scores were significantly decreased at higher levels of obesity (all $P<0.01$; Table 3.5). In addition, in every BMI category mean EQ-5D scores were significantly lower in more deprived SES quartiles (all $P<0.01$). For example, the mean EQ-5D scores in men with class II/III obesity were 0.679 in SES quartile 1 (most deprived) and 0.819, 0.884 and 0.918 in quartiles 2, 3 and 4 (least deprived), respectively. In women analogous figures were 0.689, 0.802, 0.842 and 0.850, respectively.

Similar trends were found when controlling for individual and household characteristics (Table 3.6). In every BMI category mean EQ-5D scores varied significantly by SES (all $P\leq0.01$), and EQ-5D scores were lowest in the most deprived SES groups (the mean EQ-5D score of normal weight women in the least deprived quartile was not significantly different to that of normal weight women in the most deprived quartile). In men, mean EQ-5D scores declined significantly in SES quartiles 1, 2 and 3 (all $P<0.05$), but not in quartile 4 (least deprived; $P=0.35$). In women, mean EQ-5D scores declined significantly with obesity in every SES group (all $P<0.01$). In every

model, for both men and women, the BMI/SES category with the lowest mean EQ-5D score was class II/III obesity/SES quartile 1.

Furthermore, similar results were obtained when we control for longstanding and acute illnesses (Table 3.7). In men, mean EQ-5D scores declined significantly in SES quartiles 1 and 2 (all $P<0.05$), but not in quartiles 3 and 4 (least deprived). In women, mean EQ-5D scores declined significantly with obesity in every SES group (all $P<0.01$).

To further examine variation across BMI groups and across SES groups separately we repeated the adjusted analyses by setting the individual and household characteristics to the mean values of each SES category, and then to the mean values of each obesity category (see appendix 3.2 and appendix 3.3). In these cases the variation in predicted mean EQ-5D scores between BMI or BMI/SES categories becomes more pronounced, because it also reflects differences in the values of the individual and household characteristics between categories.

3.5 Discussion

The aims of this chapter were to investigate the association between obesity and HRQL, and whether or not any observed relationship varies by SES. Our main findings are that obesity and HRQL are negatively correlated, and that the relationship between obesity and HRQL is significantly more pronounced in the lower SES groups than in the higher SES groups.

We provide evidence to show that overweight and obese men and women have significantly lower HRQL than those of normal weight. After controlling for individual and household characteristics, we obtain similar results, albeit at reduced levels. Our findings also show that the negative association between obesity and HRQL is greater in women than in men.

These results are qualitatively similar to those in other studies, which have also shown that HRQL is negatively correlated with obesity, and that the negative impact of obesity is diminished but not eliminated after controlling for individual and household characteristics (Jia & Lubetkin, 2005; Sach et al., 2007). Our findings are also quantitatively similar to Sach et al., (2007) who investigate the association between obesity and HRQL in a UK general practice population. They report that overweight, class I obesity and class II obesity are associated with a mean reduction in EQ-5D score across both genders combined of -0.023, -0.099 and -0.121, respectively. Our figures are -0.017, -0.044 and -0.089 for men, and -0.038, -0.078 and -0.130 for women. Our results controlling for individual and household characteristics but not longstanding and acute illnesses are also consistent; Sach et al. report that being overweight has a non-significant relationship with EQ-5D score and that being obese (all classes combined) is associated with a mean reduction in EQ-5D score of -0.04. Our analysis shows that being overweight has a non-significant relationship with EQ-5D score in men and a small but significant negative relationship in women (mean reduction -0.014). In men, class I obesity has a significant impact on EQ-5D score of -0.020, and class II/III obesity is associated with a mean reduction in EQ-5D score of -0.067. Comparable figures for women are -0.044, and -0.100, respectively.

This chapter also shows that the negative association between obesity and HRQL varies significantly by SES; the association is more negative in people from lower SES groups. Overweight and obese people in lower SES groups have lower HRQL than those of normal weight in the same SES group, and have lower HRQL than those in higher SES groups with the same weight. This trend is also observed after controlling for individual and household characteristics, although the statistical significance and magnitude of effects is diminished. While a number of studies have investigated the association between obesity and HRQL controlling for SES in multivariate analyses, we are not aware of any published studies that have stratified their analyses by SES groups and shown that the association between BMI groups and HRQL varies by SES.

Is the magnitude of the variation in EQ-5D scores between BMI/SES categories reported in this chapter important? Walters & Brazier (2005) investigated the smallest change in EQ-5D score that can be regarded as significant, i.e., the minimally important difference (MID; defined as the smallest difference which patients would perceive as important). Across a range of conditions (leg ulcer, back pain, early rheumatoid arthritis, limb reconstruction, osteoarthritis, irritable bowel syndrome, and chronic obstructive lung disease) the mean MID (range) for the EQ-5D was 0.074 (-0.011 to 0.140). Pickard, Neary & Cella (2007) investigated MIDs for the EQ-5D in the UK for 12 types of cancer and calculated estimates in the range 0.070 to 0.120. Applying these MIDs to our results, compared with the base category (normal weight/SES quartile 1 (most deprived)), the range of coefficients for men and women in our study was -0.160 to +0.107 and –0.152 to +0.091, respectively, in the unadjusted models (Table 3.5) and

-0.139 to +0.039 and -0.141 to +0.026 after controlling for individual and household characteristics (Table 3.6), with the largest differences occurring between SES quartiles 1 and 4 in the normal weight groups and between class II/III obesity and normal weight in SES quartile 1. Caution is required when drawing conclusions from these comparisons because the MIDs are derived from randomised controlled trials and longitudinal observational studies that were designed to assess the efficacy of treatment interventions, rather than evaluate differences in health input states. In addition, the MIDs vary by condition and none of the studies focused specifically on obesity; so MIDs for obesity are unknown, and may fall outside of the ranges given above. Nonetheless, these comparisons suggest that the variations in EQ-5D score reported in this study between BMI/SES categories are quantitatively important.

In Chapter 2 we suggest that obesity reduces health by increasing the health depreciation rate. Furthermore we suggested that the health depreciation rate might vary by SES and presented three hypotheses for this relationship. The findings in this chapter support hypothesis (b) of a more pronounced negative impact of obesity on health in the lower SES group compared with higher SES groups. Birch, Jerret and Eyles (2000) suggest that given the economic resources and favourable social circumstances, high SES individuals are more capable of tackling the harm of unhealthy lifestyles. While low SES groups are already weak and are more likely to suffer the comorbidities of unhealthy lifestyles. We ran regressions controlling for longstanding and acute illness and find that the magnitude of the associations was considerably reduced, which emphasise the role of obesity-related comorbidities. Hence, in the following

Chapter 4 we will analyse further their part in the association between obesity and HRQL.

The main implication of our findings is that more attention needs to be paid to accounting for heterogeneity in the pathways to health when undertaking research studies, in particular accounting for the modifying effect of SES when analysing the impact of obesity on health and interventions to reduce obesity. Given the common use of HRQL measures such as the EQ-5D in cost-effectiveness analyses, our findings also suggest that particular attention needs to be paid to the role of SES when undertaking economic evaluations of interventions to reduce obesity. The results show that obesity is associated more strongly with HRQL in lower SES groups compared with higher SES groups. Assuming that these associations are causal, it suggests that greater HRQL gains might be achieved by reducing obesity in lower SES groups. Such findings suggests that care is required when selecting HRQL estimates for BMI categories in cost-effectiveness analyses to ensure that the values used are appropriate to the SES group under investigation. It also suggests that the cost-effectiveness of programmes to reduce obesity may vary by SES group, and that sub-group analyses by SES may be warranted.

This chapter has a number of potential limitations related to endogeneity, making it difficult to draw inferences about the causal impact of obesity on EQ-5D. First, there may be omitted variable bias: there may be unobserved factors that affect both obesity and HRQL. One such factor might be the discount rate – the rate of time preference – which affects the weight that individuals give to their future health and

also their BMI (Morris, 2007). Another factor might be health problems. For example, individuals may be obese because they have a health problem, which precludes them from undertaking physical activity. Conversely, individuals might lose weight due to a health problem. Hence, it is difficult to establish the direction of the bias. We ran models controlling for limiting longstanding and acute illnesses. Although the magnitude of the association is reduced the results demonstrate a similar trend. The effect of obesity is still significantly more negative in the lower SES groups. Second, there may be measurement error: if BMI is mis-measured, and if the level of mis-measurement is associated with HRQL, the obesity coefficients may be biased (note that in our data height and weight were obtained during the interviewer visit and are not self-reported so the likelihood of this problem is reduced). Third, and related to this, there may be reporting bias with respect to HRQL: the EQ-5D, like any HRQL measure, is essentially subjective. It may therefore suffer from systematic reporting bias that is correlated with obesity and/or with SES. Fourth, our measure of obesity is BMI, which has been criticised, e.g., because it does not incorporate body fat, which is an independent predictor of ill health (Burkhauser & Cawley, 2008). Fifth, the HSE data are cross-sectional. We therefore model the impact of current obesity on current HRQL, when obesity in previous time periods might be more significant.

Most previous studies also suffer from these limitations. Nonetheless, we believe that our dataset offers a number of advantages over previous studies. Our BMI measure is based on actual height and weight measurements obtained during the interviewer visit, not on self-reported values. Thus, the likelihood of systematic measurement error is reduced. The number of observations in our sample is sufficiently large to permit

analyses stratified by gender, SES and BMI; our findings show that the impact of obesity on HRQL varies by all these factors. Also, we have better information on individual and household characteristics and so can argue that it is less likely that the estimated effects of obesity on HRQL in our models are due to their correlation with omitted variables.

To summarise, this chapter has shown that, as in previous studies, obesity is negatively correlated with HRQL. In addition, we have shown that the negative association between obesity and HRQL is greater in people from lower SES groups: overweight and obese people in lower SES groups have lower HRQL than those of normal weight in the same SES group, and have lower HRQL than those in higher SES groups of the same weight.

CHAPTER 4

Socioeconomic variation in the relationship between obesity and health related quality of life in individuals with obesity-related comorbidities

4.1 Introduction

In Chapter 3 we compared models with and without controlling for longstanding and acute illnesses. We illustrate that the negative association between obesity and HRQL is substantially reduced when we control for illnesses. However, obesity is still significant negatively associated with HRQL. This means that the loss in HRQL due to obesity is predominantly due to increased prevalence of illnesses in the obese (the risk factor effect). In this chapter we want to investigate further the role of the risk factor effect of obesity. First, we investigate how obesity-related comorbidities modify the association between obesity and HRQL. Second, we investigate how socioeconomic status influences this. The analysis allows us to explore the relationship between obesity and HRQL in individuals with obesity-related comorbidities. In addition, this analysis allows us to explore possible causes of socioeconomic variation in the association between BMI and HRQL identified in Chapter 3.

To approach the first aim we present a model where we have an obesity-related comorbidity (C_t) variable that can take on two values (no comorbidity=0, has

comorbidity=1). This obesity-related comorbidity could be diabetes, for example. In that case the C_t variable will take the value one if the person has diabetes and zero otherwise. We interact the obesity-related comorbidity with obesity (O_t). In this example we assume that the obesity variable can take on two values (normal weight=0, obese=1):

$$H_t = \beta_0 + \beta_1 O_t + \beta_2 C_t + \beta_3 O * C_t + \beta_4 X_t + u_t \quad (4.1)$$

H_t represents health, the X_t denotes all other covariates like age and other lifestyles (smoking), u_t is the error term, t denotes the time period and the β's are vectors of parameters to be estimated. To test if there is variation in the association between obesity and HRQL by obesity-related comorbidities we can test the hypothesis H0: $\beta_3=0$. If $\beta_3\neq0$ it means that H0 is rejected and the association between obesity and HRQL is not equal in individuals with and without comorbidities. Hence, the association between obesity and HRQL varies by obesity-related comorbidities (or vice versa).

The second aim of this chapter is to explore SES variation in the association between obesity-related comorbidities and HRQL. In Chapter 3 we found that the association between obesity and HRQL is more negative and more significant in lower SES individuals. One possible explanation for this is that obesity-related comorbidities are more negatively associated with HRQL in lower SES groups compared with higher SES groups. It might be that the favourable social circumstances that high SES individuals enjoy will act as a buffer against the negative consequences of comorbidities. For instance, favourable circumstances in higher SES individuals might increase their

diabetes control, which then again reduces the probability of suffering negative consequences of diabetes.

To investigate this we need a method for looking at variation across SES groups in the effect of obesity-related comorbidities on HRQL. We use a model with a third level interaction between obesity, SES and the comorbidity. The motivation behind using such an approach is that we have two interactions that need to be taken account of and we want to test a third interaction. The first interaction we need to take account of is the interaction between obesity and SES indentified in Chapter 3 ($O*SES_t$). The second is the interaction in equation 4.1 between obesity-related comorbidities and obesity ($O*C_t$). The third interaction, and the one we want to test, is between obesity-related comorbidities and SES ($C*SES_t$). Hence, we run the model below:

$$\begin{aligned} H_t &= \alpha_0 + \alpha_1 O_t + \alpha_2 C_t + \alpha_3 SES_t \\ &+ \alpha_4 O*C_t \\ &+ \alpha_5 O*SES_t \\ &+ \alpha_6 C*SES_t \\ &+ \alpha_7 O*C*SES_t + \alpha_8 X_t + u_t \end{aligned} \qquad (4.2)$$

In this example we assume that the SES variable (SES_t) can take on two values (high SES= 0, low SES=1). In the equation above α_1 represents the effect of obesity when both C_t=0 and SES_t=0. α_2 represents the effect of the comorbidity when O_t=0 and SES_t=0. α_3 represents the effect of SES when O_t=0 and C_t=0. The second level interactions α_4 controls for the effect of obesity in individuals with comorbidities if SES_t= 0, α_5 controls for the effect of obesity when SES_t=1 and C_t=0, and α_6 controls for

the effect of comorbidities when SES_t=1 and O_t=0. The third level interaction α_7 captures all variation in H_t due to obesity if an individual is low SES (SES_t=1) and has a comorbidity (C_t=1). We want to test if there is variation in the association between obesity-related comorbidities and HRQL by SES group. Hence, we want to test the hypothesis H0: α_6=0. However, this is not straight forward as we need to take into account the third level interaction (α_7) and we will later discuss how to do this. If $\alpha_6 \neq 0$ across obesity groups it means that H0 is rejected and the association between obesity-related-comorbidities and HRQL is not equal across SES groups. Hence, the association between obesity-related comorbidities and HRQL varies by SES (or vice versa).

Little is known about the associations between obesity and HRQL in individuals suffering from various comorbidities. We have found two papers that investigate the association between obesity and HRQL in individuals with diabetes. However, only one of these compares the results in patients with and without diabetes to investigate whether or not there are significant differences. Furthermore, both studies use a sample of patients recently discharged from hospital. Hence, they might not represent the general population as individuals who do not have diabetes have been to the hospital for some other reason than diabetes. The more recent of the two studies, by Lee et al. (2005), investigates the association between obesity and EQ-5D score in individuals with type 1 diabetes, type 2 diabetes and no diabetes in Wales controlling for age. They find that obesity is negatively correlated with EQ-5D score for individuals with and without diabetes. Furthermore they found that there were no differences in the association between obesity and HRQL between those with and without diabetes

(non-significant interactions between obesity and diabetes status). Redekop et al. (2002) investigate the association between obesity and EQ-5D score in Dutch patients with type 2 diabetes. They find that obesity (BMI>30 kg/m^2) is negatively associated with EQ-5D score compared with none obese patients (BMI<30 kg/m^2) controlling for age, gender, diabetes related complications, and insulin therapy. However, this study limits their sample to individuals with diabetes and does not compare the results to individuals without diabetes. We are not aware of any studies investigating this effect across SES groups.

To recap, first we investigate if the association between obesity and HRQL varies by obesity-related comorbidities. Second, we investigate if the association between obesity-related comorbidities and HRQL varies by SES.

4.2 Methods

4.2.1 Data and variables

We use the same dataset as in Chapter 3 (HSE: 2003-2006). We focus on 4 of 42 possible longstanding illnesses in HSE: diabetes; stroke; heart attack or angina; and, hypertension.

There are numerous studies that link obesity to a number of comorbidities. The four comorbidities in this chapter are chosen based on two key reports: Foresight: Tackling Obesities: Future Choices Project Report (2nd Ed.) (Butland et al., 2007); and, Tackling

Obesity in England (National Audit Office, 2001). These reports have conducted extensive reviews of international literature with an aim of making the results relevant to the English population.

In Table 4.1 we see the rationale for the chosen longstanding illnesses. The reports only describe comorbidities that are associated with obesity; hence we cannot draw any conclusions regarding causality. For a list of all the longstanding illness categories used in HSE see Appendix 4.1.

Table 4.1: Longstanding illnesses categories in HSE that are related to obesity

Longstanding illness	**Rationale, taken from: National Audit Office*, Tackling obesity in England,* 2001. The Foresight report, Tackling Obesities: Future Choices – Project Report**
Diabetes	Obese men and women are 12.7 and 5.2 times more likely to have diabetes type 2, respectively.
Stroke	Obese men and women are 1.3 times more likely to have suffered a stroke.
Heart attack or angina	Obese men and women are 1.8 times more likely to have angina.
Hypertension/high blood pressure	Obese men and women are 4.2 and 2.6 times more likely to have hypertension, respectively.

Note
Obese and non obese groups vary between the groups used to estimate the relative risks. Hence, the risks are not comparable across illnesses.

The reports identify additional obesity-related comorbidities like various cancers, osteoarthritis and digestive complaints. For example, obese men and women are 2.7 and 3 times more likely to have cancer of the colon, respectively. However, in HSE for these diseases the categories are too wide and includes a number of longstanding illnesses that have not been associated with obesity, which will dilute the effect overall

and make the coefficients smaller and less significant. In Appendix 4.2 we provide a description of each illness category.

Consistent with Chapter 3, EQ-5D is our measure of HRQL and BMI is divided into five BMI categories (but we do not report the results for the underweight category). Furthermore, we use the same predicted SES measure and cut points as in Chapter 3 (see Table 3.4, Chapter 3). In addition, we run models controlling for: age (cubic function); ethnicity (nine categories); marital status (six categories); smoking status (four categories); Government Office Region of residence (nine categories); predicted equivalised total annual household income (continuous variable); and, survey year (four categories). We do not stratify by gender in this analysis because of the low number of individuals in each disease category. We control for gender by including a dummy variable taking the value of one if the individual is a man and zero otherwise[12].

4.2.2 Statistical analysis

We explain the following methods in three parts. We start with the methods used to investigate whether or not the association between BMI and HRQL varies by obesity-related comorbidities. We then explain the methods used to investigate socioeconomic variation in the association between obesity-related comorbidities and HRQL. Lastly, we explain the methods used to investigate the association between obesity and HRQL in individuals without any longstanding illnesses. Later the results will be explained under the same headings.

[12] We investigated if interacting gender and age as a cubic function improved model fit using a likelihood ratio test. The test did not indicate that interacting age and gender improved the model fit (P=0.2692)

4.2.2.1 Variation in the association between BMI and HRQL by obesity-related comorbidities: methods

We regress EQ-5D score against BMI categories including interaction terms between BMI categories and obesity-related comorbidity (*ORC*). In equation 4.4 below the ORC_i variable represents one of the obesity-related comorbidities (diabetes; stroke; heart attack or angina; and, hypertension). For example, in the case when it represents diabetes it is a dummy taking on the value of one if the person has diabetes and zero otherwise. The omitted categories are individuals not suffering from the comorbidity in question (ORC_i=0) and normal weight (BMI_i=1). Z_i denotes the other covariates. We fit four models, one for each of the comorbidities.

$$h_i = \beta_0 + \beta_1 \sum_{a=1}^{5} BMI_{ai} + \beta_2 ORC_i + \beta_3 \sum_{a=1}^{5} BMI_{ai} * ORC_i + \beta_3 Z_i + u_i \qquad (4.4)$$

β_1 provides the impact of BMI categories in individuals without the comorbidity (note that individuals can have other comorbidities but the one that is analysed), β_2 provides the impact of the comorbidity in individuals in the normal weight category. β_3 is the interaction between BMI and the comorbidity in individuals who are either overweight or obese (i.e., whether or not the association between obesity and EQ-5D score depends on the obesity-related comorbidity, or vice versa). We test if this interaction is statistically significant in models using two-way ANOVA.

After every regression we compute the predicted mean EQ-5D score for each BMI/comorbidity category. We do this by fixing the covariates at their whole-sample mean values so that the variation in predicted mean EQ-5D scores is a function of BMI categories and how this varies by obesity-related comorbidities, and is not affected by the values of the other covariates.

We test for equality in predicted mean EQ-5D scores between BMI categories (across each comorbidity and within each comorbidity) using Wald tests.

As in Chapter 3 we apply survey weights reported in the HSE to each observation to adjust for the fact that different observations have different probabilities of being selected for participation in the survey. In addition, we control for clustered sampling within PSUs using unique PSU/year identifiers that produce Huber/White/sandwich robust variance estimators that allow for within-group dependence (Kish & Frankel, 1974).

4.2.2.2 Socioeconomic variation in the association between obesity-related comorbidities and HRQL: Methods

We now introduce SES explicitly into this analysis and use a third level interaction approach as it allows us to jointly test the interaction between SES and comorbidity across all BMI groups to see if we get a significant interaction between the comorbidity and SES.

We rerun model 4.1 including triple categorical variable interaction terms between the BMI categories, comorbidities and the SES quartiles, with "normal weight/no comorbidity/SES quartile 1 (most deprived)" as the omitted category. As above we have four comorbidities and run one model for each.

$$
\begin{aligned}
h_i &= \beta_0 + \beta_1 \sum_{a=1}^{5} BMI_{ai} + \beta_2 ORC_i + \beta_3 \sum_{b=1}^{4} SES_{bi} \\
&+ \beta_4 \sum_{k=1}^{5} BMI_{ai} * ORC_i \\
&+ \beta_5 \sum_{a=1}^{5} \sum_{b=1}^{4} BMI_{ai} * SES_{bi} \\
&+ \beta_6 \sum_{b=1}^{4} SES_{bi} * ORC_i \\
&+ \beta_7 \sum_{a=1}^{5} \sum_{b=1}^{4} BMI_{ai} * SES_{bi} * ORC_i \\
&+ \beta_8 Z_i + u_i
\end{aligned}
\tag{4.6}
$$

This will provide EQ-5D values specific to each BMI/SES/comorbidity combination. Here β_1 is the effect of BMI_i when ORC_i=0 (not having the comorbidity in question) and SES_i=1, β_2 is the effect of the comorbidity when BMI_i=1 (normal weight) and SES_i=1 and β_3 is the effect of SES when BMI_i= 1 and ORC_i=0. β_4 is the interaction between BMI_i and ORC_i when SES_i=1, β_5 is the interaction between BMI_i and SES_i when ORC_i=0 and β_6 is the interaction between SES_i and ORC_i when BMI_i=1. β_7 is the effect of the interaction when BMI_i>1 (individuals are overweight or obese), SES_i>1 (individuals are not in the lowest SES quartile) and ORC_i=1 individuals have the comorbidity in question.

We want to test the interaction between SES and BMI. To do this we follow the steps by UCLA Academic Technology Services (2011) for testing three-way interactions using

ANOVA models[13]. This means that we test the interaction SES*comorbidity for the normal weight, overweight, obese class I and obese class II/III and calculate a mean across these groups.

We compute the predicted mean EQ-5D score for each BMI/SES/comorbidity category by fixing the covariates at their whole-sample mean values, so that the variation in predicted mean EQ-5D scores is a function of BMI categories and how this varies by obesity-related comorbidities and SES.

Table 4.2 summarise the models used in this chapter.

Table 4.2: Summary of models

Model	Comorbidity	Interactions	Covariates
1	Diabetes	BMI/comorbidity	Individual and household characteristics
2	Stroke	BMI/comorbidity	Individual and household characteristics
3	Heart attack and angina	BMI/comorbidity	Individual and household characteristics
4	Hypertension	BMI/comorbidity	Individual and household characteristics
5	Diabetes	BMI/SES/comorbidity	Individual and household characteristics
6	Stroke	BMI/SES/comorbidity	Individual and household characteristics
7	Heart attack and angina	BMI/SES/comorbidity	Individual and household characteristics
8	Hypertension	BMI/SES/comorbidity	Individual and household characteristics

[13] First, we run the ANOVA regression. Second, we generate predicted means based on the ANOVA. Third, we test the predicted means across SES and comorbidity groups for different values of the BMI variable. A significant value means that the association between the comorbidity and HRQL varies by SES groups across BMI groups.

4.3 Results

The total number of respondents having each of the chosen illnesses can be found in Table 4.3. We see that for some of the illnesses caution is needed due to low numbers especially when stratified by SES.

Table 4.3: Number of observations by BMI category, illness category and SES quartile

	Total	Normal weight	Over-weight	Class I obesity	Class II/III obesity
Diabetes					
All	1234	191	425	382	236
SES quartile					
1 (most deprived)	514	66	168	170	110
2	287	40	96	98	53
3	249	45	93	65	46
4 (least deprived)	184	40	68	49	27
Stroke					
All	219	55	85	57	22
SES quartile					
1 (most deprived)	121	29	53	27	12
2	46	9	17	16	4
3	31	10	9	8	4
4 (least deprived)	21	7	6	6	2
Hypertension					
All	2340	389	922	666	363
SES quartile					
1 (most deprived)	850	152	319	251	128
2	565	92	221	152	100
3	471	61	200	137	73
4 (least deprived)	454	84	182	126	62
Heart attack and angina					
All	684	131	285	190	78
SES quartile					
1 (most deprived)	379	75	154	107	43
2	157	26	61	49	21
3	91	20	40	19	12
4 (least deprived)	57	10	30	15	2

Source: HSE 2003-2006

We see descriptive statistics of each illness across each BMI category in Table 4.4. Almost four percent of the individuals in the sample have diabetes (either type 1 or type 2), 0.7% of the individuals in the sample have had a stroke, 7.1% have hypertension and 2.1% have had a heart attack or angina. The percentage suffering from each of the illnesses was lowest in the normal weight categories and highest in the obesity class I and obesity class II/III categories across all illnesses. There were significant variations between BMI categories with respect to all of the illnesses. In

Appendix 4.3 we have stratified the statistics in Table 4.4 by SES groups. We see that a larger share of individuals suffer from comorbidities in the lower SES groups. Also the difference between the obese and the normal weight is larger in the lower SES groups than in the higher. This suggests that the prevalence of obesity related comorbidities is higher in the lower SES groups.

Table 4.4: Descriptive statistics by BMI category

	Whole sample	Normal weight	Overweight	Class I Obesity	Class II/III obesity	*P value*
Diabetes						
Yes	3.7	1.6	3.3	6.8	9.9	
No	96.3	98.4	96.7	93.3	90.1	*P<0.01*
Stroke						
Yes	0.7	0.5	0.7	1.0	0.9	
No	99.3	99.5	99.3	99.0	99.1	*P<0.01*
Hypertension						
Yes	7.1	3.3	7.2	11.8	15.3	
No	92.9	96.7	92.8	88.2	84.7	*P<0.01*
Heart attack and angina						
Yes	2.1	1.1	2.2	3.4	3.3	
No	97.9	98.9	97.8	96.6	96.7	*P<0.01*
No illnesses						
Yes	54.7	63.1	53.6	45.3	39.4	
No	45.3	36.9	46.4	54.7	60.6	*P<0.01*

Source: HSE 2003-2006

Note

All statistics are %.

We will present the result of the analysis in two parts, similar to the methods section. First, we will present the results of variation in the association between obesity and HRQL by obesity-related comorbidities. Then we will present the findings on socioeconomic variation in the association between obesity-related comorbidities and HRQL.

4.3.1 Variation in the association between BMI and HRQL by obesity- related comorbidities: results

The predictions based on the EQ-5D regressions controlling for individual and household characteristics across each illness category are in Table 4.5. EQ-5D score decreases significantly at higher levels of BMI across individuals in each comorbidity category: diabetes; hypertension; stroke; and, heart attack and angina. For example the mean EQ-5D score among normal weight individuals with diabetes was 0.834 and for those in the overweight, class I obesity and class II/III obesity categories it was 0.864, 0.817, and 0.701, respectively.

Across the four comorbidities the association between BMI groups and EQ-5D score is most pronounced in individuals suffering from heart attack or angina, where the mean in EQ-5D score falls from a mean of 0.782 in the normal weight category to 0.561 in the obese class II/III category. The association between BMI groups and EQ-5D score was least pronounced in individuals suffering from hypertension, where the mean EQ-5D score falls from a mean of 0.888 in the normal weight category to 0.774 in the obese class II/III category.

When we look at individuals with and without each of the four comorbidities, we find the largest disparities between individuals who have suffered a stroke and individuals who have not suffered a stroke[14]. In the normal weight category the mean EQ-5D score was 0.895 in individuals who have not suffered a stroke and 0.682 in individuals who have had a stroke. Conversely, we identify the smallest disparities between individuals

[14] Note that both individuals who have suffered a stroke and individuals who have not suffered a stroke can have other comorbidities

with and without hypertension, where the mean EQ-5D score across normal weight individuals was 0.888 and 0.895, respectively.

The comparison between individuals with and without each comorbidity shows variation in the association between obesity and HRQL. The association between obesity and HRQL depends on obesity-related comorbidities. The interactions between obesity and: diabetes; heart attack or angina; and, stroke were statistically significant. The relationship between obesity and HRQL was more pronounced in individuals with each of these comorbidities than individuals without. E.g. in the diabetes group the predicted mean EQ-5D score falls from a value of 0.834 in the normal weight group to 0.701 in the class II/III category. In the comparator group (diabetes: no) similar number were 0.895 to 0.819, respectively.

The interaction between hypertension and BMI was not significant. This means that there were no significant variations in the associations between obesity and EQ-5D score between individuals with and without hypertension. However, a similar trend was identified.

Table 4.5: The relationship between obesity and HRQL by comorbidities

	Normal weight	Over-Weight	Obese Class I	Obese Class II/III	*Equal means*
Diabetes					
Yes	0.834	0.864	0.817	0.701	*P<0.01*
No	0.895	0.886	0.864	0.819	*P<0.01*
Equal means	*P<0.01*	*P=0.03*	*P<0.01*	*P<0.01*	
Significant interaction			*P<0.01*		
Hypertension					
Yes	0.888	0.875	0.84	0.774	*P<0.01*
No	0.895	0.887	0.864	0.814	*P<0.01*
Equal means	*P=0.54*	*P=0.08*	*P<0.01*	*P<0.01*	
Significant interaction			*P=0.17*		
Heart attack and angina					
Yes	0.782	0.77	0.715	0.561	*P<0.01*
No	0.896	0.888	0.865	0.816	*P<0.01*
Equal means	*P<0.01*	*P<0.01*	*P<0.01*	*P<0.01*	
Significant interaction			*P<0.01*		
Stroke					
Yes	0.682	0.722	0.706	0.461	*P<0.01*
No	0.895	0.887	0.863	0.811	*P<0.01*
Equal means	*P<0.01*	*P<0.01*	*P<0.01*	*P<0.01*	
Significant interaction			*P<0.01*		

Notes

Individuals can have other longstanding illnesses in addition to the ones in each comorbidity group.

The dependent variable is EQ-5D score. The values are predicted mean EQ-5D scores, for each BMI or BMI/SES category computed by setting the individual and household characteristics to their whole-sample mean values. The individual and household characteristics are age, age squared and age cubed, predicted equivalised total annual household income, ethnicity, marital status, smoking status, Government Office Region of residence and survey year. The underlying model the predictions are based on is OLS.

4.3.2 Socioeconomic variation in the association between obesity-related comorbidities and HRQL: results

Table 4.6 shows the predicted mean EQ-5D score in each BMI category across the four obesity-related comorbidities by SES quartile.

In each BMI category across each obesity-related comorbidity the mean EQ-5D score was lower in more deprived SES quartiles. For example, the mean EQ-5D score in diabetic individuals with type II/III obesity were 0.627 in SES quartile 1 (most deprived) and 0.712, 0.756 and 0.82 in quartiles 2, 3 and 4 (least deprived), respectively.

In the introduction we stated that we aimed to investigate socioeconomic variation in the impact of obesity-related comorbidities on HRQL. Hence, we now look at the variation in EQ-5D score between individuals with and without the comorbidity within the same BMI group across SES groups. The analysis by SES quartile reveals socioeconomic variation in the relationship between obesity-related comorbidities and HRQL: the interaction between obesity-related comorbidities and SES was statistically significant for each comorbidity (diabetes; stroke; and, hypertension) (all: $P<0.01$). However, this was not found when the comorbidity in question was the 'heart attack and angina' category, where the interaction was not significant ($P=0.2$). We see that the association between obesity-related comorbidities and HRQL is more pronounced in the lower SES quartiles than in the higher SES quartiles. For example, obese class II/III individuals in SES quartile 1 (most deprived) with and without diabetes has a mean EQ-5D score of 0.627 and 0.759, respectively. This makes up a mean difference in EQ-5D score of 0.132 (0.759-0.627=0.132). Similar numbers for SES quartile 4 (least deprived) was 0.82 and 0.849, respectively. This makes up a mean difference in EQ-5D score of 0.028 (0.849-0.82=0.028). This trend persist in the obesity classes across all comorbidities, except for some groups with hypertension where the association was more negative in the higher SES groups. In the normal weight category we find similar trends of a more pronounced association between: diabetes; stroke and hypertension and HRQL in the lower SES quartiles. In general this trend shows that association between obesity-related comorbidities and HRQL is more pronounced in lower SES individuals.

These findings support the hypothesis of a more pronounced association between obesity-related comorbidities and HRQL in lower SES groups. This means that not only do lower SES individuals have a lower HRQL in general; they also have a larger reduction in EQ-5D score when they get obesity-related comorbidities.

Table 4.6: Impact of comorbidities on the relationship between obesity and HRQL across SES quartiles

	Normal weight	Over-Weight	Obese Class I	Obese Class II/III	*Equal means*
Diabetes					
SES group 1 (most deprived)					
Has diabetes	0.774	0.845	0.786	0.627	*P<0.01*
No diabetes	0.882	0.858	0.829	0.759	*P<0.01*
Equal means	*P=0.01*	*P=0.50*	*P=0.04*	*P<0.01*	
SES group 2					
Has diabetes	0.806	0.847	0.828	0.712	*P=0.05*
No diabetes	0.914	0.901	0.871	0.832	*P<0.01*
Equal means	*P=0.01*	*P=0.01*	*P=0.08*	*P=0.01*	
SES group 3					
Has diabetes	0.866	0.9	0.808	0.756	*P<0.01*
No diabetes	0.902	0.898	0.885	0.858	*P<0.01*
Equal means	*P=0.28*	*P=0.93*	*P<0.01*	*P=0.01*	
SES group 4 (least deprived)					
Has diabetes	0.91	0.857	0.876	0.82	*P=0.02*
No diabetes	0.885	0.885	0.873	0.849	*P<0.01*
Equal means	*P=0.07*	*P=0.16*	*P=0.873*	*P=0.34*	
*Test diabetes*SES = 0*			*P<0.01*		
Heart attack and angina					
SES group 1 (most deprived)					
Has heart attack or angina	0.77	0.724	0.677	0.491	*P<0.01*
No heart attack and angina	0.882	0.862	0.833	0.754	*P<0.01*
Equal means	*P<0.01*	*P<0.01*	*P<0.01*	*P<0.01*	
SES group 2					
Has heart attack or angina	0.751	0.806	0.742	0.567	*P<0.01*
No heart attack and angina	0.914	0.901	0.873	0.831	*P<0.01*
Equal means	*P=0.01*	*P<0.01*	*P<0.01*	*P<0.01*	
SES group 3					
Has heart attack or angina	0.802	0.77	0.753	0.688	*P=0.63*
No heart attack and angina	0.902	0.9	0.883	0.854	*P<0.01*
Equal means	*P=0.06*	*P<0.01*	*P=0.01*	*P=0.02*	
SES group 4 (least deprived)					
Has heart attack or angina	0.796	0.871	0.761	0.794	*P=0.11*
No heart attack and angina	0.887	0.885	0.874	0.848	*P<0.01*
Equal means	*P=0.01*	*P=0.54*	*P=0.03*	*P=0.60*	
*Test 'h*eart attack and angina*'*SES = 0*			*P=0.2*		
Stroke					
SES group 1 (most deprived)					
Has had a stroke	0.634	0.687	0.727	0.355	*P<0.01*
No stroke	0.882	0.86	0.827	0.748	*P<0.01*
Equal means	*P<0.01*	*P<0.01*	*P=0.01*	*P<0.01*	
SES group 2					
Has had a stroke	0.755	0.777	0.746	0.421	*P=0.08*
No stroke	0.913	0.9	0.87	0.825	*P<0.01*
Equal means	*P=0.1*	*P=0.03*	*P=0.09*	*P<0.01*	
SES group 3					
Has had a stroke	0.629	0.733	0.425	0.609	*P=0.12*
No stroke	0.902	0.899	0.884	0.852	*P<0.01*
Equal means	*P<0.01*	*P=0.08*	*P<0.01*	*P=0.07*	
SES group 4 (least deprived)					
Has had a stroke	0.816	0.76	0.813	0.842	*P=0.57*
No stroke	0.885	0.885	0.873	0.847	*P<0.01*
Equal means	*P=0.01*	*P<0.01*	*P=0.41*	*P=0.96*	
*Test stroke*SES = 0*			*P<0.01*		
Hypertension					
SES group 1 (most deprived)					
Has hypertension	0.841	0.833	0.832	0.715	*P<0.01*
No hypertension	0.882	0.861	0.824	0.747	*P<0.01*
Equal means	*P=0.05*	*P=0.05*	*P=0.62*	*P=0.24*	
SES group 2					
Has hypertension	0.929	0.876	0.81	0.754	*P<0.01*
No hypertension	0.912	0.901	0.875	0.835	*P<0.01*
Equal means	*P=0.28*	*P=0.09*	*P<0.01*	*P<0.01*	

Table continues on the next page

Table continued

SES group 3					
Has hypertension	0.895	0.914	0.847	0.85	*P<0.01*
No hypertension	0.901	0.898	0.885	0.85	*P<0.01*
Equal means	*P=0.79*	*P=0.15*	*P=0.04*	*P=0.99*	
SES group 4 (least deprived)					
Has hypertension	0.901	0.883	0.865	0.816	*P<0.01*
No hypertension	0.885	0.885	0.873	0.852	*P<0.01*
Equal means	*P=0.30*	*P=0.90*	*P=0.53*	*P=0.08*	
*Test hypertension*SES = 0*			*P<0.01*		

Notes

The dependent variable is EQ-5D score. The values are predicted mean EQ-5D scores, for each BMI or BMI/SES category computed by setting the individual and household characteristics to their whole-sample mean values. The individual and household characteristics are age, age squared and age cubed, predicted equivalised total annual household income, ethnicity, marital status, smoking status, Government Office Region of residence and survey year. The underlying regression model is OLS.

4.4 Discussion

The aim of this chapter was twofold. First, we investigate the association between obesity and HRQL in individuals with obesity-related comorbidities. Second, we explore whether or not there are socioeconomic variation in the association between obesity-related comorbidities and HRQL.

We provide evidence to show that people with obesity-related comorbidities have lower HRQL than those without obesity-related comorbidities in the same BMI group. Furthermore, overweight and obese people with: diabetes; hypertension; stroke; and, heart attack and angina, have significantly lower HRQL than those of normal weight. I.e. there is an association between obesity and HRQL in a population with obesity-related comorbidities. Although, we would expect this result we are not aware of any studies that has investigated this, except for diabetes (See: Lee et al., 2005; Redekop et al., 2002).

The association between obesity and HRQL was more negative in individuals with obesity-related comorbidities than in individuals without. The interactions between

BMI groups and obesity-related comorbidities (diabetes; heart attack or angina; and, stroke) were significant (*P<0.01*). Hence, the association between BMI groups and HRQL depends on obesity-related comorbidities (or vice versa). This means that the associations between obesity and HRQL are more negative in individuals with obesity-related comorbidities than individuals without obesity-related comorbidities. However, this was not identified in individuals with hypertension as the interaction was not significant.

This part of the analysis reveals a number of potential endogeneity issues. There may be omitted variable bias: one such variable might be disease severity. For example, imagine we have two types of stroke "severe" and "less severe". Individuals who suffer a "severe" stroke may be more precluded from undertaking physical activity than individuals who have suffered a "less severe" stroke. Hence, individuals who suffer a more "severe" stroke will provide lower EQ-5D scores (because of the more severe stroke) and be more likely to be obese (because of the more severe stroke), which would mean that we overestimate the impact of obesity on HRQL. In the case of other diseases this effect might be opposite, so that individuals who have a more severe version of the illness lose weight and are less likely to be obese. In this case we would underestimate the impact of obesity on HRQL.

These results are qualitatively similar to those in other studies, which have also shown that HRQL is negatively correlated with obesity for individuals suffering from diabetes. Lee et al. (2005) demonstrate a significant negative association between obesity and EQ-5D score across individuals with diabetes type 1, diabetes type 2 and without

diabetes controlling for age. They include individuals (six weeks) post discharged from an outpatient or inpatient clinic in Wales. They find that the age standardised EQ-5D score in normal weight (BMI: 18.5-24.9 kg/m^2) individuals with Type 1 diabetes, Type 2 diabetes and no diabetes were 0.706, 0.616 and 0.720, respectively. In this chapter values in individuals with and without diabetes were 0.834 and 0.895, respectively. Hence, we predict higher values in general, which might be because we look at the general population while they look at individuals who have had an inpatient stay or outpatient visits the last six weeks. Redekop et al. (2002) shows that being obese (BMI>30 kg/m^2) compared to not being obese (BMI<30 kg/m^2) reduce the EQ-5D score in diabetic individuals with -0.044 (P<0.01) controlling for age, gender, diabetes related complications, and insulin therapy. Our results show that, in diabetic individuals, the reduction in mean EQ-5D score going from normal weight to being obese class I or obese class II/III was -0.017 and -0.133, respectively.

The result also contradicts the findings of Lee et al. (2005). They find an insignificant interaction term between BMI and diabetes status, and conclude that there is no variation in the association between obesity and HRQL by diabetes status. We find this interaction to be significant. We are not aware of any studies looking at this effect across other obesity-related comorbidities than diabetes.

Our second analysis revealed that the negative association between obesity-related comorbidities and HRQL varies by SES. The association between obesity-related comorbidities and HRQL was more negative in individuals in lower SES quartiles than in higher SES quartiles. The interactions between SES and obesity-related comorbidities

(diabetes; hypertension; and, stroke) were significant (*P<0.01*). Hence, the association between obesity-related comorbidities and HRQL depends on SES (or vice versa). However, this was not found for the 'heart attack and angina' category, where the interaction was not significant (*P=0.2*), but the trend showed a more pronounced association in the lower SES groups. This could explain why we, in Chapter 3, found that lower SES individuals have more a more pronounced association between BMI and HRQL than higher SES individuals. Obesity increases the probability of obesity-related comorbidities; however the comorbidities cause more problems in lower SES individuals. Note that this also depends on the prevalence of obesity-related comorbidities across obesity categories and SES. In Appendix 4.3 we illustrated that the prevalence of obesity-related comorbidities increased more in lower SES groups than in higher SES groups going from normal weight to obese. Hence, this is also a reason for our findings in Chapter 3.

It is not clear why obesity-related comorbidities would cause more problems in lower SES groups. We offer some suggestions. It has been shown that lower SES is associated with lower adherence to medication treatments including lipid lower medication (Avorn et al., 1998). For instance for diabetes, it could be the case that favourable economic and social circumstances improve individuals' ability to monitor and control glycemic status. Poor glycemic control is associated with development of diabetic complications (Golden et al., 2003) and therefore makes diabetes worse in lower SES groups. However, one must be aware that it is difficult to establish causality as this analysis might suffer from endogeneity issues. We can not exclude the possible reverse causality. For example, individuals who have more problems with controlling diabetes

have a tendency to become lower SES (e.g., loss in income due to poor health) as a result of the poorly controlled diabetes.

