The Role of Economic Factors in Obesity Prevalence and Diet Quality in Spain

Amr Radwan Ahmed Radwan

PhD in Economics

The Role of Economic Factors in Obesity Prevalence and Diet Quality in Spain

Amr Radwan Ahmed Radwan

Universitat de Barcelona

I dedicate this thesis

To the soul of my beloved mother and father which makes me what I am today

To the soul of my heart Radwan and Mariam as they paid for it much more than what I paid

To Sharif Gamal Siam and all Martyrs of the glorious revolution of January 25th and all the innocent victims of the brutal military coup as they irrigated the freedom tree with their immaculate blood.

The tree did not bear fruit yet, but we are confident that a day it will yield freedom, justice and prosperity.

Acknowledgments

The work on this thesis has been an inspiring, often exciting, sometimes challenging, but always interesting experience. It wouldn't have been possible without the cooperation and assistance of various people and institutions that deserve my acknowledgment.

First and foremost, I would like to express my deepest sense of gratitude to Prof. José Maria Gil Roig, the supervisor of this thesis and CREDA director, for his constant encouragement, his support which was far beyond duty, his constructive criticism, his friendly treatment, and for many helpful comments which were worth gold. He taught me many things not only in words but also in actions. A few of those that I have learned from him are how to set my benchmark even higher and to look for solutions to problems rather than focus on the problem, how he handled and arranged his works and duties, how he find interesting research problems, and how to manage your team work in a very human manner. Really I was extraordinarily fortunate in having you as a supervisor.

Thanks also go to The Agency for Administration of University and Research (AGAUR) for the financial support during the realization of this doctoral thesis.

I gratefully acknowledge Professor Rodolfo M. Nayga and Dr. Jayachandran Variyam for co-authoring two papers of this thesis which was a great enrichment for me. I am grateful in every possible way and hope to keep up our collaboration in the future.

I am thankful to all my colleagues in CREDA for providing perfect and friendly working atmosphere.

I would like to express my sincere thanks to my friends in Spain and Egypt, There are so many that it would be impossible to name them all, for always being there when I need them most. Especially I am grateful for my soul mate, Faical Akaichi, for being always ready to lend a hand from my first day in Barcelona until the last letter of this thesis.

I would also like to thank my professors and friends from the department of agricultural economics and the faculty of agriculture, Cairo University, who built the foundation of my professional values system, knowledge, and culture, and made possible my further education abroad, especially those who know that they play a very important role in my life.

Where would I be without my family? I am deeply and forever indebted to my parents, brothers, sisters, my beloved wife, my eternal sun shine, Radwan and Mariam, and all my family for their love, support and encouragement throughout my entire life. Their love and support in all that I have done till now are the key of all my achievements.

The chain of my gratitude would be definitely incomplete if I would forget the first cause of this chain. My deepest and sincere gratitude to Allah: who gives me the power and guide my way, not only during the realisation of this thesis but during my whole life.

Cairo, July, 2014
Amr Radwan

Table of contents

Chapter 1: Introduction

"A problem clearly stated is a problem half solved."

Dorothea Brande

Obesity is considered as a complex, multi factorial, chronic disease involving genetic, prenatal, socioeconomic, dietetic and environmental components. Worldwide prevalence of obesity nearly doubled between 1980 and 2008. In Europe, the regional office of the World Health Organization (WHO) estimated that in 2008, over 50% of both men and women were overweight, and roughly 23% of women and 20% of men were obese. These estimates indicate that obesity prevalence in Europe in the last two decades has tripled affecting more than 150 million adults and 15 million children and adolescents in the region.

In Spain, the last National Health Survey (NHS) for 2011-12 (INE, 2013) indicated that the prevalence of overweight and obesity among Spanish adults aged 18 or over was 36.7% and 17.0% respectively. While the prevalence of obesity was quite similar between men and women (18% among men and 16% among women), the overweight prevalence was significantly higher among men (45%) than among women (28%). Obesity was found to be associated with age as in the age segment between 18 and 24 its prevalence merely reached 5.5% for both genders. On the contrary, in the segment between 65 and 74 the prevalence of obesity reached 25.6% and 27.9% among men and women, respectively (although it is true that it decreased for the eldest segments). There is also a significant negative relationship between education level and obesity. In fact, 30.0% of obese people were found to be illiterate persons.

From a historical perspective, it is worth mentioning that, in spite of the up to now relative low percentages in relation to other EU countries, the prevalence of obesity in Spain has increased with a very alarming rate in 25 years moving from 6.9% and 7.9% among men and women, respectively, in 1987 to 18% and 16%, in 2012.

The Spanish National Survey of Dietary Intake (ENIDE) (2011), concluded that obesity rates in Spain was not due to eating too much (daily energy intake was 2482 kcal, slightly lower than the recommended level between 2550 and 2600 calories, depending on the individual's physical activity), but to an unbalanced diet characterized by the overconsumption of red meat, sodas and pastries. According to the Spanish Food Safety Agency (AESA), food habits have changed with a significant reduction of both family meals and the time allocated to eating during weekdays.

Noticeably, the prevalence of obesity has increased during the financial crisis that started to affect Spanish households in 2009 and more intensively during 2010. Comparing the data from the last two National Health Surveys (2006 and 2012), the obesity rate significantly increased from 15.6% to 18%, among males but not as much among women (from 15.2% to 16%). This result has to do with a lower consumption of fresh foods, fruits and vegetables and a higher consumption of fast food, ready-to-eat meals and fatty foods, which have been relatively much cheaper (Rao et al., 2013). This trend seems to continue in the future as the OECD predicted that the number of overweight and obese people in Spain will rise by a further 10% over the next decade.

In fact, an alarming 30% of teenagers are overweight, putting Spain just behind USA and Scotland. Moreover, a stunning 40% of youths aged between 13 and 18 never practiced sport.

Although the WHO characterizes overweight and obesity as diseases, it is also well known that both (together with smoking) are key determinants in the incidence of the most important contemporary chronic diseases, such as cancer, cardiovascular problems, certain types of diabetes, etc. The most worrying is that health disorders , such as type II diabetes, that were exclusively associated with elderly people are now being diagnosed in children, mainly due to the increasing prevalence of childhood obesity. In fact, one in every five adolescents in Spain runs the risk of suffering major cardiovascular problems in later life (Mora et al., 2012). There is a Positive significant association between obesity and the prevalence of all chronic illness in Spain (Costa-Font and Gil, 2005).

The economic costs associated with obesity are non-trivial as well. Obesity accounts for 7% of total health care costs (WHO, 2005) without considering other economic externalities which are difficult to estimate. In Spain, the Spanish Society for the Study of Obesity (SEEDO) estimated that direct and indirect obesity costs account for 7% of the total health care costs (2.5 billion Euros/year).

Mora et al. (2012) investigated the impact of BMI, obesity and overweight on direct medical costs using a longitudinal dataset of medical records for patients followed up over seven consecutive years (2004-2010).

They found that one unit increase in individual BMI increased direct medical costs by 5 to 10 € per patient and year. Similarly, obesity (overweight) increased direct medical costs by 50 to 96€ (17 and 79€) per patient and year. They concluded that, if half of the Spanish population experienced the same BMI increase, then the annual rise in direct healthcare costs would represent around 0.025% of GDP (256 million €).

In Spain, no systematic research has been undertaken to analyze economic factors affecting food consumption, the quality of diet and obesity and to assess the effectiveness of potential market intervention policies. Up to our knowledge, only there are some studies about the prevalence of obesity and the importance of socio-demographic factors that are affecting it (SEEDO and results from the project PORGROW)) in which the role of prices and income was neglected. From an economic point of view, we are only aware about three studies dealing with the relationship between chronic diseases and obesity, the socio-economic inequalities in obesity, and the relationship between maternal employment and childhood obesity in Spain (Mora et. al., 2012; Costa-Font and Gil, 2005; Garcia et. al., 2006).

Objectives

This Ph.D. thesis aims at analyzing the relevance of economic factors (mainly income and other socioeconomic characteristics of Spanish households and market prices) on the prevalence of obesity in Spain and to what extent market intervention policies (e.g. taxing less healthy foods) are effective to reduce obesity and improve the quality of the diet, and under which circumstances. In relation to the existing worldwide literature, this project is the first attempt in Spain trying to get an overall picture on the effectiveness of public policies on both food consumption and the quality of diet, on one hand, and on the prevalence of obesity on the other hand. To achieve this general objective, we propose to tackle the following specific objectives:

1. Presenting an overall panorama of the research that dealt with economic causes and consequences of obesity, the role of economic factors on the growing obesity prevalence and the role of different

economic intervention policies in combating this crisis and incentivizing healthy eating among consumers.

2. Analyzing the relevance of economic factors (mainly income and other socioeconomic characteristics of Spanish households and market prices) on the prevalence of obesity in Spain, this is one of the main contributions of this thesis.

3. Building a new Obesity Specific- Healthy Eating Index (OS-HEI) based on available information.

4. Analyzing the effect of obesity, as an approximation for diet quality, on the consumption patterns of the main food categories among Spanish consumers and assessing the differences between obese and non-obese consumers in their consumption patterns.

5. Analyzing the effectiveness of market intervention policies (e.g. taxes and subsidies on foods and income) on reducing the prevalence of obesity in Spain.

6. Providing guidelines about data needs for further research

Thesis structure

The dissertation is consisted of four papers. The first paper which follows the thesis' introduction includes a critical review of the literature on the economic approach used to understand and deal with obesity prevalence epidemic, diet quality and public intervention policies. Although an important body of obesity literature is dealing with physical exercise, in this paper, however, we focus on those studies related to food consumption and diet quality.

Through the literature review, we also identified the gaps the other papers of the thesis tried fill. The main gaps identified in the exiting literature are as follows:

1- Data availability is the main limitation to carry out any economic analysis that relates obesity, food consumption and food prices.

This thesis filled this gap through merging the two main sources of secondary data on food consumption, food prices and the diet's health aspects in Spain. The first one is the National Health Survey (NHS). The NHS is a cross-section survey that provides ample data on the health status of citizens and its determinants. The survey collects information on the individual socioeconomic characteristics, morbidity, food habits and the demand for health care. Food habits refer to two main issues: type of breakfast and frequency of consumption of selected food groups. However, the data set does not provide information on quantities consumed (or purchased) neither on prices. The second main source that mainly refers to consumption data is the Spanish Household budget Survey. This survey provides annual information on the expenditure and quantity consumed of various classes of food products consumed by a stratified random sample of around 24,000 households. Since prices are not explicitly recorded, unit values for each group are calculated dividing expenditures by quantities. The survey also gathers information on a limited number of household characteristics including the level of education and main activity of the head of the household, household income, household size, age and sex of family members and town size, among others. As similar characteristics are found in both data sets, our methodological approach consisted of merging the two databases by defining different segments of population, using the socioeconomic characteristics of households. Then for each segment we calculated average values for the relevant variables contained in each database.

2- Clear geographical concentration is observed as most published literature that deal with the economics of obesity come from North America and some concentrating in Europe. Our thesis represents the first systematic research that was undertaken in Spain to analyze economic factors affecting food consumption, the quality of diet and obesity as well as to get information about the effectiveness of potential market intervention policies.

3- From a methodological point of view, the literature review showed that most of the studies that dealt with the economic analysis of obesity used a parametric approach. The complex and multifactorial nature of obesity urge the use of nonparametric techniques capable of capturing the nonlinear effect of some covariates and the interaction between them. Our analysis in the third chapter of this thesis could be considered exception of this trend since we applied the Multivariate Adaptive Regression Splines MARS model, as a nonparametric alternative of the multinomial logit model in analyzing most determinant factors that affect obesity prevalence in Spain. The obtained results are quite different from those obtained using traditional parametric models, more consistent with the literature and offering specific recommendations for subgroups of the population. This encourages more use of such nonparametric models in studying economics of obesity

4- Another issue related with intervention policies to enhance diet quality and combat obesity is which food groups could be considered healthy (unhealthy) and how to measure the effect of specific policy on diet quality. Diet quality appears to have no official definition in the literature. Definitions vary widely depending on the tools used to measure it. A main contribution of our thesis is the development of the new Obesity Specific-Healthy Eating Index (OS-HEI) which is presented in chapter 4.

5- A main drawback of the articles that dealt with the effect of intervention policies to improve diet quality and combat obesity is that they mostly concentrated on the effect of such intervention policies on a specific food group (e.g. Fruit and vegetable; SSB; fast food … etc.) which shadow doubts on reliability its results as they omitted the income and the substitution effect and did not take into account the holistic nature of the diet. Gao et al. (2013) tried to fill this gap applying household production theory to systematically estimate consumer demand for diet quality using the Healthy Eating Index (HEI). Their results indicated that consumers have insufficient consumption of food containing dark green and orange vegetables, legumes and whole grains. Combining the taxation of sugar sweetened beverages (SSBs) and/or fats with targeted unsweetened beverages and/or fruit and vegetable subsidies, could have a higher positive effect on enhancing consumer's diet quality. Moreover, this combination offer a more politically viable option as it is expected to have a higher social acceptance and the revenue of the tax could be used in financing the subsidy without any extra burden on the public expenditure. Another gap observed in the literature is the need to use advanced demand systems capable of taking into account the unobserved heterogeneity, among individuals, which is especially important in the case of obesity. Those demand systems also should be quite flexible to fit the complexity and multi-dimensional nature of obesity. The recent developed Exact Affine Stone Index (EASI) could be considered more than appropriate in doing so. Zhen et. al. (2013) used EASI demand system in analyzing the effect of SSBs tax on obesity prevalence in United States. Up to our knowledge no published article succeed in dealing with the two aforementioned shortcomings by using flexible demand systems such as EASI in analyzing the demand for diet quality but in a framework where sugar sweetened beverages the holistic nature of the diet is taken into account. This could be considered the main contribution of this thesis. In fact, in the fifth chapter an EASI demand system is estimated to assess the Spanish consumers demand for diet quality taking into account the person BMI.

6- The literature review also showed that there is need to apply behavioral and experimental economics techniques. While traditional economic analyses of obesity can tell us who is obese, they are incapable of telling us why people are obese or explaining the heterogeneity within the different demographic groups. On the other hand, behavioral economics techniques could play a vital role in doing so. Incorporating behavioral aspects into the economic analysis of obesity is quite challenging as behavioral factors are determined by complex socio-psychological determinants such as habits, emotions, attitudes, parental feeding practices, beliefs, etc.

How these gaps were filled by the present thesis and the aforementioned objectives were addressed is the subject the next four chapters of this thesis. The thesis is organized as follows. Chapter 2 presents an overall panorama of the economic literature that dealt with obesity concentrating on the role of economic factors in obesity prevalence and the different intervention policies that were implemented to combat it. In Chapter 3, the determinant factors that affect obesity prevalence were analyzed using a nonparametric approach, the Multivariate Adaptive Regression Splines (MARS). The development and validation process of our new Obesity Specific- Healthy Eating Index (OS-HEI) is discussed in Chapter 4. Chapter 5 deals with the estimation of Spanish consumers demand for food and diet quality, approximated by the BMI, using the EASI demand system and the simulation of the effect of the different price intervention policies on food demand and diet quality in Spain. The thesis ends with some concluding remarks and some suggestions for further research.

Chapter 2:
The Obesity Fighting: Can Economics Help?

"A problem is a solution yet to be discovered"

Marc Torra

1. Introduction

Obesity is a growing wide spread epidemic all over the world. Overweight and obesity are defined as “abnormal or excessive fat accumulation that may impair health” (WHO, 2013). Body Mass Index (BMI) is the most used measure of estimating whether a person is overweight or obese. BMI is defined as person’s weight in kilograms divided by the square of their height in meters. A person is considered overweight (obese) if his/her BMI ranges between 25 and 29.9 (greater than 30). Overweight and Obesity have become the first public enemy not only in developed countries but also in many developing countries. According to the World Health Organization (WHO), obesity prevalence doubled between 1980 and 2008. In the past three decades, there have been considerable changes in lifestyle around the world mainly because of the globalization of the food supply and urbanization (Swinburn et al., 2011). These changes have affected people’s diet and decreased levels of their physical activity, resulting in a significant increase in body mass index (BMI) (Chaput and Tremblay, 2009). Obesity has reached epidemic proportions globally, resulting in around 3 million yearly deaths due to being overweight or obese, surprisingly more than deaths due to under nutrition. In 2008, more than 1.4 billion adults, 20 years old and older, were overweight. Of these over 200 million men and nearly 300 million women were obese. About third of the world population was overweight and one tenth was obese. More than 40 million children under the age of five were overweight in 2011. Stevens et al. (2012) reviewed obesity prevalence worldwide and concluded that by 2008 one third of adults were overweight and that the third of those overweight people were obese. Moreover, 20 percent of women and men in 117 and 73 countries respectively were obese. Obesity prevalence rate notably has accelerated during the last decade. Logically this high obesity prevalence rate was found to be associated with a higher mean BMI. Although obesity increased in most countries, levels and trends varied substantially.

Senauer and Gemma (2006) tried to answer the question why is the obesity rate so low in Japan and high in the U.S from an economic point of view. Comparing obesity rates in the Japan and the United States is especially relevant as Japan has one of the lowest rates and the United States one of the highest rates of obesity in the world. They concluded that the prevalence of obesity is lower in Japan partially because the Japanese walk more during their daily lives than Americans do. The other reason behind this difference is that Japanese are more likely to use public transportation due to the high cost of owning and operating an automobile in Japan. The use of public transit usually requires more walking than driving one's own car.

Most of the papers that studied economic causes of obesity generally examine only one or a few factors at a time. The literature therefore lacks a clear answer to the big-picture question of how well "the economic explanation" of people responding to changing incentives can explain the rise in obesity.

Although obesity and overweight were previously considered a problem of high income countries, they are now spreading fast in developing countries and have reached world record levels in some of them (e.g. Mexico). Furthermore, they resulted in a double burden on these countries wherein population, especially children, is suffering from high prevalence of malnutrition (e.g. India) (Malik et al., 2013).

Clear geographical concentration is observed as most published literature that deal with the economics of obesity come from North America and some concentrating in Europe. Future research should estimate food demand and its effect on obesity in low and middle income countries. Therefore the question is to what extent the results can be transferred to other countries. This can be explained through three main reasons. First, the obesity epidemic has started earlier in the United States. Recent trends have shown that other OECD countries follow the development in the US with a time lag. There is good reason to believe that at least some of the underlying trends are not only a local but a universal phenomenon. However, there are certainly cultural and other differences that are important and that need to be considered urge for more geographical cover of studies on the economics of obesity. Special attention should be given to developing countries as they are progressively suffering from a double burden of under and over nutrition which is not the case in developed countries. Another factor contribute to the big production of papers in USA is the data availability and the relative higher quality of such data comparing with other countries. Availability of the necessary fund to conduct such studies is also contributing to the observed geographical bias.

Obesity has become a major public health and public finance concern. The cost of obesity, being paid not only by individuals but also by societies and governments, is quite high. The World Health Organization (2011) estimated that 2.8 million people die each year as a result of excess weight. The obesity in United States was estimated to annually cost 112,000 lives and $190 billion (Flegal et al., 2005; Cawley and Meyerhoefer, 2012).

The skyrocketed obesity trend and its high cost led economists to examine whether obesity can be considered as an economic phenomenon. Technological progress has resulted in an obesogenic environment characterized with abundant easy to access cheap food and increasingly easy to avoid physical activity. Taking into account people response to the mentioned changes in food prices and physical exercise could be helpful in explaining the rise in overweight and obesity prevalence. Integrating economics into the obesity epidemic research could be helpful due to two main reasons. First, economics provides a set of analytical techniques and quantitative methods to analyze consumer behavior regarding food and physical exercise and consequently obesity. Furthermore, economics provides a way of thinking about individual behavior.

The main objective of this paper is to present an overall panorama of the research that dealt with economic causes and consequences of obesity, the role of economic factors on the growing obesity prevalence and the role of different economic intervention policies in combating this crisis and incentivizing healthy eating among consumers. The aim consists in figuring out whether economics can help in combating the growing obesity epidemic. The answer we found, through revising the existing literature, is that economics, considered a main cause of obesity, could be an effective cure for it. In this paper, we limit our attention to adult obesity although we are conscious that child obesity is an important and growing worldwide crisis. We discard child obesity from our scope because two review papers that reviewed the economic aspects of child obesity were recently published (Papoutsi et al., 2013; Cawely, 2010). Although obesity prevalence is affected by diet and exercise choices, in this review we concentrate on the economic aspects of diet choices. A detailed review on economic aspects of exercise choices were done by Wu et al., (2011); Bleich and Sturm, (2009); and Sturm, (2005) and (2004).

In contrast with previous literature reviews, dealt with the economic aspects of obesity, in this literature review we tried to offer a complete panorama on this issue and evaluate it critically suggesting some interesting research lines for future research. Previous literature reviews offered only one horizon of the issue such as the role of food prices on obesity prevalence (Powel and Chaloupka, 2009); impact of economic policies targeting obesity (Faulkner et al., 2011); use of economic instruments to promote dietary and physical activity behavior change (Shemilt et al., 2013). Moreover our literature review concentrated on recent literature published during last few years. This allowed us to present an up to date view of the issue. This is quite important in such a dynamic field of research with a growing body of the literature.

The rest of the paper is organized as follows. Section 2 discusses briefly the rationale of the economic approach in studying obesity. The economic causes and cost of obesity have been reviewed in sections 3 and 4, respectively. The effect of income, food prices and other socioeconomic factors on obesity prevalence has been discussed in section 5. Overview of the different interventions policies applied to combat obesity and the rationale behind it presented in section 6. The paper ends with a future research agenda wherein we have tried to figure out if behavioral and experimental economics could be considered useful tools in overcoming shortcomings of traditional economic tools used to study obesity, the application and future of some emerging disciplines such as Genoeconomics, Neuroeconomics in studying and the role of food industry in obesity prevalence is argued.

2. Rationale of the economic approach in studying obesity

Taking into account the complex multifactorial nature of obesity, an interdisciplinary approach seems to be the most fruitful to address the obesity epidemic. Looking at the biological origin of obesity, the will to survive have caused preferences for high energy dense foods, with the goal to accumulate energy reserves for an uncertain future, and for physical inactivity (Hill et al. 2004). While the environment has experienced radical changes during past decades, the genetic makeup of individuals has hardly changed. This makes it hard to explain recent hike in obesity rates by genetics as most of the rise has spanned less than a single generation. Moreover, while biology is capable of offering a sound explanation of cross sectional differences in obesity between individuals (by gender, age, etc.), it is incompetent in explaining the sharp rise in obesity that has occurred in too short time to have any significant genetic or biological causes (Anderson and Butcher, 2006). The fundamental question of our paper is why to integrate economics into research on the dynamics of obesity prevalence. Integrating economics into the obesity epidemic research could be helpful due to two main reasons. Firstly, economics provides a set of analytical techniques and quantitative methods. Furthermore, economics provides a way of thinking about individual behavior. One of the most popular definitions of the economics is “the study of choice”.

The economic view of diet and exercise choices

The skyrocketed obesity trend has provoked economists to examine whether obesity can be considered an economic phenomenon involving individuals’ responses to incentives. Technological progress has resulted in an obesogenic environment characterized with abundant easy to access cheap food and increasingly easy to avoid physical activity. Taking into account people response to the mentioned changes in relative food prices and physical exercise could be helpful in explaining the rise in overweight and obesity prevalence.

Philipson and Posner (2003), in their seminal paper, reinforce this notion by modeling weight status as the result of both eating and exercise decisions using a utility maximization framework where Individuals tradeoff the disutility from excess weight with the enjoyment of eating and relaxing lifestyle, subject to a budget constraint. The model predicts that lower food prices and reduced physical activity should increase weight, while the effect of additional income on weight varies across the income distribution. Cutler et al. (2003) pointed out that the time costs of eating could matter in addition to the monetary costs shading light on how time saving innovations such as vacuum packing and microwaves have reduced the time cost of food practices, while Courtemanche (2009) and Wehby and

Courtemanche (2011) suggested that the long-run relationship might actually be negative.

Lakdawalla and Philipson (2009) presented a dynamic theory of body weight and developed its implications. They argued that technological changes have prompted weight growth through making people's activities more sedentary and by lowering food prices through agricultural intensification and innovation.

The economic view is that people are involved in the production of their own health and happiness through combining their time with market goods such as foods and health care. Evidently, more of everything (food, health, happiness…etc.) will always be better. In a real world characterized by scarcity of resources, it is not the case as all these variables are subject to constraints. These constraints includes: time constraints (i.e. 24 hours per day), budget constraints (i.e. financial resources), and biological constraints as weight will rise when caloric intake is higher than calorie expenditure. People cannot choose a level of health or wellbeing directly nevertheless they do so indirectly through their behavior. The main two obesity related behaviors that consumers can control are physical activity and caloric intake. It is worth mentioning that while high caloric intake increases utility directly, it could decrease it indirectly by decreasing health status. The positive indirect effect of physical exercises has on utility is indubitable, its direct effect could, however, be positive or negative depending upon the nature of this physical exercise.

In economics price is defined as opportunity cost that does not only include the monetary price attached to a product or a service but also the time and other costs (e.g. risk) associated with consuming it (Propper, 2004). The price is linked to the idea of incentives since a change in the relative price of an alternative represents a change in incentives to choose that alternative.

Body weight is the result of a bulk of economic decisions affecting behavior. These decisions involve market transactions directly through the purchase of food or other goods. Taking into account the income constraint, an individual will normally buy food that represents the best value for the price. The decisions also involve market transactions indirectly through the allocation of scarce leisure time to physical activity or preparing healthy meals (Smith 2002).

3. Economic burden of obesity

Obesity has a quite high cost that is paid not only by obese persons but also by governments and societies. From an economic perspective, obesity has a severe impact in multiple additional areas beyond health costs, including productivity, transportation, and the loss of better opportunities by obese individuals (O'Grady and Capretta, 2012). While cost estimates differ, the literature shows that obesity is linked with high health care costs and economic productivity losses. Obesity accounts for 7 percent of total health care costs (WHO, 2005) without considering other economic externalities, which, on the other hand, are difficult to estimate. In Spain, the Spanish Society for the Study of Obesity (SEEDO) estimates that obesity costs account for 7 percent of total health care costs (i.e. 2.5 billion Euros/year).

Mora et al. (2012) investigated the impact of BMI variation, obesity and overweight on direct medical costs using a longitudinal dataset of medical records for patients followed up over seven consecutive years (2004-2010). They found that one unit increase in individual BMI increases direct medical costs by between 5 and 10€ per patient and year. Similarly, obesity (overweight) increases direct medical costs by between 50 and 96€ (17 and 79€) per patient and year. They concluded that, if half of the Spanish population experienced the same BMI increase, then the annual rise in direct healthcare costs would be around 0.025 percent of the Spanish GDP (i.e. 256 million €).

In USA, estimates by Finkelstein et al., (2009) showed that the annual medical burden of obesity had risen to almost 10 percent of all medical spending and could amount to $147 billion a year in 2008. They also mentioned that the increase in obesity-attributable costs between 1998 and 2006 was mainly due to 37 percent increase in obesity prevalence rate and not to the broader trend of rising health costs. A recent study by Behan and Cox, (2010) estimated the cost of obesity in United States and Canada. They found that the total economic cost of overweight and obesity was about $300 billion per year (i.e. overweight was responsible for approximately $80 billion, and approximately $220 billion was due to obesity). They also found that 90 percent of the total cost in both countries corresponds to United States.

