Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

ECONOMIA

Ciclo XXVII

Settore Concorsuale di afferenza: 13/A4 Economia Applicata

Settore Scientifico disciplinare: SECS-P/06 Economia Applicata

TITOLO TESI

ESSAYS IN HEALTH ECONOMICS

Presentata da: **Paolo Nicola Barbieri**

Coordinatore Dottorato **Relatore**

Prof. Matteo Cervellati **Prof. Davide Dragone**

Esame finale anno 2014/15

Alma Mater Studiorum - University of Bologna

Doctoral Thesis

Essays on Health Economics

Author:
Paolo Nicola Barbieri

Supervisor:
Davide Dragone

March 2015

Acknowledgements

First of all I would like to express my special appreciation and thanks to my advisor Professor Davide Dragone, you have been a unique mentor for me. I would like to thank you for encouraging my research and for allowing me to grow not just as a research scientist but also as a man and a future husband. Your advice on both research as well as on my career have been priceless. I would especially like to thank all professors, among which especially Francesca Barigozzi, Giulio Zanella, Daniele Fabbri, Chiara Monfardini, and PhD colleagues from the University of Bologna which have contributed immensely to my personal and professional time at the University of Bologna. All of you have been there to support me in all my restless search and development of research ideas for my PhD thesis. I would like to thank Franco Sassi, Marion Devaux, Annalisa Belloni and Jean-Benoit Eymeoud from the OECD, where they made my time not only full of exiting work and research but also enjoyable and life changing. I will forever be grateful for such opportunity. I also thank my friends (too many to list here but you know who you are!) for providing support and friendship that I needed, whether with words, hugs or with music. My life would never be so full and passionate if I haven't you by my side. Lastly, I would like to thank my family for all their love and encouragement. For my parents who raised me with a love of science and supported me in all my pursuits. For the presence of my sisters Vittoria and Paola wherever I was. And most of all for my loving, supportive, encouraging, and patient fiancé Anna for her faithful support during not just the Ph.D. but also for all the future challenges that we will face. No words can express how deeply and unconditionally my love is, you will be forever my true an only love, thank you.

Contents

Acknowledgements **i**

Contents **ii**

List of Figures **iv**

List of Tables **v**

Abstract **vi**

1 Obesity Traps and Incentives for weight loss **1**

Abstract **1**

1.1 Introduction . . . 2
1.2 The Model . . . 4
1.2.1 The Utility and Incentive Function . . . 4
1.2.2 The Dynamics of Individuals Body Weight . . . 5
1.3 Optimality conditions and stability . . . 6
1.3.1 Obesity Traps, $\alpha = 0$ . . . 7
1.3.2 Incentive for weight loss, $\alpha > 0$ . . . 11
1.4 Conclusion . . . 14

2 Social Distortion in Weight Perception: A Decomposition of the Obesity Epidemic **16**

Abstract **16**

2.1 Introduction . . . 17
2.2 Theoretical Model . . . 19
2.3 Data and Variables . . . 22
2.3.1 Data Source . . . 22
2.3.2 Obesity and Covariates . . . 22
2.3.2.1 Dependent Variable: Obesity Measure . . . 22
2.3.2.2 Covariates . . . 22
2.3.3 Social Distortion Proxies . . . 23
2.4 Baseline Analysis . . . 24
2.4.1 Univariate Logit . . . 24
2.4.2 Results . . . 25

2.5 Obesity increases over time: Within-Group or Across-Group Changes? . . . 26
2.5.1 Fairlie's Decomposition . . . 26
2.6 Stratification Analysis . . . 28
2.6.1 Gender . . . 28
2.6.2 SES . . . 29
2.7 Conclusion . . . 30

3 Body Weight and Employment: Is there an Overweight Premium? 32

Abstract 32

3.1 Introduction . . . 33
3.2 Methodology . . . 35
3.2.1 Baseline Estimation . . . 35
3.2.2 Local estimation and Generalized Additive Models (GAMs) . . . 36
3.3 Data and Variables . . . 39
3.4 Empirical Results . . . 40
3.4.1 Bodyweight and Employment: Logit and Semi-parametric Results . . . 41
3.4.2 Occupational Category . . . 42
3.5 Fat Free Mass and Body Fat . . . 44
3.6 Conclusion and Comments . . . 45

A Obesity Traps and Incentives for Weight Loss 47
A.1 The intertemporal problem (1.3) - (1.6) . . . 47
A.1.1 The Model with stochastic time of death . . . 47
A.2 Concavity of the Hamiltonian function . . . 48
A.3 The Derivation of Condition (1.14) . . . 49
A.4 Comparative Statics . . . 50
A.4.1 The example with explicit functions . . . 50

Bibliography 52

List of Figures

A.1 The law of motion of body weight for $\bar{w} = 0,82$. 60
A.2 Phase diagram of (15) for the parameters values: $a = 3.9$, $\beta = 0.05$, $\bar{w} = 0.82$, $w^H = 0.44$ and $\rho = 0.8$ 61
A.3 Phase diagram of (16) after the introduction of the incentive for the parameters values: $a = 3.9$, $\beta = 0.05$, $\bar{w} = 0.82$, $w^H = 0.44$, $\alpha = 0.4$ and $\rho = 0.8$ 62
A.4 Comparative Statics, pre- and post-incentive. 63
A.5 Area Incidence of Obesity in England 2002-2006 64
A.6 BMI and Employment, by gender - Semi-Parametric 65
A.7 BMI and Employment, by type of labour - Semi-Parametric 66
A.8 Men, Fat Free Mass, Body Fat and Employment - Semi-Parametric 67
A.9 Women, Fat Free Mass, Body Fat and Employment - Semi-Parametric 68

List of Tables

A.1 Summary Statistics . . . 69
A.2 Logit Results (2002-2006) . . . 70
A.3 Fairlie Decomposition 2002-2006 . . . 71
A.4 Fairlie Decomposition 2002-2006, by Gender . . . 72
A.5 Fairlie Decomposition 2002-2006, by Skill level (Male Only) . . . 73
A.6 Fairlie Decomposition 2002-2006, by Skill level (Male Only), Alternative social norm definition . . . 74
A.7 Fairlie Decomposition 2002-2006, by Skill level (Female Only) . . . 75
A.8 Fairlie Decomposition 2002-2006, by Skill level (Female Only), Alternative social norm definition . . . 76
A.9 Summary Statistics . . . 77
A.10 Probit (ME) and Semi-Parametric Results for Employment Probability . . . 78
A.11 Probit (ME) and Semi-Parametric Results for Employment Probability . . . 79
A.12 Probit (ME) and Semi-Parametric Results for Employment Probability . . . 80

Abstract

Essays on Health Economics

by Paolo Nicola Barbieri

Over the last decades increased obesity rates posed some major health concerns. Given the rapidity of such weight gain current literature started to investigate possible explanations for such pattern and for why individuals struggle to lose weight. This dissertation is intended to contribute in such effort by examining the causes for such weight gain and the consequences that they had on individual's health and economic outcomes.

In the first chapter we develop a theoretical model investigating food consumption and body weight with a novel assumption regarding human caloric expenditure (i.e. metabolism), in order to investigate why individuals can be rationally trapped in an excessive weight equilibrium and why they struggle to lose weight even when offered incentives for weight-loss. This novel metabolic assumption draws from recent medical evidence suggesting that the two main human tissues, namely fat mass and fat free mass, contribute to individual caloric expenditure in different ways. Fat free mass, for example, was identified as the main driver of human caloric expenditure, in fact to an increase in fat free mass is associated a subsequent increase in caloric expenditure; fat mass, conversely, contributes negatively to caloric expenditure, meaning that to an increase in fat mass is associated a decrease in individual's caloric expenditure. Based on this evidence we argue that since overweight individual have higher fat content with respect to leaner individuals they have also lower caloric expenditure and we develop an equation for the evolution of body weight and caloric expenditure capturing such dynamics. This assumption allows the theoretical model to have multiple equilibria and to provide an explanation for why losing weight is so difficult even in the presence of incentives, without relying on rational addiction, time-inconsistency preferences or bounded rationality. This multiplicity of steady states makes weight loss attempts difficult due to the presence of a so-called "obesity trap" threshold dividing two possible outcomes. If the individual, by means of dieting, is not able to surpass this unstable threshold any dieting effort will be unsuccessful and he will slowly converge back to his original body weight. In addition to this result we are able to characterize under which circumstances a temporary incentive can create a persistent weight loss. Given the unstable threshold a lump-sum incentive scheme is not always able to produce permanent weight loss, while a progressive incentive scheme, rewarding at a higher rate weight loss at higher level of body weight, is able to sustain persistent weight losses.

In the second chapter, in order to shed some light on the possible causes for the recent obesity epidemic, we investigate the possible contributions that social norms and peer effects had on the spread of obesity. In recent literature peer effects and social norms have been characterized as important pathways for the biological and behavioral spread of body weight, along with decreased food prices and physical

activity. We add to this literature by proposing a novel concept of social norm related to what we define as *social distortion in weight perception.* This distortion is develop through shared experience in a social environment with a high prevalence of obesity. The underlying assumption is that when the vast majority of peers are obese and such status is commonly observed, imitative obesity could represent an important contributor to weight-related norms. Obesity, in fact, spread through social ties since having obese social contacts might change the person's tolerance for being obese. Our novel social channel captures the idea that if overweight/obesity is wide spread, in the social environment where the individual interacts, this makes it less likely to be recognized as a problem mitigating health concerns. The theoretical model shows that, in equilibrium, the effect of an increase in peers' weight on i's weight is unrelated to health concerns while it is mainly associated with social concerns. Using regional data from England we prove that such social component is significant in influencing individual weight. In order to do so we exploit a markedly difference in obesity rates between 2002 and 2006 to undertake a Fairlie decomposition analysis. Of the 5% point difference in obesity rates between this two periods our decomposition analysis, when we include social norms, explains up to 80% with respect to less than 50% when we exclude such social components. These results suggest that individual body weight concerns are closely related to the actual realization of weight in his reference network.

Finally, in the last chapter we investigate the relationship between body weight and employment probability. Differently from previous literature, which found an "obesity penalty", we try to shed some new evidence on the possibility of a non-linear relationship between weight status and employment probability due to the recent increase in average weight. We argue, in fact, that an increased in the average weight had a subsequent effect on the employment prospects that overweight/obese individuals have on the job market. Given that heavier individuals are now more common they might have higher employment prospects since their type is more accepted than decades ago and fitness-based discrimination might be lower. Using a semi-parametric regression we are able to investigate the shape of the relationship between body weight and employment probabilities and if indeed heavier individual have a higher probability of employment. Our results show that men and women employment probability do not follow a linear relationship with body mass index (BMI) but rather an inverted U-shaped one, peaking at a BMI way over the clinical threshold for overweight. Instead of an "obesity penalty" we find evidence of an "overweight premium", in socially active labors, consistent with the notion developed in the second chapter, of an endogenous social norms governing body weight which was recently updated upwardly.

The findings presented in this dissertation provide some new evidence on the role that individual specific and environmental factors played in the recent increased in the average weight of individuals and on what consequences such weight gain had on individuals' employment possibilities, hence on their future socio-economic condition.

Chapter 1

Obesity Traps and Incentives for weight loss

Abstract

Given the increasing incidence of chronic diseases related to excessive weight and obesity, the promotion of efficient ways to invert this trend are major health concerns. Incentives for weight loss resulted supportive towards such goal, although few of the existing attempts were long-term successful. This chapter presents a theoretical investigation of why losing weight is so difficult even in the presence of such short-run incentives and what features an incentive scheme should have in order to sustain permanent weight-loss. The theoretical model delivers multiple steady states and a threshold characterizing a condition of "obesity traps", that the individual has to surpass in order to successfully lose weight. Given such threshold a lump-sum incentive scheme is not always able to produce a permanent weight loss. On the contrary a non-decreasing incentive scheme rewarding at a higher rate weight loss at higher level of body weight (i.e. progressive), is able to sustain steady state decreases in body weight and food consumption.

1.1 Introduction

Over the last decade the incidence of obesity increased dramatically (Ogden and Carroll, 2010, von Ruesten et al., 2011), creating increased concerns toward the promotion of efficient ways to invert this trend, by means of dieting (Baradel et al., 2009, Bish et al., 2012) or physical activity (Mozaffarian et al., 2011). However even if weight loss attempts increased, more than 50% of obese individuals still struggle to lose weight permanently (Kruger et al., 2004). Due to this unsuccessful attempts and in order to facilitate weight-loss efforts, recent insight from behavioral economics (Loewenstein et al., 2007) were used to promote excessive weight reduction by motivating individuals to follow a behavior that they were not following naturally by means of monetary and non-monetary incentives[1] (Augurzky et al., 2012, Cawley and Price, 2012, John et al., 2011, Volpp et al., 2008). Unfortunately, even if short-run results were promising and weight loss was indeed achieved during the experiment, substantial weight regain was found after the incentive was removed (John et al., 2011, Volpp et al., 2008, Cawley and Price, 2012), proving that even if sustained by incentives individuals were still failing in losing weight permanently[2]. This poses the question of understanding why weight loss is so difficult to achieve, what factors influence it and how they can even neglect the effects of short run incentives. Current literature relates weight-loss difficulties to several possible explanation such as: time-inconsistent preferences (Dodd, 2008, Ikeda et al., 2010); willpower depletion (Ozdenoren et al., 2006); the interaction between two conflicting system driving human's behavior (cognitive vs. affective) (Loewenstein and O'Donoghue, 2004, Ruhm, 2012); or more generally to the failure of individuals to rationally balance current benefits and future costs related to food consumption and body weight (O'Donoghue and Rabin, 1999).

In this chapter we add to existing literature by proposing another, complementary, explanation for why dieting efforts are so difficult, even in the absence of rational addiction, time-inconsistent preferences or bounded rationality. To do so we develop a novel assumption regarding individual's metabolism. Current economic literature consider individual metabolism to be linearly related to body weight, such that the more an individual weights the higher will be her caloric expenditure (Dragone, 2009, Levy, 2002, Harris and Benedict, 1918). The key assumption is that caloric expenditure to keep the basic functions of organism running while the individual is at rest, or in medical terms basal metabolic rate, is an increasing function of body weight, meaning that of two individual of similar age, the one with higher body weight (i.e. the one with the greater weight stock $w(t)$) burns a higher fraction of it. This assumption entails that the marginal effect of an additional kg on individual's caloric expenditure is always positive and therefore that caloric expenditure for obese/overweight individuals is higher than for leaner ones, simply because they weight more.

[1]Experimental attempts used either lump-sum payments or lotteries, in order to exploit individual fallacies in the estimation of the probability of an unlikely event, to motivate subjects to lose weight (Kahneman and Tversky, 1979).

[2]In most of the cases weight regain was so high that there was no longer a significant difference in weight loss between the treated and the control group.

This notion of body weight accumulation and expenditure contradicts recent insight from medical literature, from which metabolism actively and non-monotonically reacts to changes in body weight[3] (Catenacci et al., 2011, Ebbeling et al., 2012, Gale et al., 2004, Katan and Ludwig, 2010, Leibel and Hirsch, 1984, Leibel et al., 1995), such that obese/overweight individuals do not burn more calories simply because they weight more (Johnstone et al., 2005, Cunningham, 1991). Even for the same level of physical activity and food intake, the higher adipose content in the body composition of obese/overweight individuals contributes in reducing their caloric expenditure rather than increasing it, since body fat negatively affects individual's caloric expenditure. Mifflin et al. (1990), Cunningham (1991) report that the main predictor for total energy expenditure is the level of non-adipose mass that the individual possess and that fat-mass is not one of its main predictor. Moreover Johnstone et al. (2005) quantify that fat mass explains less than 6% of basal metabolic rate variability, while fat free-mass is able to explains almost 70%. Based on this evidence we can conclude that given the small incidence that fat mass has on individual's caloric expenditure and the fact that obese/overweight individuals have positive imbalance of fat mass over fat free mass (Bazzocchi et al., 2012), the relationship between body weight and caloric expenditure is non-linear. This means that of two individuals of the same age the heavier one (i.e. the one with a higher adipose content) is predicted to burn fewer calories in the next period, with respect to a leaner one. In order to capture this dynamics we will assume that exists an individual specific level of body weight after which any additional weight gain will decrease caloric expenditure, conversely any weight gain before this threshold will increases caloric expenditure.

The results from the theoretical model augmented with this novel metabolic assumption are two-folded. First of all we extend the rational eating literature (Dragone, 2009, Levy, 2002) to show the impact that metabolism can have on the steady state decision of a rational individual. Differently from current literature, in which the individual exhibit only a rational and stable outcome of overweight, we allow for multiple equilibria and report one characterized by a healthier body weight. This multiplicity of steady states provides the evidence of an "obesity trap" with an associated threshold dividing the two possible outcomes. Intuitively any weight-loss efforts that the individual undertake has to surpass such threshold in order to be able to creates a permanent weight loss, otherwise the individual will gradually regain weight and converges to her previous body weight. This threshold gives the rationale for any unsuccessful dieting efforts even in the absence of rational addiction, time-inconsistent preferences or bounded rationality i.e. the individual has to surpass body resistant to weight-loss. Therefore the presence of such threshold has an influence on the possible persistence of dieting effort and incentives for weight loss. In fact from this multiplicity we derive the second set of results of the chapter related to what kind of short-run incentives are able to promote a persistence weight loss and which are not. A non-decreasing incentive scheme rewarding more weight loss at higher level of body weight (i.e. progressive) is able to sustain steady state decreases in body weight and food consumption while a

[3] In fact, metabolic activity slows down when there is weight loss and/or speeds up when there is weight gain (Ebbeling et al., 2012, Gale et al., 2004, Katan and Ludwig, 2010, Leibel and Hirsch, 1984, Leibel et al., 1995).

lump-sum payment for weight loss, the most used in experimental literature, is not. This result allows us to give a theoretical explanation for the long-run weight regain reported in experimental evidence.

1.2 The Model

1.2.1 The Utility and Incentive Function

Consider an individual whose intertemporal utility depends on food consumption, $c(t) \geq 0$ and body weight, $w(t) > 0$, according to the following utility function

$$V(c(t), w(t)) = U(c(t), w(t)) + \alpha I(w^H, w(t)) \tag{1.1}$$

Where $U(.)$ represents the instantaneous utility from food consumption and the disutility from body weight, while $I(.)$ is utility related to incentives for weight loss. We will assume that the instantaneous utility $U(c, w)$ is twice continuously-differentiable, jointly concave with negative second order derivatives, U_{cc} and U_{ww} and separable, $U_{cw} = 0$.

Following Dragone and Savorelli (2011) we assume that there exists a satiation point for individual food consumption, denoted by c^{sat}, with an associated level of body weight w^{sat}, such that $\partial U(c^{sat})/\partial c = 0$, which is to be interpreted as the (steady state) body weight that an agent would reach if she always ate to satiation[4]. We say the agent is on a diet if $U_c > 0$ (she would increase utility by increasing food consumption) and is binging if $U_c < 0$ (she would increase utility by decreasing food consumption). Regarding body weight we assume that there exists BMI that maximizes each agent's health condition; denote the corresponding body weight as w^H, such that $\partial U(w^H)/\partial w = 0$. We say the agent is overweight ($w > w^H$) if $U_w < 0$ (she would increase her utility by decreasing body weight) and conversely underweight ($w < w^H$) if $U_w > 0$ (she would increase utility by increasing body weight). It is possible to distinguish between three possible cases. If $w^{sat} > w^H$, satiation induces being overweight; likewise we say that satiation induces to be underweight if $w^H > w^{sat}$. Lastly if $w^{sat} = w^H$ satiation corresponds to healthy weight.

In addition to the instantaneous utility $U(.)$ we also assume that there exists an incentive function for weight loss, which represents any type of incentive, internal and external, that the individual exhort in order to sustain weight loss. Such function, denoted by $\alpha I(w^H, w(t))$, can be either positive or equal to zero, depending whether the individual is motivated or not to lose weight. $I(.)$ is a non-decreasing incentive function which achieves a (unique) minimum when $w = w^H$, where w^H is the target body weight and $\alpha \in [0, 1]$ is a parameter related to the individual's responsiveness to incentive. The intuition behind the incentive function is that the more the individual is distant from his target

[4] c^{sat} corresponds to the solution of a standard constrained optimization problem where the agent must choose between two goods, food and non-food.

body weight the higher will be the incentive to sustain dieting effort, which will gradually decrease as the individual converges towards w^H. The $I(.)$ function captures the idea that heavier individuals should be rewarded more than leaner individuals for weight-loss of equal magnitude, whether by some intrinsic or extrinsic incentive scheme. In order to focus on healthy weight-loss incentives and to rule out any type of punishment or weight-loss below individual's healthy weight, w^H, we need to impose the following set of restrictions on $I(.)$:

Assumption 1.1. The incentive function has two properties:

1. For $w < w^H$ $I(.) = 0$,
2. For $w > w^H$ $I(.) > 0$

The first assumption excludes any incentive (or punishment) for weight lower than w^H; the second one follows from the fact that the incentive function is an increasing function in body weight. If the first case does not hold we would model cases in which individuals are motivated in losing weight over what is their healthier body weight level. Such situations can be recollected to eating disorder such as bulimia and anorexia which are not explored in this chapter.

1.2.2 The Dynamics of Individuals Body Weight

The determinants of individual's body weight are given by food consumption, $c(t)$, and current body weight, $w(t)$ which will influence the accumulation of body weight in time as follows

$$\dot{w}(t) = c(t) - g(w(t)) \tag{1.2}$$

Assumption 1.2. The $g(.)$ function is twice continuously-differentiable, positive and concave with a maximum in $g(\bar{w})$

$g(w(t))$ is a concave function determining the amount of energy expended per unit or body weight, which in medical terms is defined as basal metabolic rate (Harris and Benedict, 1918), which is the amount of energy expended by individuals at rest. Following the existing rational eating literature (Dragone, 2009, Levy, 2002) we assume that the contribution of food consumption to body weight is positive and linear. Additionally, the contribution of body weight to the rate of calories expenditure is assumed to be positive (i.e. $g_w \geq 0$) for $w \in [0, \bar{w}]$ and decreasing afterwards, with $g_w < 0$. The idea is the following: the more an individual eats the more he will gain weight, the more he will gain weight the more calories he will burn until he reaches a certain body weight, $\bar{w}$, where the effect of additional body weight is actually decreasing the rate at which he consumes calories. This explanation is related to the fact that the main predictor of basal metabolic rate is individual's lean mass (Johnstone et al., 2005, Mifflin et al., 1990), which is supposed to be lower (in proportion) in heavier individual. Therefore $\bar{w}$ can be thought as the maximum sustainable level of body weight after which the organism is storing additional calories in fat tissue and not in lean mass, resulting then in a net decreases in caloric expenditure since fat mass burn at a lower rate with respect to lean mass.

We will assume that as long as any additional weight gain is resulting in an increase in caloric expenditure, such that $g_w \geq 0$, the individual can be considered has having a positive imbalance between fat free mass and fat mass. Conversely as soon as she will surpass $\bar{w}$, resulting in a decrease in caloric expenditure associated with an increase in body weight, $g_w < 0$, the individual can be considered most certainly begin overweight or obese, since she has percentage of fat mass strictly higher than fat free mass.