There are several implications of the results in this chapter. They suggest, as earlier studies, that there is an association between obesity and HRQL in individuals with obesity-related comorbidities. Furthermore, this is the first study that we are aware of that suggests that this association is significantly different for individuals with and without obesity-related comorbidities. The association is more negative in individuals with obesity-related comorbidities. The implication is that attention needs to be paid to the role of comorbidities when undertaking economic evaluation of obesity interventions. Care is required when selecting HRQL estimates for BMI categories to ensure that the values used are appropriate to the prevalence of obesity-related comorbidities in the population of interest.

We find that obesity-related comorbidities are associated with a more pronounced loss in HRQL in lower SES groups compared with higher SES groups. This could explain the results obtained in Chapter 3. However, as mentioned this also depends on the incidence and prevalence of comorbidities across SES groups. Nevertheless, it suggests that part of the reason why there is SES variation in the association between obesity and HRQL is that lower SES individuals have more problems with the obesity-related comorbidities.

These findings can also be generalised to explain socioeconomic variation in the impact of other lifestyle-related risk factors on health. For example, Birch et al. (2000) and

Pampel and Rogers (2004) findings on socioeconomic heterogeneities in the impact of smoking on health. Smoking is associated with increased prevalence of comorbidities, which might have an increased burden in lower SES groups. There has been a focus on comorbidities as a reason for the variations in the associations between lifestyle and health. For example, Birch et al. (1997) suggest that high SES individuals might have a better 'adaptive defence' against smoking-related diseases. However, to our awareness it has never been tested. These findings suggest that it is related to the comorbidities.

This study also draws attention to a potential weakness in studies investigating the impact of obesity on HRQL. That is that we do not know if it is obesity that decreases HRQL, or if it is the severity of the comorbidity that increase/reduce BMI and reduce HRQL. Hence, it is emphasised that we have investigated *associations* between obesity and HRQL and not the *impact* of obesity on HRQL. Nevertheless, this study is the first to investigate and show significant interactions between obesity and obesity-related comorbidities. In this way it suggests that more research is needed in this area to establish causal relationships.

In addition to the endogeneity issues mentioned above, this study suffers from a number of potential limitations. There may be other omitted variables in addition to the ones mentioned above. One such factor might be the discount rate – the rate of time preference – which affects the weight that individuals give to their future health and also their BMI (Morris, 2007). There may be measurement error: if BMI is mis-measured (e.g., if it is based on self-reported height and weight), and if the level of

mis-measurement is associated with HRQL, the obesity coefficients may be biased (note that in our data height and weight were obtained during the interviewer visit and are not self-reported so the likelihood of this problem is reduced). Related to this, there may be reporting bias with respect to HRQL: EQ-5D, like any HRQL measure, is essentially subjective. It may therefore suffer from systematic reporting bias that is correlated with obesity, SES and/or with the longstanding illness. In addition, the HSE data are cross-sectional. We therefore model the impact of current obesity on current HRQL, when obesity in previous time periods might be more significant. Furthermore, our measure of obesity is BMI, which has been criticised, e.g., because it does not incorporate body fat, which is an independent predictor of ill health (Burkhauser & Cawley, 2008). Also, we measure HRQL in a single measure (EQ-5D); previous studies have provided results using both this and alternative HRQL measures, such as the SF-6D (Sach et al., 2007) and the SF-12 (Jia & Lubetkin, 2005). Previous studies investigating the impact of obesity on HRQL in a population with diabetes (Lee et al., (2005); and Redekop et al., 2002) also suffer from the limitations mentioned above.

To summarise, we illustrate that the association between obesity and HRQL is more negative in individuals with obesity-related comorbidities than in individuals without these. Furthermore, we illustrate that obesity-related comorbidities are more negatively associated with HRQL in the lower SES groups. Hence, the reason for the SES variation identified in Chapter 3 is partly due to SES variation in the association between obesity-related comorbidities and HRQL.

CHAPTER 5

Socioeconomic variation in the relationship between obesity and life expectancy

5.1 Introduction

Life expectancy has improved year-on-year over the past decade; however the socioeconomic gap continues to widen (Department of Health, 2008). As mentioned earlier, obesity levels have been rising over the last decade and there is compelling evidence showing a socioeconomic gradient in obesity. Numerous theories have tried to explain how different factors contribute to the widening socioeconomic life expectancy gap, and lifestyle related risk factors like smoking and obesity are frequently identified as contributors (see, e.g., Marmot, 2010).

In chapter 3 we found that the association between obesity and HRQL is more negative in lower SES individuals compared with higher SES individuals and suggested that the relationship between obesity (O_t) and HRQL should be modelled by including an interaction between obesity and SES:

$$H_t = b_0 + b_1 O_t + b_2 SES_t + b_3 O * SES_t + b_4 X_t + e_t \quad (5.1)$$

The aims of this chapter are to investigate whether or not we find similar SES variation in the relationship between obesity and life expectancy. We start with a review of the literature investigating the relationship between obesity and life expectancy. We then undertake an analysis using data from an individual level health survey and its longitudinal follow up, which includes height and weight collected at nurse visit, plus a set of individual and household characteristics that allows us to control for confounding factors that affect the relationship between obesity and life expectancy.

5.2 Literature review: the association between obesity and mortality/life expectancy

The objective of the review was to identify studies that investigate the association between obesity and mortality/life expectancy in order to identify the appropriate econometric specification to use. In the following analysis we investigate the impact of obesity over a lifetime. We want the studies identified in this review to inform our analysis. Hence, we choose to focus on papers using duration/survival analysis methods that take lifetime into consideration. Thus we do not look at studies who investigate outcomes like "the probability of premature mortality" or "10 year mortality risk" using binary choice models.

There are two main estimation techniques depending on the underlying survival time variable; continuous or discrete time models. In our data time is measured as a continuous variable, hence these models will be in focus in this review. Within continuous time models we have both; semi-parametric and parametric models. The advantage of the semi-parametric model is that it is less restricted by the underlying

parametric distribution. However, the parametric distributions allow one to predict median life expectancy and control for unobservable heterogeneity; this will be discussed in greater detail later.

5.2.1 Search strategy

We use similar methods to Chapter 2 and search the literature in three stages: First, we conduct a search in PubMed using controlled vocabulary:

"Obesity"[Mesh] OR "Body Mass Index"[Mesh] AND "Mortality"[Mesh] OR "Life Expectancy"[Mesh] AND "Survival Analysis"[Mesh]

The search gave 1096 hits. We screened the title and abstract of these to identify relevant papers. We ended up with 9 papers. The second stage was a reference search of the identified papers in the first stage. In the third stage we used the “cited by” option in Google scholar to identify papers that had referenced the papers identified in the first two stages. Based on the second and third stage we found 9 new papers and we ended up with a total of 18 papers.

From each paper we extracted the publication year, country, baseline year and length of follow up, methods and results, see Table 5.1.

Table 5.1: Literature review of studies on obesity and life expectancy

Authors, country, year	Participants and age at baseline	Baseline and follow up years	Statistical methods	Results
Abell et al., USA, 2008	n=27691, not given	Values collected between 1960 and 1980; follow up 15-30 years	Cox model; BMI as categorical variable; control for age, smoking status, hypertension, cholesterol, and diabetes; separate models by gender and race	Find that obesity is an independent predictor of cardiovascular disease mortality in white people.
Adams et al., USA, 2006	n=567265, age 50-71	1995; follow up 10 years	Cox model; include BMI as a categorical variable; adjusted for race or ethnic group, level of education, smoking status, physical activity, and alcohol consumption; separate by race, age, smoking, chronic disease, and duration of follow-up	The risk of death among obese (BMI> 30) subjects was two to at least three times that of participants with a BMI of 23.5 to 24.9.
Allison et al., USA, 1999	n varied by dataset, age 18+	Use six datasets. The dataset with the longest follow up was from 1971 and had a follow up of 21 years	Cox model; BMI as categorical variable; controlling for age, gender and smoking	Apply six data sources and find hazard rates between 1.33 – 1.60 of a BMI between 30-35 compared to a BMI between 23-25.
Al Snih et al., USA, 2007	n=12725, age 65+	1982; follow up for 7 years.	Cox model: BMI as categorical variable. control for age, sex, marital status, smoking status, years of formal education, and selected obesity related medical conditions	Looked at individuals 65 and older. Found higher mortality in individuals who has had a BMI of 35 and above.
Batty et al., England, 2006	N=18403, age 40-64, male only, only civil servants	1967-1970: follow up maximum 35 years	Cox model: BMI as categorical variable. Control for age, employment grade, physical activity, smoking habit, marital status, disease at entry, weight loss in the previous year, BP lowering medication, height adjusted FEV1, systolic BP, diastolic BP, plasma cholesterol, blood glucose (in normoglycaemic patients), glucose intolerance, and diabetes status.	Find a significant negative impact of obesity on mortality.
Bender, Trautner, Spraul, & Berger, Germany , 1998	N=6193, age 18-75, include only individuals with a BMI over 25	1961; follow up 33 years	Cox model; BMI as categorical variable; control for age, systolic blood pressure, cholesterol, diabetes, glucose tolerance and smoking; separate by gender	Having a BMI of 36-40 is associated with a hazard ratio of 1.31 compared to having a BMI of 25-32
Calle et al., USA, 1999	N=1,184,657, age 30+	1982; follow up 14 years	Cox model; include BMI as a categorical variable; separated population by smoking status and disease history; control for age, education, physical activity, alcohol use, marital status, aspirin use, fat consumption, vegetable consumption, and (in women) use of estrogen-replacement therapy	Find that the hazard rate of a BMI of 30-31.9 is 1.62 compared to a BMI of 23.5-24.9 in non smoking men aged 30-64
Czernichow et al., UK, 2011	N varied between the surveys, age not given	Collected between 1994 and 2003; mean follow up of 8.1 years	Cox model; BMI as a continuous variable; control for: age and sex, followed by socioeconomic status, cigarette smoking, systolic blood pressure, total blood cholesterol, diabetes status, physical activity and history of CVD	Obesity increase mortality risk using Body Mass Index, waist circumference, and waist to hip ratio as measure of obesity.
Flegal, Graubard, Williamson, & Gail, USA, 2005	N=36859, Age 25+	1971; follow up 21 years	Cox model; BMI as categorical variable; controlled for sex, smoking status, race, and alcohol use; separated by age	Calculate excess deaths from overweight and obesity combined (BMI >25) to be 25 814 in year 2000
Fontaine et al., USA, (2003)	N varied by the survey, age18-85	1971-75; follow up 17-21 years	Cox model; include BMI as a quadratic term; separate by race and gender; control for age and smoking	20 year old white male with a BMI of more than 45 is estimated to lose 13 years of life due to obesity.
Freedman, Ron, Ballard-	N=83744, age 22+	1983-1989; follow up 11-17 years	Cox model; BMI as categorical variable; control for education, race, smoking	Calculate that the hazard rate of a BMI of 30-34.9 is

Barbash, Doody, & Linet, USA, 2006			and alcohol use; separate by gender, race and age	1.47 for a man above the age of 55 compared to a man with a BMI of 21-22.9
Lawlor et al., Scotland, 2006	Men: n=11033 women: n=8327, age 45-64	1972-1976; follow up 28 to 34 years	Cox model: BMI as categorical variable. Control for age and smoking	Obesity and overweight is associated with all cause mortality
Mayhew, Richardson, & Rickayzen, UK 2009	N=7414, age 18+	1984-85; follow up 21 years	Cox model; include BMI and waist to height as a quadratic term; separate by gender and smoking status; control for age	A non-smoking 30-year old man with a BMI of 34 is expected to live 4 year less than a similar man with a BMI of 24. For women the equivalent loss is 2 years.
Seidell, Verschuren, van Leer, & Kromhout, Netherlands, 1996	N=48287, age 30-54	1974-1980; follow up 12 years	Cox model; BMI as a categorical variable; control for smoking, hypertension, hypercholesterolemia, and diabetes mellitus	Obesity (BMI>30) had a hazard rate of 1.5 compared to normal weight (18.5<BMI<4.9)
Stevens et al., USA, 1998	Men: n=62116 women: n=262019, age 30+	1959; follow up 12 years	Cox model; include BMI both as a continuous and categorical variable; exclude individuals with a smoking history; separate by age; control for age, education, physical activity and alcohol use	The hazard ratio associated with an increment of 1 in BMI was 1.1 for men aged 30-44. For women the corresponding estimate was 1.08
Tsai et al., USA, 2006	N=7139, age not given, only workers at an oil plant	1988; follow up 20 years	Cox model: BMI as categorical variable. Control for age, sex, and smoking status, cholesterol, hypertension, and fasting blood glucose	Obesity was associated with increased mortality
Vapattanawong et al., Thailand, 2010	N=15997, age 60+	2004; mean follow up:3.8 years	Cox model: BMI as categorical variable. control for marital status, urban/rural, education, living arrangement, diabetes, hypertension, smoking, and physical activity	Focused on individuals aged 60 and above and did not find a negative impact of obesity on survival.
Yan et al., USA, 2006	N=17643, age 31-64	Values collected between 1967 and 1973; mean follow up was 32 years	Cox model; BMI as categorical variable; control for age, sex, education, minor electrocardiographic abnormality, systolic blood pressure, and serum total cholesterol level. Stratified by smoking, blood pressure and cholesterol level	For individuals with no cardiovascular risk factors as well as for those with 1 or more risk factors, those who are obese in middle age have a higher risk of mortality from CHD, cardiovascular disease, and diabetes in older age than those who are normal weight

Note

Could not obtain the full version of the Seidell, Verschuren, van Leer, & Kromhout (1996) paper. The description is based on information in the abstract

5.2.2 Findings

We identified 18 studies and most of them support a negative association between obesity and life expectancy. They all use the Cox proportional hazard model and frequently test for proportional hazard of the covariates and fit models that do not violate the proportional hazard assumption. However, we are not aware of any studies that have applied other alternative parametric survival distributions.

In these studies there is a trade-off between how recent the data is (date for obesity measurement and other measures) and the length of the mortality follow-up. The longest follow up we have identified is a mean of 33 years in a German study by Bender, Trautner, Spraul, & Berger (1998) using data from 1961. None of the identified studies had repeated measures of BMI.

Smoking receives a great deal of attention in these studies as it has been showed to modify the association between BMI and mortality (Stevens, Juhaeri, Cai, Thun, 2000). The reason being that people who smoke have a tendency to have a lower BMI but also a higher mortality; hence it reduces the association between obesity and mortality. This is controlled for either by stratifying by smoking status or controlling for it as a covariate in the model.

Age has also received attention in some of the studies as two American studies found it to have an impact on the relationship between BMI and mortality. The relationship between BMI and mortality is characterised by a U or a J shape, however it has been found to flatten out as age increase (Freedman et al., 2006). BMI seems to have a protective effect in the oldest individuals (aged 75 and older) (Stevens at al., 1998).

Survival models, which look at BMI, have issues with omitted variables. This is other explanatory variables that have not been included in the model that are correlated with the dependent variable (life expectancy) and the independent variables (BMI groups). Although we would expect obesity to reduce life expectancy it is possible that

it is actually health variables that influence both life expectancy and BMI. One way to control for this is to control for longstanding illnesses[15]. However, remember that when we do this we reduce the risk factor effect of obesity[16]. Six studies have controlled for illness variables and two stratify by this, although with conflicting results. However, the studies that control for illness variables have only controlled for illness variables that are related to obesity. For example, Abell et al. (2008) control for hypertension, cholesterol and diabetes. Hence, they do not sufficiently deal with the omitted variable issue mentioned above. There is a need for a study that investigates the consequences of controlling for all illnesses, not only those related to obesity.

Although research from other countries is useful when assessing the impact of excess body fat on the UK population, research based on UK specific data would be more relevant when answering UK specific questions (Mayhew, Richardson, & Rickayzen, 2009). The majority of research is based on US data; however we have identified four studies based in the UK. Three studies using UK data have calculated hazard ratios due to obesity based on Cox models. Czernichow et al. (2011) conducted a study to investigate the association between three measures of body fat (BMI, waist circumference and waist-hip ratio) and CVD mortality risk using Health Survey for England data linked to mortality records. They find that BMI is a predictor of mortality controlling for age and sex, socioeconomic status, cigarette smoking, systolic blood pressure, total blood cholesterol, diabetes status, physical activity and history of CVD. Little difference was found between the measures of body fat (BMI, waist

[15] It has been argued that controlling for comorbidities is an overadjustment as these factors are part of the pathway between obesity and the output variable (See Zizza, 2004).

[16] Although we do not completely remove the risk factor effect as individuals can get longstanding illnesses after the date of BMI measurement.

circumference and waist-hip ratio). The mean age at baseline ranged from 49-57 years, but there is a relatively short follow up of only 8.1 years. Lawlor et al. (2006) investigate the impact of obesity controlling for age and smoking using data from two Scottish towns. They account for heterogeneity bias by omitting individuals who die within two years after the physical measurement. They find that obese individuals have a significantly higher hazard rate for death than normal weight individuals. Batty et al. (2006) use the Whitehall study and investigate the hazard rate for death in obese individuals controlling for age, employment grade, physical activity, smoking habit, marital status, weight loss in the previous year, physical and illness variables related to obesity. These three studies all use the Cox proportional hazard model. The authors do not provide information about whether or not they have tested the proportional hazard assumption. If this assumption is violated their estimates may be invalid. Hence, although it is an advantage that these studies control for a number of covariates in terms of heterogeneity, the more covariates the more likely they are to violate the proportional hazard assumption. Hence, we will test for proportional hazard in our analysis. They also report their results as hazard ratios, but do not show the consequences of obesity for life expectancy in terms of life years lost. Furthermore, both the study by Lawlor et al. (2006) and Batty et al. (2006) use does not use data that is representable of the English population. Although Czernichow et al. (2011) use a the HSE the short follow up makes it problematic to estimate the consequences of obesity in the long run, in addition, the results are extra vulnerable to heterogeneity bias.

Mayhew, Richardson, & Rickayzen (2009) investigate the effect of obesity, measured by BMI and waist to height ratio (WTH), on life expectancy controlling for age and stratifying by smoking status and conclude that higher BMI and WTH reduce life expectancy. They test for proportional hazard and allow for interactions between BMI/WTH and age. They do not find an interaction between age and BMI; however they do find an interaction between age and WTH. To generate predictions of life expectancy they calculate age and BMI specific hazard ratios based on the Cox model. They then get the probability of death in each year of adult life from life tables. Then multiply the BMI specific hazard ratio with the probability of death in each year. To control for smoking the analysis was conducted on two datasets: non-smokers and all participants (smokers and non-smokers together). This study does not control for any other covariates like SES, ethnicity, marital status and comorbidities/longstanding illnesses, which can cause heterogeneity. For example, we know that obesity is related to socioeconomic status. Furthermore, longstanding illnesses might affect both obesity and life expectancy. Hence, we will control for these in our study. In addition, rather than stratifying by smoking status we will control for it in the analysis.

Relatively little interest has been given to socioeconomic variables. Six studies have controlled for SES variables. Though, we have not found any studies that focus specifically on SES and analyse whether or not the impact of obesity varies by SES groups.

To conclude, there is evidence of a relationship between obesity and mortality/life expectancy, in that obesity increases mortality and reduces life expectancy. Based on

the identified literature we argue that there are still issues yet to be resolved. None of the studies has controlled for all comorbidities, only comorbidities related to obesity. We will in the following look at the consequences of controlling for comorbidities. Except from the Cox model, no other alternative parametric distribution has been tried out so none have predicted life expectancy based on parametric distributions. None of the studies consider the influence of unobservable heterogeneity related to duration dependence on estimated mortality, which might produce incorrect results (Jones, 2005). There is uncertainty regarding the effect of age, only one study has looked at this in England and the findings were inconsistent. In addition, there has been relatively little focus on how socioeconomic status influences this analysis and we have not been able to identify studies that investigate whether or not the impact of obesity on life expectancy varies by SES. We will account for these issues in the following chapter.

5.3 Methods

5.3.1 Data

The analysis is based on data from the first wave of the *Health and Lifestyle Survey* (HALS1)[17] (1984/1985; *n*=9003) and the longitudinal follow up in June 2009. HALS1 surveys a representative sample of the population of Great Britain aged 18 years and

[17] There is also a follow up survey HALS2 1991/2. The benefit of using this second round would be that we could get repeated BMI measurement. However, this has not been used in this analysis due a relatively high dropout rate (41%) (n=5352). We analyse our data by BMI and SES groups and need a certain amount of individuals in each group for our analysis to be meaningful. This is even more important in survival analysis as only part of the individuals in the sample have died. In addition, with a high dropout rate there are problems with attrition bias.

over. Data for HALS1 were collected between autumn 1984 and summer 1985, in two home visits, the second by a research nurse during which a range of physiological measurements (including height and weight) were taken. In the longitudinal follow up in June 2009 the original participants in HALS1 were flagged on the NHS Central Register (NHSCR) (now part of "The NHS Information Centre for Health and Social Care"), which is a computerised record of everyone registered with a GP in Great Britain that also collects demographic information, including mortality (*http://www.ic.nhs.uk/*). The linkage between HALS1 and NHSCR meant that it was possible to assess whether or not participants in HALS1 had died by June 2009 and if so, their date of death. Ninety eight percent of participants in HALS1 were identified in the NHSCR.

In our analysis we excluded pregnant women at the date when BMI was measured, because their BMI is temporarily increased, and individuals with missing BMI data; this reduced the sample from 8803 to 7,289. We also excluded individuals under the age of 40 years at the time of HALS1. This was for two reasons. First, by 2009 low mortality rates were observed among those below 40 years of age, which meant that a high proportion of the survival data were censored in this group. Second, there is some uncertainty about the SES of those under 40 years of age; our SES measures are based on educational, occupation and social class, which may change over time, especially among younger groups. We are unable to account for such changes in our analysis, which means that the observed impact of SES on survival could be biased. A similar approach has been taken in previous analyses of HALS data (Balia & Jones, 2008). This reduced the sample from 7,289 to 4,062.

5.3.2 Variables and data handling

Survival time is measured as the time to death or censoring in months from the date at which height and weight were measured in HALS1. Thus, the 'entry date' is the date of BMI measurement in HALS1 and the 'exit date' is either date of death or date last recorded alive (which is June 2009) if the respondent did not die during the follow-up period. Survival time is right censored because actual survival time is not observed for those who died beyond the censoring point. It is also left truncated because it is observed only for those individuals who survived up to the HALS1 interview date.

Our obesity measure is based on body mass index (BMI), measured as weight in kilograms divided by height in metres squared (kg/m^2). BMI is computed from the height and weight measurement obtained during the nurse visit; it is not based on self-reported height and weight, which means that the likelihood of systematic measurement error is reduced. Obesity is measured as a categorical variable based on three BMI categories: normal weight 18.5-25 kg/m^2; overweight 25-29.9 kg/m^2; and, obese $\geq$30 kg/m^2. Our data also has a small number of respondents in the underweight category. As earlier chapters we include this as a separate category, but do not report the results.

To generate our measure of SES we follow the approach in earlier chapters and use a composite measure. The main reason being that it is designed to capture multiple dimensions of SES (individual and household variables) rather than focusing on a single aspect (e.g., income). We use interval regression methods (because the income

variable is banded) to construct the SES measure and regress total weekly household income reported in 12 income bands (including an open-ended top category) against a set of socioeconomic variables. In this interval regression we do not exclude individuals under the age of 40. I.e. we include all individuals over the age of 18 in this prediction and control for age. This is done to include the maximum amount of individuals in the model. The socioeconomic variables used as covariates in the interval regression are: education qualifications (measured in six categories); social class of household reference person (HRP; seven categories); individual economic activity status (seven categories); owning a house (yes/no); bedrooms in household (four categories); age (as a quadratic function); and, gender interacted with age. We calculate predicted values from this model. In case of the income variable in HALS there is a high non-response rate, so to maximise the sample size, we included individuals in the duration models with missing income data for whom we could compute a predicted value. In the duration models we included an indicator for missing income (yes/no), to control for the possibility that income may not be missing at random. By doing this we compute SES values for 1313 (over the age of 40) individuals with missing income values. The predicted SES measure is then divided into quartiles for use in the duration models.

In the duration models we also control for a range of individual and household characteristics. Age was included as a quadratic function (based on likelihood ratio tests for functional form). We control for smoking status (three categories) as it has been showed to modify the association between BMI and mortality (Stevens, Juhaeri, Cai, Thun, 2000). In addition we control for marital status (five categories), geographical area (ten categories) and ethnicity (three categories) to reduce the risk of

heterogeneity bias. We also run a model controlling for number (five categories) and whether or not the longstanding illness limits activities in some way (yes/no).

5.3.3 Econometric methods

We start by describing the theory behind the econometric methods applied before we describe how they will be used in this analysis.

5.3.3.1 Parametric vs. semi-parametric models

The basic concepts of duration analysis are the survival function and the hazard rate. The survival function (*S(t)*) is the probability that someone survives to at least time t. The probability of survival is equal to 1 at the entry in the state of interest. In this analysis we follow the individuals from the time of height and weight measure (sometime between 1984-1985) until June 2009. At this point those who are not dead are censored. The hazard rate is the instantaneous rate of failure per unit of time, so it shows the probability that someone fails at time t, given that they have survived up to that point. It can be written as:

$$h(t) = \frac{f(t)}{S(t)} \tag{5.2}$$

Where *f(t)* is the probability density function for the probability of failing at time t. The density function will capture uncensored observations, where we observe their actual time of failure, and the survival function is used for censored observations where we only know that they have survived to at least time t.

We follow our participants for up to 300 months (1984/85 - 2009). We consider these intervals (in months) to be sufficiently small so that it is reasonable to treat the observed times occurring on a continuous time scale; hence we apply continuous time models (Masyn, 2003). We focus on two approaches to duration analysis; *parametric* and *semi-parametric* approaches. Both of these can be expressed as:

$$h(t_i) = h_0(t) * \exp(x_i \beta) \qquad (5.3)$$

The dependent variable is the hazard function at a given time, for individual *i*. The first term on the right hand side ($h_0(t)$) is the baseline hazard and corresponds to the probability of dying when all explanatory variables (x_i) are equal to zero. The shape of the baseline hazard allows us to estimate how the hazard function changes with time. In the *semi-parametric* Cox model this term is left unspecified and treated as an unknown function of time. If this term is specified it yields a parametric model. A parametric model assumes particular functional forms for the survival function *S(t)* and the probability density function *f(t)* and then also the hazard function *h(t)*. For a further description of each functional form for *S(t)* and *f(t)* available in Stata 11 see StataCorp (2009), table 1, p. 358. The second part of function 5.3 above ($\exp(x_j \beta)$) allows us to estimate the proportional change that can be expected in the baseline hazard related to changes in the explanatory variables (the *x's*).

The *semi-parametric* Cox proportional hazard[18] model is more flexible and robust than the *parametric* models as it does not specify the baseline hazard but treats it as an unknown function of time. However, "a parametric version, if adequate, would lead to more efficient estimates of survival probabilities and contribute to a better understanding of the survival/duration phenomenon under study" (Hjort, 1992). A further advantage of the parametric models is that they allow controlling for unobservable heterogeneity by including a frailty distribution, which we explain in more detail below. This cannot be done in Cox models without repeated measurements for each individual[19]. In addition, parametric models allow one to predict median life expectancy based on the survival function.

Above we explained how we could calculate the impact of the coefficients ($\exp(x_j\beta)$) on the baseline hazard, which is the instantaneous rate of failure per unit of time. This could also be seen as a multiplicative risk of dying in each time period and we refer to it as the relative mortality. The relative mortality between obese and normal weight is: $\frac{Number\ of\ deaths\ in\ the\ obese\ at\ time\ t}{Number\ of\ deaths\ in\ the\ normal\ weigth\ at\ time\ t}$. Since the hazard ratio is relative similar hazard ratios can have unequal consequences in additive terms across groups depending on the baseline mortality in that group.

[18] The Cox model assumes that time is continuous and there are no tied survival failures. Tied failures are failures that fall within the same time period but that did not happen exactly at the same point in time. These do occur and the Cox model needs a method to handle these. We use the Breslow method for tied failures, which assumes that failures in the same period did actually happen in the same period. It has been shown to work well if the number of events is small relative to the sample size.

[19] The stcox command for the Cox proportional hazard model includes an option for estimating models with shared frailty, assuming a Gamma mixture. However, in this data that we use, there is only a single record per person, and so the command is not applicable (Jenkins, 2008).

While the relative mortality risk can give an indication about relative severity of obesity it does not quantify the implications in terms of magnitude of difference (such as how much sooner an obese person will die). Hence, we also predict life expectancy, which can be considered more useful in terms of information. We describe this in more detail later.

5.3.3.2 Unobservable heterogeneity

Unobservable heterogeneity has received some attention in the literature on duration analysis (Jones, 2005). The potential problem can be explained by a scenario where the population is made up of two subpopulations: 'frail' individuals' with a high hazard rate and 'strong individuals' with a low hazard rate. Whether or not one is considered to be in the frailty group is decided by unobserved variables. Over time the frailer individuals die first and the proportion of frail individuals in the population will decrease and the overall hazard will decrease.

This can bias the results in two ways. First, if the frail individuals are equally distributed across the sample it can decrease all hazard ratios of each explanatory variable in the model (decreasing duration dependence). Second, if frailty is a confounder for some variables it could lead to false hazard ratios for these estimates (Masyn, 2003). For example, terminally ill individuals might lose weight and die early, which would underestimate the impact of obesity on life expectancy.

To account for unobservable heterogeneity (frailty) we add an extra error term (v) to the parametric models as shown below.

$$h(t_i) = h_0(t) * \exp(x_i\beta) * v = h_0(t) * \exp(x_i\beta + u) \quad (5.4)$$

Where $h_0(t)$ is the baseline hazard function and the 'error' term $u = \ln(v)$ is a random variable with a mean of zero. The random variable *v*, or equivalently *u*, can be interpreted as the impact of 'omitted variables' on the hazard rate. We must assume a distribution for the added error term (frailty distribution) and the Gamma distribution and Inverse Gaussian distribution have been most commonly used (Jenkins, 2008). We follow Forster & Jones (2001) and assume a gamma distribution for this error term (gamma frailty).

5.3.3.3 The proportional hazard assumption

A main assumption in a proportional hazard model is proportionality, which means that the hazard rate is independent of time t. In the Cox model we use Schoenfeld and scaled Schoenfeld residuals to test for proportional hazard, both as a whole and for each individual predictor. The null hypothesis is, effectively, that the log hazard ratio is constant over time; a rejection of the null hypothesis indicates that the proportional hazards assumption is violated.

Note that proportional hazard does not mean that the BMI of individuals does not vary across time. It measures the average hazard rate for the individuals depending on their BMI category at the time of the survey. Hence, most individuals change their BMI status during follow up. So our model measures the impact of being obese at one point

in time and then it takes into account mean variations from this obesity status across all individuals in the sample.

5.3.3.4 Predicting median life expectancy

As mentioned above we aim to calculate *life expectancy* and see how this varies by BMI and SES. Life expectancy for an individual is a result of the hazard rate (probability of dying in each time period) for an individual across time. Hence, life expectancy is a useful and interpretable index. We focus on median life expectancy which is the time beyond which 50% of the subjects are expected to survive.

Estimating median life expectancy consists of two stages. First, we plug in the mean covariate values and the parameter estimates into the survival function *S(t)*, we denote this $\hat{S}(t)$. The second stage is to set $\hat{S}(t)$=1/2 to get the predicted value when 50% of the subjects are alive.

$$Median\_t = \{t : \hat{S}(t) = 1/2\} \quad (5.5)$$

Hence, to predict the median survival time in this way the survival function must be specified. As mentioned earlier the *semi-parametric* Cox model leaves the baseline function unspecified so we do not have this option in *semi-parametric* models.

5.3.3.5 Choosing the parametric distribution

There are several parametric distributions and we want to choose the parametric distribution that best fits the data. We test the following distributions of the survival function *S(t)* available in Stata 11: Weibull, generalised gamma, Gompertz, exponential, lognormal and loglogistic. We decide upon the parametric specification that best fit our data in two stages. First, we use plots of cumulative Cox-Snell residuals for the observed failures in the sample to assess the general fit of the model for those who fail. This compares the predictions of the hazard accumulated over time with the actual failures over time. A correctly fitted model should yield a straight line with a slope of 1. We pick out the distribution based on visual examination of the residuals. Hence, we might end up with more than one distribution and use the second stage to choose between these. In the second stage we compute Akaike's Information Criterion (AIC). The AIC is a measure that rewards general goodness of fit of a model, which penalises overfitting. The preferred model is the one with the lowest AIC value (Forster & Jones, 2001).

5.3.3.6 Estimation strategy

We start by running all the diagnostic tests to identify the parametric distribution that best fits our data. To determine whether the results are sensitive to the parametric assumptions we follow an approach suggested by Newey, Powell & Walker (1990) and reanalyse the data using the semi-parametric Cox model. Relating to the Cox model we test for proportional hazard.

We regress the probability that someone fails (dies) at time t against BMI categories, SES quartiles, age and other control variables (*Z*) using both semi-parametric and parametric methods:

$$h(t_i) = h_0(t) * \exp(\beta_1 overweight_i + \beta_2 obese_i + \beta_3 SES2_i + \beta_4 SES3_i + \beta_5 SES4_i + \beta_6 age_i + \beta_7 agesquared_i + \beta_8 Z_i + u_i) \quad (5.6)$$

Where the omitted categories are normal weight and SES groups 1 (most deprived).

We test for unobservable heterogeneity (frailty) by a likelihood ratio test of the null hypothesis that the variance in the heterogeneity parameter (u_i) is zero. A significant test means that the frailty term accounts for unobservable heterogeneity in the model.

The aim in this chapter was not only to investigate the impact of obesity on mortality/life expectancy, but also to investigate if this impact varies by SES. Furthermore, we want to investigate how age influences this. One can investigate if the impact of obesity on relative *mortality* varies by SES and age by investigating if there are interactions between these. We use the likelihood ratio test to compare larger models (with interactions) to more sparsely fitted models. We fit three models with interactions.

First, a model interacting BMI groups with SES groups:

$$h(t_i) = h_0(t) * \exp(\beta_1 overweight_i + \beta_2 obese_i + \beta_3 SES2_i + \beta_4 SES3_i + \beta_5 SES4_i + \beta_6 overweight * SES2_i + \ldots + \beta_{11} obese * SES4_i + \beta_{12} age_i + \beta_{13} agesquared_i + \beta_{14} Z_i + u_i) \quad (5.7)$$

Second, a model interacting BMI groups with age[20]:

$$h(t_i) = h_0(t) * \exp(\beta_1 overweight_i + \beta_2 obese_i + \beta_3 SES2_i + \beta_4 SES3_i + \beta_5 SES4_i + \beta_6 age_i + \beta_7 agesquared_i + \beta_8 overweight * age_i + \beta_9 obese * age_i + \beta_{10} Z_i + u_i) \quad (5.8)$$

Third, we fit a model interacting BMI with both age and SES:

$$h(t_i) = h_0(t) * \exp(\beta_1 overweight_i + \beta_2 obese_i + \beta_3 SES2_i + \beta_4 SES3_i + \beta_5 SES4_i + \beta_6 overweight * SES2_i + \ldots + \beta_{11} obese * SES4_i + \beta_{12} age_i + \beta_{13} agesquared_i + \beta_{14} overweight * age_i + \beta_{15} obese * age_i + \beta_{16} Z_i + u_i) \quad (5.9)$$

As mentioned life expectancy is a prediction based on the underlying mortality model. We base the predictions of life expectancy on the models with the best fit based on the likelihood ratio tests.

Care is required when interpreting the interactions in equation 5.7, 5.8 and 5.9. An interaction in a multiplicative survival model is neither sufficient nor necessary to meaningfully decide whether or not there are interactions in the predicted life years. Hence, we do not use the significance in the multiplicative interaction terms to decide whether or not there are interactions in the association between obesity and *life expectancy*, rather we look at the implications of the predicted values (Ai & Norton, 2003; Greene, 2010; and Berry, DeMerit & Esarey, 2010).

[20] We interact BMI groups with age as a linear variable based on likelihood ratio tests.

We calculate predicted median survival in each BMI group across the SES groups for men and women setting the age to 65, 75 and 85 and the covariates to their mean value based on the preferred parametric model. Hence, the variation in predicted median life expectancy is a function of the impact of obesity on life expectancy and how this varies by SES, and is not affected by the values of the other covariates.

We run each model twice using first semi-parametric and then parametric methods. We run separate models for men and women. And separate models with and without controlling for longstanding illnesses. Hence in total we run 8 models for each gender as shown in Table 5.2.

Table 5.2: summary of models

Model	BMI, SES and age	Covariates	Distribution
1	Separate dummy variables	Household and individual characteristics*	Cox proportional hazard
2	Separate dummy variables	Household and individual characteristics*	Parametric distribution
3	Interacted models	Household and individual characteristics*	Cox proportional hazard
4	Interacted models	Household and individual characteristics*	Parametric distribution
5	Separate dummy variables	Household, individual characteristics* and longstanding illness variables**	Cox proportional hazard
6	Separate dummy variables	Household, individual characteristics* and longstanding illness variables**	Parametric distribution
7	Interacted models	Household, individual characteristics* and longstanding illness variables**	Cox proportional hazard
8	Interacted models	Household, individual characteristics* and longstanding illness variables**	Parametric distribution

* Smoking status, marital status, geographical area and ethnicity
**Number of longstanding illnesses and whether or not the longstanding illness limits activities in some way

P values for the hazard ratios below the 5% level (z scores higher than ≈1.9) are regarded as statistically significant. Values between 5% and 10% (z scores between ≈1.6 and ≈1.9) are regarded as weakly significant.

5.4 Results

After exclusions, our estimation sample consisted of 1,832 men and 2,230 women of whom 1,052 (57%) and 1,023 (46%) respectively, were reported as deceased at the censoring point (Table 5.3). Comparable figures by BMI category for men were 463 of 817 (57%) in the normal weight group, 445 of 807 (55%) in the overweight group, and 107 of 167 (64%) in the obese group. For women these figures were 471 of, 1,150

(41%) in the normal weight group, 335 of 687 (49%) in the overweight group, and 172 of 311 (55%) in the obese group. A higher percentage was reported deceased in lower SES quartiles compared with higher SES quartiles across each BMI group for both men and women.

Table 5.3: Summary statistics of the individuals used in the duration model

	Whole sample		Normal weight		Overweight		Obese	
	Total N	**Failures (%)**	**Total N**	**Failures (%)**	**Total N**	**Failures (%)**	**Total N**	**Failures (%)**
Men								
N	1832	1052 (57)	817	463 (57)	807	445 (55)	167	107 (64)
SES								
1 (most deprived)	395	345 (87)	186	162 (87)	152	132 (87)	39	33 (85)
2	469	337 (72)	204	137 (67)	205	147 (72)	44	37 (84)
3	459	223 (49)	181	89 (49)	222	105 (47)	53	28 (53)
4 (least deprived)	509	147 (29)	246	75 (30)	228	61 (27)	31	9 (29)
Women								
N	2230	1023 (46)	1150	471 (41)	687	335 (49)	311	172 (55)
SES								
1 (most deprived)	619	500 (81)	275	213 (77)	210	171 (81)	104	91 (88)
2	546	280 (51)	260	129 (50)	187	90 (48)	84	51 (61)
3	555	161 (29)	304	78 (26)	159	56 (35)	78	22 (28)
4 (least deprived)	510	82 (16)	311	51 (16)	131	18 (14)	45	8 (18)

Source: HALS1 1984/1985 and the longitudinal follow up in June 2009

Note

These figures do not control for age at entry

Selected descriptive statistics are in Table 5.4, while full summary statistics on each of covariates in our model across each BMI group can be found in Appendix 5.1. Forty eight percent of participants were normal weight at measurement, thirty seven percent were overweight and twelve percent were obese. The mean age at measurement was 58, the mean time to censoring (at date of measurement) was 291 months and the mean unadjusted time to death was 149 months. The highest predicted income was in the normal weight category and the percentage obese was highest in SES quartile 1 (most deprived).

Table 5.4: Selected descriptive statistics by BMI category

	Whole sample	Normal weight	Overweight	Obese
Mean time to censoring at date of measurement (months)*	291	291	291	291
Mean time to death at date of measurement (months)**	149	150	152	152
Predicted income (Weekly)	136	141	135	123
SES groups				
1 (most deprived)	25	23	24	30
2	25	24	26	27
3	25	25	26	27
4 (least deprived)	25	28	24	16
Missing income				
Yes	17	17	17	17
No	83	83	83	83
Age	58	57	59	58
Gender				
Male	46	42	54	35
Female	54	58	46	65

Source: HALS1 1984/1985 and the longitudinal follow up in June 2009

Notes:

All statistics are either mean (for continuous variables) or % (categorical variables)

* includes only individuals who had survived up until censoring

** includes only individuals who died before censoring

Table 5.5 shows the results of the interval regression model of banded total weekly household income against the individual, household and area socioeconomic variables. Education, social class of HRP, accommodation tenure and bedrooms in household are all positively correlated with income. Compared to being in paid employment all other types of economic activity are negatively correlated with income. We used these coefficients to predict income for each respondent. Respondents were categorised into quartiles based on the predicted income values. The range of values for predicted weekly income was 3 to 299; the cut-points used to set the quartiles were 91, 132 and 178.

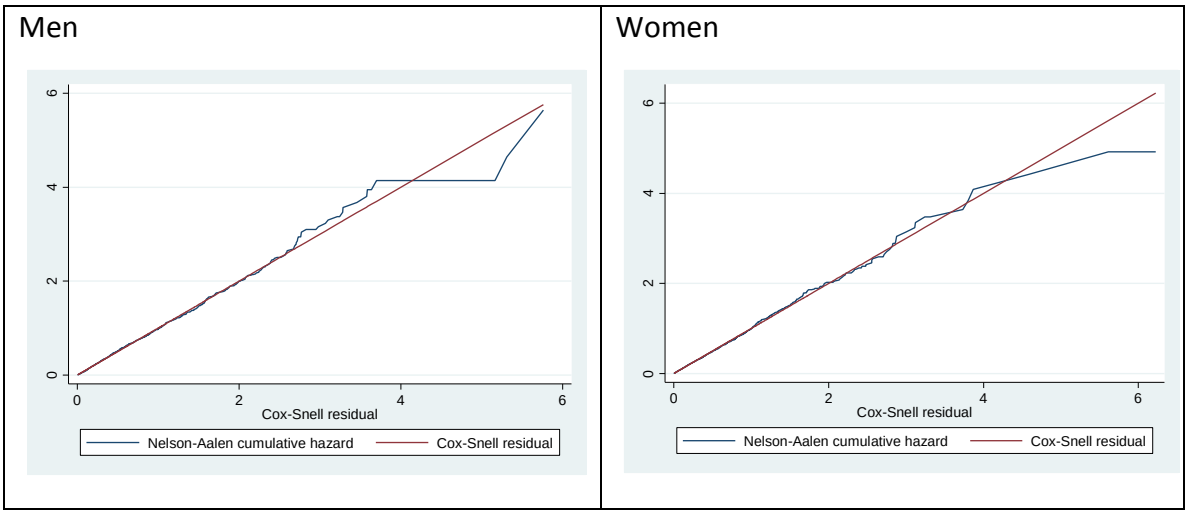
Table 5.5: Interval regression of weekly total household income against SES indicators

	Coef.	Z
Educational qualifications		
Degree or equivalent	Base category	
Higher education below degree	-14.55	-3.3
NVQ3/GCE A Level or equivalent	0.60	0.1
NVQ2/GCE O Level or equivalent	-13.11	-3.2
Other	-19.00	-3.9
No qualification	-29.13	-7.5
Social Class of HRP		
Professional	Base category	
Managerial technical	-25.13	-4.9
Skilled non-manual	-45.13	-8.0
Skilled manual	-47.91	-9.1
Semi-skilled manual	-56.97	-10.1
Unskilled manual	-56.45	-8.2
Other	-33.22	-3.2
Economic activity status for last week		
Working full time	Base category	
Working part time	-18.05	-4.6
Unemployed	-68.95	-13.5
Permanently sick or disabled	-54.11	-7.3
Retired	-48.93	-10.1
Keeping house	-39.09	-10.1
Full time student	-52.96	-4.1
Bedrooms in household		
One	Base category	
Two	7.52	1.6
Three	21.21	4.7
Four or more	70.17	13.1
Housing tenure		
Own accommodation	35.37	14.4
Rent	Base category	
Gender		
Male	55.70	3.2
Female	Base category	
Age	2.33	4.3
Age squared	-0.03	-5.2
Male*age	-2.65	-3.5
Male*age squared	0.03	3.7
Constant	147.09	10.9
Observations	5787	
Adjusted R^2	0.08	

In Appendix 5.4 and Appendix 5.5 we see the Cox-Snell residuals for men and women for each parametric distribution. These indicate that our data fits models with monotonic hazard rates for both men and women. Appendix 5.4 and Appendix 5.5 also show the AIC for men and women for each parametric distribution. These show that

the Gompertz model has the best fit for both men and women. The Cox-Snell residuals for men and women in the Gompertz model with gamma frailty are showed in Figure 5.1. We see that the fit is good as the residuals lie close to the 45 degree line. However, the fit is less good at higher values. It is common for models with censored data to have less good fit at large values of time; this is because fewer individuals have the highest hazard values due to prior failures (UCLA Academic Technology Services, 2011).