Cawley and Meyerhoefer (2013) have solved what is called the causality problem applying more sophisticated modeling techniques than those used in previous studies. Most of their estimates measured how obesity and other medical conditions are associated with spending, but not the relation of causality between obesity and the other medical condition. They segmented obese people into two subpopulations: 1) people for whom obesity was the effect but not the cause, and 2) people for whom obesity was the cause but not the effect. In the subpopulation where obesity was the cause, they found that the per-person cost was 249, 226 and 846 percent higher for drug prescription and inpatient spending, respectively (Cawley and Meyerhoefer, 2013).

It is noteworthy that while cost estimates of obesity are available for developed countries, we could not find such estimates in the case of transitional and developing countries.

4. Economic causes of obesity

It is a given that a person gains weight when their caloric intake exceeds the calories expended through basic metabolism and physical activity. Each excess 3,500 consumed calories consumed could result in one pound excess in person's weight (Zahour, 2006). While this is true, it masks the more complicated picture about how energy is acquired and used. Despite recognizing the role of personal responsibility, it is widely accepted that the causes of obesity are complex and multi-factorial, with a sizeable part of the literature focusing on environmental rather than individual factors. Currently, people are suffering from being surrounded with a tremendously and progressively obesogenic environment that promotes a high energy intake as well as a sedentary life style (Swinburn et al., 1999). This environment in which we live and which often makes unhealthy choices more accessible was termed obesogenic by the World Health Organization in 1998. Obesogenic environments are defined as 'the sum of influences that the surroundings, opportunities, or conditions of life have on promoting obesity in individuals or populations' (Swinburn and Egger, 2002). The complex and multifactorial nature of overweight and obesity makes it quiet difficult to attribute to a single factor the growing obesity epidemic, but it is doubtless that economic policy is considered one of the main areas for intervention. Over time the total price, taking into account both money and time, of consuming different foods and beverages has declined leading to a notable reduction in calories' prices and increase purchasing power. At the same time, Physical exercise has become more expensive because of higher wage rates and longer hours spent in sedentary employment. Taking into account the aforementioned phenomena and economic theory, it can be predicted that these price changes would rationally lead individuals to increase their calories intake and reduce their caloric expenditure. Madore (2007) highlighted that the increasing prevalence of obesity come accompanied with a decrease in the relative price of consuming a calorie over time and a rise in the opportunity cost of burning a calorie. Concluding that, economic instruments, capable of changing the price of food and exercise, could affect food consumption and physical activity and consequently the overweight and obesity prevalence. This situation is worsening nowadays by having a virtual obesogenic environment beside the real obesogenic environment we are surrounding with. The rise of social networking websites and computer games has resulted in making the Internet a popular destination especially for the young generation. This trend make the surrounding environment more obesogenic through encouraging a more sedentary life style and allowing less time for physical exercise and other calories consuming activities. Moreover, fast food companies have become active in social networking by extending its marketing strategies into cyberspace which resulted in increasing the consumption of junk food especially among adolescents.

Economists have sought to explain the increase in obesity rates over time by the decreasing relative cost of food, in particular the cost of calories and fat, and the increasing cost of physical activity (Philipson and Posner, 1999 and Cutler et al., 2003). The relative price of food has declined substantially due to technological changes, which have increased the efficiency of production in agriculture and distribution throughout the food system. Furthermore, technology has eliminated much of the need for physical activity during work or for mobility. Nowadays, getting significant physical activity requires a conscious commitment to exercise which becomes timely and monetary expensive. In the not so distant past, the majority of jobs entailed heavy physical activity, so people were in a sense being paid to exercise.

Most of the papers that studied economic causes of obesity generally examine only one or a few factors at a time. The literature therefore lacks a clear answer to the overall question of to what extent the economic explanation of people responding to changing incentives can explain the rise in obesity.

From a methodological point of view, it can be observed that most of the studies dealt with the economic analysis of obesity, used a parametric approach. The complex and multifactorial nature of obesity urge the use of nonparametric techniques capable of capturing the nonlinear effect of some covariates and the interaction between them. Third chapter of this thesis could be considered exception of this trend where the MARS model applied, as a nonparametric counterpart of the multinomial logit models in analyzing most determinant factors that affect obesity prevalence in Spain. Results are quite different from those obtained from traditional parametric models and more consistent with the literature and offering specific recommendations for subgroups of the population. This encourages more use of such nonparametric models in studying economics of obesity.

5. Effects of income, food prices and other socio economic factors on obesity prevalence

Previous economic studies have analyzed the influence of income on health. In general, there seems to be a consensus about the positive effect of income increase on health (Smith, 1999). Consequently, we would expect, all things being equal, an association between low income and obesity due to: 1) the unavailability of healthy foods in low income neighborhoods (Beaulac et al., 2009; Larson et al., 2009); 2) when healthy foods are available they are usually more expensive, so poor families cannot afford to buy them (Drewnowski, 2010; Monsivais and Drewnowski, 2009); 3) Lower income neighborhoods have fewer physical activity resources making it difficult to lead a physically active lifestyle (Moore et al., 2008); and 4) Low income families are more expected to face high levels of stress increasing the likelihood of being overweight or obese (Gundersen et al., 2011; Moore and Cunningham, 2012). Costa-font et al. (2014) studied the effect of income inequalities on unhealthy behaviors in Spain and found that obesity is concentrating in poorer classes in recent decades. Lakdawalla and Philipson (2009) characterized how body weight varies depending upon income. They found that, within a country, income might have an inverted U shaped relationship with bodyweight due to the offsetting effects of the demand for food and the demand for an ideal body weight. This can have important implications for the bodyweight impacts of public transfer programs. Across countries, they, however, found that the mean weight is likely to be higher in richer countries. To measure the extent to which income affects obesity, Moran et al., (2013) exploited the data from a natural experiment of the Social Security Benefits Notch. The Notch is the result of a legislative accident that created variation in retirement income that was large, unanticipated, and beyond the control of the individual, making it a suitable instrument. They estimated models of instrumental variables (IV) using data from the National Health Interview Survey and find little evidence that income affects weight. For instance, they found that a permanent $1,000 increase in Social Security income (in 2006 dollars) resulted in a change in weight of more than 1.4 pounds in either direction for men or women.

Apart from income, food prices also play an important role in obesity prevalence. Powell and Chaloupka (2009) in their literature review concluded that food prices had a significant but small effect on obesity and overweight prevalence. As a result, they concluded that fiscal pricing policies could help in reversing obesity trends and those small taxes or subsidies were not able to do so while nontrivial pricing interventions might have measurable effects on weight outcomes, especially for those belonging to low social and economic class.

Beside income and food prices, several recent economic studies explain the role played by different cultural and socio-demographic factors on obesity rates. Leaving genetics aside, obesity is caused by consumption of too much calories and/or low expenditures of calories (i.e. low physical activity). For example, Schlosser (2002) showed that the rapid growth of fast food and soda drinks consumption has increased the dietary intake of saturated fats, sugars, and calories and accordingly, the prevalence of obesity. Other researchers argue that female labor participation is a leading factor in increasing obesity rates (Cawley and Liu, 2012; Garcia et al., 2006), mainly in children.

6. Intervention policies to combat obesity

The role of government in reducing obesity and overweight rates is quite debatable. While public health experts normally favor government intervention because of the government responsibility of protecting citizen from the many diseases related to overweight and obesity such as diabetes and cardiovascular diseases. The main reason justifying why most mainstream economists agree on the importance of government intervention to tackle obesity crisis is the market failure due to information asymmetry, self-control problems and externalities (Cawley, 2004). Information deficiency refers to the fact that health consequences of diet and exercise choices are not clear to most consumers. This assumes that emphasizing more these health consequences could lead to reducing obesity and increasing consumer awareness about the health consequences of his/her diets. Improving consumer awareness could be achieved through nutritional labeling and some educational campaigns. As noticed by Cash et al. (2007) nutritional labeling may not be sufficient because being aware of the nutritional content of a product does not necessarily means being aware of the health consequences of this nutritional content. The second type of market failure that justifies intervention is the self-control problems. Many obese people or those with weight problems have difficulty controlling their food intake or devoting sufficient time to physical activity, despite knowledge of the consequences of the excess of caloric intake. Behavioral economists suggest that this self-control problem could be grounds for government intervention. These economists think of individuals as having two ‘selves’: a relatively myopic self, which is the one that makes diet and exercise decisions, and a relatively far-sighted ‘future’ self, which lives with the health consequences. There is sometimes a conflict between the two selves as Today’s self may not adequately take into account future self’s welfare resulting in attraction to calorie rich foods and sedentary lifestyles. Excise taxation of unhealthy foods or physical activity subsidies can be thought of as a commitment devise to improve the ‘future-selves’ welfare (Gruber and Kozegi, 2004; O'Donoghue and Rabin, 2006). Another justification for economic intervention is that obesity results in large health care costs and these costs are borne collectively, so that obesity imposes financial externalities on those who are able to exercise self-control. It could be considered as the thin are subsidizing the health care costs of the obese. Also, obese people seem receive comparable wages although they have lower productivity. Excise taxes on unhealthy foods would make obese individuals internalize the costs they are imposing on others.

The discussion has, so far, focused on the desirability of tax and subsidy policies to correct various types of market failure. Another rationale for government intervention is to improve the welfare of low income households. Those with limited means often economize by purchasing calorie-dense and processed foods and drinks. The reason is that, although these items are not particularly nutritious, they may provide the cheapest calories per dollar. Transferring income to such households would enable them to purchase more costly nutritious foods. The advantage of income transfers is that they can be targeted at those who are less affluent (most excise taxes and subsidies, on the other hand, target everyone.) One limitation is that there is no guarantee that less affluent households will use the subsidy to purchase healthy foods. In fact, they might have different priorities. From the above discussion, intervention policies to combat obesity can be defined as policies that aim to change individuals’ behaviors and weight outcomes through changing relative prices and through income transfer programs.

Mazzocchi and Traill (2006) classified the wide range of potential instruments available to public authorities in four groups according to their expected impacts on economic agents: 1) policies addressed to change consumer utility function; 2) those aimed to better inform consumers’ choices without changing the utility function; 3) market measures addressed to affect actual choices without changing the utility function; and 4) supply-side policies affecting food availability. They show that the number of potential alternatives is very large and very heterogeneous in nature, which, on the other hand, merely reflects the complexity of the problem and the number of factors influencing dietary habits and intakes (i.e. individuals’ socioeconomic characteristics and lifestyles). Moreover, it is also true that food policies addressed to the emerging nutrition challenges need to coexist with agricultural and trade policies, which have traditionally regulated the agro-food activities with very different objectives. Such coexistence may reduce the effectiveness and complicate the implementation of some of the instruments.

Dealing with the alarming and growing public health burden of obesity need a comprehensive and a well-designed combination of regulatory, educational, agricultural and economic policies. Without any doubt economic interventions by themselves are not the magic solution for the obesity puzzle but should be considered one of the most important components of such integrated approach.

Although there is enormous number of papers dealing with the economic aspects of obesity, shifting from empirical evidences to policy recommendations is still challenging. This raises the need to learn from tobacco control policies. Initial tobacco control interventions were not evidence based but represented sound judgment at the time.

Faulkner et al. (2011) conducted a scoping review with the aim of synthesizing existing evidence regarding the impact of economic policies targeting obesity and its causal behaviors (i.e. diet and physical activity). Their results showed that there is consistent evidence that weight outcomes are responsive to food and beverage prices and, hence, the impact of any specific economic instrument on obesity independently would be relatively modest. Shemilt et al. (2013) conducted a scoping review analyzing the use of economic instruments to promote dietary and physical activity behavior changes. They defined economic instrument as follows: “it encompasses fiscal or legislative government policies designed to change the relative prices of goods or services or people’s disposable income, and promotional practices used by retailers to change the relative prices of goods and services”.

A tax on foods, especially junk foods, has been suggested. People have developed unhealthy habits in response to the low price of foods, especially calorie-dense foods, and the low relative price of driving a car for transportation. However, these habits could be changed by altering the structure of economic incentives on which people base their decisions.

Economic policies could be used to create incentives to both reduce excess calories consumption and excessive reliance on the automobile (Senauer and Gemma, 2006). Modeling a tax rate similar to that applied on tobacco products, of approximately 58%, the researchers found that a greater weight loss would be achieved. The acceptability of taxes at this level is not Clear (Fletcher et al., 2010). Imposing such high tax level in the case of food could be considered a quite challenging task as eating is both an absolute necessity and intrinsically healthy, whereas smoking have been shown to pose serious health risks. Because of that, a direct tax on food, even on high-calorie foods that are low in nutrients, for the purpose of reducing obesity is not politically feasible. Additionally, a tax on food is regressive, since people with lower incomes spend a large share of their budget on food. A solution for this social and political unacceptable taxing food policy could be solved through combining such food tax with a food subsidy of healthy foods which assure reducing the effect of food taxes on poor people and granted a higher political acceptance.

Lakdawalla and Philipson (2001) inspected how food taxation may impact BMI. They found that high food taxation (relative to the taxation of other goods) increases the price of food, and this increase in the relative food price will, in turn, decrease BMI. Schroeter et al. (2008) simulated the impact that changes in taxes or subsidies for foods and drinks would have on body weight. They found that a tax on food away from home or a subsidy of vegetables and fruits would increase body weight while a tax on regular soft drinks or a subsidy of diet soft drinks would lower body weight. The result that an increase in the tax on food away from home would increase the body weight appears surprising. However, the tax on food away from home does in fact decrease away-from-home food consumption, but it increases at home food consumption (e.g. processed food) due to the fact that the two categories are substitutes. Because many of the foods consumed at home are energy-rich, total consumption actually increases. Etilé (2008) estimated, using French household survey data, the impact of a 10 percent price decrease in fruits and vegetables on body weight and the impact of a 10 percent price increase in soft drinks, pastries, deserts, snacks and ready-meals on body weight. He then simulated the impact of five policy scenarios (taxes and subsidies) on the overweight and obesity prevalence. The author found that with each of the five policy scenarios the distribution of BMI in the population is clearly more favorable in a public-health sense.

Healthy eating interventions may be more effective if consumers perceive their eating habits as a more serious personal health risk. Hoefkens et al., 2013 investigated European consumers' perceived seriousness of their eating habits, its determinants and relative importance among other potential personal health risks including weight, stress and pollution. They found that European consumers underestimated the seriousness of their eating habits for personal health. Eating habits were perceived as more serious among women, Italians, obese, and younger individuals with a stronger health motive .This population group may therefore be more receptive to healthy eating interventions. Nevertheless, other potential health risks may still be considered more important than personal eating habits. Additional efforts are needed to raise Europeans' awareness of the seriousness of their eating habits for personal health, especially among population groups who perceive their eating habits less serious.

In the next paragraphs different intervention policies targeting specific food groups and other policies discussed.

Fruit and vegetable subsidy

The empirical evidence clearly demonstrates a link between lower obesity risk and greater fruit and vegetable consumption. In particular, lower prices of fruits and vegetables are associated with lower child weight.

'5-a-day' public information campaigns have been implemented in many countries to promote the consumption of at least five portions of fruits and vegetables per day per person as part of a healthy diet. It was originally designed by the California Department of Health in late 1980s, and then implemented in other states in the United States (US). In 1992, the state of Victoria (Australia) implemented a version of the 5-a-day campaign. At the beginning of the last decade, the campaign expanded to the United Kingdom (UK), France and many other countries. Today, more than twenty-five countries around the world have implemented versions of the 5-a-day campaign (Silva et al., 2013).

Several studies have attempted to quantify the campaign’s impact and there is consensus in recognizing a positive impact on awareness, but only a small increase in actual fruits and vegetables (F&Vs) consumption (Bremner et al., 2006; Cullum, 2003; Douarin and Di Falco, 2008). A detailed review of existing evaluations of policies promoting F&Vs consumption in Europe and the US indicated that the average effect on consumption is between 0.2 and 0.6 portion per day (Pomerleau et al., 2005).

Capacci and Mazzocchi (2011) evaluated the impact of the 5- a-day campaign on the food purchase behavior of UK households between 2002 and 2006. They exploited different cross-sections of the Expenditure and Food Survey to identify the change in purchased quantities and controlling for the price effects and other observed confounders of the policy impact. They found that the 5-a-day program lifted F&Vs consumption, on average, by 0.3 portions. They also provided quantitative evidence of a differentiated impact by income group, ranging from 0.2 to 0.7 portions. All impacts were larger than those observed by simply comparing pre-policy and post-policy intakes.

Silva et al., (2013) extended their evaluation framework to identify the consequences of the 5-a-day campaign in France by implementing a dynamic fixed-effect tobit panel data model. Their estimates revealed that the 5-a-day public information campaign led to an increase of 0.38 portions of F&Vs. In addition, the largest increase was observed for the fresh fruits (0.16 portions), the processed vegetables (0.08 portions) and the natural fruit juices (0.08 portions). As a negative nutritional consequence, they also found an increase in fruit drinks with added sugar (0.05 portions). These findings are in line with previous evaluation studies.

Caloric sweetened beverage tax

It is well accepted that excess intake of sugar sweetened beverages (SSBs) has been shown to result in weight gain. To address the growing obesity epidemic, one option is to combine programs that target individual behavior change with a fiscal policy such as taxing SSBs (Escobar et al., 2013). Sugary drinks have raised concerns as some evidence suggests that US children are gaining more calories from drinks than from food. In the US, an economic review found that soft drinks tax resulted in weight loss at different levels for different groups. However, weight loss was generally quite low at current tax rates and insufficient to counter obesity (Fletcher et al., 2010). A tax on caloric sweetened beverages is justified for many reasons. Unlike fast foods, caloric sweetened beverages serve no nutritional value. In addition, empirical evidence showed no indication that such a tax would be regressive and unfairly penalize low income individuals and households. For instance, Escobar et al. (2013) reviewed papers that studied the effect of tax on SSBs on obesity prevalence and its effectiveness in reducing the consumption of SSBs and move it toward more healthy substitutes. They reviewed nine studies: six from USA and only one from France, one from Mexico and one from Brazil. Negative own price elasticity were detected in all the papers (pooled own-price elasticity equals – 1.3), indicating that 10% increase in the price of SSBs could result in about 13% reduction in its consumption. They also found that higher prices for SSBs were associated with an increased demand for alternative beverages such as fruit juice and milk and a reduced demand for diet drinks. Additionally, the studies from USA revealed that a higher price could also lead to reducing the prevalence of overweight and obesity.

Recently, Zhen et al., (2013) studied the effect of SSBs tax in United States using a censored Exact Affine Stone Index incomplete demand system for 23 packaged foods and beverages. Instrumental variables were used to control for endogenous prices. A half-cent per ounce increase in sugar-sweetened beverage prices was predicted to reduce total calories from the 23 foods and beverages but increase sodium and fat intakes as a result of product substitution. The predicted decline in calories was larger for low-income households than for high-income households, although welfare loss was also higher for low-income households.

The main drawback of Zhen et al., (2013)'s articles is that it is mostly concentrated on the effect of intervention policies on a specific food group (e.g. Fruit and vegetable; SSBs; fast food … etc.) which shadow doubts on its reliability as it is omitting the income and the substitution effect and did not take into account the holistic nature of the diet. Gao et al. (2013) tried to fill this gap through applying household production theory to systematically estimate consumer demand for diet quality using the Healthy Eating Index (HEI). Their results indicated that insufficient consumption of food containing dark green and orange vegetables, legumes and whole grains. Moreover, Age and education were found to have a significant impact on consumer demand for diet quality. On the other hand, income did not have a significant effect on the demand for diet quality. They also mentioned that own-price elasticities of demand for diet quality were found to be inelastic. Simulation of tax scenarios revealed that a tax on SSBs may be more efficient than a tax on fats, oils and salad dressing in improving consumer diet quality.

Combining the implementation of such a tax on SSBs and/or fats with and subsidizing targeted unsweetened beverages and/or fruit and vegetable, could have a higher positive effect on enhancing consumer's diet quality. Moreover, this combination offers a more politically viable option as it is expected to have a higher social acceptance and the revenue of the tax could be used in financing the subsidy without any extra burden on the public expenditure.

Another gap observed in the literature is the need to use advanced demand systems (e.g. Exact Affine Stone Index (ESAI)) capable of taking into account the unobserved heterogeneity between individual which is especially important in the case of obesity. Those demand systems are also quite flexible allowing fitting the complexity and multi-dimensional nature of obesity. The recent developed of EASI demand system could be considered more than appropriate in doing so. Zhen (2013) used EASI demand system in analyzing the effect of SSBs tax on obesity prevalence in United States.

Up to our knowledge no published article succeed in dealing with the two aforementioned shortcomings by using flexible demand systems such as EASI in analyzing the demand for diet quality but in a framework that takes into account the holistic nature of the diet. This could be an interesting future research line. We do so in the fifth chapter of this thesis by estimating an EASI demand system to assess the Spanish consumers for diet quality taking into account the person BMI.

Despite all of the ink spilled on this issue, no significant action has been taken yet. The debate on the use of food taxes and subsidies to address obesity should now shift to how best to address practical issues in designing such policies. Existent research evidence largely corroborates the public health case for using economic instruments to the case of discouraging the purchase and use of tobacco and alcohol products. However, the corresponding case for the use of these instruments to encourage healthier eating and physical activity remains controversial from both evidence and policy perspective (Shemilt et al., 2013).

Agricultural subsidy

The role of agricultural subsidies in rising obesity prevalence remains debatable due to the difficulty in isolating the role of specific agricultural policy on food prices. Empirical evidence indicated that agricultural R&D subsidies have increased agricultural output and lowered farm commodity prices, and thereby have had an important impact on consumption and consequently on obesity.

Loureiro and Nayga (2005) estimated the effect of consumer support for the agricultural sector on overweight and obesity prevalence in OECD countries. They find that the lower is overweight and obesity prevalence in the country the higher is the required contributions from consumers (via higher agricultural prices). Cutler et al. (2003) examined the association between the obesity rate and the ratio of agricultural prices in the country to worldwide prices as a measure of producer protection and the frequency of price controls across OECD countries. Both determinants were found to be negatively associated with obesity rates. Vosti (2006) discussed the possible impact of U. S. agricultural subsidies on obesity. The essential argument is that farm subsidies reduce the cost of food and hence encourage over-consumption. Moreover, healthier foods such as fruits and vegetables are not subsidized and are substantially more expensive than calorie-dense foods. Most of the subsidization goes to a few commodities, particularly wheat, corn and cotton.

Income transfer programs

While food-based transfer programs have the effect of improving nutrition for the poor, they might have an adverse effect of encouraging excessive energy consumption that leads to overweight and obesity for adult women.

A number of studies investigated whether food stamps lead to obesity and found mixed results (Baum, 2011; Beydoun et al., 2008; Chen et al., 2005; Fan, 2010; Gibson, 2003 and 2006; Meyerhoefer and Pylypchuck, 2008; Kaushal, 2007; and Ver Ploeg et al., 2007).

7. Future research agenda

An important avenue for future research may be to focus on the limitations of the rational choice model in the case of obesity and ways to deal with it. The use of behavioral and experimental economics techniques and the emerging fields of Genoeconomics and neuroeconomics could play a crucial role in doing so.

Importance of behavioral and experimental economics in studying obesity is related with the controversial role of information and educational programs in combating obesity. Recent advances and application in behavioral and experimental economics shadow some light on this arguable relationship by explaining the role of risk aversion, time preferences and personality traits in the different response of consumers to the same informational and educational content.

While traditional economic analyses of obesity can tell us who is obese, it is incapable of telling us why they are obese or explain heterogeneity within the different demographic groups. On the other hand, behavioral economics techniques could play a vital role in doing so. Incorporating behavioral aspects into the economic analysis of obesity is quite challenging as behavioral factors are determined by complex socio-psychological determinants such as habits, emotions, attitudes, parental feeding practices, beliefs, etc.

Experimental economics

Experimental economics techniques could be a very useful tool in studying obesity. Besides taking into account the behavioral aspects that play a significant role in overweight and obesity prevalence, experimental techniques allow overcoming the problem of the scarcity of secondary data, especially in the case of developing countries, suitable for analyzing economic aspects of obesity. Experimental economics produce a kind of rich reliable primary data.

Ehmke et al. (2008) identified experimental economic tools that can be employed to explain the role of economic behavior in overweight and obesity in the household. For instance, loss aversion experiments were suggested as a tool to understand challenges that some individuals face in achieving a healthy diet. Furthermore, test bed experiments were introduced as a tool to test and understand new policies and incentives for better health.

Kawley et al. (2013) emphasized the importance of natural experiments in determining the extent to which economic variables such as food prices, income and technological change affect the risk of obesity, and estimating the various economic consequences of obesity. This importance is due to scarcity and low quality of data on calories consumed and calories expended.

The biological nature of obesity put emphasis on the use of emerging disciplines such as Genoeconomics and Neuroeconomics wherein biological principals combined with the economic rational offering a unique multidisciplinary tool capable of dealing with the multifactorial complex nature of obesity epidemic.

Food industry and obesity (Market makers)

A clear shortcoming observed in the published studies deal with the economics of obesity prevalence is that, it is concentrating only on the consumer role in combating obesity and forgetting the important role food market makers, including food industry and food retailers, playing in the issue. This approach could be considered misleading as the consumer is the weakest part through the food chain. Swinburn (2011) considered the powerful global food industry as a main culprit responsible for the hiking obesity rates worldwide as it is producing more processed, affordable, and effectively marketed food than ever before. He argues that the increased supply of cheap, palatable, energy dense foods, joined with better distribution and marketing, has led to ‘passive overconsumption’. Problems with the food industry, especially the growth of processed food culture, include: producing recipes high in sugar, salt and fat; large portion sizes; poor nutritional labeling; and Aggressive marketing of unhealthy food, especially to children.

To successfully address obesity, people must consume fewer calories, which means eating less food, or at least different types of food. This implies less industry profit, as the foods most at risk are the most processed, with the highest profit margins, often made by the biggest industry players (Brownel and Warner, 2009). A key option to combat the growing obesity epidemic come through convincing food industry and food retailers that they can make more profit selling healthier food and assure the sufficient incentives to do so.

This problem is aggravated by the lake of awareness among consumers about the role food industry play in the obesity epidemic. Lusk and Ellison (2013) tried to answer the question “how is blame for obesity?”, from the public point of view and trying to identify the determinants of such perceptions, using a nationwide survey among 800 US individuals. Respondents were asked to place each of seven entities (food manufacturers, grocery stores, restaurants, government policies, farmers, individuals, and parents) into three categories: primarily, somewhat, and not to blame for the rise in obesity. Surprisingly, eighty percent said individuals were primarily to blame for the rise in obesity. Parents were the next-most blameworthy group, with 59% ascribing primary blame.