1.3 Optimality conditions and stability

The individual's goal is to maximize his intertemporal utility by choosing the amount of food consumption, which in turn will affect her body weight. Given an infinite time horizon and a positive discount rate, ρ, the individual's problem can be written as follows (See Appendix A.1):

$$\max_{c(t)} \int_0^\infty e^{-\rho t}[U(c(t), w(t))] + \alpha I(w^H, w(t))dt \tag{1.3}$$

$$\text{subject to } \dot{w}(t) = c(t) - g(w(t)) \tag{1.4}$$

$$w(t), c(t) \geq 0 \tag{1.5}$$

$$w(0) = w_0, \text{ given} \tag{1.6}$$

Where we introduced the non-negativity constraint on c and w since they represents respectively food consumption and body weight which by definition cannot become negative. The associated current value Hamiltonian (dropping the time indexes for convenience) is

$$H(c, w, \lambda) = U(c, w) + \alpha I(w^H, w) + \lambda(c - g(w)) \tag{1.7}$$

1.3.1 Obesity Traps, $\alpha = 0$

We start by presenting the results excluding the presence of any incentives for weight-loss. The aim of this subsection is provide evidence on the existence of a multiplicity of steady states and of threshold creating the "obesity trap", due to the novel assumption that we included on individual's metabolism.

Given joint concavity (See Appendix), the first-order conditions for this problem are

$$H_c = 0 \qquad U_c + \lambda = 0 \tag{1.8a}$$
$$\dot{\lambda} = \lambda(\rho + g_w) - U_w \tag{1.8b}$$
$$\dot{w} = c - g(w) \tag{1.8c}$$

If we time-differentiate equation (1.8a) and substitute it in (1.8b) along with (1.8a) we derive an expression for $\dot{c}$, which, along with (1.8c) results in the following system of differential equations characterizing the equilibrium

$$\begin{aligned} \dot{c} &= \frac{U_c}{U_{cc}}(\rho + g_w) + \frac{U_w}{U_{cc}} \\ \dot{w} &= c - g(w) \end{aligned} \tag{1.9}$$

For an internal steady state the following conditions must be satisfied

$$\dot{c} = 0 \qquad U_w = -U_c(\rho + g_w) \tag{1.10}$$
$$\dot{w} = 0 \qquad c^{ss} = g(w^{ss}) \tag{1.11}$$

Plugging equation (11) into (10) we get

$$U_c(g(w), w)\,(g_w + \rho) + U_w(g(w), w) = 0 \tag{1.12}$$

In order to establish the multiplicity of steady states we need to prove that equation (1.12) has more than one possible solution. The number of steady states will be given by the number of solutions that equation (1.12) admits. However given the implicit form of equation (1.12) it is not straightforward to explicitly derive the number of solutions. Thus we go beyond and differentiating it with respect to w in order to get

$$g_{ww}U_c(g(w)) + g_w\,(g_w + \rho)\,U_{cc}(g(w)) + U_{ww} \lesseqgtr 0 \tag{1.13}$$

Due to the fact that the marginal utility from food consumption (U_c) and the marginal contribution of body weight to calorie consumption (g_w) are, respectively, changing sign two times, this creates a sufficient condition for the non-monotonicity of equation (13) resulting in the possibility of multiple steady states.

Proposition 1.3. *The intertemporal problem (2)-(5) is associated with multiplicity of steady states characterized by the sufficient condition for an unstable threshold (see Wirl and Feichtinger (2005))*

$$-\,g_w(\rho + g_w) > \frac{H_{ww}}{U_{cc}} \tag{1.14}$$

Associated with the following trade-offs between body weight and food consumption according to equation (10)

1. *Being overweight and on a diet,* $U_w < 0$ *and* $U_c > 0$

2. *Being underweight and on binging,* $U_w > 0$ *and* $U_c < 0$

(See Appendix)

Interestingly the only case in which eating up to satiation is a steady state outcome is associated to the case in which $c^{sat} = c^H$, with c^H being the amount of food consumption associated with w^H, because $U_c = -U_w$, thus there is, although under specific circumstances, a possible steady state in which the individual is perfectly healthy and consuming up to satiation, with no loss of utility whatsoever. More generally the steady state will not be associated with satiation, but either with binging or dieting, since $U_c \neq -U_w$ but either $U_c > 0$ and $U_w < 0$ (binging and being underweight) or $U_c < 0$ and $U_w > 0$ (dieting and being overweight).

From condition (1.14) we can also derive an additional (sufficient) condition for multiple steady states. Additionally to the unstable threshold in (1.14), which makes the determinant positive, we can notice that equation the equation

$$-g_w(\rho + g_w) - \frac{H_{ww}}{U_{cc}}$$

which is the determinant of the Jacobian associated with the intertemporal problem (1.3)-(1.6) (See Appendix) can be negative for two different cases, thus allowing two possible stable steady states. The firs one is when $g_w > 0$, while the second one is for $g_w < 0$. This characterization of steady states based on the sign of the derivative of the $g(.)$ function allow us to say that our model entails two stable steady states, one healthier than the other, since $g_w > 0$ is associated, by definition to a condition of healthy weight while $g_w < 0$ with overweight/obesity. More generally we will assume that there will be *at most* three steady states, divided by the threshold defined in equation (1.14).

Corollary 1.4 (Obesity Traps)**.** *Among the two possible steady states associated with the intertemporal problem (2)-(5) one is characterized by an healthy body weight (i.e.* $g_w > 0$*) while the other one is characterized by a condition of overweight/obesity (i.e.* $g_w < 0$*).*

Proposition 4.1 allows us to state that if there are indeed multiple steady states, and thus condition (1.14) is satisfied, instability becomes an issue and there exists a threshold such that the individual is trapped in an excessive weight steady state. This result characterizes a situation in which there is body resistance to weight-loss that the individual has to surpass in order to permanently lose weight, which can be reconciled with the concept of "homeostasis" reported in medical evidence (Ebbeling et al., 2012, Gale et al., 2004, Katan and Ludwig, 2010, Leibel and Hirsch, 1984, Leibel et al., 1995). Such that the organism contrast dieting effort since losing weight results in a decrease in energy available and in order to keep the organism in a stable energy state metabolism contrast weight-losses.

Since a more formal proof of multiple steady states is quite cumbersome we will resort to explicit functions to show more explicitly such multiplicity. In order to do so we will follow Wirl and Feichtinger

(2005) and Wirl (2004) in order to prove that our result holds under the most common and general functional specification, so to resort to explicit formulation for computational purpose without losing too much generality. For what concerns the utility function we will assume a quadratic formulation:

$$U(c,w) = c\left(1 - a\frac{c}{2}\right) - \beta\frac{(w - w^H)^2}{2}$$

where $c^{sat} = a$.

As for the law of motion of body weight we assume that individual body weight is composed by two main components: (1) fat mass (FM) and (2) fat-free mass (FFM). Each one is supposed to represent a fraction μ and $1-\mu$ with $\mu \in [0,1]$, respectively, of total body weight w, such that

$$\begin{aligned} FM &= (1-\mu)w \\ FFM &= \mu w \\ w &= FFM + FM = \mu w + (1-\mu)w = w \end{aligned}$$

As medical literature suggests caloric expenditure for each of this component is different and assumed to be given by the following functions

$$\begin{aligned} CE_{FFM} &= \mu w \\ CE_{FM} &= -(1-\mu)\frac{w^2}{2} \end{aligned}$$

Such that: (1) fat-free mass has a linear, positive and monotonically increasing contribution to caloric expenditure and (2) fat-mass has a quadratic and negative effect on caloric expenditure. The intuition is the following: initially there is a positive imbalance of fat-free mass over fat mass when, after $\bar{w}$ the (negative) contribution of fat-mass tops the positive of the fat-free-mass and caloric expenditure starts to decrease (See Figure A.1). Overall caloric expenditure is deviated by combining CE_{FFM} with CE_{FM}, such that

$$CE_{TOT} = \frac{1}{2}w(2\mu + (\mu - 1)w)$$

Which, given its quadratic characteristic, can be approximated with a general logistic function

$$CE_{TOT} = \frac{1}{2}w(2\mu + (\mu - 1)w) \approx w\left(\bar{w} - \frac{w}{2}\right) = g(w)$$

where $\bar{w}$ is assumed to represent the body weight level after which there is a positive imbalance of fat-mass over fat-free-mass or more intuitively the starting point after which the individual is becoming overweight and experience a reduction in caloric expenditure due to this.

[Figure 1 about here.]

To summarize utility and body weight law of motion are the following

$$\begin{aligned} U(c,w) &= c\left(1-\alpha\frac{c}{2}\right) - \beta\frac{(w-w^H)^2}{2} \\ \dot{w} &= c - g(w) = c - w(\bar{w} - \frac{w}{2}) \end{aligned} \tag{1.15}$$

From which is clear that we deliberately choose simple and standard specification for both equations in order to stress out that the results are present even when using very simple and compact functional form. In Appendix A.4.1 is provided a full solution for this specific example. Fig.2 summaries the number and the dynamic orientation of the three steady states of the system in equation (15), which is characterized by two saddle point points, w_1^{ss} and w_3^{ss}, divided by an unstable node, w_2^{ss}. In the figure it is also represented the saddle path leading to the two steady states and passing through the unstable node.

[Figure 2 about here.]

The multiplicity of steady states gives us a rationale for why, in general, dieting efforts are so difficult and why weight-loss is not always guaranteed after a diet is finished. The reason behind this is the presence of a weight threshold after which any weight-loss will be permanent. This means that if by dieting the individual is not able to surpass such threshold she will start to gain weight once the diet is finishes and she will slowly converge back to her previous weight, nullifying all her dieting effort.

1.3.2 Incentive for weight loss, $\alpha > 0$

Including the incentive for weight-loss the system of differential equations characterizing the equilibrium is similar to the one we derived in Section 1.3.1

$$\begin{aligned} \dot{c} &= \frac{U_c(\rho + g_w) + U_w + \alpha I_w}{U_{cc}} \\ \dot{w} &= c - g(w) \end{aligned} \tag{1.16}$$

With the only difference being the presence of the incentive function in the $\dot{c}$ equation. For an internal steady state the following conditions must be satisfied

$$\dot{c} = 0 \qquad U_w + \alpha I_w = -U_c(\rho + g_w) \tag{1.17}$$

$$\dot{w} = 0 \qquad c = g(w) \tag{1.18}$$

In order to investigate what consequences had the introduction of the incentive on the steady states value of body weight and food consumption we will follow Dragone et al. (2013) from which we compute the change in the steady state body weight as a response to a change in the sensitivity to the incentive is given by the following expression (See Appendix A.4):

$$\frac{\partial w^{ss}}{\partial \alpha} = -\frac{|P|}{|J|} = -\frac{I_w/U_{cc}}{|J|} \tag{1.19}$$

Where $|J|$ is the determinant of the Jacobian associated with the system of differential equations (1.16) and $|P|$ is the determinant of the following matrix

$$P = \begin{bmatrix} \frac{\partial \dot{c}}{\partial \alpha} & \frac{\partial \dot{c}}{\partial c} \\ \\ \frac{\partial \dot{w}}{\partial \alpha} & \frac{\partial \dot{w}}{\partial c} \end{bmatrix}$$

The sign of equation (1.19) leads to the following proposition.

Proposition 1.5. *A sufficient condition for an incentive for weight loss to sustain a reachable steady state with a lower body weight is* $I_w > 0$*.*

Proof. In a steady state the determinant of the Jacobian is negative and thus the denominator of equation (1.19) is negative, so for $I_w > 0$ and $U_{cc} < 0$ the presence of the incentive ($\alpha > 0$) is associated with stable decreases in body weight. □

Corollary 1.6. *Incentives scheme unconditional of body weight (i.e.* $I_w = 0$*) are not able to lead to stable changes in body weight*

Proof. For $I_w = 0$ the effect of incentives on the equilibrium level of body weight is nil, since equation (1.19) is equal to zero. □

As far as it concerns steady states food consumption we can derive a condition similar to (1.19) (See Appendix A.4)

$$\frac{\partial c^{ss}}{\partial \alpha} = -\frac{|K|}{|J|} = -\frac{g_w I_w/U_{cc}}{|J|} \tag{1.20}$$

$|K|$ is the determinant of

$$K = \begin{bmatrix} \frac{\partial \dot{c}}{\partial \alpha} & \frac{\partial \dot{c}}{\partial w} \\ \\ \frac{\partial \dot{w}}{\partial \alpha} & \frac{\partial \dot{w}}{\partial w} \end{bmatrix}$$

Such that

Proposition 1.7. *Steady state decreases in food consumption are achieved by* $I_w > 0$*, for overweight/obese individuals.*

Proof. By definition excessive body weight, or a positive imbalance between fat mass and fat free mass, is characterized by the condition $g_w < 0$, which will make equation (20) negative resulting in a decreases in steady states food consumption. $\square$

The Jacobian matrix associated with the system (1.19)

$$J = \begin{pmatrix} \rho + g_w & \dfrac{U_{ww} + g_{ww}U_c + \alpha V_{ww}}{U_{cc}} \\ \\ 1 & -g_w \end{pmatrix} \tag{1.21}$$

The trace is still positive and equals to $tr(J) = \rho > 0$, so the steady state can be at most a saddle point[5]. However the determinant now is

$$\det(J) = -g_w\left(\rho + g_w\right) - \frac{(U_{ww} + g_{ww}U_c + \alpha I_{ww})}{U_{cc}} \tag{1.22}$$

If we take a closer look at the determinant in (1.22) we arrive to the following proposition

Proposition 1.8. *"Progressivity" in the incentive scheme, $I_{ww} > 0$, or high level of sensitivity to incentives scheme are necessary conditions in order to sustain persistent weight loss.*

Proof. Rewriting the determinant equation (1.22) in terms of the instability condition (14) we have that

$$-g_w(\rho + g_w) > \frac{H_{ww}}{U_{cc}} + \alpha\frac{I_{ww}}{U_{cc}} \tag{1.23}$$

High level of either I_{ww} or α significantly reduce the conditions under which the inequality for instability is verified, since the second term in equation (23) is negative due to $U_{cc} < 0$. $\square$

Proposition 4.4. proves that incentive schemes associating higher rewards for weight loss at higher body weight level are able to promote persistent weight loss. Intuitively weight loss is possible whenever there is an incentive scheme which is more rewarding for weight loss at higher level of body weight with respect to weight loss achieved when body weight is relatively low ($I_{ww} > 0$), or when the incentive is so salient that is able to be an active motivation for the dieter (high α). This progressivity characteristic that incentives scheme should have in order to create permanent results might arise the controversy that agents should stay fat in order to gain higher rewards. However one additional assumption has to be considered: *the incentives scheme should be designed to reward weight loss at higher level of body weight, not higher body weight per se'*[6].

The results on the canonical system (16) of an introduction a quadratic incentive function, $\alpha(w - w^H)^2$ are summarized in Fig.3, from which the downward shift of the $\dot{c} = 0$ isocline reduces the number

[5] At least one eigenvalues must be positive

[6] This is usually done in experimental literature by providing the reward if and only if the individual were able to decreases his body weight.

of steady states from three to a single one. Pre-incentives individuals which were converging to the excessive weight steady state are now attracted by the only steady state reachable, which is represented by a lower body weight and hence an improved health condition. However the convergence to the lower steady state it is not trivial and it can be either just temporary or permanent depending if the individual is able, once the incentive is removed, to have reached the stable manifold converging to the lower study state, thus surpassing the intermediate unstable steady state.

[Figure 3 about here.]

[Figure 4 about here.]

Following Fig. 4 we can characterize the following outcomes

1) **Temporary Result:** The convergence towards the lowest steady state is such that the individual is not able to surpass the unstable steady state. Once the incentive is removed the individual will re-converge to the initial steady state.

2) **Permanent Result:** The convergence towards the lowest steady state is such that the individual is able to surpass the unstable steady state. Once the incentive is removed the individual will keep in converging to the healthier steady state.

In the first case the individual is not able to escape the “obesity trap” and thus he will not permanently lose weight. In the second case, due to the presence of the incentive, the individual is able to escape from the “obesity trap” and he can permanently converge to a lower level of body weight. The results from the model shows how the design of the incentives is crucial in order to produce stable results; however it has to be pointed out that since the intervention it is introduced temporarily its design it is only a *necessary* condition for the convergence to an improved weight condition.

1.4 Conclusion

Given the increasing incidence of obesity, the promotions of efficient ways to invert this trend are major health concerns. A reduction in caloric intake or an increase in caloric expenditure, via physical activity, is the two main prescribed strategies normally utilized. Even though there is an increase weight loss attempt using both this recommended strategies, their beneficial effects are late to be seen. Due to these unsuccessful attempts and in order to facilitate weight-loss efforts, recent insight from behavioral economics, were used to promote excessive weight reduction by motivating individuals to follow a behavior that they were not following naturally by means of monetary and non-monetary incentives. Although the results seem to be supportive of the ability of the scheme to promote weight loss during the time of the experimentation, few of the existing experiments were able to sustain a long-lasting effect and almost all weight was regained post-intervention. This poses the question of understanding why weight loss is so difficult to achieve and what factors influence weight loss and how they can neglect even the effect of short-run incentives. Thus in order to design more effective intervention for future implementations a theoretical investigation on the effects that these incentives have on individuals' behavior and why weight loss maintenance after the cessation of the incentive was not sustained is needed. In order to do so we add to existing literature by proposing another, complementary, explanation for why dieting efforts are so difficult, even in the absence of rational addiction, time-inconsistent preferences or bounded rationality.

In this chapter we present a theoretical model, with a novel assumption on individual metabolism, in order to investigate why weight-loss attempts are so difficult to achieve even if short-run incentives are present. The novelty of the model is related to the presence of a non-monotonic relationship between caloric expenditure and body weight, resulting in lower caloric expenditure for overweight/obese individuals. The theoretical analysis helps to understand why losing weight is so difficult without relying on social pressure, rational addiction, time-inconsistent preferences or bounded rationality and most importantly the conditions under which an incentive for weight loss is designed such that the individual is able to permanently lose weight. Multiple steady states and a threshold characterizing a condition of "obesity traps", from which the individual try to escapes by means of incentives, are derived. However the presence of incentives is not always able to enable a permanent effect, such effect is related to the design of the incentive scheme. A non-decreasing incentive scheme rewarding more weight loss at higher level of body weight (i.e. progressive) is able to sustain steady state decreases in body weight and food consumption, while a lump-sum is not.

Chapter 2

Social Distortion in Weight Perception: A Decomposition of the Obesity Epidemic

Abstract

This chapter examines the influence of social factors on obesity. We develop a concept of social norm related to *social distortion in weight perception* developed through shared experiences in a common social environment with a high prevalence of obesity. The theoretical model shows that when obesity is common it less likely to be recognized as a problem by mitigating individual's health concerns. We prove that our empirical measures of such social component are significant in influencing individual weight, using regional data from Health Survey for England. We exploit a markedly difference in obesity rates between 2002 and 2006 to undertake a Fairlie decomposition analysis. Our findings suggest that when we exclude social norms our estimates explain less than 50% of the obesity gap. When we include the social norms our estimates explain between 50% to 80% of the overall obesity gap. By stratifying the result by gender we notice that men had a higher obesity rates increased with respect to women, and they are more susceptible to social distortion, especially low-skilled ones. Medium- and low-skilled women, similarly, are more susceptible to environmental pressure than high-skilled ones, which are completely unaffected by it. While men are affected by a broader set of environmental pressure, women are found to be mainly affected by their closer peers. Overall these results suggest that individual body weight concerns are closely related to the actual realization of weight in his reference network.

> *"The social aspects of obesity may have a [social] multiplier effect on the growth of obesity. When obesity is relatively rare, it is considered abnormal and repulsive, and this negative response helps to keep it in check. As obesity begins to rise, the negative image of obesity becomes less intense because obesity is now more common."*

Philipson and Posner, JEL (2008)

2.1 Introduction

Although obesity, over the last two decades, has been a wide policy concern the reasons for its long-lasting incidence are yet to be clearly unraveled (Popkin, 2007). Since the average weight gain was too abrupt to being the result of a major genetic evolution (Rosin, 2008) several scholars have recently focused on some other possible explanations such as: the decrease in real food prices (Goldman et al., 2011, Philipson and Posner, 2008, 2003), the reduction of the time cost of food (Cutler et al., 2003, Ruhm, 2012), in technological changes due to economic growth (Huffman et al., 2010, Lakdawalla and Philipson, 2009) or in the introduction of welfare-improving technological change (Lakdawalla et al., 2006). However they give rise to several skepticism. In fact food prices, and most important eating time and physical activity on the job, declined substantially from early 1970s though mid 1980s due to technological innovations which allowed for cheap and energy dense food, to increases in restaurant supply (Ruhm, 2012) and to less physical demanding jobs (Philipson and Posner, 2003). These changes might have been the starting causes of an imbalance between calorie intake and expenditure however, given the fact that prices varied little afterwards they cannot, alone, entirely explain the entire spread of obesity (Christian and Rashad, 2009).

Recently peer effects and social norms have been characterized as important pathways for the biologic and behavioral spread of excessive body weight. Norms and customs developed through shared experiences in a common social environment and can exert important influences on behavior. When the majority of peers are obese and such status is common, imitative obesity could represent an important contributor to weight-related norms. This view argues that weight gain exhibits a sort of *'social contagion'* within networks of friends and family (Christakis and Fowler, 2007, Trogdon and Allaire, 2014). Such contagion may occur because judgments on body size depend on the actual size of one's friends or on the prevalence of overweight/obese peers. Obesity, in fact, spread through social ties and peers since having obese social contacts might change a person's tolerance for being obese or might influence his or her adoption of specific behaviors (Christakis and Fowler, 2007, Halliday and Kwak, 2009, Trogdon et al., 2008, Ali et al., 2012). Also eating patterns and relative satiation seems are affected by the surroundings in which individuals live, adjusting consumption according to the body type of the other (McFerran et al., 2010a,b).

For these reasons, social interaction has been explored as one of the possible channels for the obesity epidemic. Empirical literature has been mostly concentrated in disentangle the relationship between individual's perception of their own body weight with respect to their reference group, either related to a social norms (Etilé, 2007, Gil and Mora, 2011) or because overweight perceptions and dieting are influenced by a person's relative BMI (Blanchflower et al., 2009, Oswald and Powdthavee, 2007). Theoretical literature tried to provide some general predictions in order to define the form of such social component. In order to do so several attempts were made to provide a possible mechanism trough which peers could influence individual weight actions such as: exogenous social norm (Burke and Heiland, 2007, Etilé, 2007, Dragone and Savorelli, 2012), intertemporal social pressure and appearance evaluations (Wirl and Feichtinger, 2010, Strulik, 2014) or more general interplay between economic, social and psychological factors (Reich and Weibull, 2012).