Figure 5.1: Cox-Snell residuals for men and women based on the Gompertz distribution with gamma frailty distribution



Note:
The dependent variable is survival time and the independent variables are age, SES, BMI, smoking, ethnicity, marital status, area and missing income

Table 5.6 show the output of Gompertz model accounting for unobservable heterogeneity (gamma frailty distribution) for men and women with and without interactions. We also see the results of the Schoenfeld and scaled Schoenfeld residual test for proportional hazard for both men and women for all variables. We see that we fail to reject the "global test" of proportional hazard in each of the models for both men and women. Although not shown, we also fail to reject proportional hazard for

each of the covariates (except from the dummy variable for East Midlands for men in the main effects model). Failing to reject the proportional hazard test means that we do not have any problems with our models not fulfilling the proportional hazard assumption; hence we can apply proportional hazard models. The upshot of this is that the preferred model is the Gompertz model, but we also rerun the analysis using a Cox model.

The coefficients in the tables are hazard ratios. A value larger than 1 should be interpreted as an upward shift in the baseline hazard function, hence it increases the likelihood of dying in each time period. Conversely, a value smaller than 1 should be interpreted as a downward shift in the baseline hazard function, hence it decreases the likelihood of dying in each time period.

The main effects models in Table 5.6 illustrate a significant correlation between obesity and mortality, in that obesity increases hazard rate in each time period, controlling for age and a range of other individual characteristics. This association is more pronounced in women where the hazard rate in the Gompertz model is 1.42 versus 1.35 in men. In the main effects models we also find a significant positive relationship between SES and mortality controlling for age and a range of other individual characteristics. I.e. higher SES groups have lower mortality in each time period independently of other factors. The SES effect on mortality is similar for men and women; however the impact of SES group 2 is only significant in men. The hazard rates for SES groups 2, 3 and 4 (least deprived) in men are 0.82, 0.68 and 0.54, respectively. For women similar figures are 0.9, 0.73 and 0.54, respectively. Table 5.6

also shows that the unobservable heterogeneity parameter (frailty distribution) is significant in women but not in men in the main effects models.

The likelihood ratio tests, which compare the fit of the models with interactions to the main effect model, were insignificant in men for: the BMI interacted with SES model (*P=0.31*), the BMI interacted with age model (*P=0.44*) and for the fully interacted model (*P=0.43*). Hence, the interactions do not do much to help explain variation in the dependent variable in men.

In women, the likelihood ratio tests were insignificant comparing the main effects model with the BMI interacted with SES model (*P=0.29*) and the BMI interacted with age model (*P=0.11*), but significant in the fully interacted model (*P=0.01*). Furthermore, comparisons illustrate that the fully interacted model had a significantly better fit than: the BMI interacted with SES model (*P<0.01*) and the BMI interacted with age model (*P=0.02*). Hence, the fully interacted model has the best fit in women.

Table 5.6: Gompertz models controlling for age and additional household and individual characteristics

	Main effects model				BMI interacted with SES model				BMI interacted with age model				Fully interacted model			
	Men		Women		Men		Women		Men		Women		Men		Women	
	Haz. ratio	z	Haz. ratio	z	Haz. ratio	z	Haz. ratio	z	Haz. Ratio	z	Haz. ratio	Z	Haz. ratio	z	Haz. ratio	z
BMI																
Normal weight	Base category		Base category		Base category		Base category		Base category		Base category		Base category		Base category	
Overweight	1.043	0.6	0.977	-0.3	0.993	-0.06	1.063	0.55	0.621	-1.1	1.295	0.54	0.672	-0.76	2.720	1.61
Obese	1.346	2.65	1.418	3.46	1.232	1.04	1.486	2.88	1.381	0.45	4.484	2.68	1.550	0.5	19.296	3.98
SES																
1 (most deprived)	Base category		Base category		Base category		Base category		Base category		Base category		Base category		Base category	
2	0.82	-2.4	0.904	-1.12	0.683	-3.11	1.009	0.07	0.793	-2.7	0.935	-0.77	0.677	-3.15	1.111	0.86
3	0.68	-3.82	0.718	-2.92	0.683	-2.64	0.752	-1.91	0.675	-3.81	0.745	-2.64	0.668	-2.7	0.882	-0.81
4 (least deprived)	0.543	-5.07	0.532	-4.29	0.612	-3.09	0.656	-2.36	0.541	-4.99	0.565	-3.86	0.594	-3.11	0.798	-1.22
BMI*SES																
Overweight*SES 2					1.306	1.55	0.782	-1.36					1.331	1.64	0.702	-1.87
Overweight*SES 3					1.011	0.06	1.092	0.42					1.064	0.31	0.918	-0.37
Overweight*SES 4					0.801	-1.04	0.661	-1.39					0.860	-0.65	0.533	-1.93
Obese*SES 2					1.455	1.35	1.049	0.22					1.437	1.28	0.749	-1.26
Obese*SES 3					0.984	-0.05	0.756	-0.99					0.972	-0.09	0.453	-2.54
Obese*SES 4					0.864	-0.36	0.675	-0.96					0.823	-0.44	0.358	-2.34
Age																
Age	1.179	5.34	1.160	4.62	1.178	5.30	1.163	4.70	1.180	5.22	1.158	4.58	1.180	5.23	1.170	4.93
Age squared	0.999	-2.31	1.000	-1.47	1.000	-2.29	1.000	-1.55	0.999	-2.33	1.000	-1.38	0.999	-2.31	1.000	-1.6
BMI*age																
Overweight*age									1.008	1.23	0.996	-0.62	1.006	0.77	0.987	-1.56
Obese*age									1.000	-0.04	0.982	-2.11	0.996	-0.28	0.964	-3.49
LR-test compared with main effects model					*P=0.31*		*P=0.29*		*P=0.44*		*P=0.11*		*P=0.43*		*P=0.01*	
*LR-test compared with SES*BMI model*													*P=0.65*		*P>0.01*	
*LR-test compared with age*BMI model*													*P=0.39*		*P=0.02*	
Test for unobservable heterogeneity	*P=0.38*		*P=0.04*		*P=0.38*		*P=0.26*		*P=0.3*		*P=0.39*		*P=0.3*		*P=1.00*	
*Test for proportional hazard**	*P=0.94*		*P=0.36*		*P=0.73*		*P=0.57*		*P=0.97*		*P=0.5*		*P=0.82*		*P=0.66*	

The dependent variable is survival time. Controlling for smoking, ethnicity, marital status, area and missing income. There are 1791 participants in the models for men and 2147 in the models for women.

*Significant tests means that the proportional hazard assumption is violated

In Appendix 5.2 we see the output of the same regressions run in the Cox models. In general there are only minor differences in the output from the Gompertz and Cox models. However, the coefficients for obesity show a stronger impact in the Gompertz models for both genders. For example, the hazard rate of obesity for women in the main effect model is measured in the Cox model to be 1.37 while it is 1.42 in the Gompertz model. The differences between the models are more pronounced in women than in men.

Table 5.7 show the output of the Gompertz models controlling for longstanding illnesses. We can see that we fail to reject the global tests across all models in both men and women. We also fail to reject the proportional hazard in the Schoenfeld and scaled Schoenfeld residual test for both men and women for all variables[21] (except from the dummy variable for East Midlands for men and SES quartile 2 in women). This means that we can apply a proportional hazard model.

The main effects models in Table 5.7 illustrate that obesity still has a significant negative impact on mortality. However, we can see that the impact is reduced going from 1.35 in the main effects model not controlling for illness to 1.31 in the main effects model controlling for illnesses in men. In women, corresponding figures were 1.42 and 1.30, respectively. So controlling for illnesses reduces the hazard ratio relatively more in women and it is now slightly more pronounced in men. The positive impact of SES in the main effects models are also slightly reduced, however it is still

[21] Not shown in Table 5.7.

significant in SES quartile 3 and 4 (least deprived) in both men and women. We also find that the unobservable heterogeneity parameter (frailty distribution) is significant in women but not in men, in the main effects models.

The comparisons of the models based on the likelihood ratio tests provide the same results as in the model not controlling for illnesses. Hence, the interactions do not improve the fit of the model in men. However, in women the fully interacted model has a significantly better fit than the other models.

In Appendix 5.3 we see the output of the Cox models controlling for illnesses. We find also here that the mortality rates are weaker in the Cox model than in the Gompertz model, however the differences are small and the trend is similar in both models.

Table 5.7: Gompertz models controlling for age and additional household and individual characteristics controlling for illnesses

	Main effects model				BMI interacted with SES model				BMI interacted with age model				Fully interacted model			
	Men		Women		Men		Women		Men		Women		Men		Women	
	Haz. ratio	z	Haz. ratio	z	Haz. ratio	z	Haz. Ratio	Z	Haz. Ratio	z	Haz. Ratio	z	Haz. ratio	z	Haz. ratio	z
BMI																
Normal weight	Base category		Base category		Base category		Base category		Base category		Base category		Base category		Base category	
Overweight	1.045	0.62	0.95	-0.64	1.004	0.03	1.009	0.08	0.672	-0.92	1.219	0.41	0.760	-0.52	2.202	1.25
Obese	1.314	2.41	1.304	2.59	1.233	1.01	1.346	2.1	1.498	0.56	4.735	2.74	1.765	0.64	19.746	3.93
SES																
1 (most deprived)	Base category		Base category		Base category		Base category		Base category		Base category		Base category		Base category	
2	0.868	-1.68	0.958	-0.48	0.724	-2.59	1.031	0.25	0.835	-2.06	0.982	-0.2	0.721	-2.6	1.128	0.97
3	0.741	-2.91	0.757	-2.45	0.759	-1.87	0.786	-1.59	0.734	-2.96	0.784	-2.18	0.749	-1.9	0.914	-0.58
4 (least deprived)	0.596	-4.25	0.571	-3.79	0.67	-2.49	0.68	-2.15	0.595	-4.18	0.604	-3.4	0.659	-2.47	0.818	-1.08
BMI*SES																
Overweight*SES 2					1.292	1.46	0.827	-1.04					1.309	1.52	0.750	-1.49
Overweight*SES 3					0.981	-0.1	1.138	0.6					1.017	0.08	0.976	-0.1
Overweight*SES 4					0.797	-1.05	0.686	-1.25					0.838	-0.75	0.571	-1.71
Obese*SES 2					1.418	1.23	1.115	0.48					1.391	1.15	0.771	-1.1
Obese*SES 3					0.912	-0.3	0.733	-1.09					0.886	-0.38	0.436	-2.64
Obese*SES 4					0.906	-0.24	0.756	-0.68					0.839	-0.39	0.388	-2.13
Age																
Age	1.168	5.02	1.153	4.4	1.168	4.99	1.155	4.46	1.171	4.96	1.155	4.48	1.172	4.96	1.166	4.76
Age squared	1.000	-1.93	1.000	-1.28	1.000	-1.93	1.000	-1.35	1.000	-2	1.000	-1.31	1.000	-1.96	1.000	-1.47
BMI*age																
Overweight*age									1.007	1.04	0.996	-0.55	1.004	0.55	0.989	-1.27
Obese*age									0.998	-0.18	0.980	-2.33	0.995	-0.43	0.963	-3.59
LR-test compared with main effects model					*P=0.39*		*P=0.19*		*P=0.52*		*P=0.07*		*P=0.47*		*P=0.01*	
*LR-test compared with SES*BMI model*													*P=0.71*		*P<0.01*	
*LR-test compared with age*BMI model*													*P=0.39*		*P=0.03*	
Test for unobservable heterogeneity	*P=0.32*		*P=0.05*										*P=0.3*		*P=0.42*	
*Test for proportional hazard**	*P=0.71*		*P=0.14*		*P=0.52*		*P=0.25*		*P=0.8*		*P=0.23*		*P=0.62*		*P=0.29*	

The dependent variable is survival time. Controlling for smoking, ethnicity, marital status, area, missing income, number of longstanding illnesses and whether or not the illnesses is limiting. There are 1791 participants in the models for men and 2147 in the models for women.

*Significant tests means that the proportional hazard assumption is violated

We now show the predicted life expectancy based on the hazard models. The model with the best fit in men was the main effects model while in women it was the fully interacted model; hence we base the predictions on these.

The predicted life expectancy of individuals aged 55, 65 and 75 across the BMI groups for men and women are in Table 5.8. Across all men we see that obese individuals have a lower predicted life expectancy than normal weight individuals. The life expectancy among men of normal weight aged 55 was 25 years, and for those being overweight, and obese it were 24.6, and 22.1, respectively. In women we see that obesity reduce life expectancy for individuals aged 55 and 65. However, in women aged 75 obesity is associated with increased life expectancy.

Table 5.8: Predicted survival time based on the Gompertz distribution with gamma frailty

	Age 55			Age 65			Age 75		
	Normal weight	Overweight	Obese	Normal weight	Overweight	Obese	Normal weight	Overweight	Obese
Men									
All	25	24.6	22.1	15.9	15.6	13.5	9.5	9.2	7.7
SES									
1 (most deprived)	21.9	21.5	19.2	13.4	13.1	11.2	7.6	7.4	6.1
2	23.8	23.4	21	15	14.6	12.6	8.7	8.5	7.1
3	25.7	25.3	22.8	16.5	16.2	14.1	9.9	9.6	8.1
4 (least deprived)	28	27.5	25	18.5	18.1	15.9	11.4	11.1	9.4
Women									
All	31	30.8	26.4	20.1	21.1	19.3	11.6	13.4	13.6
SES									
1 (most deprived)	30.5	27.7	21	19.6	18.3	14.5	11.2	11.1	9.6
2	29.4	30.2	22.7	18.7	20.6	16	10.5	12.9	10.8
3	31.8	29.8	30	20.8	20.2	22.6	12.1	12.6	16.5
4 (least deprived)	32.8	36.5	33.5	21.7	26.4	25.9	12.8	17.9	19.5

Note

The predictions are based on the models in Table 5.6 where we control for age, age squared, smoking status, marital status, geographical area and ethnicity. The predictions are in years, although they are based on models where the time variable is in months.

Table 5.8 also shows the results by SES groups. Less deprived SES groups had a higher predicted life expectancy in both men and women across all BMI groups. For example, normal weight men aged 55 have a predicted life expectancy of 21.9 years in SES quartile 1 (most deprived), and 23.8, 25.7 and 28 years in quartiles 2, 3 and 4 (least deprived), respectively. In normal weight women aged 55 the predicted life expectancy in SES quartile 1 (most deprived) was 30.5, in quartiles 2, 3 and 4 (least deprived) it was 29.4, 31.8 and 32.8 years, respectively.

We did not find an interaction effect between obesity and SES in the latent variable model for men. This means that the relative impact of obesity on *mortality* does not vary by SES. Hence, obesity has an equally negative impact in each time period in men. Individuals who live longer will have a more negative impact of obesity. The reason is that the negative impact of obesity "works" for more time periods. Higher SES men live longer than lower SES men, so they will lose more life in absolute terms due to obesity. For example, a man aged 65 SES quartile 1 (most deprived), will on average lose 2.2 years due to obesity, while and a man aged 65 in SES quartile 4 (least deprived) would on average lose 2.6 years due to obesity.

The likelihood ratio tests supported the fully interacted models in women. The results of the predictions based on this model reveal an opposite effect of obesity in high and low SES groups. Our results show SES variation in the impact of obesity on life expectancy in women. The predictions illustrate that obesity reduces life expectancy for women in SES quartile 1 (most deprived) across women aged 55, 65 and 75.

However, in SES group 4 (least deprived) obesity is associated with increased life expectancy across women aged 55, 65 and 75.

In Table 5.9 we see predictions based on the survival models controlling for longstanding illnesses. We see similar trends of a lower life expectancy in obese compared with normal weight in men and women (except in women aged 75 where the trend is positive). We also see that lower SES individuals have a lower life expectancy than higher SES individuals. In addition, we find a similar SES gradient in the impact of obesity on life expectancy. Hence, obesity has a negative impact on life expectancy in low SES women but positive impact on life expectancy in high SES women.

Table 5.9: Predicted survival time based on the Gompertz distribution with gamma frailty: Controlling for longstanding illnesses

	Age 55			Age 65			Age 75		
	Normal weight	Overweight	Obese	Normal weight	Overweight	Obese	Normal weight	Overweight	Obese
Men									
All	25	24.6	22.4	16	15.6	13.8	9.4	9.1	7.8
SES									
1 (most deprived)	22.6	22.2	20.1	14	13.6	11.9	7.9	7.6	6.5
2	23.9	23.5	21.4	15.1	14.7	12.9	8.7	8.4	7.2
3	25.5	25	22.8	16.4	16	14.2	9.7	9.4	8
4 (least deprived)	27.6	27.2	24.9	18.2	17.9	15.9	11.1	10.8	9.3
Women									
All	30.5	30.7	26.6	20.1	21.2	20.0	11.8	13.5	14.5
SES									
1 (most deprived)	30.2	28.3	21.7	19.8	19.1	15.5	11.5	11.8	10.7
2	29.0	29.9	23.0	18.7	20.6	16.6	10.7	13.0	11.6
3	31.0	29.4	30.5	20.6	20.1	23.5	12.2	12.6	17.7
4 (least deprived)	32.1	35.9	32.7	21.6	26.1	25.7	13.0	17.7	19.7

Note

The predictions are based on the models in Table 5.7 where we control for age, age squared, smoking status, marital status, geographical area, ethnicity number of longstanding illnesses and whether or not the illnesses is limiting. The predictions are in years, although they are based on models where the time variable is in months.

5.5 Discussion

In this chapter we use the *Health and Lifestyle Survey* from 1984/1985 and its longitudinal follow up in 2009 to investigate the relationship between obesity and mortality/life expectancy and whether or not this varies by SES. We have conducted a comprehensive analysis using semi-parametric and parametric methods to investigate this relationship, applying different combinations of covariates accounting for unobservable heterogeneity. To decide upon the correct parametric distribution we conduct extensive analysis and based on this choose a Gompertz distribution. We run both semi-parametric Cox models and parametric Gompertz models, which show similar results.

In men, obesity is negatively associated with life expectancy across all ages. In women, obesity reduces life expectancy at the ages of 55 and 65; while it increases life expectancy in women aged 75. This is consistent with Stevens et al. (1998) who found that obesity increased mortality in American women aged 30 to 74, while the women over the age of 75 had a reduced risk of death[22].

We also investigate the impact of overweight as well as obesity on mortality/life expectancy. We do not find the hazard rate of overweight to be significant. This is consistent with the existing literature where some have found a positive effect (see, e.g., Peeters et al., 2003; Adams et al., 2006), but not others (see, e.g., Flegal et al., 2005).

[22] They also found this in men. Hence, our study contradicts these results.

Our results illustrate a lower mortality rate and longer life expectancy in higher SES groups than in lower SES groups in both men and women. We run models with and without interactions between BMI and SES groups. The likelihood ratio tests support a model with no interactions in men. Hence, the impact of SES and obesity on *mortality* does not vary across SES groups in men. When we look at *life expectancy* we look at the consequences of *mortality* over a lifetime. Our models indicate that obesity leads to a greater loss in life years due to obesity in higher SES men than in lower SES men.

We predict the number of years survived after the measurement date at age 55, 65 and 75. A 55 year old obese man in the lowest SES group will on average live 19 years, while a 55 year old obese man in the highest SES group will on average live for 25 years. An obese man in the highest SES group has a higher predicted life expectancy (25 years) than a normal weight man in the lowest SES group (22 years). Men in higher SES groups have a longer predicted life expectancy and will have a greater predicted absolute loss in life years due to obesity. The reason is that the negative impact of obesity will "work" for more time periods in these models.

In women we find that models fitting separate interactions between age and BMI did not improve the fit compared with the main effects model. Neither did models that fit interactions between SES and BMI. However, when we fit a model with both interactions in one model they significantly improve the fit of the model. Hence, a fully interacted model with interactions between BMI and SES, and BMI and age was the model with the best fit in women. This means that the mortality rate due to obesity varies by both SES and age.

A 55 year old obese woman in the lowest SES group will on average live 21 years, while a 55 year old normal weight woman in the lowest SES group will on average live for 30.5 years. Similar results for women in the highest SES group are 33.5 and 32.8, respectively. Hence, obesity reduces life expectancy in the low SES women aged 55 and increase life expectancy in the high SES women aged 55. A similar SES gradient was found in women aged 65 and 75.

We fit a model controlling for number of longstanding illnesses and whether or not these are limiting. The models show that the impact of obesity and SES is slightly reduced. When we control for longstanding illnesses we must keep in mind that we only control for longstanding illnesses at the time of obesity measurement and not between the time of measurement and death/censoring. Hence, longstanding illnesses can still have an impact on the hazard ratio, however this is after obesity measurement hence it follows the obesity status. In this way we control for endogeneity issues related to omitted illness/health variables.

Mayhew, Richardson and Rickayzen (2009) have calculated expected life years lost from obesity by applying hazard ratios (based on Cox models) for obese individuals to life tables. They tested an interaction between BMI and age and found it to be insignificant in both men and women. We find an interaction between age and BMI in women. They found that for non smoking males, the expected years of life lost across all ages would be 4 to 16 years for individuals with BMI in excess of 35 compared to having a BMI of 24. For non smoking females this was 2 to 10 years compared to

having a BMI of 26. Hence, they have found a more pronounced effect of obesity in males. We find a more pronounced obesity effect in females in the model controlling for individual and household characteristics and a more pronounced obesity effect in males in the model that also control for longstanding illnesses. The main differences between our studies are as mentioned that we control for a range of covariates (e.g., ethnicity) and account for unobservable heterogeneity (frailty). We have a longer follow up period and do not stratify by smoking status (we control for it in the regression). Hence, we can argue that our study is less likely to be affected by omitted variables bias (e.g. unobserved health variables, discount factor or general frailty), decreasing duration dependence and sample selection bias. In addition, we demonstrate that the hazard rate varies by both age and SES in women.

In this chapter we have presented both hazard ratios and life expectancy. The relative mortality or the absolute loss in predicted life years answers different questions and which of these one wants to focus on depends on the question that one wants to answer. The results in Chapter 3 showed that obesity leads to a greater loss in HRQL in lower SES groups than in higher SES groups. This supports the hypothesis of a more negative impact of obesity on health in lower SES groups compared with higher SES groups. If we view life expectancy as another dimension of health these results support this hypothesis in women, but not in men. Hence, a question arises regarding why we do not get these results in men. HRQL is a measure of perceived health status. It might be that lower SES men perceive the consequences of obesity to be worse as they have less favourable economic and social circumstances that will increase the burden of

obesity. But that the actual physical impact of obesity and its related comorbidities is similar across SES groups.

The result in women serves to further emphasize the heterogeneous nature of obesity. Attention needs to be paid to the pathways in which obesity influences health and life expectancy. This has consequences for decision making in health and care is required when selecting estimates for BMI categories in simulation models.

We are not aware of any studies that have used other alternative parametric distributions than the Cox proportional hazard model. We suggest more attention should be given to parametric distributions in this type of analysis as there are two advantages with this approach. First, one can predict life expectancy based on the parametric distribution. Second, one can control for unobservable heterogeneity in datasets without repeated measurements for the same individual. We find that the unobservable heterogeneity is significant in the main effects model in women even after controlling for a range of individuals characteristics including longstanding illnesses. The downside of using parametric distributions is that their validity relies on finding a distribution that fits the data.

Our findings have a number of limitations. First, our obesity data are measured once for each individual; it would be preferable to have repeated measures over time as we do not capture obesity onset timing which might have an impact on survival. Second, the relatively small sample size at higher BMI levels does not allow us to divide our sample into additional obesity categories. It is likely that the effect of obesity class III

(BMI>40 kg/m^2) on life expectancy differs from the effect of obesity class I (30 kg/m^2>BMI>35 kg/m^2). Third, the data are based on measures from 1984/1985, when the prevalence of obesity was lower than it is today; the impact of obesity may change as population norms change. Fourth, our measure of obesity is BMI, which has been criticised, e.g., because it does not incorporate body fat, with body fat content being what is actually the independent predictor of ill health (Burkhauser & Cawley, 2008). Hence, measurement error might arise.

Nonetheless, we believe this chapter has a number of advantages over previous studies. Our BMI measure is based on height and weight measurements obtained during the nurse visit, not on self-reported values. Thus, the likelihood of systematic measurement error is reduced. We are the first study to investigate the relationship between BMI and mortality/life expectancy in the UK by including BMI as a categorical variable. This allows us to conclude that there is significant difference in the mortality of obese individuals compared to normal weight individuals. We are the first to apply the HALS longitudinal follow up from June 2009, hence we have a longer follow up period. We are the first to apply both semi-parametric and parametric models in our estimation. This makes our results more robust as the semi-parametric models are not constrained by a parametric distribution. Earlier studies have important limitations in terms of endogeneity. The most important is omitted variable bias[23]: for example there might be health variables that reduce/decrease weight and increase mortality. We account for the issue by including a frailty distribution. Furthermore, we fit models

[23] Reverse causality is not an issue as the dependent variable and the independent variable are determined at different points in time.

both with and without controlling for longstanding illnesses and show that obesity has an impact in both. Also, we have information on individual and household characteristics so we can argue that it is less likely that the estimated effects of obesity on life expectancy in our models are due to their correlation with other omitted variables. Controlling for unobservable heterogeneity also ensures that our results do not suffer from decreasing duration dependence.

To summarise, this chapter has shown that, as in previous studies, obesity is negatively correlated with life expectancy in men and younger women. In older women obesity is associated with increased life expectancy. We have found that these results are robust across semi-parametric and parametric distributions controlling for a range of different covariates both with and without controlling for unobservable heterogeneity. In addition, we find that the impact of obesity on life expectancy depends on SES and that the impact is more negative in lower SES women than in higher SES women. However, this was not found in men.

CHAPTER 6

Socioeconomic variation in the relationship between obesity and health service use

6.1 Introduction

It has been shown that there is a socioeconomic gradient in obesity (see, e.g., Wardle, 2002) and socioeconomic status (SES) has been showed to have an impact on how the health service are used (Allin, Masseria & Mossialos, 2010; Benzeval & Judge, 1994; Morris, Sutton & Gravelle, 2005; Saez, Saurina, Coenders & Gonzalez-Raya, 2006; Vallejo-Torres & Morris, 2010; Veugelers & Yip, 2003). For example, lower SES individuals have been shown to use more primary care and higher SES individuals use more secondary care (See, e.g., Morris, Sutton, & Gravelle, 2005). In this chapter we develop the current approach to allow for interactions between obesity and SES. We look at if the association between a current health risk factor (obesity) and health service use differs across SES groups.

Based on the theoretical model in Chapter 2 we presented the following empirical model for health service use:

$$m_t = a_0 + a_1 L_t + a_2 SES_t + a_3 L * SES_t + a_4 X_t + a_5 S_t + e_t \quad (6.1)$$

Where health service use (m_t) is determined by lifestyle (L_t), SES, other non-need variables (such as ethnicity) and supply. The a's are parameters to be estimated and the e_t is the error term. We have also included an interaction between lifestyle and SES, which allows the association between obesity and use to vary by SES or vice versa. The aims of this chapter are to investigate the relationship between obesity and health service use, and to investigate whether or not this relationship varies by SES. We start with a review of the literature of the relationship between obesity and health service use. We then undertake the analysis using one panel dataset and one pooled cross sectional dataset. Both include height and weight, plus a comprehensive set of individual and household characteristics that allows us to control for factors that affect the relationship between obesity and health service use. It is an advantage to use a panel dataset as it allows us to account for time invariant heterogeneity. In addition, this dataset has good geographical identifiers. However, the cross sectional dataset has objectively measured height and weight, which is less likely to be biased compared to the self-reported height and weight in the panel dataset.

6.2 Literature review of the association between obesity and health service use

The objective of the review was to identify studies that investigate the impact/association between obesity and health service use. We have restricted this search to studies presenting statistics on health service use; hence we exclude cost of illness studies.

6.2.1 Search strategy

We followed similar strategies as in Chapter 3 and searched the literature in three stages: First, we search the journals[24]: *Obesity Reviews; Health Economics; Journal of Health Economics; and, Social Science & Medicine*[25]. The search term was:

obesity OR obese OR overweight OR weight loss AND GP OR physicians OR specialist OR outpatient OR inpatient OR hospital OR length of stay OR medication OR drugs OR preventive services OR nurse OR primary care OR secondary care

From the first stage we identified a total of 6 studies. The second stage was a reference search of the identified papers in the first stage. In the third stage we used the "cited by" option in Google scholar to identify papers that had referenced the papers identified in the two first stages. In total we identified 27 studies.

From each paper we extract data on year, country, type of health service use studied, methods and results. This is shown in Table 6.1. We restricted the search to papers written in English.

[24] Tried to conduct a search in MEDLINE using the search term: (((((("Obesity"[Mesh]) OR "Body Mass Index"[Mesh]) AND "Inpatients"[Mesh]) OR "Outpatients"[Mesh]) OR "Health Personnel"[Mesh]) OR "Preventive Health Services"[Mesh]. However, this gave 676990 hits so we decided to use this second best solution.

[25] *Obesity Reviews* is the highest ranked obesity journal by impact factor (Thompson ISI, 2010). *Health Economics* and *Journal of Health Economics* were chosen as they are the two top ranked journals in the field of health economics in the *Thomson Reuters Journal Citation Reports*. *Social Science & Medicine* is the world most cited social science journal (Thomson ISI, 2010) and it has a focus on health economics.

Table 6.1: Literature review of studies examining the associations between obesity and health service use

Authors, country, year	Dataset	Statistical methods	Type of use	Results
Bertakis, Azari, USA, 2005	Patients at a primary care clinic (year not given), n= 509, age not given	Used log-transformed utilisation variables to investigate the impact of a dummy for obesity (BMI $\geq$ 30) controlling for health status, depression, age, education, income, and sex.	Primary care visits, specialty clinic visits, emergency department visits, hospitalisations, and laboratory, diagnostic, and radiological tests	Obesity was significantly related to the use of primary care and diagnostic services
Chen, Jiang, Mao, Canada, 2007	National survey (2003), n=113603, age 20+	Used logistic regressions methods to investigate the impact of four BMI groups controlling for age, marital status, household size, number of bedrooms, income adequacy, educational level, immigrant status, visible minority, marital status, smoking status, alcohol use, exercise and allergy history	Hospital admissions and length of stay	Obese individuals have more hospital admissions and longer length of stay
Chu, Wang, MS, Chang, and Wu, Taiwan, 2010	Cross sectional national survey (2001), n=12283, age 20-85	Use a two-stage approach with linear regression and tobit models to investigate the effect of four BMI groups on use controlling for age, smoking, drinking and co morbidities	Outpatient visits and hospitalization	Medical utilization was found to increase with higher BMI
Counterweight project team, UK, 2005	Patients at 80 general practises around in the UK (2000-2003), n=7300, age 18-75	Logistic regression comparing normal weight (BMI 18.5–<25) and obese (BMI>30) controlling for age, sex, deprivation category, country, and the presence of a comorbidity	Medication	Obesity increased prescribing in most categories
Counterweight Project Team, UK, 2005	Patients at 80 general practises around in the UK (2000-2003), n=7300, age 18-75	Used logistic regression methods to investigate the correlation between health service use and obesity (BMI>30) compared with normal weight (18.5<BMI<24.9). Controlled for age, sex, social deprivation category, country, and group and additionally for the presence of comorbidities.	GP visits, practice nurse visits, inpatient and outpatient visits	Found a significant relationship between obesity and; GP visits, practice nurse visits, inpatient and outpatient visit
Dzien et al., Austria, 2003	Outpatients (year of measurement not given), n=3360, only mean age given, which was 57	Provided mean and standard deviations of BMI categories across medication categories. Used significant tests to test for variation in the means.	Medication	Found that more medication was given to individuals who were overweight or obese
Folmann et al., Denmark, 2006	Two subsamples, one with measured WC, n=5151 and one with measured BMI, n=4048 (1999-2001), age 30-60	Used two-part models and Poisson regression investigating the impact on use	Hospital contacts	Obese men and women had an increased number of hospital contacts compared with normal weight individuals
Guallar-Castillon et al., Spain, 2002	Cross sectional Spanish National Health Survey (1993), n=13244, age 16+	Logistic regression models were used to calculate the impact of three BMI groups on use controlling for age, education level, occupation, civil status, social support, tobacco use, alcohol consumption, physical activity at work and during leisure time, job status and town of residence. Does also investigate an interaction	Hospitalisations, medical visits, utilisation of hospital emergency services, and medication use	Overweight and obese women visited the physician, used hospital emergency services and took medication with greater frequency than did women of normal weight. They do not find the interaction terms between education and

		between BMI as a linear term and education.		BMI to be significant.
Gupta & Greve, Denmark, 2011	National health survey combined with register data (2000 and 2005), n=10013, age 25-60	Use a finite mixture model to investigate the impact of three BMI groups on use controlling for gender, education, marriage, age, and smoking	GP use, Hospitalisation and Bed days,	Estimated bodyweight effects vary across latent classes and show that being obese or overweight does not increase the utilization of GP services among infrequent users but does so among frequent users
Han, Truesdale, Taber, Cai, Juhaeri and Stevens, USA, 2009	Cross sectional study from four US communities (1987-1989), n=15792, age 45-64	Used negative binomial models to investigate the impact of three BMI groups adjusted for race, gender, field centre, age, physical activity, education level, smoking status, alcoholic beverage consumption and health insurance. Adjusted numbers of hospitalizations were calculated after setting covariates to the mean value.	Hospitalisations	Obese men and women had more hospitalisations than non obese men and women. Found no race differences.
Hauck & Hollingsworth, Australia, 2010	Data from 122 Australian hospitals (2005/6), episodes=435 147, age 17+	Quantile Regression analysis is used to generate 19 estimates of the difference between severely obese and other patients across the whole range of length of stay,	Hospital length of stay across specialities	There are significant differences in average length of stay for almost all specialties.
Luchsinger, Lee, Carrasquillo, Rabinowitz, Shea, USA, 2003	Medicare Beneficiary Survey (1992-94), n=8754, age 65-100	Poisson regression was used for multivariate analyses relating five BMI groups and BMI as a continuous variable to number of hospitalizations, adjusting for sex, age, smoking status, and heart disease. Only in individuals aged 75 and older.	Hospitalisations	Does not find a significant association between BMI and hospitalizations
Pearson, Bhat-Schelbert, Ford and Mokdad, USA, 2009	National Ambulatory Medical Care survey (2006), n=9280, age 18+	Used logistic regression models to investigate the impact of two BMI groups (exclude underweight) controlling for age, sex, race, payment source, major reason for the visit, and the total number of co-morbid chronic conditions	Duration of outpatient visit and prescribed medication	Time spent with the provider was found to be greater among visits with obese patients, but not significantly different from visits with non-obese patients. The number of medications for each visit was found to be significantly greater for visits where the patient was considered to be obese.
Peytremann-Bridevaux & Santos-Eggimann, Europe, 2007a	Data from 10 European countries (SHARE survey) (2004), n=16695, age 50-79	Used logistic regression to investigate the impact of three BMI groups (excluded underweight) on use controlling for age, socioeconomic status, smoking, physical activity, alcohol, country of residence, and chronic conditions	Ambulatory care, high use of a GP, visits to specialists, high use of medication, hospitalization, high number of times hospitalized and nights spent in the hospital, surgery, home healthcare and domestic help	Demonstrated an increased risk of using ambulatory care and visiting general practitioners, as well as taking ≥2 medication categories.
Peytremann-Bridevaux & Santos-Eggimann, Europe, 2007b	Data from 10 European countries (SHARE survey) (2004), n=16695, age 50-79	Used logistic regressions to investigate the impact of three BMI groups controlling for age, sex, education, income, smoking, alcohol, physical activity, and country	Influenza immunization, colorectal and breast cancer screening, discussion and recommendation about physical activity, and	Overweight and obesity were associated with higher odds of receiving influenza immunization but not with receipt of breast or colorectal cancer screening. Overweight and obese individuals mentioned

			weight measurement	more frequently that their general practitioner discussed physical activity or checked their weigh
Popoola, UK, 2004	Health Survey for England (1998), n=4102 , age 50-69	Used logistic regression methods to investigate the correlations between health service use and across three BMI groups in a population between 50 and 69. Controlling for age, gender, ethnicity, SES, education, alcohol and smoking.	GP visits, medication use, outpatient visits and inpatients stays	Significant association between BMI and outpatient services and medication use but not inpatient services and GP visits
Quesenberry, Caan, Jacobson, USA, 1998	Members of an health maintenance organisation (1993), n=17118, age 20+	Used Poisson regression to investigate the impact of four BMI groups (excluding underweight) on use controlling for age sex, smoking, alcohol consumption, educational level, and race	hospitalisations, laboratory services, outpatient visits, outpatient pharmacy and radiology services	There was an association between BMI and annual rates of inpatient days, number and costs of outpatient visits, costs of outpatient pharmacy and laboratory services,
Reidpath et al. Australia, 2002	Cross sectional Australian health survey (1995), men: n=17033 women: n=17174, age 20+	Fit separate logistic models using continuous and categorical BMI (four groups) as the dependent variable controlling for age and income.	Medication use, hospitalization, emergency room visits, outpatient clinic visits, doctor visits, other health professional visits, and in women only regular pap smear tests, regular breast examinations, regular mammograms	Found a positive relationship with medication use, outpatient and doctor visits for men and women in Australia. In addition, they found a significant relationship for women for other health professionals. They also looked at the relationship between BMI and preventive service for women, and found that this relationship was negative.
Saez, Saurina, Coenders & Gonzalez-Raya, Spain, 2006	Using a survey of the population in one regian in Spain (2002), n=734, all ages	Three BMI groups. Use a regional multilevel model with random effects. Controls for age, gender, marital status, co morbidities, smoking, medical insurance tenancy, income, labour status and profession, place of birth, people in household, number of cars, number of owned dwellings. Fit separate models across income groups.	GP and specialist visits	Find that obese individuals have a higher probability of having GP visit, however a lower probability of a specialist visit.
Sansone RA, Sansone LA, Wiederman, USA, 1998	Women in one health maintenance organisation (year not given), n=194, age 17-52	Used linear regression and logistic models to investigate the impact of five BMI groups on use controlling for age, race, family income, education, smoking status, and health insurance status	Mammography, clinical breast examination, gynaecologic examination, and Papanicolaou smear and the number of physician visits	Among women, an increased BMI is associated with decreased preventive health care services, however increased physician visits
Schafer & Ferraro, USA, 2007	Hospital records (1971-1975), n=4574, age 40-77	Use various methods including interval regression and duration models to investigate the impact of the duration of obesity on use controlling for age, gender, race, education, income, live alone, health insurance, drinking, smoking, sedentary lifestyle, place of residence, morbidity and hospitalisation history	Hospitalisation and length of stay	Results reveal that obesity increased hospital admissions and length of stay over the 20 years studied.
Trakas, Lawrence, Shear, Canada,	Canadian cross sectional survey (1994), n=17626,	Used logistic regression to investigate the impact of obesity (BMI $\geq$ 27) on use controlling for	Physician visits, disability days, admissions to	Obese respondents were more likely to consult with physicians,

1999	age 12+	age, sex, level of physical activity and education	hospital and medication use	be prescribed a number of medications and to require excess disability days
Van Dijk, Otters and Schuit, The Netherlands, 2006	Patients in a Dutch general practice (2001), n=8944, age 18+	Used bivariate and multilevel analysis to investigate the impact of BMI (in two or three groups) on use controlling for age, sex, educational level, type of health insurance, residence, smoking, alcohol consumption and exercising	GP use and prescriptions	Obesity increases the workload of Dutch general practitioners and the use of prescribed medication.
Von Lengerke et al., Germany, 2005	Cross sectional German survey, n=947, age 25-74	Use a two-part model approach for the two utilization parameters. The first part is a logistic model and the second part is a zero-truncated negative binomial model. Controlling for sex, age, social class, public vs. private health insurance, and rural vs. urban place of residence Use 5 BMI groups but exclude underweight.	GP visits and inpatient stays	Found a relationship between obesity and GP use. Regarding days in hospital, only the extremely obese reported significantly more utilization than those normal weight
Wildenschild, Kjøller, Sabroe, Erlandsen, and Heitmann, Denmark, 2011	Cross sectional Danish Health Survey (1987, 1994, 2000 and 2005), men: n=19142, women: n=18335, age16+	Used logistic regression to investigate the impact of four BMI groups controlling for age, marital status, educational level, employment, and smoking status	GP, physician from the emergency service, medical specialist, industrial medical officer, emergency ward, outpatient clinic, hospitalization, and other physician	Computed a composite measure for all types of care and found a significant positive association for both men and women
Yan et al., USA, 2006	Participants recruited in the Chicago area (1967-1963), n=17643, age 31-64. Followed up the individuals and recorded hospital stays until 2002	Used logistic regression models to investigate the impact of three BMI groups on hospitalisations controlling for sex, race, age, education, minor lectrocardiographic abnormality, SBP and serum total cholesterol	Hospitalisation	Those who are obese in middle age have a higher risk of hospitalization and mortality from CHD, cardiovascular disease, and diabetes in older age than those who are normal weight
Zizza , Herring, Stevens, Popkin, USA, 2004	Used national US longitudinal survey sectional survey (1971-1992), n=14407, age 25+	Used Poisson regression to investigate the impact of five BMI groups on length of stay controlling for age, gender, race, marital status, whether respondents had children, smoking status, socioeconomic status, region of residence, and urban-rural residence in our models	Inpatient days	Obese individuals have longer hospital stays

6.2.2 Findings

We start with an overview of the studies, methods and results. We then discuss in more detail the UK specific studies and studies that measure socioeconomic variation in the impact of obesity on health service use.

We identified 27 studies and there were wide variations in the origins of the data, however many of the studies were from North America (11 studies). We also found three studies from the UK (Counterweight Project Team, 2005a; Counterweight Project Team, 2005b; Popoola, 2004).

The literature covers a variety of services including; GP visits, specialist visit, emergency services, preventive services, medication use, inpatient and outpatient visits. Although, countries, populations, age range and other control variables differ across studies it has been shown that obesity is generally associated with increased use of health services. Table 6.2 shows summary statistics of selected results of the review. Of the ten studies looking at GP visits, nine find that obesity increases the frequency of visits. The evidence for a relationship for outpatient visits is also quite robust, with six identified studies all showing a positive relationship. The evidence for hospitalisations however is more mixed and of the 18 identified studies 12 demonstrate a positive relationship between obesity and hospital admittance, while the rest shows no relationship. However, there appears to be a positive association between inpatient stays and those in obesity class III (BMI>40 kg/m^2).