The negative effect of food industry could be reduced through government interventions. Important example of these government interventions is the mandatory food labeling. In June 2011, the European Parliament voted against proposals to introduce ‘traffic-light’ style food labeling regulations. The traffic light system displays a red, amber or green light on the front of food packaging to clearly show consumers how the contents rate in terms of healthy eating criteria. The color is determined by the levels of calories, sugar, salt and fat the product contains. The European Public Health Alliance (EPHA) backed the measure, citing evidence from British and Australian research that found that: ‘a traffic light label on the front of package is the best way to facilitate accurate interpretation of key nutritional information and therefore enable consumers to make informed choices about the food they purchase.’ Studies showed that the traffic light system is understood equally by all consumers, regardless of socio-economic status, gender or ethnicity. This type of labeling is used by some supermarkets in the UK.

Chapter 3
Parametric and Nonparametric Analysis of the Role of Socio-Economic Factors on Obesity Prevalence in Spain

"In the end, your past is not my past and your truth is not my truth and your solution - is not my solution."

Zadie Smith

1. Introduction

Obesity is partly a result of an energy imbalance caused by consumption of too many calories and/or low expenditure of calories (i.e., low physical activity) over a considerable period. Consequently, most published economic research has examined the increased growth of obesity rates by analyzing several factors that may contribute to this imbalance of caloric consumption and usage (Myton et al., 2012; Lin et al.,2011; Cutler et al., 2003; Chou et al., 2004; Lakdawalla and Philipson, 2009; Loureiro and Nayga, 2005; among others).

Due to rising concerns about obesity, the availability, accessibility and choice of foods to meet an adequate diet are becoming key challenges to our food system today. Good nutrition is essential to obtain optimum health and productivity and in reducing the risk of chronic and infectious diseases. Understanding factors influencing food consumption and obesity is needed to gain a clearer picture of the mechanisms that would cause individuals to eat unhealthful or become over weighted; especially, as it is observed that consumers tend to overeat despite quite obvious future health implications (Richards et al., 2007). A better knowledge about how people make food choices and how economic and non-economic factors influence food consumption and obesity is critically important to improve policy interventions and developing agricultural and food programs that can assure a safe, affordable, reliable and nutritious food supply and promote health.

Previous economic studies have analyzed the influence of income on health. In general, there seems to be a consensus about the positive effect of income on health (Smith, 1999). Consequently, we would expect, all things being equal, a negative effect on obesity due to: 1) the unavailability of healthy food in low income neighborhoods (Beaulac et al., 2009; Larson et al., 2009); 2) when healthy food is available it is usually more expensive, so poor families try to stretch their food budget by purchasing unhealthy cheap food (Drewnowski, 2010; Monsivais and Drewnowski, 2009); 3) Lower income neighborhoods have fewer physical activity resources making it difficult to lead a physically active lifestyle (Moore et al., 2008); 4) Low income families are more expected to face high levels of stress increasing the likelihood of being overweight or obese (Gundersen et al., 2011; Moore and Cunningham, 2012). Costa-font et al. (2014) in a recent article studied the effect of income inequalities on unhealthy behaviors in Spain finding that obesity is concentrating in poorer classes in recent decades.

Apart from income, food prices also play an important role in obesity prevalence. Powell and Chaloupka (2009) in their literature review concluded that food prices had a significant but small effect on obesity and overweight prevalence concluding that fiscal pricing policies could help in reversing obesity trends and that small taxes or subsidies were not able to do so while nontrivial pricing interventions might have measurable effects on weight outcomes, especially for those belonging to low social and economic class.

Beside income and food prices, several recent economic studies explain the role played by different cultural and socio-demographic factors on obesity rates. Leaving genetics aside, obesity is caused by consumption of too much calories and/or low expenditures of calories (i.e. low physical activity). For example, Schlosser (2002) showed that the rapid growth of fast food and soda drinks consumption has increased the dietary intake of saturated fats, sugars, and calories and accordingly, the prevalence of obesity. Other researchers argue that female labor participation is a leading factor in increasing obesity rates (Cawley and Liu, 2012; Garcia et al., 2006), mainly in children.

Most of the literature in Spain has concentrated on the adequacy of alternative instruments to measure obesity or on educational and environmental factors (i.e. food consumption) affecting obesity. However, limited attention has been paid to the role of economic factors (income and prices) on food choices, physical activity and, consequently, on the prevalence of obesity.

This study aims at analyzing the relevance of economic factors (mainly income and other socioeconomic characteristics of Spanish households and market prices) on the prevalence of obesity in Spain, which is one of the main contributions of this study.

Data come from the 2011-2012 National Health Survey (NHS) (INE, 2013). Information on most relevant variables is available at household level except for prices. Categorical body mass index (CBMI) has been calculated using the reported weight and height collected from participants.

BMI has been used as a categorical variable to reduce the potential bias in BMI estimates as it is not measured but self-reported (Gil and Mora, 2011). The database has been extended by considering Food at-Home and Out-of-Home prices.

From a methodological point of view, this paper compares the results obtained from the use of parametric and nonparametric models to tackle this issue. While previous literature has focused on parametric methods, such as Multinomial Logistic Regression (MLR), this study also has considered the estimation of a Multivariate Adaptive Regression Splines (MARS) model, which is flexible enough to provide more insight on how covariates interact with the prevalence of obesity.

Results from the MARS model will be used to increase the goodness-of-fit of the traditional parametric approach by transforming independent variables using MARS results allowing nonlinear covariates and interactions. The combined model clearly outperforms parametric approach while being easier to interpret than MARS. This is the second contribution of this paper.

To achieve the paper' objective, the rest of the paper is organized as follows. Section 2 provides a brief description on the obesity prevalence in Spain. The methodological approach applied in our analysis is explained in section 3. Our empirical application and the main results are discussed in sections 4 and 5, respectively. Finally, the paper ends with some concluding remarks.

2. The prevalence of overweight and obesity in Spain

Obesity is considered as a complex, multi factorial, chronic disease involving genetic, prenatal, socioeconomic, dietetic and environmental components. Worldwide prevalence of obesity nearly doubled between 1980 and 2008. In Europe, the regional office of the World Health Organization (WHO) and according to country estimates for 2008, over 50% of both men and women were overweight, and roughly 23% of women and 20% of men were obese. These estimates indicate that obesity prevalence in Europe in the last two decades has tripled affecting more than 150 million adults and 15 million children and adolescents in the region.

In Spain, the last National Health Survey (NHS) for 2011-12 (INE, 2013) indicated that the prevalence of overweight and obesity among Spanish adults aged 18 years old or more was 36.7% and 17.0% respectively. While the prevalence of obesity was quite similar between men and women (18% among men and 16% among women), the overweight prevalence was significantly higher among men (45%) than among women (28%). Obesity is associated with age (Figure 1) as in the age segment between 18 and 24 years old its prevalence merely reaches 5.5% for both genders. On the contrary, in the segment between 65 and 74 years old the prevalence of obesity reaches 25.6% and 27.9% among men and women, respectively (although it is true that it decreases for the eldest segments). There is also a significant negative relationship between education level and obesity. In fact, the highest percentage (30.0%) is found among illiterate persons.

[Please insert figure1 here]

From a historical perspective, it is worth mentioning that, in spite of the up to now relative low percentages in relation to other EU countries, the prevalence of obesity in Spain has increased with a very alarming rate in 25 years moving from 6.9% and 7.9% among men and women, respectively, in 1987, to the above mentioned 18% and 16%, in 2012.

The Spanish National Survey of Dietary Intake (ENIDE) (2011), concluded that obesity rates in Spain was not due to eat too much (daily energy intake was 2482 kcal, slightly lower than the recommended level between 2550 and 2600 calories, depending on the individual's physical activity), but to an unbalanced diet characterized by the overconsumption of red meat, sodas and pastries. According to the Spanish Food Safety Agency (AESA), food habits have changed with a significant reduction of both family meals and the time allocated to eat during weekdays.

Noticeably, the prevalence of obesity has increased during the financial crisis that started to affect Spanish households in 2009 and more intensively during 2010. Comparing the data from the last two National Health Surveys (2006 and 2012), the obesity rate significantly increased from 15.6% to 18%, among males but not as much among women (15.2% and 16%, in 2006 and 2012, respectively). This result has to do with a lower consumption of fresh foods, fruits and vegetables and a higher consumption of fast food, ready-to-eat meals and fatty foods, which have been relatively much cheaper (Rao et al., 2013). This situation seems to continue in the future as the OECD predicted that the number of overweight and obese people in Spain will rise by a further 10 per cent over the next decade.

In fact, an alarming 30 per cent of teenagers are overweight, putting Spain just behind USA and Scotland. Moreover, a stunning 40 per cent of youths aged between 13 and 18 never practice sport.

Although the WHO characterizes overweight and obesity as diseases, it is also well known that both (together with smoking) are key determinants in the incidence of the most important contemporary chronic diseases, such as cancer, cardiovascular problems, certain types of diabetes, etc. What is most worrying is that health disorders that were once almost exclusively associated with the elderly, such as type II diabetes, are now being diagnosed in children, mainly due to the increasing prevalence of childhood obesity. In fact, one in every five adolescents in Spain now runs the risk of suffering major cardiovascular problems in later life (Mora et al., 2012). There is a Positive significant association between obesity and the prevalence of all chronic illness including obesity in Spain (Costa-Font and Gil, 2005).

The economic costs associated with obesity are non-trivial as well. Obesity accounts for 7% of total health care costs (WHO, 2005) without considering other economic externalities, which, on the other hand, are difficult to estimate. In Spain, the Spanish Society for the Study of Obesity (SEEDO) estimates that direct and indirect obesity costs account for 7% of total health care costs (2.5 billion Euros/year).

Mora et al. (2012) investigated the impact of BMI, obesity and overweight on direct medical costs on a longitudinal dataset of medical records for patients followed up over seven consecutive years (2004-2010). They found that one unit increase in individual BMI increases direct medical costs by between 5 and 10€ per patient and year. Similarly, obesity (overweight) increases direct medical costs by between 50 and 96€ (17 and 79€) per patient and year. They concluded that, if half the Spanish population experienced the same BMI increase, then the annual rise in direct healthcare costs would represent around 0.025% of GDP (256 million €).

2. The Multivariate Adaptive Regression Splines model (MARS)

Analyzing main determinants affecting the prevalence of overweight and obesity is not an easy task for two main reasons: 1) this is a complex phenomenon in which a large number of covariates could be considered; and 2) studies have shown that obesity response to socioeconomic covariates is frequently characterized by thresholds requiring flexible response functions (Cavaliere and Banterle, 2008). Moreover, the complexity of interactions between different socioeconomic covariates requires flexible multivariate models capable of dealing with the different ways covariates interact with the dependent variable. Let's take the age as an example. Figure 1 showed the relationship between age and the prevalence of overweight and obesity among men and women. As can be observed the relationship is not linear. Moreover, this nonlinear relationship differs between men and women underling an interaction between gender and age.

The traditional approach to deal with the issue addressed in this paper has been the Multinomial Logistic Regression (MLR). Like any other classical parametric regression methods, MLR is 'global' in nature; that is, when a covariate enters the model all values of such covariate are considered to be relevant in explaining the variation of the dependent variable. This is not the case in reality in which the relationship is true only for certain values of the covariate. Furthermore, in the MLR missing data are either dropped or replaced by mean values reducing the model performance.

The Multivariate Adaptive Regression Splines (MARS), first introduced as a data mining tool (Friedman, 1991), is able to address the above limitations of MLR and other classical regression methods. MARS is a nonparametric method hence, it is expected to perform as well as, or even better than, the classical regression techniques when distributional assumptions are not satisfied. It allows also for local models and thus for a more accurate function approximation. MARS is not affected by any volume of missing data since it automatically introduces indicator functions for every variable that contains missing values. Furthermore, this method is designed to capture higher-order interactions, even in high-dimensional settings. But unlike other available nonparametric methods that can capture complex relationships among the variables such as the Classification and Regression Tree (CART) or Artificial Neural Networks (ANNs), MARS produces very simple and easy-to-interpret models.

MARS performance depends on data structure (Ture et al., 2005) but is generally known for predictive accuracy, computational speed and simplicity of interpretation. Leathwick et al. (2006) compared General Additive Models (GAM) and MARS models and highlighted the advantages of MARS in cases involving large data sets. MARS models are also parsimonious and provide more extensive predictions. Muñoz and Fellicisimo (2004) used two different ecological data sets to compare MARS over other modeling techniques such as MLR, Principal Component Regression and CART observing that MARS performed consistently well. Using a motor vehicle injury data consisting of 59 cases and 689 controls and with up to 3% missing values for some of the variables, Kuhnert et al. (2000) showed that MARS outperformed CART and MLR, in terms of accuracy and flexibility as a modeling tool. Haughton and Loan (2004) compared different statistical techniques to model vulnerability from a panel of 4,272 households. They showed that MARS, together with CART, were the most parsimonious model and were able to capture nonlinearities and interaction effects.

The main advantage of MARS comparing with other regressions such as the logistic regression is that MARS is a data driven technique. Instead of fitting a single regression equation for the model, MARS get many piecewise regression equations which allow the researcher to obtain more consistent and unbiased estimates of the covariates.

The main principle of MARS is based on searching for every point where linearity breaks. These cut-off points of the covariate, where the slope of the line change, are called knots. The Knot defines the end of one domain and beginning of another. Between two knots, a linear (or cubic) regression line is fitted to that range of data. When the slope is not changing along the entire range, no knots are detected and a single linear regression is defined between the covariate and the dependent variable, as in the parametric approach. As mentioned, in MARS the data are left to reveal the variable knot locations while the user need not to input any specification into the model.

Based on knots detected in the process, basis functions are defined to re-express the relations between the dependent variable and its covariates. Basis functions in MARS, which serve as independent variables, are truncated linear functions, which address the problem of discontinuity of recursive partitioning algorithms. To model Basis functions, MARS uses the so called Hinge functions or hockey-stick functions which take the following expression:

$$(X - t_k)_+ = \begin{cases} X - t_k, & \text{if } \mathrm{X} \geq \mathrm{t_k}, \\ 0, & \text{else} \end{cases} \tag{1}$$

where $\mathrm{t_k}$ is a constant called knot.

Dissimilar with additive models, MARS allows interactions up to an order specified by the user, and trades off the interaction order and complexity of the additive functions and interactions (Frank, 1995; De Veaux et al., 1993). Not only piecewise linear functions can be formed from hinge functions, but they can be multiplied among them to form nonlinear functions.

In the case of categorical dependent variables, as in this study, the MARS model deals with this case as a classification problem so it determines a common set of Basis functions in the predictors and estimates different coefficients for each category. This method seems quite similar to some neural networks where multiple outcome variables are predicted from common basis functions with different coefficients.

The MARS model can be written as:

$$y = \sum_{i=1}^{M} \beta_i B_i(X) \tag{2}$$

where, B_i (i=1,2,…,M) are the Basis functions and β_i are the coefficients to be estimated.

MARS is a stepwise process that uses both forward and backward progresses for robust and unbiased parameter estimations. It starts by maximizing all possible effects of explanatory variables in the forward model and then removes the least effective functions in the backward model, using the Ordinary Least Squares method, in order to minimize the so called Generalized Cross Validation (GCV) indicator (Kayri, 2007), given by:

$$\mathrm{GCV} = \frac{\frac{1}{\mathrm{N}}\sum_{i=1}^{N}\left[y_i - \hat{f}_M(X_i)\right]^2}{\left[1 - \frac{d-M}{N}\right]^2} \tag{3}$$

where N is the number of observations, M is the number of Basis functions in the model and $\hat{f}$ denotes the fitted values of the current MARS model. The numerator refers to the common Residual Sum of Squares (RSS), which is penalized by the denominator, which accounts for the increasing variance as the model complexity increases. The penalizing parameter “d” is chosen by the user, although the conventional value is d = 4. A lower (higher) value of d generates a larger (smaller) model with more (less) Basis functions. Thus, the GCV can be considered as a form of regularization by trading off goodness-of-fit against model complexity. In MARS models the RSS cannot be used for comparing models, as the RSS always increases as MARS terms are dropped, which means that if the RSS were used to compare models the backward step of model construction would always choose the largest model.

The main disadvantage of MARS is the low prediction power with insufficient sample size. This is not the case in our analysis as we have a quite big data set which consists of 19069 observations. Moreover, Briand et al. (2007) mentioned that the model might suffer from multicollinearities as MARS gets interactions between predictive variables involved in the model. Also, the MARS methodology has a risk of over fitting because of the very exhaustive search that is conducted to identify nonlinearities and interactions. This drawback could be controlled through choosing the appropriate penalty term of the model.

3. Data and variables definition

Our data come from the 2011-2012 National Health Survey (NHS) (INE, 2013). The NHS is a cross-section survey that provides micro data on the health status of citizens and its determinants. It is carried out by the National Institute of Statistics (INE) in cooperation with the Spanish Ministry of Health, Social Services and Equality. The survey collects information on respondent's socioeconomic characteristics, morbidity, food habits and their demand for health care. Food habits refer to two main issues: type of breakfast and frequency of consumption of selected food groups. However, the data set does not provide information on quantities consumed (or purchased) neither on prices. In order to take into account the effect of economic factors we augment the data set with regional average food prices (FAHP) (MAGRAMA, 2012), the regional price index for Food Out-of-Home (FOHP) relative to the Consumer Price Index for Food (INE, 2012) and the regional per capita expenditure as a proxy of household' disposable income (INE, 2012). Our sample consists of 19069 adults (18 years old or more). Table 1 show some descriptive statistics of the variables used in this study.

[Please insert Table1 here]

To deal with the objective of this paper, we have followed a three-step strategy. Firs, as a benchmark, we have estimated a MRL (Greene, 2003). As a second step, and trying to identify nonlinearities and covariates interactions, a multinomial MARS model has been estimated. MARS estimates will be used in the third step to improve the performance of the parametric MRL. In the following section, we will present the results from each of the three steps.

4. Results

Results from MLR model (Table 2) suggest that all variables except Food Out-of-Home Prices (FOHP) have a significant marginal effect on the probability of being obese. The effect of gender is positive; indicating that being male increases the likelihood to be obese, while the effects of perceived health status[1] and physical exercise are negative. In the case of overweight, gender and age (doing sufficient physical exercise) have a positive (negative) significant marginal effect on its prevalence. These results are quite consistent with those reported in the literature. As mentioned above, however, one of the main shortcomings of the MLR model is that it fails to take into account the potential non-linearity that could exist between the dependent variable and its covariates. For instance, Table 2 suggest that age has a significant positive marginal effect on the prevalence of both overweight and obesity for all age groups while Figure 1 showed that this prevalence decreased from 74 years old onwards. Of course it would be possible to graph each covariate in order to detect nonlinearities but these can be conditioned by other covariates entering into the regression and by interactions among them. The MARS model simplifies this task.

[Please insert Table2 here]

In order to keep the model as simple as possible, interactions of order 2 and a penalty term equaling 4 have been chosen.[2] As it was explained earlier, in the backward step the best model is reached by minimizing the GCV. The optimal model consisted of 11 basis Functions. Table 3 shows the Basis functions estimates for the three BMI categories. As can be observed, the MARS model outperformed the MLR (adjusted R^2= 0.16).

[Please insert Table3 here]

The importance of each explanatory variable is calculated as the square root of the GCV value of a sub model from which all basis functions that involve this variable have been removed, minus the square root of the GCV value of the selected model. Resulting values are re-scaled by giving a value equals 100 for the variable with the largest importance. The variables Age, gender (Male) and perceived health status are the most important in predicting the prevalence of obesity and overweight (Table 4).

---

[1] Although the health status could be considered endogenous to obesity, this is not the case in this paper as not the real but the self-perceived health status is considered. In this context, a self-perception of a good health status could have a negative effect on obesity as this perception could lead people to care less in relation to their diet.

[2] Higher interaction orders (3 and 4) and different penalty terms (2, 3, 5 and 6) were considered. No significant differences were found in terms of Basis functions, knots and variable importance.

Economic factors seem to be less important in explaining obesity and overweight with income being the most important among these economic factors.

[Please insert Table4 here]

As mentioned above, the MARS method is more flexible as, among other issues, it does not impose linearity between the dependent variable and its covariates. In fact, when the relationship is linear, results obtained from the MARS model are quite similar to those obtained in the parametric approach (i.e. Physical exercise), but significant differences have been found when this is not the case. Moreover, we have found more consistency with expected results in the case of variables that have a non-linear relationship (i.e. age). For instance, while in the MLR the age has a significant but very small effect on the prevalence of obesity; in the MARS it is considered the most important variable.

As mentioned above, physical exercise is the only covariate in the MARS model entering without any transformation or interaction with other explanatory variable. As a consequence, both models (MLR and MARS) resulted in quite similar parameter estimates for this variable in terms of sign and magnitude. All other covariates enter the MARS model in a nonlinear way or interacting with other covariates, making this approach a very useful tool to provide better insights to policy makers about specific population segments towards they could address effective policies to tackle obesity.

Although both MARS and MLR indicate that being male increases the likelihood of being overweight and obese, the magnitude of parameter estimates are different as the MLR does not consider interactions between gender and other covariates. A self-perception of a good health status decreases the prevalence of overweight and obesity for the whole sample. However, a self-perception of a bad health status has a negative impact on the prevalence of obesity only among men. This is may be due to the fact that traditionally females are caring more about their appearance regardless their health status while having health problems could lead men to observe more their diets (Case and Menendez, 2009). Similarly, having a healthy breakfast reduces the risk of being obese or overweight only in the case of men.

Regarding economic variables, similar to the MLR model, Food Out-of-Home prices do not have any significant effect on the prevalence of overweight or obesity although this result should be interpreted with caution as we are using regional instead of household prices due to the unavailability of such prices at household level. Food at- Home prices and income seem to have a small significant effect but only on a specific groups of our sample. In fact, higher Food At Home prices have a negative (positive) effect on the risk of being overweight (obese) but only when average food prices exceed 2.14 euros (approximately a value representing 90% of the average sample food prices) and for people younger than 46 years. Income has a very small positive effect on the prevalence of overweight and obesity but only in the case of total per capita expenditure less than 9027 euro (this value represents, approximately 80% of the average sample value).

Finally, while the MLR model fails to capture the nonlinear effect of age, this is not the case for MARS model. As can be observed, age is positively related with the prevalence of overweight and obesity only for women younger than 46. This positive relationship also holds for the whole sample between 46 and 72 years old. These results are more consistent with Figure 1.

As a third step in our study, the Parametric MLR model has been re-estimated allowing for nonlinearities and interactions obtained from the MARS model. The combined model could be considered the best one as it is outperformed MLR in terms the goodness-of-fit ($R^2 = 0.084$) and offering group specific results while having a simple easier to interpret estimates comparing to MARS. Table 5 summarizes estimates and marginal effects of the different variables.

[Please insert Table5 here]

As can be observed, all parameter estimates are significant. In the case of linear covariates such as physical exercise, results from the combined model are exactly the same as in both the MLR and MARS models.

However, in general terms, in the case of nonlinear covariates, results from the combined model are quite similar to those obtained from MARS. Considering the most relevant covariates, the three models considered in this study capture the positive effect of being male on the prevalence of overweight and obesity with magnitudes from the combined model (0.22 and 0.11, respectively) and the MARS model (0.27 and 0.15, respectively) being similar, while those from the MLR model (0.19 and 0.04, respectively) are lower, especially in the case of obesity. In relation to economic factors, while the MLR fails to capture the significant effect of Food At Home prices for a specific group (people younger than 46 years paying average prices over 2.14 euro), the combined model is more flexible to account for such an effect.

5. Concluding remarks

In this paper, we have analyzed the effect of socioeconomic factors on the prevalence of obesity and overweight in Spain. Micro data from the most recent National Health Survey (2011-12) carried out in Spain has been used. The performance of three alternative models has been compared: the Multinomial Logistic Regression (MLR), the Multivariate Adaptive Regression Splines (MARS) and a parametric combination of both. Results suggest that the nonparametric approach clearly outperforms the MLR model as it is more flexible to capture the potential covariate nonlinearities. However, parameter estimates are more difficult to interpret. In any case, this paper has shown that results from the nonparametric model can be used to re-specify the parametric approach allowing for the above mentioned potential nonlinearities, which is the main contribution of this paper. Moreover, it is the first paper which focuses on this issue in Spain. In any case, results from this study strongly depend on data availability. The Survey we have used provides very detail information on health issues. However, the information on food consumption is limited to frequency (but not quantities) and information on prices is missing. Regional averages have been taken into account to tackle with this issue.

Socioeconomic factors seem to have a significant impact on the prevalence of overweight and obesity in Spain. In general, higher food prices, being male and the self-perception of having a bad health status increase the probability of being obese. On the other hand, doing regular physical exercise and having a more complete breakfast decrease the probability of being obese.

Results from this study provide valuable information to help policy makers in designing more influential policies to combat obesity through targeting specific groups of the population. For instance, targeting specific age group such as women younger than 46 years and people between 46 and 72 years could enhance the effectiveness of such policies. Moreover, as males are more affected by obesity and overweight, this raises the need for gender specific policies to combat obesity.

As mentioned, this study has provided an initial look on the effect of socioeconomic factors on the prevalence of overweight in Spain. Moreover, it provides a flexible methodological framework to tackle with this issue. As the main limitation is data availability, further research should be oriented to collect new information. The Spanish National Institute of Statistics (INE) publishes yearly the Household Expenditure Survey using a different panel of households in which information of expenditures and quantities for specific food products is available. Matching both datasets remains a challenge for providing new insights about the effects of socioeconomic factors on the prevalence of obesity in Spain.

Table 1 Descriptive statistics of variables used in this study

Variable	Units	Mean	S. deviation
Categorical BMI (CBMI)	1= normal 2=overweighed 3=obese	1.72	0.74
Age	Years	50.18	18.41
Male	1=male 0=female	0.48	0.50
Food at Home Price (FAHP)	Regional average food prices (Euro)	2.25	0.18
Relative food out of home price(FOHP)	Consumer Price Index of Food Out-of-Home relative to the Global Consumer Price Index for Food by region	0.99	0.02
Income	Total per capita expenditure by region (Euro)	11101.51	1451.88
Physical exercise	1=doing intensive or moderate physical exercise 0=not doing	0.20	0.40
Perceived health status	1= perceive to have a good or very good health 0= perceive not to have a good health	0.70	0.46
Healthy breakfast	1= having breakfast at home 0= otherwise	0.86	0.35

Table 2 Parameter estimates and marginal effects from Multinomial Logistic Regression.