Differently from all these approaches the present chapter highlights another channel: *social distortion in weight perception*, such that if overweight is wide spread this makes it less likely to be recognized as a problem by mitigating its health concerns. In the Unites States, for example, among many adults who meet the conventional body mass index (BMI) standard for overweight ($BMI \geq 25$) fewer perceived themselves so (Burke et al., 2009, Rand and Resnick, 2000, Paeratakul et al., 2002) especially between 1988-1994 to 1999-2004 when the increased in overweight/obesity was more pronounced (Johnson-Taylor et al., 2008). These underestimates are usually described as misperceptions; however this fail to recognize their social dimension and the fact that the secular increases in adult mean BMI is the main contributor to changes over time in weight perceptions. In fact these underestimates usually correspond to the actual prevalence of overweight in the representative group and rather than misperception these judgments are more likely accurate representations of their social context against the prevailing environmental distribution of weight and perception about weight are shaped by the weight of the individual himself and by the one of his peers. Additionally adult of all races, genders and weight *'tend to see weight problems everywhere but in the mirror'* (Taylor et al., 2006), so that they see the national weight problem as being greater than the weight problems of their friends and relatives[1]. Social distortions in weight perception are crucial effects to disentangle since individuals need to perceive themselves to be at risk and recognize their overweight to be a health problem in order to change their lifestyle accordingly. Mistaken perceptions have been linked to risky behaviors, whereas accurate perceptions have been associated with appropriate weight goals. We assume that individuals are less likely to perceive themselves at risk if they are exposed to overweight/obese people in their immediate environments (Johnson-Taylor et al., 2008, Taylor et al., 2006) and that the whole weight categorization evolved, due to a constant exposure to a growing population of overweight individuals current generations are more likely to misperceive their own weight status (Burke et al., 2009, Maximova et al., 2008, Ali et al., 2011).

[1] In a survey, made in 2006 in United States by the Pew Institute (Taylor et al., 2006) the percentage of individual reporting themselves as 'very overweight' is almost 25% lower than the actual trend and more than half of the respondent reported to have a weight 'just about right' when they should be categorized as 'overweight' so that they.

The present chapter analyses a theoretical and empirical model with a social component by considering an interaction between own utility and the social prevalence of excessive weight in the society at large. Depending on the prevalence of obesity the individual will have a distorted perception of which weight is supposed to represent a healthy reference and which is not, thus risking converging to an unhealthy weight due to social pressure related to the actual realization of weight in its reference network. The theoretical model provides an explanation for the recent increase in obesity relating it to the evolution of the social approval over excessive weight, which is affecting individual social image. As the society is more indulgent towards overweight/obesity the higher will be the misperception of it as a health problem and the more the society will converge to an overweight equilibrium. The empirical model using pooled data from Heath Survey of England (HSE) from 2002-2006 tests this conclusion by establishing reference group based on regional Health Authorities and several social norm measures affecting health sensitivity, all of which are increasing the odds of being obese. Additionally, using a Fairlie decomposition we are able to estimate that the 22% increase in obesity rates between 2002 (20.2%) and 2006 (24.7%), which accounts for a 5% percentage point increase, was due to social environmental components. Our findings suggest that when we exclude the social norms our estimates explain 50% of the obesity gap. When we include the social norms our estimates explain between 50% to 80% of the obesity gap. We also find that men had a higher obesity rates increase with respect to women and they are more susceptible to social distortion, especially high-skilled ones. Medium- and low-skilled women, similarly, are more susceptible to environmental pressure than low-skilled ones, which are completely unaffected by it. While men are affected by a broader set of environmental pressure, women are found to be mainly affected by their closer peers. The policy implications will be to reduce misperception and enhance health education in order to provide social-invariant weight categorization to be followed along with a higher concern for excessive weight reduction as a serious health problem.

2.2 Theoretical Model

In order to model social-identity and weight, we start from the self-identity model of Akerlof and Kranton (2000) and adapt it to our context following an approach similar to the one of Costa-Font and Jofre-Bonet (2013). Assume that individuals choose net caloric intake, meaning that they choose both food and exercise related actions, in order to maximize a utility function which depends on his action of choice but also on his social-image and health. Additionally to this personal factors, his utility function is conditioned by their peer's appearance, though individual's social image, and by sociocultural environmental factors. The utility function is

$$U_i = U_i(a_i, SI_i, H_I; z_i, Z_i) \tag{2.1}$$

Where a_i is i's net caloric intake SI_i is social image, H_i is i's health production function, z_i are i's personal characteristics while Z_i gives the environmental factors or others' actions not affecting caloric intake of i's social reference. Social image, SI_i, depends not only on i's net caloric intake, a_i, but also on others appearance and it is conditioned by i's individual characteristics and environmental factors. The equation for social-image is the following

$$SI_i = I_i(a_i, r_{-i}, z_i, Z_i) \tag{2.2}$$

Where r_{-i} is the appearance of i's group of reference.

The health production function H_i depends on net caloric intake and other individual and environmental factors

$$H = H_i(a_i; z_i, Z_i) \tag{2.3}$$

We assume that U, SI and H are well-behaved, strict concave functions.

If we maximize equation (2.1) subjected to equations (2.2), (2.3) and a budget constraint we obtain the following first-order condition:

$$\frac{\partial U_i}{\partial a_i} + \frac{\partial U_i}{\partial SI_i}\frac{\partial SI_i}{\partial a_i} + \frac{\partial U_i}{\partial H_i}\frac{\partial H_i}{\partial a_i} = \lambda P_a \tag{2.4}$$

Where λ is the Lagrange multiplier and P_a is the monetary price of net caloric intake[2]. The first term on the LHS of equation (2.4) represents the direct effect of net-caloric intake on i's utility; the second one is the indirect effect of net-caloric intake on i's utility trough the effect of social image while the third one is the health effect; in the RHS we have the monetary values of the budget constraint. The three terms on the LHS of equation (2.4) can be either positive or negative, depending on the level of net-caloric intake at which the individual is. For example if the individual is increasing his net-caloric intake beyond his optimal one, a_i^*, which is the one at which there are no negative health consequences $\dfrac{\partial H_i}{\partial a_i} = 0$, the third term in (2.4) will be negative. A similar effect can be found for SI_i, if the individual is increasing his net-caloric intake beyond the one at which $\dfrac{\partial SI_i}{\partial a_i} = 0$, the second term in (2.4) will become negative.

Since equation (2.4) implicitly define the optimal level of net-caloric intake $a_i^* = a_i(r_{-i}, SI_i, H_i; z_i Z_i)$, in order to see how this is affected by a change in the appearance of i's reference group, r_{-i}, we apply the implicit function theorem in order to get

$$\frac{\partial a_i}{\partial r_{-i}} = -\frac{\partial SI_i}{\partial r_{-i}} \cdot \frac{\partial SI_i}{\partial a_i} \cdot \kappa \tag{2.5}$$

[2] Alternatively it can be seen as the combination of food prices and exercise monetary cost including the opportunity cost of time

with $\kappa > 0$[3]. Equation (2.5) shows that the effect of others' appearance on i's net-caloric intake is related to the partial derivative of SI_i with respect to a_i ad r_{-i} without being affected by any health concerns. The second term in equation (2.5) is the one being affected by the *social distortion in weight perception*, due to the fact that i's own net-caloric intake will have a positive marginal effect on social image as long as the individual choose it lower than the socially accepted appearance and since this socially accepted appearance increased over time also the range at which a_i has a positive marginal effect in equation (2.5) increased.

Proposition 2.1 (Social Distortion in Weight Perception)**.** *Others' appearance will have a positive effect on i's net caloric intake, as long as there is a positive marginal effect of net-caloric intake on i's social image.*

The misperception presented earlier is embedded in equation (2.5) due to the fact that health concerns are irrelevant for its sign, whereas the individual is mainly concern about keeping up his social image. Due to this he will misperceive an increase in r_{-i} as beneficial and try to keep up with to it in order to maintain his social image as high as possible.

More generally an individual, for a given range of net-caloric intake, would be expected to receive a positive marginal utility from net-caloric intake, both from health and social-image. After this range the marginal impact of net-caloric intake on social-image becomes negative. The optimal level of a_i chosen in order to maximize overall utility will change depending on the relative magnitude of the positive and negative sign in equation (2.4). Given the empirical evidence, we can assume that this maximum level increased over time because the range of values for which the negative effect of eating on social image holds decreased due to an increase in the average appearances of i's reference group which affected positively $\partial SI_i/\partial a_i$ and $\partial a_i/\partial r_{-i}$ in equation (2.5). Thus the individual is pushed, in order to keep up with the social-image of his reference group, to choose an increasing level net-caloric intake a_i over his optimal one.

The sign of equation (2.5) will be useful to identify the possible pathway that increased weight had on individual net-caloric intake and ultimately weight over time in our sample and it will be also the main identifying assumption behind our decomposition analysis.

[3]Since $\kappa = \dfrac{\dfrac{\partial^2 U_i}{\partial SI_i^2}}{\dfrac{\partial^2 H_i}{\partial a_i^2}\dfrac{\partial U_i}{\partial H_i} + \left(\dfrac{\partial H_I}{\partial a_i}\right)^2 \dfrac{\partial^2 U_i}{\partial H_i^2} + \dfrac{\partial^2 SI_i}{\partial a_i^2}\dfrac{\partial U_i}{\partial SI_i} + \left(\dfrac{\partial SI}{\partial a_i}\right)^2 \dfrac{\partial^2 U_i}{\partial SI_i^2} + \dfrac{\partial^2 U_i}{\partial a_i^2}}$, for concavity of U, SI and H, $\kappa > 0$.

2.3 Data and Variables

2.3.1 Data Source

We used pooled data from five rounds (from 2002 to 2006) of the Health Survey for England (HSE), which is a cross-sectional survey with a followed nurse visit. The HSE is a nationally representative survey of individuals aged two years and over living in England. Every year a new sample is drawn and respondents are interviewed on a range of topics including demographic and socio-economic indicators, general health and psycho-social indicators. Additionally there is a follow up visit by a nurse at which various physiological measurements are taken including height and weight.

2.3.2 Obesity and Covariates

2.3.2.1 Dependent Variable: Obesity Measure

The obesity measure is computed for each respondent from the height and weight values obtained during the nurse visit. One useful feature of this dataset is that height and weight are measured by the nurse and not self-reported, reducing the likelihood of measurement errors. Due to this follow-up analysis we don't need to use any correction in order to control for reporting errors weight and height measurements as, for example, in Cawley (2004). Obesity is measured as a dummy variable taking values one if the individual has a BMI over 30 and zero otherwise.

2.3.2.2 Covariates

We use a set of covariates grouped in six categories

1. *Employment and Socio-Economic Status:* The employment variable is a binary variable taking value one if the individual is in paid employment or self-employed and zero if the individual is unemployed or out of the labor force. Socio-economic status is measured using the skill level of the individual defined either as (1) low (i.e. semi-skilled and unskilled manual); (2) medium (i.e. skilled manual and non-manual) and (3) high (i.e. professional and managerial). In order to focus on working age individuals and to avoid any biases due to retirement weight distortion, we restrict our sample to individual between 18 and 65 years old.

2. *Education Attainment:* Educational attainment is a continuous variable measuring the age at which the respondent finished his full time education at school or college minus four years; we also included its square.

3. *Health Variables:* Measures of self-reported general health, acute ill health, longstanding illness and psycho-social health. Self-reported general health is a measure of subjective general health measured in five categories from very good to very bad. Acute ill health is measured by the number of days in the last two weeks the respondent had to cut down on the thing they usually do because of illness or injury. In terms of longstanding illnesses respondents are asked whether they have an illness, disability or infirmity that has troubled them over a period of time and its type broad diseases code. Limiting longstanding illness is categorized by whether any of these illnesses limits respondents' activities in any way. Psycho-social health is measured by GHQ-12 score, where higher values indicate more severe psycho-social problems.

4. *Housing, marriage and family size:* The HSE collects information on respondents' marital status and housing tenure. We also control for the number of infants living in the household aged zero or one year and the number of children aged 2 to 15 year living in the household.

5. *Food Prices:* Monthly food and non-alcoholic beverages consumer price index (CPI) that measure changes over time in the prices of food, derived from the Food and Agriculture Organization of the United Nation Statistics (FAOSTAT).

6. *Additional control:* Such as: gender; age; ethnicity; eating habits (per week fruits and vegetables servings); rurality and, HSE year.

2.3.3 Social Distortion Proxies

Health Authorities were part of the structure of the National Health Service (NHS) in England, each one was responsible for enacting the directives and implementing fiscal policy as dictated by the Department of Health at a regional level. They were reformed in 2001[4] and then abolished in 2006[5]. During the period 2002-2006 they remained unchanged in number and size so we decide to focus on this time period in order to exploit their physical dimension and examine how the average social distortion in health in the Health Authorities affected the spread of obesity. So we construct the following three (continuous) variables in order to proxy social distortion:

- *Average BMI in the Health Authorities where the respondent lives:* Computed as the average BMI of the Health Authority where the individual lives;
- *Social Norm:* Regional and Health Authority average BMI of a reference group, defined by gender and age category, which is divided into three threshold (18-30, 31-41 and 41-65);
- *Percentage of Obese peers in the Health Authorities where the respondent lives:* Computed as the average percentage of obesity in the Health Authority where the individual lives.

[4]When the existing regional health authorities were renamed and merged to form the 28 new strategic health authorities.

[5]After 2006 the NHS decided to reorganized the Health Authorities and to reduce their number to ten.

Where we assume that social distortion in health is spread by social norm and peer effect, which we try to estimate using the three variables above. Conditional on the other covariates these effects are significant for two reasons:

1. Obesity prevalence and number of obese peers in the local area is a measure of environmental influences that affect obesity summarizing food intake and physical habits of the local population, which are affecting ego's decision. Using the terminology of Manski (1993) this is an "endogenous effect".

2. Area level obesity might also affect own obesity via what Manski (1993) identify as "endogenous" effect. This is the effect that peer obesity has on individual obesity all else equal. Holding all other characteristics constant this has an effect on individual obesity because it reflects a social norm to conform to. In this view individual obesity is related to the empirical distribution of obesity among peers (Blanchflower et al., 2009, Burke et al., 2009, Burke and Heiland, 2007, Dragone and Savorelli, 2012, Strulik, 2014)

In Figure A.6 there are the area level incidence of obesity in 2002 (top panel) and 2006 (last panel), from which we can notice that between 2002 and 2006 obesity rates grow quite substantial in almost every of the 28 Health Authority examined, accounting for a 4.5% point percentage increase in obesity between 2002 (20.2%) and 2006 (24.7%) accounting for an overall increase in obesity rates over this period by 22%.

[Figure 5 about here.]

2.4 Baseline Analysis

2.4.1 Univariate Logit

Let obesity B be a binary variable taking the value one if the individual has a BMI greater than 30 and zero otherwise. Suppose that B is a linear function of the social norm N and other variables:

$$B_i = a_0 + a_1 S_i + a_2 H_i + a_3 F_i + a_4 C_i + a_5 N_i + u_i \qquad (2.6)$$

Where u is an error term, i indexes individuals, and S, H, F and C are vectors of variables that affect employment, including human capital variables and other variables that affect tastes for work. S measures education and schooling, H measures health status, F measures home and family variables and C is a vector of additional control variables such as employment, SES, age, sex and ethnicity.

(3.1) can be estimated by a single equation univariate logit model:

$$\begin{aligned} B_i^* &= \beta_0 x_i + dN_i + \mu_{0i} \\ E[\mu_0] &= 0 \\ Var[\mu_0] &= 1 \end{aligned} \tag{2.7}$$

Where B^* is an unobserved latent variable and $x = \{S, H, F, C\}$. Empirically we observe the binary variable B that takes value one if the individual is obese ($B_i^* > 0$) and zero otherwise ($B_i^* < 0$). N is the variables representing the social norms, μ is the error term and d and β are coefficients. d is a measure of the impact of peers' appearance on individual obesity.

2.4.2 Results

The proportion of the sample in each obesity category is in Table 1. Only 38% of the sample are in the healthy weight category while 60% are either obese (24%) or overweight (36%). The overweight category has a higher probability of being employed and male[6]. Overweight and obese individual are also older, reporting lower general health but with similar marital status. Obese individuals seem to live in authorities characterized by a higher proportion of obese peers.

[Table 1 about here.]

Table A.2 presents the Logit results, where we displayed the marginal effects and the statistical significance of the peer effects of interest. In men and women all three measures of neighborhood incidence of obesity have a significant and positive effect on obesity, with the average BMI in Health Authorities being just marginally significant.

[Table 2 about here.]

These results point out to a significant and positive effect that neighborhood obesity characteristics have on individual obesity, for example a one point percentage increase in the average BMI in the health authority were the respondent lives leads to a 4% increases in the odds of being obese for men (Column (3) Table A.2) and in an almost 5% for women (Column (4) Table A.2). Similar, but not reported, analysis were run by different ethnic categories (i.e. white, black) the results are available upon request, however due to a very low sampling of non-whites they were not significant for that ethnic category.

[6]This is consistent to the inverted U-shaped result that peaks at a level of BMI level way over the healthy one, which is higher level for male with respect for female, that Caliendo and Gehrsitz (2014), Sunder and Kropfhäußer (2013) find

2.5 Obesity increases over time: Within-Group or Across-Group Changes?

2.5.1 Fairlie's Decomposition

As pointed out before over the period 2002 and 2006 obesity rates increased by almost 5 percentage point, accounting for an overall increase in obesity rates over this period by 22% [7]. Given that this increases is quite striking for a time period so short we argue that this might be due to an increases weight in the population, shifting overweight as a new standard and creating a social distortion in weight rather than to some underlying demographic makeup of the English population occurring during this period which might have caused large changes for certain demographic groups, while other groups' body composition has remained unchanged.

In order to quantify the contribution that social norm had on obesity rates over time we employ Fairlie (2005, 1999) decomposition technique as it is supposed to be more suitable in order to calculate gaps for binary variables (Font et al., 2010). Following Fairlie (2005, 1999) the decomposition for a non-linear equation of the type $p(y=1) = F(x\hat{\beta})$ can be expressed as

$$\bar{y}^E - \bar{y}^L = \left[\sum_{i=1}^{N^E} \frac{F(x^E\hat{\beta}^E)}{N^E} - \sum_{i=1}^{N^L} \frac{F(x^L\hat{\beta}^E)}{N^L}\right] + \left[\sum_{i=1}^{N^L} \frac{F(x^L\hat{\beta}^E)}{N^L} - \sum_{i=1}^{N^L} \frac{F(x^L\hat{\beta}^L)}{N^L}\right] \tag{2.8}$$

where $\bar{y}^i$ is the average obesity probability in period j ($j = E, L$ for early and late period, respectively), x_j is the set of average values of the independent variables used is period j, $\hat{\beta}^j$ is the coefficient estimated for period j, F is the cumulative distribution function from a standard normal or a logistic distribution and N refers to the sample size in each period. Differently from the threefold Blinder-Oaxaca decomposition we used in Section 2.5 in this case we have a twofold decomposition. The first term in brackets in equation (2.8) shows the part of cross-time difference that is due to group differences in the distribution of characteristics of the x_j, referred as the "explained part", while the second term in brackets represents the portion of the cross-time differences due to differences in coefficients to the exogenous covariates but it also captures differences in unobservable endowment, this is why is referred as the "unexplained part". Similarly the non-linear decomposition can be written as

$$\bar{y}^E - \bar{y}^L = \left[\sum_{i=1}^{N^E} \frac{F(x^E\hat{\beta}^L)}{N^E} - \sum_{i=1}^{N^L} \frac{F(x^L\hat{\beta}^L)}{N^L}\right] + \left[\sum_{i=1}^{N^E} \frac{F(x^E\hat{\beta}^E)}{N^E} - \sum_{i=1}^{N^E} \frac{F(x^E\hat{\beta}^L)}{N^E}\right] \tag{2.9}$$

Where in this case the estimated coefficients for the late period $\hat{\beta}^L$, are used as weights to calculate the first term of the decomposition, and the early distribution of average characteristics is used as weights

[7] This figure are in line with the OECD estimates (Sassi and Devaux, 2012)

for the second term. Given that the two decomposition in Equations (2.8) and (2.9) provide different estimates, to avoid this problem we used the coefficient estimates from a pooled sample over all cases to weight the unexplained part of the decomposition (Oaxaca and Ransom, 1994). According to Fairlie (2005), Equations (2.8) and (2.9) provide an estimation of the contribution that the explained and unexplained part have on the total difference between the two periods. We cannot rewrite those decompositions in order to match the threefold ones derived earlier, thus we will concentrate to the single contributions of individual independent variables, with a particular interest in the social norm ones.

The calculation of the separate contributions of the individual independent variables (or group of covariates) is not direct. One has to assume that $N^E = N^L$ and $\hat{\beta}^*$ is the probit coefficient estimates from a pooled sample, then the individual contribution of regressor x_k to the cross-time obesity difference can be written as,

$$\frac{1}{N^L}\sum_{i=1}^{N^L} F\left(\hat{\alpha}^* + x_{ki}^E\hat{\beta}_k^* + \sum_{m\neq k} x_{mi}^E\hat{\beta}_m^*\right) - F\left(\hat{\alpha}^* + x_{ki}^L\hat{\beta}_k^* + \sum_{m\neq k} x_{mi}^E\hat{\beta}_m^*\right) \tag{2.10}$$

Which means that the contribution of a particular variable to the difference is calculated by holding constant the contribution of the other variables[8]. Notice that the computation of equation (2.10) involves a one-to-one matching of cases between the two groups ($N_E > N_L$) and as they typically differ in size, then a large number of random sub-samples from the larger group are drawn. Each of these random sub-samples of the early sample is then matched to the late sample and finally separate decomposition estimates are calculated. The mean value of estimates from the separate decompositions is calculated and employed to derive the results for the entire early sample.

In order to control for social factors related to health authorities behind the increases in obesity rate between the two periods we include the full set of covariates. Firstly, we control for the health authority average obesity rate to capture the influence of common regional effects affecting both genders. Secondly, we insert a measure of social environment to capture the influence of peer effects or social networks. We interpret social norms as a behavioral regularity that can be measured by the behavior of a reference group, so that any deviation from the norm results in a cost. The downside is that we cannot distinguish contextual from endogenous effects (Manski, 1993).

Table A.3 display's Fairlie's decomposition of the obesity gaps between 2002 and 2006, both excluding and including the controls for social environment i.e. social norms and peer effects. Looking at the total explained difference column at the bottom of the table we see that results indicate that our model covariates altogether explain less than 50% of the overall gaps (-0.0189), when no control for social environment is included. Conversely, these determinants explain between 60% and 80% of gaps when social environment is controlled for (-0.0264;-0.348). When no social environment is included, time

[8] See Fairlie (2005) for a further discussion on how to compute standard errors

differences in age, education, drinking and smoking pattern explain the majority of the gaps. However when social environment effects are included in the model, we find that they alone override the effect of all other variable. One possible interpretation for the 40-20% change still left to be explained might be related to welfare-enhancing technological improvements which might have reduced the health impacts of chronic conditions related to obesity (Lakdawalla et al., 2006), or to the increases of less physically demanding job (Philipson and Posner, 2003). Although since we are considering a relative short time period and that these changes took place almost several years ago, it is likely that they might not have changed so much between the beginning and the end of our sample, but they might still have a higher effect on both periods due to some form of persistency effect.

[Table 3 about here.]

