Three Essays on the Social Science of Obesity

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Michelle Saksena, B.B.A, M.S.

Graduate Program in Department of Agricultural Environmental and Development Economics

The Ohio State University

2014

Dissertation Committee:

Abdoul Sam, Advisor

Anand Desai

Brian Roe

Abstract

This dissertation contains three articles pertaining to the social science of obesity. All three chapters address the inherent dynamics of the obesity problem, which is surprisingly under-represented in the current literature. The first chapter uses microsimulation to recreate the incidence of obesity in the United States. The calibrated model is then subjected to tax and access policies in order to envisage possible outcomes from such policy intervention. Results show that the effect of taxation and increased access for the poorest individuals had little effect on average weight outcomes. Since the poorest individuals were shown to be the most obese, the results from this simulation imply that the most effective policy will be ones which aim to shift individual preferences toward healthful foods. While taxes did have a slight abatement effect, perhaps the most efficacious use of tax revenue could be used to fund programs that promote healthy eating.

The second chapter utilizes panel data from the China Health and Nutrition Survey to conduct a dynamic estimation of the occurrence of obesity in China. The results show that income has a positive effect on weight, while education was negatively associated with weight gain. Upon estimating food consumption behavior, there is evidence that shows that as

individuals become richer they substitute away from carbohydrate rich foods toward proteins and fat. This behavior maybe attributable to differing perceptions of weight relative to Western societies. Countries like China, with a relatively recent history of food scarcity may perceive weight gain as a sign of health and prosperity. Therefore, consumption of calorie dense foods like meat and fat maybe thought of as health-seeking behavior.

Finally, the third chapter is a dynamic estimation of obesity in the United States. Using pseudo-panel techniques, a dynamic model is constructed at the cohort-level using repeated cross-sectional data from the National Health and Nutrition Examination Survey (NHANES). An inverse relationship between socioeconomic status and weight was only detected among women. Interestingly, the lagged BMI variables were positive and significant in most cases for women, but was generally not for men. This seems to indicate that men are resistant to a *"weight legacy"* and are much more able to change current weight despite previous weight status. When studying food consumption behavior, women are more prone to sugar consumption as a result of increased socioeconomic status and marriage. Considering that the preponderance of overweight and obesity is higher among women, it may indicate that dietary differences across genders particularly in sugar consumption may be a contributing factor to female obesity.

Acknowledgments

I wish to express my sincere thanks to Professors Sam, Desai and Roe for their guidance and counsel.

It is a pleasure to thank my friend and colleague Dr. Norman Maldonado for his many helpful comments.

Vita

January 26, 1984 .... Born - Winnipeg, MB, CAN

2006 .... B.B.A Finance and Economics

2010 .... M.A. Applied Economics

2010-present .... Graduate Teaching Associate,
The Ohio State University.

Fields of Study

Major Field: Department of Agricultural Environmental
and Development Economics

Table of Contents

Page

Abstract ii

Acknowledgments iv

Vita v

List of Tables ix

List of Figures xi

1. Simulating Meal Choices 1

1.1 Introduction 1
1.2 Theoretical Model 5
1.2.1 Ex-Ante Formation of Neighborhoods 5
1.2.2 Utility Maximization Problem 6
1.2.3 Satiety 8
1.2.4 Health 10
1.2.5 Cost Minimization Problem 10
1.3 Computational Model 11
1.3.1 Overview 11
1.3.2 Design Concepts 15
1.3.3 Details 16
1.4 Results 19
1.4.1 Weight Gain Overtime 19
1.4.2 Monte Carlo Simulations 20
1.4.3 Simulated Data 23
1.5 Limitations of Research 28
1.6 Conclusion 29

2. A Dynamic Model of Obesity in China . . . 31
2.1 Introduction . . . 31
2.2 Theoretical Model . . . 36
2.3 The Econometric Model . . . 42
2.3.1 Data . . . 45
2.3.2 Independent Variables . . . 48
2.4 Results . . . 50
2.4.1 Demand for Health . . . 50
2.4.2 Demand for Health Inputs . . . 52
2.4.3 Carbohydrates . . . 53
2.4.4 Protein . . . 54
2.4.5 Fat . . . 55
2.5 Discussion . . . 59
2.5.1 Limitations of Research . . . 59
2.5.2 Conclusion . . . 60

3. A Pseudo-Panel Analysis of Obesity Prevalence using NHANES Data . . . 62
3.1 Introduction . . . 62
3.2 Literature Review . . . 66
3.3 Theoretical and Empirical Framework . . . 68
3.3.1 The Grossman Model . . . 68
3.3.2 Pseudo–Panel Data Sets . . . 71
3.3.3 Data . . . 74
3.3.4 Descriptive Statistics . . . 78
3.4 Results . . . 80
3.4.1 Demand for Health Stock . . . 80
3.4.2 Demand for Health Inputs . . . 87
3.5 Conclusions . . . 97

Bibliography . . . 113

Appendices 114

A. Tiebout Sorting Model . . . 114

B. Monthly Food Prices . . . 116

C. Derivation of Optimal Solution for Health Stock 119

D. Demand for Health Stock and Health Inputs Excluding Food Security . . 121

List of Tables

Table		Page
1.1	Households and Restaurants	12
1.2	Individuals	13
1.3	Alpha by BMI Status	17
1.4	Summary Statistics for 1000 Monte Carlo Simulations	21
1.5	Summary Statistics for Truncated Simulated Data	24
1.6	Summary Statistics of BMI for Truncated Simulated Data	26
1.7	Correlations of Truncated Simulated Data	27
2.1	BMI Classification	46
2.2	Summary Statistics	47
2.3	Demand for Health Stock	51
2.4	Demand for Health Inputs	57
3.1	Cohort Size	76
3.2	Descriptive Statistics	79
3.3	Demand for Health Stock: Women	83
3.4	Demand for Health Stock: Men	85

3.5 Demand for Meat: Women 90

3.6 Demand for Added Sugar: Women 91

3.7 Demand for Fat: Women 92

3.8 Demand for Meat: Men 93

3.9 Demand for Added Sugar: Men 94

3.10 Demand for Fat: Men 95

B.1 U.S. City Average Price Data 117

D.1 Demand for Health Stock: Women 121

D.2 Demand for Health Stock: Men 122

D.3 Demand for Meat: Women 123

D.4 Demand for Added Sugar: Women 124

D.5 Demand for Fat: Women 125

D.6 Demand for Meat: Men 126

D.7 Demand for Added Sugar: Men 128

D.8 Demand for Fat: Men 129

List of Figures

Figure		Page
1.1	Change in Utility as a Function of Satiety	9
1.2	BMI Growth over 10 Years by Varying Excess Daily Calorie Consumption	20
A.1	Depiction of Tiebout Sorting Model	115

Chapter 1: Simulating Meal Choices

1.1 Introduction

Obesity is a pervasive problem in the United States and is associated with increased morbidity and illness. Simply by the scale of the problem (over two thirds of Americans are overweight (Flegal et al. 2012), the cost associated with obesity has potential to become economically large and diminishes human welfare. In particular, it increases medical care costs (Finkelstein et al. 2009,Cawley and Meyerhoefer 2012) and attenuates the labor supply (Cawley 2004).