Overweight				
	Estimates		Marginal effects	
Constant	-2.181**	(0.838)		
Age	0.029**	(0.001)	0.005**	(0.000)
Male	1.028**	(0.035)	0.190**	(0.007)
Food At Home Price (FAHP)	0.023	(0.147)	0.037	(0.032)
Food Out-of-Home Prices (FOHP)	0.923	(0.832)	0.148	(0.179)
Income	0.000**	(0.000)	0.000	(0.000)
Physical exercise	-0.365**	(0.043)	-0.040**	(0.010)
Perceived health status	-0.139**	(0.041)	0.010	(0.009)
Healthy breakfast	-0.148**	(0.050)	-0.006	(0.011)
Obesity				
Constant	-0.936	(1.028)		
Age	0.032**	(0.001)	0.003**	(0.000)
Male	0.782**	(0.045)	0.041**	(0.006)
Food At Home Prices (FAHP)	-0.491**	(0.191)	-0.070**	(0.024)
Food Out-of-Home Prices (FOHP)	1.089	(1.030)	0.092	(0.130)
Income	0.000**	(0.000)	0.000**	(0.000)
Physical exercise	-0.848**	(0.066)	-0.084**	(0.007)
Perceived health status	-0.610**	(0.048)	-0.082**	(0.007)
Healthy breakfast	0.405**	(0.062)	-0.050**	(0.009)
R^2	0.074			
AIC	36491			

Note: Standard Error in parentheses

** denotes statistical significance at 5 per cent significance level.

Table 3 Parameter estimates from the MARS model

Coefficients					Variable Involved	Knot Value	Variable Involved	Knot Value
Normal Weight Equation			Overweight Equation	Obesity Equation				
Intercept		0.407	0.323	0.270				
Basis functions	1	0.011	-0.007	-0.004	Age	46		
	2	-0.415	0.269	0.146	Male	0		
	3	0.119	-0.014	-0.105	Percieved Good Health	0		
	4	0.108	-0.038	-0.069	Physical Excercise	0		
	5	0.013	-0.005	-0.009	Age	72	Male	0
	6	0.003	-0.002	-0.001	Age	72	Male	0
	7	0.000	0.000	0.000	Income	9027		
	8	0.114	-0.074	-0.040	Male	0	Percieved Bad Health	0
	9	0.062	-0.003	-0.059	Male	0	Healthy Breakfast	0
	10	-0.019	-0.003	0.022	Age	46	Food at Home Prices	2.14
	11	-0.004	0.003	0.001	Age	46	Female	0

Underlined cells indicate Basis functions of type max (0, independent-knot), otherwise max (0, knot-independent)

Table 4 Variable importance for multinomial MARS model

Variables	Number of Basis functions	Variable importance
Age	5	100
Male	6	60.27
Self-Perception of Good Health	2	38.08
Income	1	27.85
Regular Physical Exercise	1	26.26
Healthy breakfast	1	12.36
Food at Home Prices (FAHP)	1	7.30
Food-out-Home Prices (FOHP)	0	0

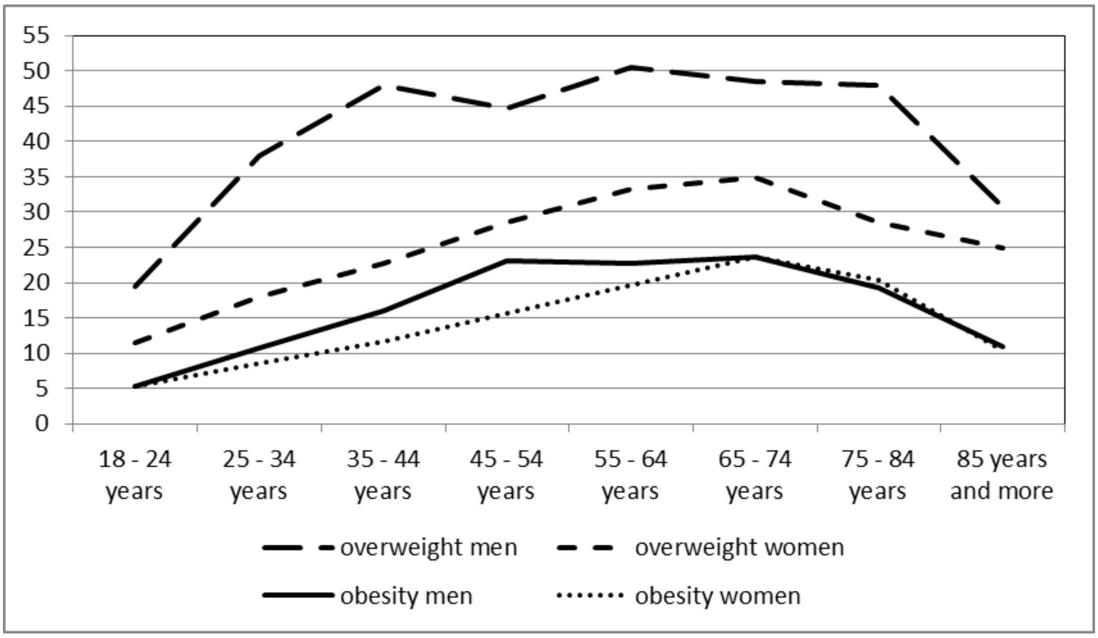
Table 5 Parameter estimates and marginal effects from Multinomial logit model with covariates obtained in the MARS model.

Overweight				
	Estimates		Marginal effects	
Constant	-0.199**	(0.063)		
Male	1.368**	(0.115)	0.223**	(0.023)
Regular Physical Exercise	-0.381**	(0.045)	-0.045**	(0.010)
Self-Perception of Good Health	-0.327**	(0.054)	-0.018	(0.012)
Income < 9027	0.000**	(0.000)	0.000	(0.000)
Age>46	-0.049**	(0.003)	-0.008**	(0.001)
Males with Age<72	-0.011**	(0.003)	-0.002**	(0.001)
Males with Age>72	-0.053**	(0.009)	-0.007**	(0.002)
Females with Age<46	0.019**	(0.002)	0.003**	(0.000)
Age<46 and FAHP>2.14	0.074**	(0.029)	0.009	(0.006)
Males having a Healthy Breakfast	-0.206**	(0.065)	-0.017	(0.014)
Males with Perceived Bad Health	-0.422**	(0.082)	-0.066**	(0.017)
Obesity				
Constant	-0.271**	(0.073)		
Male	1.439**	(0.136)	0.107**	(0.016)
Regular Physical Exercise	-0.846**	(0.067)	-0.081**	(0.006)
Self-Perception of Good Health	-0.834**	(0.064)	-0.101**	(0.010)
Income < 9027	0.000**	(0.000)	0.000**	(0.000)
Age>46	-0.060**	(0.005)	-0.005**	(0.001)
Males with Age<72	-0.015**	(0.003)	-0.001**	(0.000)
Males with Age>72	-0.089**	(0.012)	-0.009**	(0.002)
Females with Age<46	0.015**	(0.003)	0.001**	(0.000)
Age<46 and FAHP>2.14	0.138**	(0.033)	0.014**	(0.004)
Males having a Healthy Breakfast	-0.533**	(0.081)	-0.058**	(0.009)
Males with a Perceived Bad Health	-0.511**	(0.098)	-0.044**	(0.012)
R^2	0.084			
AIC	36091			

Note: Standard Error in parentheses

** denotes statistical significance at 5 per cent significance level.

Figure 1 Overweight and obesity prevalence rates (%) by age group and gender



Source: National Health Survey (2011-12)

Chapter 4
A New Obesity-Specific Healthy Eating Index (OS-HEI)

"No man fills a container worse than his stomach. A few morsels that keep him upright are sufficient for him. If he has to then, he should keep one-third for food, one-third for drink and one-third for his breathing."

Prophet Muhammad (peace and blessings be upon him)

1. Introduction

Obesity is a growing wide spread epidemic all over the world. Overweight and obesity are defined as “abnormal or excessive fat accumulation that may impair health” (WHO, 2013). Many different studies worldwide have shown association between food intake and obesity in both children and adults (Bandini et al., 2004). Promoting a healthy diet could be considered a top priority for many countries all over the world. A key step to do so is measuring the diet quality. Although diet quality is universally recognized as a key determinant of overweight and obesity prevalence, there is still a lack of consensus on how to measure it.

To achieve this objective of being able to assess diet quality, many diet quality indices have been developed during the last decades and some of them used to predict the health outcomes (such as obesity) related to food quality.

In fact most existing indices are able to predict health outcome to some extent, but the associations were generally modest for all dietary scores, casting doubts on their validity. This may be explained by the many arbitrary choices in the development of an index and the lack of insight into the consequences of these choices. The main choices relate to the components to include in the score, the cut-off values to compare intake with, and the exact method of scoring. In addition, diet quality scores may still not adequate to deal with the main reasons for a holistic approach which is the correlations between intakes of various dietary groups (Waijers et al., 2007). This weak association could also be due to a wide variety of health outcomes as some nutrients or foods could be considered determinants only for specific health outcomes while they are irrelevant for other health outcomes. That shadow doubts on the validity of these indices and raise the need to develop disease (obesity) specific diet-quality indices. To avoid subjectivity, we used the Multivariate Adaptive Regression Splines (MARS) in developing our new Obesity Specific - Healthy Eating Index (OS-HEI). The main advantage of using MARS is being a data-driven non-parametric tool, which allows avoiding the subjectivity in choosing index’s items and their cut points. Moreover, MARS allows for interaction between the different items, hence, permitting for a holistic approach that takes into account the interaction between different nutrients and food groups. In this paper, National Health and Nutrition Examination

Survey (NHANES) data set for the year 2007-2008 has been used for the development of our new index. While the same data set for the year 2009-2010 has been used to validate the index and evaluate its ability in predicting the effect of diet quality on obesity prevalence. An association was found between the new OS-HEI and obesity prevalence. This association increases significantly when controlling for some socio-demographics such as age and gender. Moreover, our new OS-HEI was found to outperform notably the HEI-2010 in predicting obesity prevalence.

To achieve the abovementioned objective, the rest of the paper is organized as follows. Section 2 provides a review of diet quality indices and its association with health outcomes, mainly obesity with special emphasize on the Healthy Eating Index (HEI). A description of the data used (NHANES) is given in section 3. The methodological approach applied in the development and validation of our new OS-HEI is explained in section 4. The main results are discussed in section 5. Finally, the paper ends with some concluding remarks.

2. Diet quality indices

Diet quality appears to have no official definition in the literature. Definitions vary widely depending on tools used to measure it. Traditionally, there is a common perception that dietary quality reflects nutrient adequacy. Nutrient adequacy refers to a diet that meets requirements for energy and all essential nutrients (IFPRI, 2002). The worrying rise, all over the world, of overweight and obesity prevalence, which are mainly due to excess of intake of certain nutrients and foods, shifted the definition of dietary quality to include both concepts of nutrient deficiency and over nutrition (WHO/FAO, 1996).

Following Drenowski (2005), in many cases, the only criteria for defining healthful foods is being free of problematic ingredients such as fat, sugar, and sodium and not by the beneficial nutrients they might contain.

Nutrient-dense foods lack a common definition (Guthrie, 1977; Lackey and Kolasa, 2004). Guthrie (1977) showed that there were only limited efforts to define the concept of a nutritious food, having only some general statements which do not depend upon clear standards or criteria.

Lackey and Kolasa (2004) affirmed that there still no agreement on the definition of a nutritious or healthful food and beverage. Past attempts to quantify the nutrient density of foods were based on a variety of calories-to-nutrient scores, nutrients-per-calorie indexes, and nutrient-to-nutrient ratios (Drenowski, 2005).

Assessment of diet quality is concerned with both the quality and variety of the holistic diet, rather than individual nutrients, and allows for the evaluation of how closely eating patterns align with dietary recommendations.

Kourlaba et al. (2009) defined Indices as "composite tools aiming to measure and quantify a variety of clinical conditions, behaviors, attitudes and beliefs that are difficult to be measured quantitatively and accurately".

Diet quality indices add an important dimension to dietary assessment as a composite measure of diet has been preferred to an index of a single nutrient or food in the area of dietary assessment (Kant, 1996; Gerber, 2001). Diet quality indices, or scores, are tools that provide an overall rating, on a numeric scale, of an individual's dietary intake in reference to nutrient and/or dietary recommendations (Wirt and Collins, 2009). In addition to the aforementioned reason, Hu (2002) emphasized the importance of using composite indices to avoid the problem of multi-collinearity (i.e. multi-collinearity occurs when a large number of highly correlated variables (i.e., the components of the index) are considered in the specification of an econometric model) that may lead to less robust estimations of the coefficients and less accurate predictions.

Many indices of overall diet quality were developed during the last decades. Two types of measurement can be identified: predefined diet quality indices, which assess compliance with prevailing dietary guidance as dietary patterns, and empirically derived diet patterns (Kant, 2004).

Predefined or theoretically defined indices consist of nutritional variables, habitually nutrients and foods or food groups that are assumed to be either healthful or detrimental. The index variables are quantified and summed to provide an overall measure of dietary quality. The definition of diet quality depends on attributes selected by the investigator. It is built upon current nutrition knowledge or theory or based on a diet that has proven healthful, such as the Mediterranean diet.

Many predefined diet quality scores have been proposed. Four of them have gained most attention: the Healthy Eating Index (Kennedy, 1995), the Diet Quality Index (Seymour et al., 2003), the Healthy Diet Indicator (Huijbregts et al., 2001), and the Mediterranean Diet Score (Trichopoulou et al., 1995). Many others have been developed carrying out several modifications to these mentioned indices.

Indices differ in several aspects, such as the items included, the cut-off values used, and the exact method of scoring, indicating that many arbitrary choices were made (Nicklas, 2012). This make development of such indices involves a high level of subjectivity.

Indices' components range from nutrients only to adherence to recommended food group servings, to variety within healthful food groups (Wirt and Collins, 2008). Kant (1996) mentioned that the construction of diet quality indices mainly based on food groups or specific foods, or nutrient intakes, or derived from combinations of both foods and nutrient intakes. Waijers et al. (2007) added that scores used national dietary guidelines or a Mediterranean pattern as a reference. Most indices, including the HEI, are based on both food groups and nutrients, while some, such as the HFI, are based on foods and food groups.

Almost all indices included five food groups: vegetables and fruits, cereals or grain, Dairy, and meat and meat products. Regarding nutrients, there is a concord on including fat in the index. Fat is introduced in different forms: total fat, and/or saturated fat (SFA), or the ratio of monounsaturated fatty acids (MUFA) to SFA. Cholesterol and alcohol are also included in many indexes. Moreover, the number of dietary variables used in each dietary index varies. It can be observed that although the vast majority of indices have been constructed using 9 or 10 components, there are some indices developed using less or more components (Waijers et al., 2007).

Once the variables have been selected to be included in an index, they need to be quantified. The simplest method is to use a cut-off value for each index item and assign a score of "0" if consumption is lower than this value (or higher) and “1” if consumption is higher (or lower) than this cut-off point, respectively.

Although very important, the relative contribution of the various index items to the total score has rarely been addressed. In most existing indices, all items contribute equally to the total score. In other words, all components have the same weight. Only few indices have been constructed assigning specific weights to some components (i.e., HEI-2010). Kourlaba et al. (2009) mentioned that, however, these weights were attributed arbitrarily. A key contribution of our paper consisted in using MARS model in weighting our new OS-HEI. This implies that weighting index components is totally objective and depends only upon the role of each component in overweight and obesity prevalence.

Development of diet quality indices include a high level of subjectivity as many arbitrary choices related to the index items or components that should be included in the index, the cut-off values that should be used for each component and the weights that should be assigned to each component take part in the composition of an index.

Empirically derived dietary patterns represents another way of examining dietary patterns using an ‘a posteriori’ approach, in which statistical methods such as factor and cluster analysis are used to generate patterns from collected dietary data.

In factor analysis dietary patterns, the so-called factors are discerned based on correlations between variables, generally foods or food groups. Correlated variables are grouped together. Then each individual is assigned a score on each factor. In contrast to factor analysis, cluster analysis does not aggregate intake variables but individuals into relatively homogenous subgroups (clusters) with similar diets. A summary score for each pattern can be derived and used in either correlation or regression analysis to examine relationships between various eating patterns and the outcome of interest, such as nutrient intake, cardiovascular risk factors, and other biochemical indicators of health (Waijers and Feskens, 2005).

Although factor and cluster analysis could be considered as ‘data-driven’ techniques, a degree of subjectivity exists as choices have to be made in each of the consecutive steps in the analytical process. These steps are somewhat alike for factor and cluster analysis. First, foods or food groups for entry into the analysis need to be selected, foods need to be assigned to food groups, and input variables can (or cannot) be adjusted for example for energy intake. The analysis itself and the identification of the dietary patterns or clusters are not straightforward either and also involve choices. In the majority of factor analysis studies in nutritional epidemiology principal components analysis has been applied, using orthogonal rotation and eigenvalues >1. In cluster analysis studies, K-Means method was most often used, but Ward's Method was also regularly applied. Both the parameters of the resulting factor and cluster solutions and the interpretability as decided on by the researcher determine which solution is finally reported. The number of derived factors reported generally ranged from 2 to 25 and for most studies the total of variance explained by all factors was limited, in general between 15 and 40 percent. The number of resulting clusters varied from 2 to 8. The researcher also gives names to the factors or clusters. And although the factor or cluster loadings are generally reported in the published results, labeling does play a critical role in the interpretation. At this moment there is not yet enough insight in to what extent outcomes are influenced by choices including treatment of the input variables and the factor or clustering method used (Waijers and Feskens, 2005; Chen et al., 2002).

2.1. The Healthy Eating Index (HEI)

The HEI (Kennedy, 1995) is considered one of the most prevalent diet quality indices. It is a summary measure of the overall quality of people's diets. The HEI is an index developed by the U.S. Department of Agriculture (USDA), based on Dietary Guidelines for Americans (1995), as a tool to measure compliance with dietary guidance. The HEI was comprised of 10 components: grains, vegetables, fruits, milk and meat intakes, total fat and saturated fat intakes as a percent of total energy intake, cholesterol and sodium intakes as mg and variety in a person's diet. Scores between 0 and 10 were assigned to each component. Where score 0 indicates non-compliance with recommended amounts or ranges while score 10 indicates intakes close to recommended amounts or ranges. The scores in between these limits were computed proportionally. Total scores are calculated simply by adding the scores assigned to each component (giving equal weight for each component) having an index with values that range between 0 and 100. A total score greater than 80 indicates that the diet is "good," a score between51-80 indicates that the diet "needs improvement," and a score less than 51 indicates that the diet is considered "poor" (USDA, CNPP, 1995).

Since 2000, the HEI has been slightly modified to reflect changes in the Dietary Guidelines (Basiotis et al., 2002). HEI-2005 was developed following the release of the 2005 Dietary Guidelines and in response to the increased emphasis on important aspects of diet quality, such as whole grains, various types of vegetables, specific types of fat and the introduction of the new concept of "discretionary calories" (Guenther et al., 2007). This updated version of HEI consisted of 12 components: total fruit, whole fruit, total vegetables, dark green and orange vegetables and legumes, total grains, whole grains, milk, meat and beans and oils, saturated fat intake as percentage of total energy intake, sodium intake as g/1000 kcal and the calories form solid fat, alcohol and added sugar as percentage of total energy intake. Individual nutrient intakes are first transformed into a base of 1,000 calories for diet groups 1 to 9 and 11. For the nutrient intakes of the first 9 diet groups, the intake of each group is compared with the corresponding recommended intake of that group. If the nutrient intake from a diet group, say Total Fruit meets the recommended quantity, it receives the maximum HEI score for that group. If the nutrient intake for diet group is zero, that group gets a zero HEI score. Intakes between zero and the recommended quantity (maximum level) are scored proportionately. For the 11th diet group, Sodium, the maximum score will be given if the Sodium intake is less than the recommended amount. For diet groups such as Saturated Fat and SoFAAS, the HEI scores are received based on the percentage of energy obtained from those groups to the total energy from food consumption. If the energy from Saturated Fat (the 10th diet group) is less than or equal to 7% of the total energy from the food consumption, the Saturated Fat diet group receives the highest HEI score. If the energy from SoFAAS (the 12th diet group) is less than or equal to 20% of the total energy, the SoFAAS diet group gets the maximum HEI score (for more details, see Guenther et al., 2007; Patricia et al., 2008). While, the minimum HEI score for all diet groups is zero, the maximum HEI scores of different diet groups vary. The first six food groups receive the maximum

HEI scores of 5, the SoFAAS group receives a maximum score of 20, and the rest 5 diet groups receive maximum scores of 10. The total score ranging from 0 to 100, similarly to, the original HEI is simply the sum of all HEI individual scores and can be used to assess the overall diet quality for food.

Publication of the 2010 Dietary Guidelines for Americans provoked a second major update of the HEI. The HEI-2010 holds several characteristics of the HEI-2005: (a) it has 12 components, most of them unchanged, including nine adequacy and three moderation components; (b) it uses a density approach to set standards, per 1000 calories or as a percentage of calories; and (c) it employs less restrictive standards. On the other hand, main Changes to HEI-2005 include: (a) the Greens and Beans component replaces Dark Green and Orange Vegetables and Legumes; (b) specific choices from the protein group reflected through adding the Seafood and Plant Proteins ; (c) Fatty Acids, a ratio of polyunsaturated and monounsaturated to saturated fatty acids, replaces Oils and Saturated Fat to acknowledge the recommendation to replace saturated fat with monounsaturated and polyunsaturated fatty acids; and (d) a moderation component, Refined Grains, replaces the adequacy component, Total Grains, to assess overconsumption (Guenther et al., 2010).

In our paper we used the same components of food groups and nutrients and cut-off points of the HEI-2010, as the last available version of HEI, in developing our novel OS-HEI.

2.2. Low association between diet quality indices and health outcomes

There is a continuing need to examine the relationship between diet quality and health in the population. The traditional approach of investigating the association between single nutrients or food and the risk of related health outcomes, is fraught with problems due to the complexity of people’s diet, the possible correlations in nutrients intake and the possible interactions in the effect of several foods/nutrients (Mertz, 1984). It is widely accepted that individuals do not consume isolated nutrients or foods but complex combinations of foods consisted of several nutrients and nonnutrients (Mertz, 1984). In response to such need, diet quality indices are progressively being used to examine epidemiological associations between dietary intake and nutrition related health outcomes (Wirt and Collins, 2008).

An inverse association of healthful dietary patterns with all-cause mortality and cardiovascular disease risk was reported in several studies. However, the magnitude of risk reduction was found to be modest and was attenuated after control for confounders (Kant, 2004). Diet quality scores were found to be weakly associated with lower risk of cardiovascular disease in men (McCullough et al., 2000 a) but not associated with a reduced chronic disease risk in women (McCullough et al., 2000 b). Furthermore, measures of overall diet quality were found to be associated with biomarkers of chronic disease risk and health outcomes (Zamora et al., 2005; Drewnowski et al., 2009). Large cohort studies, however, often showed conflicting results of the correlation between diet quality scores and chronic disease.

A systematic literature review of 30 observational studies found that the significant association between a diet index score and BMI and obesity were consistently negative (Togo et al., 2001). Other studies, however, did not find significant relationships between diet quality and weight measures (Quatromoni et al., 2002; Villegas et al., 2004). Guo et al. (2004) examined the association between HEI and obesity using a cross-sectional analysis of data from 10930 adults who participated in the Third National Health and Nutrition Examination Survey. They found that a low HEI score was associated with overweight and obesity. However, the overall effectiveness of these guidelines in disease prevention needs to be studied further. In a recent study, Marshall et al. (2014) examined the associations between diet quality and weight status in populations at risk of over-nutrition reviewing 26 studies. They reported that in eighteen of the 26 studies the association between diet quality and weight was found insignificant (Feskanich et al., 2004; Mirmiran et al., 2004; Kranz et al., 2008; Lazarou et al., 2008; Chiplonkar and Tupe, 2010; Kontogianni et al., 2010; Manios et al., 2010; Vitolo et al., 2010; Golley et al., 2011; Jennings et al., 2011). For instance, in a study of French adults, Drewnowski et al., (2009) found that higher diet quality scores were weakly associated with lower Body Mass Index (BMI) and lower blood pressure only for men and were not associated with plasma lipid profiles. Asghari et al. (2012) investigated the performances of the priori dietary pattern including Mediterranean Diet Scale (MDS), HEI-2005 and diet quality index international DQI-I to predict overweight or obesity. No significant relationship between diet quality indices and obesity and abdominal obesity were found, indicating that the ability of diet quality indices to predict obesity and abdominal obesity depends on how well these indices correlate with changes in energy balance as the primary focus in obesity.

Several studies have examined the association between HEI and morbidity. A weak association was detected between HEI score and risk of chronic disease with the exception of cancer risk (McCullough et al., 2000 a and b). Moreover, while Kant et al. (2005) reported that the HEI score was associated with obesity and biomarkers of CVD and diabetes, Fung et al. (2005) found that the HEI is not significantly correlated with any of biomarkers for CVD. Reedy et al. (2008) revealed that the (HEI-2005) is also inversely correlated with colorectal cancer risk.

From the abovementioned literature it can be observed that, the already proposed indices are adequate tools for the evaluation of diet quality, but they have shown moderate predictive ability of chronic diseases and health determinants such as obesity, casting doubts on the validity of the indexes an emphasizing the need for Obesity-specific healthy eating index (OS-HEI).

3. Data: NHANES

The National Health and Nutrition Examination Survey (NHANES) is a program of studies designed to assess the health and nutritional status of adults and children in the United States. The survey is characterized by combining interviews and physical examinations. The NHANES program began in the early 1960s and has been conducted as a series of surveys focusing on different population groups or health topics. In 1999, the survey became a continuous program that has a changing focus on a variety of health and nutrition measurements to meet emerging needs. The survey examines a nationally representative sample of about 10,000 persons each year.

The NHANES interview includes demographic, socioeconomic, dietary, and health-related questions. The examination component consists of medical, dental, and physiological measurements, as well as laboratory tests administered by highly trained medical personnel.

The data from this survey is used to determine the prevalence of major diseases and risk factors for diseases. It is also used to assess nutritional status and its association with health promotion and disease prevention. NHANES findings are also the basis for national standards for measurements such as height, weight, and blood pressure. Data from this survey is also used in epidemiological studies and health sciences research that help developing sound public health policy, direct and design health programs and services, and expand the health knowledge for the Nation.

The data on two-day food consumption and nutrient intakes for 2007-2008 (for the development of OS-HEI) and 2009-2010 (for the validation of OS-HEI) are obtained from the NHANES databases, including data of Dietary Interview of Individual Foods (DIIF) and Dietary Interview of Total Nutrient Intakes (DITN). DIIF provide detailed information on the types (corresponding to USDA food codes) and amount (in gram) of food and beverages consumed by NHANES participants in two days. DITN has the information on individual nutrient intakes based on the data from DIIF and USDA Food and Nutrient Database for Dietary Studies (FNDDS, USDA-ARS, 2006). The FNDDS provides information on nutrient information for each food listed in USDA food codes. The nutrient information helps transform individual food intake to nutrient intake. The total calorie intake from food consumption from DITN is used to transform the nutrient intake from absolute amount into intake per 1,000 calories. The MyPyramid Equivalents Database (Bowman et al., 2008) is used to transform individual food and nutrient intakes into cup or ounce equivalents of diet groups corresponding to those in Dietary Guidelines for Americans, 2010, which helps calculate the HEIs of different diet groups.

4. Development and validation of the Obesity Specific Healthy Eating Index (OS-HEI)

We adapt a semi empirical approach in developing the novel OS-HEI where we used the same food and nutrients groups as in HEI-2010 with the same cut-off points as it is following my pyramid guidelines and then we used the MARS model to determine which food groups should build up our index and the weight of each group.

The process for developing the new Obesity Specific – Healthy Eating Index (OS-HEI) can be summarized in the following steps:

Step 1: Calculating diet scores for each participant in the NHANES 2007-2008 using HEI, and OS-HEI.

We calculate the scores for each component. While these scores in HEI-2010 ranged from 0 to 5, 10 or 20, our scores ranged from 0 to 100 for all components.

We followed the same methodology of calculating scores in HEI-2010 (for details see Guenther et al., 2010)

Step 2: We run MARS model using the OS-HEI individual scores as independent variables and BMI and Waist circumstance as dependent variable to calculate the variable importance which is used as weights for each component.