2.6 Stratification Analysis

2.6.1 Gender

The results in Table A.3 point out to a significant contribution that social norms had, through the channel of social distortion in weight perception, to the increase in obesity rates between 2002 and 2006, for the overall population. However since there is a well-established literature reporting a markedly difference between the effect that peer pressure has among men and women (Christakis and Fowler, 2007, Etilé, 2007, Font et al., 2010, Halliday and Kwak, 2009, Mora and Gil, 2012, Trogdon et al., 2008, Trogdon and Allaire, 2014), we decide to stratify our results for gender. In order to check whether such effect is different among gender we estimate the same decomposition of Table A.3 but splitting the sample in men and women. Results are shown in Table A.4.

[Table 4 about here.]

The results show that obesity rates between 2002 and 2006 increased for 6% for men and 3,2% for female, thus leading to a more pronounced increased in obesity rates for male rather than for female. Similarly to Table A.3, Table A.4 shows also that after we include social environmental variables, the gap explained by our model increases going from 35% to 66-83% for men and from 40% to 34-60% for women. These results show that men are more subjected to social distortion and environmental factors with respect to women. Although the total explained difference increased for both men and women after adding the environmental covariates we can notice that while for men those variables are always statistically significant (Columns (3) and (5) Table A.4), for women only the social norm one is significant (Columns (4) and (6) Table A.4), pointing out to the fact that women are influenced by a specific social norm channel specifically related to their closer peer reference.

2.6.2 SES

Additionally to gender we now investigate if there is a possible difference in subjectivity to social norms with respect to socio-economic status. As for gender differences in obesity and health behaviors we do so since there is a well-established literature connecting health behaviors to socio-economic status differences (e.g. income, job status, education) (Cutler and Lleras-Muney, 2010, Cawley, 2004, Finkelstein et al., 2005, Han et al., 2009, Rosin, 2008). Therefore we decided to divide our sample by socio-economic status using the type of labor that the individual is currently doing, rather than using income or education. To do so we stratify our sample based on occupational category by creating two main clusters:

1. *High*: This includes professional and managerial workers;

2. *Medium-Low*: This includes skilled, semi-skilled and unskilled manual works.

We decide to divide our sample in this way in order to isolate workers who work in non-manual versus manual labors and we additionally decided to pool medium and low skilled individual due to data scarcity.

Tables A.5 and A.6 show the results for men, using our two main social norm definitions. In both tables the obesity rates increase are higher for medium-low skilled men (7%) rather than for high skilled (4%). However the percentage of this increased explained by our covariates is marginally higher for highs-skilled workers (60-84%) rather than medium-low skilled ones (60-72%). However after a closer look at the coefficients we notice that the effect of the social norm covariates between the two groups is higher and more significant for medium-low skilled (-0.0209) individual rather than for high skilled (-0.0118) and all the other social environmental coefficients are more significant for medium-low skilled (See columns (1) and (2) of Table A.5 and A.6). This points out to the fact that our decomposition analysis might be more efficient in capturing the overall gap between obesity rates for high-skilled individuals between 2002 and 2006, but medium-low skilled individuals were the one more subjected to environmental pressure and therefore more affected by social distortion in weight perception.

[Table 5 about here.]

[Table 6 about here.]

Tables A.7 and A.8 show the results for women. As for men obesity rates were higher for medium-low skilled workers (4%) with respect to high-skilled (2%). Similarly from men our decomposition seems to be able to explain this gap just for medium-low skilled women (50-70%). These results point out to the fact that increasing obesity rates among high skilled women, between 2002 and 2006, are not due to environmental factors, but to some other socio-economical explanations unrelated to social norm and peer pressure. As for the results in Table A.4 the only social environmental variable that was significant for medium-low skilled women was social norm (Column (2) Table A.7), meaning that women are mainly affected by their closer peers.

[Table 7 about here.]

[Table 8 about here.]

2.7 Conclusion

The World is getting heavy and at frightening pace. More and more research was devoted in trying to disentangle the main components that over time contributed to the spread of obesity. The fall in effective and real food prices and a decrease in physical activity on the job are among the prominent explanations provided in the current literature for such abrupt change in human phenotype. However obesity spread long after these changes on eating patter and food prices occurred. A possible explanation for such long-lasting effect can be related to some social dynamics between individuals' reference groups.

This chapter explores the consequences on the equilibrium decision of a rational individual of a social pressure to conform (or not) to a given standard arising in the society, which will modify his judgment over what represents a healthy weight. Depending on the prevalence of obesity the individual will have a distorted perception of which weight is supposed to represent a healthy reference and which is not, thus risking converging to an unhealthy weight due to social pressure related to the actual realization of weight in its reference network. The theoretical model provides an explanation for the recent increase in obesity relating it to the evolution of the social approval over excessive weight, which is affecting individual social image. As the society is more indulgent towards overweight/obesity the higher will be the misperception of it as a health problem and the more the society will converge to an overweight equilibrium. The empirical model using pooled data from Heath Survey of England (HSE) from 2002-2009 tests this conclusion by establishing reference group based on regional Health Authorities and several social norm measures affecting health sensitivity, all of which are increasing the odds of being obese. Additionally, using a Fairlie decomposition we are able to estimate that the 22% increase in obesity rates between 2002 (20.2%) and 2006 (24.7%), which accounts for a 5% percentage point increase, in our sample was due to social environmental components. Our findings suggest that the when we exclude the social norms our estimates less than 50% of the obesity gap. When we include the social norms our estimates explain between 60% to 80% of the obesity gap. We also find that men had a higher obesity rates increase with respect to women and they are more susceptible to social distortion, especially low-skilled ones. Medium- and low-skilled women, similarly, are more susceptible to environmental pressure than high-skilled ones, which are completely unaffected by them. While men are affected by a broader set of environmental pressure, women are found to be mainly affected by their closer peers. The policy implications will be to reduce misperception and enhance health education in order to provide social-invariant weight categorization to be followed along with a higher concern for excessive weight reduction as a serious health problem.

Chapter 3

Body Weight and Employment: Is there an Overweight Premium?

Abstract

Using pooled data from ten waves of the Health Survey of England (HSE) and a semi-parametric regression models, this chapter tries to shed some light on the relationship between body weight, BMI and labor market outcomes. We find consistent evidence that such relationship is best represented by a General Additive Model (GAM) rather than a linear or quadratic specification. Men and women employment probability do not follow a linear relationship and are highest at a body weight way over the clinical threshold for overweight. Such results indicate that one of the main drivers for both gender employment outcomes seems to be looks rather than health. Since for men employment peaks at a higher level of BMI than women, looks is proven to be an even more powerful driving force, explaining the adverse labor market outcomes that overweight women are subjected to. Employment penalties for overweight and obese women are mainly observed in manual works, in which leaner women are treated as less fit to operate. On the other hand for men employment prospects increase with weight, at a higher rate in manual works, in which weight might be associated with strength. Instead of an "obesity penalty" we find evidence of an "overweight premium", in socially active labors, consistent with the notion of an endogenous social norm governing body weight standards.

3.1 Introduction

There is a well establishes negative association between body weight and employment prospects in the labor economics literature. A large number of studies provide evidence of a clear link between obesity and employment both in men and women, and reach the conclusion that the association reflects a causal effect of obesity on labor market outcomes (Morris, 2007, Tunceli et al., 2006). On the whole, however, the balance of evidence points to a negative influence of obesity on employment, especially, but not exclusively, for women, which seem to have a stronger negative association between body weight and employment (Mocan and Tekin, 2011). In Ireland, for instance, the probability of being employed for obese women (aged over 50) is estimated to be around 8% lower than for normal weight women, while the gap is around 5% in men (Mosca, 2013). Evidence of especially severe outcomes for women is also available from Germany, in terms of both employment and wages, where women are found to have lower employment prospects with respect to men even if they exert higher effort and/or they have lower their reservation wage (Caliendo and Lee, 2013).

More generally heavy persons are less likely to have a job than normal-weight persons (Cawley, 2000, Lindeboom et al., 2010, García Villar and Quintana-Domeque, 2009). They have fewer chances of success when they seek employment and they tend to spend longer periods of time unemployed or on welfare (Cawley and Danziger, 2005). The probability of regaining employment after a period of unemployment is similarly lower for those with a higher BMI. There is some evidence that the heavier the individual the more disadvantaged he will be in finding employment in occupations involving direct personal contact with customers (Rooth, 2007). The obese are also more likely to be inactive[1] either because they are in poor health and unable to work or because they are discriminated (Cawley and Danziger, 2005, Klarenbach et al., 2006). Discrimination against heavier job applicants is partly driven by common negative stereotypes about their attitude, self-discipline, and competence; however health concerns are also another major impediment on the hiring of heavy individual. Health problems associated with obesity may lead to temporary or permanent disability. In the United States, the odds of short-term disability episodes are increased by 76% in the obese, and by 26% in those who are overweight but not obese (Arena et al., 2006). The recent growth in obesity rates is a leading cause of increases in disability, accounting for about one third of increases in 30 to 45-year-old (Bhattacharya et al., 2008).

The pattern described by the vast majority of the labor economics literature seems to point out to a negative linear link. Even if these results are well establish two recent contributions tried to expand them and to shed some new light on the relationship between body weight and labor market outcomes. The first one is Sunder and Kropfhäußer (2013) which develops a dynamic model, differently from previous literature mostly focuses on cross-sectional models or static panel data models, in order to capture any earnings persistence (i.e. the full adjustment of wages could take some time). To do so they

[1] Defined as unemployed and not actively seeking employment.

used panel data from Germany along with an augmented dynamic Mincer regression. They showed that there is an inverse U-shaped association between BMI and wages among young workers, with the optimal BMI at a relatively high value (i.e. overweight region). Similarly, Caliendo and Gehrsitz (2014) find conclusive evidence that the relationships between BMI and labor market outcomes are poorly described by linear or quadratic OLS specifications, which have been the main approaches in previous studies, and it is best described using a semi-parametric regression. They found a similar inverted U-shaped pattern between wages, employment probability and body weight, which peaks very close to the obesity threshold for men. After some stratification analysis they concluded that looks, rather than health, is the driving force behind the adverse labor market outcomes to which overweight women are subject. Both works suggest a puzzling result pointing out to the conclusion that the movement from normal weight into the overweight category that has recently spread to a considerable share of the workforce may not jeopardize their productivity and/or employment prospects. Thus, the recent rise in weight might have not constituted a major limitation to productivity and employment yet, due to labor market adjustments.

We intend to start from these puzzling results and try to expand it in several dimensions. First of all using data from ten waves of the Health Survey of England (HSE) we extend Caliendo and Gehrsitz (2014), Sunder and Kropfhäußer (2013) to a novel dataset in order to observe if their puzzling results still holds in different populations, since they both used German data from the same survey[2]. Second, following Caliendo and Gehrsitz (2014) methodology, we apply semiparametric regression which allow us to not specify a functional form for our main covariates of interest, body mass index (BMI), but to "let the data speaks for itself" since the semiparamatric regression allows for a more flexible function form. Third we use a nationally representative sample with in-house nurse interviews, which reported true values for weight and height and not self-reported ones as in the majority of surveys. Therefore we don't need to apply any correction mechanism (Cawley, 2004), without worrying on potential reporting errors which might have caused the puzzling results. Fourth we stratify our analysis, dividing the sample in occupation where additional strength, expressed in additional weight, might be rewarded, from occupation in which, due to social interactions, a certain look is more reward. Lastly, since BMI bears the problem of not distinguish between muscle and fat mass (Burkhauser and Cawley, 2008), we followed a technique pioneered by Wada and Tekin (2010) in order to construct two more explicit measures of body composition, fat-free mass (FFM) and body fat (BF), which will be then used in order to validate our main results for BMI.

We find evidence that employment probability peaks for men and women way after the overweight threshold, with the peak for men closer to the obesity one. We also compare our results with a parametric Logit model which would have suggested that women employment probability would have continuously declined with body weight, while for men it would continuously increases (although not statically significant). These results can be interpreted as the fact that is physical appearance the driving force behind employment prospect rather than health concerns, since we also control for a

[2]The German Socio-economic Panel (GSOEP).

wide range of health variables. When we stratify the sample for occupational categories we find that health concerns are not the main drivers in different employment prospects based on body weight. For women in socially active jobs, looks-based discrimination seems to be the leading employment driver, while in manual labor women are subjected to possible strength discrimination with respect to men. Men are also subjected to this look-based discrimination in socially active while in manual labor physical fitness and strength are the major drivers. These results are also confirmed when we use an alternative measures of body composition, namely fat free mass (FFM) and body fat (BF).

Overall we find no trace of an "obesity penalty", but the relationship point out to a sort of "overweight premium". Such result is in line with the notion of an endogenous social norm which has been recently updated due to an increase in the average body weight (Burke and Heiland, 2007, Dragone and Savorelli, 2012), forming a sort of *imitative obesity* (Blanchflower et al., 2009). This increases in the threshold for overweight during the last 20 years, might have influenced not only how individuals feel about themselves but also their chances in the labor market. As the incidence of overweight increased overtime so did the social norm and the threshold at which it was previously identified as a health problem increased (Johnson et al., 2012). This in turn created a new overweight standard (Burke et al., 2009) which is also reflected in the employment prospects that individual have on the labor market, reflected in a higher employment rate for individual meeting this new body weight standard. Therefore "a certain look" seems to matter mainly in employment prospects, however the whole categorization of which is the look to converge to seems to have increased over time.

3.2 Methodology

3.2.1 Baseline Estimation

Let employment Y be a binary variable taking the value one if the individual is paid employment and zero otherwise. Suppose that Y is a linear function of BMI (B) and other variables:

$$Y_i = a_0 + a_1 S_i + a_2 H_i + a_3 F_i + a_4 C_i + a_5 B_i + u_i \tag{3.1}$$

Where u is an error term, i indexes individuals, and S, H, F and C are vectors of variables that affect employment, including human capital variables and other variables that affect tastes for work. S measures education and schooling, H measures health status, F measures home and family variables that affect tastes for work and C is a vector of additional control variables that affect employment such as age, sex and ethnicity. Equation (3.1) can be estimated by a single univariate Logit model

$$\begin{aligned} y_i^* &= \beta_0 X_i + dB_i + \mu_{0i} \\ E[\mu_0] &= 0 \\ Var[\mu_0] &= 1 \end{aligned} \tag{3.2}$$

Where y^* is an unobserved latent variable and $X = \{S, H, F, C\}$. Empirically we observe the binary variable y that takes value one if the individual is employed ($y_i^* > 0$) and zero otherwise ($y_i^* < 0$). B is a continuous variable, measuring individual's BMI. From the Logit model then we compute the marginal effects (ME) of BMI, which captures the marginal impact of a one-unit change in BMI on the probability of employment.

3.2.2 Local estimation and Generalized Additive Models (GAMs)

With local estimation we refer to the description of the relationship between two variables by means of a series of local estimates rather than with a single parameter. Such local estimator produces an estimate such as a mean or regression, between Y and X, for some restricted range of X and then repeated it across the whole range of values. This series of local estimates is then aggregated with a line to summarize the relationship between the two variables. The power of this resulting non-parametric estimate is that it does not impose a particular functional form on the relationship between the two variables and due to the local nature of its estimation process it provides very flexible fits. Local models such as non-parametric and semiparametric regression estimate the functional form between two variables while global models impose a functional form on the data. Semiparametric regression models are often referred to as either additive or generalized additive models (GAMs) since they incorporate local estimation models into standard linear models. They allow modeling some predictor variables with non-parametric regression, while other predictor variables are estimated in a standard fashion. This means that if we suspect that a single continuous variable could have a nonlinear form we can model it non-parametrically, while the rest of the model specification is estimated parametrically. Given that GAMs rely on non-parametric regression, the assumption of a global fit between X and Y is replaced with local fitting, without dispense from the assumption of additive effects.

For categorical dependent variables the GAMs are designed to make use on the strengths of Generalized Linear Models (GLMs) without requiring the problematic steps of a priori estimation of response curve shape or a specific parametric response function[3]. They employ a class of equations called "smoothers" in order to generalize data into smooth curves, by local fitting to a subset of the data. The idea behind GAMs is to "plot" the value of the dependent variable along a single independent variable, and then to calculate a smooth curve that goes through the data as well as possible. It would be possible to do a similar thing by using a polynomial of high enough order to get a curve that went through every point. It is likely, however, that the curve would "wiggle" excessively, and not represent a smooth fit. The approach generally employed with GAMs is to divide the data into some number of sections, using "knots" at the ends of the sections. Then a low order polynomial or spline function is fit to the data in the section, with the added constraint that the second derivative of the function at the knots

[3]It should be noted that GLMs such as probit or logit models are still linear and parametric in their functional form. Only the application of the link function, such as the normal cumulative distribution in a probit model, induces some degree of non-linearity.

must be the same for both sections sharing that knot. This eliminates kinks in the curve, and ensures that it is smooth and continuous at all points.

For our purpose of interest GAM allows for a semi-parametric estimation of a regression model of the following form

$$Y_i = X_i\beta + f(BMI_i) + \epsilon_i \tag{3.3}$$

Where Y_i is an employment status dummy variable of person i. Our explanatory variable of interest, BMI_i, enters the model in a non-parametrical form, whereas we assume a linear, parametric functional form for the Xs. This method allows us to accommodate for pattern that cannot be observed by simple linear regressions, even if they include higher order terms (e.g. quadratic or cubic terms). To specify a GAM, we allow the linear predictor to be some smooth function estimated from the data.

In the logistic GAM, the basic idea is to replace the linear predictor with an additive one. To explain the mechanism let's focus on a purely additive model of the form:

$$\begin{aligned} g(p(x)) &= f_0 + \sum_{j=1}^{p} f_j(x_{ij}) \\ p(x) &= \frac{\exp(f_0 + \sum_{j=1}^{p} f_j(x_{ij}))}{1 + \exp(f_0 + \sum_{j=1}^{p} f_j(x_{ij}))} \end{aligned} \tag{3.4}$$

Where $g(.)$ is a known link function, in our case a logistic, f_i are smooth unknown functions.

Generally let $E(Y|X) = \mu$

$$\nu(x) = g(\mu) \tag{3.5}$$

Where ν is a function of p variables. Assume $Y = \nu(x) + \epsilon$ given some initial estimate of $\nu(x)$, construct the adjusted dependent variable (i.e. pseudo data)

$$Z_i = \nu_i + (y_i - \mu_i)\frac{\partial \nu_i}{\partial \mu_i} \tag{3.6}$$

Instead of fitting an additive model to Y, we fit an additive model to the Z's, treating it as the response variable. Then we apply the iterated reweighed least square (IRLS) (Hastie and Tibshirani, 1990, Wood, 2006), by initiating $f_o = g(E(Y))$ and $f_1^0 = ... = f_p^0 = 0$. Then iterate with $m = m + 1$.

From the adjusted dependent variable

$$Z_i = \nu_i^{m-1} + (y_i - \mu_i^{m-1})\frac{\partial \nu_i}{\partial \mu_i^{m-1}} \tag{3.7}$$

where

$$\nu^{m-1} = f_0 + \sum_{j=1}^{p} f_j^{m-1}(X_j)$$
$$\nu^{m-1} = g(\mu^{m-1})$$
$$\mu^{m-1} = g^{-1}(\nu_i)$$

Then form the weights $W = \left(\frac{\partial \nu_i^{m-1}}{\partial \mu_i}\right)^2 V_i^{-1}$ with $V_i = Var(Y_i)$. Fit an additive model to Z using the backfitting algorithm[4] with weight W, to obtain estimated function f_j^m, additive predictor, ν_m and fitted values $\mu_i^m = p_i$. Repeat the iteration process until the change in deviance is sufficiently small. The last step of this iteration process is simply the additive regression backfitting algorithm[5] In this semi-parametric setting, the regression in step two is fitted using a smoother, in our case we used penalized cubic regression splines to smooth the estimated residuals of BMI. In the case of additional (linear) covariates, as in the case of (3.3), the same procedure described above is used, and a linear lest square fit would be used to "smooth" a binary covariate or a continuous covariate for which a linear fit was desired (Wood, 2006).

All the estimations of equation (3.3) were made using the statistical software R. In R the smoothing parameter determining the number of knots is by default is chosen via generalized cross-validation (GCV)[6]. The idea is simple; let the data speak, and draw a simple smooth curve through the data. The problem is determining goodness-of-fit and error terms for a curve fit by eye. GAMs make this unnecessary, and fit the curve algorithmically, using the GCV, in a way that allows error terms to be estimated precisely, using either IRLS or back fitting algorithm depending on the type of the dependent variable. Some researchers point out to the fact that such automated smoothing, since it is not manually controlled by the users, could lead to overfitting. Loader (1996) suggests to adjust the smoothing parameter using visual inspection i.e. if the model displaying implausible wiggles everywhere, that indicates that it is overfitted. This issue and overfitting turns not to be a concern. Available upon request we can provide graphs where we manually chose the smoothing parameter by rounding the parameter provided by GCV to the closest integer, from which is clear that the difference between the automated GCV bandwidth selection and the manual are negligible. None of the two

[4]This back fitting algorithm is generally used to fit (3.3) with continuous independent variable, we refer to (Hastie and Tibshirani, 1990, Keele, 2008, Caliendo and Gehrsitz, 2014), for more in depth explanation of such mechanism.

[5]Such back fitting algorithm (Hastie and Tibshirani, 1990, Keele, 2008) involves an iterative process based on partial residuals. We use as starting values $\hat{\beta}_0 = E(Y)$ and $\hat{f}_j = X_j$ for all j, which are collected in matrix $\mathbf{R}_j$. In the first step, partial residuals are obtained for each variable using these starting values. For example, $\bar{\epsilon}(X_1)$ is obtained as $\hat{\epsilon}(X_1) = Y_j - \sum_{j=2}^{k} R_j - (X_j) - E(Y)$. In the second step, each partial residuals is regressed on the corresponding X-column. This means that $\hat{\epsilon}(X_1)$ is regressed on X_1, $\hat{\epsilon}(X_2)$ is regressed on X_2 and so on and so forth. The resulting coefficients are used to update matrix $\mathbf{R}_j$ before the iteration starts over from the first step with weights, so that $\hat{R}_j^m(X_i)$ denotes the estimate of $R_j(.)$ at the mth. The procedure is repeated until the model converges in term of infinitesimally small changes in the residual sum of squares i.e. when $RSS = E(Y_j - \sum_{j=2}^{k} R_j^m(X_j) - E(Y))^2$ fails to decreases.

[6]Which is a general technique foe assessing model fit based on resampling that can be applied to most statistical models

approaches yielded any excessive spikes or wiggles. If anything the GCV algorithm leads to slightly more conservative estimates of the relationship between body weight and employment.

Before concluding there is a final methodological concerns to be discussed, *endogeneity* and reverse causality. The most common strategy used in order to deal with these problems is instrumental variable, using a variety of well-established instruments such as: area level of obesity (Morris, 2007), family member's weight (Brunello and d'Hombers, 2007, Cawley, 2000) or genetic predisposition (Han et al., 2009). However since we intend to extend the first attempt made by Caliendo and Gehrsitz (2014), in the road of disentangling a possible non-linear relationship between body weight and employment due to a mis-specification of the functional form relating them, we also decided to abstain from applying any instrumental variable approach, even if in our sample we could have easily used area incident of obesity as in Morris (2007) in order to get causally sound effects. Therefore, similar to Caliendo and Gehrsitz (2014), our results should not be interpreted as *fully* causal.