The literature review also identified ten studies investigating the association between obesity and medication use. These studies all provide evidence of a positive relationship between obesity and medication use. We have found two studies that investigate the effect of obesity on medication use disaggregated into therapeutic classifications (van Dijk, Otters and Schuit, 2006; Counterweight Project Team, 2005b). The Counterweight Project Team (2005b) demonstrate that obese (BMI>30 kg/m^2)

individuals compared with normal weight (18.5<BMI<25 kg/m^2) individuals have a significantly increased use of medications for the cardiovascular system; central nervous system; endocrine system; musculoskeletal and joint disease; infections; gastrointestinal; and skin and respiratory system controlling for age, sex, deprivation category, and country (England/Scotland). In addition, van Dijk, Otters and Schuit (2006) show a significantly higher odds of using medicines for; alimentary tract and metabolism, cardiovascular system, dermatological, systemic hormonal preparations, anti-infective for systemic use, musculoskeletal system and the respiratory system for obese (BMI>30 kg/m^2) compared normal weight (BMI<25 kg/m^2) men and women in the Netherlands.

Table 6.2: Summary of identified studies examining the evidence on the association between obesity and health service use

Type of use	**Evidence of relationship between obesity and use of obesity**		
	Obesity increase use	*Obesity decrease use*	*No association*
GP visit	9	0	1
Practice nurse	1	0	0
Hospitalisations	12	0	6
Outpatient	6	0	0
Medication	10	0	0

A range of different econometric methods have been applied and the most common method is a standard binary choice model (usually a logistic regression model) looking at the relative odds of utilising the health service *P(HSU=1)*, where *HSU* denotes health service use. However, two part models are becoming more popular in more recent studies. The two part model investigates the decision process regarding use in two parts where the first is a standard binary choice model modelling the probability of use *P(HSU=1)*, and the second step equation models the level of utilisation among users

E(HSU|HSU>0). In this way they capture not only the binary decision to use the health service but also the intensity of use.

Most of the studies have presented their results in odds ratios or relative risk (*RR*). The relative risk is the risk of health service use for obese (*O*) vs. non obese is:

$$RR = \frac{P(HSU = 1|O = 1)}{P(HSU = 1|O = 0)} \tag{6.2}$$

Few have calculated the absolute difference, which is the difference in risk:

$$ME = P(HSU = 1|O = 1) - P(HSU = 1|O = 0) \tag{6.3}$$

We will discuss the implications of these different interpretations later in this chapter.

6.2.2.1 UK evidence

The function and the structure of the health care system vary across countries and this may influence people's opportunities and choice of health service use. For example, in some countries GP's act as gatekeepers while in other countries patients can go straight to a specialist. Hence, country specific evidence is important. The Counterweight project (2005a) team looks at obese and normal weight individuals who have had their height and weight registered at 80 practices spread across England and Scotland. They found that obese patients made significantly more visits to the general practitioner, practice nurse, and hospital outpatient units than normal weight patients,

and they were admitted to the hospital more often. For both GP and practice nurse visits, the relationship remained after adjusting for age, sex, social deprivation category, country, and number of comorbidities. As described above the Counterweight Project Team (2005b) also conducted a study to investigate the impact of obesity on prescribing in primary care. They find that obese individuals are prescribed more drugs than normal weight individuals in primary care. Furthermore, a study has been conducted using the Health Survey for England (1998) to examine the association between obesity and health service use in the population between 50 and 69 in England (Popoola, 2004). Popoola found that obese (BMI>30 kg/m^2) individuals had a significantly higher relative odds of using outpatient services and using two or more prescription drugs compared with normal weight (BMI: 18.5 - 25 kg/m^2). However, no significant relationship was found between the use of inpatient services and GP visits for this group.

Our study will differ from these studies in a number of respects. They include only one category for obesity. We will include more categories as it is likely that the effect of obesity class II/III (BMI>35 kg/m^2) on health service use differs from the effect of obesity class I (30 kg/m^2>BMI>35 kg/m^2). These studies do not take account of supply variables and it has been indicated that supply has significant effects on health service utilisation (Morris, Sutton & Gravelle, 2005). Hence, we will take account of this in our analysis. They use cross-sectional data; hence they cannot exclude the influence of time invariant heterogeneities, which we to some extent address with panel data. They do not focus on the gender difference by stratifying by gender. Instead they control for it in the analysis. We illustrated in earlier chapters there may be a separate gender

effect of BMI on health. Hence, it is interesting to see if this effect materialises itself in this analysis as well. They use binary regression models to investigate relative odds of positive outcomes. Hence, they do not look at the number of visits/drugs used. They present their results in odds ratios and it would be useful with a study that uses absolute values rather than relative. Furthermore, none of these studies have stratified the results by SES groups. Hence, this is an area that has not received enough attention.

6.2.2.2 Socioeconomic variation in the association between obesity and health service use

We have identified two studies that investigate SES variation in the association between obesity and health service use. The first by Saez, Saurina, Coenders & Gonzalez-Raya (2006) use Spanish data to look at relative risks of GP and specialist visits. They find that obese individuals have a higher relative risk of having a GP visit, however a lower probability of a specialist visit. They also look at the relative risk of GP and specialist visits across income groups. They find that relative risk of GP use decreased with income and that the relative risk of specialist services increased with income. Then they stratify income across BMI groups. They state that they identify the same income-use trend in the obesity and overweight groups.

The second is another Spanish study by Guallar-Casillon et al. (2002) that looks at the association between BMI and health service use (Hospitalisations, medical visits, utilisation of hospital emergency services, and medication use) in women interacting BMI and education. The interaction was non-significant. They use a binary choice

model and therefore look at if the health service is used or not. They do not investigate the number of contacts.

The aim of both these studies was to present their results in relative terms (odds ratios and relative risks), hence they do not analyse, e.g., the number of visits and drugs used or the percentage increase in the number of visits in absolute terms. As will be explained later they might have reached a different conclusion if this was their aim. Also, neither of these studies was undertaken in the UK, and results might vary between countries, for example, Spain has relatively high involvement of the private sector in their health care system.

To conclude, there is evidence of an association between obesity and a wide range of services. Although UK specific evidence exists we will apply a number of new methods and use a range of new utilisation and control variables. Furthermore, limited research has been done on socioeconomic variation in the association between obesity and health service use.

6.3 Methods

This chapter uses data from the *Health Survey for England* (HSE) and the *British Household Panel Survey* (BHPS). These have been chosen as they are British nationally representative surveys that provide data on obesity, SES and health service use. In the following sections we will present the data and variables separately for HSE and then BHPS. We then present the econometric methods.

6.3.1 Health Survey for England: Data and variables

From the HSE we include data for individuals over the age of 16 for four different combinations of the datasets depending on the health service utilisation variable. We apply the following health service use variables from HSE: GP visits, practice nurse visits, inpatient stays, outpatient visits and a range of medication use categories.

Information on GP visits for all causes is available for four rounds (1999-2002, N=60 091), practice nurse visit data are available for two rounds (2001-2002, N=38 038) and inpatient visits and outpatient visits data are available for two rounds (1999-2000, N=22 053). Lastly, medication use data are available for ten rounds (1999-2008, N=158 703).

Respondents were asked about the number of visits to the GP in the last two weeks. However, less than 3% had more than one visit in the last two weeks. Hence, we look at whether or not the individual had a GP visit as a binary dependent variable taking the value 1 if the person has had GP visits in the last two weeks and 0 otherwise(<16% had one or more visits). This is similar for practice nurse visits (less than 1% had more than one visit and 6% one or more visits) and therefore this variable is also only included as a binary variable.

For inpatient stays and outpatient visits respondents were asked if they had an inpatient stay/outpatient visit in the last year. However, in this case the HSE only provides information on whether or not the individual used inpatient/outpatient

services and not the number of visits. Therefore we include these as binary variables taking the value 1 if the person has made use of inpatient/outpatient services the last year and 0 otherwise.

We also look at current medication use. Information about current medication use was only obtained for individuals who were visited by a nurse during the survey. The years included in this analysis was 1999-2008. HSE has information on type of medication and has categorised them by 12 medication categories (+ other medication taken) based on therapeutic classifications in the British National Formula (BNF), see Appendix 6.1. In addition, HSE records the number of prescribed medications taken.

The obesity measure is based on height and weight obtained during the interviewer visit; it is not self-reported. Obesity is measured as a categorical variable based on four BMI categories, defined according to World Health Organization guidelines (WHO Expert Committee on Physical Status, 1995): normal weight, BMI 18.5-24.9 kg/m^2; overweight, BMI 25-29.9 kg/m^2; class I obesity, BMI 30-34.9 kg/m^2; and, class II/III obesity BMI >35 kg/m^2. As in previous chapters, we combine obesity classes II and III to ensure an adequate number of observations in each cell when we interact BMI and SES. Also we do not combine underweight (BMI <18.5 kg/m^2) and normal weight and include underweight as a separate category, but do not report the results for this group.

SES is a composite measure, and measured in quartiles of a continuous variable based on the linear prediction from an interval regression model of total annual household

income reported in 31 income bands (including an open-ended top category), against a set of individual and household characteristics. The predicted values were equivalised using the McClements household score provided in the HSE to account for household size and composition, and then divided into quartiles. The individual and household characteristics used in the interval regression were: year (ten categories), education qualifications (measured in seven categories); social class of household reference person (HRP; seven categories); cars or van owned by household (four categories); housing tenure (five categories); bedrooms in household (five categories); economic activity (eleven categories); and, whether or not the person was an income support claimant (yes/no). We predict income based on all individuals in the years 1999-2008 and divide this into quartiles. As we use data over the year 1999-2008 it is likely that individuals in the earlier years have a lower income because of general income growth over time. To control for this we recode every individual in the sample to the most recent year before we make the prediction.

To capture regional differences in supply we control for District Health Authority (95 categories) in the health service use regressions. By controlling extensively for health area we capture differences in supply of health services in the area, such as GPs per head, other medical staff per head, distance to hospital or GP practice. There is evidence of PCT level variation in obesity (Moon, Quarendon, Bernard, Twigg & Blyth, 2007). Hence, when we control for health area level data we possibly reduce some of the effect of both obesity and SES on health service use. We have fitted regressions without these variables, finding that the positive impact of obesity and the negative

impact of SES is slightly stronger. However, it does not have an impact on the conclusion of the analysis.

The District Health Authority variable is only available for the years 1999-2002. Hence, it is not available for all the years included in the medication analysis. In this analysis we control for area with the broader area level variable Government Office Region of residence (nine categories). This variable might also pick up regional variation in obesity and SES, we discuss this further below.

In the health service use equations we also include the following covariates: age (as a cubic function); survey year (one dummy for each year depending on the dataset); marital status (six categories); smoking status (four categories); ethnicity (ten categories); and, missing income (yes/no). We also stratify by gender.

6.3.2 British Household Panel Survey: Data and variables

The BHPS is a longitudinal cohort survey of adult members of a nationally representative sample of British households. Respondents are interviewed on a range of topics including their age, socio-economic indicators, health status and a range of health service use indicators. The latest wave with available obesity data was wave 16 collected from 1 September 2006 to April 2007. The survey collects data from adult members of the household. Those in the initial sample are followed until they refuse to participate, die or are lost to follow-up. The present study includes everyone over the age of 16 from wave 14 (1st September 2004 to 11th May 2005) and wave 16 (1st September 2006 to 3rd April 2007), which are the only waves measuring height and

weight. We included respondents who are not observed at every wave (unbalanced panel). The total number of respondents in both waves are 13 194. In addition, there were 2552 individuals who were only included in wave 14 and 1486 individuals who were only included in wave 16. This constitutes a total of 30 426 observations.

From the BHPS we apply the following health service use variables: GP visits, hospital outpatient use, hospital inpatient use, seen a chiropodist, had x-ray of lungs and chest, checked blood pressure, had a cholesterol test and had a blood test. For GP visits and outpatient use the health services use variable was measured by the following question: 'how many times have you made use of *a health service* since the 1st of September the previous year?'. The answers were grouped into 5 groups (none, 1-2, 3-5, 6-10, more than 10). For the remaining health service use variables the use was measured by a binary response to the following question 'have you made use of *a health service* since the 1st of September the previous year?'.

The participants were interviewed in a period from the 1st of September to six to seven months after. This means that the time period differs across the population. We controlled for this by including a time elapsed variable. This variable controls for the time period in days from the 1st of September to the interview date.

BMI is calculated by dividing each respondent's weight in kilograms by their height in meters squared. We constructed a measure of BMI based on the self-reported measures on height and weight obtained during the interview. We divide BMI into the same four categories (+ underweight) as in the HSE data.

SES is based on a continuous variable from a linear prediction from a random effects linear regression model of total annual household income[26], against a set of socioeconomic variables. The predicted values were equivalised using McClements household score provided in the BHPS to account for household size and composition and then divided into quartiles. The variables used in the regression were: year (two waves), education qualifications (six categories); social class of household reference person (HRP; seven categories); car or van owned by household (yes/no); housing tenure (five categories); bedrooms in household (five categories); economic activity status (9 categories); and, whether or not the person was an income support claimant (yes/no).

To capture regional differences in supply we include indicators for Primary Care Trusts (PCT) in England, Local Health Boards (LHB) in Wales, Community Health Partnerships (CHP) in Scotland, and an indicator for individuals who live in Northern Ireland, Chanel Islands, and Isle of Man (216 categories). As mentioned earlier, when we control for health area level data we possibly reduce some of the effect of both obesity and SES on health service use. As in HSE we fit regressions with and without these variables. Generally we find a reduced effect of obesity and a slightly reduced effect of SES across the different types of use when we include them. However, as in HSE it does not alter the conclusion of the analysis.

[26] We use the total household income variable provided in the BHPS. This has been imputed using a regression based imputation technique; see Taylor, 2004 for further information. Hence, we do not have any missing income values.

In the health service use regressions as well as obesity and SES we also included covariates for: age (as a cubic function), survey year, marital status (7 categories), smoking status (yes/no) and ethnicity (9 categories).

6.3.3 Econometric methods

The HSE is a repeated cross-sectional dataset while the BHPS is a panel dataset. Hence, in the following analysis we use both cross-sectional methods and panel data methods. We start by describing the regression models and the estimation strategy, followed by an explanation of the adjustments that have been made to these methods when we conduct the analysis in the BHPS.

6.3.3.1 Regression models

As mentioned above we have three types of dependent variables: binary, banded-count and count variables. When we have binary dependent variables we use nonlinear binary dependent variable models instead of the Linear Probability Model (LPM) due to the well known shortcomings of the LPM[27]. We chose the probit model over the logit[28] model as economists have a tendency to prefer the normality assumption of the error term in the probit model (see, e.g., Wooldrige, 2002; Greene, 2010).

[27] Most importantly the LPM does not restrict its predicted probabilities to a value between 0 and 1. In addition, and related to the previous problem, the partial effects are constant, which can provide unlikely values at extreme values of continuous variables.

[28] The main difference between the logit and probit models lies in the distribution of the errors; however they tend to provide very similar predicted probabilities.

For some types of use (medication use in the HSE; GP and outpatient visits in the BHPS) we have count data on the number of visits/drugs used. In the BHPS the variables are *grouped* count data and to account for this we use interval regression methods. In the HSE the variables are not grouped and we use a negative binomial (negbin) count model. The negbin model is chosen over the more common Poisson count model as we have a large proportion of zeros, which could cause the conditional mean variance of the dependent variable to differ from the conditional variance. The negbin model accounts for this potential problem by allowing for a variance that is greater than its mean (Jones, 2005).

What follows is a two-part model. "The decision to contact a physician and the decision about how often to contact a physician are determined by different decision makers" (Pohlmeier & Ulrich, 1995). Hence, in the case of GP/outpatient visits, but not medication use, we treat the decision process as consisting of two separate probability functions. We analyse the decision to use or not to use the health service by a binary probit model as above, $P(GP=1)$. We fit a model for the subsample of individuals who have used the type of health service at least once, $E(GP|GP>0)$. This model investigates the number of GP/outpatient visits.

We do not treat the number of medications used as a decision consisting of two parts. In this case we do not see the decision about *use* and *frequency* of medications to consist of different decision makers.

6.3.3.2 Interpretation of the output

In this chapter we use both linear and nonlinear models. As we discussed in the literature review, the outputs of nonlinear models are often interpreted in two ways. The first is the *absolute effect* and is the effect of the independent variables on *E(y)*. The second is the *relative effect*, which is reported in the form of risk ratios or relative difference.

Whether one should focus on absolute or relative effects depends on the question the researcher wants to answer. In this research we want to quantify the consequences of the findings and focus on absolute values. The absolute effect is the coefficient in a linear regression model. However, the interpretation of the coefficient change when one uses a nonlinear model. This has consequences for the interpretation of the interaction terms, which we explain below.

Assuming we have two SES groups high and low and we want to measure heterogeneities in the impact of obesity on a dependent variable (m_i). We can estimate the following linear model:

$$m_i = \beta_0 + \beta_1 O_i + \beta_2 y_i + \beta_3 O y_i + u_i \quad (6.4)$$

Where y_i is a dummy taking the value of one if the individual is high SES, O_i is a dummy taking the value one if the individual is obese and the Oy_i is an interaction between high SES and obese. The β's are parameters to be estimated and u_i is the error term. If the model is linear the hypothesis test of the interaction term (β_3=0) would be a test of

whether or not there are absolute differences in the impact of obesity on health and health service use. If the models are nonlinear the hypothesis tests of the interaction term (β_3=0) would be a test of whether or not there are interactions in the impact of O and Y on the latent variable m*. The latent variables are unbound unobserved variables and the hypotheses about these latent variables are about whether or not there are relative interactions in the models. The marginal effect of each variable in a nonlinear model is the cross partial derivative. In contrast to the linear model the marginal effect changes depending on the values of the other covariates, even in models without interaction terms.

Clarity about whether one focuses on absolute or the relative effects is especially relevant in our case where we seek to compare the effect of obesity across subgroups. One could end up in a situation where the relative interaction effect and the absolute interaction effect tell different stories and have different signs. See, e.g., Ai & Norton (2003); Greene (2010); and Berry, DeMerit & Esarey (2010) for further discussions.

Hence, in a linear model we focus on the interaction term. While in a nonlinear model we need an alternative method for establishing whether or not there are different consequences of obesity across SES groups. There has been much debate around how to interpret interactions in nonlinear models when one wants to interpret the interactions as marginal effects (Ai & Norton, 2003; Norton et al., 2004; Greene, 2010; Kolasinski and Siegel, 2010; Huang & Shields, 2000; Karaca-Mandic, Norton & Dowd, 2011; Berry, DeMerit & Esarey, 2010). We follow an approach by Greene (2010) who argues that hypothesis tests about interactions and other effects in nonlinear models

should be about the model coefficients and about the structural aspects of the model specifications. Marginal effects are neither coefficients nor elements of specification of the model they are implications of the specified model. Greene (2010) suggests investigating interactions in nonlinear models in two steps:

1. Build the latent variable model based on appropriate statistical procedures. Conduct hypothesis testing about the model coefficients and interactions at this stage. Hypothesis tests are about the model coefficients and the structural aspects of the model specifications.
2. Use the model to inform the reader about partial effects using graphical presentations, with no need for any testing in stage two.

We follow these steps although we do not provide graphical interpretations. Instead we provide tables of predicted means across groups.

6.3.3.3 Estimation strategy

We regress the health service use variables against the BMI categories and SES quartiles using probit/count/interval regression models for cross sectional data including interaction terms between BMI and SES categories, with normal weight and SES quartile 1 (most deprived) as the omitted categories. Hence, we fit the following latent variable model:

$$m_i = \beta_0 + \beta_1 \sum_{a=1}^{4} SES_{ai} + \beta_2 \sum_{b=1}^{5} BMI_{bi} + \beta_3 \sum_{a=1}^{4} \sum_{b=1}^{5} SES_{ai} * BMI_{bi} + \beta_4 Z_i + u_i \qquad (6.5)$$

Where health service use (m_i) is a function of SES, the BMI groups, interactions between BMI and SES, and other variables (Z_i), u_i is the error term and i indexes the individuals. The β's are vectors of coefficients to be estimated

When we follow the approach suggested by Greene (2010) we conduct our analysis in two steps. We start with step one where we decide upon the functional form that best fits our data. To decide upon the variable specification we follow previous research by Birch, Jerret & Eyles (1997 & 2000) and Birch et al., 2005 and use a stepwise backward elimination process, which is a process where one starts with a full model including all variables and interactions and then remove variables until we are left with a model that only includes significant variables. We use the likelihood ratio[29] test as criteria for removal. We run separate models for each gender controlling for individual and household characteristics.

When we have specified the model we move on to the second step, which is to estimate predicted mean health service use across SES and BMI groups, based on the specified model.

When we have binary dependent variables we want to estimate the predicted probability of use of the health services *P(m=1)* and compute predicted mean health service use for each BMI or BMI/SES category after every regression based on specified

[29] The likelihood in a clustered analysis is not a true likelihood, i.e. it is not the distribution of the sample. When clustering is used individual observations are no longer independent, thus the likelihood-ratio test should not be used, and Wald tests should be used instead (Sribney, 2005). Hence, in the analysis where we clustering by PSU we base our selection of the variables on Wald tests only.

latent variable models. We fix the other covariates at the whole sample mean values. Hence, the variation in the predicted mean health service use is a function of the impact of obesity on probability of health service use and how this varies by SES, and is not affected by the values of the covariates, which vary between individuals.

In the interval regression models and negbin models we compute predicted mean number of consultations (*E(m)*) fixing the covariates at the whole-sample mean values in the same way as explained for the probit models above.

To maximise the sample size, we included individuals in the models with missing income data for whom we could compute a predicted income value. In the models we included an indicator for missing income (yes/no), to control for the possibility that income may not be missing at random.

We apply survey weights reported in the HSE to each observation. This adjusts for the fact that different observations have different probabilities of selection and participation in the survey. Medication use data was obtained only for those who were visited by a nurse during the survey; hence we apply nurse weights reported in the HSE. This adjusts for the non response bias to the nurse section of the survey. It is also possible that, due to the sampling strategy used in the HSE, observations are independent across Primary Sampling Units (PSUs), but not within PSUs. If this is the case then if we use estimators that assume independence within these clusters the standard errors on our regression coefficients will be too small and we will overestimate the statistical significance of the independent variables in our models.

We control for clustered sampling within PSUs using unique PSU/year identifiers that produce Huber/White/sandwich robust variance estimators that allow for within-group dependence (Kish & Frankel, 1974).

P values below the 5% level (t statistics higher than ≈1.9) are regarded as statistically significant.

6.3.3.4 Panel data adjustment in BHPS

In the following part we explain the changes made to account for the BHPS being a panel dataset. There are essentially two panel data methods: fixed effects and random effects. In cases where key variables do not vary sufficiently over time we are forced to use random effects[30] (Wooldrige, 2002). In our case we have explanatory variables like SES groups and BMI groups that do not vary sufficiently over time, the reason being that we only have two time periods. Alternatively, we could use a pooled panel model. However, not only does the pooled panel model have problems with the error terms being serially correlated, the random effects model removes a fraction of the time invariant heterogeneity (Wooldrige, 2002).

We ran the models in the previous paragraph using random effects binary probit and random effects interval regression models:

$$m_i = \beta_0 + \beta_1 \sum_{a=1}^{4} SES_{ait} + \beta_2 \sum_{b=1}^{5} BMI_{bit} + \beta_3 \sum_{a=1}^{4} \sum_{b=1}^{5} SES_{ait} * BMI_{bit} + \beta_4 Z_{it} + \beta_5 W_t + v_{it} \quad (6.6)$$

[30] We also recognise that there are more advanced random effects models like Mundlak and Chamberlain. However, these models do also need sufficient variation in the dependent variables over time.

Where *t = 1, 2* denotes the time period for individual *i*. W_t is a dummy variable taking the value 1 if *t=2* and zero otherwise. It does not change across individuals so it does not have an *i* subscript. The error term in a random effects model (v_{it}) consists of a_i, which captures unobservable individual heterogeneity that does not vary over time, hence it does not have a t subscript. u_{it} is a random error that captures time variant errors. In random effects probit models one assumes that u_{it} is $IN(0,\sigma_u^2)$. To be able to compute marginal effects and predicted means one must make an assumption about the distribution of the unobserved effect a_i. We follow Wooldrige (2002) & Arulampalam (1999) and assume that the a_i's are $IN(0,\sigma_a^2)$ [31]. Further, one needs to make an assumption about the distribution of the error term v_{it}. We follow Arulampalam (1999) who showed that to calculate the predicted probabilities one must calculate the marginal effect on the latent variable and multiply this with the variance in the error term $\sqrt{1-\rho}$.

Table 6.3 shows a summary of the 14 models we run using BHPS and HSE. Each of these models is fitted separately for men and women.

[31] Based on this; the correlation between error terms for the same individual across time is: $\rho = cor(v_{it}, v_{it-1}) = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_u^2}$

Table 6.3: Summary of regression models for analysing the association between obesity and health service use

Number	Dataset	Dependent variables	Regression method	With or without interactions between BMI and SES groups	Covariates
1	BHPS (2004/5 and 2006/7)	Binary: GP visits; outpatient visit; inpatient stay; chiropodist; x-ray of chest and lungs; had blood pressure taken; had a cholesterol test; Had a blood test	Random effects probit models	No interactions	Individual and household characteristics including health area variables
2	BHPS (2004/5 and 2006/7)	Binary: GP visits; outpatient visit; inpatient stay; chiropodist; x-ray of chest and lungs; had blood pressure taken; had a cholesterol test; Had a blood test	Random effects probit models	BMI groups interacted with SES groups	Individual and household characteristics including health area variables
3	BHPS (2004/5 and 2006/7)	Banded: Number of GP visit; number of outpatient	Random effects interval regression models	No interactions	Individual and household characteristics including health area variables
4	BHPS (2004/5 and 2006/7)	Banded: Number of GP visit; number of outpatient	Random effects interval regression models	BMI groups interacted with SES groups	Individual and household characteristics including health area variables
5	HSE (1999-2002)	Binary: GP visits	Probit models for cross sectional data	No interactions	Individual and household characteristics including health area variables
6	HSE (1999-2002)	Binary: GP visits	Probit models for cross sectional data	BMI groups interacted with SES groups	Individual and household characteristics including health area variables
7	HSE 2001-2002)	Binary: Practice nurse visits	Probit models for cross sectional data	No interactions	Individual and household characteristics including health area variables
8	HSE (2001-2002)	Binary: Practice nurse visits	Probit models for cross sectional data	BMI groups interacted with SES groups	Individual and household characteristics including health area variables
9	HSE (1999-2000)	Binary: Inpatient stays and outpatient visits	Probit models for cross sectional data	No interactions	Individual and household characteristics including health area variables
10	HSE (1999-2000)	Binary: Inpatient stays and outpatient visits	Probit models for cross sectional data	BMI groups interacted with SES groups	Individual and household characteristics including health area variables
11	HSE (1999-2008)	Binary: 12 medication categories	Probit models for cross sectional data	No interactions	Individual and household characteristics
12	HSE (1999-2008)	Binary: 12 medication categories	Probit models for cross sectional data	BMI groups interacted with SES groups	Individual and household characteristics
13	HSE (1999-2008)	Count: Number of prescribed medications	Negative binomial count model	No interactions	Individual and household characteristics
14	HSE (1999-2008)	Count: Number of prescribed medications	Negative binomial count model	BMI groups interacted with SES groups	Individual and household characteristics

As mentioned earlier, the advantage of the analysis in BHPS is that we apply panel models hence it can to some extent account for time invariant heterogeneity bias and serial correlation of the error terms. In addition, the data is more recent (in all cases

except for medication use) and we have good data on health area variables. However, in HSE we have nurse measured height and weight, which is less likely to be biased compared to the self-reported height and weight in BHPS.

6.4 Results

We present the results in two parts first the results of the cross-sectional models in the HSE and then of the panel models in BHPS. We then compare the results across the datasets.

6.4.1 Health Survey for England: Results

Table 6.4 summarises the HSE data and shows the total number of individuals who were over the age of 16, had valid health service use data and BMI values across the three versions of the dataset. For data used for the GP visits analysis (years: 1999-2002) 36 766 individuals had valid data, while 22 645 had valid data for the practice nurse analysis (years: 2001-2002) and 17 834 were included in the outpatient and inpatient use (years: 1999-2000) analysis. The mean share of individuals who had a GP visit in the last two weeks across the whole sample was 16% increasing from a mean of 15% in the normal weight category to 22% in the class II/III obesity category. Similar trends were found for practice nurse visits with 5% in the normal weight category and 9% in the class II/III obesity category. For outpatient visits the last year equivalent numbers were 30% and 39%, respectively. For the inpatient stays the last year similar numbers were 7% and 8%, respectively.

Table 6.4: summary statistics for the three versions of the HSE

	Whole sample	Normal weight	Overweight	Class I Obesity	Class II/III obesity
GP visit last two weeks (1999-2002)					
Observations					
Number (%)	36766 (100)	14649 (40)	13762 (37)	5489 (15)	2198 (6)
GP visit %	16	15	16	18	22
Practice nurse visit last two weeks (2001-2002)					
Observations					
Number	22645 (100)	9059 (40)	8349 (37)	3386 (15)	1410 (6)
Practice nurse visit %	6	5	7	8	9
Inpatient and outpatient visits last year (1999-2000)					
Observations					
Number	17834 (100)	6835 (38)	6770 (38)	2810 (16)	1123 (6)
Outpatient %	32	30	32	35	39
Inpatient %	8	7	8	9	8

Note

The years from the HSE included in each analysis varies by the dependent variable. Other than for "Observations" all statistics are percentages

In Appendix 6.2 we present summary statistics of all the variables for the years 1999-2002 across the BMI groups. We see that as in previous chapters the mean predicted SES measure was highest in the overweight category and lowest in class II/III obesity category; the modal SES quartile was 4 (least deprived) in the normal weight category and quartile 1 (most deprived) in the obese class II/III category. Forty six percent of the sample was men and the mean age was 46 years.

In the medication use analysis we pool ten years and the total number of respondents was 158 703, of these 67 041 were over the age of 16, had valid medication use data (had a nurse visit), and BMI values. The numbers of observations in each BMI category are in Table 6.5. Summary statistics of the variables are in Appendix 6.3. Thirty seven percent of the individuals in the sample were normal weight, 38% (the modal category) were overweight and 24% were obese (17% in class I and 7% in class II/III). The mean number of medications used per person across the whole sample was 1.43, increasing from a mean number of 0.99 in the normal weight category to 2.45 in the class II/III

obesity. The mean share of individuals who use medication across the whole sample was 46% increasing from a mean of 37% in the normal weight category to 63% in the class II/III obesity. The mean predicted SES measure was highest in the overweight category and lowest in class II/III obesity category; the modal SES quartile was 4 (least deprived) in the normal weight category and quartile 1 (most deprived) in the obese class II/III category. Forty six percent of the sample was men and the mean age was 48 years.

Table 6.5: Summary statistics of medication use across each BMI category (HSE: 1999-2008)

	Whole sample	Normal Weight	Overweight	Class I obesity	Class II/III Obesity
Medication (1999-2008)					
Observations					
Number (%)	66435 (100)	24313 (37)	25489 (38)	11040 (17)	4626 (7)
Medication categories					
Cardio-vascular	22	13	24	33	38
Gastrointestinal	9	6	9	12	13
Respiratory	10	9	9	10	14
CNS	15	11	15	20	26
Endocrine	11	8	11	14	19
Musculoskeletal	7	4	6	10	13
Medication excluding contraceptives	46	37	48	56	63
Number of medications	1.4	1	1.5	2	2.5

Note

Other than for "Observations" all statistics are either mean (for number of medications) or % (all other variables). Number of medications are total number of medications not only the number of categories.

Table 6.6 shows the results of the interval regression of total annual household income against the SES indicators using HSE data. The variables have the expected signs. Higher education, higher social class of HRP, owning a car or van and bedrooms in household are all positively correlated with income. Compared to being in paid employment all other types of economic activity have a negative effect on income. Claiming income support is negatively correlated with income. We used these coefficients to generate our predicted SES measures. Respondents were categorised into quartiles based on the predicted SES values. The range of values for equivalised

predicted income were -94,805[32] to 16,2396; the cut-points used to set the quartiles were: 18,692; 28,243; and, 39,655.

[32] We convert the predicted income into quartiles and use these as our SES measure and do not use the actual income values in the further analysis. Hence, it is not a problem that we obtain negative predicted income values

Table 6.6: Interval regression of total annual household income against SES indicators in HSE (HSE: 1999-2008)

	Coef.	z
Educational qualifications		
Degree or equivalent	Base category	
Higher education below degree	-9629.7	-24.5
NVQ3/GCE A Level or equivalent	-8095.01	-20.25
NVQ2/GCE O Level or equivalent	-10075.4	-27.18
NVQ1/CSE other grade or equivalent	-11708.5	-25.71
Foreign/other	-10792.8	-20.93
No qualification	-11416.1	-29.97
Social Class of HRP		
Professional	Base category	
Managerial technical	-4278.77	-7.89
Skilled non-manual	-9578.07	-17.28
Skilled manual	-11772.2	-21.32
Manual unskilled	-12141.6	-20.34
Not described SES	-9611.56	-4.83
Missing SES	-9598.16	-14.04
Cars owned by household		
Household has no car	Base category	
One car	1273.366	4.77
Two cars	9716.553	25.73
Three or more cars	19546.32	25.6
Economic activity status for last week		
In paid employment or self-employed	Base category	
Going to school or college full time	-12474.6	-11.59
On a government training scheme	-15741.4	-11.81
Doing unpaid work	-5255.6	-1.4
Waiting to take up paid work already obtained	-8964.46	-3.72
Looking for paid work or a government training scheme	-15227.8	-29.35
Intending to look for work but prevented by temporary sickness	-14834.2	-11.64
Permanently unable to work because of long-term sickness	-11953.9	-29.01
Retired from paid work	-11211.2	-33.83
Looking after home or family	-10718.8	-21.48
Doing something else	-16765.5	-14
Housing tenure		
Own outright	Base category	
Buying with help of a mortgage or loan	4178.557	11.63
Pay part rent and part mortgage	-2922.75	-2.99
Rent	-1910.84	-5.47
Live rent free and/or squatting	-5286.34	-5.6
Bedrooms in household		
One	Base category	
Two	510.8716	1.63
Three	2263.775	6.86
Four	11496.58	23.13
Five or more	25732.5	24.49
Income support claimant		
No	Base category	
Yes	-2347.2	-7.06
Percent within each year		
1999	-8052.89	-16.03
2000	-7142.33	-13.65
2001	-6125.48	-13.02
2002	-4814.98	-8.96
2003	-4034.08	-7.91
2004	-3292.6	-4.71
2005	-2685.33	-4.37
2006	-2014.44	-3.84
2007	-629.859	-0.93
2008	Base category	
Constant	45347.65	53.5
Observations	84414	
Adjusted R^2	0.076	

Notes

HRP = household reference person

Appendix 6.5 shows the latent variable models before we have excluded any variables for GP visit, practice nurse visits, inpatient and outpatient visits in HSE. The tables show the coefficients, the z values and joint tests for interactions. In general, across the models the effects of BMI and SES are significant; however the joint tests for the interactions are not. In appendix 6.6 we see the refitted models after we have conducted the stepwise backward elimination procedure. These are the models we base the predictions on.

The predicted probabilities of health service use across each BMI category can be found in Table 6.7. We find that obesity increases the probability of having a GP visit for both men and women. Furthermore, we find that obese and overweight men have an increased probability of a practice nurse visit; however, there is no variation in outpatient visits across BMI groups for men. For overweight and obese women, there is no increase in use of practice nurses, however we see an increased probability of an outpatient visit and inpatient stays. The variation in probability of use is more pronounced in women than in men.

Table 6.7: Predicted probability of health service use across BMI categories

	Men				Women			
	Normal weight	Over-weight	Obese Class I	Obese Class II/III	Normal Weight	Over-weight	Obese Class I	Obese Class II/III
GP visit	0.127	0.130	0.141	0.187	0.169	0.190	0.200	0.232
Practice nurse visit	0.037	0.042	0.048	0.082	0.069	"	"	"
Outpatient visit	0.318	"	"	"	0.288	0.297	0.351	0.392
Inpatient stay	0.066	"	"	"	0.078	0.092	0.124	0.098

Note

" Predicted probability equal to the normal weight category

The predictions are based on probit models using the HSE. The dependent variables are the health service use variables. The mean values are predicted probability of health service use for each BMI or BMI/SES category computed by setting the individual and household characteristics to their whole-sample mean values. The individual and household characteristics are age, age squared and age cubed, ethnicity, marital status, smoking status, health area variables, survey year, time elapsed and missing income.

Table 6.8 also shows the predicted probabilities of using health services across each SES quartile. Across all health services for both men and women lower SES groups use the health service more than higher SES groups. There is one exception for practice nurse visits for women, where we do not find any SES variation. In general, the SES gradient in use seems to be somewhat more pronounced in men. We see similar trends within each of the quartiles, in that higher levels of obesity is associated with higher predicted probability of use.

There were no clear SES variation in the association between obesity and health service use. However, in women, the increase in GP visits in the last two weeks for overweight and obese were somewhat more pronounced in the lower SES groups compared with the higher SES groups. For example, the predicted mean share of individuals who have had a GP visit in the lowest SES quartile 1 in women increased from a predicted mean of 20% in the normal weight category to 26% in the class II/III obesity, which is a discrete difference of 6%. For the least deprived SES quartile 4 in women similar numbers were 15% and 20%, which constitutes a discrete difference of 5%. For GP visits in men there was no clear gradient however the increase in visits was more pronounced in least deprived SES quartile 4 than in the most deprived SES quartile 1.

Table 6.8: Predicted probability of health service use across BMI and SES categories

	Men				Women			
	Normal weight	Over-Weight	Obese Class I	Obese Class II/III	Normal weight	Over-weight	Obese Class I	Obese Class II/III
GP visit								
1 (most deprived)	0.180	0.180	0.183	0.236	0.203	0.224	0.231	0.263
2	0.136	0.136	0.138	0.183	0.169	0.187	0.194	0.222
3	0.125	0.125	0.127	0.169	0.160	0.178	0.184	0.212
4 (least deprived)	0.092	0.114	0.142	0.171	0.151	0.168	0.175	0.201
Practice nurse visit								
1 (most deprived)	0.042	0.048	0.054	0.091	0.085	"	"	"
2	0.044	0.049	0.056	0.093	0.063	"	"	"
3	0.034	0.039	0.044	0.076	0.063	"	"	"
4 (least deprived)	0.029	0.033	0.038	0.066	0.068	"	"	"
Outpatient visit								
1 (most deprived)	0.370	"	"	"	0.288	0.297	0.351	0.392
2	0.324	"	"	"	"	"	"	"
3	0.312	"	"	"	"	"	"	"
4 (least deprived)	0.268	"	"	"	"	"	"	"
Inpatient stay								
1 (most deprived)	0.092	"	"	"	0.097	0.112	0.146	0.115
2	0.067	"	"	"	0.065	0.077	0.103	0.079
3	0.056	"	"	"	0.079	0.093	0.122	0.095
4 (least deprived)	0.053	"	"	"	0.072	0.084	0.112	0.087

Note

The predictions are based on probit models using the HSE. The dependent variables are the health service use variables. The mean values are predicted probability of health service use for each BMI or BMI/SES category computed by setting the individual and household characteristics to their whole-sample mean values. The individual and household characteristics are age, age squared and age cubed, ethnicity, marital status, smoking status, health area variables, survey year, time elapsed and missing income.

Appendix 6.7 shows the latent variable models before we have excluded any variables for medication use visits in HSE. The tables show the coefficients the, z values and joint tests for interactions. In appendix 6.8 we see the refitted models after backward elimination procedure. These are the models we base the predictions on.

The predicted probabilities of medication use across each BMI group can be found in Table 6.9. The probability of using cardio-vascular medicine, gastrointestinal medicine, respiratory medicine, CNS medicine, endocrine medicine, musculoskeletal medicine increased for overweight and obese men and women compared with normal weight. For both genders CNS medicine was the modal type of medication in the normal weight group, while cardio vascular medicine was the modal type of medicine in the

obese class II/III category. On average normal weight men used 0.65 medications while the in the overweight, obese class I and obese class II/III categories the mean number of medications used were 0.74, 1.01 and 1.46, respectively. In women, analogous figures were 0.89, 1.06, 1.38 and 1.85.

Table 6.9: Predicted probability of medication use across BMI categories

	Men				Women			
	Normal weight	Over-weight	Obese Class I	Obese Class II and III	Normal weight	Over-weight	Obese Class I	Obese Class II and III
Cardio-vascular medicine	0.057	0.096	0.157	0.257	0.071	0.105	0.169	0.254
Gastrointestinal medicine	0.046	0.049	0.057	0.071	0.049	0.058	0.083	0.093
Respiratory medicine	0.083	0.082	0.085	0.108	0.082	0.090	0.107	0.142
CNS medicine	0.073	0.084	0.106	0.154	0.121	0.151	0.184	0.240
Endocrine medicine	0.022	0.030	0.051	0.097	0.092	0.102	0.115	0.151
Musculoskeletal medicine	0.026	0.033	0.057	0.073	0.035	0.045	0.070	0.095
Average number of prescribed medicines taken	0.649	0.739	1.013	1.457	0.889	1.063	1.377	1.853

Note

The predictions are based on negbin and probit models using the HSE. The dependent variables are the medication categories. The mean values are predicted probability of medication use for each BMI or BMI/SES category computed by setting the individual and household characteristics to their whole-sample mean values. The individual and household characteristics are age, age squared and age cubed, ethnicity, marital status, smoking status, area variables, survey year, time elapsed and missing income.

In Table 6.10 we see that similar trends were found in every SES quartile for both men and women, where the predicted probability of using medication was higher in the obesity categories than in the normal weight category

More deprived SES groups have higher probability of using each of the medication types across all BMI categories for both men and women. Except for endocrine medicine for women where there was no variation in use across SES groups. When we look at the variable for all types of medication except from contraceptives we see that both men and women have a significantly higher predicted mean number of medications prescribed across all BMI categories.

For medication use we identify a SES gradient in the association between obesity and use in both men and women. Compared with normal weight overweight and obese men and women in the lower SES groups had a more pronounced effect on number of medications used than in the higher SES groups. For example, the predicted mean number of medications used in the lowest SES quartile 1 in men increased from a predicted mean of mean of 0.91 in the normal weight category to 1.97 in the class II/III obesity, which is a discrete difference of 1.06. For the least deprived SES quartile 4 in men similar numbers were 0.5 and 1.08, which constitutes a discrete difference of 0.58. For women, although the difference is less pronounced, we identify the same gradient, in which lower SES women have a more pronounced association between obesity and number of medications used than higher SES women.

We discussed how the absolute and the relative interaction effects could differ. We see that there is no SES variation in the relative difference in the use of medications. For example, the relative difference in medications used in the lowest SES quartile 1 between normal weight and obesity class II/III obesity is 1.967/0.912=2.16. For the least deprived SES quartile 4 in men similar numbers were 1.076/0.499=2.16. The interpretation of these results is that low SES individuals use more medications than high SES individuals when they are normal weight. When they become obese the difference in number of medications used increases. Hence obesity costs more in terms of medications used in the lower SES group. However, the relative difference between the SES groups does not differ.