Step 3: The total OS-HEI is calculated by multiplying each score by its importance and sum all the resulted weighted scores having a total score ranged from 0 to 100.

In the following section we will explain the MARS model.

4.1. Multivariate Adaptive Regression Splines (MARS)

Our paper is the first attempt to use MARS model in developing diet quality indices. MARS model is capable of overcoming the high level of subjectivity involved in the development of predefined diet quality indices. MARS also outperform traditional techniques such as factor analysis and cluster analysis which normally used in developing empirical diet quality indices. Moreover, MARS is capable of taking into account the interactions between the different components of diet quality indices.

The MARS, first introduced as a data mining tool (Friedman, 1991), is able to address the above limitations of factor analysis, cluster analysis and other classical methods. MARS is a nonparametric method hence, it is expected to perform as well as, or even better than, the classical regression techniques when distributional assumptions are not satisfied. It allows also for local models and thus for a more accurate function approximation. MARS is not affected by any volume of missing data, since it automatically introduces indicator functions for every variable that contains missing values. Furthermore, this method is designed to capture higher-order interactions even in high-dimensional settings. But unlike other available nonparametric methods that can capture complex relationships among the variables such as the Classification and Regression Tree (CART) or Artificial Neural Networks (ANNs), MARS produces very simple and easy-to-interpret models.

MARS performance depends on data structure (Ture et al., 2005) but is generally known for its predictive accuracy, computational speed and simplicity of interpretation. Leathwick et al. (2006) compared General Additive Models (GAM) and MARS models and highlighted the advantages of MARS in cases involving large data sets. MARS models are also parsimonious and provide more extensive predictions. Muñoz and Fellicisimo (2004) used two different ecological data sets to compare MARS over other modeling techniques such as MLR, Principal Component Regression and CART. They found that MARS outperformed the other modeling techniques. Using a motor vehicle injury data consisting of 59 cases and 689 controls and with up to 3% missing values for some of the variables, Kuhnert et al. (2000) showed that MARS outperformed CART and MLR, in terms of accuracy and flexibility as a modeling tool. Haughton and Loan (2004) compared different statistical techniques to model vulnerability from a panel of 4,272 households. They showed that MARS, together with CART, were the most parsimonious model and were able to capture nonlinearities and interaction effects.

The main advantage of MARS comparing with other regressions such as the logistic regression is that MARS is a data driven technique. Instead of fitting a single regression equation for the model, MARS get many piecewise regression equations which allow the researcher to obtain more consistent and unbiased estimates of the covariates.

The main principle of MARS is based on searching for every point where linearity breaks. These cut-off points of the covariate, where the slope of the line change, are called knots. The Knot defines the end of one domain and beginning of another. Between two knots, a linear (or cubic) regression line is fitted to that range of data. When the slope is not changing along the entire range, no knots are detected and a single linear regression is defined between the covariate and the dependent variable, as in the parametric approach. As mentioned, in MARS the data are left to reveal the variable knot locations while the user need not to input any specification into the model.

Based on knots detected in the process, basis functions are defined to re-express the relations between the dependent variable and its covariates. Basis functions in MARS, which serve as independent variables, are truncated linear functions, which in turn address the problem of discontinuity of recursive partitioning algorithms. To model Basis functions, MARS uses the so called Hinge functions or hockey-stick functions which take the following expression:

$$(X - t_k)_+ = \begin{cases} X - t_k, & \text{if } \mathrm{X} \geq \mathrm{t_k}, \\ 0, & \text{else} \end{cases} \tag{1}$$

where $\mathrm{t_k}$ is a constant called knot.

Dissimilar with additive models, MARS allows interactions up to an order specified by the user, and trades off the interaction order and complexity of the additive functions and interactions (Frank, 1995; De Veaux et al., 1993). Not only piecewise linear functions can be formed from hinge functions, but they can be multiplied among them to form nonlinear functions.

The MARS model can be written as:

$$y = \sum_{i=1}^{M} \beta_i B_i(X) \tag{2}$$

where, B_i (i=1,2,…,M) are the Basis functions and β_i are the coefficients to be estimated.

MARS is a stepwise process that uses both forward and backward progresses for robust and unbiased parameter estimations. It starts by maximizing all possible effects of explanatory variables in the forward model and then removes the least effective functions in the backward model, using the Ordinary Least Squares method, in order to minimize the so called Generalized Cross Validation (GCV) indicator (Kayri, 2007), given by:

$$\mathrm{GCV} = \frac{\frac{1}{\mathrm{N}}\sum_{i=1}^{N}\left[y_i - \hat{f}_M(X_i)\right]^2}{\left[1 - \frac{d-M}{N}\right]^2} \tag{3}$$

where N is the number of observations, M is the number of Basis functions in the model and $\hat{f}$ denotes the fitted values of the current MARS model. The numerator refers to the common Residual Sum of Squares (RSS), which is penalized by the denominator, which accounts for the increasing variance as the model complexity increases. The penalizing parameter "d" is chosen by the user, although the conventional value is d = 4. A lower (higher) value of d generates a larger (smaller) model with more (less) Basis functions. Thus, the GCV can be considered as a form of regularization by trading off goodness-of-fit against model complexity. In MARS models the RSS cannot be used for comparing models, as the RSS always increases as MARS terms are dropped, which means that if the RSS were used to compare models the backward step of model construction would always choose the largest model.

The main disadvantage of MARS is its low predictive power with insufficient sample size. This is not the case in our analysis as we have a large data set of 9000 observations. Moreover, Briand et al. (2007) mentioned that the model might suffer from multicollinearity as MARS gets interactions between predictive variables involved in the model. Also, the MARS methodology has a risk of over fitting because of the very exhaustive search that is conducted to identify nonlinearities and interactions. This drawback could be controlled through choosing the appropriate penalty term of the model.

The importance of each explanatory variable is calculated as the square root of the GCV value of a sub model from which all basis functions that involve this variable have been removed, minus the square root of the GCV value of the selected model.

5. Results

Table 1 summarizes the descriptive statistics for the individual scores of both the HEI-2010 and OS-HEI. For both indices the minimum for all scores is zero. The maximum for all components in the case of OS-HEI is constant and equals 100 while in the case of HEI-2010 the maximum varies across components being 5 for Total vegetable, Greens and Beans, Total fruit, Whole fruit, Total protein food and Seafood and plant protein; 10 for Whole grain, Dairy, Fatty acids, Refined grains and sodium; and 20 for empty calories components.

The lowest mean score is observed for greens and beans for both indices with a value of 0.94 and 18.76 for HEI-2010 and OS-HEI respectively. The highest mean score is detected for empty calories (total protein food) with a value of 10.97 (76.31) in the case of HEI-2010 (OS-HEI).

[Please insert Table 1 here]

Table 2 shows the Basis functions estimates of MARS model with the BMI as a dependent variable and the 12 individual scores of the OS-HEI as explanatory variables. In order to keep the model as simple as possible, interactions of order 2 and a penalty term equals 4 have been chosen.[3] As it was explained earlier in the backward step, the best model is reached by minimizing the GCV. In our case, the optimal model consisted of 14 basis Functions beside the intercept (adjusted R^2= 0.09).

[3] Higher interaction orders (3 and 4) and different penalty terms (2, 3, 5 and 6) were considered. No significant differences were found in terms of Basis functions, knots and variable importance.

[Please insert Table 2 here]

Results of the MARS model seems consistent and with the expected sign. For instance, the first Base function indicates that for those with a high consumption of dairy with a very high score greater than 99.7 increasing the consumption of dairy products increase also the BMI index which indicates that the cut point for dairy consumption is quite suitable for analyzing obesity prevalence. While for most of the participants with dairy score less than 99.7, the increase of the individual quality score of dairy consumption by 10% resulted in 0.3 reduction in the BMI. The total protein quality does not affect the BMI for those with a score less than 83 while it increases BMI by around 1.3 for 10 points increase in the total protein individual quality index. In the case of Fruit, almost for all participants with individual quality score greater than 3.5, the increase of fruit individual quality index by 10 points decreases the BMI by 0.3.

Our findings show that all components have a combined effect resulted from the interaction with other diet components beside its sole effect on obesity prevalence. This emphasizes the importance of taking into account the interaction between the different components of the index.

It is worth mentioning that we re-estimated the same model using the waist circumference as a dependent variable instead of BMI with the aim of checking the consistency of our results when different measures of obesity are used. This model resulted in results quite similar to those of the BMI model indicating that the findings are consistent regardless of the way of measuring the weight status.[4]

The importance of each explanatory variable is calculated as the square root of the GCV value of a sub model from which all basis functions that involve this variable have been removed, minus the square root of the GCV value of the selected model. Individual scores for dairy, total fruit, total protein and total vegetables are the most important in predicting the prevalence of obesity and overweight. Whole fruit, whole grains and empty calories individual scores also play a role in predicting obesity (Table 3). These results seem to be consistent with the literature. Dairy was found to reduce risk for several chronic diseases, including osteoporosis,

[4] Results of this model are available upon request from the authors.

hypertension, obesity and type 2 diabetes (Pereira et al. 2002; Zemel and Miller, 2004; Choi et al. 2005). Authors and experts used to suggest to those who are obese or at risk of obesity to eat more fruits and vegetables (Khan et al., 2009). The association of meat consumption with health can be described as U-shaped. In moderate quantities they are assumed to be beneficial. However, their intake should not be too high, as high consumption levels are considered unfavorable (Waijers et al., 2007). Cereal products are not the only foods that provide dietary fiber. Moreover, the health effect of whole grains is not attributed to fiber alone, but also to the micronutrients antioxidants and non-nutritive dietary constituents such as phytoestrogens in wheat bran, and beta-glucans in oats (King, 2005). Because of that, whole grain products should be distinguished from refined grains in diet quality indices.

[Please insert Table 3 here]

The other five individual scores seem to be irrelevant in predicting obesity. Some of these scores such as Sodium consumption have no importance in predicting obesity because there is no direct link that can be identified between sodium consumption and obesity prevalence. Other components have no importance in predicting obesity although its consumption could affect obesity prevalence such as refined grains may be because its effect is overlooked by the greater effect of similar components such as whole grains. This is highlight that, while in many cases healthy foods are defined by the absence of problematic ingredients such as fats, sugar, and sodium rather than the presence of any beneficial nutrients that might contain (Drewnowski, 2005). Our results suggest that more attention should be given to beneficial nutrients.

To get the OS-HEI total score that ranges from 0 to 100, each individual score is multiplied by the variable importance. The summary statistics for the total scores of HEI-2010, OS-HEI, BMI and waist circumference using NHANES database (2007-2008) are displayed in Table4. It can be observed that the mean value of OS-HEI is 10 points greater than HEI-2010 which indicates that HEI-2010 is underestimating the diet quality in terms of obesity. Also, the observed higher standard deviation for OS-HEI compared with HEI-2010 indicates that OS-HEI is more capable of capturing variability in intakes of food and nutrients regarding obesity prevalence.[5]

[Please insert Table 4 here]

Table 5 presents weights for the different individual scores of HEI-2010 and OS-HEI. It can be observed that while only four groups which are dairy, total fruit, total protein and total vegetable represents around 80% of the components important in OS-HEI, it represents only 20% of the weight of HEI-2010. On the other hand, empty calories component has importance of 20% in HEI-2010 and only 5% in OS-HEI. This emphasizes the importance of using MARS in weighting the different components as an objective tool in doing so.

[Please insert Table 5 here]

5.1. Validation of the OS-HEI

We validate the OS-HEI using the 2009-2010 NHANES data base. Then we score each individual's diet in NHANES (2009-2010) to determine their HEI-2010 and OS-HEI scores. After that, we estimate the correlation between the BMI and waist circumference as measures of obesity prevalence and HEI-2010 and OS-HEI as indices of diet quality. The results are displayed in Table 6. While no significant correlation detected between HEI-2010 and the two obesity prevalence measures, a negative significant correlation of - 0.053 and – 0.066 between OS-HEI and BMI (-0.053) and OS-HEI and waist circumference (-0.066) indicating that increasing diet quality as measured by OS-HEI by 10% could decrease the obesity prevalence by approximately 0.5%. This low magnitude of the correlation coefficient is expected as many other factors beside the diet quality play a role in obesity prevalence.

[Please insert Table 6 here]

Then we estimated a MARS model using HEI or OS-HEI as the predictor of BMI (Waist circumference) to determine the correlation between HEI (OS-HEI) and obesity prevalence. He results are displayed in

[5] Same differences in terms of both mean and standard deviation are detected using the 2009-2010 NHANES databases.

Table 7. The HEI-2010 failed in predicting the obesity prevalence with an adjusted R^2 equals 0.000. While the OS-HEI was capable of predicting obesity prevalence in both cases with an adjusted R^2 equals 0.01. This low value of the adjusted R^2 is expected as many other factors besides diet quality play a vital role in obesity prevalence. For instance, controlling for age and gender, considered as the most important determinants in obesity prevalence (Chapter 3), increases the adjusted R^2 to 0.39. This result put emphasis on the importance of designing gender and age group specific diet quality indices which is one of our future research lines.

[Please insert Table 7 here]

Although it can be argued that total calories intake could be used instead of OS-HEI, single nutrient or food group analysis omits the synergistic nature of whole diet. Knowledge was shown to be a stronger predictor of overall diet quality than of any single nutrient or diet quality. Keeping with the doctrine that there are no good or bad foods only bad diets (Guthrie, 1986); measures of nutritional quality should focuses on total diets and not on single foods or nutrient. A ‘holistic’ approach would be more realistic though, as people have diets (i.e. they do not consume nutrients but combinations of foods).

Furthermore, calories labeling was found to be inefficient in reducing consumption of less healthy foods and enhancing diet quality in many studies (Ellison et al., 2013; Elbel et al., 2009)

Moreover, we estimated the MARS model with the 12 individual scores and total calories intake with the aim of constructing an energy adjusted diet quality index. Surprisingly our results indicated that total calories intake has a significant effect on the obesity prevalence while empty calories become insignificant. Also the importance of total calories in this model and the empty calorie in the other one with only components are quite similar which indicates that empty calorie component is capturing the total calorie intake of individuals. To keep our OS-HEI comparable with the HEI-2010, we decide not to include the total calorie intake in our model as its effect is already reflected by the empty calories component.[6]

[6] Results of this model are available from the author upon request.

6. Concluding Remarks

In this paper, we have developed a new Obesity Specific Healthy Eating Index (OS-HEI). Data from the National Health and Nutrition Examination Survey (NHANES) data set for the year 2007-2008 has been used for the development of our novel index. The NHANES data set for the year 2009-2010 was used to validate the index and evaluate its ability in predicting the effect of diet quality on obesity prevalence. In order to avoid shortcomings of previous diet quality indices, we followed a semi empirical approach using the Multivariate Adaptive Regression Splines (MARS) in developing the OS-HEI to avoid subjectivity in choosing the food groups to be considered into the index, its weight and cut points. A high association was found between the new OS-HEI and obesity prevalence. Moreover, our new OS-HEI outperformed notably the HEI-2010 and total calories consumed in predicting obesity prevalence. While HEI-2010 includes 12 components, OS-HEI includes only 7 components since seafood and plant protein, fatty acids refined grains and sodium were assumed not to have an effect on obesity prevalence. While the weight of food groups in HEI-2010 was found to be equals 5, 10 or 20, food group weights in the OS-HEI ranged from 5% for empty calories and 22% for dairy. This study has provided an initial look on the development and validation of disease (obesity) specific diet quality indices offering an OS-HEI as an alternative index capable of predicting obesity prevalence. Moreover, it provides a flexible methodological framework to develop other disease specific diet quality indices avoiding subjectivity in doing so. The capability of OS-HEI in predicting obesity increased significantly by using it jointly with age and gender. This shows the need to develop gender and age group specific healthy eating indices, especially in the case of obesity.

Table 1 Summary statistics for individual scores of HEI-2010 and OS-HEI using NHANES database (2007-2008)

Index Components	HEI-2010				OS-HEI			
	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
Total vegetable	2.46	1.71	0	5	52.85	34.21	0	100
Greens and Beans	0.94	1.81	0	5	18.76	36.12	0	100
Total fruit	2.49	2.12	0	5	49.83	42.47	0	100
Whole fruit	2.14	2.29	0	5	42.96	45.81	0	100
Whole grain	1.93	2.86	0	10	19.32	28.60	0	100
Dairy	5.29	3.61	0	10	52.88	36.15	0	100
Total protein food	3.82	1.57	0	5	76.31	31.50	0	100
Seafood and plant protein	1.55	2.05	0	5	30.95	40.98	0	100
Fatty acids	4.57	3.56	0	10	45.67	35.60	0	100
Refined grains	5.25	3.64	0	10	52.52	36.39	0	100
Sodium	5.59	3.69	0	10	55.93	36.94	0	100
Empty Calories	10.97	6.37	0	20	54.84	31.86	0	100

Table 2 Parameter estimates from the MARS model

Bases functions	Coefficients	Variable Involved	Knot Value	Variable Involved	Knot Value
0	22.5758				
1	-2.5663	Total Dairy	<u>99.72</u>		
2	0.0327	Total Dairy	99.72		
3	0.1033	Total Protein	<u>83.08</u>		
4	-0.0326	Total Protein	83.08		
5	-0.0256	Total Fruit	<u>3.51</u>		
6	0.5278	Total Fruit	3.51		
7	0.0002	Total Vegetable	<u>10.93</u>		
8	0.0020	Total Vegetable	10.93	Total Fruit	<u>3.51</u>
9	-0.0042	Whole Fruit	5.13	Total Fruit	<u>3.51</u>
10	0.0207	Whole Grain	<u>9.73</u>	Total Dairy	99.72
11	0.0918	Whole Grain	9.73		
12	0.0152	Total Vegetable	<u>0.00</u>		
13	0.0007	SOFAAS	<u>31.23</u>	Total Protein	83.08
14	0.0009	SOFAAS	31.23	Total Protein	83.08

Underlined cells indicate Basis functions of type max (0, independent-knot), otherwise max (0, knot-independent)

Table 3 Variable importance for MARS model

Index components	Number of Basis functions	Variable importance (%)
Total Dairy	3	22
Total Fruit	4	21
Total Protein	4	20
Total vegetables	3	15
Whole Fruit	1	10
Whole Grains	2	7
Empty Calories	2	5

Table 4 summary statistics for total scores of HEI-2010 and OS-HEI, BMI and waist circumference using NHANES database (2007-2008)

Variable	Mean	Std. Dev.	Min	Max
Total score (HEI-2010)	47.18	13.90	0	100
Total score (OS-HEI)	53.68	18.46	0	100
BMI	25.72	7.58	12.50	73.43
Waist circumference	87.64	22.24	37.80	178.20

Table 5 weights for individual scores of HEI-2010 and OS-HEI

Index components		HEI-2010 (%)	OS-HEI (%)
1	Total vegetable	5	15
2	Greens and Beans	5	0
3	Total fruit	5	21
4	Whole fruit	5	10
5	Whole grain	10	7
6	Dairy	10	22
7	Total protein food	5	20
8	Seafood and plant protein	5	0
9	Fatty acids	10	0
10	Refined grains	10	0
11	Sodium	10	0
12	Empty Calories	20	5

Table 6 correlation and significance estimates between obesity prevalence (BMI and waist circumference) and diet quality measured by HEI-2010 and OS-HEI

Variables	BMI		Waist Circumference	
	Correlation	significance	Correlation	significance
HEI-2010	-0.003	0.748	0.005	0.640
OS-HEI	-0.053	0.000	-0.066	0.000

Table 7 Goodness of fit estimates for the different MARS models

Explanatory variable	BMI		Waist Circumference	
	GCV	Adjusted R^2	F value	Adjusted R^2
HEI-2010	59.53	0.00	492.53	0.00
OS-HEI	59.10	0.01	487.35	0.01
OS-HEI controlling for Age and Gender	36.12	0.39	200.65	0.59

Chapter 5
Obesity and Food Demand in Spain: Can Price Intervention Policies Improve Diet Quality

"Let food be thy medicine and medicine be thy food"

Hippocrates

1. Introduction

Poor diets and rising obesity rates dominate the current food, nutrition and health policy debate in many countries including Spain. Obesity is partly a result of an energy imbalance caused by the excessive consumption and/or low expenditures (i.e., low physical activity) of calories over a considerable period of time. Consequently, most published economic research has examined the increased growth of obesity rates by analyzing several factors that may contribute to this imbalance of caloric consumption and usage (see Cutler et al., 2003; Chou et al., 2004; Lakdawalla and Philipson, 2007; Philipson and Posner, 1999; Loureiro and Nayga, 2005).

Due to rising concerns about obesity, the availability, accessibility and choice of foods to meet an adequate diet are becoming key challenges to our food system today. Good nutrition is essential to obtaining optimum health and productivity and in reducing the risk of chronic and infectious diseases (Eastwood et al. 1984; Vining 2008). Understanding the factors influencing food consumption and obesity is needed to gain a clearer picture of the mechanisms that would cause individuals to eat unhealthful or become overweight. Hence, knowledge about how people make food choices and how economic and non-economic factors influence food consumption and obesity is critically important to improve policy interventions and developing agricultural and food programs that can assure a safe, affordable, reliable and nutritious food supply and promote health.

Although some papers have analyzed the effect of economic or non-economic factors on the prevalence of obesity, up to our knowledge, none of the published papers on obesity attempted to quantify the effect of being obese or not, measured by the Body Mass Index (BMI), on the consumption patterns of the different food categories. This relation is quite important as obese individuals appear to make systematically different food choices relative to others (Richards and Hamilton, 2012). Understanding why obese people make different choices is essential to developing a reasoned policy approach to obesity. Then, the main objectives of this paper is to: 1) analyze the effect of obesity, as an approximation for diet quality, on the consumption patterns of the main food categories among Spanish consumers; 2) assess the differences between obese and non-obese consumers in their consumption patterns; and 3) study the impacts of taxing (subsidizing) different food groups on diet quality and consequently on obesity.

Only few studies assessed consumer demand for diet quality due to the unavailability of food prices and adequate measures of diet quality (Gao et al., 2013). In contrast to Gao et al. (2013), we used the BMI as an approximation for diet quality instead of using the Healthy Eating index (HEI) or any other index of diet quality. The rationale behind this decision is the weak associations between diet quality index and health outcomes such as obesity (Waijers et al., 2007). As obesity is mainly a result of diet quality and physical activity, by assuming physical activity to be constant BMI could be considered a better approximation for diet quality allowing us to avoid the low association problem.

In the estimation of our demand system we used a unique data base of hypothetical average consumers. Those hypothetical consumers were obtained through the merge of the Spanish National Health Survey and the Spanish Household Budget survey for the year 2012.

Bonnet (2013) argued that, the failure of public information campaigns have in reversing the mounting trend in obesity, lead economists to support food taxes to potentially force individuals to change their eating behavior and make the agro-food industry think more about healthy food products. Our paper is considered the first attempt to evaluate the effect of such food taxes (subsidies) on food demand in Spain.

Furthermore, our paper stands out by being one of the first empirical applications of the new and promising demand system, EASI (Lewbel and Pendakur, 2009). In our paper we estimate an EASI demand system for the full diet of the Spanish consumers incorporating the weight status of the individual approximated by the Body Mass Index (BMI). The main advantage of EASI is the derivation of Implicit Marshallian demands, which allow benefiting from desirable features of both Hicksian and Marshallian demands. Similar to the Almost Ideal Demand system (AIDS), budget shares in EASI are linear in parameters. In deference with AIDS, EASI can, however, have any rank and its Engel curves can have any shape over real expenditures. Moreover, EASI error terms can be interpreted as unobserved preference heterogeneity. In this study, price, income, BMI, age and gender elasticities were calculated and the results were used to assess the potential impact of market intervention policies on food demand and the prevalence of obesity in Spain.

To attain our objective, the rest of the paper is organized as follows. Section 2 provides a brief description on the obesity prevalence and food consumption in Spain.in section 3, a brief review of the literature on food price intervention policies presented. The methodological approach applied in the analysis is explained in section 4. Our empirical application and the main results are discussed in sections 5 and 6, respectively. Finally, the paper ends with some concluding remarks.

2. Obesity prevalence and food consumption in Spain

Obesity is considered as a complex, multi factorial, chronic disease involving genetic, prenatal, dietetic, socioeconomic, and environmental components. Worldwide prevalence of obesity nearly doubled between 1980 and 2008. In Europe, the regional office of the World Health Organization (WHO) estimated that, in 2008, over 50% of both men and women were overweight, and roughly 23% of women and 20% of men were obese. These estimates indicate that obesity prevalence in Europe in the last two decades has tripled affecting more than 150 million adults and 15 million children and adolescents in the region.

In Spain, the last National Health Survey (NHS) for 2011-12 (INE, 2013) indicated that the prevalence of overweight and obesity among Spanish adults aged 18 and over was 36.7% and 17.0%, respectively. While the prevalence of obesity was quite similar between men and women (18% among men and 16% among women), the overweight prevalence was significantly higher among men (45%) than among women (28%). Obesity was found to be associated with age. In the age segment between 18 and 24 years old, its prevalence merely reached 5.5% for both genders. On the contrary, in the segment between 65 and 74 years old the prevalence of obesity reached 25.6% and 27.9% among men and women, respectively (although it is true that it decreases for the eldest segments). There was also a significant negative relationship between education level and obesity. In fact, the highest percentage (30.0%) was found among illiterate persons. Also significant differences were found among geographical location and urbanization. For instance, the prevalence of obesity was found to be more important in Galicia, Andalucía and the Canary Islands and in rural areas.

From a historical perspective, it is worth mentioning that, in spite of the up to now relative low percentages in comparison with other EU countries, the prevalence of obesity in Spain has increased with a very alarming rate over the last 25 years moving from 6.9% and 7.9% among men and women, respectively, in 1987, to 18% and 16%, in 2012.

The Spanish National Survey of Dietary Intake (ENIDE) (2011) concluded that obesity rates in Spain was not due to eating too much (daily energy intake was 2482 kcal, slightly lower than the recommended level between 2550 and 2600 calories, depending on the individual's physical activity), but to the unbalanced diet characterized by the overconsumption of red meat, sodas and pastries. Furthermore, according to the Spanish Food Safety Agency (AESA), food habits have changed with a significant reduction of both family meals and the time allocated to eat during weekdays which lead to a higher consumption of unhealthy calorie dense poor nutritious foods.

Noticeably, the prevalence of obesity has increased during the financial crisis that started to affect Spanish households in 2009 and more intensively during 2010. Comparing the data from the last two National Health Surveys (2006 and 2012), the obesity rate significantly increased from 15.6% to 18%, among males but not as much among women (from15.2% to 16%). This result has to do with a lower consumption of fresh foods, fruits and vegetables and a higher consumption of fast food, ready-to-eat meals and fatty foods, which have been relatively much cheaper (Rao et al., 2013). This situation is likely to continue in the future as the OECD predicted that the number of overweight and obese people in Spain will rise by a further 10 per cent over the next decade.