3.3 Data and Variables

We use data from the Health Survey of England (HSE) as the basis for our analysis. The HSE is a nationally representative survey of individual aged two years and over living in England. Our sample consist of pooled data from ten rounds (from 2002 to 2011) HSE. The BMI (i.e. adult's weight in kilograms divided by the square of his height in meters) measure is computed for each respondent from height and weight values obtained during nurse interviews, reducing the likelihood of self-reporting biases.

[Table 9 about here.]

The dependent variable in our employment specification is a binary variable taking the value one if the individual is in paid employment or self-employed, and works at least one hour per week, and zero if the individual is unemployed or out of the labor force. In order to avoid issue related either to retirement or schooling, we limit our sample to persons between the ages of 20 and 65. Pooling data from different waves we end up with a final sample of more than 55,000 observations.

The average BMI in our sample is 26.5 for women and 27.3 for men (see Table A.9). Almost 70% of the men in our sample is either overweight or obese, compared to a little more than 55% for women. Our descriptive statistics doesn't show any considerable difference between overweight and normal people in relation to employment status.

In our empirical analysis we control for educational attainment measured by year of schooling[7] (and its square), marital status, type of housing, the number of infants living in the household aged zero or

[7]Continuous variables measured as the age at which the respondent finished their full time continuous education at school or college minus four years.

one year and the number of children aged 2 to 15 year living in the household. We also include health covariates such as measured of self-reported general health, acute ill health, longstanding illness and psycho-social health. Self-reported general health is a measure of subjective general health measured in five categories from very good to very bad. Acute ill health is measured by the number of days in the last two weeks the respondent had to cut down on the thing they usually do because of illness or injury. In terms of longstanding illnesses respondents are asked whether they have an illness, disability or infirmity that has troubled them over a period of time and its type broad diseases code. Limiting longstanding illness is categorized by whether any of these illnesses limits respondents' activities in any way. Comorbidities are measured by the number of longstanding illnesses. Psycho-social health is measured by GHQ-12 score, where higher values indicate more severe psycho-social problems. We also include a healthy eating variable (Eating Habits) measuring the total portion of fruits and vegetables ate in the last week. Additional control variables that may affect employment included in the analysis are: gender; age; ethnicity; rurality; region of residence; eating habits; and, HSE year.

3.4 Empirical Results

We divide our empirical analysis as follows. First of we compare the results from a Logit regression on employment status with our semi-parametric model in order to check whether the linear function form describe the relationship between body weight and employment or if there are some additional non-linearity that cannot be accounted for. Another reason for which we compare these two results is to investigate whether or not our semi-parametric estimation leads to distorted coefficients with respect to the univariate Logit. After this we run some stratification analysis on our semi-parametric results, by occupational category. Furthermore we run a final robustness checks, by applying an estimation technique pioneered by Wada and Tekin (2010) and recently adopted by Caliendo and Gehrsitz (2014) in order to derive body composition measures from survey data, and to compare our results against Logit model on Fat Free Mass (FFM) and Body Fat (BF).

3.4.1 Bodyweight and Employment: Logit and Semi-parametric Results

The main results of our semi-parametric analysis can be found in Table A.10. We can note that coefficients from the control variables are very similar in size across the Logit (column (1) and (2)) and the GAM model (column (3)), also the standard errors are quite similar, with only minor inflation in the GAM model. Additionally, the likelihood ratio tests for the specifications for both men and women suggest that our semiparametric regressions differ significantly from parametric linear or quadratic setups (see p-values in bottom row of Table A.10). Accordingly, it required the flexible functional form of our semiparametric model to uncover the true relationship between hourly wages and body weight.

[Table 10 about here.]

Our results suggest that the linear functional form assumptions mask important relationships between the dependent variable and our main explanatory variable of interest. A regular Logit model, whose marginal effects are provided in columns (1), (2), (4) and (5) in Table A.10, would have suggested that the employment probability for women continuously declines with increased weight (column (4)). However, our GAM suggests that the pattern is not linear. However, given that body mass index enters our estimation non-parametrically and the GAM estimation process is local as opposed to global, it is impossible to use a single parameter to describe the relationship between body mass index and employment and so we cannot report any coefficient for body mass index. Thus we interpret its relationship with employment graphically, using Figure A.6.

[Figure 6 about here.]

Our model indicates that the probability of being in employment is highest for women with a BMI of around 27, and then decreases in additional body weight. Instead of speaking of an "obesity penalty", the Figure A.6 suggests that the relationship between employment probability and body weight could be better defined as "overweight premium". Another interesting pattern that we can observe is that there seems to be a higher penalty for underweight women with respect to heavier ones. Similarly for men, employment probability is highest at an even higher BMI than the one observed for women (around 28), with overall similar relationship shape. The major difference between the two genders is that employment prospects for men peak very close to the obesity threshold.

Given that the body weight penalty seems to have shifted and soften up we can rule out any major health concerns driving employment decisions. In fact the difference between the peak region and the region the relationship starts level out cannot be entirely attributed to health effects given also the fact that we control for health using multiple variables. This can be interpreted as the fact that is looks and physical appearance to be the driving force behind these results. However, differently from previous literature, there seems to be a new established look standard, closer to overweight (Burke et al., 2009), in England.

One interpretation might be related to a new established social norm which is updated as the average body weight in the society evolves overtime, forming a sort of imitative obesity as overweight perceptions and dieting are influenced by a person's relative BMI (Blanchflower et al., 2009). As overweight rates increased over time, so did this new social norm and the threshold at which the penalty was set disappeared. Thus employment prospects are higher for overweight women and men since overweight become a new standard (Burke et al., 2009) and thus a certain "look" became a more prominent employment reasons with respect to health concerns.

Therefore, we conclude that the GAM model has the appropriate functional form, suggesting that body weight is positively associated with employment probabilities, at a diminishing rate, with a peak reached in the overweight region for women and close to the obesity threshold for men. One might argue that such relationship could also being spotted using a quadratic GLM (column (2) and (5) in Table A.10) since the introduction of a quadratic BMI term turns out to be significant and negative. However a regular GLM model, would have suggested that employment prospects would steadily increases for men and decreases for women without the introduction of the quadratic term. This means that if we miss the introduction of such quadratic term in a regular GLM we could have missed the true link between body weight and employment rates. Moreover, given that using a GAM, we let the data speak for themselves we don't have this risk since the estimation of the model automatically fits the best relationship without forcing us to manually choose it (even if we could do it) at the risk of misinterpreting it.

3.4.2 Occupational Category

In section 3.4.1 we already established that physical appearances rather than health concerns are the main driver through body weight is affecting individuals' employment prospects. We now try to stratify these results for occupational categories dividing our sample in manual labor and non-manual labor. In order to do so we use code from the social class of individual, which divide occupation in professional, manual and non-manual divided in six different level of skills. The reason behind this separation is that a certain looks is supposed to be more important for non-manual and professional labors in which social interactions with customers and co-workers is more frequent and a certain look might be associated with a higher degree of trust both from a professional and a human point of view. Thus we expect to find that looks, especially for women, to be more important for non-manual and professional occupation. To make the results more compact we decide to include professional workers in the non-manual cluster. Available upon request we have tables and estimates were we created an additional cluster just for professional workers. However due to the fact that professional workers are in small numbers, with respect to the other two cluster[8] and their results are very close to the one for non-manual workers we decided to cluster them together. As in the previous case results are reported graphically since BMI enters our regression non-parametrically (see Figure A.7).

[Table 11 about here.]

[Figure 7 about here.]

For woman our results suggest that physical appearances is more important in manual labor, for which employment peaks at a BMI close to 24; while in non-manual labor we observe a similar patter as the one reported in the previous section. To support these results a likelihood ration test (see Table A.11) shows that the semiparametric specification is statistically different from the parametric one, although

[8]In our final sample we have 2,682 professional workers, 24,957 non-manual workers and 16,697 manual workers.

just marginally for manual labor occupation (p-values = 0.0444). These results suggest that the new established physical appearance observed in the results in the previous section is more important for women in labor with higher social interaction; while for manual labor women are subject to a lower standard of body weight in which a heavier appearance might be associated with lower strength.

For men we find a similar although less steep results. For both occupational categories employment prospect peaks at a BMI close to obesity as in the previous section. A likelihood ratio test (see Table A.11) show that the semiparametric model is statistically different from a parametric results, but that it does not add much value in non-manual labor occupation (p-value = 0.0148). Similarly for women, look seems to be a more important driver for non-manual labor, while in manual labor a certain level of BMI seems to be associated with a higher employment due to strength and physical fitness concerns. In fact for manual labor there are steeper decreases in employment probability in moving from the peak to higher/lower body weight.

3.5 Fat Free Mass and Body Fat

As a final robustness check, using the method pioneered by Wada and Tekin (2010), we construct two measures of body compositions: fat free mass (FFM) and body (BF). The method used by Wada and Tekin (2010) consisted in accessing the NHANES II survey, which collected data on bioelectrical impedance analysis (BIA), from which they constructed exact measures of FFM and BF; then they subsequently regressed these two measures on respondent age, weigh, height, as well as second and third order and polynomials of these variables. From these regressions it turned out that these variables are very good predictors of FFM and BF, so we can construct pseudo FFM and BF using other datasets and employing the coefficients reported in Wada and Tekin (2010)[9].

It is noteworthy to say that while BF consists just in fat tissue, FFM mostly consists in muscles, lean mass and bone. Therefore BF is associated with the ill effect of obesity; while FFM is associated with health and physical fitness. This means that increases in BF can be associated with an incremental effect of obesity, while increases in FFM can be considered as the effect of healthy body grown (Allison et al., 2002, Heitmann et al., 2000). Obviously this two-compartment model is able to isolated the two diverging effect of FFM and BF, which in a single-compartment model, using for example BMI, are likely to result in a situation where they cancel each other resulting in an insignificant coefficient on BMI. We expect also that, with respect to beauty standards an increase in BF is likely to report physical unattractiveness, especially for women, while an increase in FFM is associated with physical fitness and beauty (Popkin, 2007, Stearns, 2002).

[Table 12 about here.]

[9]For a deeper discussion on the method, see Wada and Tekin (2010).

In Table A.12 we can see that, as in Section 3.4.1, the coefficients in the GAMs and the GLMs model are very similar in terms of magnitude; also the standard errors are quite similar. A likelihood ration test shows again that the semiparametric model is statistically different from a parametric ones and it is able to capture the true relationship between FFM and BF. The results are presented graphically in the Figures A.8 and A.9.

[Figure 8 about here.]

[Figure 9 about here.]

For men the relationship between employment and body composition is steeper for both measures of body composition in manual labor with respect to non-manual ones. This is somehow intuitive since in this type of jobs physical fitness is essential. A likelihood ratio test cannot reject the hypothesis that the parametric and the semiparametric model come to same results for FFM (p-value= 0.1) and BF (p-value= 0.0242) in non-manual occupations, meaning that a subtle difference between body composition is crucial for employment prospects in heavy duty labor with respect to more socially active ones. In fact figure A.8 shows men employed in manual labor the peak is at a higher value for FFM, giving reasoning to a strength-based penalty in manual job.

For women the results are, in general, less steep than the one for men and confirm the results of a higher tolerance for overweight in non-manual labor. A likelihood ratio test cannot reject the hypothesis that the parametric and the semiparametric model come to same results for FFM (p-value= 0.293) and BF (p-value= 0.15) in manual occupations, given reasoning to the fact that women are considered less fit to manual job by default. For what concerns non-manual occupation figure A.9 confirms the results of section 3.4.1.

It should be noted that this application of Wada and Tekin (2010) methodology, even if it is confirming our main results, should be treated with cautions. This is because we are using estimates from a US based study to construct the measure of FFM and BF for an entire different population. In particular even if our FFM measures are just marginally smaller than the one in Wada and Tekin (2010) (61 against 63 for men and 43 against 44 in women) the BF for women might be a too large since it averages at 27 with respect to 23 in Wada and Tekin (2010)[10]. However this robustness check yields some interesting results, consistent with our results in section 3.4.1. Namely that looks drive the results for women in non-manual labor and physical strength for men in manual labor.

[10]No such difference was found for men.

3.6 Conclusion and Comments

The aim of this chapter was to apply a semiparametric regression model in order to uncover the employment effects of body weight which are not observable employing linear regressions. Differently from previous studies which highlighted the presence of an "obesity penalty" for both gender, but higher for women, we find a different interpretation. Similar to Caliendo and Gehrsitz (2014), Sunder and Kropfhäußer (2013) we find that the notion of a penalty is somewhat misleading; in fact we find that there is an "overweight premium". By stratifying our results we found that such new arising standard is more pronounced for women in socially active occupations. While the results for men tend to point out to the fact that strength rather than look seems to matter the most in both occupations.

The interpretation of this result can be interpreted in the light of the notion of an endogenous social norm, governing body weight judgments. With respect to body weight, when the majority of peers are overweight and heavy body weight is common, imitative overweight (Blanchflower et al., 2009) can be a prominent feature of the social environment and an important contributor to weight-related norms (Burke and Heiland, 2007, Dragone and Savorelli, 2012). If overweight is wide spread, and the agent captures this spread as a positive reinforcement, this might make it less likely to recognize overweight as a problem by mitigating the health consequences of body weight. Due to this misperception, many adults who meet the conventional BMI standard for overweight ($BMI \geq 25$) might underestimate their weight category, and feel "just about right" (Burke et al., 2009). These underestimates of BMI categories are usually described as misperceptions, however this fails to recognize their social dimension. In fact these underestimates usually correspond to the actual prevalence of overweight in their representative group. Rather than misperception these judgments are more likely accurate representations of their social context against the prevailing environmental distribution of weight.

Therefore this increases in the threshold for overweight during the last 20 years, might have influenced not only how individuals feel about themselves but also their chances in the labor market. As the incidence of overweight increased overtime so did the social norm and the threshold at which it was previously identified as a health problem increased (Johnson et al., 2012). This in turn created a new overweight standard (Burke et al., 2009) which is also reflected in the employment prospects that individual have on the labor market, reflected in a higher employment rate for individual meeting this new body weight standard. Therefore "a certain look" seems to matter mainly in employment prospects, however the whole categorization of which is the look to converge to seems to have increased over time, creating an "overweight premium". Similarly to Caliendo and Gehrsitz (2014), Sunder and Kropfhäußer (2013), we find that employment prospects picks at a relatively high level of BMI. The relatively high BMI, at which the relationship becomes negative, is puzzling and somehow different front the general findings in the current literature. It suggests that the movement from normal weight into the overweight category by a considerable share of the workforce may not compromise so much employment prospects, due to an increased weight standard arising as a social norm in the society.

Appendix A

Obesity Traps and Incentives for Weight Loss

A.1 The intertemporal problem (1.3) - (1.6)

In this appendix we will prove that the intertemporal problem (1.3) - (1.6), with an infinite time horizon and constant discount rate, is an approximation of a similar model with stochastic time of death.

A.1.1 The Model with stochastic time of death

In order to take into account that individual's longevity is stochastic and impossible to known before time of death, we assume that the agent's life is finite but with an uncertain terminal date, T, (Yaari, 1965). $F(T)$ represents the probability o dying at time T and $f(T)$ is the associated density function. Given initial body weight w_0, the agent must choose the path of food consumption satisfying the following intertemporal problem

$$\max_{c(t)} E\left[\int_0^T e^{-\tilde{\rho}t}[U(c(t),w(t))] + \alpha I(w^H, w(t))dt\right] \tag{A.1}$$

$$\text{subject to } \dot{w}(t) = c(t) - g(w(t)) \tag{A.2}$$

$$w(t), c(t) \geq 0 \tag{A.3}$$

$$w(0) = w_0, \text{ given} \tag{A.4}$$

where $\tilde{\rho} > 0$ is the intertemporal discount rate representing the agent's impatience. The objective function (A.1) , differently from (1.3) represents the expected lifetime utility function of an agent with stochastic terminal time. Following Dragone et al. (2013), Yaari (1965) we can exploit a very useful result in order to prove the equivalency of the two. Equation (A.1) can be equivalently represented in terms of the following objective function

$$\int_0^T [1 - F(t)]e^{-\tilde{\rho}t}U(.) + \alpha I(.)dt$$

where $1 - F(t)$ is the probability of living beyond t (Yaari, 1965, Levy, 2002). In order to prove the equivalency of (A.1) and (1.3) we will focus on the special case where $f(T) = \hat{\rho}e^{-\hat{\rho}t}$ so that the density function is exponential. Under this assumption the expected intertemporal utility of the agent can be equivalently written as follows

$$E\left[\int_0^T e^{-\tilde{\rho}t}U(.) + \alpha I(.)dt\right] = \int_0^T e^{-\hat{\rho}t}e^{-\tilde{\rho}t}U(.) + \alpha I(.)dt = \int_0^T e^{-\rho t}U(.) + \alpha I(.)dt \tag{A.5}$$

where $\rho = \hat{\rho} + \tilde{\rho}$ and it is the overall discount rate depending on impatience ($\tilde{\rho}$) and on the hazard rate ($\hat{\rho}$). Equation (A.5) represents the discounted stream of utility of an infinitely-lived agent, however, an alternative interpretation is possible, whereby the objective function (A.5) represents the expected intertemporal utility of an agent with stochastic life and whose hazard rate is constant. Thus, providing a bridge between finite and infinite horizon models.

A.2 Concavity of the Hamiltonian function

In this appendix a formal analysis on the concavity of the Hamiltonian functions of the paper is provided. Assuming that the Utility function is separable in c and w such that $U_{cw} = 0$ we can write the Hessian of the Hamiltonian as

$$\tilde{H} = \begin{bmatrix} H(.)_{\lambda\lambda} & H(.)_{\lambda c} & H(.)_{\lambda w} \\ H(.)_{c\lambda} & H_{cc} & H(.)_{cw} \\ H(.)_{w\lambda} & H(.)_{wc} & H(.)_{ww} \end{bmatrix} = \begin{bmatrix} 0 & 1 & -g_w \\ 1 & H_{cc} & 0 \\ -g_w & 0 & H_{ww} \end{bmatrix}$$

with $H_{cc} = U_{cc}$ and $H_{ww} = U_{ww} - \lambda g_{ww}$.

If the last n minors alternate in sign provided that the $\det \tilde{H}$ itself has the sign of $(-1)^n$, where n is the number of variable in the utility function, then the point is a local maximum i.e. the Hessian is concave. On the other hand if the largest n minors all have positive sign and that sign is $(-1)^m$ where m is the number of multipliers in the problem the point is a local minimum. Thus, concavity requires that the principal largest leading minor of $\tilde{H}$ to be mixed in signs:

$$|\tilde{H}_2| = \begin{vmatrix} 0 & 1 \\ 1 & H_{cc} \end{vmatrix} = -1$$

$$|\tilde{H}_3| = \begin{vmatrix} 0 & 1 & -g_w \\ 1 & H_{cc} & 0 \\ -g_w & 0 & H_{ww} \end{vmatrix} = -g_w^2 H_{cc} - H_{ww} > 0$$

Given the assumption on $U(.)$ and $g(.)$ joint concavity is satisfied as long as $g_w^2 U_{cc} + U_{ww} + \alpha I_{ww} > \lambda g_{ww}$.

A.3 The Derivation of Condition (1.14)

The Jacobian associated with the intertemporal problem (1.3)-(1.6) is

$$J = \begin{pmatrix} \rho + g_w & \dfrac{U_{ww} + g_{ww} U_c}{U_{cc}} \\ 1 & -g_w \end{pmatrix} \tag{A.6}$$

Given that the trace is positive and equals to $tr(J) = \rho > 0$, the steady state can be at most a saddle point[1]. The determinant is harder to sign

$$\det(J) = -g_w\left(\rho + g_w\right) - \frac{(U_{ww} + g_{ww} U_c)}{U_{cc}} \tag{A.7}$$

A negative determinant, $\det(J) < 0$ implies that one of the eigenvalues is negative, characterizing a saddle point[2] A positive determinant, $\det(J) > 0$, is then a necessary and sufficient conditions to exclude a saddle (i.e. conditional) stability, both eigenvalues are positive so that the associated steady state is an unstable node. To characterize the cases in which such instability is achieved we can rewrite the determinant as

$$\det(J) = -g_w\left(\rho + g_w\right) - \frac{H_{ww}}{U_{cc}} \tag{A.8}$$

[1] At least one of the two eigenvalues of J

$$e_{1,2} = \frac{1}{2}\left[tr(J) \pm \sqrt{tr(J)^2 - 4\det(J)}\right]$$

is positive (or has positive real parts)

[2] Which even if unstable has a stable manifold (of dimension one) determining optimal policy. Given the positive trace with a little but of abuse of notation I will define *'stable'* such outcomes, since it is the maximum achievable stability given by the canonical equation in (8).

given the assumptions on $U(.)$ and $g(.)$ H is concave and the last term of equation (26) is negative meaning that a positive determinant is only possible when, $g_w < 0$, providing the following sufficient condition for instability

$$-g_w(\rho + g_w) > \frac{H_{ww}}{U_{cc}} \tag{A.9}$$

A.4 Comparative Statics

In order to compute the effect of a change in a parameter x on steady-state levels of body weight (w^{ss}) and food consumption (c^{ss}), we apply the Cramer's rule to the following system (all functions are evaluated at the steady state)

$$\begin{pmatrix} \frac{\partial \dot{c}}{\partial c} & \frac{\partial \dot{c}}{\partial w} \\ \\ \frac{\partial \dot{w}}{\partial c} & \frac{\partial \dot{w}}{\partial w} \end{pmatrix} \begin{pmatrix} \frac{\partial c^{ss}}{\partial x} \\ \\ \frac{\partial w^{ss}}{\partial x} \end{pmatrix} = \begin{pmatrix} -\frac{\partial \dot{c}}{\partial x} \\ \\ -\frac{\partial \dot{w}}{\partial x} \end{pmatrix}$$

where the first term is the Jacobian matrix ($|J|$) of the system of differential equations (16) evaluated at the steady state. Applying Cramer's rule for a change in α we have

$$\begin{pmatrix} \frac{\partial c^{ss}}{\partial x} \\ \\ \frac{\partial w^{ss}}{\partial x} \end{pmatrix} = \begin{pmatrix} -\frac{|K|}{|J|} \\ \\ -\frac{|P|}{|J|} \end{pmatrix}$$

where $|P|$ and $|K|$ are the determinant of the following matrices

$$K = \begin{bmatrix} \frac{\partial \dot{c}}{\partial \alpha} & \frac{\partial \dot{c}}{\partial w} \\ \\ \frac{\partial \dot{w}}{\partial \alpha} & \frac{\partial \dot{w}}{\partial w} \end{bmatrix} \qquad P = \begin{bmatrix} \frac{\partial \dot{c}}{\partial \alpha} & \frac{\partial \dot{c}}{\partial c} \\ \\ \frac{\partial \dot{w}}{\partial \alpha} & \frac{\partial \dot{w}}{\partial c} \end{bmatrix}$$

A.4.1 The example with explicit functions

Threshold and history dependence are possible in a perfectly governed economy (i.e. strictly concave). The first term of the utility function specifies that $c^{sat} = 1/\alpha$ where α captures individual characteristics (e.g. income., gender, age) which affects food satiation; the second term represents the consequence of begin either overweight ($w > w^H$) or underweight ($w < w^H$) with $\beta > 0$ representing the scaling parameter of the (quadratic) appreciation of a healthy body weight. (12) is determining the following canonical equations

$$\begin{aligned}\dot{c} &= \frac{(\alpha c - 1)(\rho - (w - \bar{w}))}{\alpha} + \beta\frac{(w - w^H)}{\alpha} \\ \dot{w} &= c - w\left(\bar{w} - \frac{w}{2}\right)\end{aligned} \tag{A.10}$$

From which we can derive the $\dot{c} = 0$-isocline and the $\dot{w} = 0$-isocline, which are respectively

$$\begin{aligned}\dot{c} = 0 \qquad & c = -\beta\frac{(w - w^H)}{a[\rho - (w - \bar{w})]} + \frac{1}{a} \\ \dot{w} = 0 \qquad & c = w\left(\bar{w} - \frac{w}{2}\right) = g(w)\end{aligned} \tag{A.11}$$

From which we can verify the arrows of motion of Fig.2. For the $\dot{w} = 0$-isocline we have $\dot{w} < 0$ for $c < g(w)$ and $\dot{w} > 0$ otherwise; while for the $\dot{c} = 0$-isocline we have $\dot{c} < 0$ for $c > -\beta\frac{(w-w^H)}{a[\rho-(w-\bar{w})]} + \frac{1}{a}$ and $\dot{c} > 0$ otherwise.