The obstacles in studying obesity are that the determinants of obesity are numerous and often interact with each other. In fact, the Foresight Programme of the Government of Science Office in the United Kingdom propose a systems dynamics model of obesity containing over 100 contributing factors with over 300 connections and an additional 100 endogenous processes.[1] Moreover, obesity is well suited to be framed as a complex adaptive, public health issue.[2]

Healthful foods often cost more than their less nutritious counterparts (Rao et al. 2013), which are typically calorically dense yet devoid of nutrients. Additionally, more

[1]For an illustration refer to https://www.gov.uk/government/publications/reducing-obesity-obesity-system-map

[2]See Hammond (2009), Prince (2009) , Finegood 2011 and Bruzzone, Novak, and Madeo (2012).

nutritious foods are also often capital and labor intensive to store and prepare which may be particularly burdensome to poorer individuals who may not have the requisite resources for meal preparation. It is therefore reasonable to hypothesize that poorer individuals will have a higher propensity to be overweight since they maybe out priced by health, but have access to an abundance of low quality calories. Furthermore, it maybe the case that overconsumption by poorer individuals may in fact be a rational response to income scarcity and nutritional deficiencies in their diets despite leading to weight gain.

The purpose of this paper is to characterize individual's optimal food choices. To do so, this paper utilizes microsimulation to address the relationship between income and obesity and whether consumptive behavior changes as a result of an individual's environment. In this case, simulation methods are an appropriate treatment and framework of the issue for three reasons. First, simulation is a platform that can be used to forecast many future scenarios in a much smaller time frame than that of real life. This is of particular relevance to obesity as informed policies have the potential to circumvent further spread of obesity. Second, the obesity epidemic possesses many of the characteristics of a complex system (Hammond 2009; Finegood 2011). Particularly, the pathways in which an individual becomes obese are numerous and are often studied across many disciplines. This high level of heterogeneity among agents is further compounded by an equally varied set of environmental factors. Finally, interactions between agents and their environments are complex in that exchanges are numerous and often endogenous. However, these non-linear, cyclical interactions make it quite challenging to study obesity in an empirical setting and as a result are not well understood.

Most articles that tie the application of complex systems to the obesity epidemic have largely been expositional in nature and do not present a formal structural model (Hammond 2009; Prince 2009; Bruzzone, Novak, and Madeo 2012). The exception is Auchincloss et al. (2011) whereby they present an agent based model, which looks at how the presence of food deserts in poor communities may attribute to higher preponderance of obesity. This paper is similar with Auchincloss et al. (2011)'s formation of local environments and more specifically, mobility constraints. A Tiebout (1956) sorting model is used to geographically distribute individuals by income class to form relevant income based neighborhoods. As a result, individuals will exhibit consumption behavior that optimizes their recursive utility maximization problem, which may not necessarily preclude them from unhealthy choices, but may differ with the level of access to meals. By subjecting agents to different price environments and access levels, the results may become an informative resource to constructing effective obesity policy.

To gain early insight, a deterministic model controlling for calorie consumption and mobility was run to illuminate the growth rate of BMI over time. Because of a lagged metabolic response to caloric intake, it was shown that unlike conventional wisdom, which states that weight gain increases linearly over time, BMI actually increased logarithmically. This relationship implies that the short-term response to high caloric intake has much larger effects on body weight and tapers off over time as the individual acclimatizes to higher food intake.

The results from the status quo model with no policy changes show that average BMI was approximately within one standard deviation of average BMI from the National Health and Nutrition Survey 2009-2010 (NHANES) dataset. The simulations did in fact show that the poorest individuals were on average more obese then their richer counterparts, but were also more likely to be underweight, a phenomenon which is more commonly seen in developing nations. The average mortality rate attributed to malnutrition was much higher in the simulation than what is quoted by the World Health Organization (WHO), which calculates a mortality rate of about 0.045% in the United States, but this maybe in part due to the simplification of the model, which does not include adaptive production behavior of households and restaurants.[3] Middle and the richest classes where not far behind as average BMI was still in the overweight category, but variance was much smaller for these subpopulations where there was little preponderance of underweight individuals. When implementing a tax on foods there was a slight decrease in both average BMI and standard deviation, but increased mortality rates. Upon increasing access for the poorest individuals in the taxed model I found no significant difference in average BMI, but found that increasing access promoted a are larger inverse correlation between income class and BMI. This suggests policies, which aim to increase access to food will also need to work toward changing individual preferences for meals, which at least partially is determined by income.

Section 1.2 presents a theoretical model. Section 1.3 describes the computational model using the standardized ODD (Overview, Design Concepts and Details) protocols (Grimm et al. 2006; Grimm et al. 2010; Rand and Rust 2011). Section 1.4

[3] The mortality rate was calculated based on biological limitations of BMI established by the WHO which state that a minimum BMI of 16 is considered to be severely underweight and associated with increased mortality (Berrington de Gonzalez et al. 2010; WHO 1995). Both the NHANES 2009-2010 and simulated datasets excluded individuals with BMI's less than 16 to first increase symmetry between datasets and because very low BMI's maybe correlated to other non-obesity related diseases. The mortality rate was calculated as the percent of individuals below a BMI of 16.

discusses results from Monte Carlo simulations. Section 1.5 highlights the limitations and scope of this simulation. Finally, the section 1.6 contains conclusions.

1.2 Theoretical Model

The model proposed assumes that individuals exhibit optimal behavior. At each time increment individuals must make meal choices. This choice is a function of an individual's desire for satiety and health, which often result in divergent in weight outcomes. And, the preferences for satiety and health are themselves functions of socioeconomic status and weight, which changes over time. Individuals are subjected to a mobility constraint such that individuals are relegated to meals only within their permitted travel distance. The outcomes from these meal choices result in weight changes of the individual, which subsequently impact choices of future meals.

1.2.1 Ex-Ante Formation of Neighborhoods

There is much discussion surrounding the implications of limited access to healthful foods and its affects on obesity among the poor. It would be unreasonable to assume individuals face equally randomized environmental effects. For this reason, I use the Tiebout sorting model Tiebout (1956) to form neighborhoods of like income households.[4] The rationale behind this is as Tiebout illuminates; households will locate to communities with like attributes of their household. In this model I use income to distribute the geography of households. Households with similar income ranges will inevitability face similar, but not identical environmental constraints. The result is the formation of poor, middle class and rich neighborhoods where agents are

[4] See appendix A for an illustration of the Tiebout Sorting model.

either restaurants or households producing meals or individuals who must choose what meals to consume at every t.

1.2.2 Utility Maximization Problem

In this model individuals maximize their utility by consuming calories, but are also conflicted by their preferences to maintain weight. As such, an individual's biological desire to consume a maximal amount of calories at each meal is diminished by her desire to also stay healthy. And, these preferences of satiety versus health differ among individuals and over time. Comparably, Richards and Patterson (2006) assume that ex-ante, each individual decides his or her idealized nutrient profile[5] . In their model a consumer's cost minimization problem is subject to some technology of producing nutrients from a time-dependent condition of food stock, which defines the depreciation of the individual's nutrient stock over time.[6]

Individual behavior relies on utility maximization over time and is shown in equation 1.1. Individuals derive a Cobb-Douglas utility by choosing levels of two goods: satiety and health which are functions of calories and weight. I ignore genetic and physiological factors. In other words, a calorie from fat is metabolically indistinguishable from a calorie of protein or carbohydrate. A simplification of Richards and Patterson (2006)'s model is that rather than having people decide their preferred nutrient profiles, in this model each individual inherits an initial body mass index

[5] A nutrient profile is considered to be the combination of the three macronutrients: carbohydrates, fat and protein and each person's preferred nutrient profile is a reflection of his or her preferences for satiety and health. The total amount of calories is calculated as a linear combination of all three nutrients.