According to data from the Household National Expenditure Survey for 2012, the share for food and nonalcoholic beverages represents 14.71% of the total expenditure by the Spanish household with an average annual expenditure of 4140 (1617) euros per household (person). The average budget shares of different food groups in relation to total food expenditure were: bread and cereals, 15%; meat, 25%; milk and dairy products, 12%; fruits and vegetables, 19%; fish, 13%.; fat and vegetable oils, 4% sugar and sweets, 4% ; and, finally, other food, 8%. However, important family differences appear in relation to certain household characteristics. In larger towns, households spend a relative higher percentage of their expenditure on fish, fruits and vegetables and meat, while the consumption of cereals and potatoes, dairy products and vegetable oils are lower. In relation to the education level, it is interesting to note that as the level of education increases, the relative importance of the consumption of cereals and potatoes and vegetable oils diminishes, being more significant in the first case. On the opposite side, higher education levels were found to be associated with higher budget shares allocated to meat, fish and fruits and vegetables.

In general, households with children had a higher budget share for cereals and potatoes, meat and dairy products. On the other hand, the percentage allocated to vegetable oils and fruits and vegetables was higher in one-person households and in households without children. In relation to the age of the head of the household, there existed a positive relationship between age and the consumption of fruits and vegetables and vegetable oils, while younger households were associated with higher budget shares allocated to cereals and potatoes, meat and fish. Finally, no big differences were found when accounting for the sex of the head of the household.

3. Food prices intervention policies

Dealing with the alarming and growing public health burden of obesity need a comprehensive and a well-designed combination of regulatory, educational agricultural and of course economic policies. Without any doubt economic interventions by themselves are not the magic solution for the obesity puzzle but should be considered one of the most important components of such integrated approach.

Mazzocchi and Traill (2006) classify the wide range of potential instruments available to public authorities in four groups according to their expected impacts on economic agents: 1) policies addressed to change consumer utility function; 2) those aimed at a better-informed choice without changing the utility function; 3) market measures addressed to affect actual choices without changing the utility function; and 4) supply-side policies affecting food availability. As shown by these authors, the number of potential alternatives is very large and, at the same time, they are very heterogeneous in nature, which, on the other hand, merely reflects the complexity of the problem and the number of factors influencing dietary habits and intakes (individuals' socioeconomic characteristics and lifestyles). Moreover, it is also true that food policies addressed to the emerging nutrition challenges need to coexist with agricultural and trade policies, which have traditionally regulated the agro-food activities with very different objectives. Such coexistence may reduce the effectiveness and complicate the implementation of some of the instruments.

Bonnet (2013) argued that, the failure of public information campaigns have in reversing the mounting trend in obesity, lead economists to support food taxes to potentially force individuals to change their eating behavior and make the agro-food industry think more about healthy food products.

Faulkner et al. (2011) conducted a scoping review with the aim of synthesizing existing evidence regarding the impact of economic policies targeting obesity and its causal behaviors (diet, physical activity). Their results proposed that, consistent evidence that weight outcomes are responsive to food and beverage prices and the relatively modest impact any specific economic instrument would have on obesity independently. Shemilt et al. (2013) conducted a scoping review analyzing the use of economic instruments to promote dietary and physical activity behavior change. Shemilt et al. (2013) defined economic instruments as " it is encompasses fiscal or legislative government policies designed to change the relative prices of goods or services or people's disposable income, and promotional practices used by retailers to change the relative prices of goods and services".

A tax on food, especially junk, has been suggested. People have developed unhealthy habits in response to the low price of food, especially calorie dense foods, and the low relative price of driving a car for transportation. However, these habits could be changed by altering the structure of economic incentives on which people base their decisions. Economic policies could be used to create incentives to both reduce excess calorie consumption and excessive reliance on the automobile (Senauer and Gemma, 2006). Modeling a tax rate similar to that on tobacco products, of approximately 58%, the researchers found a greater weight loss would be achieved. The acceptability of taxes at this level is not Clear (Fletcher et al., 2010). Imposing such high tax level in the case of food could be considered a quite challenging task as eating is both an absolute necessity and intrinsically healthy, whereas smoking have been shown to pose serious health risks. Because of that, a direct tax on food, even on high calorie foods low in nutrients, for the purpose of reducing obesity is not politically feasible. Additionally, a tax on food is regressive, since those with lower incomes spend a large share of their budget on food. A solution for this social and political unacceptability could be solved through combining such food tax with a food subsidy on healthy food which assure reducing the effect on the poor and granted a higher political acceptance.

Lakdawalla and Philipson (2001) inspect how food taxation may impact BMI. They find that higher food taxation (relative to the taxation of other goods) increases the price of food, and this increase in the relative food price will decrease BMI. Schroeter et al. (2008) simulate the impact that changes in taxes or subsidies for food and drinks would have on body weight finding that a tax on food away from home or a subsidy of vegetables and fruits would increase body weight while a tax on regular soft drinks or a subsidy of diet soft drinks would lower body weight. The result that an increase in the tax on food away from home would increase the body weight appears surprising. However, the tax on food away from home does in fact decrease away-from-home food consumption, but it increases at home food consumption (e.g. processed food) due to the fact that the two categories are substitutes. Because many of the foods consumed at home are energy- rich, total consumption actually increases. Etilé (2008) estimates for French household survey data the impact of a 10% price decrease in fruits and vegetables on body weight and the impact of a 10% price increase in soft drinks, pastries, deserts, snacks and ready-meals on body weight. He then simulates the impact of five policy scenarios (taxes and subsidies) on the overweight and obesity prevalence. The authors find that with every of the five policy scenarios the distribution of BMI in the population is clearly more favorable in a public- health sense.

It is well accepted that, Excess intake of sugar sweetened beverages (SSBs) has been shown to result in weight gain. To address the growing obesity epidemic, one option is to combine programs that target individual behavior change with a fiscal policy such as excise tax on SSBs (Escobar et al., 2013). Sugary drinks have raised concerns as some evidence suggests that US children are gaining more calories from drinks than from food. In the US, an economic review found that soft drinks tax resulted in weight loss at different levels for different groups. However, weight loss was generally quite low at current (low, mean 3.3%) tax rates and insufficient to counter obesity (Fletcher et al., 2010). A tax on caloric sweetened beverages is justified for many reasons. Unlike fast foods, caloric sweetened beverages serve no nutritional value. In addition, empirical evidence shows no indication that such a tax would be regressive and unfairly penalize low income individuals and households. In a recent paper, Escobar et al. (2013) reviewed papers studied the effect of tax on SSB on obesity prevalence and its effectiveness in reducing the consumption and move it for more healthy substitutes. Nine articles were studied, six from USA and only one from France, Mexico and Brazil. Negative own price elasticity detected in all papers (pool own price elasticity equals – 1.3). Indicating that 10% increase in the price of SSBs could result in about 13% reduction in its consumption. Studies showed also that, higher prices for SSBs were associated with an increased demand for alternative beverages such as fruit juice and milk and a reduced demand for diet drinks. Additionally, studies from USA revealed that a higher price could also lead to reduce the prevalence of overweight and obesity.

Recently Zhen et al., (2013) studied the effect of SSBs tax in United States using a censored Exact Affine Stone Index incomplete demand system for 23 packaged foods and beverages. Instrumental variables are used to control for endogenous prices. A half-cent per ounce increase in sugar-sweetened beverage prices is predicted to reduce total calories from the 23 foods and beverages but increase sodium and fat intakes as a result of product substitution. The predicted decline in calories is larger for low-income households than for high-income households, although welfare loss is also higher for low-income households.

A main drawback of the mentioned articles is that it is mostly concentrate on the effect of intervention policies on a specific food group (e.g. Fruit and vegetable; SSB; fast food … etc.) which shadow doubts on its reliability as it is omitting the income and the substitution effect and did not take into account the holistic nature of the diet. Gao et al. (2013) tried to fill this gap through applying household production theory to systematically estimate consumer demand for diet quality using the Healthy Eating Index (HEI). Their results indicated that, consumers have insufficient consumption of food containing dark green and orange vegetables, legumes and whole grains. Moreover, Age and education found to have a significant impact on consumer demand for diet quality. On the other hand, income does not have a significant effect on the demand for diet quality. They mentioned that own-price elasticities of demand for diet quality found to be inelastic. Simulation of tax scenarios revealed that a tax on SSB may be more efficient than a tax on fats, oils and salad dressing in improving consumer diet quality.

Combining the implementation of such a tax on SSBs and/or fats with targeted unsweetened beverages and/or fruit and vegetable subsidies, could have a higher positive effect on enhancing consumer's diet quality. Moreover, this combination offer a more politically viable option as it is expected to have a higher social acceptance and the revenue of the tax could be used in financing the subsidy without any extra burden on the public expenditure.

Another gap observed in the literature is the need to use advanced demand systems capable of taking into account the unobserved heterogeneity between individual which is especially important in the case of obesity. Those demand systems also should to be quite flexible allowing fitting the complexity and multi-dimensional nature of obesity. The recent developed Exact Affine Stone Index (EASI) could be considered more than appropriate in doing so. Zhen et al. (2013) used EASI demand system in analyzing the effect of SSBs tax on obesity prevalence in United States.

Up to our knowledge no published article succeed in dealing with the two aforementioned shortcomings by using flexible demand systems such as EASI in analyzing the demand for diet quality but in a framework taking into account the holistic nature of the diet. This could be an interesting future research line. We do so in the fifth chapter of this thesis by estimating an EASI demand system to assess the Spanish consumers for diet quality taking into account the person BMI. This draws another novelty of our paper.

4. Theoretical and econometric frameworks

In the neoclassical economic theory, consumer's utility is specified as a function of quantities of goods and services purchased, assuming that consumer's tastes and preferences and health status are constant, which is an assumption that does not always hold in the real world. In this context, incorporating consumers' health status in the derived demand function can be misleading as the demand itself is derived from assuming tastes and preferences to be constant. Only in the case where the effect is temporary, we could accept introducing such information as a demand shifter. However, if the effect is more permanent, we should modify the consumer's maximization problem.

Philipson and Posner (2003), in their seminal paper, modeled weight status as the result of both eating and exercise decisions using a utility maximization framework where Individuals tradeoff the disutility from excess weight with the enjoyment of eating and relaxing lifestyle, subject to a budget constraint.

Let us assume that we have n food categories and that consumers are maximizing their utility from consuming these n food categories, taking also into account their health status which we are approximating by the Body Mass Index (BMI). The consumer' optimization problem may be stated as:

$$Max\ U = U\ [x, h(BMI)]$$
$$S.t.\ p'x \leq y \tag{1}$$

To solve this maximization problem we define the following Lagrange function:

$$L = U\ [x,\ h(BMI)] + \lambda(y - p'x) \tag{2}$$

where λ is the Lagrange multiplier. By deriving (2) with respect to x and λ, we get the first order conditions from which the Marshallian demand functions are obtained.

$$x^m = f\,(p, y, h(BMI)) \tag{3}$$

where the demand of each product depends on total expenditure, prices and consumer' BMI.

The dual cost minimization problem is:

$$\min p'x + \mu(u - U(x,h)) \tag{4}$$

where μ is the Lagrange multiplier.

The Marshalian demand function can be written as

$$x^m = f(p, y, h(BMI)) \Rightarrow x^m = f(p, y, BMI) \tag{5}$$

Capturing the obesity impact on food consumption requires the specification of the demand system shown in (6). The Deaton and Muellbauer's (1980) Almost Ideal Demand System (AIDS) has been widely used due to its desirable characteristics. It is a plausible demand system, easy to estimate and the imposition of theoretical restrictions is straightforward.

In spite of its desirable characteristics, a main shortcoming of AIDS model is that the Engel curves are assumed to be linear in real expenditures. To avoid this problem and keeping the estimation process simple, we estimate an EASI demand system for the full diet of the Spanish consumers incorporating the weight status of the individual approximated by the Body Mass Index (BMI). The main advantage of EASI is the derivation of Implicit Marshallian demands which combine desirable features of both Hicksian and Marshallian demands. Similar to AIDS model, EASI budget shares are linear in parameters given real expenditures. However, EASI is superior in that the derived demands can have any rank and the Engel curves can have any shape over real expenditures. Moreover, EASI error terms can capture the unobserved preference heterogeneity among consumers.

Equation 6 can be rewritten following the cost function proposed by Lewbel and Pendakur (2009) which is particularly convenient for empirical estimation.

$$\ln C(p,y,z,\varepsilon) = y + \sum_{j=1}^{J} m^j(y,z)\ln p^j + \frac{1}{2}\sum_{j=1}^{J}\sum_{k=1}^{J} a^{jk}(z)\ln p^j \ln p^k + \frac{1}{2}\sum_{j=1}^{J}\sum_{k=1}^{J} b^{jk}\ln p^j \ln p^k y + \sum_{j=1}^{J}\varepsilon_j \ln p^j \quad (6)$$

Where y is the implicit utility function and corresponds to an affine function of the Stone index deflated by the log nominal expenditures, P is the price vector, z is a vector of demographic variables (includes BMI, age and gender) which proxy observable preference heterogeneity and ε a vector of error terms which include unobservable presences heterogeneity.

The implicit Marshalian budget shares for each $j \in 1...J$ is then given by:

$$w^j = \sum_{r=1}^{R} b_r^j y^r + \sum_{t=1}^{T} g_t^j z_t + \sum_{k+1}^{T}\sum_{t=1}^{T} a_j k_t z_t \ln p^k + \sum_{k=1}^{J} b_{jk} \ln p^k y + \sum_{t=2}^{T} h_t^j z_t y + \varepsilon_j \quad (7)$$

Where y is defined as follows:

$$y = \frac{\ln x - \sum_{j=1}^{J} w_j \ln p^j + \frac{1}{2}\sum_{j=1}^{J}\sum_{k=1}^{J} a_{jkt} z_t \ln p^j \ln p^k}{1 - \frac{1}{2}\sum_{j=1}^{J}\sum_{k=1}^{J} b_{jk} \ln p^j \ln p^k} \quad (8)$$

Marshalian price, income and demographic elasticities are calculated following Lewbel and Pendakur (2009).

5. Data and empirical application

Data availability is the main limitation to carry out any economic analysis which relates obesity, food consumption and food prices. Currently, in Spain, there are two main sources of secondary data related with this issue. The first one is the National Health Survey (NHS). The NHS is a cross-section survey that provides ample data on the health status of citizens and its determinants. It is carried out by the National Institute of Statistics (INE) in cooperation with the Ministry of Health and Consumption. The survey collects information on the individual socioeconomic characteristics, morbidity, food habits and the demand for health care. Food habits refer to two main issues: type of breakfast and frequency of consumption of selected food groups. However, the data set does not provide information on quantities consumed (or purchased) neither on prices. The second main source that mainly refers to consumption data is the Spanish Household budget Survey. This survey provides annual information on the expenditure and quantity consumed of various classes of food products consumed by a stratified random sample of around 24,000 households. Since prices are not explicitly recorded, unit values for each group are calculated dividing expenditures by quantities. The survey also gathers information on a limited number of household characteristics including the level of education and main activity of the head of the household, household income, household size, age and sex of family members and town size, among others. As similar characteristics are found in both data sets, our methodological approach has consisted of merging the two databases by defining different segments of population, using the socioeconomic characteristics of households. Then for each segment we calculated average values for the relevant variables contained in each database.

We made these segments using geographical variables (province and district size) and socio-demographic characteristics (age and gender). First, we divided each sample into 52 subsamples. Each subsample represented a province of the 52 Spanish provinces. Data for each province were divided into five segments of districts depending upon the district size. These five segments are: districts with more than 100,000 habitants; districts with 50,000 to 100,000 habitants; districts with 20,000 to 50,000 habitants; districts with 10,000 to 20,000 habitants; and districts with less than 10,000 habitants. After that, each subsample was divided using the socioeconomic characteristics (age and gender) of the family head. Age and gender were chosen since our analysis, in chapter 3 of this thesis, revealed that age and gender were found to be the most determinant factors of obesity prevalence. Using gender as a segmentation criterion, each subsample was divided into two subsamples: one included males family heads while the other one included female family heads. In the case of age, we created four groups: family head with less than 25 years; family head between 25 and 45 years; family head between 45 and 65 years; and family head older than 65 years. The cutoff points used to obtain the four age groups were determined based on the cut points (knots) obtained from our analysis in chapter 3. By doing so, we could get a set of about 2080 segments (52 provinces multiplied by 5 district size groups multiplied by 2 gender groups multiplied by 4 age groups). However, due to the absence of some segments in at least one of the original data sets, the final data set included 753 hypothetical representative average consumers.

For each segment, the constructed dataset provided micro data on expenditure, prices and the BMI together with some socio-demographic and economic characteristics such as age and gender. Food products were aggregated into the following eight food groups: Carbohydrates (includes grain and potatoes); Protein (includes meat, fish, egg and legumes); Dairy; Fat; Oils; Fruits and Vegetables; Sweets; and other foods (includes food and nonalcoholic beverages not included in the previous groups). Unlike previous studies we separate oils from fat because most of the oils consumed in Spain are olive oil which is perceived as a very healthy product while other fats are perceived as unhealthy products.

Since prices are not explicitly recorded, unit values for each group are calculated dividing expenditures by quantities. Unit values may reflect not only spatial variations caused by supply shocks (i.e., transportation costs, cost of information, seasonal variations, etc.) but also differences in quality which can be attributed to brand loyalty or marketing services among other factors. Then, unit values were adjusted following Gao et al. (1997). The quality-adjusted price is defined as the difference between the unit price and the expected price, given its specific quality-related characteristics.

6. Results

Table 1 presents the descriptive statistics of the demographic variables, shares and prices of the eight food groups. It can be observed that protein and carbohydrates are the two biggest food groups with shares of around 38% and 17%, respectively, of the total food expenditure. On the other hand fat and oil are the two smallest food groups with shares of only 0.22% and 2.6%, respectively. The lowest average price was observed for the dairy group (1.67 euro) while the highest price corresponded to other foods group (10.86 euro).

[Please insert Table 1 here]

Then, the EASI model given in (8) was estimated using BMI, age and gender as demographic variables[7]. The equation corresponding to "other foods" group was dropped in order to avoid singularity. It is well known that the demand system is invariant to which equation is dropped and the parameters of the dropped equation are derived from the adding up conditions. The system is estimated by an iterated 3SLS. This estimator is similar to the estimator suggested by Blundell and Robin (1999) for the QUAIDS.

The most important economic information in demand systems is provided by elasticities. Expenditure, price and demographic elasticities were calculated following Lewbel and Pendakur (2009). The estimated conditional (as we assumed weak separability) expenditure, own-price and demographic (BMI, age and gender) elasticities are presented in table 2. As can be observed, all expenditure elasticities are positive and statistically significant (except for fat and other foods). Sweets (1.26), oil (1.15) and fat (1.12) are considered luxury products. The elasticities of fruit and vegetables, protein and other foods groups were found to be not significantly different from 1; while carbohydrates (0.92) and dairy (0.91) can be defined as necessity.

[Please insert Table 2 here]

All own-price elasticities are negative and inelastic except in the case of dairy (-1.01) and fat (-1.05). The food groups with the most inelastic demand are oil (-0.20) and carbohydrate (-0.33). This could be explained by the fact that Spanish consumers have strong positive preferences for olive oil and bread which represent the main food products in oil and carbohydrates groups. This inelastic price elasticity is consistent with the nature of food products and that Spanish consumer, as all developed countries; spend a small share of their income on food.

Most demographic elasticities are statistically non-significant and with a quite small magnitude. Increasing the age, increases slightly the consumption of protein and fruits and vegetables while decreases the

[7] Estimation results are presented in annex1.

consumption of carbohydrates and sweets. This emphasized that the elder are consuming healthier diet may be due to that they are watching more their diets to avoid health consequences which is more observed among old people. In respect to gender being a male increases the consumption of protein while decreases consumption of sweets. The BMI has a small positive effect on oil consumption and on the other hand it decreases the consumption of healthy foods such as fruit and vegetable.

After estimating the parameters and calculating the elasticities, we simulated price change scenarios by modifying the level of indirect taxes (Value Added Tax (VAT)). In Spain, VAT is set at 21% for most products. However, necessities like bread, flour, eggs, milk, fruits and vegetables, legumes, potatoes and grains are taxed at 4%, while for the rest of food products the VAT is 10%. Different price scenarios were simulated: 1) Decreasing taxes on healthy products such as Fruits and Vegetables (from 4% to 1%); 2) Increasing taxes on sugary unhealthy products such as sweets (from 10% to 21%); 3) increasing taxes on fats (from 10% to 21%); 4) increasing taxes on both sugary and fatty products as fat and sweets at the same time (from 10% to 21%); 5) decreasing taxes on fruits and vegetables (from 4% to 1%) and increase taxes on fat (from 10% to 21%) at the same time; 6) decreasing taxes on fruits and vegetables (from 4% to 1%) and increase taxes on sweets (from 10% to 21%) at the same time; 7) decreasing taxes on fruits and vegetables (from 4% to 1%) and increase taxes on both fat and sweets (from 10% to 21%) at the same time. In all scenarios, we assumed that the food supply is competitive and that there are not specialized inputs (i.e. marginal and average costs remain constant). Under these assumptions, any price change will be fully passed forward to consumers. Own- and cross-price elasticities were used to make the simulations assuming that total food expenditures remain constant.

Results of the different simulation scenarios are presented in table 3. It worth mentioning that, in the case of sole intervention policies, the smallest effect observed in the case of subsidizing fruit and vegetable. Even a tax on fat or fat and sweets could resulted in a higher consumption of fruit and vegetable than the increase obtained through subsidizing fruit and vegetable, due to substitution effect. Generally a tax on fat has a higher impact, in motivating healthy consumption, than similar tax on sweets. Fat tax has an unintended effect on protein consumption. Although mixed intervention policies has a quite similar effect to taxing unhealthy products (e.g. fat; sweets), these mixed intervention policies could have higher social acceptability.

[Please insert Table 3 here]

7. Concluding remarks

This paper is one of the first attempts to analyze the relationship between BMI and food consumption patterns and how food price intervention policies can improve diet quality and combat obesity prevalence in Spain. We attain this objective through merging two data sets which were initially designed to collect information for different purposes. The methodological approach followed here consisted in the specification and estimation of an EASI demand system from which expenditure; price, BMI and demographic variables elasticities were calculated. Our results indicate that changes in the BMI have a quite small positive (negative) effect on oil (fruit and vegetable) consumption in Spain. Taxing unhealthy products (e.g. fat and sweets) found to have higher impact on food consumption than subsidizing healthy foods (e.g. fruit and vegetable). Mixed intervention policies could have a higher social acceptance. Further research is needed in order to calculate the effect of changes in calorie consumption and consequently on BMI. Examining experimentally the social acceptance of the different price intervention policies is considered as an interesting future research line.

Table 1 Structure of food expenditure in Spain by socio-economic groups (2006)

Variable	Mean	St. Deviation	Min.	Max
Age	48	18.81	18	84
Gender	0.6	0.48	0	1
BMI	26.02	2.39	18.22	36.13
Share (carbohydrate)	16.81	3.79	3.95	46.55
Share (Protein)	38.04	5.98	8.16	56.74
Share (dairy)	13.27	3.06	4.05	25.39
Share (fat)	0.22	0.24	0	2.23
Share (oil)	2.6	1.75	3.91	29.23
Share (fruit and vegetables)	13.84	3.23	1.99	31.98
Share (sweets)	8.48	3.08	1.67	23.29
Share (other food)	6.75	2.33	18.22	36.13
Price (carbohydrate)	2.27	0.41	0.17	3.64
Price (Protein)	7.04	1.14	1.61	11.35
Price (dairy)	1.67	0.34	0.79	3.5
Price (fat)	2.39	0.5	0.53	5.37
Price (oil)	9.73	29.96	0.86	49.47
Price (fruit and vegetables)	1.59	0.28	0.52	2.68
Price (sweets)	1.75	0.61	0.63	5.79
Price (other food)	10.86	5.78	2.93	50.98

Source: Encuesta Continua de Presupuestos Familiares (INE) and own elaboration

Table 2 Calculated own-prices, expenditure and demographic elasticity

Marshallian own-price elasticity	
Carbohydrate	-0.33 (0.02)
Protein	-0.45 (0.04)
Dairy	-1.01 (0.02)
Fat	-1.05 (0.00)
Oil	-0.20 (0.01)
Fruit and Vegetables	-0.87 (0.02)
Sweets	-0.79 (0.01)
Other	-0.76 (0.51)
Expenditure elasticity	
Carbohydrate	0.92 (0.09)
Protein	1.00 (0.07)
Dairy	0.91 (0.11)
Fat	1.12 (0.72)
Oil	1.15 (0.36)
Fruit and Vegetables	0.97 (0.11)
Sweets	1.26 (0.16)
Other	1.01 (0.61)
Demographic elasticity	
BMI	
Carbohydrate	0.0009(0.0009)
Protein	-0.0009 (0.001)
Dairy	0.000 (0.0008)
Fat	-0.000 (0.000)
Oil	0.001 (0.000)
Fruit and Vegetables	-0.002 (0007)
Sweets	0.0004 (0.0007)
Other	0.000 (0.0022)
Age	
Carbohydrate	-0.0003(0.0001)
Protein	0.0008 (0.0002)
Dairy	-0.0001 (0.0001)
Fat	0.000 (0.0000)
Oil	0.000 (0.000)
Fruit and Vegetables	0.0007 (0.0001)
Sweets	-0.0007 (0.0001)
Other	-0.000 (0.0003)
Gender	
Carbohydrate	0.003 (0.003)

Protein	0.019 (0.005)
Dairy	-0.001 (0.0032)
Fat	-0.000 (0.000)
Oil	-0.000(0.000)
Fruit and Vegetables	-0.005 (0.0030)
Sweets	-0.006 (0.0027)
Other	-0.000 (0.0086)

Note: Standard Error in parentheses

Table 3 Impact of price intervention policies on quantities consumed

Food Groups		Subsidizing healthy foods	Taxing unhealthy foods			Mixed intervention policies		
		Fruit and Vegetable	Sweets	Fat	Sweet and Fat	Fruit and Vegetable + Sweet	Fruit and Vegetable + Fat	Fruit and Vegetable + Sweet and Fat
Change in the quantity of	Carbohydrate	0.24	-2.61	-6.60	-9.21	-6.36	-2.38	-8.98
	Protein	0.53	-3.81	-19.91	-23.72	-19.38	-3.28	-23.19
	Dairy	-0.34	0.78	6.52	7.30	6.18	0.43	6.96
	Fat	-0.03	0.21	-10.54	-10.33	-10.56	0.19	-10.35
	Oil	0.07	-1.21	2.63	1.41	2.69	-1.15	1.48
	Fruit and Vegetables	2.60	2.19	5.41	7.60	8.01	4.79	10.20
	Sweets	-0.48	-7.89	8.27	0.39	7.80	-8.36	-0.09
	Other	0.33	-0.35	3.08	2.74	3.42	-0.02	3.07