The system (A.11) has up to three steady states since substituting the stationary solution for $\dot{c} = 0$ into $\dot{w} = 0$ results in a cubic equation in w. The Jacobian as well as the determinant simplify to

$$j = \begin{pmatrix} \rho + g_w & \dfrac{\beta - (\alpha c - 1)}{\alpha} \\ & \\ 1 & -g_w \end{pmatrix} \tag{A.12}$$

$$\det(j) = -g_w(\rho + g_w) + \frac{(\alpha c - 1) - \beta}{\alpha}\frac{\beta}{\alpha} - cg''(w) + \frac{g''(w)}{\alpha} - \rho g'(w) - g'(w)^2 \tag{A.13}$$

The $\dot{w} = 0$, $c = w(\bar{w} - w/2)$ is determined by the metabolic level of caloric consumption, $c = g(w)$ with a maximum in $\bar{w}$. The $\dot{c} = 0$ hits this function in three points corresponding to three steady state characterized by an incremental level of steady state body weight.

Bibliography

Akerlof, G. A. and R. E. Kranton (2000). Economics and identity. *The Quarterly Journal of Economics 115*(3), 715–753.

Ali, M. M., A. Amialchuk, S. Gao, and F. Heiland (2012). Adolescent weight gain and social networks: is there a contagion effect? *Applied Economics 44*(23), 2969–2983.

Ali, M. M., A. Amialchuk, and F. Renna (2011). Social network and weight misperception among adolescents. *Southern Economic Journal 77*(4), 827–842.

Allison, D. B., S. Zhu, M. Plankey, M. S. Faith, and M. Heo (2002). Differential associations of body mass index and adiposity with all-cause mortality among men in the first and second national health and nutrition examination surveys (nhanes i and nhanes ii) follow-up studies. *International Journal of Obesity & Related Metabolic Disorders 26*(3).

Arena, V. C., K. R. Padiyar, W. N. Burton, and J. J. Schwerha (2006). The impact of body mass index on short-term disability in the workplace. *Journal of occupational and environmental medicine 48*(11), 1118–1124.

Augurzky, B., T. K. Bauer, A. Reichert, C. Schmidt, and H. Tauchmann (2012). Does money burn fat? evidence from a randomized experiment.

Baradel, L., C. Gillespie, J. Kicklighter, M. Doucette, M. Penumetcha, H. Blanck, et al. (2009). Temporal changes in trying to lose weight and recommended weight-loss strategies among overweight and obese americans, 1996-2003. *Preventive medicine 49*(2-3), 158.

Bazzocchi, A., D. Diano, F. Ponti, A. Andreone, C. Sassi, U. Albisinni, G. Marchesini, and G. Battista (2012). Health and ageing: A cross-sectional study of body composition. *Clinical Nutrition.*

Bhattacharya, J., K. Choudhry, and D. Lakdawalla (2008). Chronic disease and severe disability among working-age populations. *Medical care 46*(1), 92–100.

Bish, C., H. Blanck, M. Serdula, M. Marcus, H. Kohl, and L. Khan (2012). Diet and physical activity behaviors among americans trying to lose weight: 2000 behavioral risk factor surveillance system. *Obesity Research 13*(3), 596–607.

Blanchflower, D. G., B. Landeghem, and A. J. Oswald (2009). Imitative obesity and relative utility. *Journal of the European Economic Association 7*(2-3), 528–538.

Brunello, G. and B. d'Hombers (2007). Does body weight affect wages?: Evidence from europe. *Economics & Human Biology 5*(1), 1–19.

Burke, M., F. Heiland, and C. Nadler (2009). From "overweight" to "about right": evidence of a generational shift in body weight norms. *Obesity 18*(6), 1226–1234.

Burke, M. A. and F. Heiland (2007). Social dynamics of obesity. *Economic Inquiry 45*(3), 571–591.

Burkhauser, R. V. and J. Cawley (2008). Beyond bmi: The value of more accurate measures of fatness and obesity in social science research. *Journal of Health Economics 27*(2), 519–529.

Caliendo, M. and M. Gehrsitz (2014, February). Obesity and the labor market: A fresh look at the weight penalty. *IZA Dp No. 7947*.

Caliendo, M. and W.-S. Lee (2013). Fat chance! obesity and the transition from unemployment to employment. *Economics & Human Biology 11*(2), 121–133.

Catenacci, V., P. Maclean, L. Ogden, S. Polsky, H. Wyatt, and J. Hill (2011). *Chapter 46 in: John Cawley (editor), The Oxford Handbook of the Social Science of Obesity,*. Oxford University Press, USA.

Cawley, J. (2000). An instrumental variables approach to measuring the effect of body weight on employment disability. *Health Services Research 35*(5 Pt 2), 1159.

Cawley, J. (2004). The impact of obesity on wages. *Journal of Human Resources 39*(2), 451–474.

Cawley, J. and S. Danziger (2005). Morbid obesity and the transition from welfare to work. *Journal of Policy Analysis and Management 24*(4), 727–743.

Cawley, J. and J. Price (2012). Financial incentives for weight loss: Results from a workplace wellness program.

Christakis, N. A. and J. H. Fowler (2007). The spread of obesity in a large social network over 32 years. *New England journal of medicine 357*(4), 370–379.

Christian, T. and I. Rashad (2009). Trends in us food prices, 1950–2007. *Economics & Human Biology 7*(1), 113–120.

Costa-Font, J. and M. Jofre-Bonet (2013). Anorexia, body image and peer effects: evidence from a sample of european women. *Economica 80*(317), 44–64.

Cunningham, J. J. (1991). Body composition as a determinant of energy expenditure: a synthetic review and a proposed general prediction equation. *The American journal of clinical nutrition 54*(6), 963–969.

Cutler, D., E. Glaeser, and J. Shapiro (2003). Why have americans become more obese? Technical report, National Bureau of Economic Research.

Cutler, D. M. and A. Lleras-Muney (2010). Understanding differences in health behaviors by education. *Journal of health economics 29*(1), 1–28.

Dodd, M. (2008). Obesity and time-inconsistent preferences. *Obesity Research & Clinical Practice 2*(2), 83–89.

Dragone, D. (2009). A rational eating model of binges, diets and obesity. *Journal of health economics 28*(4), 799–804.

Dragone, D., F. Manaresi, and L. Savorelli (2013). Obesity and smoking: can we catch two birds with one tax?

Dragone, D. and L. Savorelli (2011). Thinness and obesity: A model of food consumption, health concerns, and social pressure. *Journal of Health Economics*.

Dragone, D. and L. Savorelli (2012). Thinness and obesity: A model of food consumption, health concerns, and social pressure. *Journal of Health Economics 31*(1), 243–256.

Ebbeling, C., J. Swain, H. Feldman, W. Wong, D. Hachey, E. Garcia-Lago, and D. Ludwig (2012). Effects of dietary composition on energy expenditure during weight-loss maintenance. *JAMA: The Journal of the American Medical Association 307*(24), 2627–2634.

Etilé, F. (2007). Social norms, ideal body weight and food attitudes. *Health Economics 16*(9), 945–966.

Fairlie, R. W. (1999). The absence of the african-american owned business: An analysis of the dynamics of self-employment. *Journal of Labor Economics 17*(1), 80–108.

Fairlie, R. W. (2005). An extension of the blinder-oaxaca decomposition technique to logit and probit models. *Journal of economic and social measurement 30*(4), 305–316.

Finkelstein, E., C. Ruhm, and K. Kosa (2005). Economic causes and consequences of obesity. *Annu. Rev. Public Health 26*, 239–257.

Font, J. C., D. Fabbri, and J. Gil (2010). Decomposing cross-country differences in levels of obesity and overweight: does the social environment matter? *Social Science & Medicine 70*(8), 1185–1193.

Gale, S., V. Castracane, and C. Mantzoros (2004). Energy homeostasis, obesity and eating disorders: recent advances in endocrinology. *The Journal of nutrition 134*(2), 295–298.

García Villar, J. and C. Quintana-Domeque (2009). Income and body mass index in europe. *Economics & Human Biology 7*(1), 73–83.

Gil, J. and T. Mora (2011). The determinants of misreporting weight and height: the role of social norms. *Economics & Human Biology 9*(1), 78–91.

Goldman, D., D. Lakdawalla, and Y. Zheng (2011). *Food prices and the dynamics of body weight*. University of Chicago Press.

Halliday, T. J. and S. Kwak (2009). Weight gain in adolescents and their peers. *Economics & Human Biology 7*(2), 181–190.

Han, E., E. C. Norton, and S. C. Stearns (2009). Weight and wages: fat versus lean paychecks. *Health economics 18*(5), 535–548.

Harris, J. A. and F. G. Benedict (1918). A biometric study of human basal metabolism. *Proceedings of the National Academy of Sciences of the United States of America 4*(12), 370.

Hastie, T. J. and R. J. Tibshirani (1990). *Generalized additive models*, Volume 43. CRC Press.

Heitmann, B., H. Erikson, B. Ellsinger, K. Mikkelsen, and B. Larsson (2000). Mortality associated with body fat, fat-free mass and body mass index among 60 year old swedish men 22-year follow-up. the study of men born in 1913. *International Journal of Obesity & Related Metabolic Disorders 24*(1).

Huffman, W. E., S. Huffman, K. Rickertsen, and A. Tegene (2010). Over-nutrition and changing health status in high income countries. *Frontier in Health Policy Research 13*(1), 2.

Ikeda, S., M. Kang, and F. Ohtake (2010). Hyperbolic discounting, the sign effect, and the body mass index. *Journal of health economics 29*(2), 268–284.

John, L., G. Loewenstein, A. Troxel, L. Norton, J. Fassbender, and K. Volpp (2011). Financial incentives for extended weight loss: a randomized, controlled trial. *Journal of general internal medicine 26*(6), 621–626.

Johnson, W. G., R. Stewart, and A. T. Pusser (2012). The perceptual threshold for overweight. *Eating behaviors 13*(3), 188–193.

Johnson-Taylor, W., R. Fisher, V. Hubbard, P. Starke-Reed, and P. Eggers (2008). The change in weight perception of weight status among the overweight: comparison of nhanes iii (1988–1994) and 1999–2004 nhanes. *International Journal of Behavioral Nutrition and Physical Activity 5*(1), 9.

Johnstone, A. M., S. D. Murison, J. S. Duncan, K. A. Rance, and J. R. Speakman (2005). Factors influencing variation in basal metabolic rate include fat-free mass, fat mass, age, and circulating thyroxine but not sex, circulating leptin, or triiodothyronine. *The American journal of clinical nutrition 82*(5), 941–948.

Kahneman, D. and A. Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, 263–291.

Katan, M. and D. Ludwig (2010). Extra calories cause weight gain - but how much? *JAMA: the journal of the American Medical Association 303*(1), 65–66.

Keele, L. J. (2008). *Semiparametric regression for the social sciences*. John Wiley & Sons.

Klarenbach, S., R. Padwal, A. Chuck, and P. Jacobs (2006). Population-based analysis of obesity and workforce participation. *Obesity 14*(5), 920–927.

Kruger, J., D. Galuska, M. Serdula, D. Jones, et al. (2004). Attempting to lose weight: specific practices among us adults. *American journal of preventive medicine 26*(5), 402.

Lakdawalla, D. and T. Philipson (2009). The growth of obesity and technological change. *Economics & Human Biology 7*(3), 283–293.

Lakdawalla, D. N., T. J. Philipson, and J. Bhattacharya (2006). Welfare-enhancing technological change and the growth of obesity. In *American Economic Review: AEA Papers and Proceedings*, Volume 95, pp. 253–257.

Leibel, R. and J. Hirsch (1984). Diminished energy requirements in reduced-obese patients. *Metabolism 33*(2), 164–170.

Leibel, R., M. Rosenbaum, and J. Hirsch (1995). Changes in energy expenditure resulting from altered body weight. *New England Journal of Medicine 332*(10), 621–628.

Levy, A. (2002). Rational eating: can it lead to overweightness or underweightness? *Journal of Health Economics 21*(5), 887–899.

Lindeboom, M., P. Lundborg, and B. van der Klaauw (2010). Assessing the impact of obesity on labor market outcomes. *Economics & Human Biology 8*(3), 309–319.

Loader, C. R. (1996). Local likelihood density estimation. *The Annals of Statistics*, 1602–1618.

Loewenstein, G., T. Brennan, and K. Volpp (2007). Asymmetric paternalism to improve health behaviors. *JAMA: The Journal of the American Medical Association 298*(20), 2415–2417.

Loewenstein, G. and T. O'Donoghue (2004). Animal spirits: Affective and deliberative processes in economic behavior. *Available at SSRN 539843*.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The review of economic studies 60*(3), 531–542.

Maximova, K., J. McGrath, T. Barnett, J. O'Loughlin, G. Paradis, and M. Lambert (2008). Do you see what i see? weight status misperception and exposure to obesity among children and adolescents. *International Journal of Obesity 32*(6), 1008–1015.

McFerran, B., D. Dahl, G. Fitzsimons, and A. Morales (2010a). Ill have what shes having: Effects of social influence and body type on the food choices of others. *Journal of Consumer Research 36*(6), 915– 929.

McFerran, B., D. Dahl, G. Fitzsimons, and A. Morales (2010b). Might an overweight waitress make you eat more? how the body type of others is sufficient to alter our food consumption. *Journal of Consumer Psychology 20*(2), 146.

Mifflin, M. D., S. St Jeor, L. A. Hill, B. J. Scott, S. A. Daugherty, and Y. Koh (1990). A new predictive equation for resting energy expenditure in healthy individuals. *The American journal of clinical nutrition 51*(2), 241–247.

Mocan, N. and E. Tekin (2011). oobesity, self&esteem, and wages, pin m. grossman and n. mocan (eds.) economic aspects of obesity.

Mora, T. and J. Gil (2012). Peer effects in adolescent bmi: Evidence from spain. *Health Economics*.

Morris, S. (2007). The impact of obesity on employment. *Labour Economics 14*(3), 413–433.

Mosca, I. (2013). Body mass index, waist circumference and employment: Evidence from older irish adults. *Economics & Human Biology*.

Mozaffarian, D., T. Hao, E. B. Rimm, W. C. Willett, and F. B. Hu (2011). Changes in diet and lifestyle and long-term weight gain in women and men. *New England Journal of Medicine 364*(25), 2392–2404.

Oaxaca, R. L. and M. R. Ransom (1994). On discrimination and the decomposition of wage differentials. *Journal of econometrics 61*(1), 5–21.

O'Donoghue, T. and M. Rabin (1999). Doing it now or later. *American Economic Review*, 103–124.

Ogden, C. and M. Carroll (2010). Prevalence of overweight, obesity, and extreme obesity among adults: United states, trends 1960–1962 through 2007–2008. *NCHS Health and Stats*.

Oswald, A. J. and N. Powdthavee (2007). Obesity, unhappiness, and the challenge of affluence: Theory and evidence.

Ozdenoren, E., S. Salant, and D. Silverman (2006). Willpower and the optimal control of visceral urges. Technical report, National Bureau of Economic Research.

Paeratakul, S., M. A. White, D. A. Williamson, D. H. Ryan, and G. A. Bray (2002). Sex, race/ethnicity, socioeconomic status, and bmi in relation to self-perception of overweight. *Obesity research 10*(5), 345–350.

Philipson, T. J. and R. A. Posner (2003). The long-run growth in obesity as a function of technological change. *Perspectives in biology and medicine 46*(3), S87–S107.

Philipson, T. J. and R. A. Posner (2008). Is the obesity epidemic a public health problem? a decade of research on the economics of obesity. *Journal of Economic Literature 46*(4), 974–82.

Popkin, B. M. (2007). The world is fat. *Scientific American 297*(3), 88–95.

Rand, C. S. and J. L. Resnick (2000). The good enough body size as judged by people of varying age and weight. *Obesity Research 8*(4), 309–316.

Reich, B. and J. W. Weibull (2012). Obesity as a social equilibrium phenomenon.

Rooth, D.-O. (2007). Evidence of unequal treatment in hiring against obese applicants: a field experiment.

Rosin, O. (2008). The economic causes of obesity: a survey. *Journal of Economic Surveys 22*(4), 617–647.

Ruhm, C. J. (2012). Understanding overeating and obesity. *Journal of Health Economics*.

Sassi, F. and M. Devaux (2012). Obesity update 2012. *OECD 2012; 1-7*.

Stearns, P. N. (2002). *Fat history: Bodies and beauty in the modern west.* NYU Press.

Strulik, H. (2014). A mass phenomenon: The social evolution of obesity. *Journal of Health Economics 33*(0), 113 – 125.

Sunder, M. and F. Kropfhäußer (2013). A weighty issue revisited: the dynamic effect of body weight on earnings and satisfaction in germany.

Taylor, P., C. Funk, and P. Craighill (2006). Americans see weight problems everywhere but in the mirror. *Philadelphia: Pew Foundation Social Trends Report*.

Trogdon, J. G. and B. T. Allaire (2014). The effect of friend selection on social influences in obesity. *Economics & Human Biology*.

Trogdon, J. G., J. Nonnemaker, and J. Pais (2008). Peer effects in adolescent overweight. *Journal of health economics 27*(5), 1388–1399.

Tunceli, K., K. Li, and L. Williams (2006). Long-term effects of obesity on employment and work limitations among us adults, 1986 to 1999. *Obesity 14*(9), 1637–1646.

Volpp, K., L. John, A. Troxel, L. Norton, J. Fassbender, and G. Loewenstein (2008). Financial incentive–based approaches for weight loss. *JAMA: the journal of the American Medical Association 300*(22), 2631–2637.

von Ruesten, A., A. Steffen, A. Floegel, G. Masala, A. Tjønneland, J. Halkjaer, D. Palli, N. Wareham, R. Loos, T. Sørensen, et al. (2011). Trend in obesity prevalence in european adult cohort populations during follow-up since 1996 and their predictions to 2015. *PLoS One 6*(11), e27455.

Wada, R. and E. Tekin (2010). Body composition and wages. *Economics & Human Biology 8*(2), 242–254.

Wirl, F. (2004). Thresholds in concave renewable resource models. *Ecological economics 48*(2), 259–267.

Wirl, F. and G. Feichtinger (2005). History dependence in concave economies. *Journal of Economic Behavior & Organization 57*(4), 390–407.

Wirl, F. and G. Feichtinger (2010). Modelling social dynamics (of obesity) and thresholds. *Games 1*(4), 395–414.

Wood, S. (2006). *Generalized additive models: an introduction with R.* CRC press.

Yaari, M. E. (1965). Uncertain lifetime, life insurance, and the theory of the consumer. *The Review of Economic Studies 32*(2), 137–150.

FIGURE A.1: The law of motion of body weight for $\bar{w} = 0,82$.