Table 6.10: Predicted probability of medication use across BMI and SES categories

	Men				Women			
	Normal weight	Over-weight	Obese Class I	Obese Class II and III	Normal weight	Over-weight	Obese Class I	Obese Class II and III
Cardio-vascular medicine								
All	0.057	0.096	0.157	0.257	0.071	0.105	0.169	0.254
1 (most deprived)	0.078	0.127	0.199	0.311	0.098	0.145	0.189	0.270
2	0.060	0.101	0.163	0.265	0.082	0.109	0.186	0.280
3	0.050	0.085	0.141	0.235	0.063	0.091	0.178	0.249
4 (least deprived)	0.045	0.078	0.130	0.220	0.047	0.078	0.125	0.221
Gastrointestinal medicine								
All	0.046	0.049	0.057	0.071	0.049	0.058	0.083	0.093
1 (most deprived)	0.068	0.071	0.081	0.099	0.065	0.076	0.105	0.116
2	0.049	0.052	0.060	0.074	0.050	0.059	0.083	0.092
3	0.041	0.044	0.051	0.063	0.045	0.053	0.076	0.084
4 (least deprived)	0.034	0.036	0.042	0.053	0.038	0.045	0.065	0.073
Respiratory medicine								
All	0.083	0.082	0.085	0.108	0.082	0.090	0.107	0.142
1 (most deprived)	0.099	0.101	0.127	0.180	0.100	0.109	0.128	0.167
2	0.084	0.085	0.077	0.117	0.084	0.092	0.109	0.143
3	0.075	0.075	0.074	0.063	0.075	0.082	0.098	0.130
4 (least deprived)	0.077	0.071	0.069	0.078	0.073	0.080	0.095	0.126
CNS medicine								
All	0.073	0.084	0.106	0.154	0.121	0.151	0.184	0.240
1 (most deprived)	0.147	0.166	0.200	0.266	0.198	0.222	0.264	0.328
2	0.077	0.089	0.112	0.158	0.154	0.175	0.213	0.270
3	0.058	0.068	0.086	0.125	0.122	0.161	0.164	0.233
4 (least deprived)	0.041	0.049	0.063	0.094	0.094	0.124	0.170	0.183
Endocrine medicine								
All	0.022	0.030	0.051	0.097	0.092	0.102	0.115	0.151
1 (most deprived)	0.027	0.037	0.061	0.115	"	"	"	"
2	0.025	0.034	0.057	0.108	"	"	"	"
3	0.020	0.027	0.046	0.089	"	"	"	"
4 (least deprived)	0.017	0.024	0.041	0.081	"	"	"	"
Musculoskeletal medicine								
All	0.026	0.033	0.057	0.073	0.035	0.045	0.070	0.095
1 (most deprived)	0.037	0.047	0.077	0.097	0.042	0.054	0.082	0.109
2	0.028	0.036	0.061	0.078	0.038	0.050	0.076	0.101
3	0.024	0.031	0.053	0.069	0.033	0.044	0.067	0.091
4 (least deprived)	0.018	0.023	0.041	0.053	0.026	0.034	0.054	0.074
Average number of prescribed medicines taken								
All	0.649	0.739	1.013	1.457	0.889	1.063	1.377	1.853
1 (most deprived)	0.912	1.050	1.410	1.967	1.113	1.312	1.664	2.192
2	0.675	0.777	1.043	1.456	0.935	1.102	1.398	1.841
3	0.562	0.647	0.869	1.212	0.808	0.952	1.207	1.591
4 (least deprived)	0.499	0.575	0.771	1.076	0.753	0.887	1.125	1.482

Note

The predictions are based on negbin and probit models using the HSE. The dependent variables are the medication categories. The mean values are predicted probability of medication use for each BMI or BMI/SES category computed by setting the individual and household characteristics to their whole-sample mean values. The individual and household characteristics are age, age squared and age cubed, ethnicity, marital status, smoking status, area variables, survey year, time elapsed and missing income.

Due to low use of some of the services the results would be sensitive to outliers and other random variations. We exclude the results for the services for which the predicted probability of use was below 10% for all BMI/SES groups. This meant that we

excluded the results for: gynae/urinary medicine; cytotoxic medicine; medicine for nutrition/blood; eye/ear etc.; infection; medicine for skin; other medicine; and, day case visits.

6.4.2 British Household Panel Survey: Results

The total number of observations in BHPS across the two periods was 30 426, of these 20 824 had valid health service use data, were over the age of 16 and had valid BMI data.

Descriptive statistics of the utilisation variables are in Table 6.11 and for the rest of the variables in Appendix 6.4. Forty one percent of the sample were normal weight (the modal category), 38% were overweight and 19% were obese (14% class I obesity, 5% class II/III obesity). Across the whole sample 74% had a GP visit, 41% had an outpatient visit, 10% had an inpatient stay, 10% had seen a chiropodist, 14% had x-ray of chest and lungs, 52% had their blood pressure checked, 23% had a cholesterol test, and 41% had a blood test. The percentage having used the different health services increased for overweight and obese individuals across all health services. The mean share of individuals who had outpatient and inpatient visits in HSE were 32% and 8%, respectively, which is lower than in the BHPS where it was 41% and 10%, respectively. Note that the year of measurement differed between the surveys. The mean predicted SES measure was highest in the overweight category and lowest in the class II/III obesity category; the modal SES quartile was quartile 4 (least deprived) in the normal weight category and quartile 1 (most deprived) in class I and II/III obesity categories. Forty seven percent of the sample was men and the mean age was 47 years.

Table 6.11: Summary statistics of health service use across each BMI category in BHPS (Waves: 2004/5 and 2006/7)

	Whole sample	Normal weight	Overweight	Class I obesity	Class II/III obesity
Health service use BHPS					
Observations					
Number	20776 (100)	8515 (41)	7844 (38)	2900 (14)	1086 (5)
Health service use variables					
GP visit	74	71	74	78	82
Outpatient visit	41	39	42	45	53
Inpatient stay	10	10	10	11	14
Chiropodist	10	8	10	14	17
X-ray of chest and lungs	14	12	15	15	18
Blood pressure	52	45	54	60	65
Cholesterol test	23	16	27	31	35
Blood test	41	34	43	48	55

Note

Other than for "Observations" all statistics are % (categorical variables)

Table 6.12 shows the results for the regression model of total household income against the individual and household socioeconomic variables. Education, social class of HRP, owning a car, and number of bedrooms in household are all positively correlated with income. Compared to being employed other types of economic activity have a negative effect on income, except for maternity leave which is not significant. Claiming income support is negatively correlated with income. We used these coefficients to generate our predicted SES measure. Respondents were categorised into quartiles based on the predicted SES values. The range of the values for the equivalised predicted income were -3,387[33] to 86,065; the cut points used to set the quartiles were 19,823, 26,847 and 34,774.

[33] As earlier we convert the predicted income into quartiles and use these as our SES measure and do not use the actual income values in the further analysis. Hence, it is not a problem that we obtain negative predicted income values.

Table 6.12: Random effects regression of total annual household income against SES indicators in the BHPS

	Coef.	z
Degree or equivalent	Base category	
NVQ3/GCE A Level or equivalent	-1485.67	-2.44
NVQ2/GCE O Level or equivalent	-3356.48	-6.7
NVQ1/CSE other grade or equivalent	-4868.89	-11.25
No qualification	-5012.46	-8.85
Social Class of HRP		
Professional	Base category	
Managerial technical	-3540.44	-4.24
Non manual labour	-7347.77	-8.12
Manual skilled	-8334.93	-9.84
Partly skilled	-8846.56	-9.38
Manual unskilled	-9251.66	-10.35
Never had a job	-8874.88	-7.58
Missing social class	-7318.11	-8.48
Car owned by household		
Owns car	Base category	
No car	-1267	-4.16
Housing tenure		
Own outright	Base category	
Shared	-2947.66	-1.33
Rent	-7473.69	-11.66
Live rent free and/or squatting	-7425	-6.28
Other accommodation	-6751.6	-2.79
Bedrooms in household		
One	Base category	
Two	2652.567	1.86
Three	4608.093	3.2
Four	7256.092	4.88
Five or more	13838.63	8.8
Income support		
Yes	-1394.22	-3.33
No	Base category	
Economic activity		
Employed	Base category	
Unemployed	-5584.63	-9.44
Retired	-14690.2	-27.09
Maternity leave	846.0586	0.6
Family care	-6494.85	-9.75
FT studt, school	-2640.05	-4.03
LT sick, disabld	-9440.1	-20.06
Gvt trng scheme	-5478.29	-3.48
Other economic activity	(omitted)	
Wave		
Wave 16	-1638.37	-5.31
_cons	38661.22	24.06
Observations	28494	
Test for random effects: Var(u) = 0	0.0000	

Note

The test for random effect is a *Breusch and Pagan Lagrangian multiplier test for random effects* a significant results show that u is not zero and a pooled OLS will provide biased estimates.

Appendix 6.9 shows the coefficients and z values of the probit and interval regression models fitted with interactions between BMI groups and SES groups in BHPS. Based on the test for joint significance of the interactions we refit the models with or without interactions. There were significant interactions between BMI and SES in the interval regression model for GP visits for women. The interval regression model is a linear

model hence the significant interactions mean that there is a significant difference in the predicted mean utilisation and that the impact of obesity is more pronounced in lower SES women than in higher SES women. In Appendix 6.10 we see the models after the backwards elimination procedure, these are the models we base the predictions on.

The predicted probabilities of health service use across each BMI category can be found in Table 6.13. We find that overweight and obese men and women have an increased *probability* of having a GP visit and have an increased *number* of GP visits. Similarly, for all men and all women the predicted *probability* of having an outpatient visit increases at higher levels of obesity. The predicted *number* of outpatient visits increases at higher levels of obesity for women, but not in men. Furthermore, compared with normal weight, overweight and obese men and women have increased predicted probability of inpatient stays, having seen a chiropodist, had blood pressure taken, had a cholesterol test and having had a blood test at higher levels of obesity. In addition, we find an increased predicted probability of having an x-ray of the chest and lungs for overweight and obese women compared with normal weight women, we do not find this in men. The predicted probability of using most types of the health service was higher in women than in men, except from having a cholesterol test where men have a higher predicted probability.

Table 6.13: Predicted probability of health service use across BMI categories in BHPS (Waves: 2004/5 and 2006/7)

	Men				Women			
	Normal weight	Over-weight	Obese Class I	Obese Class II/III	Normal weight	Over-weight	Obese Class I	Obese Class II/III
GP visit	0.656	0.690	0.712	0.763	0.809	0.839	0.866	0.907
Number of GP visit if number of visits>0	3.542	3.689	4.098	4.804	4.094	4.566	5.004	5.741
Outpatient visit	0.355	0.370	0.387	0.457	0.417	0.437	0.478	0.563
Number of outpatient visit if number of visits>0	3.176	"	"	"	3.540	3.419	3.830	3.885
Inpatient stay	0.035	0.040	0.044	0.070	0.085	0.095	0.106	0.124
Chiropodist	0.029	0.041	0.077	0.113	0.055	0.068	0.088	0.111
X-ray of chest and lungs	0.113	"	"	"	0.098	0.108	0.107	0.155
Had blood pressure taken	0.387	0.467	0.515	0.603	0.539	0.591	0.646	0.698
Had a cholesterol test	0.133	0.186	0.224	0.313	0.122	0.157	0.188	0.221
Had a blood test	0.299	0.353	0.389	0.479	0.389	0.453	0.512	0.581

Note

" Predicted mean equal to the normal weight category

Predictions based on random effects probit and interval regression models. The dependent variables are the health service use variables. The mean values are predicted probability of health service use for each BMI or BMI/SES category computed by setting the individual and household characteristics to their whole-sample mean values. The individual and household characteristics are age, age squared and age cubed, ethnicity, marital status, smoking status, health area variables, survey year and time elapsed.

Table 6.14 shows the predicted probability of using the health service across each SES quartile. We see similar trends within each of the quartiles, in that higher levels of obesity are associated with higher predicted probability of use. Lower SES men and women have a higher utilisation of GP, outpatient, inpatient, X-ray and blood test services than higher SES men and women. We also find a social gradient in the percent having had a cholesterol test in women but not in men.

Although, a higher *percentage* of the low SES individuals have had a GP visit, the impact of obesity and overweight on percentage share of individuals who have had a GP visit is relatively similar across the SES groups. However, as mentioned we identified an interaction in the *number* of GP visits for women. The predicted mean number of visits increases more in the lower SES women than in the higher SES women. For example, the predicted mean number of GP visit in the lowest SES quartile 1 women increased from a predicted mean of 4.8 in the normal weight category to 5.5

in the class II/III obesity, which is a discrete difference of 1.7. For the least deprived SES women quartile 4 similar numbers were 3.9 and 4.6, which constitutes a discrete difference of 0.7.

We also identify an SES gradient in inpatient visits, where the impact of obesity on the predicted share of individuals who have had an inpatient stay increased more for both men and women in the lower SES groups than in the higher SES groups. The predicted mean share of individuals who have had an inpatient stay in the lowest SES quartile 1 men increased from a predicted mean of 5% in the normal weight category to 10% in the class II/III obesity, which is a discrete difference of 5%. For the least deprived SES men quartile 4 similar numbers were 3% and 6 %, which constitutes a discrete difference of 3%. Similar numbers were found for women.

Table 6.14: Predicted probability of health service use across BMI and SES categories in BHPS (Waves: 2004/5 and 2006/7)

	Men				Women			
	Normal weight	Over-weight	Obese Class I	Obese Class II/III	Normal weight	Over-weight	Obese Class I	Obese Class II/III
GP visit								
1 (most deprived)	0.703	0.737	0.755	0.800	0.823	0.850	0.875	0.913
2	0.647	0.683	0.703	0.753	0.817	0.845	0.870	0.910
3	0.645	0.681	0.701	0.751	0.786	0.817	0.845	0.891
4 (least deprived)	0.634	0.671	0.691	0.742	0.808	0.837	0.863	0.905
Number of GP visit if number of visits > 0								
1 (most deprived)	4.321	4.487	4.838	5.499	4.779	5.195	5.921	6.512
2	3.607	3.773	4.124	4.786	3.967	4.465	5.248	5.629
3	3.182	3.347	3.698	4.360	3.810	4.223	4.147	5.230
4 (least deprived)	3.109	3.275	3.626	4.287	3.873	4.226	4.196	4.606
Outpatient visit								
1 (most deprived)	0.395	0.412	0.426	0.494	0.417	0.437	0.478	0.563
2	0.361	0.378	0.392	0.458	"	"	"	"
3	0.336	0.352	0.366	0.431	"	"	"	"
4 (least deprived)	0.336	0.353	0.367	0.432	"	"	"	"
Number of outpatient visit if number of visits > 0								
1 (most deprived)	3.425	"	"	"	3.829	3.690	4.085	4.117
2	3.158	"	"	"	3.575	3.435	3.830	3.862
3	3.031	"	"	"	3.320	3.181	3.575	3.608
4 (least deprived)	3.069	"	"	"	3.438	3.299	3.693	3.726
Inpatient stay								
1 (most deprived)	0.054	0.063	0.066	0.097	0.104	0.117	0.127	0.146
2	0.041	0.049	0.051	0.077	0.073	0.084	0.092	0.106
3	0.027	0.032	0.034	0.053	0.077	0.087	0.096	0.111
4 (least deprived)	0.029	0.034	0.036	0.056	0.078	0.089	0.097	0.112
Chiropodist								
1 (most deprived)	0.029	0.041	0.077	0.113	0.055	0.068	0.088	0.111
2	"	"	"	"	"	"	"	"
3	"	"	"	"	"	"	"	"
4 (least deprived)	"	"	"	"	"	"	"	"
X-ray of chest and lungs								
1 (most deprived)	0.150	"	"	"	0.119	0.130	0.127	0.177
2	0.122	"	"	"	0.101	0.110	0.108	0.153
3	0.101	"	"	"	0.080	0.088	0.086	0.125
4 (least deprived)	0.092	"	"	"	0.089	0.098	0.095	0.137
Had blood pressure taken								
1 (most deprived)	0.387	0.467	0.515	0.603	0.539	0.591	0.646	0.698
2	"	"	"	"	"	"	"	"
3	"	"	"	"	"	"	"	"
4 (least deprived)	"	"	"	"	"	"	"	"
Had a cholesterol test								
1 (most deprived)	0.133	0.186	0.224	0.313	0.119	0.112	0.164	0.180
2	"	"	"	"	0.124	0.117	0.168	0.210
3	"	"	"	"	0.121	0.122	0.160	0.161
4 (least deprived)	"	"	"	"	0.122	0.090	0.132	0.207
Had a blood test								
1 (most deprived)	0.331	0.389	0.423	0.511	0.419	0.483	0.540	0.605
2	0.307	0.365	0.398	0.486	0.392	0.455	0.512	0.578
3	0.275	0.329	0.361	0.448	0.370	0.432	0.489	0.555
4 (least deprived)	0.289	0.345	0.377	0.464	0.379	0.442	0.498	0.565

Note

The predictions are based on random effects probit and interval regression models. The dependent variables are the health service use variables. The mean values are predicted probability of health service use for each BMI or BMI/SES category computed by setting the individual and household characteristics to their whole-sample mean values. The individual and household characteristics are age, age squared and age cubed, ethnicity, marital status, smoking status, health area variables, survey year and time elapsed.

Similarly to the analysis in HSE we exclude results where the utilisation across all BMI/SES groups is below 10% for both men and women separately. Hence, we exclude physiotherapist use, cervical smear tests (in women) and breast cancer screens (in women).

6.4.3 Comparison of the results in HSE and BHPS

In general the findings are consistent across both datasets, where we find that overweight and obese individuals utilise the health service more than normal weight and that lower SES individuals utilise more than higher SES groups. However, where direct comparison is possible the associations between obesity and use are more positive in BHPS than in HSE. This could possibly be due to BHPS being panel data and therefore partly controls for heterogeneity from time invariant variables. Such variables could include discount rate (the weight individuals put on future benefits) or impact of increased health service use on obesity (reverse causality). Another, possibility is that the BHPS data is more recent (except in the medication use analysis). It could also be due to different time periods across the datasets and different populations. However, this will be discussed further for each dependent variable in the conclusion of this chapter.

The actual predicted probability of a GP visit differs for the HSE and BHPS. The reason is that the dependent variable in the HSE measures GP visits in the last two weeks while BHPS measures it since the 1st of September last year. In both HSE and BHPS we

find that the predicted probability of GP visits increases when BMI increase. Furthermore we find that the predicted mean number of visits increases when BMI increases for both men and women in BHPS. We also find a higher GP utilisation in more deprived SES quartiles than in less deprived SES quartiles in both men and women across all BMI groups in both datasets. Furthermore there are indications of a more pronounced association between obesity and mean share of individuals who have had a GP visit in the lower SES groups in women in HSE. These findings are supported by a significantly more pronounced association between BMI and number of GP visits in women in the BHPS. In men an opposite result was found of a somewhat more pronounced association in higher SES men across both datasets.

We find that obesity increases the use of outpatient services in women across both datasets. In men we only find that obesity increases use in the BHPS, but not in the HSE. In both datasets lower SES men use significantly more outpatient services than higher SES men. In women we do not detect an SES gradient in the probability of having an outpatient visit in either the HSE or the BHPS. However, we find that lower SES women who have had at least one outpatient visit have a higher number of visits than higher SES women.

For inpatient visits we find conflicting results in men as we do not find variation in predicted probability of a visit across BMI groups in the HSE, however we do find this in the BHPS. In women we find variation in the predicted share of users across BMI groups in both datasets. In both men and women we find that lower SES individuals use inpatient services more than higher SES individuals. In addition, we observe a more

pronounced association between obesity and inpatient stays in the lower SES groups across both datasets in women; we also find this in the BHPS in men.

Table 6.15: Summary of findings

Utilisation variable	Gender	Association between BMI and use	Association between SES and use	SES variation in the impact of obesity (absolute)
GP visits	Men	Positive	Negative	Variation across SES groups, but no clear trend
	Women	Positive	Negative	More pronounced in lower SES women than in higher SES women.
Outpatient visits	Men	Positive association in BHPS no association in HSE	Negative	No variation
	Women	Positive	No association in HSE. Negative association for number of visits in BHPS.	No variation
Inpatient visits	Men	No association in HSE. Positive association in BHPS	Negative	More pronounced in lower SES than in higher SES groups.
	Women	Positive	Negative	More pronounced in lower SES than in higher SES groups.
Practice nurse visits	Men	Positive	Negative	More pronounced association in lower SES than in higher SES groups
	Women	No association	Negative	No variation
Chiropodist	Men	Positive	No association	No variation
	Women	Positive	No association	No variation
X-ray of chest and lungs	Men	No association	Negative	No variation
	Women	Positive	Negative	More pronounced association in lower SES than in higher SES groups
Had blood pressure taken	Men	Positive	No association	No variation
	Women	Positive	No association	No variation
Had a cholesterol test	Men	Positive	No association	No variation
	Women	Positive	Negative	Variation across SES groups, but no clear trend.
Had a blood test	Men	Positive	Negative	More pronounced association in lower SES than in higher SES groups
	Women	Positive	Negative	No variation
Medication use	Men	Positive	Negative	More pronounced in lower SES than in higher SES
	Women	Positive	Negative	More pronounced in lower SES than in higher SES groups

6.5 Discussion

The aims of this chapter were to investigate the relationship between obesity and health service use, and whether or not any observed relationship varies by SES. Our main finding is that obesity and health service use are positively correlated. However, the SES stratified results are mixed. The relationship between obesity and use is more

pronounced in the lower SES groups than in the higher SES groups for some types of use, but not all.

Using the BHPS we find that, controlling for individual and household characteristics and a set of comprehensive area level variables, obesity is associated with increased probability of: GP visits, outpatient visits, inpatient stays, chiropodist visits, blood pressure tests, cholesterol tests and blood tests. In addition, we find that obesity will increase the use of x-ray of chest and lungs for women but not for men.

We rerun the analysis in HSE for GP visits; inpatient stays; and, outpatient visits and find that the results were consistent across the datasets except from inpatient visits in men, where we only found a significant result in the BHPS but not in HSE. In HSE we also investigate the relationship between obesity and practice nurse visits and find that this is positive and significant in men, but not in women.

We provide evidence to show that overweight and obese men and women have a greater predicted mean number of medications used than those of normal weight after controlling for a range of individual and household characteristics. Our findings are also disaggregated by therapeutic classification and show that obesity is associated with higher utilisation of cardio-vascular medicine, gastrointestinal medicine, respiratory medicine, CNS medicine, endocrine medicine, and musculoskeletal medicine for both men and women. It is to be expected that obesity increases the use of medication for cardiovascular disease and the endocrine system since obese have a higher risk of comorbidities like stroke, CHD and diabetes. We would also expect to

find an increase in the use of central nervous system medication as this is the category of drugs for treatment of obesity. The NICE guidelines states that drugs aimed at weight reduction (Orlistat or Sibutramine[34]) should be considered treatment in overweight and obese individuals if comorbidities are present. In addition, drugs should be considered to all individuals in obesity class II and III even when comorbidities are not present (NICE, 2006). The musculoskeletal medicine category includes medication for soft tissue inflammation which might be more of a problem in obese. In addition, obesity is possibly related to respiratory diseases like asthma (Shore & Johnston, 2006) which could explain why our analysis shows an association with respiratory medicine.

BHPS has an advantage over HSE as it is a panel dataset one can use panel models to some extent to control for unobserved time invariant heterogeneity. In addition, a further advantage of the BHPS is that we have controlled for extensive PCT level area variables, which reduces the likelihood that supply will have an influence on the results. We have only been able to control extensively for health area in the part of the HSE analysis that looks at GP visits, practice nurse visits, inpatient and outpatient visits, but not in the part that investigate medication use. In the medication analysis part of the impact of supply will be picked up by clustering at the primary sampling unit; however this will only affect the variance and therefore only the significance of the variables and not the coefficients. However, part of the variation might also be picked up by the broad area level variables Government Office Region of residence (nine

[34] Although Sibutramine have later been withdrawn from market in September 2010

categories). Nevertheless, an advantage of the HSE dataset is that we have nurse measured height and weight. Thus, the likelihood of systematic measurement error is reduced. Although, the results are consistent across the datasets the association between obesity and use is generally more pronounced in BHPS. This result might arise as HSE might be more vulnerable to reversed causality bias (if this bias is time invariant). For example, some types of health service might make individuals less obese, hence individuals who have frequent health service visits might be less likely to be obese. Furthermore, it might be because the BHPS data is more recent (except from the medication analysis), and that more interventions to handle obesity and obesity related comorbidities are available. However, it could also be a bias in BHPS as we use self reported height and weight compared with nurse measured values in HSE.

We provide evidence to show that overweight and obese men and women use more health services than those of normal weight. Our findings also show that women use the health services more than men. These results are qualitatively similar to those of other studies, which have also shown that health service use is positively associated with obesity (see, e.g., Reidpath et al., 2002). However, most of the studies in our literature review have presented their results in terms of relative odds ratios and it is difficult to directly compare them with our predicted probabilities. Hence, to enable comparison we have calculated odds ratios. We have chosen to compare our results with the results by Reidpath et al., (2002). The reason is that this study investigates the association between obesity and health service use separately for men and women. They find that compared with normal weight men (20 kg/m^2>BMI>25 kg/m^2) the odds ratio of having had a doctor visit in overweight men (25 kg/m^2>BMI>30 kg/m^2) was

1.06 and obese (BMI>30 kg/m^2) was 1.2 using Australian data controlling for age and income. In women comparable figures were 1.15 and 1.3. In HSE we find that the odds ratio of having had a GP visit the last two weeks in men was 1.03 in the overweight, 1.13 in the obese class I group and 1.58 in the obese class II/III. In women comparable odds ratios were 1.15, 1.22 and 1.48, respectively. In the BHPS we find that comparable odds ratios in men were 1.2, 1.43 and 2.04, and in women 1.31, 1.74 and 2.96, respectively. As we can see the results in the HSE are comparable to the results obtained by Reidpath et al., while the results in BHPS are more pronounced. This might be due to both HSE and the Reidpath study using cross sectional data, which might suffer from omitted variables bias to a greater extent than the BHPS analysis. Another explanation could be that the BHPS data (years: 2004/5-2006/7) is more recent than the data from the HSE (years: 1999-2002) and Reidpath (year: 1995) and that more treatments to tackle obesity related complications have been developed.

Conflicting results has been found regarding inpatient visits, for example Reidpath et al., (2002) do not find a significant effect in either men or women. Our results support a positive impact of obesity on inpatient stays in women in both datasets. In men we find variation across BMI groups on the probability of inpatient stays only in BHPS and not in HSE. These conflicting results in men might arise for similar reasons as mentioned above.

We are not aware of any studies that have investigated the association between obesity and chiropodist visits, blood pressure tests, cholesterol tests, blood tests and x-ray of chest and lungs. We find that obesity is positively related to use of these services

(x-ray in women only). Obesity is associated with hypertension and increased risk of having high cholesterol levels (National Audit Office, 2001). Hence, obese individuals should have more frequent blood pressure tests and cholesterol tests and this study shows that they do. In addition, obese individuals are likely to have more problems with their lower limbs, which could explain the significant increase in chiropodist visits. Increased blood tests can stem from increased liver problems in the obese, which can be detected by blood tests. However, there is no apparent direct link that would explain why obese women have an increased use of x-ray of chest and lungs compared with normal weight women.

As mentioned in the literature review we have identified two studies that investigate the effect of obesity on medication use disaggregated into therapeutic classifications (van Dijk, Otters and Schuit, 2006; Counterweight Project Team, 2004). Our results support the results of the Counterweight Project Team (2004) that show a significantly increased relative log odds of using medications for the cardiovascular system; central nervous system; endocrine system; musculoskeletal and joint disease; Infections; gastrointestinal; and skin and respiratory system controlling for age, sex, deprivation category, and country (England/Scotland). Reidpath et al., (2002) find that the compared with normal weight the relative odds ratio of having used medication in the last two weeks in men for the overweight was 1.13 and for the obese it was 1.46, respectively. Similar figures in women were in the overweight 1.32 and in the obese 1.68. Comparable figures in our study in men were (showing odds ratios) 1.17 in the overweight group, 1.58 in the obese class I group, and 2.27 in the obese class II/III group. In women, they were 1.24, 1.67 and 2.53, respectively. As we can see from

these numbers our study shows a more pronounced association between obesity and medication use. This might be because our data is more recent, hence more drugs have been developed to counteract obesity and the related comorbidities. It might also be that in general obese individuals are prescribed more medication in the England than in Australia.

Across both datasets we find that lower SES individuals have higher health service utilisation than higher SES individuals across all BMI groups. This is expected as we have showed in earlier studies that in general individuals in lower SES groups have a lower health related quality of life and lower life expectancy within each BMI group, see Chapter 3, 4 and 5.

The second aim of this chapter was to investigate SES variations in the relationship between obesity and health service use. We find a greater positive relationship between obesity and number of GP visits in the lower SES groups than in higher SES groups in women in the BHPS. Obese women in the lower SES group have on average 1.7 more GP visits per year than in the normal weight group. Obese women in the higher SES group have on average 0.7 more GP visits per year than in the normal weight group. Hence, the absolute difference in GP visits is more than double in the lowest SES group compared with the highest SES group. Furthermore, in HSE the predictions based on the probit models show a more pronounced positive relationship between obesity and GP visits in lower SES groups than higher SES groups in women.

Hence, in the case of women both analyses support the same hypothesis although the dependent GP visit variable differs across the datasets[35].

We find in the BHPS a more pronounced association between obesity and the probability of an inpatient visit in the last year in the less deprived SES groups in men and women. The results are consistent in the HSE for women, but not in men where we did not find variation in use across BMI groups. Furthermore, we find that the association between obesity and use is stronger in lower SES groups than in higher for: chiropodist and blood test in men, and x-ray in women.

For medication use we find that overweight and obese individuals use more medication than those of normal weight in the same SES group, and use more medications than those in higher SES groups within the same BMI group. Our analyses also illustrate a more positive association between obesity and medication use in lower SES groups compared to higher SES groups. Obese men in the lowest SES groups use on average 1.08 drugs more than normal weight men in the lowest SES group. Obese men in the highest SES group use on average 0.56 more drugs than normal weight men in the highest SES group. Hence, the association between obesity and medications use in the lowest SES group is nearly double that of the highest SES group. Similar trends, although less pronounced, where found in women.

[35] In the BHPS the variable is number of visits the last year while in the HSE it is a binary variable taking the value one if a person has had a visit in the last two weeks and zero otherwise.

This means not only is there a health service use gap between high and low SES individuals, this gap increases for overweight and obese individuals compared to normal weight individuals. These results support hypothesis (b) in Chapter 2 of a more pronounced impact of obesity in the lower SES groups. Our results are different to those found by Guallar-Castillon et al. (2002) who investigate health service use (hospitalisations, medical visits, hospital emergency services and medication) and do not find any interaction effect between BMI and education in Spanish women. However, our studies are not directly comparable as we look at predicted mean utilisation while they look at relative odds of use.

In this chapter we have focused on the absolute effect of obesity on use and not the relative. We illustrated with an example how absolute and relative interactions have different interpretations and clarity is needed in what one wants to express with the results. The results in this chapter illustrate that in some cases there is no interaction term between BMI and SES in the relative effects, but there is still SES variation in the absolute consequences of obesity on use. For example, in the medication use analysis we found no relative interaction effect however, the number of medications used increased more in lower SES groups.

These results have consequences for health service planning. Other studies have showed that obesity is related to SES, in which lower SES groups are over-represented in the higher BMI groups (Butland et al., 2007; Marmot, 2010). The results of this chapter suggests that obesity in lower SES groups leads to greater resource utilisation than in higher SES groups, for some types of use. This means that not only do lower

SES groups have a higher prevalence of obesity, but obesity also demands more health care resources in these groups. This provides an argument for paying particular attention to counteracting obesity in lower SES groups. In addition, this suggests that attention needs to be paid to the role of SES when undertaking predictions of future utilisation and costs. If one fails to take into consideration the interface between SES and obesity one might fail to reflect the actual relationship between obesity and health service use.

A number of studies have investigated horizontal inequity in health service use (See, e.g., Morris, Sutton & Gravelle, 2005) and found that lower SES groups use more primary care services and less secondary care services. Our study finds more use of both primary and secondary care services. However, we have not fully controlled for need variables (as this was not the aim of our study).

Our study has some potential limitations. First, obesity might be endogenous making it difficult to draw inferences about the impact of obesity on health service use. Endogeneity might arise for the following reasons. There might be reverse causality: for example, the GP might tell obese individuals to exercise more and eat less energy dense food. Hence, obesity is then a function of health service use. This would mean that we have underestimated the impact of obesity on use. There might be omitted variable bias, i.e., variables that affect both health service use and obesity. One such variable might be health status. Health status may increase health service use and, as discussed in Chapter 3, it may reduce physical activity which will increase weight. Conversely, reduced health status in terms of diseases may also reduce weight. Hence,

it is difficult to establish the direction of this causality. Another omitted variable might be supply of health services. We have not been able to control for supply factors as extensively in the medication analysis as in the two other analyses. However, we argue that clustering by PSU and controlling for the broad area level variables to some extent can control for this. There may be measurement error: BMI in itself might be mis-measured when it is based on self-reported height and weight. If the level of mis-measurement is associated with the dependent variable the obesity coefficients might be biased. This is an issue in the BHPS but not in the HSE where BMI is based on nurse measured height and weight. Another potential measurement error is related to using BMI as a measure of obesity. This measure can be considered an imperfect measure, and it has been criticised, e.g., because it does not incorporate body fat, which is an independent predictor of ill health (Burkhauser & Cawley, 2008).

This chapter has a number of advantages compared with earlier studies. The number of observations in our sample is sufficiently large to permit analyses stratified by both gender and SES; our findings show that the impact of obesity on health service use varies by both these factors. BHPS is a panel dataset, which makes our results less vulnerable to time invariant heterogeneity bias. Also, we have rich information on individual and household characteristics and so can argue that it is less likely that the estimated effects of obesity on health service use in our models are due to their correlation with omitted variables. Our extensive health area variables (216 groups in the BHPS and 95 categories in the HSE) let us control for area variation in supply. This has not been done in previous studies.

To summarise, first this chapter has shown that obesity is positively correlated with health service use. Second, we have analysed how SES influences this relationship and found mixed results. Individuals in lower SES groups use more health care than individuals in higher SES groups across all obesity categories. In addition, the association between obesity and GP visits (in women), inpatient stays (in women) and medication use (for both men and women) is more pronounced in lower SES groups compared with higher SES groups.

CHAPTER 7

Conclusion and implications

7.1 Introduction

The two general aims of this thesis were:

1. Examine the relationships between obesity and health outcome; and obesity and health service use in Great Britain
2. Investigate whether or not there are socioeconomic variations in these relationships

To meet the general aims the preceding chapters we have provided a comprehensive analysis of socioeconomic variation in the consequences of obesity. In Chapter 2, we outlined the relationship between socioeconomic status, lifestyle and health in light of a theoretical model, an economic model and a literature review. We also mapped the relationship between socioeconomic status, obesity and health service use. Based on the economic models we generated econometric models for testing. Chapters 3, 4 and 5 provided extensive analysis of the association between obesity and health and how this varies by SES. We measured health in terms of Health Related Quality of Life (HRQL) and life expectancy. Chapter 6 examined the association between obesity on a range of health services and medications and how these varied by SES.

In this final chapter we summarise and discuss policy implications of the findings. We then discuss some limitations of these studies and offer some suggestions for further research.

7.2 Main findings

Chapter 2 provided a framework for examining socioeconomic variation in the relationships between: obesity and health; and, obesity and health service use. We show by what means the CSDH framework use to illustrate the pathways and mechanism through which social determinants of health influences health. We present a human capital model for health (the Grossman model) and use this model to analyse how the impact of risk factors on health might vary by SES. The models predict that individuals who are obese will have a lower equilibrium level of health. We further include SES and predict that higher SES groups have a higher equilibrium health than lower SES groups. We then take into account SES variation in the impact of obesity and its consequences for equilibrium health. We develop econometric models suitable for empirical testing based on these models. We conduct a literature review and find that limited information exists on this topic. Most of the evidence is on smoking and few studies have looked at obesity. We develop hypotheses to analyse SES variation in the impact of obesity on health service use, using a model based on the concept of need and non-need factors.

We will now discuss the empirical findings in this thesis under the heading of each of the general aims, starting by the first aim followed by the second aim.

7.2.1 The associations between; obesity and health, and obesity and health service use

In Chapter 3 we examined the relationship between HRQL and obesity and found that obesity and HRQL are negatively correlated. After controlling for individual and household characteristics, we obtain similar results, albeit at reduced levels. Our findings also show that the negative association between obesity and HRQL is greater in women than in men. These findings are consistent with previous findings using British data (see, e.g., Macran, 2004; Sach et al., 2007).

In Chapter 4 we further analyse the relationship between HRQL and obesity in individuals with obesity-related comorbidities. We look at four longstanding illness categories related to obesity: diabetes; hypertension; stroke; and, heart attack and angina. We find that there is a significant negative association between obesity and HRQL in individuals with each of these four comorbidities. Furthermore, the relationship between obesity and HRQL varies by these obesity-related comorbidities. These results contradict the findings by Lee et al., (2005) who investigate the association between BMI and HRQL and find insignificant interaction terms between BMI and diabetes status. Except from diabetes we are not aware of any studies that investigate if the association between obesity and HRQL varies by other obesity-related comorbidities.

In Chapter 5 we investigated the relationship between obesity and mortality/life expectancy controlling for a number of covariates. Obesity increased mortality and reduced life expectancy for men and younger women. In older women we found obesity to increase life expectancy. We also investigated the impact of overweight as well as obesity on life expectancy. We did not find a clear trend and could not conclude upon how overweight affects mortality. Furthermore, we fitted models controlling for a number of longstanding illnesses and whether or not these were limiting. These models showed that the impacts of obesity and SES are still significant and the magnitudes were only slightly reduced. In general, our findings support earlier findings on the relationship between obesity and life expectancy. However, this is the first study we are aware of that has: identified the parametric distribution that best fits the data; controlled for unobservable heterogeneity; predicted median life expectancy based on a parametric survival function; and, identified significant interactions between obesity as a category and age in women. In addition, this is the first study to control for all types of longstanding illnesses and whether or not they are limiting and we still show that there is an effect.

In Chapter 6 we analysed the relationship between health service use and obesity, and whether or not this relationship varied by SES. We found that obesity was associated with increased probability of; GP visits, outpatient visits, inpatient stays, chiropodist visits, blood pressure tests, cholesterol tests and blood tests, using the BHPS (wave 14 (1st September 2004 to 11th May 2005) and wave 16 (1st September 2006 to 3rd April 2007)) controlling for individual and household characteristics and a set of comprehensive area level variables. In addition, we found that obesity increases the

use of x-ray of chest and lungs for women but not for men. Although numerous studies have investigated the associations between obesity and several types of health service use, we are not aware of any studies that have investigated the association between obesity and; chiropodist visits, blood pressure tests, cholesterol tests, blood tests and x-ray of chest and lungs. We found that obesity is positively related to use of these services. This serves to further explain the burden obesity puts on the health service and suggests that leaving out these services in studies of the cost of obesity will lead to an underestimate of the costs.

We also investigated the relationship between obesity and practice nurse visits and found a positive association in men but not in women. We provided evidence to show that overweight and obese men and women have a greater predicted mean number of medications used than those of normal weight after controlling for a range of individual and household characteristics. Our findings were also disaggregated by therapeutic classification and showed that obesity was associated with a significantly higher utilisation of cardio-vascular medicine, gastrointestinal medicine, respiratory medicine, CNS medicine, endocrine medicine, musculoskeletal medicine for both men and women. This supports the current literature on obesity and medication use.

Although a number of studies have investigated the association between obesity and health service use, our study has numerous advantages over earlier British studies. We are the first to apply count and linear models to investigate number of visits/medications used, while the others have used binary models. By doing this we capture more of the consequences of obesity. We are the first to control for supply

variables which has been showed to affect use. We are the first to use panel data methods hence we do to some extent control for unobservable heterogeneity. In addition, we are not aware of any British studies that have generated predicted probabilities (instead of odds ratios) to investigate the absolute (quantitative) increase in health service use and number of medications used.

In Table 7.1 and Table 7.2 below we see a summary of some of the findings of the relationships between; obesity and health, and obesity and health service use.

Table 7.1: Extracts of some of the findings of the relationship between obesity and health outcome

- Men who are obese class II / III have a mean EQ-5D score of 0.831 compared with 0.893 in normal weight men
    - In women similar numbers were 0.882 and 0.740, respectively
- Obese class II / III individuals with diabetes had a mean EQ-5D score of 0.701, falling from a mean of 0.834 in the normal weight group
    - For individuals with stroke similar numbers were 0.461 to 0.682, respectively
    - For individuals with hypertension similar numbers were 0.774 to 0.888, respectively
    - For individuals who have had heart attack or who has angina similar numbers were 0.561 to 0.782, respectively
- Obese (BMI>30kg/m^2) men aged 55 will on average life 22.1 years while normal weight men aged 55 will on average live for another 25 years.
    - Similar numbers for women were 26.4 and 31.0, respectively.

Table 7.2: Extracts of some of the findings of the relationship between obesity and health service use

- Obese class II / III men will on average have 1.4 more GP visits per year than normal weight men
  - 0.5 more outpatient visits per year
  - 3.5% more likely to have an inpatient stay each year
  - 4% more likely to have a chiropodist visit each year
  - 21.4% more likely to have blood pressure taken each year
  - 17.7% more likely to have a cholesterol test each year
  - 18% more likely to have a blood test each year
  - Use 0.81 more types of medication per year
- Obese class II / III women will on average have 1.9 more GP visits per year than normal weight women
  - 0.7 more outpatient visits per year
  - 3.9% more likely to have an inpatient stay each year
  - 5.5% more likely to have a chiropodist visit each year
  - 15.8% more likely to have blood pressure taken each year
  - 9.9% more likely to have a cholesterol test each year
  - 19.2% more likely to have a blood test each year
  - Use 0.96 more types of medication per year

7.2.2 Socioeconomic variation in the associations between; obesity and health, and obesity and health service use

The second aim of this thesis was to investigate if there is socioeconomic variation in the associations between: obesity and health, and obesity and health service use. Furthermore we wanted to establish the direction of these relationships and quantify them.

Rather than focusing on a single SES indicator we developed a composite measure based on a prediction of income from a range of SES indicators. We use this measure of SES in Chapter 3 to examine whether or not the association between obesity and HRQL varies by SES. We find that the negative association between obesity and HRQL differs significantly across SES groups; the association is more negative in people from lower SES groups compared with individuals from higher SES groups. This is also

observed after controlling for individual and household characteristics, although the statistical significance and magnitude of the effects is diminished. A number of studies have investigated the impact of obesity on HRQL controlling for SES in multivariate analyses. In addition, the study by Laaksonen et al., (2005) found the association between obesity and SF-36 score to vary by working conditions (self evaluated job control and job demand). However, we are not aware of any published studies that have stratified their analyses by SES groups and showed that the association between BMI groups and HRQL varies by SES.

In Chapter 4 we focus on obesity-related comorbidities and conduct an analysis of SES variation in the association between obesity-related comorbidities and HRQL. We find that the association between the obesity-related comorbidities on HRQL is more negative in lower SES groups compared with higher SES groups; this could explain why we identify the gradient in Chapter 3.

In Chapter 5 we investigated whether or not the association between obesity and mortality/life expectancy varies by SES. We found that higher SES groups had a lower mortality and a longer predicted life expectancy than individuals in lower SES groups in both men and women. The hazard rate from obesity did not vary across SES groups in men. However, the predicted loss in life years due to obesity was greater in higher SES men than in lower SES men. In women, the hazard rate varies by both age and SES. We found that obesity had a negative impact on life expectancy in lower SES women (aged 55, 65 and 75) and a positive impact on life expectancy in higher SES women (aged 55, 65 and 75).