Annex 1. Estimates of the EASI demand system

Equations and variables	Estimate	Std. Error	t value	Pr(>|t|)
eq1_Constante	0.1576	0.0164	9.58	0
eq1_y^1	0.1121	0.0633	1.77	0.0769
eq1_y^2	-0.069	0.0369	-1.87	0.0618
eq1_y^3	-0.0906	0.0818	-1.11	0.268
eq1_y^4	0.2779	0.1129	2.46	0.0139
eq1_y^5	0.278	0.1524	1.82	0.0682
eq1_age	-0.0003	0.0001	-3.56	0.0004
eq1_gender	0.0033	0.0028	1.17	0.2439
eq1_BMI	0.001	0.0007	1.33	0.1826
eq1_y*age	0	0.0003	-0.15	0.8812
eq1_y*gender	0.0191	0.0104	1.83	0.0675
eq1_y*BMI	-0.0048	0.0028	-1.7	0.0884
eq1_pCarbohydrate	0.1228	0.0415	2.96	0.0031
eq1_pProtein	-0.0889	0.0424	-2.09	0.0363
eq1_pDairy	0.0305	0.028	1.09	0.2763
eq1_pFat	0.0022	0.0033	0.64	0.5196
eq1_pOil	-0.0173	0.0187	-0.92	0.3564
eq1_pFV	-0.0212	0.0287	-0.74	0.4592
eq1_pSweet	-0.0092	0.0212	-0.43	0.6652
eq1_y*pCarbohydrate	-0.1966	0.0375	-5.24	0
eq1_y*pProtein	0.2222	0.0446	4.99	0
eq1_y*pDairy	-0.0485	0.028	-1.74	0.0825
eq1_y*pFat	0.0044	0.0031	1.41	0.1599
eq1_y*pOil	0.0004	0.0179	0.02	0.9828
eq1_y*pFV	0.0228	0.0284	0.8	0.4226
eq1_y*pSweet	-0.0066	0.0224	-0.29	0.77
eq1_age*pCarbohydrate	-0.0002	0.0008	-0.32	0.7482
eq1_age*pProtein	0	0.0008	0.02	0.9805
eq1_age*pDairy	-0.0002	0.0005	-0.43	0.6685
eq1_age*pFat	-0.0001	0.0001	-1.17	0.2435
eq1_age*pOil	-0.0002	0.0004	-0.52	0.6011
eq1_age*pFV	0.0002	0.0005	0.37	0.7084
eq1_age*pSweet	-0.0002	0.0004	-0.48	0.6327
eq2_Constante	0.3534	0.0251	14.08	0
eq2_y^1	-0.1192	0.0972	-1.23	0.2203
eq2_y^2	0.0086	0.0596	0.14	0.8858

eq2_y^3	-0.1006	0.1264	-0.8	0.4263
eq2_y^4	-0.0388	0.1992	-0.19	0.8458
eq2_y^5	0.1705	0.2582	0.66	0.509
eq2_age	0.0008	0.0001	6.05	0
eq2_gender	0.0185	0.0043	4.32	0
eq2_BMI	-0.0009	0.0011	-0.83	0.4074
eq2_y*age	-0.0003	0.0005	-0.64	0.5195
eq2_y*gender	-0.0357	0.016	-2.24	0.0253
eq2_y*BMI	0.0064	0.0043	1.48	0.1397
eq2_pCarbohydrate	-0.0889	0.0424	-2.09	0.0363
eq2_pProtein	0.2589	0.0787	3.29	0.001
eq2_pDairy	-0.0228	0.0408	-0.56	0.5769
eq2_pFat	-0.0139	0.0044	-3.16	0.0016
eq2_pOil	-0.0432	0.0252	-1.71	0.087
eq2_pFV	-0.0295	0.04	-0.74	0.4597
eq2_pSweet	-0.0652	0.0302	-2.16	0.0311
eq2_y*pCarbohydrate	0.2222	0.0446	4.99	0
eq2_y*pProtein	-0.3367	0.1035	-3.25	0.0012
eq2_y*pDairy	0.0137	0.0539	0.25	0.7994
eq2_y*pFat	0.0095	0.0064	1.49	0.1374
eq2_y*pOil	0.001	0.0351	0.03	0.9778
eq2_y*pFV	0.0001	0.0556	0	0.9982
eq2_y*pSweet	0.0467	0.0415	1.13	0.2599
eq2_age*pCarbohydrate	0	0.0008	0.02	0.9805
eq2_age*pProtein	-0.001	0.0014	-0.72	0.4745
eq2_age*pDairy	-0.0006	0.0008	-0.81	0.4202
eq2_age*pFat	0.0002	0.0001	2.43	0.0153
eq2_age*pOil	0.001	0.0005	2.21	0.0273
eq2_age*pFV	0.0001	0.0007	0.1	0.9208
eq2_age*pSweet	0.0009	0.0006	1.52	0.1288
eq3_Constante	0.133	0.0147	9.04	0
eq3_y^1	-0.0161	0.0571	-0.28	0.7775
eq3_y^2	-0.1059	0.0344	-3.08	0.0021
eq3_y^3	0.0458	0.0743	0.62	0.5372
eq3_y^4	0.2215	0.1127	1.96	0.0495
eq3_y^5	0.0444	0.1481	0.3	0.7642
eq3_age	-0.0002	0.0001	-2.05	0.0407
eq3_gender	-0.0016	0.0025	-0.63	0.5268
eq3_BMI	0.0005	0.0007	0.8	0.4264

eq3_y*age	0.0003	0.0003	0.88	0.3815
eq3_y*gender	0.0065	0.0094	0.7	0.4855
eq3_y*BMI	-0.0007	0.0025	-0.27	0.7856
eq3_pCarbohydrate	0.0305	0.028	1.09	0.2763
eq3_pProtein	-0.0228	0.0408	-0.56	0.5769
eq3_pDairy	-0.0534	0.0371	-1.44	0.1507
eq3_pFat	0.0023	0.0034	0.69	0.4927
eq3_pOil	0.035	0.0182	1.93	0.054
eq3_pFV	0.039	0.0279	1.4	0.162
eq3_pSweet	-0.0219	0.0195	-1.13	0.2601
eq3_y*pCarbohydrate	-0.0485	0.028	-1.74	0.0825
eq3_y*pProtein	0.0137	0.0539	0.25	0.7994
eq3_y*pDairy	0.0617	0.0488	1.26	0.2063
eq3_y*pFat	-0.0065	0.0045	-1.43	0.1536
eq3_y*pOil	-0.0242	0.0243	-1	0.3182
eq3_y*pFV	0.0453	0.0367	1.24	0.2164
eq3_y*pSweet	-0.0241	0.0271	-0.89	0.3745
eq3_age*pCarbohydrate	-0.0002	0.0005	-0.43	0.6685
eq3_age*pProtein	-0.0006	0.0008	-0.81	0.4202
eq3_age*pDairy	0.001	0.0007	1.5	0.1331
eq3_age*pFat	0	0.0001	-0.28	0.7832
eq3_age*pOil	-0.0006	0.0003	-1.72	0.0853
eq3_age*pFV	-0.0005	0.0005	-0.93	0.3534
eq3_age*pSweet	0.0006	0.0004	1.76	0.0779
eq4_Constante	0.0009	0.0014	0.63	0.529
eq4_y^1	0.0064	0.0053	1.22	0.2208
eq4_y^2	0.0007	0.0033	0.2	0.8415
eq4_y^3	-0.002	0.0069	-0.28	0.7766
eq4_y^4	0.0052	0.0111	0.47	0.6399
eq4_y^5	0.0054	0.0141	0.38	0.701
eq4_age	0	0	0.99	0.3206
eq4_gender	-0.0009	0.0002	-3.73	0.0002
eq4_BMI	0.0001	0.0001	0.93	0.3545
eq4_y*age	0	0	0.63	0.5259
eq4_y*gender	0.0009	0.0009	1.01	0.3121
eq4_y*BMI	-0.0003	0.0002	-1.21	0.227
eq4_pCarbohydrate	0.0022	0.0033	0.64	0.5196
eq4_pProtein	-0.0139	0.0044	-3.16	0.0016
eq4_pDairy	0.0023	0.0034	0.69	0.4927

eq4_pFat	-0.0021	0.001	-2.13	0.0333
eq4_pOil	0.0006	0.0027	0.23	0.8162
eq4_pFV	0.007	0.0037	1.88	0.0595
eq4_pSweet	0.004	0.0022	1.78	0.0756
eq4_y*pCarbohydrate	0.0044	0.0031	1.41	0.1599
eq4_y*pProtein	0.0095	0.0064	1.49	0.1374
eq4_y*pDairy	-0.0065	0.0045	-1.43	0.1536
eq4_y*pFat	-0.0038	0.0018	-2.17	0.0304
eq4_y*pOil	-0.0033	0.0037	-0.89	0.3762
eq4_y*pFV	-0.0005	0.0051	-0.09	0.9276
eq4_y*pSweet	0.0019	0.0032	0.6	0.5513
eq4_age*pCarbohydrate	-0.0001	0.0001	-1.17	0.2435
eq4_age*pProtein	0.0002	0.0001	2.43	0.0153
eq4_age*pDairy	0	0.0001	-0.28	0.7832
eq4_age*pFat	0	0	1.97	0.0488
eq4_age*pOil	0	0.0001	-0.02	0.9877
eq4_age*pFV	-0.0001	0.0001	-1.68	0.0936
eq4_age*pSweet	0	0	-1.02	0.3072
eq5_Constante	-0.0038	0.0086	-0.44	0.6605
eq5_y^1	-0.025	0.0332	-0.75	0.4528
eq5_y^2	0.036	0.0203	1.77	0.0764
eq5_y^3	0.0145	0.0434	0.34	0.7373
eq5_y^4	-0.1093	0.0681	-1.61	0.1085
eq5_y^5	-0.1171	0.0866	-1.35	0.1763
eq5_age	0.0001	0	1.19	0.2322
eq5_gender	-0.0033	0.0015	-2.25	0.0243
eq5_BMI	0.001	0.0004	2.69	0.0072
eq5_y*age	0	0.0002	-0.09	0.9317
eq5_y*gender	0.0001	0.0055	0.02	0.9879
eq5_y*BMI	0.0012	0.0015	0.83	0.4088
eq5_pCarbohydrate	-0.0173	0.0187	-0.92	0.3564
eq5_pProtein	-0.0432	0.0252	-1.71	0.087
eq5_pDairy	0.035	0.0182	1.93	0.054
eq5_pFat	0.0006	0.0027	0.23	0.8162
eq5_pOil	0.0187	0.0185	1.01	0.3115
eq5_pFV	0.0184	0.0193	0.96	0.3394
eq5_pSweet	-0.025	0.0127	-1.97	0.0493
eq5_y*pCarbohydrate	0.0004	0.0179	0.02	0.9828
eq5_y*pProtein	0.001	0.0351	0.03	0.9778

eq5_y*pDairy	-0.0242	0.0243	-1	0.3182
eq5_y*pFat	-0.0033	0.0037	-0.89	0.3762
eq5_y*pOil	0.0328	0.0251	1.31	0.1914
eq5_y*pFV	-0.0517	0.0261	-1.98	0.0473
eq5_y*pSweet	0.009	0.0179	0.5	0.6145
eq5_age*pCarbohydrate	-0.0002	0.0004	-0.52	0.6011
eq5_age*pProtein	0.001	0.0005	2.21	0.0273
eq5_age*pDairy	-0.0006	0.0003	-1.72	0.0853
eq5_age*pFat	0	0.0001	-0.02	0.9877
eq5_age*pOil	0	0.0004	0.12	0.9043
eq5_age*pFV	-0.0004	0.0004	-1.21	0.2271
eq5_age*pSweet	0.0003	0.0002	1.3	0.1945
eq6_Constante	0.1595	0.014	11.36	0
eq6_y^1	0.0126	0.0544	0.23	0.817
eq6_y^2	0.0592	0.033	1.8	0.0727
eq6_y^3	0.1562	0.0704	2.22	0.0266
eq6_y^4	-0.1862	0.1078	-1.73	0.0842
eq6_y^5	-0.3091	0.1399	-2.21	0.0272
eq6_age	0.0008	0.0001	10.04	0
eq6_gender	-0.006	0.0024	-2.48	0.0132
eq6_BMI	-0.0021	0.0006	-3.38	0.0007
eq6_y*age	-0.0001	0.0003	-0.31	0.7549
eq6_y*gender	0.0018	0.0089	0.21	0.8374
eq6_y*BMI	-0.0012	0.0024	-0.51	0.6108
eq6_pCarbohydrate	-0.0212	0.0287	-0.74	0.4592
eq6_pProtein	-0.0295	0.04	-0.74	0.4597
eq6_pDairy	0.039	0.0279	1.4	0.162
eq6_pFat	0.007	0.0037	1.88	0.0595
eq6_pOil	0.0184	0.0193	0.96	0.3394
eq6_pFV	-0.0365	0.0402	-0.91	0.3644
eq6_pSweet	0.0443	0.02	2.22	0.0264
eq6_y*pCarbohydrate	0.0228	0.0284	0.8	0.4226
eq6_y*pProtein	0.0001	0.0556	0	0.9982
eq6_y*pDairy	0.0453	0.0367	1.24	0.2164
eq6_y*pFat	-0.0005	0.0051	-0.09	0.9276
eq6_y*pOil	-0.0517	0.0261	-1.98	0.0473
eq6_y*pFV	-0.0482	0.055	-0.88	0.3809
eq6_y*pSweet	0.0141	0.0282	0.5	0.6169
eq6_age*pCarbohydrate	0.0002	0.0005	0.37	0.7084

eq6_age*pProtein	0.0001	0.0007	0.1	0.9208
eq6_age*pDairy	-0.0005	0.0005	-0.93	0.3534
eq6_age*pFat	-0.0001	0.0001	-1.68	0.0936
eq6_age*pOil	-0.0004	0.0004	-1.21	0.2271
eq6_age*pFV	0.0011	0.0008	1.47	0.1422
eq6_age*pSweet	-0.0005	0.0004	-1.22	0.2208
eq7_Constante	0.1135	0.0128	8.84	0
eq7_y^1	0.0089	0.0497	0.18	0.8579
eq7_y^2	0.0372	0.0303	1.23	0.2185
eq7_y^3	0.0082	0.0636	0.13	0.8977
eq7_y^4	-0.0877	0.0973	-0.9	0.3671
eq7_y^5	-0.0824	0.1246	-0.66	0.5081
eq7_age	-0.0008	0.0001	-11.12	0
eq7_gender	-0.0063	0.0022	-2.89	0.0038
eq7_BMI	0.0004	0.0006	0.72	0.4693
eq7_y*age	0	0.0003	-0.11	0.9112
eq7_y*gender	-0.0069	0.0081	-0.85	0.3941
eq7_y*BMI	0.0008	0.0022	0.35	0.7229
eq7_pCarbohydrate	-0.0092	0.0212	-0.43	0.6652
eq7_pProtein	-0.0652	0.0302	-2.16	0.0311
eq7_pDairy	-0.0219	0.0195	-1.13	0.2601
eq7_pFat	0.004	0.0022	1.78	0.0756
eq7_pOil	-0.025	0.0127	-1.97	0.0493
eq7_pFV	0.0443	0.02	2.22	0.0264
eq7_pSweet	0.0823	0.0211	3.9	0.0001
eq7_y*pCarbohydrate	-0.0066	0.0224	-0.29	0.77
eq7_y*pProtein	0.0467	0.0415	1.13	0.2599
eq7_y*pDairy	-0.0241	0.0271	-0.89	0.3745
eq7_y*pFat	0.0019	0.0032	0.6	0.5513
eq7_y*pOil	0.009	0.0179	0.5	0.6145
eq7_y*pFV	0.0141	0.0282	0.5	0.6169
eq7_y*pSweet	-0.015	0.0298	-0.5	0.6154
eq7_age*pCarbohydrate	-0.0002	0.0004	-0.48	0.6327
eq7_age*pProtein	0.0009	0.0006	1.52	0.1288
eq7_age*pDairy	0.0006	0.0004	1.76	0.0779
eq7_age*pFat	0	0	-1.02	0.3072
eq7_age*pOil	0.0003	0.0002	1.3	0.1945
eq7_age*pFV	-0.0005	0.0004	-1.22	0.2208
eq7_age*pSweet	-0.0013	0.0004	-3.2	0.0014

Chapter 6
Concluding Remarks

"A conclusion is simply the place where you got tired of thinking."

Dan Chaon

Obesity is a growing wide spread epidemic all over the world, in developed as well as developing countries. Taking into account its nature as a complex phenomenon affected by different aspects including economic ones, a growing body of literature has examined the effect of economic factors on obesity prevalence and the effectiveness of economic intervention policies in combating it. Many papers argue that economics could be a cure as well as one of the main causes of obesity. Despite the increasing obesity rate in Spain, to my knowledge, no known published research studied the economic factors affecting obesity prevalence.

The objective of this thesis has been to analyze the relevance of economic factors (mainly income and other socioeconomic characteristics of Spanish households and market prices) on the prevalence of obesity in Spain and to what extent market intervention prices are effective to reduce obesity and improve the quality of the diet, and under what circumstances. In relation to the existing literature, this thesis project is the first attempt in Spain trying to get an overall picture on the effectiveness of public policies on both food consumption and the quality of diet, on one hand, and on the prevalence of obesity on the other hand.

After four years of research and a careful literature review, a number of decisions have been adopted in order to achieve the above mentioned objective. Although we believe that the adopted decisions have been found in sound with scientific background, we recognize that in some cases other alternatives would have been also plausible; thus generating, on the other hand, new areas for further research. In the following lines we will try to summarize the main decisions adopted and the conclusions obtained. In some cases, we will make some self-criticism and outline further possibilities for future research.

Many papers argued that economics could be a cure, as well as, it is considered one of the main causes of obesity. The first decision has been to conduct a critical review of the economic literature that dealt with the obesity epidemic with the aim of presenting an overall panorama of the research that examined the economic causes and consequences of obesity, the role of economic factors on the growing obesity prevalence and the role of different economic intervention policies in combating this crisis and incentivizing healthy eating among consumers. Therefore, the main aim of the first chapter of the thesis was to figure out if economics can help in combating the growing obesity epidemic. The extensive review of the literature showed that economics, which is considered a main cause for obesity, could also be an effective cure for it.

Another decision was to limit our focus in this thesis to adult obesity, although we are conscious that child obesity is an important and growing worldwide health crisis. Furthermore, we discard child obesity from the scope of the literature review because two recently published papers reviewed the economic aspects of child obesity (Papoutsi et al., 2013; Cawely, 2010). Although obesity prevalence is affected by diet and physical activity choices, in this thesis we mainly concentrate on the economic aspects of diet choices. A detailed review on economic aspects of physical activity choices were done by Wu et al., (2011); Bleich and Sturm, (2009); and Sturm, (2005) and (2004).

It is noteworthy that the review of the literature worth mentioning Different gaps observed in the existing economic literature studied obesity epidemic. We choose to respond to these identified gaps through the thesis.

The third decision was related to how to analyze the effect of economic factors on overweight and obesity prevalence. The literature review showed that most of the studies that dealt with the economic analysis of obesity used a parametric approach. The complex and multifactorial nature of obesity urge the use of nonparametric techniques capable of capturing the nonlinear effect of some covariates and the interaction between them. Our analysis in the third chapter of this thesis could be considered an exception of this trend. In fact, we applied the MARS model considered as a nonparametric counterpart of the multinomial logit model in analyzing most determinant factors that affect obesity prevalence in Spain. The obtained results were found to be quite different from those obtained using traditional parametric models, more consistent with the literature and offering specific recommendations for specific subgroups of the population. These findings are expected to encourage the use of such nonparametric models to study economics of obesity in future research studies.

Our results, generally, suggest that higher food prices, being male and the self-perception of having a bad health status increase the probability of being obese (with marginal effects around 0.11, 0.02 and 0.10, respectively). On the other hand, doing regular physical exercise and having a more complete breakfast decrease the probability of being obese (with marginal effects around -0.08 and -0.06, respectively).

The fourth decision consisted in responding to the unavailability of a reliable measure of diet quality. As a result, we decided to develop a new disease (obesity) specific healthy eating index. We proposed the use of the Multivariate Adaptive Regression Splines (MARS) to reduce subjectivity and develop our new Obesity Specific - Healthy Eating Index (OS-HEI), which is a data driven non-parametric tool and allows for interaction between the different items. The data used come from the 2007-2008 and 2009-2010 National Health and Nutrition Examination Surveys (NHANES). While the first one was used to develop the new index, the second dataset was used to validate the results. The traditional HEI-2010 index was used as a benchmark. Results indicate that the OS-HEI outperforms significantly the HEI-2010 in predicting obesity prevalence.

A main drawback of the articles that dealt with the effect of intervention policies to improve diet quality and combat obesity is the fact of focusing on the effect of such intervention policies only on a specific food group (e.g. Fruit and vegetable; SSB; fast food … etc.) which, in turn, shadow doubts on its reliability as it is omitting the income and the substitution effect and did not take into account the holistic nature of the diet. Gao et al. (2013) tried to fill this gap through applying household production theory to systematically estimate consumer demand for diet quality using the Healthy Eating Index (HEI). Another gap observed in the literature is the need to use advanced demand systems capable of taking into account the unobserved heterogeneity between individual which is especially important in the case of obesity. The demand systems also should to be quite flexible and allow fitting the complexity and multi-dimensional nature of obesity. The recent developed Exact Affine Stone Index (EASI) has these desirable proprieties.

Therefore, our final decision, in terms of the methodological approach, was to overcome the two aforementioned shortcomings using the EASI demand system to analyzing the demand for diet quality, but, in a framework where the holistic nature of the diet is taken into account. Moreover, the effect of different price intervention policies that aim to improve diet quality was examined. These intervention policies include taxing unhealthy foods (e.g. sweets; fat), subsidizing healthy foods (e.g.

Fruit and vegetable) or taxing unhealthy foods and subsidizing healthy foods at the same time. Our results were consistent with the literature suggesting that taxes (subsidies) could have a significant but small effect on diet quality. In addition such policies can be used to facilitate the necessary funds to finance educational campaigns and complementary health policies.

Although, this study has been proved to be useful in better understanding obesity prevalence and food demand in Spain, it can be extended through applying recent advances in experimental and behavioral economics. The majority of the papers dealing with the role of economic factors on obesity prevalence, including our papers, found economic factors to have a significant but a quite small effect on obesity prevalence. This suggests that non-economic factors are playing the major role in obesity prevalence. Just and Payne (2009) mentioned that because food decisions are made with little cognitive involvement, food policies designed to appeal to highly cognitive thought (e.g., fat taxes, detailed information labels) are likely to have little impact. With the aim of understanding this little impact of such policies and why consumers sometimes behave in ways that contradict standard assumptions of economic analysis and make decisions that prevent them from reaching rationally intended goals, the use of experimental economics tools in future research could be helpful in this pending question.

References

Agencia espanola de seguridad alimentaria y nutricion (AESAN), 2011, Encuesta nacional de ingesta dietetica espanola (ENIDE), http://www.west-info.eu/files/Report188.pdf. acced on October,1,2013.

Anderson, P.M., Butcher, K.E., 2006. Childhood obesity: trends and potential causes. Future Child, 16(1), 19-45

Asghari, G., Mirmiran, P., Rashidkhani, B., Asghari-Jafarabadi, M., Mehran, M., Azizi, F., 2012. The association between diet quality indices and obesity: Tehran Lipid and Glucose Study. Archives of Iranian Medicine, 15, 599 – 605.

Bandini, L. G., Vu, D., Must, A., Cyr, H., Goldberg, A., Dietz, W.H., 1999. Comparison of high-calorie, low-nutrient-dense food consumption among obese and non-obese adolescents. Obes Res, 7, 438–443.

Basiotis, PP., Carlson, S., Gerrior, S.A., Juan, W.Y., Lino, M., 2002. The Healthy Eating Index: 1999–2000.

Baum, Charles L., 2011. The Effects of Food Stamps on Obesity. Southern Economics Journal.

Beaulac, J., Kristjansson, E., Cummins, S., 2009. A systematic review of food deserts, 1966-2007. Preventive Chronic Disease, 6(3), 105-110.

Behan, D. and Cox, S., 2010. Obesity and its Relation to Mortality and Morbidity Costs. Society of Actuaries, 59.

Beydoun, M.A., et al., 2008. Reduced Away-from-home Food Expenditure and Better Nutrition Knowledge and Belief Can Improve Quality of Dietary Intake among U.S. Adults. Public Health Nutrition ,10(1017),1-13.

Bleich, S.N. and Sturm, R., 2009. Developing policy solutions for a more active nation: Integrating economic and public health perspectives. Preventive Medicine, 49(4),306-8.

Blundell, R., Robin, J. M., 1999. Estimation in Large and Disaggregated Demand Systems: An Estimator for Conditionally Linear Systems. Journal of Applied Econometrics, 14(3), 209-32.

Bonnet, C.,2013. How to Set up an Effective Food Tax? Comment on "Food Taxes: A New Holy Grail?". Int J Health Policy Manag, 1(3),233-4.

Bowman, S.A., Gortmaker, S.L., Ebbeling, C.B., Pereira, M.A., Ludwig, D.S., 2004. Effects of fast-food consumption on energy intake and diet quality among children in a national household survey. Pediatrics, 113, 112–118.

Bremner, J. D., 2006. Stress and brain atrophy. CNS Neurol, 5, 503–512.

Briand, L.C., Freimut, B., Vollei, F., 2007. Using multiple adaptive regression splines to understand trends in inspection data and identify optimal inspection rates. ISERN TR 00-07 2007. Avaliable from: http://www.salfordsystems.com.

Brownell, K and Warner, K., 2009. The perils of ignoring history: Big tobacco played dirty and millions died. How similar is big food?. The Millbank Quarterly, 87(1), 259-494.

Capacci, S., Mazzocchi, M., 2011. Five-a-day, a price to pay: An evaluation of the UK program impact accounting for market forces. Journal of Health Economics, 30(1), 87-98.

Case, A., Menendez, A., 2009. Sex differences in obesity rates in poor countries: Evidence from South Africa. Economics & Human Biology, 7(3), 271-282.

Cash, S., Lacanilao, R., 2007. Taxing food to improve health: Economic evidence and arguments. Agricultural and Resource Economics Review, 36(2), 174.

Cavaliere, A., Banterle, A., 2008. Economic factors affecting obesity: an application in Italy. International Congress, August 26-29, 2008, Ghent, Belgium, European Association of Agricultural Economists.

Cawley, J., 2010. The economics of childhood obesity. Health Affairs, 29(3), 364-371.

Cawley, J., 2004. The Impact of Obesity on Wages. Journal of Human Resources, 39(2), 451-74.

Cawley, J., Liu, F., 2012. Maternal employment and childhood obesity: A search for mechanisms in time use data. Economics & Human Biology, 10(4), 352-364.

Cawley, J. and Meyerhoefer, C., 2012. The Medical Care Costs of Obesity: An Instrumental Variables Approach. Journal of Health Economics, 31(1), 219-30.

Cawley, J., Frisvold, D., Meyerhoefer, C., 2013.The Impact of Physical Education on Obesity among Elementary School Children. Journal of Health Economics, 32(4), 743-55.

Chaloupka, F.J., 2009. Alcoholic beverage taxes, prices and drinking. Addiction. ,104(2),191-2.

Chaput J.P., Tremblay, A., 2009. Obesity and physical inactivity: the relevance of reconsidering the notion of Sedentariness. Obesity Facts.

Chen H, Ward MH, Graubard BI, Heineman EF, Markin RM, Potischman NA, Russell RM, Weisenburger DD, Tucker KL. Dietary patterns and adenocarcinoma of the esophagus and distal stomach. Am J Clin Nutr. 2005;75:137-144.