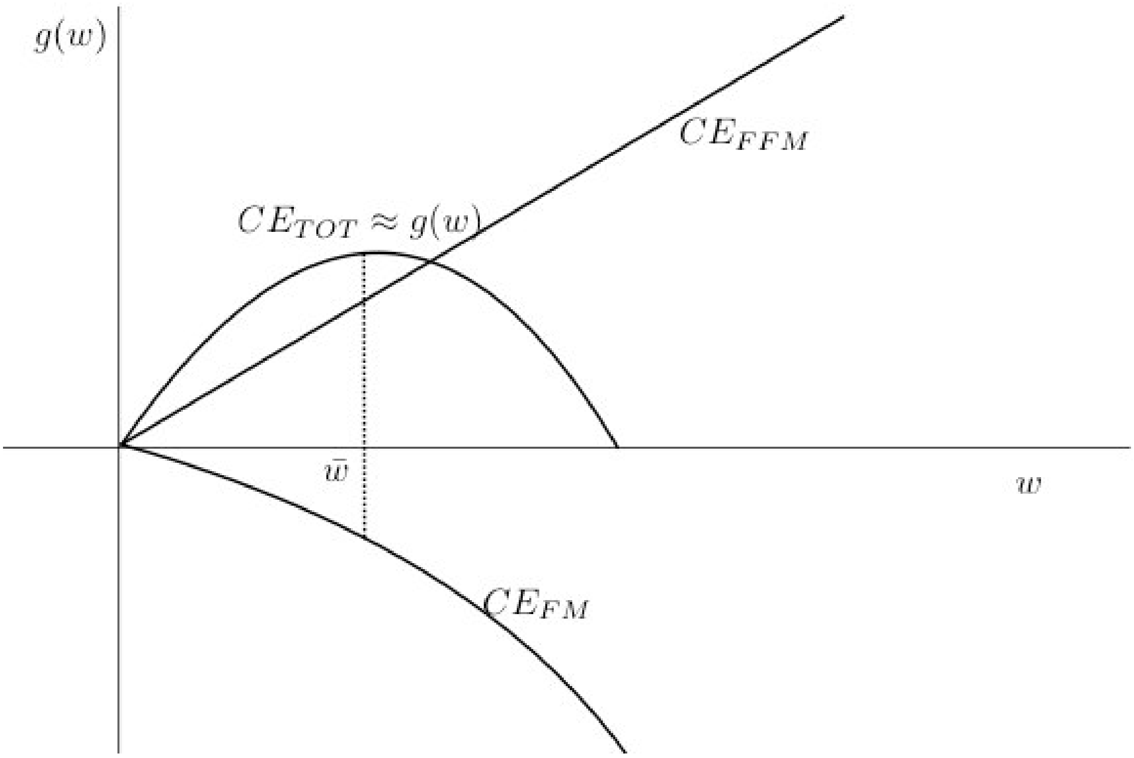


FIGURE A.2: Phase diagram of (15) for the parameters values: $a = 3.9$, $\beta = 0.05$, $\bar{w} = 0.82$, $w^H = 0.44$ and $\rho = 0.8$

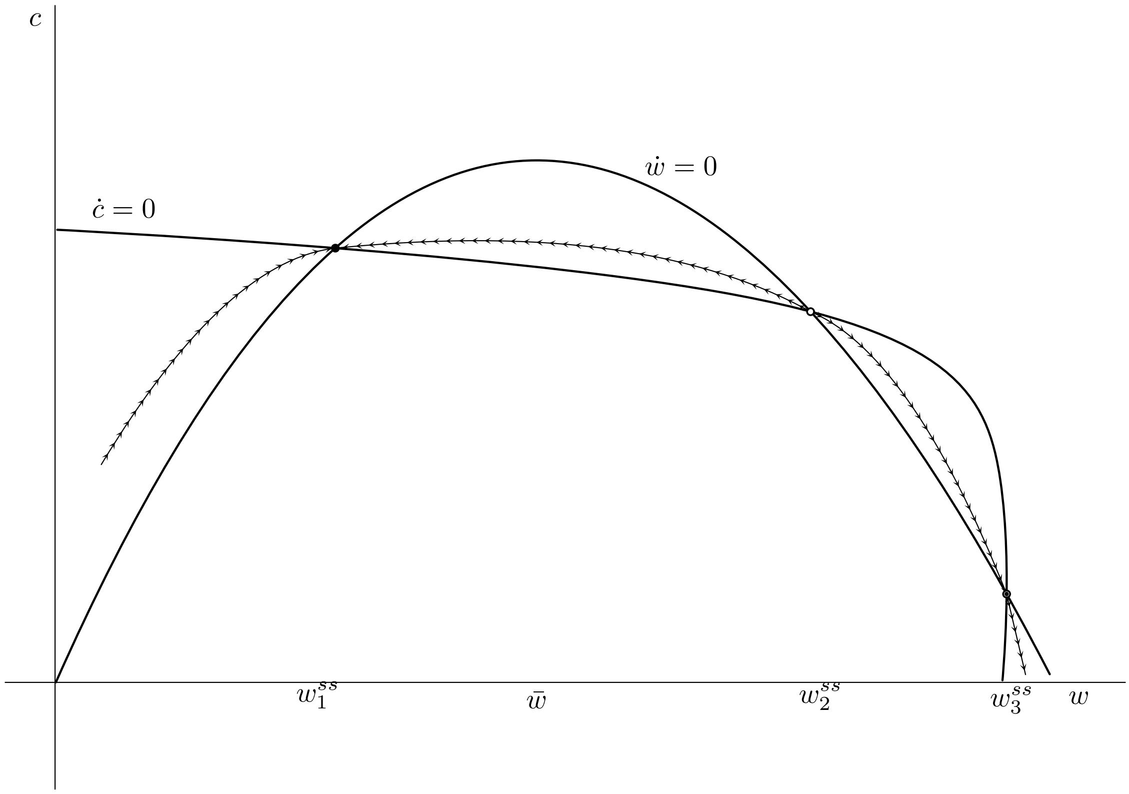


FIGURE A.3: Phase diagram of (16) after the introduction of the incentive for the parameters values: $a = 3.9$, $\beta = 0.05$, $\bar{w} = 0.82$, $w^H = 0.44$, $\alpha = 0.4$ and $\rho = 0.8$

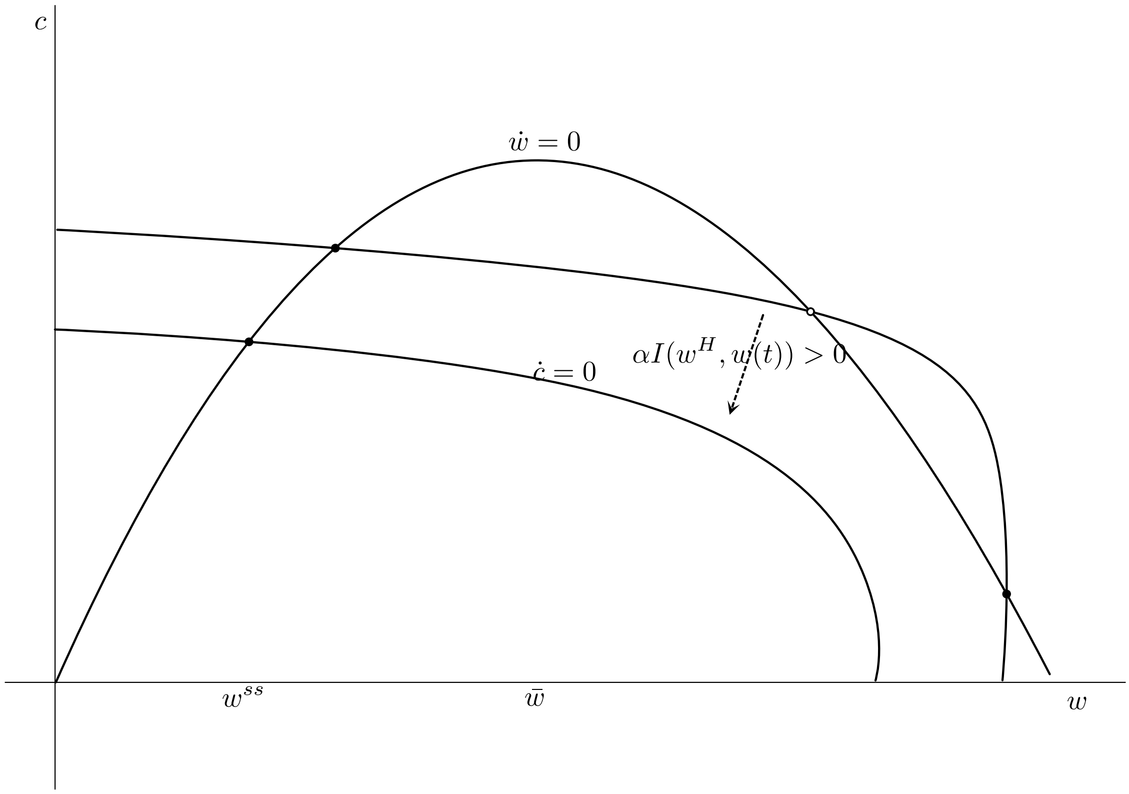


FIGURE A.4: Comparative Statics, pre- and post-incentive.

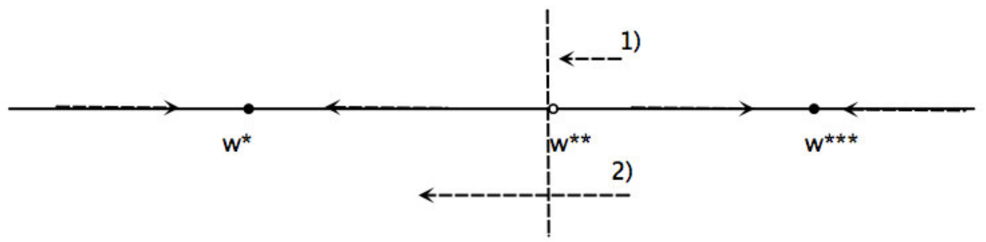
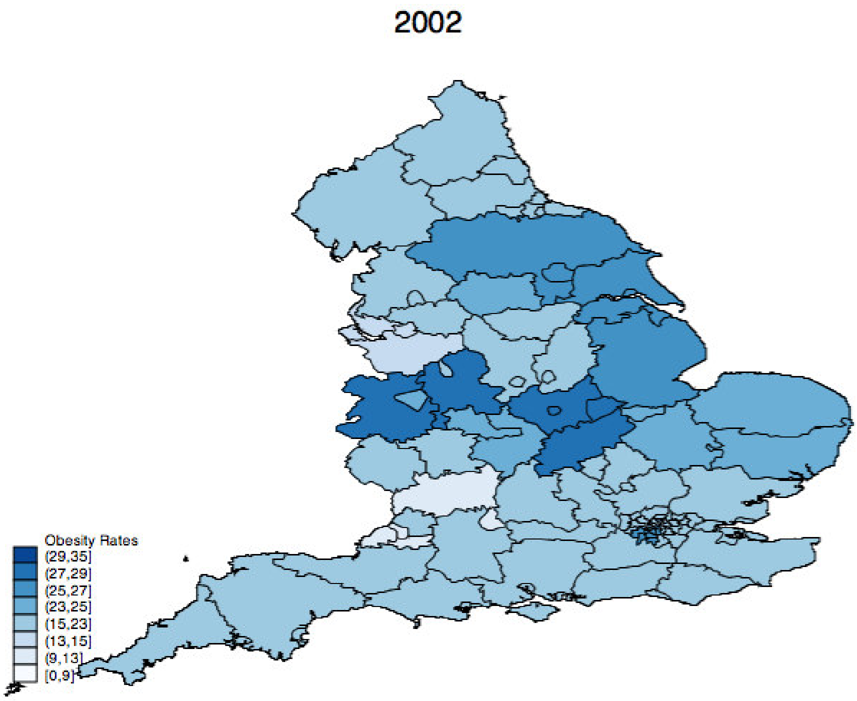


FIGURE A.5: Area Incidence of Obesity in England 2002-2006



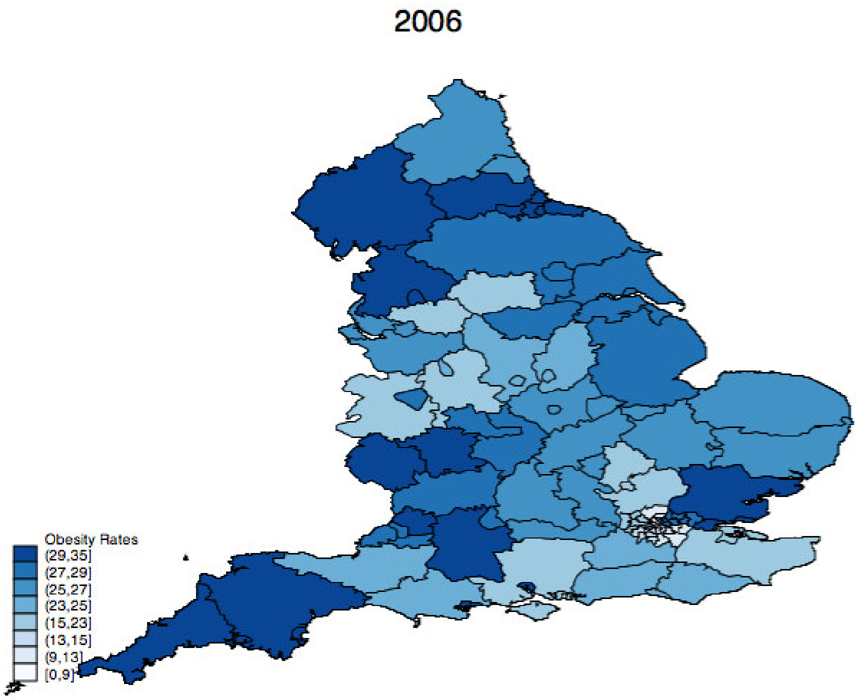
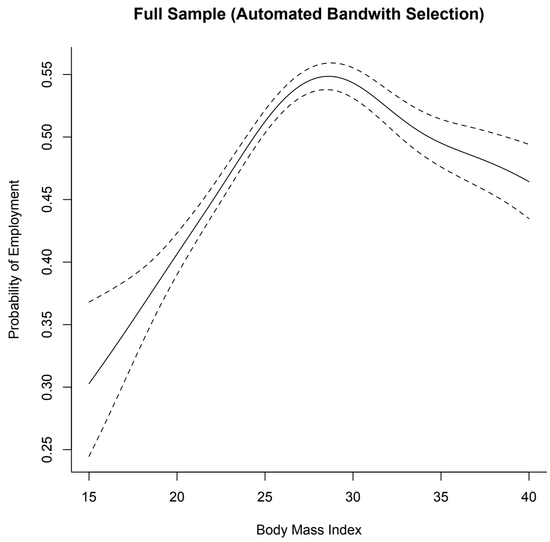
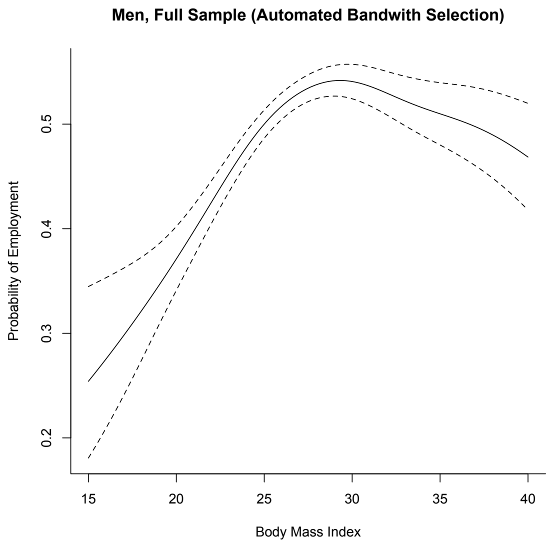


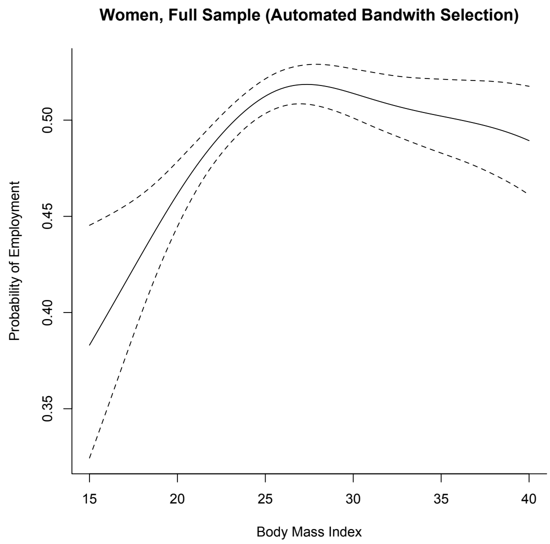
FIGURE A.6: BMI and Employment, by gender - Semi-Parametric



(a) BMI and Employment



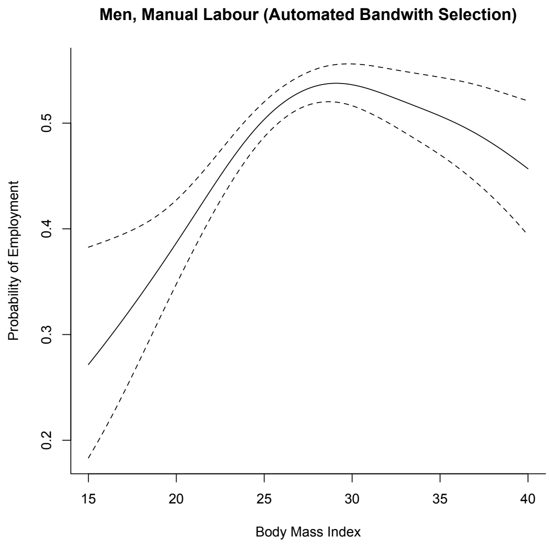
(b) BMI and Employment



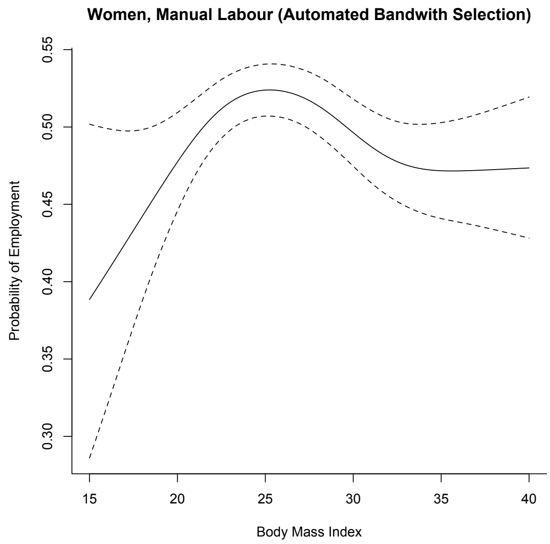
(c) BMI and Employment

Note: Graphs display the results of a semiparametric regression of an employment dummy on body mass index and a full set of control variables.
Solid line shows the effect of BMI on employment probabilities, stratified by gender. Dashed lines are 95 percent confidence bands. The semiparametric results, centered in zero, were transformed into probability scale for an easier interpretation. Smoothing parameters were obtained using the automated cross-validation algorithm implemented in R's mgcv library.

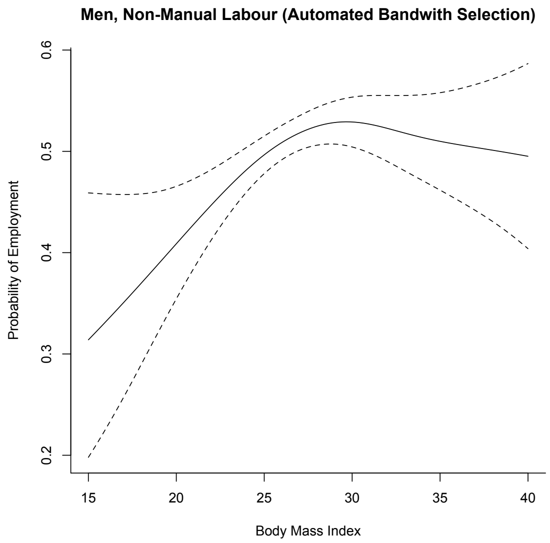
Figure A.7: BMI and Employment, by type of labour - Semi-Parametric



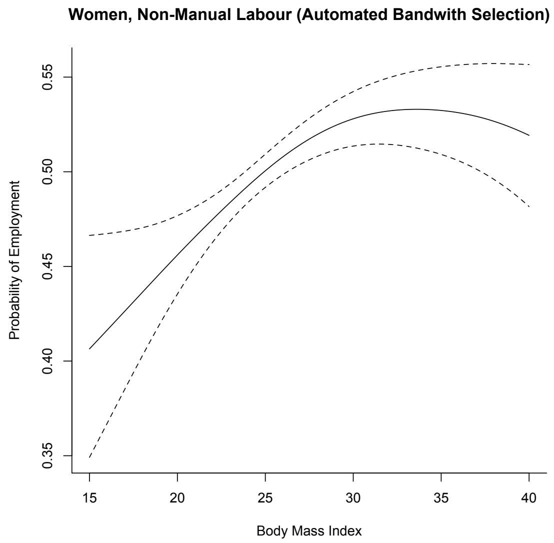
(a) BMI and Employment



(b) BMI and Employment



(c) BMI and Employment

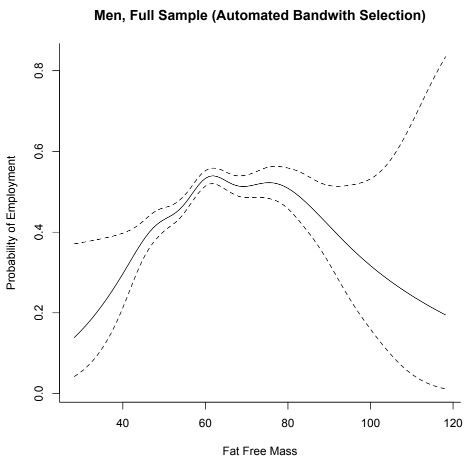


(d) BMI and Employment

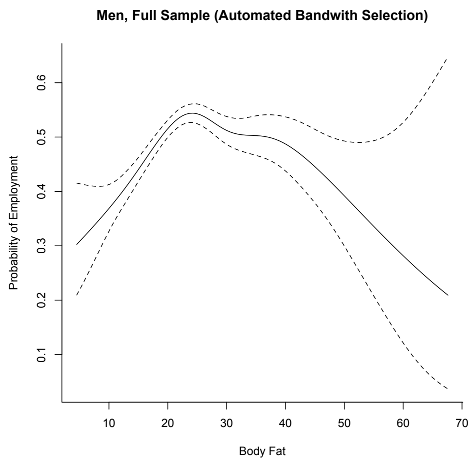
Note: Graphs display the results of a semiparametric regression of an employment dummy on body mass index and a full set of control variables.

Solid line shows the effect of BMI on employment probabilities, stratified by types of work. Dashed lines are 95 percent confidence bands. The semiparametric results, centered in zero, were transformed into probability scale for an easier interpretation. Smoothing parameters were obtained using the automated cross-validation algorithm implemented in R's mgcv library.

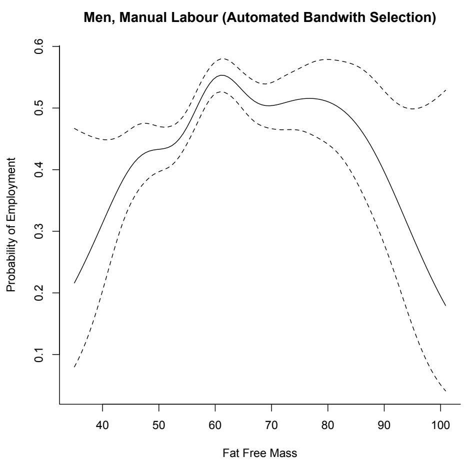
Figure A.8: Men, Fat Free Mass, Body Fat and Employment - Semi-Parametric



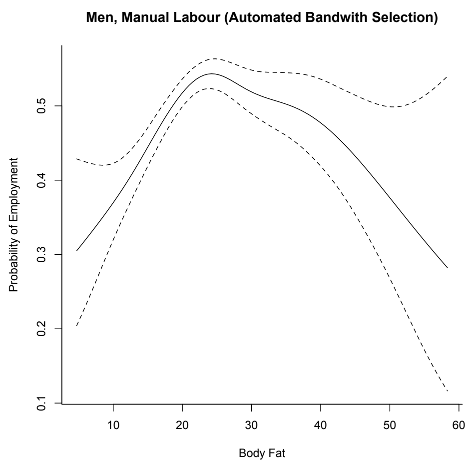
(a) FFM and Employment



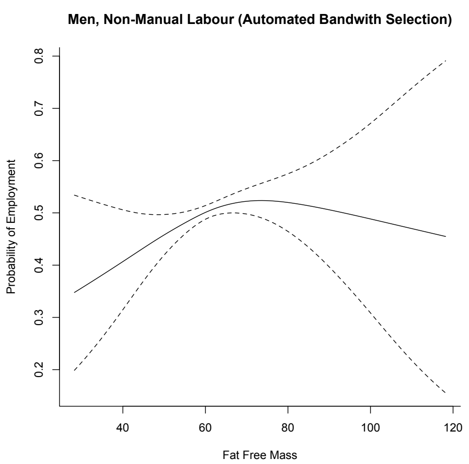
(b) BF and Employment



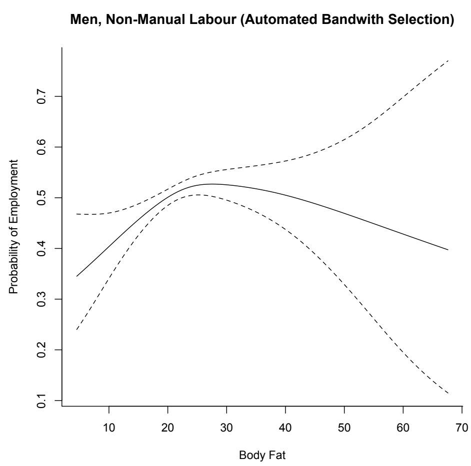
(c) FFM and Employment



(d) BF and Employment



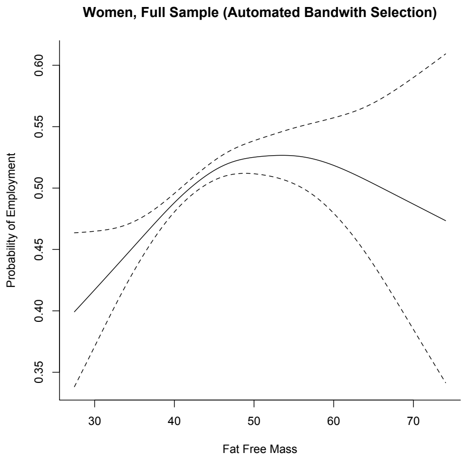
(e) FFM and Employment



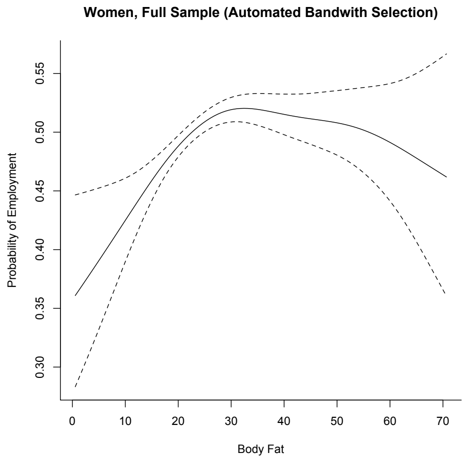
(f) BF and Employment

Note: Graphs display the results of a semiparametric regression of an employment dummy on body mass index and a full set of control variables.
Solid lines show the effect of fat free mass (FFM) and body fat (BF), respectively, on employment for men, stratified by types of work. Dashed lines are 95 percent confidence bands. The semiparametric results, centered in zero, were transformed into probability scale for an easier interpretation. Smoothing parameters were obtained using the automated cross-validation algorithm implemented in R's mgcv library.