In Chapter 6 we investigated SES variation in the association between obesity and health service use. We found that the relationship between obesity and GP visits was more positive in the lower SES groups in women. In addition, we found a trend of a more pronounced association between obesity and the probability an inpatient stay in the lower SES groups. Our analyses also showed a more pronounced association between obesity and medication use in lower SES groups compared to higher SES groups in both men and women. This trend was more pronounced in men. Furthermore, the analysis shows a tendency for a more pronounce association in the lower SES groups between obesity and; having had a blood test in men, having had a practice nurse visit in men and having x-ray of chest and lungs in women. However, we also investigated a range of services where we did not identify this trend.

7.3 Policy implications

There are two main messages resulting from this research. First, in this thesis we have provided evidence of associations between obesity and health related quality of life, life expectancy and health service use in Great Britain. Not only do our findings show that obesity is related to all of these, we also quantify the consequences. This serves to further explain the burden that obesity puts on the individual and the health care system. Hence, it emphasises that effort is needed to avoid the current trend of increasing obesity rates in Great Britain.

The second main message is that the consequences of obesity are likely to vary by SES. Hence, more attention needs to be paid to analysing heterogeneities in pathways to health and health service use. Socioeconomic context is an element in understanding the production of health across populations.

In the introduction we illustrated that there has been a focus on public health interventions aimed at vulnerable groups and that there is a call for interventions targeted at lower SES groups. The Department of Health's White paper (2010) has suggested "a radical new approach" in counteracting health inequalities, where the government should consider different approaches for different groups. Furthermore, the Foresight Project Report suggested targeted interventions as a strategy to counteract obesity (in itself and not related to counteracting inequalities in health) (Butland et al., 2007). Hence, obesity interventions may be seen as being particularly important in deprived groups. However, to be able to measure the effectiveness and cost-effectiveness of targeted interventions information about SES specific outcomes are needed. Therefore this research suggests that to evaluate interventions across SES groups one would need to use estimates of costs and effects specific to the groups in question. Failure to do so might lead to misallocation of resources. The intervention being most cost-effective across the whole population may not be the most cost-effective intervention for any of the SES groups when we stratify by SES.

There is a focus on identifying factors that explain socioeconomic inequalities in health. Inspired by frameworks like the Commission for Social Determinants of Health (CSDH) framework, the focus has mainly been on identifying prevalence of risk factors

across SES groups. The impression has been that higher prevalence of a lifestyle related risk factor in lower SES groups contributes to inequalities in health. Hence, obesity will contribute to inequalities in health simply based on it being more prevalent in lower SES groups than in higher SES groups. This view has lead a number of government reports to suggest specifically counteracting lifestyle related risk factors in lower SES groups as means of preventing socioeconomic health inequalities (See, e.g., Department of Health, 2010; and, Marmot, 2010). However, this is based on an assumption that lifestyle related risk factors have a relatively similar impact on health across socioeconomic groups. This thesis challenges this assumption, and we suggest that information specific to the SES group in question is needed.

The *Acheson Report* (1998) highlighted as one of its key recommendations "that as part of health impact assessment, all policies likely to have a direct or indirect effect on health should be evaluated in terms of their impact on health inequalities, and should be formulated in such a way that by favouring the less well-off they will, wherever possible, reduce such inequalities". A policy maker who wants to assess how a public health intervention will influence the social gradient in health will need two pieces of information. First, one must know if there are variations across SES groups in the effect of the intervention on the prevalence of the risk factor. Second, one must know the relationship between the risk factor and health across SES groups. If one focuses only on the first part one might fail to identify variations in outcomes within populations and risk misinterpreting the consequences of the intervention and misallocate resources.

To recap, failure to identify heterogeneities could lead to inappropriate policy responses by inappropriate use of resources and increases in inequalities in health within populations. However, care is needed when we interpret the findings as we have not established causality.

7.4 Research implications

The aim of this thesis was not to develop new methods, but to apply existing methods investigating new hypotheses. However, we still have some implications for further research.

The vast majority of obesity research investigates the relationship between obesity and an outcome variable (health or health service use) by controlling for SES factors. Our analysis serves to illustrate that by using this approach one risks representing only partially the relationship between obesity and the outcome variable. The prevalence of obesity has a socioeconomic gradient and researchers must take care to ensure that they in their research on pathways to health in fact represent the real relationship between the determinants of health and health/health service use. Hence, the main implication is that researchers must allow for possible interactions between obesity and SES. Below we briefly discuss a number of other implications for future research.

Predicted SES

Throughout this thesis we have based the measure of SES on a prediction of income. In this way we capture the multidimensional aspects of SES rather than focusing on a

single aspect like education or income. SES is a summary measure consisting of many factors and an advantage of our approach is that we can allow for a range of factors to influence in which group an individual is placed. Hence, we suggest using combined measures of SES when investigating this concept.

Parametric survival distributions

In Chapter 5 we have used fully parametric survival distributions alongside the commonly used semi-parametric distribution. We noted that this had a number of advantages. First, it allows for an easy interpretation of the consequences of the results in terms of predicting life expectancy. Second, one can control for unobservable (duration) heterogeneity by adding a frailty term. The downside of the parametric approach is that the results might be sensitive to the underlying parametric distribution. To investigate this we suggest reanalysing the data using a Cox model.

To decide upon the "correct" parametric distribution we suggest using plots of cumulative Cox-Snell residuals and the Aikake's information criterion. We also suggest to use the Cox proportional hazard model to run tests of proportional hazard of the coefficients using Schoenfeld and scaled Schoenfeld residuals.

Using the BHPS and the HSE to model associations between obesity and use

We have used both the BHPS and the HSE to model the associations between obesity and use. One advantages of the HSE is that the height and weight measurements are conducted by a nurse. Hence, it does not have the bias of self-reported height and weight measurement. In addition, the HSE has detailed information about medication

use. However, there is less information about other types of use. In the BHPS there is more detailed information about a range of health service use categories. In addition, there is more information about number of visits. Another advantage with the BHPS is that we have repeated measurements of the same individuals. This provides opportunities for a number of methods to eliminate the influence of omitted variables.

Generating predicted absolute values

In this thesis we have generated predictions of variation in health outcomes and health service use based on variation in BMI and SES. We have used a number of different approaches to quantify the findings depending on the dependent variable. When using interactions it can be difficult to understand the implications of the estimated coefficients, especially when applying to nonlinear models. Hence, we recommend that one provides predicted values to show the implications of the findings.

Testing for heterogeneity

In this thesis we have used a number of methods to test for socioeconomic heterogeneity across linear and nonlinear models. We discussed reasons for including interactions in linear and nonlinear models as well as the interpretation of those. In a linear model an interaction term can be interpreted directly as an absolute effect. In a nonlinear model the coefficients have a different interpretation and so has the interaction term.

There has been much debate around how to interpret interactions in nonlinear models when one wants to interpret them as absolute effects. The reason is as illustrated by Ai

& Norton (2003) that the interaction term in the latent variable model cannot be used to say whether or not there are significant interactions in the absolute effects. We recommend following the approach suggested by Greene (2010), where one focuses on what the predicted values reveal about the problem in context, instead of the coefficients in the latent variable models.

7.5 Some limitations and issues

In each chapter it is acknowledged that there are a number of limitations. These arise mainly from restrictions in available data. We now discuss these in the context of offering some suggestions for further research.

Endogeneity

An independent variable is endogenous if it is correlated with the error term and this can arise for three reasons (Wooldrige, 2002). The first is omitted variables. Omitted variables occur when there are variables that affect both the dependent and the independent variable but, because of data availability, we can not include them in the regression. The second type of endogeneity is simultaneity or reverse causality. This is when the independent variable is determined partly by the dependent variable. Third there may be measurement error, which is when we can only observe an imperfect measure of one or more of the variables in the regression.

In this thesis we have essentially used three types of dependent variables; health related quality of life, survival time and health service use. For each analysis it is

possible that there are different versions of the endogeneity issues mentioned above. In the following we will discuss in more detail why there might be endogeneity issues and how these might have influenced the results for each dependent variable.

We start with the relationship between obesity and HRQL. While obesity may affect HRQL, the reverse may also be true and/or there may be omitted variables that affect both obesity and HRQL. One such variable might be functional status. Individuals may be obese because they have a reduced functional status. This reduced functional status might also be the cause of the reduced HRQL. The result of this would be that we overestimate the impact of obesity on HRQL. Conversely, we might have underestimated the impact of obesity on HRQL if, for example, an illness which reduce HRQL also reduce appetite. Hence, the bias might go both directions and it is difficult to establish whether we over- or underestimate the impact.

Omitted variables can be mitigated if a proxy variable is available. Hence, to mitigate the issues above we run a separate analysis where we control extensively for longstanding and acute illnesses, in addition to the analysis where we only control for individual and household characteristics. The downside of this approach is that it is an overadjustment as health factors are part of the pathway between obesity and the output variable (See Zizza, 2004). Hence, when we use this approach we underestimate the impact of obesity on HRQL. It would be preferable to use an approach where one can look at causal impact of obesity accounting for the complex interface between obesity and its comorbidities. It is a general weakness in these studies, including our own, that this is not properly accounted for.

The second dependent variable is survival time. Newman et al. (2001) illustrated that weight loss is a predictor of mortality. If individuals loose weight shortly before they die one will underestimate the impact of obesity on mortality. Hence, this is a potential omitted variable issue. Furthermore, there might be other omitted variables that can affect both mortality and BMI. One such variable is health status. However, in our analysis we have accounted for this issue. First, we control for a range of variables and fit regressions both with and with out controlling for illness variables. Second, we add a frailty term to account for unobservable factors affecting the relationship between obesity and mortality. Hence, we argue that omitted variables are not an issue in our analysis. Furthermore, this analysis does not suffer from reverse causality as the dependent and the independent variables are measured at a different point in time.

The third dependent variable discussed is health service use, which may suffer from omitted variable bias. For example, this association is an overestimate of the causal effect of obesity on use if some individuals becomes obese because of functional limitations, and have higher health service use because of these functional limitations. On the contrary, the causal effect might be an underestimate if, for example, individuals with a lower propensity to use the health service also are more likely to be obese (Cawley & Meyerhoefer, 2011). In addition, this analysis may suffer from reverse causality, i.e., health service use might reduce BMI. For example, a GP might motivate an obese person to loose weight, which would lead to an underestimate of the causal impact. Based on this it is difficult to establish the direction of the bias. Nevertheless, we limit the bias by controlling for a range of individual and household characteristics.

I each chapter there might be measurement error. Our measure of obesity is BMI, which can be considered to be an imperfect measure of obesity. For example, BMI has been criticised because it does not distinguish body fat from fat-free mass such as muscle and bone (Burkhauser & Cawley, 2008). This means that we include some individuals in the higher BMI categories which have a high BMI due to a high muscle mass. This would mean that we underestimate the negative consequences of higher BMI and reduce the significance of the BMI variables. Although we recognise this issue we have chosen to use BMI as it is the most commonly used measure of obesity, which enabled us to compare our results to those of other studies. Furthermore, it is more often available than other measures of obesity.

BMI might also be mis-measured (e.g., if it is based on self-reported height and weight). If the level of mis-measurement is associated with the dependent variables, the obesity coefficients may be biased. In BHPS the values are self-reported. However, the HSE and HALS obtained their height and weight measurement during the interviewer visit and it is not self-reported so the likelihood of this problem is reduced.

It is also likely that the dependent variables suffer from measurement error. For example, there may be reporting bias with respect to HRQL and health service use. It may be systematic reporting bias that is correlated with obesity and/or with SES in the HRQL measurement or the health service utilisation.

Little attention has been provided to these issues in earlier literature. However, some researchers have used an instrumental variable approach. This will be discussed in more detail below.

Missing data

We might have issues with missing data. For the covariates the missing values are included as a part of the excluded category. To allow for those values not missing at random we included dummy variables to indicate non-response. However, we have not done this for missing BMI. These values may be missing for a number of reasons which might not be random. This means that the data which is missing is correlated with a number of variables and our results fail to take these observations into account. There are a number of methods available for handling missing data. Some have for example imputed the missing variables. However, there are issues with this approach. If there is a unique set of values attached to the missing data it is difficult to argue that one can capture these based on predictions of non-missing individuals.

Utilisation measures

There are issues related to the measures of health service use. When we measure the probability of having a GP visits per year we do not capture the full extent of the implications of obesity for GP use. When we look at the number of visits per year our measure improves. However, we still do not capture how much time and resources are used on each defined group. For example, it might be that obese individuals get more treatment every time they go to the GP than normal weight individuals. If this was the case we would underestimate the impact of obesity on health service use.

Obesity onset and long versus short term consequences of obesity

We have not been able to measure obesity onset which might have an impact on the consequences of obesity in terms of health. Furthermore, we have not been able to distinguish between long and short term consequences of obesity. It might take a number of years in the obesity state before the negative consequences of obesity emerge.

7.6 Suggestions for future research

Instrumental variable approach

To address the endogeneity issues some researchers have used an instrumental variable approach. In this approach one use variables that are good predictors of obesity but are not independently related to the outcome variable (health or health service use). The validity in the instrumental variable approach relies in finding the appropriate instruments and these could often be hard to find.

There are essentially two types of variables that have been used as instruments for obesity and we discuss each in the following.

Some have applied are area level measures e.g. mean BMI and percentage obesity by area (see Vallejo-Torres & Morris, 2010; Morris, 2007). These have showed to be good predictors of obesity. However, there are issues with using this as an instrument. The

reason being that this instrument also suffers from similar endogeneity issues to the obesity variable. Imagine we look at the impact of obesity on health and want to remove the possible reverse causality in that health might affect obesity, and to do this we apply area level BMI as instruments. An issue with this is that area level health might affect obesity status in the area; hence we would not fully remove the reverse causality problem. One could control for area level health in the regressions, however this might then again be correlated with individual health. Hence, we remove part of the pathway between obesity and HRQL.

Another approach is to exploit genetic variation in weight. This is to use the weight of a biological relative as an instrument for weight of the respondent. This has been showed to be a powerful predictor of the weight of the respondent (Cawley & Meyerhoefer, 2011). However, the instrument must also be uncorrelated with the respondents health or health service use. This means that a common environment must be uncorrelated with health and health care expenditures, which is difficult to prove.

The validity of both instruments above is threatened as both are based on quite strong assumptions. In addition it has in general proved to be difficult to find other instruments for obesity. Hence, rather than going into this discussion in this PhD we have chosen to completely ignore this approach and we see it as a question for further research.

Panel data methods

Panel data methods gives the opportunity to control for unobservable individual effects which remain constant over time. Hence, this might deal with the omitted variable bias and measurement error. However, due to limited time periods with obesity measurement in the BHPS and limited variation in the dependent variables we were forced to use random effects methods.

If more time periods with BMI values in BHPS become available it would be interesting to redo the studies using a fixed effects model or a more advanced random effect model, which strictly looks at deviations from the mean. This would allow us to be confident that we have fully omitted time invariant heterogeneity.

Using other measures of obesity

To avoid the biases related to BMI one could use other measures of obesity. Candidates include: total body fat; percent body fat, which is total body fat divided by total mass; waist circumference; and waist-to-hip ratio. However, one must note that these measures also have weaknesses (Burkhauser & Cawley, 2008). The main reason why these have not been applied in this research is that they are less frequently available and literally all studies that investigate obesity use BMI and we aimed to compare our results with other studies.

In addition, it would be interesting to measure the consequences of more extreme obesity categories. In our research we merged obesity class II and obesity class III to assure an adequate number of individuals in each category. However, the

consequences across these groups might differ and it would be interesting to investigate this further.

Reason for SES variation

It would be interesting to investigate reasons for why we get SES variation. We have to some extent investigated this in Chapter 4 by investigating SES variation in the association between obesity-related comorbidities and HRQL. Another related question is: Does the prevalence of obesity-related comorbidities vary by SES? Another interesting question could be; is obesity across SES groups caused by similar behaviour? This is interesting because it would give an indication on how to develop obesity interventions targeted at specific SES groups.

Cost-effectiveness analysis of an obesity intervention across SES groups

The Foresight report mentions as one of their "top five policy responses to obesity" targeting health interventions for those at increased risk, such as low income groups (Butland et al., 2007). However, little is known about the cost effectiveness of targeted interventions. It would be interesting to investigate whether or not this would yield good value for money.

7.7 What does this thesis add?

This thesis serves to increase focus on heterogeneous consequences of a homogenous condition. To inform a decision making process the evidence used must be designed

around the context of the problem. By failing to do so one risks misinforming policy makers and misallocating resources.

In meeting the aims of this thesis we have made original contributions to the literature in a number of respects. In Chapter 3 we show that there is socioeconomic variation in the association between obesity and HRQL, this is the first time this has been shown. In Chapter 4, we identify a significantly more pronounced association between obesity and HRQL in individuals suffering from a range of obesity-related comorbidities. To our knowledge, this is the first time this has been shown. Furthermore, we demonstrate that the association between obesity-related comorbidities and HRQL is more pronounced in lower SES groups compared with higher SES groups. This is the first time this has been show. In Chapter 5 we investigate socioeconomic variation in the association between obesity and life expectancy. This is the first time this has been done. In addition, it is the first time the relationship between obesity and life expectancy has been investigated using a parametric distribution controlling for unobservable heterogeneity. Furthermore, it is the first time the predicted loss in life expectancy due to obesity has been predicted based on a parametric distribution. In Chapter 6 we analyse socioeconomic variation in the relationship between obesity and health service use. We are the first to show that obesity has a more pronounced association with health service use in lower SES groups than in higher SES groups. Furthermore, this is the first time the relationship between obesity and health service use has been analysed controlling for extensive health area variables. It is also the first study in Great Britain to provide “additive” (quantitative) estimates of the association between obesity and health service use. Lastly, we have discussed and provided a

framework for interactions in linear and nonlinear models in health economics research.

References

Abell, J.E., Egan, B.M., Wilson, P.W.F., Lipsitz, S., Woolson, R.F., & Lackland, D.T., (2008). Differences in Cardiovascular Disease Mortality Associated With Body Mass Between Black and White Persons. *American Journal of Public Health*, 98, 63-66

Accounts Commission for Scotland. (1999). *Supporting prescribing in general practice*. Edinburgh: Accounts Commission for Scotland

Adams, K.f., Schatzkin, A., Harris, T.B., Kipnis, V., Mouw, T., Ballard-Barbash, R., Hollenbeck, A., & Leitzmann, M.F. (2006). Overweight, Obesity, and Mortality in a Large Prospective Cohort of Persons 50 to 71 Years Old. *The New England Journal of Medicine,* 355:763-78

Adda, J., & Lechene, V. (*2001)* Smoking and Endogenous Mortality: Does Heterogeneity in Life Expectancy Explain Differences in Smoking Behaviour? *University College London, Unpublished Paper*

Ai, C., &, Norton, E. (2003). Interaction terms in logit and probit models. *Economics Letters,* 80, 123–129

Allin, S., Masseria, C., & Mossialos, E. (2010). Equity in health care use among older people in the UK: an analysis of panel data. *Applied economics*, 1–11, iFirst

Allison, D.B., Fontaine, K.R., Manson, J.E., Stevens, J., VanItallie, T.B., (1999). Annual Deaths Attributable to Obesity in the United States. *JAMA*, 282, 1530-1538

Al Snih, S., Ottenbacher, K.J., Markides, K.S., Kuo, Y., Eschbach, K.,Goodwin, J.S. (2007). The Effect of Obesity on Disability vs Mortality in Older Americans. *Arch Intern Med.* 167, 774-780

Arulampalam, W., (1999). A note on estimated coefficients in a random effects probit models. *Oxford Bulletin of Economics and Statistics*, 61, 597–602

Avorn, J., Monette, J., Lacour, A., Bohn, R.L., Monane, M., Mogun, H., LeLorier, J. (1998). Persistence of use of lipid-lowering medications: a cross-national study. *JAMA*, 13, 1458-6

Balia, S., & Jones, A., (2008). Mortality, lifestyle and socio-economic status. *Journal of Health Economics*, 27, 1-26

Batty, G.D., Shipley, M.J., Jarrett, R.J., Breeze, E., Marmot, M.G., Davey Smith, G. (2006). Obesity and overweight in relation to disease-specific mortality in men with and without existing coronary heart disease in London: the original Whitehall study. *Heart,* 92, 886-92

Birch, S., Jerret, M., Wilson, K., Law, M., Elliot, S., & Eyles J. (2005) Heterogeneities in the production of health: smoking, health status and place. *Health Policy*, 72, 301-310

Birch, S., Jerret, M., & Eyles J. (2000) Heterogeneity in the determinants of health and illness: the example of socioeconomic status and smoking. *Social Science & Medicine,* 51, 307-317

Birch, S., Jerret, M., & Eyles J. (1997) Heterogeneity in the determinants of health? An empirical analysis of the smoking-health relationship in socioeconomic context. McMaster University, Centre for Health Economics and Policy Analysis, Working Paper #97-8

Bender, R., Trautner, C., Spraul, M., & Berger, M. (1998). Assessment of Excess Mortality in Obesity. *Am. J. Epidemiol*, 147: 42 - 48.

Benzeval, M., Judge, K. (1994). The determinants of hospital utilisation: implications for resource allocation in England. *Health Econ*, 32,105–116

Berry, W.D., DeMerit, J.H.R., & Esarey, J. (2010) Testing for Interaction in Binary Logit and Probit Models: Is a Product Term Essential?. *American Journal of Political Science*, 54, 248–266

Bertakis, K.D., Azari, R. (2005). Obesity and the use of health care services. *Obes Res*, 13, 372–379

Blaxter, M. 1990. *Health and Lifestyles*. London: Tavistock.

Brown, W.J., Dobson, A.J., & Mishra, G. (1998). What is a healthy weight for middle aged women? *Int J Obes Relat Metab Disord*, 22, 520–528.

Burkhauser, R.V., & Cawley, J. (2008). Beyond BMI: The Value of More Accurate Measures of Fatness and Obesity in Social Science Research. *Journal of Health Economics*, 27, 519-529

Butland, B., Jebb, S., Kopelman, P., McPherson, K., Thomas, S., Mardell, J., et al. (2007). Foresight: Tackling obesities: Future choices e Project report (2nd ed.). Government Office for Science. Retrieved from www.foresight.gov.uk.

Calle, E.E., Thun, M.J., Petrelli, J.M., Rodriguez, C., & Heath, Jr., C.W. (1999). Body-Mass Index and Mortality in a Prospective Cohort of U.S. Adults. *N Engl J Med*, 341, 1097-1105

Carbone, J., Kverndokk, K., & Røgeberg, O., J. (2003) Smoking and Health Investments: Impacts of Health Adaptation and Damage Reversibility. *HEALTH ECONOMICS RESEARCH PROGRAMME Working Paper:* 12

Cawley, J., & Meyerhoefer, C. (2011). The medical care cost of obesity: An instrumental variable approach. *Journal of Health Economics.* Web version

Chen, Y., Jiang, Y., Mao, Y. (2007). Hospital admissions associated with body mass index in Canadian adults. *International Journal of Obesity,* 31, 962–967

Chu, N., Wang, S., Chang, H., & Wu, D. (2010). Medical Services Utilization and Expenditure of Obesity-Related Disorders in Taiwanese Adults. *Value in Health,* 13, 829-836

Contoyannis, P., Jones, A.M. (2004). Socio-economic status, health and lifestyle. *Journal of Health Economics*, 23, 965–995

Czernichow, S., Kengne, A.P., Stamatakis, E., Hamer, M., Batty, G.D. (2011) Body mass index, waist circumference and waist–hip ratio: which is the better discriminator of cardiovascular disease mortality risk? Evidence from an individual-participant meta-analysis of 82 864 participants from nine cohort studies. O*besity Reviews*, web version

The Counterweight Project Team. (2005a). Obesity Impacts on General Practice Appointments, *Obesity Research,* 13, 1442-1449

The Counterweight Project Team. (2005b). The impact of obesity on drug prescribing in primary care. *Br J Gen Prac*, 55, 743-749

The Counterweight Project Team. (2008). The Counterweight programme: Prevalence of CVD risk factors by body mass index and the impact of 10% weight change. *Obesity Research & Clinical Practice*, 2, 15-27

The Counterweight Project Team. (2010). Long-term cost-effectiveness of weight management in primary care. *Int J Clin Pract,* 64, 828–29

CSDH. (2008). *Closing the gap in a generation: health equity through action on the social determinants of health. Final Report of the Commission on Social Determinants of Health.* Geneva, World Health Organization

Davey Smith, G. and Shipley, M. (1991) Confounding of occupation and smoking: Its magnitude and consequences. *Social Science and Medicine,*32, 1297–1300

Department of Health (1998). *Independent Inquiry into Inequalities in Health – Report*. London: HMSO.

Department of Health. (2008). *Health inequalities: Progress and next steps*, London

Department of Health (2010) Healthy lives, healthy people: our strategy for public health in England. Available at: http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/@ps/documents/digitalasset/dh_122347.pdf

Department of health (2011). Obesity General Information. Available at: www.dh.gov.uk

Dolan, P. (1997). Modelling valuation for EuroQol health states. *Medical Care,* 35, 351–363

Doll, H.A., Petersen, S.E.K., & Stewart-Brown, S.L. (2000). Obesity and physical and emotional well-being: associations between body mass index, chronic illness, and the physical and mental components of the SF-36 questionnaire. *Obes Res,* 8, 160–170

Dzien, A., Pfeiffer, K.P., Dzien-Bischinger, C., Hoppichler, F.,& Lechleitner, M. (2003). The Influence of Obesity on the Frequency and Distribution of Medication. *Acta Medica Austriaca*, 30, 51-54

Euroqol Group. (1990). Euroqol – a new facility for the measurement of health-related quality-of-life. *Health Policy,* 16, 199–208

Fine, J.T., Colditz, G.A., Coakley, E.H., Moseley, G., Manson, J.E., Willett, W.C., . . . Kawachi, I. (1999). A prospective study of weight change and health-related quality of life in women. *JAMA,* 282, 2136–2142

Flegal, K.M., Graubard, B.I., Williamson, D.F., & Gail, M.H. (2005). Excess Deaths Associated With Underweight, Overweight, and Obesity, *JAMA*, 293(15):1861-1867

Folmann, N.B., Skovgaard Bossen, K., Willaing, I., Sørensen, J., Sahl Andersen, J., Ladelund, S., Jørgensen T. (2006). Obesity, Hospital Services use and Costs, in Kristian Bolin, John Cawley (ed.) *The Economics of Obesity (Advances in Health Economics and Health Services Research, Volume 17)*, Emerald Group Publishing Limited, pp.319-332

Fontaine, K.R., Redden, D.T., Wang, C., Westfall, A.O., Allison, D.B. (2003). Years of Life Lost Due to Obesity. *JAMA*, 289(2):187-193

Ford, E.S., Moriarty, D.G., Zack, M.M., Mokdad, A.H., & Chapman, D.P. (2001). Self-reported body mass index and health-related quality of life: findings from the Behavioural Risk Factor Surveillance System. *Obes Res,* 9, 21–31

Forster, M., & Jones, A.M. (2001). The role of tobacco taxes in starting and quitting smoking: duration analysis of British data. *Journal of the Royal Statistical Society Series A*, 164, 517–547

Freedman, D.S., Perry, G., (2000). Body composition and health status among children and adolescents. *Preventive Medicine*, 31, S34–S53

Freedman, D.M., Ron, E., Ballard-Barbash, R., Doody, M.M., & Linet, M.S. (2006). Body mass index and all-cause mortality in a nationwide US cohort. *International Journal of Obesity,* 30**,** 822–829

Golden, S., Boulware, L.E., Berkenblit, G., Brancati, F., Chander G., & Marinopoulos S., *et al.*, (2003). Use of glycated hemoglobin and microalbuminuria in the monitoring of diabetes mellitus, *Evid Rep Technol Assess*, 84, 1–6

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

Graham, H. (2004). "Socioeconomic inequalities in health in the UK: Evidence on patterns and determinants". A short report for the Disability Rights Commission. Can be found at: *http://www.leeds.ac.uk/disability-tudies/archiveuk/graham/socioeconomic_inequalities.pdf*

Greene, W. (2010). Testing hypotheses about interaction terms in nonlinear models. *Economics Letters*, 7, 291-296

Groessl, E.J., Kaplan, R.M., Barrett-Connor, E., & Ganiats, T.G. (2004). Body mass index and quality of well-being in a community of older adults. *Am J Prev Med,* 26, 126–129

Grossman, M. (1972). On the Concept of Health Capital and the Demand for Health. *Journal of Political Economy*, 80, 223-255

Grossman, M. (2000). Chapter 7: The human capital model, in: Culyer A.J., & Newhouse, J.P., eds. Handbook of health economics, Vol. 1, *Elsevier Science B. V*, 347-408

Guallar-Castillon, P., Lopez Garcia, E., Lozano Palacios, L., Gutierrez-Fisac , J.L., Banegas Banegas, J.R., Lafuente Urdinguio, P.J., *et al*. (2002). The relationship of overweight and obesity with subjective health and use of health-care services among Spanish women. *Int J Obes Relat Metab Disord,* 26, 247-52

Gupta N.D., & Greve, J. (2011). Overweight and obesity and the utilization of primary care physicians. *Health Economics*, online version

Han, E., Truesdale, K.P., Taber, D.R., Cai, J., Juhaeri, J., & Stevens, J. (2009). Impact of overweight and obesity on hospitalization: race and gender differences. *International Journal of Obesity,* 33, 249–256

Han, T.S., Tijhuis, M.A., Lean, M.E., & Seidell, J.C. (1998). Quality of life in relation to overweight and body fat distribution. *Am J Public Health*, 88, 1814–1820

Hauck, K., & Hollingsworth, B. (2010). The Impact of Severe Obesity on Hospital Length of Stay. *Medical care*, 48, 335-340

Health and Lifestyle Survey – University of Cambridge Clinical School 2009. Deaths and Cancer 2009. *University of Cambridge*. [distributor]

The Health and Social Care Information Centre (2011). *Provisional Monthly Patient Reported Outcome Measures (PROMs) in England: A guide to PROMs methodology.* Can be found at: *http://www.hesonline.nhs.uk/*

Health Reform Evaluation Programme (2011). *Publications.* [online] Available at: *http://hrep.lshtm.ac.uk/publications/index.htm*

Heo, M., Allison, D.B., Faith, M.S., Zhu, S., & Fontaine, K.R. (2003). Obesity and quality of life: mediating effects of pain and comorbidities. *Obes Res,* 11, 209–216

Hjort, N.L. (1992). On inference in parametric survival data models. *Int. Statist. Rev*. 60, 355-387

House of Commons Health Select Committee. (2004). Obesity: Third Report of Session 2003/04. London: The Stationery Office

Huang, C., & Shields.T. (2000). Interpretation of Interaction Effects in Logit and Probit Analyses: Reconsidering the Relationship between Registration Laws, Education, and Voter Turnout. *American Politics Research,* 28, 80–95.

Jenkins, S.P. (2008). Survival Analysis with Stata: Module EC968. Accessed: May, 2011. Can be found at: *http://www.iser.essex.ac.uk/resources/survival-analysis-with-stata-module-ec968*

Jia, H., & Lubetkin, E.I. (2005). The impact of obesity on health-related quality-of-life in the general adult US population. *J Public Health (Oxford)*, 27, 156–164

Jones, A.M. (2005). *Applied Econometrics for Health Economists A Practical Guide, 2nd Edition.* Office of Health Economics, London

Jones, A.M., & Shurer, S. (2011) How does heterogeneity shape the socioeconomic gradient in health satisfaction?. *Journal of Applied Econometrics*, 549-579

Kish, L., & Frankel, MR. (1974). Inference from complex samples. *Journal of the Royal Statistical Society Series B*, 36, 1–37

Kooiker, S., & Christiansen, T. (1995) Inequalities in health: the interaction of circumstances and health related behaviour. *Sociology of Health & Illness*, 17, 495-524

Krieger, N., Williams, D.R., & Moss, N.E. (1997). MEASURING SOCIAL CLASS IN US PUBLIC HEALTH RESEARCH: Concepts, Methodologies, and Guidelines. *Annu. Rev. Public Health, 18, 341–78*

Krueger, P. M., & Chang, V. W. (2008) Being poor and coping with stress: health behaviours and the risk of death. *American Journal of Public Health*, 98, 889-896

Kolasinski, A.C., & Siegel, A.F. (2010). On the Economic Meaning of Interaction Term Coefficients in Non-Linear Binary Response Regression Models. *SSRN eLibrary*

Korn, E. L., & Graubard, B.I. (1999). Analysis of Health Surveys. New York, USA. Wiley

Kortt, M.A., & Clarke, P.M. (2005). Estimating utility values for health states of overweight and obese individuals using the SF-36. *Quality of Life Research*, 14, 2177–2185

Laaksonen, M., Sarlio-Lahteenkorva, S., Leino-Arjas, P., Martikainen, P. & Lahelma, E. (2005). Body weight and health status: importance of socioeconomic position and working conditions. *Obesity Research*, 13, 2169-2177

Larsson, U., Karlsson, J., & Sullivan, M. (2002). Impact of overweight and obesity on health-related quality of life – a Swedish population study. *Int J Obes Relat Metab Disord,* 26, 417–424

Lawlor, D.A., Hart, C.L., Hole, D.J., Davey Smith G. (2006). Reverse causality and confounding and the associations of overweight and obesity with mortality. *Obesity* 14, 294-304

Lean, M.E., Han, T.S., & Seidell, J.C. (1998). Impairment of health and quality of life in people with large waist circumference. *Lancet,* 351, 853–856

Lean, M.E., Han, T.S., & Seidell, J.C. (1999). Impairment of health and quality of life using new US federal guidelines for the identification of obesity. *Arch Intern Med*, 159, 837–843

Lee, A.J., Morgan, C.L., Morrissey, M., Wittrup-Jensen, K.U., Kennedy-Martin, T., Currie, C.J. (2005). Evaluation of the association between the EQ-5D (health-related utility) and body mass index (obesity) in hospital-treated people with Type 1 diabetes, Type 2 diabetes and with no diagnosed diabetes. *Diabet Med,* 22, 1482–86

Le Pen, C., Levy, E., Loos, F., Banzet, M.N., & Basdevant, A. (1998). 'Specific' scale compared with 'generic' scale: a double measurement of the quality of life in a French community sample of obese subjects. *J Epidemiol Community Health,* 52, 445–450

Luchsinger, J.A., Lee, W.N., Carrasquillo, O., Rabinowitz, D., Shea, S. (2003). Body mass index and hospitalization in the elderly. *Journal of the American Geriatrics Society,* 11, 1615-1620

Macran, S. (2004). The relationship between Body Mass Index and Health-Related Quality of Life. *CHE Discussion Paper 190,* Centre for Health Economics, University of York

Makela, P., & Paljarvi, T. (2008). Do consequences of a given pattern of drinking vary by socioeconomic status? A mortality and hospitalisation follow-up for alcohol-related causes of the Finnish Drinking Habits Surveys. *Epidemiol Community Health,* 62, 728-733

Marang-van de Mheen, P. J., Davey Smith, G., & Hart, C.L. (1999) The health impact of smoking in manual and non-manual social class men and women: a test of the Blaxter hypothesis. *Social Science & Medicine*, 48, 1851-1856

Marmot, M., Rose, G., Shipley, M. and Hamilton, P., 1978. Employment grade and coronary heart disease in British civil servants. *Journal of Epidemiology and Community Health,* 32, 244–249

Marmot, M., (chair). (2010). *Strategic Review of Health Inequalities in England Post 2010 (Marmot Review)*. London, UK. The Marmot review

Masyn, K.E. (2003). Discrete-Time Survival Mixture Analysis for Single and Recurrent Events Using Latent Variables, *A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy*

Mayhew, L., Richardson, J., & Rickayzen, B. (2009). A study into the detrimental effects of obesity on life expectancy in the UK. Retrieved from: *http://www.actuaries.org.uk/*

McGuire, A., Henderson, J., & Mooney, G. (1994). *The economics of health care: An introductory text.* London, Routledge and Kegan Paul.

McLaren, L. (2007). Socioeconomic status and obesity. *Epidemiologic Reviews*, 29, 29–48

McPherson, K., Marsh, T., & Brown, M. (2007). Tackling Obesities: Future Choices – Modelling Future Trends in Obesity and the Impact on Health. Government Office for Science. Retrieved from: *www.foresight.gov.uk*

Mond, J. M., & Baune, B. T. (2009). Overweight, Medical Comorbidity and Health-related Quality of Life in a Community Sample of Women and Men. *Obesity*, 27, 1-8

Moon, G., Quarendon, G., Barnard, S., Twigg, L., & Blyth′ B. (2007) Fat nation: Deciphering the distinctive geographies of obesity in England. *Social Science & Medicine*, 65, 20-31

Morris, S. (2007). The impact of obesity on employment. *Labour Economics*, 14, 413-433

Morris, S., Sutton, M., & Gravelle, H. (2005). Inequity and inequality in the use of health care in England: an empirical investigation. *Social Science & Medicine*, 60, 1251-1266

Muurinen, J. (1982). Demand for health: a generalised Grossman model, *Journal of Health Economics,* 1:5-28

MVH Group. (1995). *The Measurement and Valuation of Health: Final Report on the Modelling of Valuation Tariffs*. York, UK. Centre for Health Economics, University of York

National Audit Office. (2001). *Tackling Obesity in England*. London: National Audit Office

National Centre for Social Research and Department of Epidemiology and Public Health, University College London (UCL), *Health Survey for England, 1998 - 2008* [data file] Colchester, Essex: UK Data Archive [distributor]

National Centre for Social Research and University College London. Department of Epidemiology and Public Health. (2009). Health Survey for England 2008 – trend tables. The NHS Information Centre. Retrieved from: *www.ic.nhs.uk*

National Institute of Clinical Excellence (NICE). (2006). *Obesity: guidance on the prevention, identification, assessment and management of overweight and*

obesity in adults and children. London, UK. National Institute for Health and Clinical Excellence

National Institute of Clinical Excellence (NICE). (2008). *Guide to the Methods of Technology Appraisal*. London, UK. National Institute for Health and Clinical Excellence

National Obesity Observatory. (2010). *Adult obesity and socioeconomic status*. NOO data briefing. Can be found at: *http://www.noo.org.uk/uploads/doc/vid_7929_Adult%20Socioeco%20Data%20Briefing%20October%202010.pdf*

Newey, W.K., Powell, J.L., Walker, J.R. (1990). Semiparametric estimation of selection models: some empirical results. *Am Econ Rev*, 80: 324–328

Newman, A., Yanez, D., Harris, T., Duxbury, A., Enright, P.L., Fried, L.P. (2001). Weight change in old age and its association with mortality. *Journal of American Geriatrics Society*, 49, 1309 - 1318

NHLBI [National Heart, Lung and Blood Institute]. (1998). *Clinical guidelines on the identification, evaluation and treatment of overweight and obesity in adults*. NIH Publication Number 98-4083. New York: National Institute of Health.

Norton, E.C., Wang, H., & Ai, C. (2004). Computing interaction effects and standard errors in logit and Probit models. *The Stata Journal*, 4, 154-167

Pampel, F. R., & Rogers, R. G. (2004). Socioeconomic status, smoking and health: A test of competing theories of cumulative advantage. *Journal of Health and Social Behaviour,* 45, 306-321

Pearson, W., Bhat-Schelbert, K., Ford, E., Mokdad, A. (2009). The impact of obesity on time spent with the provider and number of medications managed during

office-based physician visits using a cross-sectional, national health survey. *BMC Public Health*, 9, 436

Peytremann-Bridevaux, I., & Santos-Eggimann, B. (2007a). Healthcare utilization of overweight and obese Europeans aged 50–79 years. *J Public Health,* 15, 377–384

Peytremann-Bridevaux, I., & Santos-Eggimann, B. (2007b). Use of Preventive Services of Overweight and Obese Europeans Aged 50–79 Years. *Journal of General Internal Medicine,* 22, 923-929

Pickard, A. S., Neary, M. P., & Cella, D. (2007). Estimation of minimally important differences in EQ-5D utility and VAS scores in cancer. *Health and Quality of Life Outcomes*, 5, 70

Pohlmeier, W., & Ulrich, V., (1995). An econometric model of the two-part decision making process in the demand for health care. *J. Human Resources,* 30, 339–361

Popoola, F.A. (2004). Increase in Obesity and Health-Care Use, from the Health Survey for England, Ages 50 to 69, 1998, *Obesity Surgery, 14, 1258-1262*

Preamble to the Constitution of the World Health Organization as adopted by the International Health Conference, New York, 19-22 June, 1946; signed on 22 July 1946 by the representatives of 61 States (Official Records of the World Health Organization, no. 2, p. 100) and entered into force on 7 April 1948

Quesenberry Jr., C.P., Caan, B., Jacobson, A. (1998). Obesity, health services use, and health care costs among members of a Health Maintenance Organization. *Archives of internal medicine,* 158, 466-472

Redekop, W.K., Koopmanschap, M.A., Stolk, R.P., Rutten, G.E., Wolffenbuttel, B.H., Niessen, L.W. (2002). Health-related quality of life and treatment satisfaction in Dutch patients with type 2 diabetes. *Diabetes Care,* 25, 458–463

Reidpath, D.D., Crawford, D., Tilgner, L., & Gibbons, C. (2002). Relationship between body mass index and the use of health care services in Australia. *Obes Res*, 10, 526–531

Sach, T.H., Barton, G.R., Doherty, M., Muir, K.R., Jenkinson, C., & Avery, A.J. (2007). The relationship between body mass index and health-related quality of life: comparing the EQ-5D, EuroQol VAS and SF-6D. *International Journal of Obesity,* 31, 189–196

Saez, M., Saurina, C., Coenders, G., & González-Raya, S. (2006). Use of primary health care services according to the different degrees of obesity in the Girona Health Region, Spain. *Health Economics*, 15, 173–193

Sansone, R.A., Sansone, L.A., & Wiederman, M.W. (1998). The relationship between obesity and medical utilization among women in a primary care setting. *The international journal of eating disorders,* 23, 161-167

Schafer, M.H., & Ferraro, K.F. (2007). Obesity and hospitalization over the adult life course: does duration of exposure increase use? *J Health Soc Behav*, 48, 434–449

Schafer, M. H., Ferraro, K. F., & Williams, S. R. (2011). Low socioeconomic status and body mass index as risk factors for inflammation in older adults: conjoint influence of C-reactive protein? *Journal of Gerontology: MEDICAL SCIENCES,* doi: 10.1093/gerona/glr009

Seidell, J.C., Verschuren, W.M., van Leer, E. M., Kromhout, D. (1996). Overweight, underweight, and mortality. A prospective study of 48,287 men and women. *Arch Intern Med*. 13, 958-63

Singh-Manoux, A., Britton, A., Kivimaki, M., Gueguen, A., Halcox J., & Marmot, M. (2008). Socioeconomic status moderates the association between carotid intima-media thickness and cognition in midlife: Evidence from the Whitehall II study. *Atherosclerosis,* 197, 541-548

Solar, O., & Irwin, A. (2007). *A conceptual framework for action on the social determinants of health*. Discussion paper for the Commission on Social Determinants of Health. Geneva, World Health Organization

Sribney, W. (2005) Likelihood-ratio test after survey/robust ML estimation. *StataCorp,* Can be found at: *http://www.stata.com/support/faqs/stat/lrtest.html*

StataCorp. (2009). *Stata: Reference manual: Release 11. Stata survival analysis and epidemiological tables*. College Station, TX: StataCorp LP.