[6] Stigler and Becker (1977), Becker and Murphy (1988) and Iannaccone (1986) previously mention the idea of depreciation of nutrient stocks over time. Richards and Patterson (2006) incorporate this into their model.

(BMI) a priori from a distribution calibrated to data[7]. A person's preferences for satiety and health are informed by their weigh status and income class. Since overall an individual desires to maintain his or her weight it is assumed that as weight status increases an individual will prefer health, $h_{i,t}$ over satiety, $s_{i,t}$ and thus $\alpha_{i,t}$ is decreasing in weight status, $w_{i,t}$.[8] Likewise, as income increases, food scarcity decreases thus poorer individuals will tend to prefer satiety over health and thus $\alpha_{i,t}$ is also decreasing with income. Utility in the current period is also a function of a previous utility as it is reasonable to assume that utility from previous meals influences the choices of an individual in the current period.

$$U_{i,t} = \delta U_{i,t-1} + s_{i,t}(.,.)^{\alpha} h_{i,t}(.,.)^{1-\alpha} \tag{1.1}$$

$$0 < \alpha_{i,t}(w_{i,t-1}, Y_i)$$

$$\frac{\partial \alpha_{i,t}}{\partial w_{i,t}}, \frac{\partial \alpha_{i,t}}{\partial Y_i} < 0 \text{ and } 0 < \delta < 1$$

where

$$\frac{\partial \alpha_{i,t}}{\partial w_{i,t}}, \frac{\partial \alpha_{i,t}}{\partial Y_i} < 0$$

s.t.

$$Y_i \geq p_{pr} pr + p_f f + p_c c$$

$$m_i \geq TCF_i * d_j \tag{1.2}$$

[7] National Health and Nutrition Examination Survey (NHANES) 1999-2000 dataset

[8] In reality this assumption may not hold for everyone. A person's biological state may in fact induce them to seek out more calorie dense foods. In other words, obesity may induce individuals to seek less nutritious foods because of a physiological and metabolic compulsion. The justification for making such an assumption is that broadly speaking it is safe to assume that most individuals strive to be at a healthy weight.

Each individual is subjected to prices of each macronutrient p_{pr},p_c,p_f which were calculated using average monthly food price data for metro areas gathered from the Bureau of Labor Statistics (BLS).[9] The total price of a meal was calculated as the composition of protein, carbohydrate and fat times their prices. Since the price of protein was highest, meals that were protein rich tended to be the most expensive.

Each individual was also subjected to a mobility constraint, $m_{i,t}$. Mobility was decided as a function of income class. Poorer individuals are most restricted since it is reasonable to assume that access to transportation is also a function of income hence their travel cost factor should be higher than richer individuals. The mobility constraint essentially creates a travel radius for each individual and is centered on each person's household location.

1.2.3 Satiety

Equation 1.3 illustrates the functional form of satiety. Satiety is modeled as a sinodial curve with dampening amplitude. The functional form is such that the argmax is exactly one third of the individual's Basal Metabolic Rate (BMR) or the daily amount of calories burned by the individual at time t.[10]

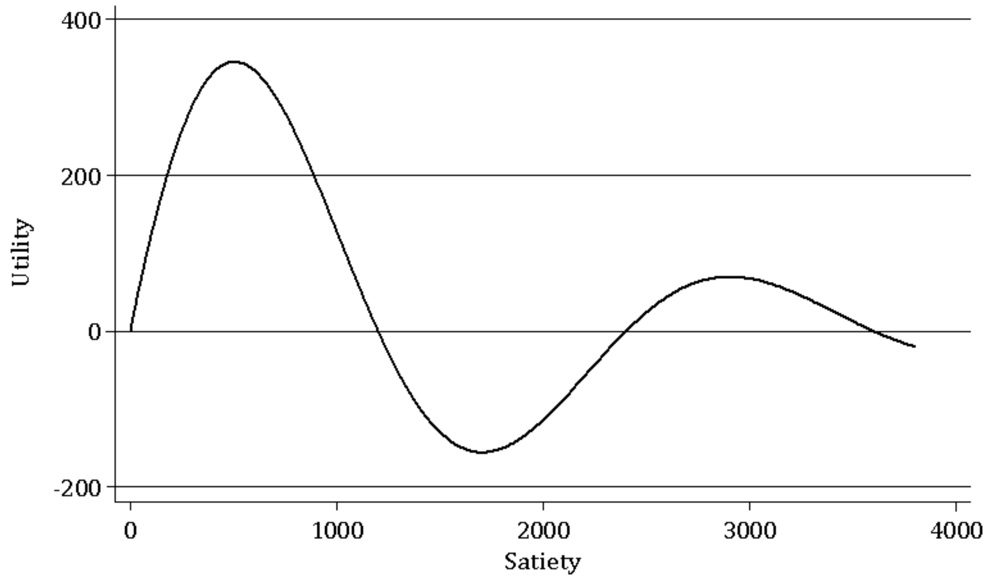
$$s_{i,t} = \frac{BMR_{i,t-1}}{3} e^{\frac{-kcals_j}{BMR_{i,t-1}}} sin\left(\frac{1}{\frac{4BMR_{i,t-1}}{5}}\right)\pi kcals_j \tag{1.3}$$

Figure 1.1 is a depiction of a person's satiety curve whose BMR is precisely 1500 calories. The first segment of the curve can loosely be interpreted as a Laffer curve. The more calories an individual consumes the more utility she receives up until a global maximum where argmax $\equiv$ BMR/3. Furthermore, the argmax is set where

[9] Refer to appendix B for a list of food prices used to calculate macronutrient prices.

[10] For a more thorough definition of BMR refer to section 1.4.1.

Figure 1.1: Change in Utility as a Function of Satiety



BMR/3 because it is assumed that individuals perfectly smooth calorie consumption throughout the day. Specifically, ideally individuals eat three meals a day and eat equal amounts of calories at each meal. Beyond this point the individual receives decreasing levels of utility, as she would prefer to maintain her weight. Utility from excess calories further decreases to the point of negative utility levels, however at some point, utility levels start to increase again. The justification for this is that at some point, excess calories are so much larger than what is ideal that the individual changes his preferences in favor of more satiety. The minimum point can be thought of as the *point of resignation* and consumption of calories at or beyond this point will increase utility because individuals value less of maintaining weight in favor of meal satisfaction, but the maximal amount will not exceed the utility derived from eating the calorie equivalent of a third of ones BMR.