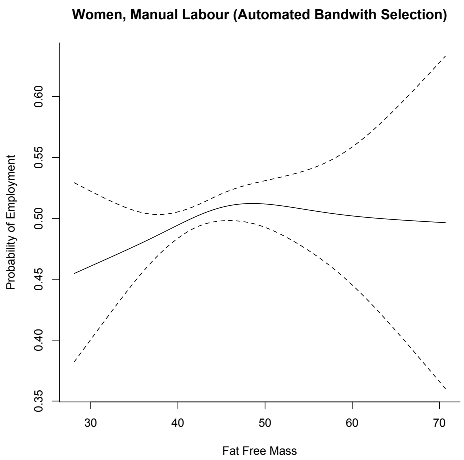
Figure A.9: Women, Fat Free Mass, Body Fat and Employment - Semi-Parametric



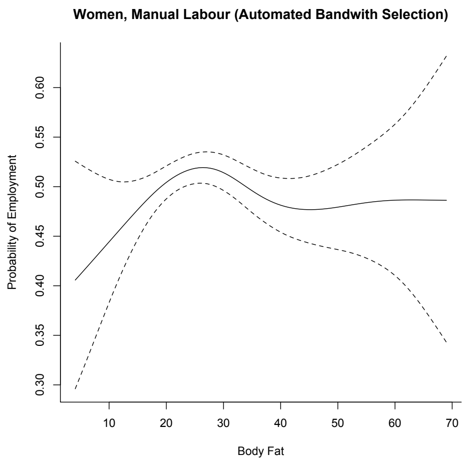
(a) FFM and Employment



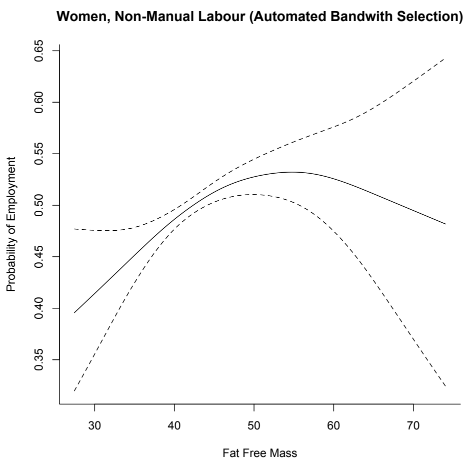
(b) BF and Employment



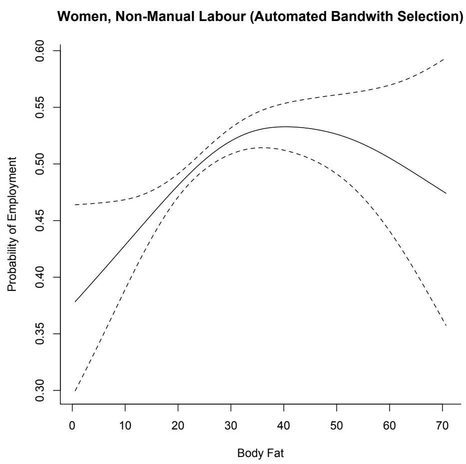
(c) FFM and Employment



(d) BF and Employment



(e) FFM and Employment



(f) BF and Employment

Note: Graphs display the results of a semiparametric regression of an employment dummy on body mass index and a full set of control variables.
Solid lines show the effect of fat free mass (FFM) and body fat (BF), respectively, on employment for women, stratified by types of work. Dashed lines are 95 percent confidence bands. The semiparametric results, centered in zero, were transformed into probability scale for an easier interpretation. Smoothing parameters were obtained using the automated cross-validation algorithm implemented in R's mgcv library.

Table A.1: Summary Statistics

	(1) All mean	(2) Healthy mean	(3) Overweight mean	(4) Obese mean
employed	.7748081	.7560348	.8211641	.7623977
male	.4641116	.3957453	.5391814	.4640507
age last birthday	40.33567	36.94468	42.43423	43.87618
self-assessed general health	1.881972	1.754542	1.850936	2.1446
marital	.830317	.8269464	.8425545	.8359011
Education	12.6586	12.6589	12.88492	12.63401
rural	.1100278	.1085531	.115796	.1077676
BMI	26.83866	22.4693	27.27556	34.15576
Drinking	.915298	.916215	.9223644	.9055686
N. of Cigarettes (daily)	3.646654	4.035484	3.322479	3.392224
Eating Habits	3.50379	3.481657	3.590429	3.453392
Social Norm (BMI)	26.85911	26.46513	27.14844	27.25364
Average BMI in HA	26.88817	26.87436	26.88893	26.90935
Perc. of Ob. in HA	.2343358	.2333838	.2342294	.235995
Obese	.2327749			
Overweight	.3632357			
Healthy	.3845464			
Observations	53541	20589	19448	12462

Table A.2: Logit Results (2002-2006)

	Obesity (ME)					
	(1) Male	(2) Male	(3) Male	(4) Female	(5) Female	(6) Female
Average BMI in HA	0.0397* (0.0156)			0.0483* (0.0189)		
Social Norm (BMI)		0.0252 (0.0167)			0.0429* (0.0179)	
Perc. of Ob. in HA			0.0493*** (0.00776)			0.0376*** (0.00910)
Ed. Controls	Yes	Yes	Yes	Yes	Yes	Yes
Age	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity	Yes	Yes	Yes	Yes	Yes	Yes
Health Controls	Yes	Yes	Yes	Yes	Yes	Yes
Family composition	Yes	Yes	Yes	Yes	Yes	Yes
Eating Habits	Yes	Yes	Yes	Yes	Yes	Yes
Food Prices	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes
Observations	9622	9622	9622	11226	11226	11226

Marginal effects; Standard errors in parentheses

(d) for discrete change of dummy variable from 0 to 1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.3: Fairlie Decomposition 2002-2006

	Obesity		
	(1) Fairlie	(2) Fairlie	(3) Fairlie
Demographic	-0.0237***	-0.00194	-0.0244***
	(0.00326)	(0.00474)	(0.00332)
SES	0.00380	0.000965	0.00196
	(0.00341)	(0.00310)	(0.00347)
Education	0.00210***	0.000122	0.00207***
	(0.000526)	(0.000463)	(0.000540)
Health	-0.000225	-0.00390***	-0.000692
	(0.000722)	(0.00105)	(0.000710)
Drinking	-0.000479*	-0.00106**	-0.000700*
	(0.000231)	(0.000403)	(0.000293)
N. of Cigarettes (daily)	-0.000258	-0.000451*	-0.000436
	(0.000215)	(0.000226)	(0.000227)
Eating Habits	-0.0000889	-0.000152	-0.000234
	(0.00132)	(0.00134)	(0.00143)
Social Norm (BMI)		-0.0152***	
		(0.00278)	
Average BMI in HA		-0.00480	
		(0.00334)	
Perc. of Ob. in HA			-0.0124**
			(0.00379)
Observations	13521	13521	13521
Food Prices	Yes	Yes	Yes
Obesity (Early)	0.202	0.202	0.202
Obesity (Late)	0.247	0.247	0.247
Difference	-0.0451	-0.0451	-0.0451
Tot Expl. Difference	-0.0189	-0.0264	-0.0348

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.4: Fairlie Decomposition 2002-2006, by Gender

	Obesity					
	(1) Men	(2) Women	(3) Men	(4) Women	(5) Men	(6) Women
Demographic	-0.0284***	-0.0211***	-0.00529	-0.000775	-0.0306***	-0.0213***
	(0.00480)	(0.00454)	(0.00802)	(0.00636)	(0.00505)	(0.00456)
SES	0.00875	0.00362	0.00409	0.000877	0.00533	0.00283
	(0.00536)	(0.00461)	(0.00482)	(0.00427)	(0.00544)	(0.00467)
Education	0.00187*	0.00203**	0.000208	0.000707	0.00151	0.00219**
	(0.000871)	(0.000667)	(0.000531)	(0.000552)	(0.000814)	(0.000702)
Health	-0.00233*	0.00322**	-0.00391*	-0.00112	-0.00242*	0.00305**
	(0.00114)	(0.00120)	(0.00161)	(0.00159)	(0.00119)	(0.00117)
Drinking	-0.000274	-0.000577	-0.00139*	-0.000807	-0.000841	-0.000626
	(0.000297)	(0.000372)	(0.000702)	(0.000487)	(0.000480)	(0.000397)
N. of Cigarettes (daily)	0.000432	-0.000756	-0.00116	-0.000370	-0.000257	-0.000824*
	(0.000468)	(0.000386)	(0.000598)	(0.000259)	(0.000417)	(0.000418)
Eating Habits	-0.00121	0.000515	-0.00160	0.000488	-0.00183	0.000498
	(0.00186)	(0.00184)	(0.00209)	(0.00179)	(0.00222)	(0.00192)
Social Norm (BMI)			-0.0187***	-0.0126***		
			(0.00539)	(0.00324)		
Average BMI in HA			-0.0124**	0.00231		
			(0.00446)	(0.00492)		
Perc. of Ob. in HA					-0.0205***	-0.00535
					(0.00528)	(0.00534)
Observations	6295	7226	6295	7226	6295	7226
Food Prices	Yes	Yes	Yes	Yes	Yes	Yes
Obesity (Early)	0.194	0.209	0.194	0.209	0.194	0.209
Obesity (Late)	0.254	0.241	0.254	0.241	0.254	0.241
Difference	-0.0600	-0.0321	-0.0600	-0.0321	-0.0600	-0.0321
Tot Expl. Difference	-0.0213	-0.0130	-0.0401	-0.0110	-0.0495	-0.0194

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

TABLE A.5: Fairlie Decomposition 2002-2006, by Skill level (Male Only)

	Obesity	
	(1) High	(2) Medium-Low
Demographic	-0.00443	-0.00454
	(0.0103)	(0.0106)
SES	0.00484	0.00420
	(0.00885)	(0.00548)
Education	0.000447	-0.00133
	(0.000866)	(0.000811)
Health	-0.00241	-0.00340
	(0.00273)	(0.00215)
Drinking	-0.000118	-0.00240*
	(0.000767)	(0.00109)
N. of Cigarettes (daily)	-0.00118	-0.000687
	(0.00107)	(0.00102)
Eating Habits	-0.000888	-0.00105
	(0.00299)	(0.00183)
Social Norm (BMI)	-0.0118*	-0.0209**
	(0.00599)	(0.00697)
Average BMI in HA	-0.0120	-0.0113*
	(0.00842)	(0.00515)
Observations	2419	3876
Food Prices	Yes	Yes
Obesity (Early)	0.196	0.193
Obesity (Late)	0.239	0.264
Difference	-0.0432	-0.0716
Tot Expl. Difference	-0.0272	-0.0415

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.6: Fairlie Decomposition 2002-2006, by Skill level (Male Only), Alternative social norm definition

	Obesity	
	(1) High	(2) Medium-Low
Demographic	-0.0220***	-0.0334***
	(0.00650)	(0.00690)
SES	0.00688	0.00660
	(0.00921)	(0.00602)
Education	0.000467	-0.000192
	(0.000884)	(0.000513)
Health	0.000423	-0.00305
	(0.00225)	(0.00205)
Drinking	-0.000195	-0.00131
	(0.000932)	(0.000707)
N. of Cigarettes (daily)	-0.000881	0.00137
	(0.000928)	(0.000798)
Eating Habits	-0.00131	-0.00133
	(0.00352)	(0.00216)
Perc. of Ob. in HA	-0.0204*	-0.0192**
	(0.00999)	(0.00607)
Observations	2419	3876
Food Prices	Yes	Yes
Obesity (Early)	0.196	0.193
Obesity (Late)	0.239	0.264
Difference	-0.0432	-0.0716
Tot Expl. Difference	-0.0366	-0.0506

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.7: Fairlie Decomposition 2002-2006, by Skill level (Female Only)

	Obesity	
	(1) High	(2) Medium-Low
Demographic	-0.00990	0.00288
	(0.0129)	(0.00725)
SES	0.00227	-0.00203
	(0.00843)	(0.00505)
Education	0.00211	-0.000424
	(0.00171)	(0.000375)
Health	0.00374	-0.00468*
	(0.00270)	(0.00199)
Drinking	0.00198	-0.00102
	(0.00206)	(0.000558)
N. of Cigarettes (daily)	-0.000222	0.000878
	(0.00253)	(0.000531)
Eating Habits	0.00210	-0.000189
	(0.00336)	(0.00114)
Social Norm (BMI)	-0.00635	-0.0147***
	(0.00748)	(0.00364)
Average BMI in HA	0.0145	-0.000986
	(0.0115)	(0.00536)
Observations	2382	4844
Food Prices	Yes	Yes
Obesity (Early)	0.194	0.216
Obesity (Late)	0.213	0.256
Difference	-0.0194	-0.0407
Tot Expl. Difference	0.0104	-0.0201

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.8: Fairlie Decomposition 2002-2006, by Skill level (Female Only), Alternative social norm definition

	Obesity	
	(1)	(2)
	High	Medium-Low
Demographic	-0.0196**	-0.0213***
	(0.00698)	(0.00590)
SES	0.00195	0.00210
	(0.00872)	(0.00539)
Education	0.00351**	-0.000311
	(0.00131)	(0.000347)
Health	0.00516*	-0.000310
	(0.00229)	(0.00159)
Drinking	0.00235	-0.000652
	(0.00248)	(0.000472)
N. of Cigarettes (daily)	-0.000163	0.000473
	(0.00279)	(0.000413)
Eating Habits	0.00238	-0.000225
	(0.00405)	(0.00180)
Perc. of Ob. in HA	0.00755	-0.00845
	(0.0116)	(0.00594)
Observations	2382	4844
Food Prices	Yes	Yes
Obesity (Early)	0.194	0.216
Obesity (Late)	0.213	0.256
Difference	-0.0194	-0.0407
Tot Expl. Difference	0.00334	-0.0284

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

TABLE A.9: Summary Statistics

	Men			Women		
	(1)	(2)	(3)	(4)	(5)	(6)
	Full Sample	BMI<25	BMI> 25	Full Sample	BMI<25	BMI> 25
	mean	mean	mean	mean	mean	mean
Employed	.8718931	.8609534	.8785376	.7393797	.7575709	.728663
BMI	27.43346	22.87404	29.5283	26.90789	22.43596	30.47886
Years of schooling	13.05274	13.28707	12.95148	13.04825	13.31611	12.84351
Age	42.58827	39.22078	44.14162	41.75934	39.69506	43.44113
General Health	1.844398	1.759509	1.877325	1.835631	1.69042	1.941731
Married	.8738879	.8837977	.8695853	.8230457	.8390283	.810324
Number of infants (0-1)	.091522	.1039218	.085711	.1097008	.1204049	.1010034
Number of childrens (2-15)	.5551167	.5155156	.5729467	.7081926	.7336032	.6905827
Living in rural area	.1115899	.0982438	.1181703	.1132494	.1212146	.1075314
Eating Habits	3.416058	3.352084	3.454781	3.694575	3.738354	3.672333
Has a Manual Job	.4637033	.4587287	.4648645	.2997155	.2677558	.3238326
Obese	.2469579			.2431842		
Overweight	.4438859			.3213226		
Healthy	.3021345			.421401		
Observations	25065	7573	17314	29307	12350	16544

Table A.10: Probit (ME) and Semi-Parametric Results for Employment Probability

	Men			Woman		
	Logit	Logit	GAM	Logit	Logit	GAM
	(1)	(2)	(3)	(4)	(5)	(6)
BMI	0.001	0.348***		−0.006*	0.128***	
	(0.005)	(0.035)		(0.003)	(0.022)	
BMI (Squared)		−0.006***			−0.002***	
		(0.001)			(0.0004)	
Age	0.099***	0.088***	0.084***	0.117***	0.115***	0.114***
	(0.014)	(0.015)	(0.015)	(0.011)	(0.011)	(0.011)
Age (Squared)	−0.001***	−0.001***	−0.001***	−0.002***	−0.002***	−0.002***
	(0.0002)	(0.0002)	(0.0002)	(0.0001)	(0.0001)	(0.0001)
Education (Year)	2.072***	2.063***	2.084***	2.220***	2.197***	2.192***
	(0.302)	(0.303)	(0.304)	(0.229)	(0.229)	(0.229)
Education (Square)	−0.073***	−0.073***	−0.074***	−0.080***	−0.079***	−0.078***
	(0.012)	(0.012)	(0.012)	(0.009)	(0.009)	(0.009)
Married	0.545***	0.543***	0.542***	−0.079*	−0.084*	−0.084*
	(0.061)	(0.061)	(0.061)	(0.043)	(0.043)	(0.043)
Infants in hh	0.030	0.034	0.028	−1.361***	−1.371***	−1.372***
	(0.082)	(0.082)	(0.083)	(0.047)	(0.047)	(0.047)
Kids in hh	−0.020	−0.028	−0.029	−0.470***	−0.471***	−0.472***
	(0.027)	(0.027)	(0.027)	(0.018)	(0.018)	(0.018)
Eating Habits	0.079***	0.077***	0.075***	0.026***	0.025***	0.025***
	(0.010)	(0.010)	(0.010)	(0.006)	(0.006)	(0.006)
SES	0.775***	0.765***	0.766***	0.672***	0.671***	0.671***
	(0.033)	(0.033)	(0.033)	(0.023)	(0.023)	(0.023)
Ethnic Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes	Yes	Yes
Health Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Region Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	21,059	21,059	21,059	24,594	24,594	24,594
Adjusted R^2			0.296			0.234
Log Likelihood	−6,746.816	−6,697.864	−6,709.020	−11,991.300	−11,971.940	−11,988.490
p-Value LR-Test	0.000	0.000	0.000	0.000	0.000	0.000

Note: *p<0.1; **p<0.05; ***p<0.01

Table A.11: Probit (ME) and Semi-Parametric Results for Employment Probability

	Men		Women	
	Manual Labour	Non-Manual Labour	Manual Labour	Non-Manual Labour
	(1)	(2)	(3)	(4)
Age	0.124***	0.105***	0.117***	0.134***
	(0.021)	(0.030)	(0.020)	(0.016)
Age (Squared)	−0.002***	−0.001***	−0.001***	−0.002***
	(0.0002)	(0.0003)	(0.0002)	(0.0002)
Education (Year)	0.592	1.678**	1.724***	1.144***
	(0.459)	(0.675)	(0.432)	(0.348)
Education (Square)	−0.016	−0.059**	−0.059***	−0.038***
	(0.018)	(0.026)	(0.017)	(0.013)
Married	0.483***	0.389***	−0.019	−0.324***
	(0.088)	(0.127)	(0.074)	(0.066)
Infants in hh	−0.159	0.055	−1.310***	−1.576***
	(0.116)	(0.165)	(0.094)	(0.062)
Kids in hh	−0.117***	0.050	−0.373***	−0.596***
	(0.037)	(0.059)	(0.032)	(0.026)
Eating Habits	0.064***	0.007	0.056***	−0.006
	(0.014)	(0.018)	(0.013)	(0.009)
Ethnic Dummies	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes
Health Dummies	Yes	Yes	Yes	Yes
Region Dummies	Yes	Yes	Yes	Yes
Observations	8,817	10,179	6,453	15,122
Adjusted R^2	0.305	0.172	0.195	0.151
Log Likelihood	−3,077.603	−1,967.651	−3,517.299	−6,313.269
p-Value LR-Test	0.000	0.0148	0.044	0.000

Note: *p<0.1; **p<0.05; ***p<0.01

Table A.12: Probit (ME) and Semi-Parametric Results for Employment Probability

	Men		Women	
	Logit	GAM	Logit	GAM
	(1)	(2)	(3)	(4)
FFM	0.017**		0.028***	
	(0.007)		(0.007)	
BF	−0.007		−0.009**	
	(0.008)		(0.004)	
Age	0.123***	0.114***	0.127***	0.125***
	(0.017)	(0.017)	(0.012)	(0.012)
Age (Squared)	−0.002***	−0.001***	−0.002***	−0.002***
	(0.0002)	(0.0002)	(0.0001)	(0.0001)
Education (Year)	1.317***	1.316***	2.060***	2.044***
	(0.354)	(0.355)	(0.249)	(0.249)
Education (Square)	−0.047***	−0.047***	−0.075***	−0.074***
	(0.014)	(0.014)	(0.010)	(0.010)
Married	0.433***	0.429***	−0.185***	−0.187***
	(0.071)	(0.072)	(0.047)	(0.047)
Infants in hh	0.005	0.002	−1.455***	−1.461***
	(0.093)	(0.093)	(0.050)	(0.050)
Kids in hh	−0.068**	−0.076**	−0.518***	−0.519***
	(0.030)	(0.030)	(0.019)	(0.019)
Eating Habits	0.038***	0.037***	0.007	0.006
	(0.011)	(0.011)	(0.007)	(0.007)
SES	0.737***	0.733***	0.626***	0.626***
	(0.038)	(0.038)	(0.025)	(0.025)
Ethnic Dummies	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes
Health Dummies	Yes	Yes	Yes	Yes
Region Dummies	Yes	Yes	Yes	Yes
Observations	19,248	19,248	22,513	22,513
Adjusted R^2		0.296		0.234
Log Likelihood	−5,198.697	−5,201.556	−10,405.170	−10,422.090
p-Value LR-Test	0.000	0.02	0.000	0.000

Note: *p<0.1; **p<0.05; ***p<0.01