Stevens, J., Cai, J., Pamuk, E.R., Williamson, D.F., Thun, M.J., & Wood, J.L. (1998). The Effect of Age on the Association between Body-Mass Index and Mortality. *N Engl J Med*, 338, 1-7

Stevens, J., Juhaeri, Cai, J., Thun, M. (2000). The impact of smoking and pre-existing illness on the relationship between body-mass index and mortality. *Nutrition Research*, 20, 9, 1259-1277

Sutton, M., Gravelle, H., Morris, S., Leyland, A., Windmeijer, F., & Dibben, C., *et al.* (2002). Allocation of resources to English areas; individual and small area determinants of morbidity and use of healthcare resources. *Report to the Department of Health, Information and Statistics Division*, Edinburgh

Taylor, M.F., Brice, J., Buck, N., & Prentice-Lane, E. (2004). *British household panel survey user manual volume A: introduction, technical report and appendices.* Colchester: University of Essex

Thrane, C. (2006). Explaining educational-related inequalities in health: Mediation and moderator models, *Social Science & Medicine*, 62, 467-478

Thomson ISI (2010) *Thomson Reuters Journal Citation Reports,* Can be found at: *http://science.thomsonreuters.com/*

Trakas, K., Lawrence, K., Shear, N.H. (1999). Utilization of health care resources by obese Canadians. *CMAJ*, 160: 1457–1462

Tsai, S.P., Donnelly, R.P., Wendt, J.K. (2006). Obesity and mortality in a prospective study of a middle-aged industrial population. *J Occup Environ Med,* 48, 22-7

UCLA Academic Technology Services. Survival Analysis with Stata. [Accessed 15th of June, 2011]

Vallejo-Torres, L., & Morris, S. (2010). Factors associated with the use of primary care services: the role of practice nurses, *Eur J Health Econ*

Vallejo-Torres, L., & Morris, S. (2010). The contribution of smoking and obesity to income-related inequalities in health in England. *Social Science & Medicine*, 71, 1189-1198

van Dijk, L., Otters, H.B., & Schuit, A.J. (2006). Moderately overweight and obese patients in general practice: a population based survey. *BMC Family Practice,* 7

van Doorslaer, E. (1987) E.K.A. van Doorslaer, Health, Knowledge and the Demand for Medical Care, Assen, New York (1987).

Vapattanawong, P., Aekplakorn, W., Rakchanyaban, U., Prasartkul, P., Porapakkham, Y., (2010). Obesity and mortality among older Thais: a four year follow up study. *BMC Public Health*, 10, 604

Veugelers, P.J., & Yip, A.M. (2003). Socioeconomic disparities in health care use: Does universal coverage reduce inequalities in health?. *J Epidemiol Community Health*, 57, 424-428

Von Lengerke, T., Happich, M., Reitmeir, P., & John, J. (2005). Utilization of Out- and Inpatient Health Services by Obese Adults: a Population-Based Study in the Augsburg Region, Germany *Gesundheitswesen*, 1, S150–S157

Walters, S. J., & Brazier, J., E. (2005) Comparison of the minimally important difference for two health state utility measures: EQ-5D and SF-6D. *Quality of Life Research,* 14, 1523-1532

Wardle, J., Waller, J., Jarvis, M.J. (2002). Sex Differences in the Association of Socioeconomic Status with Obesity. *American Journal of Public Health,* 92, 1299-1304

Wee, H-L., Cheung, Y-B., Loke, W-C., Tan, C-B., Chow, M-H., Li, S-C., . . . Thumboo, J. (2008). The Association of Body Mass Index with Health-Related Quality of Life: An Exploratory Study in a Multiethnic Asian Population. *Value in Health,* 11, 105-114

Wiczinski, E., Doring, A., John, J., & von Lengerke, T. (2009). Obesity and health-related quality of life: Does social support moderate existing associations?. *British Journal of Health Psychology*, 14, 717–734

Wildenschild, C., Kjøller, M., Sabroe, S., Erlandsen, M., & Heitmann, B.L. (2011). Change in the prevalence of obesity and use of health care in Denmark: an observational study. *Clin Epidemiol*., 3, 31–41

Williamson, D. L. (2000) Health behaviours and health: evidence that the relationship is not conditional on income adequacy. *Social Science & Medicine*, 51, 1741-1754

Wooldrige, J.M. (2002). *Econometric Analysis of Cross Section and Panel Data*. The MIT Press Cambridge, Massachusetts London, England

World Health Organization (WHO) Expert Committee on Physical Status. (1995). *The use and interpretation of anthropometry. Report of a World Health Organization Expert Committee. Geneva: WHO*. Technical Support, Series 854

World Health Organisation (WHO) (2011). Health Topics. Retrieved from: http://www.who.int/topics/en/

World Health Organisation (WHO) (2010). Obesity and overweight. Retrieved from: *http://www.who.int/mediacentre/factsheets/fs311/en/index.html*

Yan, L.L., Daviglus, M.L., Liu, K., Pirzada, A., Garside, D.B., Schiffer, L. . . . Greenland, P. (2004). BMI and health-related quality of life in adults 65 years and older. *Obes Res,* 12, 69–76

Yan, L.L., Daviglus, M.L., Liu, K., Stamler, J., Wang, R., Pirzada, A., *et al*. (2006). Midlife body mass index and hospitalization and mortality in older age. *JAMA*, 295, 190–198

Zizza, C., Herring, A.H., Stevens, J., & Popkin, B.M. (2004). Length of hospital stays among obese individuals. *Am J Public Health*, 94, 1587–1591

Appendices

Appendices to Chapter 2

Appendix 2.1: Search terms and number of identified studies

Ovid MEDLINE(R) Daily Update, Ovid OLDMEDLINE(R) 1946 to 1965, Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) 1948 to Present

Search	Query	Results
#1	"Health Status Indicators"[Mesh]	151880
#2	"Life Expectancy"[Mesh]	12027
#3	"Mortality"[Mesh]	232329
#4	#1 OR #2 OR #3	383415
#5	"Socioeconomic Factors"[Mesh]	284754
#6	#1 OR #2 OR #3 AND #5	17942
#7	"Obesity"[Mesh]	107834
#8	#1 OR #2 OR #3 AND #5 AND #7	224

Ovid MEDLINE(R) Daily Update, Ovid OLDMEDLINE(R) 1946 to 1965, Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) 1948 to Present

Search	Query	Results
#1	"Health Status Indicators"[Mesh]	151880
#2	"Life Expectancy"[Mesh]	12027
#3	"Mortality"[Mesh]	232329
#4	#1 OR #2 OR #3	383415
#5	"Socioeconomic Factors"[Mesh]	284754
#6	#1 OR #2 OR #3 AND #5	17942
#7	"Smoking"[Mesh]	102909
#8	#1 OR #2 OR #3 AND #5 AND #7	693

Ovid MEDLINE(R) Daily Update, Ovid OLDMEDLINE(R) 1946 to 1965, Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) 1948 to Present

Search	Query	Results
#1	"Health Status Indicators"[Mesh]	151880
#2	"Life Expectancy"[Mesh]	12027
#3	"Mortality"[Mesh]	232329
#4	#1 OR #2 OR #3	383415
#5	"Socioeconomic Factors"[Mesh]	284754
#6	#1 OR #2 OR #3 AND #5	17942
#7	"Alcohol Drinking"[Mesh]	42601
#8	#1 OR #2 OR #3 AND #5 AND #7	263

Ovid MEDLINE(R) Daily Update, Ovid OLDMEDLINE(R) 1946 to 1965, Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) 1948 to Present

Search	Query	Results
#1	"Health Status Indicators"[Mesh]	151880
#2	"Life Expectancy"[Mesh]	12027
#3	"Mortality"[Mesh]	232329
#4	#1 OR #2 OR #3	383415
#5	"Socioeconomic Factors"[Mesh]	284754
#6	#1 OR #2 OR #3 AND #5	17942
#7	"Drug Users"[Mesh]	538
#8	#1 OR #2 OR #3 AND #5 AND #7	3

EMBASE 1980 to 2011 Week 25

Search	Query	Results
#1	exp health status/	94565
#2	exp survival/	386229
#3	#1 OR #2	475718
#4	exp socioeconomics/	131838
#5	#1 OR #2 AND #4	11649
#6	exp obesity/	187472
#7	#1 OR #2 AND #4 AND #6	544

EMBASE 1980 to 2011 Week 25

Search	Query	Results
#1	exp health status/	94565
#2	exp survival/	386229
#3	#1 OR #2	475718
#4	exp socioeconomics/	131838
#5	#1 OR #2 AND #4	11649
#6	exp smoking/	163264
#7	#1 OR #2 AND #4 AND #6	846

EMBASE 1980 to 2011 Week 25

Search	Query	Results
#1	exp health status/	94565
#2	exp survival/	386229
#3	#1 OR #2	475718
#4	exp socioeconomics/	131838
#5	#1 OR #2 AND #4	11649
#6	exp drinking behaviour/	27439
#7	#1 OR #2 AND #4 AND #6	114

EMBASE 1980 to 2011 Week 25

Search	Query	Results
#1	exp health status/	94565
#2	exp survival/	386229
#3	#1 OR #2	475718
#4	exp socioeconomics/	131838
#5	#1 OR #2 AND #4	11649
#6	exp drug abuse/	51597
#7	#1 OR #2 AND #4 AND #6	83

Appendices to Chapter 3

Appendix 3.1: Descriptive statistics by BMI category

	Whole sample	Normal weight	Overweight	Class I obesity	Class II / III obesity	*P* value
Observations						
Number	33,105	11,796	12,814	5,656	2379	
%	100	36	39	17	7	
EQ-5D score	0.871	0.9	0.876	0.84	0.782	<0.01
Predicted SES measure	29,652	30,914	30,159	27,831	25,821	<0.01
Missing income	14	14	14	14	15	0.46
SES quartile						
1 (most deprived)	25	22	24	29	35	
2	25	25	25	26	26	
3	25	26	26	24	23	
4 (least deprived)	25	28	26	21	17	<0.01
Gender						
Men	46	38	53	50	34	
Women	54	62	47	50	66	<0.01
Age (years)	48	43	50	52	50	<0.01
Ethnicity						
White	94	93	94	95	95	
Black Caribbean	1	1	1	1	1	
Black African	1	1	1	1	1	
Indian	2	2	2	1	1	
Pakistani	1	1	1	1	1	
Bangladeshi	<1	<1	<1	<1	<1	
Chinese	<1	<1	<1	<1	<1	
Other ethnic group	2	2	1	1	1	
Missing ethnicity	<1	<1	<1	<1	<1	<0.01
Marital status						
Married	57	48	62	65	61	
Single	17	25	13	10	14	
Separated	2	2	2	2	2	
Divorced	6	6	6	6	6	
Widowed	7	6	7	7	8	
Cohabiting	11	12	10	10	9	<0.01
Smoking status						
Never smoked	44	47	44	42	44	
Ex occasional smoker	6	6	6	5	5	
Ex regular smoker	26	19	29	33	32	
Current cigarette smoker	24	28	22	20	19	<0.01
GOR of residence						
North East	6	6	6	7	7	
North West	14	14	14	14	14	
Yorkshire and The Humber	10	10	10	11	11	
East Midlands	10	9	10	10	12	
West Midlands	11	10	11	12	13	
East of England	12	11	12	12	11	
London	10	12	9	8	8	
South West	10	10	10	10	10	
South East	17	18	17	16	14	<0.01
Survey year						
2003	35	36	35	35	32	
2004	15	15	16	15	15	
2005	17	17	17	17	18	
2006	32	32	32	33	34	0.03
Educational qualifications						
Degree or equivalent	19	23	19	15	12	
Higher education below degree	12	11	13	12	12	
NVQ3/GCE A Level or equiv.	13	15	12	11	11	
NVQ2/GCE O Level or equiv.	24	25	23	23	25	
NVQ1/CSE other grade or equiv.	5	7	5	5	6	
Foreign/other	3	3	3	3	3	
No qualification	25	21	25	31	32	<0.01

Table continued

Social Class of HRP						
Professional	8	9	8	6	5	
Managerial technical	36	38	37	33	30	
Skilled non-manual	15	16	15	15	14	
Skilled manual	24	21	24	27	28	
Semi-skilled manual	13	13	13	15	17	
Unskilled manual	4	4	4	4	5	<0.01
Cars owned by household						
Household has no car	16	17	14	15	19	
One	43	41	43	44	44	
Two	33	33	34	32	28	
Three or more	9	9	9	9	9	<0.01
Economic activity status last week						
In paid employment/self-employed	67	71	66	63	62	
Going to school or college full time	1	2	1	1	1	
On a government training scheme	<1	<1	<1	<1	<1	
Doing unpaid work	<1	<1	<1	<1	<1	
Waiting to take up paid work	<1	<1	<1	<1	<1	
Looking for paid work	1	1	1	1	1	
Temporary sickness	<1	<1	<1	<1	1	
Long-term sickness	4	3	3	5	7	
Retired from paid work	23	18	26	27	23	
Looking after home or family	4	4	3	3	5	
Doing something else	<1	<1	<1	<1	<1	<0.01
Housing tenure						
Own outright	32	29	34	34	27	
Buying with mortgage or loan	46	48	46	43	43	
Pay part rent part mortgage	<1	<1	<1	<1	1	
Rent	21	22	18	21	27	
Live rent free and/or squatting	1	1	1	1	1	<0.01
Bedrooms in household						
One	6	6	5	5	7	
Two	21	21	21	21	23	
Three	50	48	51	54	53	
Four	18	19	18	16	14	
Five or more	5	6	5	4	3	<0.01
IMD quintile (SOA level)						
1 (least deprived)	23	23	24	20	17	
2	22	22	23	22	20	
3	20	20	21	21	20	
4	20	20	18	21	23	
5 (most deprived)	15	15	14	16	20	<0.01
Income support claimant						
No	95	95	95	94	92	
Yes	5	5	5	6	8	<0.01
No. longstanding illnesses						
None	55	62	54	45	39	
One	28	24	29	31	29	
Two	12	9	12	15	18	
Three	4	3	4	6	10	
Four or more	2	1	2	3	4	<0.01
Limiting longstanding illness	24	23	24	30	37	<0.01
Type of longstanding illness						
Neoplasms and benign growths	2	2	2	2	2	0.144
Endocrine and metabolic	5	1	7	11	15	<0.01
Mental disorders	3	5	3	4	6	<0.01
Nervous system	4	4	4	4	4	0.256
Eye complaints	1	2	2	2	2	0.039
Ear complaints	2	2	3	2	2	<0.01
Heart and circulatory system	14	7	13	19	22	<0.01
Respiratory system	12	12	9	9	13	<0.01
Digestive system	7	5	5	6	6	<0.01
Genito-urinary system	3	2	2	3	3	<0.01
Skin complaints	3	1	2	2	3	<0.01
Musculoskeletal system	42	13	20	25	30	<0.01
Infectious disease	<1	<1	<1	<1	<1	0.338
Blood and related organs	2	1	1	1	1	0.544
Acute illnesses days the last two weeks						

Table continued

One to three	5	5	5	5	5	0.324
Four to six	3	2	2	3	4	<0.01
Seven to thirteen	3	2	3	3	4	<0.01
Fourteen	6	5	5	7	10	<0.01

Source: HSE 2003-2004

Notes

Other than for "Observations" all statistics are either mean (for continuous variables) or % (categorical variables). "<1" indicates non-zero values less than one; "0" indicates zero values.

GOR = Government Office Region. HRP = household reference person. IMD = Index of Multiple Deprivation. SOA = super output area.

Appendix 3.2: Association between obesity and EQ-5D score (controlling for individual and household characteristics); predicted EQ-5D scores are computed by setting the individual and household characteristics to the mean values of their SES group

	Normal weight				Overweight				Class I obesity				Class II / III obesity				Equal
	Coef.	**t**	**Mean**	**SE**	**Coef.**	**t**	**Mean**	**SE**	**Coef.**	**t**	**Mean**	**SE**	**Coef.**	**t**	**Mean**	**SE**	**means**
Men																	
All	Base category		0.898	0.002	-0.002	-0.45	0.897	0.002	-0.02	-4.25	0.878	0.003	-0.067	-7.2	0.831	0.007	*P*<0.01
SES quartile																	
1 (most deprived)	Base category		0.822	0.006	-0.012	-1.22	0.810	0.006	-0.044	-3.35	0.779	0.009	-0.139	-5.81	0.683	0.019	*P*<0.01
2	0.039	4.52	0.904	0.004	0.037	4.01	0.902	0.004	0.013	1.17	0.878	0.006	-0.036	-1.87	0.829	0.014	*P*<0.01
3	0.033	3.63	0.919	0.003	0.036	3.89	0.923	0.003	0.025	2.35	0.911	0.005	0.001	0.09	0.888	0.010	*P*=0.04
4 (least deprived)	0.021	2.00	0.937	0.003	0.023	2.13	0.939	0.002	0.017	1.48	0.933	0.004	0.004	0.27	0.920	0.009	*P*=0.35
Equal means	*P*<0.01				*P*<0.01				*P*<0.01				*P*<0.01				
Women																	
All	Base category		0.889	0.002	-0.014	-4.35	0.875	0.002	-0.044	-9.53	0.845	0.004	-0.100	-14.06	0.789	0.007	*P*<0.01
SES quartile																	
1 (most deprived)	Base category		0.829	0.006	-0.028	-3.31	0.801	0.006	-0.061	-5.81	0.768	0.009	-0.141	-9.99	0.687	0.012	*P*<0.01
2	0.026	3.72	0.903	0.003	0.006	0.76	0.882	0.005	-0.03	-2.91	0.846	0.008	-0.074	-5.14	0.802	0.012	*P*<0.01
3	0.011	1.42	0.911	0.003	0.004	0.43	0.903	0.005	-0.018	-1.81	0.881	0.007	-0.054	-3.87	0.845	0.012	*P*<0.01
4 (least deprived)	-0.007	-0.69	0.924	0.003	-0.011	-1.07	0.919	0.004	-0.033	-2.61	0.898	0.007	-0.064	-4.10	0.866	0.012	*P*<0.01
Equal means	*P*<0.01				*P*<0.01				*P*<0.01				*P*<0.01				

Notes

The dependent variable is EQ-5D score. The coefficients and t-statistics show the impact of the obesity variables on EQ-5D scores based on OLS. The mean and SE values are predicted mean EQ-5D scores, and their standard errors, for each BMI or BMI/SES category computed by setting the individual and household characteristics to the mean values of their SES group. The individual and household characteristics are age, age squared and age cubed, predicted equivalised total annual household income, ethnicity, marital status, smoking status, Government Office Region of residence and survey year.

Appendix 3.3: Association between obesity and EQ-5D score (controlling for individual and household characteristics); predicted EQ-5D scores are computed by setting the individual and household characteristics to the mean values of their obesity group

	Normal weight				Overweight				Class I obesity				Class II / III obesity				Equal
	Coef.	**t**	**Mean**	**SE**	**Coef.**	**t**	**Mean**	**SE**	**Coef.**	**T**	**Mean**	**SE**	**Coef.**	**t**	**Mean**	**SE**	**means**
Men																	
All	Base category		0.910	0.002	-0.002	-0.45	0.894	0.002	-0.02	-4.25	0.866	0.003	-0.067	-7.20	0.821	0.008	*P*<0.01
SES quartile																	
1 (most deprived)	Base category		0.887	0.006	-0.012	-1.22	0.858	0.007	-0.044	-3.35	0.819	0.009	-0.139	-5.81	0.726	0.019	*P*<0.01
2	0.039	4.52	0.926	0.004	0.037	4.01	0.908	0.004	0.013	1.17	0.875	0.006	-0.036	-1.87	0.829	0.039	*P*<0.01
3	0.033	3.63	0.919	0.003	0.036	3.89	0.907	0.003	0.025	2.35	0.887	0.005	0.001	0.09	0.867	0.033	*P*<0.01
4 (least deprived)	0.021	2.00	0.908	0.004	0.023	2.13	0.894	0.004	0.017	1.48	0.879	0.005	0.004	0.27	0.869	0.021	*P*<0.01
Equal means	*P*<0.01				*P*<0.01				*P*<0.01				*P*<0.01				
Women																	
All	Base category		0.904	0.002	-0.014	-4.35	0.866	0.003	-0.044	-9.53	0.827	0.004	-0.1	-14.06	0.775	0.007	*P*<0.01
SES quartile																	
1 (most deprived)	Base category		0.897	0.007	-0.028	-3.31	0.846	0.007	-0.061	-5.81	0.804	0.009	-0.141	-9.99	0.727	0.013	*P*<0.01
2	0.026	3.72	0.923	0.004	0.006	0.76	0.880	0.005	-0.03	-2.91	0.835	0.008	-0.074	-5.14	0.795	0.012	*P*<0.01
3	0.011	1.42	0.908	0.003	0.004	0.43	0.877	0.005	-0.018	-1.81	0.847	0.007	-0.054	-3.87	0.815	0.012	*P*<0.01
4 (least deprived)	-0.007	-0.69	0.890	0.005	-0.011	-1.07	0.862	0.006	-0.033	-2.61	0.832	0.009	-0.064	-4.10	0.804	0.013	*P*<0.01
Equal means	*P*<0.01				*P*<0.01				*P*=0.01				*P*<0.01				

Notes

The dependent variable is EQ-5D score. The coefficients and t-statistics show the impact of the obesity variables on EQ-5D scores based on OLS. The mean and SE values are predicted mean EQ-5D scores, and their standard errors, for each BMI or BMI/SES category computed by setting the individual and household characteristics to the mean values of their obesity group. The individual and household characteristics are age, age squared and age cubed, predicted equivalised total annual household income, ethnicity, marital status, smoking status, Government Office Region of residence and survey year.

Appendices Chapter 4

Appendix 4.1: Longstanding illnesses in the HSE

1. Cancer (neoplasm) 2. Diabetes 3. Other endocrine/metabolic 4. Mental illness/anxiety/depression/nerves 5. Mental handicap 6. Epilepsy/fits/convulsions 7. Migraine/headaches 8. Other problems of nervous system 9. Cataract/poor eye sight/blindness 10. Other eye complaints 11. Poor hearing/deafness 12. Tinnitus/noises in the ear 13. Menieres disease/ear complaints causing balance problems 14. Other ear complaints 15. Stroke/cerebral haemorrhage/cerebral thrombosis 16. Heart attack/angina 17. Hypertension/high blood pressure/blood pressure 18. Other heart problems 19. Piles/haemorrhoids including Varicose Veins in anus 20. Varicose veins/phlebitis in lower extremities 21. Other blood vessels/embolic	22. Bronchitis/emphysema 23. Asthma 24. Hayfever 25. Other respiratory complaints 26. Stomach ulcer/ulcer/abdominal hernia/rupture 27. Other digestive complaints 28. Complaints of bowel/colon 29. Complaints of teeth/mouth/tongue 30. Kidney complaint 31. Urinary tract infection 32. Other bladder problems/incontinence 33. Reproductive system disorders 34. Arthritis/rheumatism/fibrositis 35. Back problems/slipped disc/spine/neck 36. Other problems of bones/joints/muscles 37. Infectious and parasitic disease 38. Disorders of blood and blood forming organs and immunity disorders 39. Skin complaints 40. Other complaints 41. Unclassifiable (no other codable complaint) 42. Complaint no longer present

Appendix 4.2: Description of illness categories

Longstanding illness	**Description based on HSE user guide**
Diabetes	Incl. Hyperglycaemia
Stroke	Incl. stroke victim - partially paralysed and speech, Difficulty Hemiplegia, apoplexy, cerebral embolism, Cerebro - vascular accident
Heart attack or angina	Incl. coronary thrombosis, myocardial infarction
Hypertension/high blood pressure/blood pressure	

Appendix 4.3: Share of individuals with comorbidities by SES group

	Normal weight	Over-weight	Class I obesity	Class II/III obesity
Diabetes				
SES quartile				
1 (most deprived)				
Yes	2.5%	5.5%	10.4%	13.3%
No	97.5%	94.5%	89.6%	86.7%
2				
Yes	1.4%	3.1%	6.7%	8.7%
No	98.6%	96.9%	93.3%	91.3%
3				
Yes	1.5%	2.8%	4.8%	8.5%
No	98.5%	97.2%	95.2%	91.5%
4 (least deprived)				
Yes	1.2%	2.0%	4.1%	6.6%
No	98.8%	98.0%	95.9%	93.4%
Stroke				
SES quartile				
1 (most deprived)				
Yes	1.1%	1.7%	1.6%	1.5%
No	98.9%	98.3%	98.4%	98.5%
2				
Yes	0.3%	0.5%	1.1%	0.7%
No	99.7%	99.5%	98.9%	99.3%
3				
Yes	0.3%	0.3%	0.6%	0.7%
No	99.7%	99.7%	99.4%	99.3%
4 (least deprived)				
Yes	0.2%	0.2%	0.5%	0.5%
No	99.8%	99.8%	99.5%	99.5%
Hypertension				
SES quartile				
1 (most deprived)				
Yes	5.8%	10.4%	15.3%	15.5%
No	94.2%	89.6%	84.7%	84.5%
2				
Yes	3.2%	7.0%	10.4%	16.5%
No	96.8%	93.0%	89.6%	83.5%
3				
Yes	2.0%	6.1%	10.2%	13.5%
No	98.0%	93.9%	89.8%	86.5%
4 (least deprived)				
Yes	2.6%	5.5%	10.4%	15.2%
No	97.4%	94.5%	89.6%	84.8%
Heart attack and angina				
SES quartile				
1 (most deprived)				
Yes	2.9%	5.0%	6.5%	5.2%
No	97.1%	95.0%	93.5%	94.8%
2				
Yes	0.9%	1.9%	3.3%	3.5%
No	99.1%	98.1%	96.7%	96.5%
3				
Yes	0.7%	1.2%	1.4%	2.2%
No	99.3%	98.8%	98.6%	97.8%
4 (least deprived)				
Yes	0.3%	0.9%	1.2%	0.5%
No	99.7%	99.1%	98.8%	99.5%

Source: HSE 2003-2006

Appendices Chapter 5

Appendix 5.1: Summary statistics for individuals used in the duration studies

	Whole sample	Normal weight	Overweight	Obese
Predicted income	136	141	135	123
SES groups				
1 (most deprived)	25	23	24	30
2	25	24	26	27
3	25	25	26	27
4 (least deprived)	25	28	24	16
Missing income				
Yes	17	17	17	17
No	83	83	83	83
Age	58	57	59	58
Gender				
Male	46	42	54	35
Female	54	58	46	65
Ethnicity				
White European	97	98	97	97
Other	2	2	2	2
Not answered	1	1	1	0
Regular smoker				
Yes	29	35	24	22
No	71	65	76	78
Ex-smoker				
Yes	2	2	2	1
No	98	98	98	99
Marital status				
Married	75	75	77	70
Single	6	6	4	5
Separated	2	2	2	1
Divorced	4	4	3	5
Widowed	14	12	14	18
Area				
London	10	10	9	7
Wales	5	5	5	8
North	6	6	7	7
North West	13	13	13	12
Yorks/Humber	9	9	9	10
West Midlands	8	8	8	8
East Midlands	8	9	7	8
East Anglia	4	4	4	4
South West	9	9	9	10
South East	18	20	18	14
Scotland	10	8	10	13
Educational qualifications				
Degree or equivalent	12	14	10	7
Higher education below degree	8	9	8	8
NVQ3/GCE A Level or equivalent	4	4	3	3
NVQ2/GCE O Level or equivalent	9	9	9	8
Other	62	58	65	70
No qualification	5	5	5	4
Social Class of HRP				
Professional	6	7	5	2
Managerial technical	25	26	25	20
Skilled non-manual	11	12	11	10
Skilled manual	36	33	37	42
Semi-skilled manual	17	16	17	18
Unskilled manual	5	6	5	7
Other	1	1	1	0
Economic activity status for last week				
Working full time	36	35	39	28
Working part time	13	15	10	13
Unemployed	3	3	3	3
Permanently sick or disabled	3	3	3	4

Table continued

Retired	35	33	37	37
Keeping house	0	0	0	0
Full time student	10	11	7	15
Bedrooms in household				
One	7	8	7	8
Two	26	25	28	25
Three	51	51	50	56
Four or more	15	16	14	11
Housing tenure				
Own accommodation	65	67	65	57
Rent	35	33	35	43
No. longstanding illnesses				
One illness	29	29	28	31
Two illnesses	10	9	10	14
Three illnesses	3	2	3	4
Four or more illnesses	0	0	0	1
Limiting longstanding illness	23	21	22	29

Source: HALS1 1984/1985 and the longitudinal follow up in June 2009

Notes:

All statistics are either mean (for continuous variables) or % (categorical variables)

Appendix 5.2: Cox model controlling for individual and household characteristics

	Main effects model				BMI interacted with SES model				BMI interacted with age model				Fully interacted model			
	Men		Women		Men		Women		Men		Women		Men		Women	
	Haz. ratio	z	Haz. ratio	z	Haz. ratio	z	Haz. ratio	z	Haz. ratio	z	Haz. ratio	z	Haz. ratio	z	Haz. ratio	z
BMI																
Normal weight	Base category		Base category		Base category		Base category		Base category		Base category		Base category		Base category	
Overweight	1.043	0.61	0.974	-0.36	0.988	-0.1	1.054	0.5	0.647	-1.03	1.316	0.59	0.693	-0.71	2.677	1.58
Obese	1.337	2.63	1.365	3.42	1.225	1.03	1.452	2.9	1.350	0.43	4.514	2.73	1.511	0.48	18.961	3.96
SES																
1 (most deprived)	Base category		Base category		Base category		Base category		Base category		Base category		Base category		Base category	
2	0.823	-2.42	0.901	-1.25	0.684	-3.17	1.015	0.13	0.796	-2.76	0.938	-0.75	0.680	-3.19	1.113	0.87
3	0.688	-3.89	0.73	-2.96	0.688	-2.67	0.757	-1.92	0.687	-3.88	0.749	-2.65	0.675	-2.71	0.882	-0.81
4 (least deprived)	0.55	-5.22	0.543	-4.33	0.615	-3.15	0.650	-2.44	0.552	-5.13	0.563	-3.96	0.599	-3.15	0.783	-1.32
BMI*SES																
Overweight*SES 2					1.306	1.58	0.785	-1.38					1.328	1.66	0.700	-1.88
Overweight*SES 3					1.018	0.1	1.101	0.47					1.067	0.33	0.919	-0.36
Overweight*SES 4					0.812	-0.99	0.679	-1.31					0.866	-0.63	0.544	-1.86
Obese*SES 2					1.440	1.34	1.041	0.19					1.424	1.28	0.750	-1.25
Obese*SES 3					0.987	-0.05	0.768	-0.96					0.976	-0.08	0.456	-2.52
Obese*SES 4					0.867	-0.35	0.697	-0.9					0.830	-0.42	0.367	-2.28
Age																
Age	1.169	5.11	1.162	4.91	1.170	5.05	1.164	4.93	1.169	5.02	1.155	4.55	1.170	5.01	1.166	4.82
Age squared	1.000	-2.08	1.000	-1.85	1.000	-2.03	1.000	-1.85	0.999	-2.13	1.000	-1.38	0.999	-2.08	1.000	-1.51
BMI*age																
Overweight*age									1.007	1.15	0.995	-0.67	1.005	0.71	0.987	-1.53
Obese*age									1.000	-0.02	0.982	-2.17	0.997	-0.26	0.965	-3.47

Note

The dependent variable is survival time. Controlling for smoking, ethnicity, marital status, area and missing income.

Appendix 5.3: Cox model controlling for individual, household characteristics and illness variables

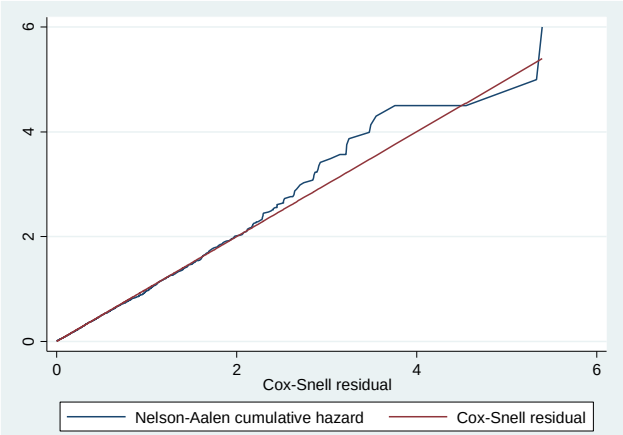
	Main effects model				BMI interacted with SES model				BMI interacted with age model				Fully interacted model			
	Men		Women		Men		Women		Men		Women		Men		Women	
	Haz. ratio	z	Haz. ratio	z	Haz. ratio	z	Haz. ratio	z	Haz. ratio	z	Haz. ratio	z	Haz. ratio	z	Haz. ratio	z
BMI																
Normal weight	Base category		Base category		Base category		Base category		Base category		Base category		Base category		Base category	
Overweight	1.045	0.64	0.948	-0.73	0.999	-0.01	1.000	0	0.698	-0.85	1.241	0.46	0.780	-0.48	2.183	1.25
Obese	1.302	2.38	1.254	2.47	1.229	1.04	1.306	2.06	1.448	0.52	4.736	2.79	1.713	0.62	19.133	3.94
SES																
1 (most deprived)	Base category		Base category		Base category		Base category		Base category		Base category		Base category		Base category	
2	0.873	-1.66	0.949	-0.63	0.729	-2.62	1.036	0.3	0.840	-2.09	0.984	-0.19	0.726	-2.63	1.130	0.99
3	0.754	-2.89	0.764	-2.54	0.770	-1.84	0.789	-1.63	0.750	-2.92	0.789	-2.18	0.762	-1.85	0.915	-0.58
4 (least deprived)	0.607	-4.29	0.578	-3.87	0.677	-2.5	0.674	-2.24	0.609	-4.21	0.602	-3.48	0.667	-2.46	0.804	-1.17
BMI*SES																
Overweight*SES 2					1.291	1.51	0.828	-1.07					1.306	1.55	0.750	-1.5
Overweight*SES 3					0.989	-0.06	1.149	0.67					1.021	0.1	0.979	-0.09
Overweight*SES 4					0.811	-0.99	0.707	-1.17					0.848	-0.72	0.584	-1.64
Obese*SES 2					1.391	1.2	1.095	0.43					1.367	1.13	0.773	-1.11
Obese*SES 3					0.907	-0.33	0.754	-1.03					0.883	-0.4	0.441	-2.63
Obese*SES 4					0.905	-0.24	0.785	-0.6					0.844	-0.38	0.400	-2.07
Age																
Age	1.156	4.75	1.152	4.63	1.16	4.7	1.153	4.64	1.158	4.71	1.152	4.46	1.159	4.69	1.162	4.68
Age squared	1.000	-1.67	1.000	-1.58	1.000	-1.64	1.000	-1.57	1.000	-1.75	1.000	-1.32	1.000	-1.7	1.000	-1.42
BMI*age																
Overweight*age									1.006	0.97	0.996	-0.6	1.004	0.5	0.989	-1.27
Obese*age									0.998	-0.15	0.980	-2.39	0.995	-0.4	0.963	-3.6

Note

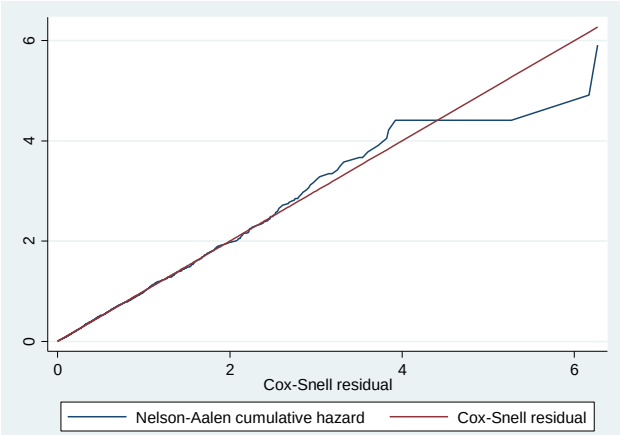
The dependent variable is survival time. Controlling for smoking, ethnicity, marital status, area, missing income, number of longstanding illnesses and whether or not the illness is limiting.

Appendix 5.4: Cox-Snell and AIC residuals for men

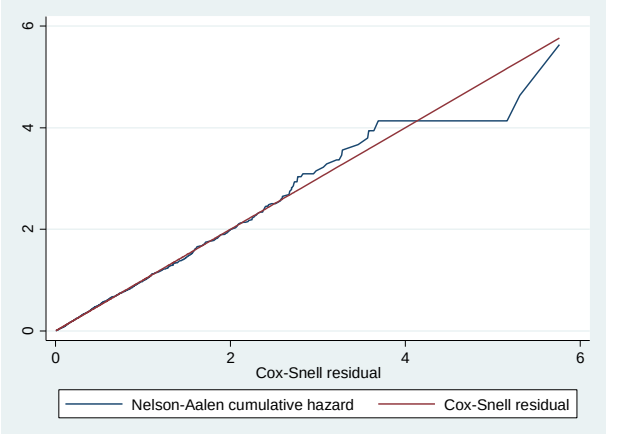
Weibull (AIC = 3433.77)



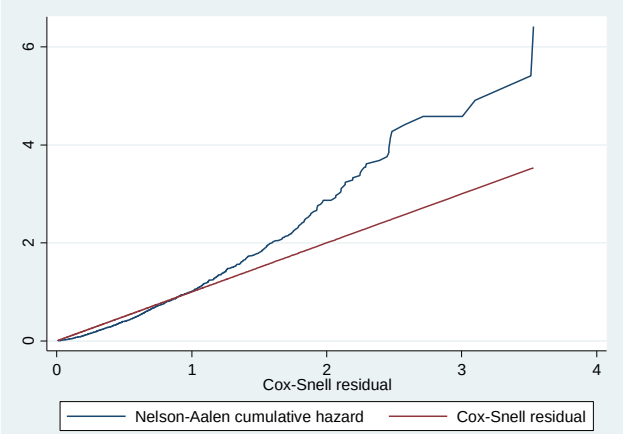
Generalised gamma (AIC= 3427.088)



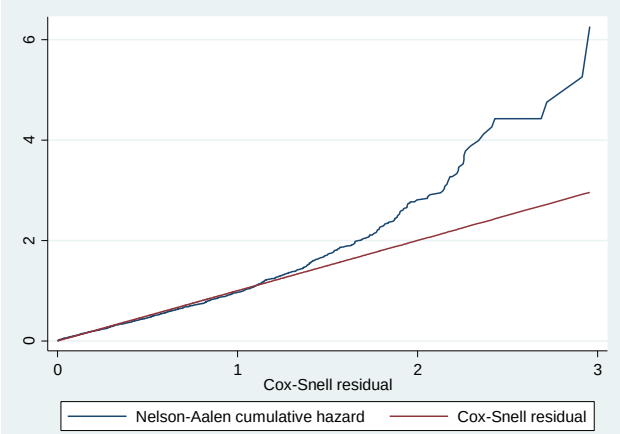
Gompertz (AIC= 3404.631)



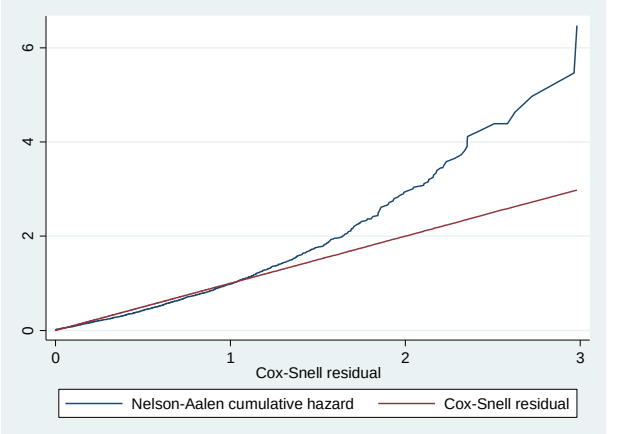
Exponential (AIC= 3709.575)



Loglogistic (AIC= 3549.274)

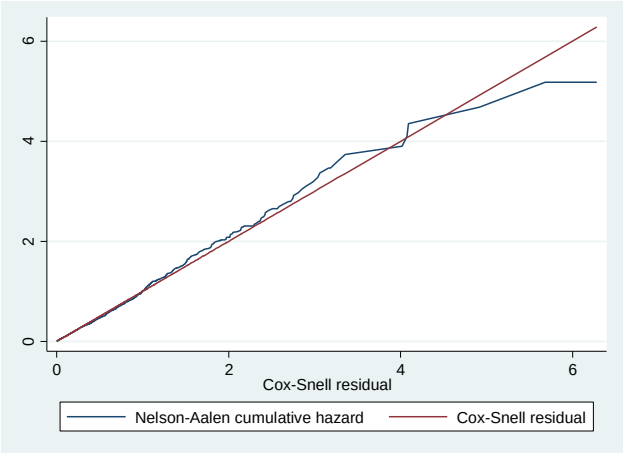


Lognormal (AIC= 3661.926)

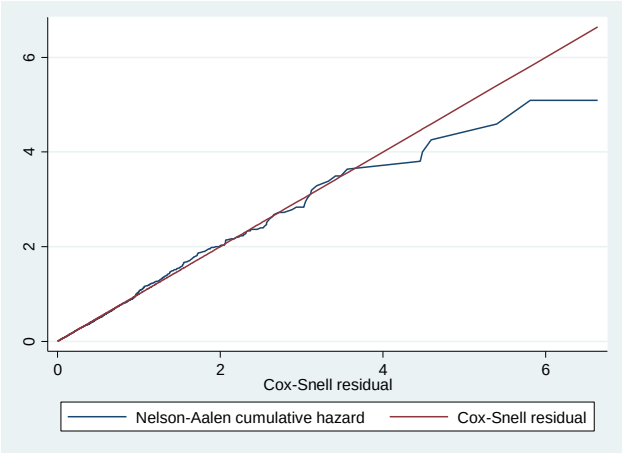


Appendix 5.5: Cox-Snell residuals and AIC for women

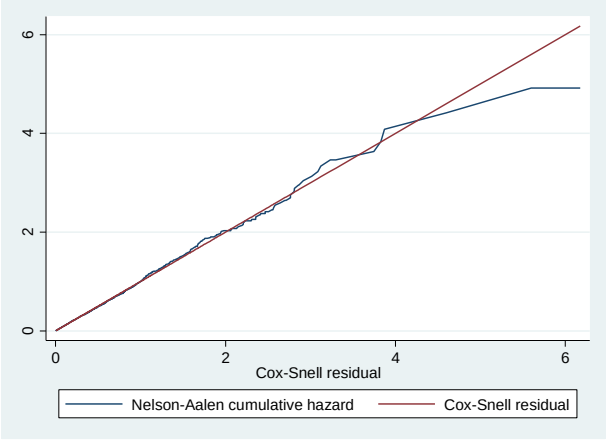
Weibull (AIC= 3671.249)



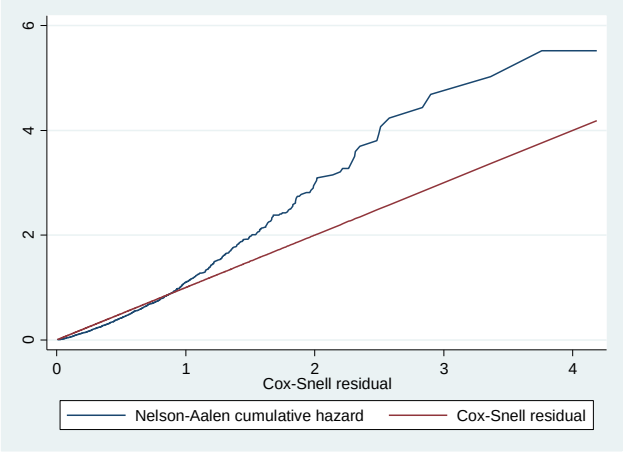
Generalised gamma (AIC= 3666.321)



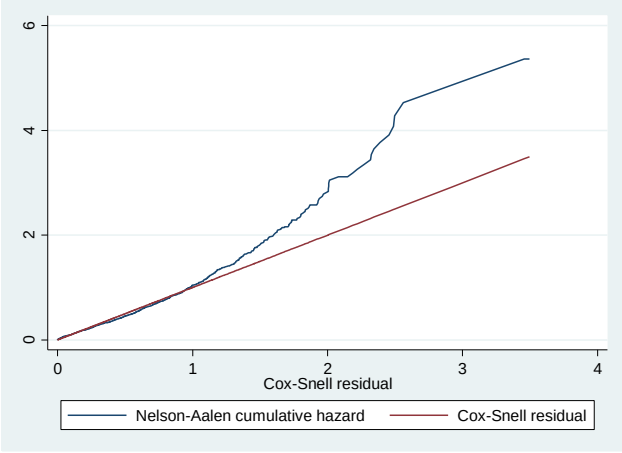
Gompertz (AIC= 3618.839)



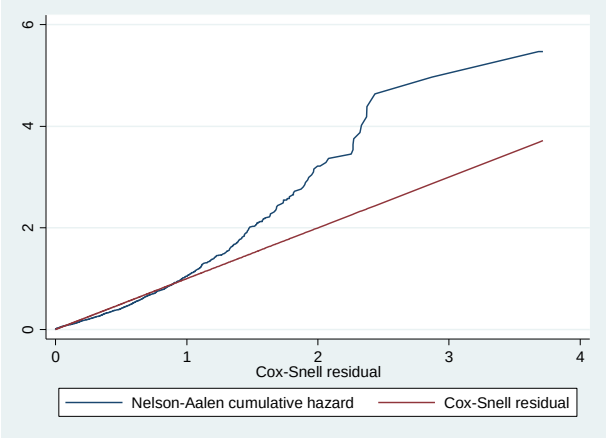
Exponential (AIC= 3932.148)



Loglogistic (AIC= 3765.378)



Lognormal (AIC= 3879.428)



Appendices Chapter 6

Appendix 6.1: Prescribed medication categories

1: Cardio-vascular medicine
2: Gastrointestinal medicine
3: Respiratory medicine
4: CNS medicine
5: Medicine for infection
6: Endocrine medicine
7: Gynae/Urinary medicine
8: Cytotoxic medicine
9: Medicine for nutrition/blood
10: Musculoskeletal medicine
11: Eye/Ear etc medicine
12: Medicine for skin
13: Other medicine

Appendix 6.2: Descriptive statistics by BMI category of the HSE variables for the years 1999-2002

	Whole sample	Normal weight	Overweight	Class I obesity	Class II/III obesity
Observations					
Number	36766	14649	13762	5489	2198
%	100	40	37	15	6
Predicted SES measure	25704	26071	26408	24715	23203
Missing income	18	19	18	17	14
SES quartile					
1 (most deprived)	25	24	23	28	33
2	25	25	25	25	23
3	25	25	25	24	24
4 (least deprived)	25	26	27	23	20
Gender					
Men	46	40	53	49	31
Women	54	60	47	51	69
Age (years)	45	40	49	50	47
Ethnicity					
White	93	92	94	94	94
Mixed	0	0	0	0	0
Caribbean	1	1	1	1	2
African	1	1	1	1	1
Indian	2	2	2	1	1
Pakistani	1	1	1	1	1
Bangladeshi	0	1	0	0	0
Chinese	0	1	0	0	0
Asian other	0	0	0	0	0
Other ethnicity	1	1	1	1	1
Marital status					
Married	52	43	60	62	57
Single	23	34	16	13	16
Separated	2	2	2	2	3
Divorced	5	5	5	6	7
Widowed	7	5	8	8	8
Cohabiting	10	11	9	9	9
Smoking status					
Never smoked	44	45	42	42	45
Ex occasional smoker	6	6	6	5	5
Ex regular smoker	24	17	28	30	28
Current cigarette smoker	27	31	24	22	21
GOR of residence					
North East	6	6	6	7	8
North West	14	14	15	14	14
Yorkshire and The Humber	11	11	11	12	11
East Midlands	10	9	10	11	12
West Midlands	10	10	10	11	12
East of England	12	12	12	11	12
London	11	12	10	10	10
South West	11	11	11	10	9
South East	15	16	15	14	13
Survey year					
1999	19	19	19	18	17
2000	20	19	20	20	19
2001	37	34	39	40	39
2002	25	28	22	22	25
Educational qualifications					
Degree or equivalent	15	17	15	12	9
Higher education below degree	11	10	12	10	11
NVQ3/GCE A Level or equiv.	14	16	12	11	11
NVQ2/GCE O Level or equiv.	24	26	22	23	24
NVQ1/CSE other grade or equiv.	6	5	6	6	6
Foreign/other	4	4	5	5	4
No qualification	27	21	28	34	35
Social Class of HRP					

Table continued

Professional	30	29	32	28	26
Managerial technical	24	27	23	22	24
Skilled non-manual	35	32	37	39	38
Skilled manual	5	5	5	7	8
Semi-skilled manual	0	0	0	0	0
Unskilled manual	5	7	3	4	4
Cars owned by household					
Household has no car	20	21	18	20	23
Household has car	80	79	82	80	77
Economic activity status last week					
In paid employment/self-employed	65	68	64	61	60
Going to school or college full time	2	4	2	1	1
On a government training scheme	0	0	0	0	0
Doing unpaid work	0	0	0	0	0
Waiting to take up paid work	0	0	0	0	0
Looking for paid work	1	2	1	2	2
Temporary sickness	0	0	0	0	0
Long-term sickness	4	3	4	5	9
Retired from paid work	20	15	24	25	21
Looking after home or family	5	6	4	5	6
Doing something else	0	0	0	0	0
Housing tenure					
Own outright	27	23	30	31	24
Buying with mortgage or loan	46	48	47	43	42
Pay part rent part mortgage	0	1	0	0	0
Rent	25	26	21	25	33
Live rent free and/or squatting	1	1	1	1	1
Bedrooms in household					
One	6	6	6	6	7
Two	21	21	21	23	24
Three	50	48	51	52	53
Four	17	18	17	14	12
Five or more	5	6	5	4	3
Income support claimant					
No	93	93	94	93	88
Yes	7	7	6	7	12

Notes

Other than for "Observations" all statistics are either mean (for continuous variables) or % (categorical variables). "<1" indicates non-zero values less than one; "0" indicates zero values. HRP = household reference person.