1.2.4 Health

The functional form of health is modeled much simpler and is measured as the inverse of the distance between calories consumed from the meal and a third of the individual's BMR at time t. Health is represented simplistically due to how the meals are defined in the simulation. Meals are characterized by their macronutrient content only and as a result, there is no consistent way of assigning micronutrient levels to meals without fully disclosing what foods are in each meal. As a result, more accurate health measures such as the Healthy Eating Index (HEI) cannot be used in the context of this simulation. Equation 1.4 is the functional form for health.

$$health_{i,t} = \frac{1}{\left| kcal_j - \frac{BMR_{i,t-1}}{3} \right|} \tag{1.4}$$

1.2.5 Cost Minimization Problem

The caloric content of meals produced by restaurants and households is determined by a cost minimization problem subject to p_{pr},p_c,p_f. The cost minimization problem and input factor demand functions for protein, carbohydrate and fat are as follows.

$$\min_{c,pr,f} = p_c c + p_{pr} pr + p_f f \tag{1.5}$$

s.t.

$$\bar{Q}_j = F(c, pr, f) = Ac^{\omega} pr^{\beta} f^{\gamma}$$

The optimal input factor demands:

$$c^* = \frac{\bar{Q}_j}{\left(\frac{\beta}{\omega} * \frac{p_c}{p_{pr}}\right)^{\beta}} \times \left(\frac{\gamma}{\omega} * \frac{p_c}{p_f}\right)^{\gamma} \tag{1.6}$$

$$pr^* = \frac{p_c}{p_{pr}} \times \frac{\beta}{\omega} \times c^* \tag{1.7}$$

$$f^* = \frac{p_c}{p_f} \times \frac{\gamma}{\omega} \times c^* \tag{1.8}$$

1.3 Computational Model

The structure of the computational model follows the ODD protocols laid out by Grimm et al. (2006). ODD protocols where created to fulfill a need for standardized methods to effectively describe the components of social simulation models, which are often numerous and can be overwhelming to explain as the complexity of the model increases.

1.3.1 Overview

This model was designed to explore questions about obesity outcomes. In this simulation I am interested in unraveling obesity outcomes of populations under different price and access environments.

There are four kinds of entities: individuals, households, restaurants and patches of land. The patches make up a square grid landscape that is 25 × 25. Each patch is categorized by type. Type 0 or null type is designated as an empty patch, type 1 is a

Table 1.1: Households and Restaurants

State Variables	Units	Distribution/Equation
$\bar{Q}$		N(45,4)
p_{pr}, p_c, p_f	USD	BLS data
ω, β, γ		[4/17, 4/17, 9/17]
Calculated Variables		
$Carbohydrates_j$	g	1.6
$Protein_j$	g	1.7
Fat_j	g	1.8
$kcals_j$		$4(Protein_j + Carbohydrate_j) + 9 \times Fat_j$
$Price_j$	USD	$Protein_j \times p_{pr} + Carbohydrate_j \times p_c + Fat_j \times p_f$
Price Class		$\begin{cases} 1, & if\ price_j \in first\ tertile \\ 2, & if\ price_j \in second\ tertile \\ 3, & if\ price_j \in third\ tertile \end{cases}$

residential patch and type 2 is a commercial patch. Households can only exist on type 1 patches while restaurants may only occupy type 2 patches. Each household and restaurant offers one representative meal, which is characterized by the following state variables illustrated in table 1.1: output $\bar{Q}$, prices, p_{pr},p_f,p_c and output elasticities ω, β, γ. $\bar{Q}$ is determined by random normal draws such that the average caloric content of each meal is approximately 800 calories. Input prices for the macronutrients are imputed using average monthly food prices of metropolitan areas collected by the BLS. Output elasticities are taken from normalized grams to calories conversion rates for each macronutrient. Each gram of carbohydrate and protein is equivalent to four calories while one gram of fat is equal to nine calories.

Table 1.2: Individuals

State Variable	Units	Distribution/Equation
Sex_i	$1 = female$	$B(n, 0.5)$
$Age_{i,t=0}$	years	U[18..80]
$BMI_{i,t=0}$		Gumbel Distribution] $(\lambda = 20, \kappa = 5)$
$Height_i$	cm	U[152..183]
$Income_i$	US Dollars	$f(price_j)$
$Leisure_i$	Hours (weekly)	U[5..20]
Calculated Variable		
$Activity\ Level_i$		$\begin{cases} U[1.2\ 1.375] \\ if\ leisure_i \in first\ tertile \\ U[1.2\ 1.375\ 1.55\ 1.725] \\ if\ leisure_i \in second\ tertile \\ U[1.2\ 1.375\ 1.55\ 1.725] \\ if\ leisure_i \in third\ tertile \end{cases}$
$Weight_i$	kg	$BMI_{i,t} \times height_i^2$
$BMR_{i,t}$		1.9 and 1.10
$IncomeClass_i$		$\begin{cases} 1, & if\ income_i \in\ first\ tertile \\ 2, & if\ income_i \in\ second\ tertile \\ 3, & if\ income_i \in\ third\ tertile \end{cases}$
$Travel\ Cost\ Factor_i$		[high, medium, low]
$Travel\ Radius_i$		$\begin{cases} 5, & if\ income\ class_i = 1 \\ 7, & if\ income\ class_i = 2 \\ 10, & if\ income\ class_i = 3 \end{cases}$

Levels of carbohydrates, fat and protein is in grams and each are calculated as the optimal level of their input factor demands given by equations 1.6, 1.7 and 1.8. Total calories is equal to each macronutrient in grams multiplied by the appropriate gram to calorie conversion rate. Price is calculated similarly as the sum of each macronutrient times the price while price class is determined by the household or restaurant's percentile score. Similarly, individuals are also characterized by many state variables. Table 1.2 shows the distributions for each state variable. Each individual is assigned a sex, age, body mass index (BMI), height, income and leisure hours. Sex is determined by a binomial distribution. Age is bounded between 18 and 80 and draws are from a discrete uniform distribution. Using the National Health and Nutrition Examination Survey (NHANES) 1999-2000 dataset I use maximum likelihood methods to fit a maximum extreme value distribution (λ= 20, κ= 5) to BMI data. Starting BMI levels were drawn from this distribution. Height was determined from a uniform distribution and is bounded by 152 to 183 centimeters. Leisure hours are determined from a discrete uniform distribution bounded by 5 and 20 hours. Lastly, income is endogenous to household prices of meals. Since the household is viewed as a producer of meals the price of the meal reflects the household income devoted to the production of that meal. And, at each time period each individual then decides whether to use that income to produce meals at home or to forego production and seek meals outside of the household.[11]

The timeline for each simulation is ten years. Each time increment (tick) is interpreted as one meal event and three consecutive ticks are considered to be one day.

[11] In order to rid rounding errors, income is initially calculated to be 1.05*$price_j$.

Therefore, the simulation runs for 10 years or 10950 ticks[12]. Prior to the optimization, each household and restaurant must migrate to the appropriate neighborhood. The order in which agents move is irrelevant during this step as the process assumes that agents move simultaneously. The aggregate result is the formation of neighborhoods by income. During the optimization only individuals move. Each individual selects his optimal choice and moves to that destination. After each individual moves to his optimal destination he "consumes" the meal and moves back to his household. Each meal has a caloric content, which is inherited by the individual and incrementally changes the person's weight. Here BMR also adjusts to accommodate the caloric intake. Over time each person's weight changes as a consequence of his or her historical caloric intake and as a result, his or her preferences for satiety and health also change over time. Again, the succession of individuals is unimportant since it is assumed that all individuals consume meals simultaneously at each time increment while the scarcity assumption is relaxed.[13]

1.3.2 Design Concepts

This model tries to understand how food choices and food environments affect the long-run weight status of individuals. Individual behavior is based on economic foundations of utility maximization and cost minimization. When faced with varying food environments individuals make food choices that affect their health differently. As a consequence of individual local food environments individuals make food choices in order to maximize their utility differently from one environment to the next. These

[12] $(3 \times 365 \times 10) = 10950$

[13] It is assumed that each restaurant/household can supply infinite number of meals. Therefore, there are no restrictions on the number of patrons per restaurant/household at any time.

preferences also evolve overtime in response to their changing weight status. Variability of the agents is determined using pseudo-random generators to create relevant variables, which characterize each agent and influence their decision process and response to calorie consumption. Sensing in this model is particularly important for the individual agents. At each time increment individuals are assumed to have knowledge of nutrient profiles and prices of all meals within their travel radius. The level of sensing affects the food choices of individuals as mobility restrictions may force individuals to choose sub-optimal (non-weight-maintaining meals), which lead to weight gain or decline.