Chiplonkar, S.A., Tupe, R., 2010. Development of a diet quality index with special reference to micronutrient adequacy for adolescent girls consuming a lacto-vegetarian diet. J Am Diet Assoc, 110(6),926-31.

Choi ,H.K., Atkinson, K., Karlson, E.W., Curhan, G., 2005. Obesity, weight change, hypertension, diuretic use, and risk of gout in men: the health professionals follow-up study. Arch Intern Med,165(7),742-8.

Chou, S-Y., Grossman, M., Saffer, H., 2004. An Economic Analysis of Adult Obesity: Results from the Behavioral Risk Factor Surveillance System. Journal of Health Economics, 23, 565-587.

Costa-Font, J., Hernández-Quevedo, C., Jiménez-Rubio, D., 2014. Income inequalities in unhealthy life styles in England and Spain. Economics & Human Biology, 13(1), 66-75.

Costa-Font, J., Gil, J., 2005. Obesity and the incidence of chronic diseases in Spain: A seemingly unrelated probit approach. Economics & Human Biology, 3(2), 188-214.

Courtemanche, C., 2009. Longer hours and larger waistlines? The relationship between work hours and obesity. Forum for Health Economics & Policy, 12 (2).

Cullum, A., McCarthy, A., Gunnell, D., Davey Smith, G., Sterne, J.A., Ben-Shlomo, Y.,2004. Dietary restraint and the mis-reporting of anthropometric measures by middle-aged adults. Int J Obes Relat Metab Disord,28(3),426-33.

Cutler, D.M., Glaeser, E.L., Shapiro, J.M., 2003. Why have Americans become more obese?. Journal of Economic Perspectives, 17, 93-118.

Deaton, A. and Muellbauer, J., 1980. An Almost Ideal Demand System. American Economic Review, 70, 312–26.

De Veaux, R. D., Psichogios, D. C., Ungar, L. H., 1993. A Comparison of Two Nonparametric Estimation Schemes: MARS and Neural Networks. Computers and Chem. Engineering, 17(8), 819-837.

Douarin, E., Di Falco, S., 2008. On the Impact of the Five a Day Campaign: Evidence from a Dynamic Engel Curve Analysis. UMRSMART Seminars, INRA, Rennes.

Drewnowski, A., 2010. Healthier foods cost more. Nutrition Reviews, 68, 184–185.

Drewnowski, A., 2005.Concept of a nutritious food: toward a nutrient density score. Am J Clin Nutr, 82(4),721–32.

Drewnowski , A., Fiddler, E.C., Dauchet, L., et al., 2009. Diet quality measures and cardiovascular risk factors in France: applying the Healthy Eating Index to the SU.VI.MAX study. J Am Coll Nutr, 28(1),22-29.

Eastwood, M. A., Brydon, W. G. et al., 1984. Faecal weight and composition, serum lipids, and diet among subjects aged 18 to 80 years not seeking health care. American Journal of Clinical Nutrition 40, 628–634.

Ehmke, M. D. T., T. Warziniack, Schroeter, C. and Morgan, K., 2008. Applying Experimental Economics to Obesity in the Family Household. Journal of Agricultural and Applied Economics 40(2), 539–549.

Elbel, B., Kersh, R., Brescoll, V.L., Dixon, L.B.,2009. Calorie labeling and food choices: a first look at the effects on low-income people in New York City. Health Affairs, 28, 1110 – 1121.

Ellison, B., Jayson L., and David, D., 2013. Looking at the Label and Beyond: The Effects of Calorie Labels, Health Consciousness, and Demographics on Caloric Intake in Restaurants. International Journal of Behavioral Nutrition and Physical Activity, 10-21.

Escobar, M.C. , Veerman, J.L., Tollman, S.M., Bertram, M.Y.and Hofman,.J., 2013. Evidence that a tax on sugar sweetened beverages reduces the obesity rate: a meta-analysis. BMC Public Health , 13,1072.

Etilé, F., 2007. Social Norms, Ideal Body Weight and Food Attitudes. Health Economics, 16, 945–966.

Fan, M., 2010. Do Food Stamps Contribute to Obesity in Low-Income Women? Evidence from the National Longitudinal Survey of Youth 1979. American Journal of Agricultural Economics, 92(4), 1165-1180.

Finkelstein, E.A., Trogdon, J.G., Brown, D.S., Allaire, B.T., Dellea, P.S., Kamal-Bahl, S.J., 2008. The lifetime medical cost burden of overweight and obesity: implications for obesity prevention. Obesity. ,16,1843–1848.

Flegal, K. M., Carroll, M. D., Kit, B. K., Ogden, C. L., 2012. Prevalence of obesity and trends in the distribution of body mass index among U.S. adults, 1999–2010. Journal of the American Medical Association, 307(5) ,E1–E7.

Fletcher, J. et al., 2010. Can soft drink taxes reduce population weight?. Contemporary Economic Policy, 28 (1).

Frank, I. E., 1995. Tutorial: Modern Non-linear Regression Methods. Chemometrics and Intelligent Lab. Sys., 27, 1-19.

Frieadman, J.H., 1991. Multivariate additive regression splines. Ann. Of Stat., 19, 1-67.

Fung, T.T., Rexrode, K.M., Mantzoros, C.S., Manson, J.E., Willett, W.C., Hu, F.B., *2009.* Mediterranean diet and incidence of and mortality from coronary heart disease and stroke in women. Circulation , 119, 1093 –1100.

Gao, Z., Yu, X. and Lee, J.-Y., 2013. Consumer demand for diet quality: evidence from the healthy eating index. Australian Journal of Agricultural and Resource Economics, 57, 301–319.

Gao, X.M., Richards, T.J., Kagan, A., 1997. A latent variable model of consumer taste determination and taste change for complex carbohydrates. Applied Economics, 29, 1643-1654.

Garcia, E., Labeaga, J.M.. and Ortega, C., 2006. Maternal Employment and Childhood Obesity in Spain. Documento de Trabajo, 2006-17. Fundación de Estudios de Economía Aplicada (FEDEA).

Gerber, M., 2001. The comprehensive approach to diet: a critical review. J. Nutr. 131(suppl. 11), 3051S–3055S.

Gibson, D., 2006. Long-term Food Stamp Program Participation is Positively Related to Simultaneous Overweight in Young Daughters and Obesity in Mothers. Journal of Nutrition, 136,1081-1085.

Gibson, D., 2003. Food Stamp Program Participation is Positively Related to Obesity in Low Income Women. Journal of Nutrition, 133, 2225-2231.

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

Golley, R.K., Hendrie, G.A., McNaughton, S.A., 2011. Scores on the dietary guideline index for children and adolescents are associated with nutrient intake and socioeconomic position but not adiposity. J Nutr, 141(7),1340-1347.

Greene, W.H., 2003. Econometric Analysis. Fifth Edition. Prentice Hall.

Gruber, J. and Botond K., 2004. Tax Incidence When Individuals Are Time-Inconsistent: The Case Of Cigarette Excise Taxes. Journal of Public Economics, 88(9-10), 1959-1987.

Guenther, P.M., Reedy, J., Krebs-Smith, S.M., Reeve, B.B., Basiotis, P.P., 2007. Development and Evaluation of the Healthy Eating Index-2005: Technical Report. Center for Nutrition Policy and Promotion, U.S. Department of Agriculture. Available at http://www.cnpp.usda.gov/HealthyEatingIndex.htm

Guenther, P.M., Casavale, K.O., Reedy, J., Kirkpatrick, S.I., Hiza, H.A., Kuczynski, K.J., Kahle, L.L., Krebs-Smith, S.M., 2013. Update of the Healthy Eating Index: HEI-2010. J Acad Nutr Diet, 113(4),569–580.

Gundersen, C., Mahatmya, D., Garasky, S., Lohman, B., 2011. Linking psychosocial stressors and childhood obesity. Obesity Reviews, 12(5), 54–63.

Guo, S. S., Huang, C., Maynard, L. M., Demerath, E., Towne, B., and Chumlea, W. C., 2000. Body Mass Index during Childhood, Adolescence and Young Adulthood in Relation to Adult Overweight and Adiposity: The Fels LongitudinalStudy. International Journal of Obesity and Related Metabolic Disorder, 24, 1628-1635.

Guthrie, H.,1977. Concept of a nutritious food. Jam Diet Assoc, 71, 14–19.

Guthrie, H.A., 1986. There’s no such thing as “junk food,” but there are junk diets. Health line, 5, 11–12.

Guy E.J., et al., 2011. Economic instruments for obesity prevention: results of a scoping review and modified delphi survey. International Journal of Behavioral Nutrition and Physical Activity, 8,109.

Haughton, D., Loan, L.T.T., 2004. Vulnerability of Vietnamese Households, 1992-1998. Working Paper (June 2004). Ford Foundation/General Statistics Office Project, Vietnam.

Hill, N. E., Castellino, D. R., Lansford, J. E., Nowlin, P., Dodge, K. A., Bates, J. E., And Pettit, G. S., 2004. Parent academic involvement as related to school behavior, achievement, and aspirations: Demographic variations across adolescence. Child Development, 75, 1491–1509.

Hoefkens, C., Valli,V., Mazzocchi, M., Traill, B., Verbeke, W., 2013. European consumers' perceived seriousness of their eating habits relative to other personal health risks. Preventive Medicine, 57 , 618–622.

Hu, F.B., 2002. Dietary pattern analysis: a new direction in nutritional epidemiology. Curr Opin Lipidol ,13:3–9.

Huijbregts, P., Feskens, E., Rasanen, L., Fidanza, F., Nissinen, A., Menotti ,A. and Kromhout, D., 1997. Dietary pattern and 20 year mortality in elderly men in Finland, Italy, and the Netherlands: longitudinal cohort study. BMJ, 315, 13–17.

Inestituto nacional de Estadestica (INE), 2013, Encuesta nacional de salud(ENS)http://www.ine.es/jaxi/menu.do;jsessionid=E35B0D01B1416C7691485D1448FE740A.jaxi03?type=pcaxis&path=/t15/p419&file=inebase&L=0. Acceed on October, 1, 2013.

Jennings, C.L., Lambert, E.V., Collins, M., Joffe, Y., Levitt, N.S., Goedecke, J.H., 2008. Determinants of insulin-resistant phenotypes in normalweight and obese Black African women. Obesity , 16: 1602–1609.

Jin, R., Chen, W., Simpson, T. W., 2000. Comparative Studies of Metamodeling Techniques Under Multiple Modeling Criteria. AIAA-2000-4801. American Institute of Aeronautics and Astronautics, Inc.

Just, D. R. and Payne, C. R., 2009. Obesity: can behavioral economics help?. Annals of Behavioral Medicine, 38, S47–S55.

Kant, A.,1996. Indexes of overall diet quality. J Am Diet Assoc, 96, 785–791.

Kant, A., 2004. Dietary patterns and health outcomes. J Am Diet Assoc 104, 615–635

Kant, A. K. and Graubard, B. I., 2005. Energy density of diets reported by American adults: association with food group intake, nutrient intake, and body weight. International Journal of Obesity, 29, 950-956

Kaushal, N., 2007. Do food stamps cause obesity?: Evidence from immigrant experience, Journal of Health Economics, Elsevier, vol. 26(5), 968-991.

Kayri, M., 2007. Two-step clustering analysis in researches: A case study. Eurasian Journal of Educational Research 7(28), 89-99.

Kennedy, E.T., Ohls, J., Carlson, S., Fleming, K., 1995. The Healthy Eating Index: design and applications. J Am Diet Assoc, 95,1103–8.

Khan, L.K., Sobush, K., Keener, D., Goodman, K., Lowry, A., Kakietek, J., et al., 2009. Centers for Disease Control and Prevention. Recommended community strategies and measurements to prevent obesity in the United States. MMWR Recomm Rep,58, 1 – 26.

King, L., 2006. The role of health promotion: between global thinking and local action. Health Promotion Journal of Australia, 17(3), 196-9.

Kontogianni, M.D., Farmaki, A.E., Vidra, N., Sofrona, S., Magkanari, F., Yannakoulia, M., 2010. Associations between lifestyle patterns and body mass index in a sample of Greek children and adolescents. J Am Diet Assoc,110(2),215–221

Kooperberg C., Bose S., Stone C.J., 1997. Polychotomous Regression. Journal of the American Statistical Association 92(437), 117 - 127.

Kourlaba, G., Kondaki, K., Grammatikaki, E., Roma-Giannikou, E., Manios, Y., 2009. Diet quality of preschool children and maternal perceptions/misperceptions: the GENESIS study. Public Health, 123,738–742.

Kranz, S., Findeis, J.L., Sundar, S., Shrestha, S., 2008. Use of the Revised Children's Diet Quality Index to assess preschooler's diet quality, its sociodemographic predictors, and its association with body weight status. J Pediatr (Rio J).,84(1),26-34.

Kuhnert, P. M., Do, K., McClure, R., 2000. Combining Non-Parametric Models with Logistic Regression: An Application to Motor Vehicle Injury Data. Comp. Stat. and Data Analysis 34, 371-386.

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

Lackey, C.J., Kolasa, K.M., 2004. Healthy eating: defining the nutrient quality of foods. Nutr Today ,39,26 –9.

Larson, N. I., Story, M. T., Nelson, M. C., 2009. Neighborhood environments: disparities in access to healthy foods in the U.S. American. Journal of Preventive Medicine, 36(1), 74-81

Leathwick, J. R., Elith, J., Hastie, T., 2006. Comparative Performance of Generalized Additive Models and Multivariate Adaptive Regression Splines for Statistical Modeling of Species Distributions. Ecol. Mod. 199, 188-196.

Lewbel, A. and Pendakur, K., 2009. Tricks with Hicks: The EASI Demand System. American Economic Review, 99(3), 827-63.

Lin, B.H., Smith, T.A., Lee, J.Y., Hall, K.D., 2011. Measuring weight outcomes for obesity intervention strategies: the case of a sugar-sweetened beverage tax. Economics & Human Biology 9, 329–341.

Loureiro, M., Nayga, R.M., 2005. International Dimensions of Obesity and Overweight Related Problems: an Economic Perspective. American Journal of Agricultural Economics, 87(5), 1147-1153.

Lusk, J., and Ellison, B., 2013. Who is to Blame for the Rise in Obesity?. Appetite, 68(1),14-20.

Madore, O., 2007. The Impact of Economic Instruments That Promote Healthy Eating, Encourage Physical Activity and Combat Obesity: Literature Review. Ottawa, Canada: Library of Parliament Publications.

Malik, V.S., Willett, W.C., Hu, F.B., 2013. Global obesity: trends, risk factors and policy implications. Nat Rev Endocrinol, 9(1), 13–27.

Manios, Y., Kourlaba, G., Grammatikaki, E., Androutsos, O., et al., 2010. Comparison of two methods for identifying dietary patterns associated with obesity in preschool children: The GENESIS study. European Journal of Clinical Nutrition 64(12),1407-1414.

Marshall S.J., Biddle S.J.H., Gorely T., Cameron N. and Murdey, I., 2004. Relationships between media use, body fatness and physical activity in children and youth: a meta-analysis. International Journal of Obesity , 28, 1238–1246.

Mazzocchi, M., Traill, B., 2006. Nutrition, health and economic policies in Europe. Food Economics, 2, 113-116.

McCullough, M.L., Feskanich, D., Stampfer, M.J., et al., 2000. Adherence to the Dietary Guidelines for Americans and risk of major chronic disease in women. Am J Clin Nutr,72,1214–22.

McCullough, M.L., Feskanich, D., Rimm, E.B., et al., 2000. Adherence to the Dietary Guidelines for Americans and risk of major chronic disease in men. Am J Clin Nutr ,72,1223–31.

Mertz, W., 1984. Foods and nutrients. J Am Diet Assoc ,84, 769–70.

Meyerhoefer, C. and Pylypchuck, Y., 2008. Does Participation in the Food Stamp Program Increase the Prevalence of Obesity and Health Care Spending?. American Journal of Agricultural Economics, 90.

Miljkovic, D., Nganje, W., de Chastenet, H., 2008. Economic factors affecting the increase in obesity in the United States: Differential response to price. Food Policy, 33(1), 48-60.

Mirmiran, P., Azadbakht, L., Esmaillzadeh, A. and Azizi, F., 2004. Dietary diversity score in adolescents- a good indicator of the nutritional adequacy of diets: Tehran lipid and glucose study. Asia Pacific Journal of Clinical Nutrition 13(1), 56-60.

Monsivais, P., Drewnowski, A., 2009. Lower-energy-density diets are associated with higher monetary costs per kilocalorie and are consumed by women of higher socioeconomic status. J Am Diet Assoc, 9(5), 814-22.

Moore, C. J., Cunningham, S. A., 2012. Social position, psychological stress, and obesity: a systematic review. Journal of the Academy of Nutrition and Dietetics, 112(4), 518-526.

Moore, L. V., Diez Roux, A. V., Evenson, K. R., McGinn, A. P., Brines, S. J., 2008. Availability of recreational resources in minority and low socioeconomic status areas. American Journal of Preventive Medicine, 34(1), 16-22.

Mora,T., Gil, J., Sicras,A., 2012. The Influence of BMI, Obesity and Overweight on Medical Costs: A Panel Data Approach. FEDEA WP 2012-08, Oct. 2012.

Moran, L. J., Sui, Z., et al., 2013. A decrease in diet quality occurs during pregnancy in overweight and obese women which is maintained post-partum. International Journal of Obesity, 37(5), 704-711.

Muñoz, J., Fellicisimo, A. M., 2004. Comparison of Statistical Methods Commonly Used in Predictive Modeling. J. of Vegetation Science , 15, 285-292.

Mytton, O., Clarke, D., Rayner, M., 2012. Taxing unhealthy food and drinks to improve health, BMJ 344, 2931.

Nicklas, J.M., Huskey, K.W., Davis, R.B., Wee, C.C., 2012. Successful weight loss among obese U.S. adults. Am J Prev Med, 42(5), 481-5.

O’Donoghue, T., Rabin, M., 2006. Optimal sin taxes. Journal of Public Economics ,90, 1825—49.

O'Grady, M.J., Capretta, J.C., 2012 Campaign to End Obesity. Robert Wood Johnson Foundation; Princeton. Assessing the Economics of Obesity and Obesity Interventions, 1–40.

Papoutsi, G.S., Drichoutis, A.C., Nayga, R.M., 2013. The causes of childhood obesity: a survey. Journal of Economic Surveys , 27(4),743–767.

Pereira, M.A., Jacobs, D.R. , Van Horn, L., Slattery, M.L., Kartashov, A.I., Ludwig, D.S.,2002. Dairy consumption, obesity, and the insulin resistance syndrome in young adults: the CARDIA Study. JAMA, 287(16), 2081-9.

Philipson, T.J., Posner, R.A., 2003. The long-run growth in obesity as a function of technological change. Perspectives in Biology and Medicine, 46(3), 87-607.

Pomerleau, O.F., Pomerleau, C.S., Mehringer, A.M., Snedecor, S.M., Ninowski, R., Sen ,A., 2005. Nicotine dependence, depression, and gender: characterizing phenotypes based on withdrawal discomfort, response to smoking, and ability to abstain. Nicotine Tob Res,7(1),91–102.

Powell, L.M., Chaloupka, F.J., 2009. Food prices and obesity: evidence and policy implications for taxes and subsidies. Milbank Q. 87(1), 229-57.

Propper, C., 2005. Why Economics is Good for your Health. Health Economics, 14(10), 987-998.

Quatromoni, P.A., Copenhafer, D.L., D'Agostino, R.B., et al., 2002. Dietary patterns predict the development of overweight in women: The Framingham Nutrition Studies. J Am Diet Assoc, 102(9), 1239-1246.

Rao, M., Afshin,A., Singh,G., Mozaffarian,D., 2013. Do healthier foods and diet patterns cost more than less healthy options? A systematic review and meta-analysis. BMJ Open. 3, 42-77.

Reedy, J., Mitrou, P.N., Krebs-Smith, S.M., et al., 2008. Index-based dietary patterns and risk of colorectal cancer: the NIH-AARP Diet and Health Study. Am J Epidemiol ,168,38–48.

Richards, T. J., Allender, W.J. and Hamilton, S. F., 2012. Social Networks and New Product Choice. Annual Meeting, August 12-14, Seattle, Washington, Agricultural and Applied Economics Association.

Richards, T.J., Patterson, P.M., Tegene, A., 2007. Obesity and nutrient consumption: a rational addiction? Contemporary Economic Policy. 25 (3), 309–324.

Schlosser, E., 2002. Fast Food Nation: The Dark Side of the All-American Meal. New York: Perennial .

Schroeter, C., Lusk, J., 2005. Economic factors and body weight: an empirical analysis. Journal of Agricultural and Applied Economics. 40(2), 523-38.

Schroeter, C., Lusk, J., Tyner, W., 2008. Determining the impact of food price and income changes on body weight. Journal of Health Economics, 27(1),45-68.

Seymour, J., Calle, E., Flagg, E., Coates, R., Ford, E., Thun, M., 2003. Diet quality index as a predictor of short-term mortality in the American Cancer Society cancer prevention study II nutrition cohort. Am J Epidemiol, 157, 980–988.

Senauer, B., Gemma, M., 2006. Why is the obesity rate so low in Japan and high in the US? Some possible economic explanations. Paper presented at the Conference of International Association of Agricultural Economists, Gold Coast, Australia, August 12–18.

Shemilt, I., Hollands, G.J., Marteau, T.M., Nakamura, R., Jebb, S.A., et al., 2013. Economic Instruments for Population Diet and Physical Activity Behaviour Change: A Systematic Scoping Review. PLoS ONE, 8(9), e75070.

Silva, A., Etile, F. and Jamet, G., 2013. Consequences Of The 5-A-Day Campaign: Evidence From French Panel Data. Article selected for presentation at the American Agricultural and Applied Economics Association, Washington D.C., August, 2013

Smith, H.L., Meldrum, D.J., Brennan, L.J.,2002. Childhood obesity: A challenge for the anaesthetist. Paediatr Anaesth, 12,750–61

Smith, J. P., 1999. Healthy Bodies and Thick Wallets: The Dual Relation Between Health and Economic Status. Journal of Economic Perspectives 13(2), 145-166.

Stevens, G.A., Singh, G.M., Lu, Y., Danaei, G., Lin, J.K., et al., 2012. National, regional, and global trends in adult overweight and obesity prevalences. Population health metrics, 10- 22.

Sturm, R.,2005. Economics and physical activity. American Journal of Preventive Medicine, 28 (2), 141-149.

Sturm, R. 2004. The economics of physical activity societal trends and rationales for interventions. American Journal of Preventive Medicine, 27 (3), 126-135.

Swinburn, B.A., Sacks, G., Hall, K.D., et al., 2011. The global obesity pandemic: shaped by global drivers and local environments. Lancet,378(9793),804-814

Swinburn, B.A., Egger, G., Raza, F., 1999. Dissecting obesogenic environments: the development and application of a framework for identifying and prioritizing environmental interventions for obesity. Prev. Med. 29, 563–570.

Swinburn, B. and Egger, G.,2002. Preventive strategies against weight gain and obesity. Obesity Reviews, 3(4), 289-301.

Swinburn, B.A., Sacks, G., Hall, K.D., McPherson, K., Finegood, D.T., Moodie, M.L., Gortmaker, S.L., 2011. The global obesity pandemic: shaped by global drivers and local environments. Lancet, 378(9793),804–814.

Trichopoulou, A., Kouris-Blazos, A., Wahlqvist, M., Gnardellis, C., Lagiou ,P., Polychronopoulos, E., Vassilakou, T., Lipworth, L., Trichopoulos, D., 1995. Diet and survival in elderly people. BMJ, 311, 1457–1460.

Togo, P., Osler, M., Sorensen, T.I., et al., 2001. Food intake patterns and body mass index in observational studies. Int J Obes Relat Metab Disord, 25(12),1741- 1751.

Ture, M., Kurt, I., Kurum, A. T., Ozdamar, K., 2005. Comparing Classification Techniques for Predicting Essential Hypertension. Expert Systems with App, 29, 583-588.

U.S. Department of Agriculture (USDA), Center for Nutrition Policy and Promotion (CNPP). (1996 sl. rev.). The Food Guide Pyramid (Home and Garden Bulletin Number 252).

U.S. Department of Agriculture (USDA), Center for Nutrition Policy and Promotion (CNPP). (1995). The Healthy Eating Index (CNPP-1).

Ver Ploeg, M., Mancino, L. and Lin, B.H., 2007. Food and Nutrition Assistance Programs and Obesity: 1976-2002. Economi Research Report No. 48. Washington, DC, U.S. Department of Agriculture Economic Research Service.

Villegas, R., Salim, A., Collins, M.M., et al., 2004. Dietary patterns in middle-aged Irish men and women defined by cluster analysis. Public Health Nutr, 7(8),1017-1024.

Vining, E. P. G., 2008 , Long-term health consequences of epilepsy diet treatments. Epilepsia, 49, 27–29.

Waijers , P., Feskens, E., Ocke, M., 2007. A critical review of predefined diet quality scores. Br J Nutr, 97, 219–231.

Waijers PMCM , Feskens EJM, 2005. Indexes of overall diet quality - A review of the literature. RIVM Report 350010003.

Warner, M. L., 2003. The incidence of overweight and obesity in southern adolescents. Presented at the Food Assistance and Nutrition Research Small Grants Program Conference, Washington, DC.

Wehby, G.L., Courtemanche, C.J., 2012. The heterogeneity of the cigarette price effect on body mass index. Journal of health economics, 31 (5), 719-729.

Wirt, A., Collins, C.E., 2009. Diet quality—What is it and does it matter? . Public Health Nutr ,12,2473–2492

World Health Organization, Global Status Report on Noncommunicable Diseases, 2010. Geneva: World Health Organization, 2011.

Wu, S., Cohen, D., Shi ,Y., Pearson, M., and Sturm, R., 2011. Economic Analysis of Physical Activity Interventions. American Journal of Preventive Medicine40(2),149–158

World Health Organization. Diet, nutrition and the prevention of chronic diseases. Report of a joint WHO/FAO consultation. Geneva, Switzerland. 2003. Technical report 916. Internet: www.who.int/dietphysicalactivity/ publications/trs916/intro/en/ (accessed 5 June 2012).

WHO, 2005. http://www.who.int/hpr/NPH/docs/gs_obesity.pdf [14 October 2013]

Zahour, K.L., Porcari, J.P., Foster, C., and Seebach, E., 2006. Walk Like the Amish"; accessed <walking.about.com/cs/measure/a/amish010704.htm> on May 12, 2012

Zemel, M.B., Miller, S.L., 2004. Dietary calcium and dairy mudulation of adiposity and obesity risk. Nutrition Reviews , 62, 125-131.

Zamora, D., Gordon-Larsen, P., He, K., et al., 2011. Are the 2005 Dietary Guidelines for Americans Associated With reduced risk of type 2 diabetes and cardiometabolic risk factors? Twenty-year findings from the CARDIA study. Diabetes Care, 34(5),1183-1185.

Zhen, C., Finkelstein, E.A., Nonnemaker, J.M., Karns, S.A., and Todd, J.E.,2014. Predicting the effects of sugar-sweetened beverage taxes on food and beverage demand in a large demand system. American Journal of Agricultural Economics, 96 (1),1-25.