Appendix 6.3. Descriptive statistics by BMI category in HSE (1999-2008)

	Whole sample	Normal weight	Overweight	Class I obesity	Class II/III obesity
Observations					
Number	66435	24313	25489	11040	4626
%	100	37	38	17	7
Predicted SES measure	30796	31378	31543	29625	27748
Missing income	17	18	17	16	16
SES quartile					
1 (most deprived)	25	23	23	28	33
2	25	25	25	25	25
3	25	25	25	24	24
4 (least deprived)	25	27	26	23	19
Gender					
Men	46	39	53	49	33
Women	54	61	47	51	67
Age (years)	48	43	51	52	50
Ethnicity					
White	93	92	93	94	93
Mixed	1	1	1	0	1
Caribbean	1	1	1	1	2
African	1	1	1	1	1
Indian	2	2	2	1	1
Pakistani	1	1	1	1	1
Bangladeshi	0	1	0	0	0
Chinese	0	0	0	0	0
Asian other	0	1	0	0	0
Other ethnicity	1	1	1	1	1
Marital status					
Married	55	46	61	63	59
Single	19	28	13	11	14
Separated	2	2	2	2	2
Divorced	6	6	6	6	7
Widowed	7	6	8	8	8
Cohabiting	10	11	9	9	9
Smoking status					
Never smoked	45	47	44	42	45
Ex occasional smoker	6	6	6	5	6
Ex regular smoker	26	19	29	33	31
Current cigarette smoker	23	28	21	19	19
GOR of residence					
North East	6	6	6	7	7
North West	14	14	15	14	14
Yorkshire and The Humber	11	11	11	12	11
East Midlands	10	9	10	10	11
West Midlands	10	9	10	11	12
East of England	12	11	12	11	12
London	10	12	9	9	8
South West	11	11	11	10	10
South East	16	17	16	16	15
Survey year					
1999	1	1	1	1	1
2000	9	9	9	8	8
2001	17	17	18	17	15
2002	11	13	10	9	10
2003	16	15	16	16	15
2004	1	1	1	1	1
2005	10	9	10	11	11
2006	14	14	14	15	16
2007	7	7	7	7	7
2008	15	14	14	16	17

Table continued

Educational qualifications					
Degree or equivalent	17	20	18	14	11
Higher education below degree	11	11	13	11	11
NVQ3/GCE A Level or equiv.	13	16	13	11	10
NVQ2/GCE O Level or equiv.	24	25	22	23	24
NVQ1/CSE other grade or equiv.	5	5	5	5	6
Foreign/other	3	3	3	4	3
No qualification	26	21	27	32	35
Social Class of HRP					
Professional	33	33	36	32	28
Managerial technical	23	26	22	21	22
Skilled non-manual	34	30	35	38	38
Skilled manual	5	5	5	6	7
Semi-skilled manual	0	0	0	0	0
Unskilled manual	5	7	3	3	4
Cars owned by household					
Household has no car	17	18	16	17	20
Household has car	83	82	84	83	80
Economic activity status last week					
In paid employment/self-employed	64	68	63	59	59
Going to school or college full time	2	3	1	1	1
On a government training scheme	0	0	0	0	0
Doing unpaid work	0	0	0	0	0
Waiting to take up paid work	0	0	0	0	0
Looking for paid work	1	1	1	1	1
Temporary sickness	0	0	0	0	0
Long-term sickness	4	3	3	5	7
Retired from paid work	24	19	28	29	25
Looking after home or family	4	5	4	4	5
Doing something else	0	0	0	0	0
Housing tenure					
Own outright	32	28	35	36	28
Buying with mortgage or loan	44	46	44	41	40
Pay part rent part mortgage	0	0	0	0	0
Rent	23	25	20	22	31
Live rent free and/or squatting	1	1	1	1	1
Bedrooms in household					
One	6	6	6	6	8
Two	21	21	21	22	24
Three	49	47	50	52	52
Four	18	19	18	16	14
Five or more	5	6	5	4	3
Income support claimant					
No	94	94	95	94	90
Yes	6	6	5	6	10
Medication					
Cardio-vascular	22	13	24	33	38
Gastrointestinal	9	6	9	12	13
Respiratory	10	9	9	10	14
CNS	15	11	15	20	26
Infection	2	2	2	2	3
Endocrine	11	8	11	14	19
Musculoskeletal	7	4	6	10	13
Medication excluding contraceptives	46	37	48	56	63
Number of medications	1	1	1	2	2

Notes

Other than for "Observations" all statistics are either mean (for continuous variables) or % (categorical variables).

Appendix 6.4: descriptive statistics by BMI category in the BHPS (2004/5 and 2006/7)

	Whole sample	Normal weight	Overweight	Class I obesity	Class II/III obesity
Observations					
Number	20776	8515	7844	2900	1086
%	100	41	38	14	5
Predicted SES measure	28350	28642	28959	27491	25890
SES quartile					
1 (most deprived)	25	25	23	27	33
2	25	25	24	26	25
3	25	24	27	25	24
4 (least deprived)	25	26	26	23	18
Gender					
Men	47	46	53	45	35
Women	53	54	48	55	65
Age (years)	47	44	49	50	48
Ethnicity					
White	97	97	98	98	96
Black Caribbean	0	0	0	0	0
Black African	0	0	0	0	0
Black other	0	0	0	0	0
Indian	1	1	1	1	2
Pakistani	0	0	0	0	0
Bangladeshi	0	0	0	0	0
Chinese	0	0	0	0	0
Other ethnic group	1	1	1	0	2
Marital status					
Married	54	47	59	61	58
Couple	12	13	11	11	10
Widowed	7	7	8	8	7
Divorced	6	5	6	6	8
Separated	2	2	2	2	2
Never married	20	27	14	12	14
Civil partnership	0	0	0	0	0
Smoking status					
Yes	25	28	23	20	22
No	75	72	77	80	78
Wave (overall)					
2004/2005	62	70	58	55	53
2006/2007	38	30	42	45	47
Educational qualifications					
Degree or equivalent	41	44	41	40	34
NVQ3/GCE A Level or equivalent	12	14	11	9	12
NVQ2/GCE O Level or equivalent	18	18	18	16	17
NVQ1/CSE other grade or equivalent	8	7	8	9	9
No qualification	20	17	21	25	28
Still in school	1	1	0	0	0
Social Class of HRP					
Professional and managerial technical	17	17	18	17	14
Non manual labour	42	46	39	39	40
Manual skilled	23	21	24	25	24
Manual unskilled	5	5	4	5	6
Missing social class	2	2	2	2	4
Never had a job	11	9	12	11	12
Cars owned by household					
Household has no car	29	32	24	26	36
Household has car	71	68	76	74	64

Table continued

Economic activity status last week					
Employed	58	59	60	57	53
Unemployed	3	4	2	3	3
Retired	22	19	25	25	20
Maternity leave	1	0	1	1	0
Family care	6	6	5	6	10
FT studt, school	5	8	3	1	2
LT sick, disabld	5	3	4	7	11
Gvt trng scheme	0	0	0	0	0
Other economic activity					
Housing tenure					
Own outright	77	76	80	75	68
Shared	1	1	0	1	1
Rent	21	22	18	23	30
Live rent free and/or squatting	1	1	1	1	1
Other accommodation	0	0	0	0	0
Bedrooms in household					
One	1	1	0	0	0
Two	4	4	4	5	4
Three	15	15	15	16	19
Four	27	26	28	29	30
Five or more	53	54	54	50	46
Income support claimant					
No	95	95	95	93	90
Yes	5	5	5	7	10
Health service use variables					
GP visit	74	71	74	78	82
Outpatient visit	41	39	42	45	53
Inpatient stay	10	10	10	11	14
Chiropodist	10	8	10	14	17
X-ray of chest and lungs	14	12	15	15	18
Blood pressure	52	45	54	60	65
Cholesterol test	23	16	27	31	35
Blood test	41	34	43	48	55

Notes

Other than for “Observations” all statistics are either mean (for continuous variables) or % (categorical variables). “<1” indicates non-zero values less than one; “0” indicates zero values. HRP = household reference person.

Appendix 6.5: Output of probit models with interactions between BMI groups and SES in HSE (1999-2002)

	GP				Practice nurse				Inpatient				Outpatient			
	Men		Women		Men		Women		Men		Women		Men		Women	
	Coef.	z	Coef.	z	Coef.	z	Coef.	z	Coef.	z	Coef.	Z	Coef.	z	Coef.	z
BMI																
Normal weight	Base category		Base category		Base category		Base category		Base category		Base category		Base category		Base category	
Overweight	0.020	0.34	0.076	1.53	0.139	1.28	0.089	1.06	-0.034	-0.37	0.083	1	0.176	2.37	-0.079	-1.16
Obese class I	0.015	0.20	0.123	2.07	0.059	0.44	0.047	0.48	-0.155	-1.23	0.234	2.34	0.013	0.14	0.094	1.12
Obese class II/III	0.386	3.34	0.206	2.87	0.715	4.02	0.277	2.53	-0.030	-0.15	0.131	1.03	0.383	2.37	0.290	2.83
SES																
1 (most deprived)	Base category		Base category		Base category		Base category		Base category		Base category		Base category		Base category	
2	-0.148	-2.41	-0.111	-2.29	0.146	1.24	-0.082	-0.98	-0.218	-2.10	-0.247	-2.67	-0.007	-0.09	-0.058	-0.84
3	-0.216	-3.30	-0.167	-3.4	-0.057	-0.44	-0.138	-1.57	-0.282	-2.50	-0.022	-0.25	-0.059	-0.74	-0.011	-0.16
4 (least deprived)	-0.407	-5.91	-0.186	-3.59	-0.158	-1.18	-0.011	-0.13	-0.428	-3.49	-0.243	-2.46	-0.304	-3.52	-0.043	-0.59
BMI*SES																
Overweight*SES 2	-0.045	-0.55	-0.009	-0.13	-0.228	-1.52	-0.110	-0.94	0.008	0.06	0.047	0.37	-0.256	-2.48	0.181	1.87
Overweight*SES 3	-0.021	-0.24	-0.005	-0.07	-0.093	-0.59	0.021	0.18	-0.030	-0.20	-0.133	-1.06	-0.229	-2.17	0.072	0.74
Overweight*SES 4	0.111	1.30	-0.003	-0.05	0.005	0.03	-0.210	-1.66	0.126	0.83	0.136	1	-0.060	-0.55	0.198	1.92
Obese class I*SES 2	-0.013	-0.12	-0.133	-1.51	0.096	0.52	-0.027	-0.18	0.266	1.43	0.060	0.38	-0.062	-0.46	0.107	0.88
Obese class I*SES 3	0.000	0.00	0.033	0.37	0.056	0.29	0.103	0.68	0.225	1.17	-0.162	-1.02	0.068	0.49	0.173	1.37
Obese class I*SES 4	0.251	2.29	0.001	0.01	0.073	0.36	0.053	0.33	0.379	1.90	0.230	1.28	0.175	1.23	0.074	0.52
Obese class II/III*SES 2	-0.363	-2.04	0.058	0.54	-0.846	-2.70	-0.315	-1.80	0.092	0.29	0.146	0.71	-0.330	-1.37	0.144	0.91
Obese class II/III*SES 3	-0.269	-1.53	-0.040	-0.35	-0.175	-0.66	-0.095	-0.52	-0.016	-0.04	-0.483	-2.06	-0.304	-1.24	-0.198	-1.19
Obese class II/III*SES 4	0.006	0.03	-0.114	-0.9	-0.586	-1.81	-0.195	-0.93	0.398	1.20	0.143	0.62	-0.098	-0.39	-0.016	-0.09
Joint prob > chi2	*0.1701*		*0.6612*		*0.1163*		*0.5332*		*0.7646*		*0.2508*		*0.1557*		*0.2975*	

Note

Controlling for age, age squared and age cubed, ethnicity, marital status, smoking status, area variables, survey year, and missing income.

Appendix 6.6: Output of probit models after exclusion of insignificant variables between BMI groups and SES in HSE (1999-2002)

	GP				Practice nurse				Inpatient				Outpatient			
	Men		Women		Men		Women		Men		Women		Men		Women	
	Coef.	z	Coef.	z	Coef.	Z	Coef.	z	Coef.	z	Coef.	z	Coef.	z	Coef.	z
BMI																
Normal weight	Base category		Base category		Base category						Base category				Base category	
Overweight	-0.001	-0.02	0.071	2.71	0.059	1.04					0.085	1.77			0.027	0.75
Obese class I	0.012	0.25	0.095	2.81	0.118	1.66					0.245	4.02			0.176	3.68
Obese class II/III	0.200	2.73	0.195	4.63	0.388	3.74					0.101	1.25			0.286	4.7
SES																
1 (most deprived)	Base category		Base category		Base category		Base category		Base category		Base category		Base category		Base category	
2	-0.190	-5.06	-0.129	-4.22	0.014	0.22	-0.159	-3.19	-0.169	-2.67	-0.210	-3.65	-0.124	-2.61		
3	-0.243	-6.13	-0.164	-5.07	-0.099	-1.43	-0.152	-2.82	-0.266	-3.83	-0.108	-1.83	-0.158	-3.16		
4 (least deprived)	-0.426	-6.71	-0.201	-5.7	-0.171	-2.29	-0.119	-2.06	-0.291	-4.05	-0.161	-2.41	-0.288	-5.49		
BMI*SES																
Overweight*SES 2																
Overweight*SES 3																
Overweight*SES 4	0.130	1.78														
Obese class I*SES 2																
Obese class I*SES 3																
Obese class I*SES 4	0.252	2.7														
Obese class II/III*SES 2																
Obese class II/III*SES 3																
Obese class II/III*SES 4	0.190	1.21														

Note

Controlling for age, age squared and age cubed, ethnicity, marital status, smoking status, area variables, survey year, and missing income.

Appendix 6.7: Output of probit and negbin models with interactions between BMI groups and SES for individuals having had a nurse visit in HSE (1999-2008)

Part 1 of 2

	Cardio-vascular medicine				**Gastrointestinal medicine**				**Respiratory medicine**				**CNS medicine**			
	Men		Women		Men		Women		Men		Women		Men		Women	
	Coef.	z	Coef.	z	Coef.	z	Coef.	z	Coef.	z	Coef.	z	Coef.	z	Coef.	z
BMI																
Normal weight	Base category		Base category		Base category		Base category		Base category		Base category		Base category		Base category	
Overweight	0.364	6.72	0.235	5.13	0.005	0.09	0.028	0.55	0.008	0.15	0.015	0.30	0.101	2.07	0.070	1.70
Obese class I	0.665	10.50	0.412	7.85	0.072	1.09	0.173	3.07	0.147	2.21	0.091	1.60	0.239	4.16	0.236	4.99
Obese class II/III	1.167	12.50	0.682	11.36	0.285	2.79	0.274	4.26	0.369	4.12	0.285	4.60	0.512	6.54	0.403	7.43
SES																
1 (most deprived)	Base category		Base category		Base category		Base category		Base category		Base category		Base category		Base category	
2	-0.008	-0.12	-0.100	-1.95	-0.118	-1.81	-0.239	-4.40	-0.091	-1.53	-0.147	-2.95	-0.405	-6.71	-0.211	-4.90
3	-0.114	-1.68	-0.236	-4.60	-0.306	-4.32	-0.216	-3.88	-0.153	-2.54	-0.182	-3.72	-0.450	-7.16	-0.348	-7.97
4 (least deprived)	-0.198	-3.06	-0.379	-7.01	-0.350	-5.11	-0.321	-5.74	-0.142	-2.35	-0.225	-4.46	-0.633	-10.77	-0.508	-11.18
BMI*SES																
Overweight*SES 2	-0.140	-1.78	-0.073	-1.11	-0.020	-0.25	0.136	1.91	-0.004	-0.05	0.082	1.22	0.057	0.75	0.050	0.85
Overweight*SES 3	-0.137	-1.72	-0.038	-0.57	0.103	1.18	0.021	0.27	-0.006	-0.08	-0.011	-0.16	-0.063	-0.82	0.116	1.95
Overweight*SES 4	-0.095	-1.26	0.020	0.29	0.025	0.29	0.073	0.95	-0.046	-0.61	0.059	0.85	-0.094	-1.27	0.101	1.66
Obese class I*SES 2	-0.190	-2.06	0.091	1.19	-0.098	-1.01	0.201	2.46	-0.193	-2.05	0.058	0.72	0.013	0.15	0.006	0.09
Obese class I*SES 3	-0.102	-1.11	0.197	2.57	0.117	1.13	0.062	0.72	-0.153	-1.62	0.100	1.25	-0.115	-1.28	-0.036	-0.50
Obese class I*SES 4	-0.108	-1.24	0.114	1.39	0.103	1.04	0.128	1.40	-0.204	-2.13	0.077	0.89	-0.041	-0.47	0.149	1.97
Obese class II/III*SES 2	-0.318	-2.33	0.129	1.45	-0.096	-0.63	0.074	0.74	-0.185	-1.38	0.026	0.28	-0.045	-0.33	0.076	0.94
Obese class II/III*SES 3	-0.424	-3.19	0.170	1.88	-0.097	-0.66	0.057	0.54	-0.459	-3.23	-0.016	-0.17	-0.281	-2.25	0.062	0.73
Obese class II/III*SES 4	-0.237	-1.80	0.223	2.24	-0.174	-1.07	0.071	0.62	-0.360	-2.44	0.127	1.25	-0.062	-0.48	0.035	0.38
Joint prob > chi2	*0.1284*		*0.0248*		*0.4286*		*0.5488*		*0.0311*		*0.6895*		*0.359*		*0.1764*	

Part 2 of 2

	Endocrine medicine				Musculoskeletal medicine				Average number of prescribed medicines taken			
	Men		Women		Men		Women		Men		Women	
	Coef.	z	Coef.	z	Coef.	z	Coef.	z	Coef.	z	Coef.	z
BMI												
Normal weight	Base category		Base category		Base category		Base category		Base category		Base category	
Overweight	0.241	3.29	-0.018	-0.36	0.161	2.37	0.058	1.02	0.155	3.02	0.135	3.62
Obese class I	0.451	5.77	0.110	2.05	0.377	4.88	0.240	3.97	0.446	7.61	0.314	7.53
Obese class II/III	0.798	8.24	0.293	4.62	0.524	5.38	0.452	6.79	0.810	10.09	0.574	12.81
SES												
1 (most deprived)	Base category		Base category		Base category		Base category		Base category		Base category	
2	0.047	0.57	-0.084	-1.68	-0.091	-1.05	-0.164	-2.70	-0.303	-4.99	-0.250	-6.11
3	-0.069	-0.79	-0.017	-0.33	-0.145	-1.55	-0.133	-2.19	-0.433	-6.73	-0.363	-8.74
4 (least deprived)	-0.040	-0.49	0.016	0.33	-0.251	-2.87	-0.271	-4.34	-0.601	-10.12	-0.459	-10.35
BMI*SES												
Overweight*SES 2	-0.164	-1.57	0.162	2.45	-0.096	-0.94	0.165	2.07	0.032	0.43	0.052	0.95
Overweight*SES 3	-0.051	-0.48	0.025	0.38	-0.044	-0.40	0.019	0.23	-0.064	-0.85	-0.004	-0.07
Overweight*SES 4	-0.241	-2.37	0.111	1.74	-0.077	-0.76	0.084	0.97	-0.026	-0.37	0.072	1.22
Obese class I*SES 2	-0.011	-0.09	0.137	1.84	0.047	0.41	0.176	1.96	-0.029	-0.34	0.160	2.62
Obese class I*SES 3	-0.171	-1.46	-0.056	-0.73	-0.045	-0.38	0.114	1.26	-0.062	-0.71	0.112	1.77
Obese class I*SES 4	-0.158	-1.40	-0.018	-0.24	-0.091	-0.80	0.136	1.40	0.043	0.52	0.084	1.21
Obese class II/III*SES 2	-0.102	-0.69	0.059	0.65	-0.004	-0.03	0.175	1.76	-0.044	-0.36	0.157	2.23
Obese class II/III*SES 3	-0.116	-0.81	0.014	0.15	-0.099	-0.65	-0.032	-0.30	-0.125	-1.01	0.112	1.6
Obese class II/III*SES 4	-0.108	-0.73	-0.098	-1.02	-0.062	-0.40	-0.006	-0.05	-0.003	-0.02	0.199	2.31
Joint prob > chi2	*0.1232*		*0.056*		*0.8379*		*0.386*		*0.8435*		*0.0979*	

Note

Controlling for age, age squared and age cubed, ethnicity, marital status, smoking status, area variables, survey year, and missing income.

Appendix 6.8: Output of refitted probit and negbin models with interactions between BMI groups and SES for individuals having had a nurse visit in HSE (1999-2008)

Part 1 of 2

	Cardio-vascular medicine				Gastrointestinal medicine				Respiratory medicine				CNS medicine			
	Men		Women		Men		Women		Men		Women		Men		Women	
	Coef.	z	Coef.	z	Coef.	Z	Coef.	z	Coef.	z	Coef.	z	Coef.	z	Coef.	z
BMI																
Normal weight	Base category		Base category		Base category		Base category		Base category		Base category		Base category		Base category	
Overweight	0.162	7.69	0.235	5.13	0.006	0.24	0.060	2.93	0.008	0.15	0.042	2.07	0.045	1.98	0.094	3.7
Obese class I	0.403	15.32	0.412	7.85	0.069	2.38	0.195	7.56	0.147	2.21	0.119	4.53	0.153	5.46	0.210	6.71
Obese class II/III	0.689	16.82	0.682	11.36	0.133	2.82	0.249	7.81	0.369	4.12	0.266	8.52	0.331	7.71	0.405	10.79
SES																
1 (most deprived)	Base category		Base category		Base category		Base category		Base category		Base category		Base category		Base category	
2	-0.092	-3.45	-0.100	-1.95	-0.109	-3.88	-0.097	-4.03	-0.091	-1.53	-0.081	-3.39	-0.306	-11.47	-0.150	-6.83
3	-0.178	-6.6	-0.236	-4.6	-0.191	-6.57	-0.146	-5.75	-0.153	-2.54	-0.142	-5.64	-0.432	-15.66	-0.258	-8.02
4 (least deprived)	-0.196	-7.29	-0.379	-7.01	-0.231	-7.8	-0.183	-6.78	-0.142	-2.35	-0.147	-5.48	-0.532	-18.93	-0.367	-10.91
BMI*SES																
Overweight*SES 2			-0.073	-1.11					-0.004	-0.05						
Overweight*SES 3			-0.038	-0.57					-0.006	-0.08					0.040	0.93
Overweight*SES 4			0.020	0.29					-0.046	-0.61					-0.004	-0.09
Obese class I*SES 2			0.091	1.19					-0.193	-2.05						
Obese class I*SES 3			0.197	2.57					-0.153	-1.62					-0.080	-1.47
Obese class I*SES 4			0.114	1.39					-0.204	-2.13					0.112	1.9
Obese class II/III*SES 2			0.129	1.45					-0.185	-1.38						
Obese class II/III*SES 3			0.170	1.88					-0.459	-3.23					-0.024	-0.34
Obese class II/III*SES 4			0.223	2.24					-0.360	-2.44					-0.074	-0.99

Part 2 of 2

	Endocrine medicine				Musculoskeletal medicine				Average number of prescribed medicines taken			
	Men		Women		Men		Women		Men		Women	
	Coef.	z	Coef.	z	Coef.	z	Coef.	z	Coef.	z	Coef.	z
BMI												
Normal weight	Base category		Base category		Base category		Base category		Base category		Base category	
Overweight	0.051	2.04	0.044	2.38	0.048	1.97	0.074	3.47	0.141	5.15	0.164	8.03
Obese class I	0.228	7.54	0.108	4.6	0.235	7.88	0.231	8.61	0.436	13.33	0.402	17.23
Obese class II/III	0.486	10.58	0.255	8.56	0.322	7.06	0.370	11.34	0.769	16.66	0.677	24.25
SES												
1 (most deprived)	Base category		Base category		Base category		Base category		Base category		Base category	
2	-0.028	-0.94			-0.085	-2.86	-0.038	-1.49	-0.301	-9.91	-0.174	-7.94
3	-0.117	-3.85			-0.151	-4.9	-0.095	-3.63	-0.484	-15.17	-0.320	-13.47
4 (least deprived)	-0.124	-3.94			-0.204	-6.44	-0.144	-5.16	-0.603	-18.91	-0.391	-15.18
BMI*SES												
Overweight*SES 2												
Overweight*SES 3												
Overweight*SES 4												
Obese class I*SES 2												
Obese class I*SES 3												
Obese class I*SES 4												
Obese class II/III*SES 2												
Obese class II/III*SES 3												
Obese class II/III*SES 4												

Note

Controlling for age, age squared and age cubed, ethnicity, marital status, smoking status, area variables, survey year, and missing income.

Appendix 6.9: Output of probit models and interval regression models with interactions between BMI groups and SES for individuals in the BHPS

Part 1 of 2

	GP visits				Number of GP visits				Outpatient visits				Number of outpatient visits				Inpatient stay			
	Men		Women		Men		Women		Men		Women		Men		Women		Men		Women	
BMI																				
Normal weight	Base category		Base category		Base category		Base category		Base category		Base category		Base category		Base category		Base category		Base category	
Overweight	0.166	1.80	0.188	2.02	0.021	0.11	0.416	2.27	-0.022	-0.26	0.126	1.76	-0.226	-1.06	-0.352	-1.70	-0.149	-1.43	-0.056	-0.67
Obese class I	0.236	1.76	0.396	3.24	1.007	3.83	1.142	5.01	0.114	0.95	0.276	3.07	0.099	0.34	0.142	0.57	0.077	0.55	0.129	1.28
Obese class II/III	0.475	2.23	0.665	3.95	1.144	2.86	1.733	5.86	0.551	2.98	0.435	3.72	0.277	0.67	0.089	0.29	0.224	1.09	0.143	1.11
SES																				
1 (most deprived)	Base category		Base category		Base category		Base category		Base category		Base category		Base category		Base category		Base category		Base category	
2	-0.261	-3.14	0.038	0.45	-0.659	-3.55	-0.812	-4.65	-0.199	-2.47	-0.068	-1.00	-0.385	-1.76	-0.657	-3.16	-0.280	-2.65	-0.293	-3.53
3	-0.162	-1.89	-0.138	-1.59	-1.084	-5.75	-0.969	-5.30	-0.164	-1.99	-0.012	-0.18	-0.726	-3.29	-0.705	-3.34	-0.589	-4.95	-0.333	-3.84
4 (least deprived)	-0.177	-2.01	-0.051	-0.58	-1.297	-6.62	-0.906	-4.95	-0.236	-2.76	-0.015	-0.21	-0.398	-1.71	-0.482	-2.28	-0.514	-4.27	-0.218	-2.58
BMI*SES																				
Overweight*SES 2	0.132	1.09	-0.074	-0.59	0.057	0.22	0.082	0.33	0.228	2.01	-0.043	-0.43	0.098	0.33	0.720	2.45	0.300	2.03	0.209	1.74
Overweight*SES 3	-0.074	-0.62	-0.080	-0.64	0.147	0.58	-0.002	-0.01	0.000	0.00	-0.120	-1.20	0.608	2.02	0.184	0.62	0.408	2.58	0.258	2.11
Overweight*SES 4	-0.160	-1.34	0.011	0.08	0.346	1.35	-0.064	-0.25	0.093	0.82	-0.080	-0.79	0.250	0.83	0.172	0.58	0.325	2.09	0.116	0.94
Obese class I*SES 2	-0.005	-0.03	-0.266	-1.58	-0.753	-2.11	0.139	0.43	0.020	0.13	-0.190	-1.48	0.343	0.83	0.762	2.05	0.010	0.05	-0.026	-0.17
Obese class I*SES 3	-0.052	-0.30	0.019	0.11	-0.666	-1.84	-0.805	-2.46	-0.089	-0.55	-0.021	-0.16	0.606	1.43	0.116	0.32	0.069	0.31	0.102	0.66
Obese class I*SES 4	-0.042	-0.24	-0.132	-0.73	-0.513	-1.41	-0.819	-2.39	0.049	0.30	-0.141	-1.04	-0.368	-0.88	0.061	0.16	0.143	0.67	-0.050	-0.30
Obese class II/III*SES 2	0.015	0.05	0.011	0.04	0.766	1.33	-0.071	-0.17	-0.253	-0.98	0.084	0.49	0.461	0.73	0.489	1.06	0.199	0.65	0.010	0.05
Obese class II/III*SES 3	-0.376	-1.32	-0.245	-1.00	-0.703	-1.19	-0.313	-0.68	-0.533	-2.02	-0.003	-0.02	-0.092	-0.14	0.907	1.89	0.410	1.27	0.441	2.18
Obese class II/III*SES 4	0.151	0.50	-0.036	-0.13	0.078	0.14	-1.000	-1.98	-0.121	-0.46	0.011	0.05	0.293	0.47	-0.592	-1.09	0.086	0.26	-0.123	-0.49
Joint prob > chi2	*0.20*		*0.70*		*0.07*		*0.04*		*0.27*		*0.77*		*0.22*		*0.06*		*0.37*		*0.25*	
*Likelihood-ratio test**	*0.19*		*0.70*		*-*		*-*		*0.26*		*0.77*		*-*		*-*		*0.35*		*0.25*	

* Likelihood-ratio test of restricted model with no interactions vs. less restrictive model with interactions. The models for "number of GP visits" and "number of outpatient visits" are linear models hence we do not run likelihood ratio tests for these models.

Part 2 of 2

	Chiropodist				X-ray of chest and lungs				Blood pressure				Cholesterol test				Blood test			
	Men		Women		Men		Women		Men		Women		Men		Women		Men		Women	
BMI																				
Normal weight	Base category		Base category		Base category		Base category		Base category		Base category		Base category		Base category		Base category		Base category	
Overweight	0.327	1.65	0.171	1	-0.075	-0.81	-0.077	-0.87	0.114	1.21	0.198	2.38	0.256	2.14	0.291	2.82	0.279	2.84	0.085	1.09
Obese class I	0.832	3.19	0.527	2.59	0.034	0.27	-0.041	-0.37	0.333	2.48	0.407	3.81	0.563	3.51	0.376	3.03	0.4	2.87	0.256	2.6
Obese class II/III	1.496	4.19	0.724	2.73	0.116	0.6	0.217	1.58	0.565	2.79	0.605	4.29	0.74	3.13	0.917	5.91	0.644	3.05	0.626	4.79
SES																				
1 (most deprived)	Base category		Base category		Base category		Base category		Base category		Base category		Base category		Base category		Base category		Base category	
2	0.017	0.08	0.039	0.22	-0.252	-2.75	-0.232	-2.6	-0.239	-2.65	-0.033	-0.43	-0.15	-1.19	0.041	0.39	-0.098	-1.03	-0.269	-3.58
3	-0.325	-1.39	0.079	0.44	-0.375	-3.92	-0.467	-4.83	-0.203	-2.22	0.058	0.72	-0.119	-0.96	0.015	0.14	-0.143	-1.46	-0.221	-2.86
4 (least deprived)	-0.234	-0.95	-0.155	-0.83	-0.454	-4.51	-0.251	-2.72	-0.194	-2.06	0.112	1.38	-0.16	-1.25	0.022	0.2	-0.161	-1.59	-0.203	-2.63
BMI*SES																				
Overweight*SES 2	-0.111	-0.4	-0.001	0	0.244	1.91	0.16	1.26	0.288	2.28	0.003	0.03	0.111	0.68	-0.017	-0.12	0.018	0.14	0.285	2.63
Overweight*SES 3	0.11	0.37	0.097	0.41	0.186	1.44	0.334	2.5	0.176	1.42	-0.048	-0.41	0.12	0.75	-0.038	-0.26	-0.151	-1.15	0.12	1.09
Overweight*SES 4	-0.008	-0.03	0.152	0.61	0.231	1.76	0.12	0.91	0.183	1.47	0.027	0.23	0.092	0.57	-0.22	-1.5	-0.058	-0.44	0.122	1.1
Obese class I*SES 2	0.003	0.01	-0.09	-0.31	-0.023	-0.13	0.146	0.9	0.15	0.85	0.102	0.68	0.104	0.48	0.122	0.7	-0.039	-0.21	0.339	2.42
Obese class I*SES 3	0.138	0.36	-0.028	-0.1	-0.022	-0.12	0.316	1.85	0.129	0.72	-0.064	-0.42	-0.104	-0.47	-0.116	-0.65	-0.109	-0.58	0.114	0.8
Obese class I*SES 4	0.241	0.63	0.055	0.17	0.059	0.32	-0.115	-0.64	0.151	0.84	-0.078	-0.49	-0.073	-0.33	0.125	0.68	-0.028	-0.15	0.184	1.25
Obese class II/III*SES 2	-0.314	-0.59	0.264	0.7	0.028	0.1	0.329	1.6	0.25	0.9	0.201	0.97	0.337	1.02	-0.34	-1.49	-0.029	-0.1	0.324	1.68
Obese class II/III*SES 3	-0.321	-0.55	0.071	0.18	0.156	0.54	0.047	0.21	0.031	0.11	-0.058	-0.27	0.013	0.04	-0.891	-3.6	-0.229	-0.75	-0.216	-1.09
Obese class II/III*SES 4	0.239	0.44	-0.138	-0.3	-0.178	-0.56	0.013	0.05	0.452	1.56	-0.196	-0.81	0.541	1.62	-0.477	-1.76	0.391	1.3	-0.213	-0.95
Joint prob > chi2	*0.98*		*0.99*		*0.65*		*0.17*		*0.53*		*0.86*		*0.72*		*0.01*		*0.67*		*0.06*	
*Likelihood-ratio test**	*0.98*		*0.99*		*0.64*		*0.16*		*0.52*		*0.86*		*0.72*		*0.01*		*0.66*		*0.06*	

Note

Controlling for age, age squared and age cubed, ethnicity, marital status, smoking status, area variables, survey year, and missing income.

* Likelihood-ratio test of restricted model with no interactions vs. less restrictive model with interactions

Appendix 6.10: output of refitted probit models and interval regression models in the BHPS

Part 1 of 2

	GP visits				Number of GP visits				Outpatient visits				Number of outpatient visits				Inpatient stay			
	Men		Women		Men		Women		Men		Women		Men		Women		Men		Women	
BMI																				
Normal weight	Base category		Base category		Base category		Base category		Base category		Base category				Base category		Base category		Base category	
Overweight	0.132	3.08	0.15	3.17	0.166	1.74	0.416	2.27	0.059	1.43	0.068	1.82			-0.089	-0.81	0.094	1.65	0.081	1.77
Obese class I	0.205	3.31	0.304	4.63	0.517	3.82	1.142	5.01	0.106	1.79	0.197	3.96			0.368	2.59	0.122	1.52	0.141	2.35
Obese class II/III	0.405	3.86	0.596	5.96	1.179	5.41	1.733	5.86	0.326	3.37	0.46	6.56			0.34	1.83	0.378	3.12	0.239	2.95
SES																				
1 (most deprived)	Base category		Base category		Base category		Base category		Base category				Base category		Base category		Base category		Base category	
2	-0.207	-3.61	-0.028	-0.48	-0.714	-5.92	-0.812	-4.65	-0.116	-2.17			-0.267	-1.94	-0.232	-1.79	-0.156	-2.29	-0.221	-4.12
3	-0.215	-3.65	-0.181	-3.03	-1.139	-9.12	-0.969	-5.3	-0.205	-3.7			-0.394	-2.76	-0.536	-4.05	-0.39	-5.21	-0.191	-3.45
4 (least deprived)	-0.252	-4.13	-0.072	-1.15	-1.212	-9.33	-0.906	-4.95	-0.201	-3.53			-0.356	-2.39	-0.414	-2.99	-0.358	-4.69	-0.184	-3.17
BMI*SES																				
Overweight*SES 2							0.082	0.33												
Overweight*SES 3							-0.002	-0.01												
Overweight*SES 4							-0.064	-0.25												
Obese class I*SES 2							0.139	0.43												
Obese class I*SES 3							-0.805	-2.46												
Obese class I*SES 4							-0.819	-2.39												
Obese class II/III*SES 2							-0.071	-0.17												
Obese class II/III*SES 3							-0.313	-0.68												
Obese class II/III*SES 4							-1	-1.98												

Part 2 of 2

	Chiropodist				X-ray of chest and lungs				Blood pressure				Cholesterol test				Blood test			
	Men		Women		Men		Women		Men		Women		Men		Women		Men		Women	
BMI																				
Normal weight	Base category		Base category				Base category		Base category		Base category		Base category		Base category		Base category		Base category	
Overweight	0.335	2.94	0.229	2.42			0.063	1.28	0.278	6.14	0.189	4.35	0.338	5.69	0.291	2.82	0.227	4.68	0.217	5.23
Obese class I	0.981	6.07	0.515	4.26			0.045	0.7	0.443	6.86	0.392	6.64	0.548	6.73	0.376	3.03	0.353	5.13	0.409	7.4
Obese class II/III	1.431	6.24	0.802	4.87			0.319	3.72	0.751	7.07	0.599	7.04	0.967	7.51	0.917	5.91	0.677	6.03	0.634	8.04
SES																				
1 (most deprived)					Base category		Base category								Base category		Base category		Base category	
2					-0.162	-2.73	-0.123	-2.18							0.041	0.39	-0.094	-1.51	-0.094	-1.93
3					-0.300	-4.82	-0.286	-4.73							0.015	0.14	-0.232	-3.58	-0.171	-3.35
4 (least deprived)					-0.360	-5.54	-0.213	-3.39							0.022	0.2	-0.171	-2.55	-0.139	-2.6
BMI*SES																				
Overweight*SES 2															-0.017	-0.12				
Overweight*SES 3															-0.038	-0.26				
Overweight*SES 4															-0.22	-1.5				
Obese class I*SES 2															0.122	0.7				
Obese class I*SES 3															-0.116	-0.65				
Obese class I*SES 4															0.125	0.68				
Obese class II/III*SES 2															-0.34	-1.49				
Obese class II/III*SES 3															-0.891	-3.6				
Obese class II/III*SES 4															-0.477	-1.76				

Note

Controlling for age, age squared and age cubed, ethnicity, marital status, smoking status, area variables, survey year, and missing income.