A simplification of the model in this case is that households and restaurants do not exhibit adaptive behavior in the production of foods. As individual preferences change over time it is perhaps more realistic for households and restaurants to alter the nutrient profiles of meals in order to attract patronage. However, in favor of tractability this complication was not added, but maybe a source of exploration in the future.

1.3.3 Details

In each simulation there are 300 individuals and households as well as an additional 201 restaurants. As stated earlier, neighborhoods are formed using a Tiebout sorting model, which establishes the locations of restaurants and households prior to the optimization. Every individual, household and restaurant is characterized by their initial state variables explained in Tables 1 and 2. First restaurants and households are categorized by income/price level. The poorest households and restaurants correspond to the first tertile of the income/price distribution while middle and rich classes correspond to the second and third tertile respectively. Initially, households and restaurants are distributed onto the landscape randomly. Each restaurant/household randomly moves to a vacant patch if less than 66% of neighboring patches are not also occupied by like households or restaurants. The sorting model converges when all agents have stopped moving. Once sorting has completed each restaurant and household inherit the patch's coordinates as its permanent location.

Table 1.3: Alpha by BMI Status

BMI Range	Obesity Classification	α
< 18.5	Underweight	0.5
$[18.5, 25)$	Normal	0.3
$[25, 30)$	Overweight	0.2
$[30, 34.9)$	Obese I	0.1
$[34.9, 39.9)$	Obese II	0.05
≥ 39.9	Obese III	0.01

Once the neighborhoods are formed the actual optimization occurs. All restaurants within the person's travel radius and his or her own household are included as possible destinations. Then, restaurants whose meal price exceeds the individual's income constraints are excluded from possible destinations. This forms the individual's feasible choice set, $\mathcal{F}_{i,t}$. For each destination a utility is calculated at time t using equation 1.1. Individual's preferences for satiety and health are determined by their current BMI and are further augmented by income class. Table **??** shows the corresponding α by BMI status. Once utilities are calculated the individuals move to the destination that offers the highest level of utility. Once each individual moves to his or her optimal destination the individual consumes the meal. The individual banks the caloric content of the meal. The net calorie amount, which is given as calories consumed minus a third of the individuals BMR incrementally updates the person's weight. BMR is also adjusted to correspond to each iterative change in weight. BMR is calculated using the Harris Benedict Equation 1.9 and 1.10 (Roza and Shizgal 1984). Since the HB equation is also dependent on age, age is also updated every representative year.

Females:

$$\begin{aligned} BMR = 88.362 + 4.799 \times height(m^2) + 13.397 \times weight(kg) \\ - 5.677 age \times activity\ level \end{aligned} \tag{1.9}$$

Males:

$$\begin{aligned} BMR = 447.593 + 3.098 \times height(m^2) + 9.247 \times weight(kg) \\ - 4.330 age \times activity\ level \end{aligned} \tag{1.10}$$

Three rounds of 1000 Monte Carlo simulations each with different initial conditions were run. The first model or status quo model was run to calibrate the simulation to existing NHANES BMI data and does not contain any policy interventions. Upon successful calibration, two subsequent models altering initial conditions are run. A taxed model is run implementing producer and consumer taxes on macronutrients. A third model then builds off of the taxed model incorporating increased access levels for the poorest individuals.

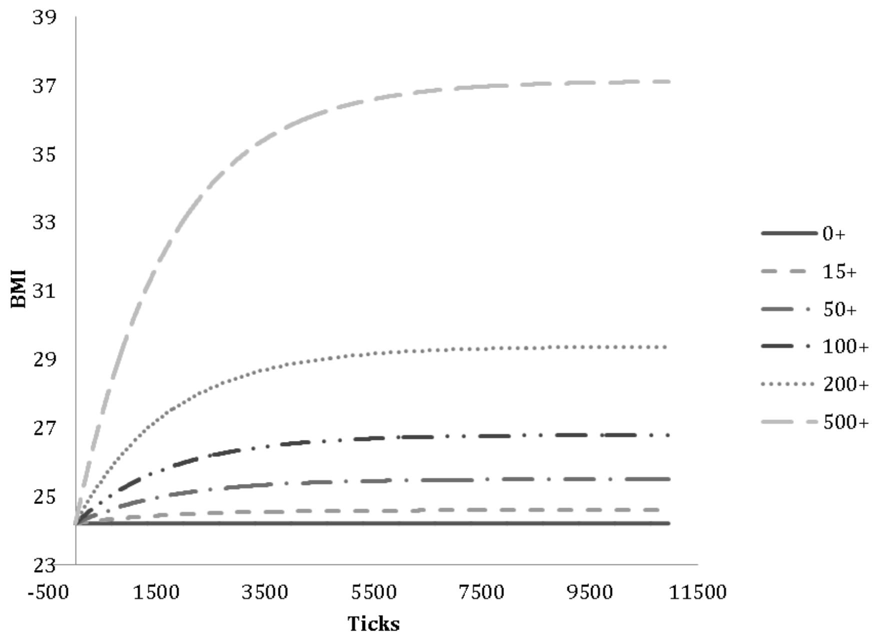
1.4 Results

1.4.1 Weight Gain Overtime

Figure 1.2 shows the results of a deterministic model, which controls for the calorie content of meals and mobility constraints and age. Using the male archetype used in the Roza and Shizgal (1984), this model depicts the weight gain of an individual who is initially 1.7m in height, 70kgs and is 50 years old. This amounts to an initial BMI of 24.2 and a resting BMR of 1558 kcals/day. Each plot shows weight gain over time from eating the corresponding daily excess calories over 10 years. Unsurprisingly, the individual experiences no weigh gain in the 0+ simulation where the individual eats exactly 1/3 of his initial BMR at each meal. More interestingly are the 200+ and 500+ plots. A daily excess of 200 calories shows quite a significant increase in BMI. In fact, over ten years the individual's BMI increased by 21.29%. This is consistent with empirical findings found by Cutler, Glaeser, and Shapiro (2003) who found that a daily excess of just 200 calories leads to significant weight gain. Lastly, the 500+ line shows a very dramatic weight gain. In ten years this person increased his BMI by 53.22%. This plot is in reference to the common adage that an extra 500 calories a day will lead to a 1 pound increase in weight per week because there are 3500 calories in one pound of fat. However, this graph quite pessimistically suggests otherwise and would propose that weight gain at least initially increases far quicker.

The logarithmic growth is attributable to lagged BMR. At t, net calories equals $kcals_j - BMR_{i,t-1}$. When the individual consumes more than his daily caloric amount he is essentially only equipped with a metabolism, which burns less and as a result gains weight. Over time the difference between BMR and caloric intake decreases as

Figure 1.2: BMI Growth over 10 Years by Varying Excess Daily Calorie Consumption



the person's BMR adjusts to the weight gain. The end result is $\frac{\partial BMI}{\partial t} < 0$. Specifically, we see that BMI grows logistically.

1.4.2 Monte Carlo Simulations

Table 1.4 shows results from the status quo model are consistent with the BMI distribution from NHANES 2009-2010 data. After truncating the data to exclude individuals with BMI's lower than 16, the results from 1000 Monte Carlo simulations show that the average mean and standard deviation from the simulated data are within one standard deviation to the average and standard deviation of the NHANES 2009-2010 data. This provides a sufficient burden of proof that the status quo model is calibrated to NHANES 2009-2010 empirical findings and justifies further exploration of different pricing environments and access levels, which may support weight loss.

Table 1.4: Summary Statistics for 1000 Monte Carlo Simulations

	J	μ	σ	Min	Max
Status Quo					
No Truncation					
Average BMI	1000	29.24	0.72	27.24	32
Standard Deviation	1000	7.13	0.58	5.64	9.46
With Truncation					
Average BMI	1000	29.6	0.73	27.55	32.26
Standard Deviation	1000	6.77	0.56	5.29	8.9
Mortalities	1000	6.8	2.61	0	18
Mortality Rate	1000	0.02	0.01	0	0.06
Tax					
No Truncation					
Average BMI	1000	28.46	0.64	26.74	30.99
Standard Deviation	1000	6.98	0.54	5.64	9.11
With Truncation					
Average BMI	1000	28.93	0.65	27.01	31.46
Standard Deviation	1000	6.54	0.52	5.26	8.52
Mortalities	1000	8.94	3.17	0	22
Mortality Rate	1000	0.03	0.01	0	0.07
Tax and Increase Access					
No Truncation					
Average BMI	1000	28.51	0.68	26.65	30.94
Standard Deviation	1000	7.03	0.55	5.62	9.17
With Truncation					
Average BMI	1000	28.99	0.69	27.02	31.44
Standard Deviation	1000	6.58	0.53	5.2	8.51
Mortalities	1000	9.06	3.04	0	20
Mortality Rate	1000	0.03	0.01	0	0.07

Note: Data from NHANES (2009-2010) show that BMI had average, μ=29.02 and standard deviations, σ=6.87.

One should make note that WHO estimates for mortality due to starvation in the United States is approximately 0.0045. The average mortality rate from the simulated data was much higher than WHO estimates. Uncharacteristically high mortality rates maybe a consequence of a simplification of the model, which does not implement adaptive production behavior of the household or restaurants. More specifically, household and restaurants are not able to alter the meals they provide as individual tastes change over time thus the restaurant environment in the simulation is static and restaurants or households that produce meals that are not attractive to individuals are not eliminated from the model nor are they replaced by more relevant restaurants or changes in meal production for the households.

The taxed model implements a producer tax on fat such that the relative price with respect to the price of carbohydrates goes from \$1.71 to \$0.47 and with respect to protein goes from \$0.88 to \$2.10. In addition, the output price of carbohydrates increased such that the relative price to protein goes from \$0.52 to \$2.59. Results show that this tax policy yields only a slight decrease in average BMI. This may indicate that producers substitute toward relatively cheaper macronutrients to maintain similar caloric content in foods while individuals also exhibit inelastic demand for calories. This suggests that taxing foods may not be the most effective policy to curb obesity, but could be a source of tax revenue, which could be used to fund programs that fight obesity. These results are similar to previous empirical work by Kuchler, Tegene, and Harris (2004), Jacobson and Brownell (2000) and Powell (2011).

The third scenario maintained the tax, but also increased mobility for the poorest individuals. Results from table 1.4 show that increasing access did not amount to an appreciable change in average BMI or standard deviation when compared to the tax-only model. Mortality rates did not change in the increased access environment and this suggests that increased access may not be enough to overcome price barriers for poor individuals.

1.4.3 Simulated Data

To understand the simulated data at an individual level, the data from 1000 Monte Carlo Simulations were aggregated for each of the three scenarios. Again, individuals with BMI's lower than 16 were dropped resulting in the number observations for each scenario where 293,201, 291,040 and 290,939 respectively. Table 1.5 shows summary statistics of all relevant variables. In the status quo model, the simulation yielded an average meal, composing 45.61 grams of carbohydrates, 23.59 grams of protein and 59.92 grams of fat resulting in the average caloric content of a meal to be 816 calories. The taxed model shows that on average the meals where composed more of carbohydrate, while average protein content went up and average fat content went down. This is to be expected as the relative producing price of fat increased in this scenario. Lastly, the average caloric intake decreased by approximately 20 calories in the taxed model. As to be expected, in the taxed and increased access model similar averages for carbohydrates, fat, protein and calorie content were observed since the price regime did not change in this scenario when compared to the tax model.

Table 1.6 shows the breakdown of BMI by income class. The overall average BMI was 29.61, higher than the average BMI from NHANES 2009-2010, but within the first standard deviation of the NHANES data. The results are consistent with previous empirical literature which indicates that the poorest individuals tend to be the

Table 1.5: Summary Statistics for Truncated Simulated Data

Coefficient	N	μ	σ	Min	Max
Status Quo					
Carbohydrates	293,201	45.61	3.97	27.69	63.44
Protein	293,201	23.59	2.05	14.32	32.81
Fat	293,201	59.92	5.22	36.38	83.36
kcals	293,201	816.04	71.08	495.44	1135.23
α	293,201	0.55	0.2	0.33	1
Travel Radius	293,201	15.11	2.82	10	20
Tax					
Carbohydrates	291,040	72.04	6.27	43.72	100.18
Protein	291,040	37.26	3.24	22.61	51.81
Fat	291,040	39.94	3.48	24.24	55.54
kcals	291,040	796.7	69.34	483.47	1107.8
α	291,040	0.55	0.2	0.33	1
Travel Radius	291,040	15.11	2.82	10	20
Tax and Increased Access					
Carbohydrates	290,939	72.05	6.26		100.18
Protein	290,939	37.27	3.24	22.61	51.81
Fat	290,939	39.95	3.47	24.24	55.54
kcals	290,939	796.81	69.21	483.47	1107.8
α	290,939	0.55	0.2	0.33	1
Travel Radius	290,939	16.6	2.33	15	20

Note: Data from NHANES (2009-2010) show that BMI had average, μ=29.02 and standard deviations, σ=6.87.

most overweight with an average BMI of 31.33. Surprisingly, the poorest individuals were also more likely to be underweight, which is expressed in the much higher standard deviation for income class 1 across all three scenarios. This phenomenon is more commonly seen in developing countries and may reflect the simplification of the model which does not include adaptive behavior of restaurants or meal production in households. Table 1.6 also shows that each income class experienced a small decrease in average BMI and standard deviation in the taxed and taxed and increased access models. However, the poorest individuals remained to be the most overweight in all three models.

Futhermore, table 1.7 shows correlation coefficients for relevant variables. The status quo model is consistent with previous findings (Sobal and Stunkard 1989; McLaren 2007), which showed that income and income class were negatively correlated with BMI. In the status quo model, income class is found to increase BMI by approximately 6% as income class goes down. Satiety, which is a function of income and weight is positively correlated with BMI and suggests that while increased BMI should promote a decrease in α, low income makes propensity toward high calorie consumption overwhelmingly positive. Calories is surprisingly not significantly correlated to BMI but was in the expected positive direction. Provision of calories in the household and income class is positively correlated. In combination with the positive correlation between BMI and α and the negative correlation with BMI and income class suggest that while richer individuals have access to more calories, poor individuals exhibit behavior which preclude them to consume weight gaining foods.

The taxed model shows a smaller negative correlation with income class and similarly α is less strongly positively correlated with BMI. Calories is now significant

Table 1.6: Summary Statistics of BMI for Truncated Simulated Data

	N	μ	σ
Status Quo			
Income Class			
Income Class 1	43,456	31.33	9.54
Income Class 2	200,036	29.2	6.23
Income Class 3	49,709	29.74	5.95
Total	293,201	29.61	6.82
Tax			
Income Class			
Income Class 1	43,271	30.32	9.09
Income Class 2	198,376	28.54	6.03
Income Class 3	49,393	29.27	5.82
Total	291,040	28.93	6.58
Tax and Increased Access			
Income Class			
Income Class 1	43,340	30.55	9.22
Income Class 2	197,748	28.58	6.06
Income Class 3	49,851	29.25	5.84
Total	290,939	28.99	6.63

Table 1.7: Correlations of Truncated Simulated Data

Status Quo				
	BMI	Income Class	α	Calories
BMI	1			
Income Class	-0.0607*	1		
α	0.0914*	-0.9212*	1	
Calories	0.0026	0.8286*	-0.7619*	1

Tax				
	BMI	Income Class	α	Calories
BMI	1			
Income Class	-0.0399*	1		
α	0.0714*	-0.9213*	1	
Calories	0.0220*	0.8299*	-0.7619*	1

Tax & Increased Access				
	BMI	Income Class	α	Calories
BMI	1			
Income Class	-0.0499*	1		
α	0.0821*	-0.9255*	1	
Calories	0.0158*	0.8350*	-0.7719*	1

and positively correlated with BMI. Interestingly, increasing access to the poorest individuals promotes the negative correlation between BMI and income class. This suggests that increasing access has a small deleterious effect on BMI. This could be attributable to the fact that increasing mobility for the poorest individuals gives them greater exposure to healthy foods, but also more calorie dense foods. Since their proclivities toward higher calorie foods (correlation between α and income class only changes marginally from the status quo model) they may be more opt to pick unhealthy foods because of an increased $\mathcal{F}$ choice set which includes more unhealthy foods.

1.5 Limitations of Research

There are two limitations of this model, which should be discussed. The model does not incorporate adaptive behavior for the restaurants and households in terms of the provision of calories. Restaurants and households do not adjust to individual's food consumption behavior and this maybe a reflection of higher than usual mortality rates. Additionally, the measure of access is incomplete in this simulation since it only measures access to preexisting meals which as discussed do not update with individuals or over time. The static food environment means that access only changes in the construct of increasing mobility for the individuals and not necessarily increasing access by propagating more restaurants that offer more relevant meals.

Second, in this model calories from each nutrient are considered equal and the idea of propensity to be converted into and stored as body fat is neglected. For example, glycemic index (GI) is a measure which indicates which foods are more readily stored as fat. For example, table sugar has a GI of 100 and would be considered to be very fattening since the body readily metabolizes and converts it to body fat despite there being no actual consumptive fat in table sugar itself. Contrastingly, protein has the same gram to calorie conversion as carbohydrates. However, the way protein is metabolized dictates that only in very larger amounts of consumption will protein be converted and stored as fat. Unfortunately, GI's of foods can only be determined by laboratory methods and so cannot be incorporated in this model.

1.6 Conclusion

This paper examines the effects of price and access environments on incidence of obesity in a simulation context. The status quo model shows many characteristics that it is well calibrated to actual NHANES 2009-2010 BMI data. Much like the predominant empirical literature, BMI was found to be higher among the poorest individuals. This serves as a good basis for applying tax and access policy. The taxed model showed a decrease in BMI both overall and for each income class, but this decrease was not dramatic. These findings are similar to what is found in existing literature and suggests that taxation may not be an effective mitigating policy. As suggested by previous papers, the tax revenue could be used to fund more effective policies. The tax and increased access model showed similar results to just the taxed model in terms of average BMI for the overall sample and by income class. However, when looking at the correlation between relevant variables, increasing access actually promoted a more negative correlation with income class. As discussed in the previous section access is defined as mobility. The results from this simulation suggest that while increasing mobility for the poorest individuals increases exposure to healthier meals, their proclivities toward calories which is a function of income does not change and thus poor individuals still choose unhealthy options despite increased exposure to healthier meals. Therefore, an effective policy would include increasing taxes to funding policies geared toward first increasing access to foods (either by increasing mobility and/or by increasing meal availability) and fundamentally changing preferences for unhealthy foods, which is at least are partially determined by income.

Chapter 2: A Dynamic Model of Obesity in China

2.1 Introduction

Thirty years ago obesity was assumed to be a problem reserved only for the developed world. In fact, concerns for impoverished nations primarily focused on malnutrition and spread of infectious diseases. The belief was that obesity was non-existent in developing countries where people were protected by poverty as low incomes did not permit a sufficient amount calorie attainment. However, this future would turn out to only be partially true. With decreasing food prices, particularly grain-based and oil foods, what we now observe is a paradoxical coexistence of undernutrition and obesity in developing countries (Sawaya et al. 1995; Sichieri, Siqueira, and Moura 2000; Doak et al. 2004; Misra and Khurana 2008). This dual phenomenon is particularly burdensome for developing worlds with inadequate medical services as they are forced to deal not only with medical legacies relating to undernutrition, but also new deleterious diseases associated with obesity like type-2 diabetes, stroke and heart disease.

This paper studies the growing adult obesity problem in contemporaneous China. While rates of obesity are still dwarfed by those observed in the western world, the rapid spread of obesity in China (Popkin 1998; Popkin 2001; Wu 2006) is cause to study this phenomenon. Technological advancements have rapidly decreased the relative prices of edible oils and refined carbohydrates. China in particular has shown a greater consumption of edible oils in recent years. We have also observed increases in consumption of sugars, animal-based foods and consumption of food outside the home (Popkin 2001; Popkin, Adair, and Ng 2012). This fact and the increased usage of motorized transportation, transition toward human-capital intensive, sedentary employment and preference shifts toward more inactive leisure activities have shaped the current obesegenic environment in China.

We are particularly interested in how income affects obesity outcomes in China. In the United States, obesity prevalence is overwhelmingly associated with low socio-economic status (Sobal and Stunkard 1989; McLaren 2007; Ogden et al. 2010). Accordingly, obesity has been shown to adversely affect cognitive ability early in life (Li 1995; Campos et al. 1996; Mikkilä et al. 2003), which may deter future educational in adulthood (Kristjánsson, Sigfúsdóttir, and Allegrante 2010). There is also increasing evidence that obese adults endure wage penalties. Several theories exist which may explain this phenomenon. Cawley (2007) finds that obese individuals are less productive and exhibit higher absenteeism in the workplace due to weight related sickness. Bhattacharya and Bundorf (2009) suggest that obese individuals may forego higher earnings for better health benefits because of the higher expected future medical costs. Obese people may also be subjected to discrimination in the workforce as Conley and Glauber (2007) found average employment status was lower among obese individuals. Additionally, obesity contributes to higher medical costs; Finkelstein et al. (2009) found that in aggregate, medical costs attributable to obesity-related conditions were about $147 billion in 2008.

In developing countries economic growth has increased access to more calories. As a result, countries like China have experienced great changes in nutritional profiles (Popkin 2001; Popkin, Adair, and Ng 2012). As noted by Cawley, Han, and Norton (2009) there is repeated evidence that the correlation between weight and wages is positive in developing countries. We hypothesize that income will be a positive contributing factor to obesity rates in China due to the relative newness of the obesity phenomenon and its correspondence with recent economic prosperity in the country. In fact, social norms pertaining to weight may actually encourage weight gain, as negative connotations associated with excess weight commonly held in western societies may not be established in Chinese society yet. Weight gain may still culturally-speaking connote better health and affluence in a culture which historically suffered from long-term food scarcity (McLaren 2007). Therefore, individuals with newfound financial autonomy and recollection of past food scarcity may increase food consumption and forgo certain investments in long-term health such as calorie restriction and physical leisure activities.

There are several challenges to conducting empirical analysis of obesity studies. First, weight outcomes are likely a result of many endogenous processes. For example, excessive carbohydrate intake may lead to weight gain, but having a high weight may also induce more carbohydrate consumption. Additionally, unobservable characteristics that determine income may also contribute to a person's weigh status. For example, a person with high levels of perseverance might receive higher wages as he is more likely to invest in higher education, but it may also contribute to more persistent weight maintenance and result in normal weight status and may attribute to negative correlations with wage and obesity in developed countries. Issues with endogeneity can be circumvented using panel data as the lags of the endogenous variables can be used as instruments to estimate consistent models.

The existing empirical literature on obesity in China is largely static and has ignored issues with endogeneity inherent in the consumptive variables (Xu et al. 2005; Shimokawa, Chang, and Pinstrup-Andersen 2009; Cai et al. 2013; Xiao et al. 2013). Du et al. (2004) and Chen and Meltzer (2008) use random effects (RE) and fixed effects (FE) models respectively to address individual heterogeneity but fail to instrument for income which is endogenous. Furthermore, the RE model is only consistent if the individual heterogeneity is assumed to be independent of the regressors, which is mostly like not the case. Additionally, while the FE model does control for the correlation between unobserved individual idiosyncrasies it also removes important time invariant (ie. gender) and nearly time-invariant variables (ie. education). Moreover, both the RE and FE does not permit the use of the lagged dependent variables of the regressors, which in the case of dynamic models is important for consistent estimation. To ameliorate this gap in the literature we estimate a dynamic model for the demand of health stock and health inputs, which allows us to make casual inference on the contributing factors of weight. [14]

As the basis for our empirical framework, we turn to Grossman's seminal paper (Grossman 1972a; Grossman 1972b) on the demand for health capital for a dynamic model of health. The critical contribution of the Grossman Model is acknowledgement that health inputs such as medical care are not in and of themselves demanded, but rather are the investment goods used to beget health, a durable capital stock that

[14] Ng et al. (2012) provide a dynamic correlative analysis of obesity in China. Their analysis does not use a multi-stage estimation as is specified by the Grossman Model and as such their results cannot be used to make causal inference.

depreciates over time. In essence, the Grossman Model partially assumes individual sovereignty over one's length of life through investment choices made on health inputs at each time period.

The purpose of this paper is to understand the long-run effects of income on obesity. To do so, we estimate the production function for health and health inputs of adults in China, which are functions of income. We use the Grossman Model as a theoretical platform to model obesity as a dynamic problem and to implement this framework we use a system GMM estimation.

This paper makes a number of contributions. First, we extend the Grossman Model to incorporate dietary regimen as a health input, which to our knowledge has not been estimated previously for developing nations.[15] Second, our use of panel data allows us to make empirical refinements to the existing, predominantly static empirical literature. Furthermore, we are able to make causal inference of the effect of income and education on BMI since we control for endogeneity present in these variables.

Our results show that income and education both have a direct impact on the demand for health stock. We find that income promotes increased BMI. Contrastingly, education was found to be negative and suggests that increased education mitigates weight gain. When analyzing results from the input demands we find evidence which shows that income contributes negatively toward the demand for carbohydrates, but is positive for both protein and fat demands. In the case of developing countries cultural perceptions of health and in particular weight maybe we different than those

[15] Goldman, Lakdawalla, and Zheng (2009) do present a dynamic model for health using nutritional inputs of elder individuals in the United States, but do not explicitly estimate the demands for these nutritional inputs.

held in developed countries. As a result, health seeking behavior may illustrate itself differently in developing countries. Our results show evidence of this behavior as measures of socio-economic status (income) where shown to increase the consumption of protein and fat; macronutrients which are more likely to be perceived as weight gaining.

The paper is organized as follows. Section 2.2 presents a derivation of the Grossman Model using nutritional regime as the health inputs. Section 2.3 provides the empirical model, which addresses the dynamic nature of obesity. Section 2.4 describes the results from our model and section 2.5 provides limitations of research and discussion.

2.2 Theoretical Model

The customary approach to modeling long-run health is to use the Grossman Model, which assumes that individuals act as producers of their own health stock by investing in inputs such as medical care and dietary regimen. Grossman also defines health as a durable capital stock different from other forms of human capital such as education. Specifically, knowledge stock is used to increase market and non-market productivity while health stock controls the amount of time that can be devoted to producing income. Equations 2.1 - 2.4 presents the individuals optimization problem.

$$\int_0^T e^{\rho t} U[s(H(t)), Z(t))], \tag{2.1}$$

$$\dot{H}(t) = I(D(t), t^p) - \delta(t, R(t))H(t) \tag{2.2}$$

$$T = \{t : H_t \leq H_{death}\} \tag{2.3}$$

$$\dot{A}(t) = rA(t) + Y\left[sH(t)\right] - \pi^H(t)I(t) - \pi^Z(t)Z(t) \tag{2.4}$$

2.1 presents the individual's inter-temporal utility as a function of sick days, s(.) and consumption of non-health input goods, Z(t). ρ is a time discount factor and H(t) is health stock at t. It is assumed that utility decreases as the number of sick days increases $\frac{\partial U(t)}{\partial s(t))} < 0$, utility increases with increases in non-health good consumption, $\frac{\partial U(t)}{\partial Z(t)))} > 0$ and the number of sick days decreases as health stock increases, $\frac{\partial U(t)}{\partial H(t))}$ < 0. Each individual receives initial endowments of health stock H_0. H_t is the net investment of health and is considered to be a durable capital stock and varies over time according to 2.2, where, I(.) captures gross investments in dietary regimen, D(t) used to augment health and t^p is time spent toward preventing illness. $\delta(.)$ is the depreciation rate and is a function of time and environmental factors R(t) that impact health.

We assume that depreciation of health increases over time, $\dot{\delta} > 0$. 2.3 shows that death occurs at T, when the depreciation is so high and the cost of investment to replenish health is so excessive that the health stock depreciates to H_{death}, the death stock. In a sense this model grants autonomy to the individual by allowing him or her to choose the length of life through investment choices made at each time interval leading to death.

Over their lifetimes, individuals accumulate pecuniary assets A(t) given by 2.4. For every time period t, individuals accrue income from investment earnings, rA(t) and wages, Y(.) minus outlays of health and consumption goods. We assume that r is exogenous and Y(.) decreases as the number of sick days increase $\frac{\partial Y(t)}{\partial s(t))} < 0$; π^H and π^Z are marginal/average costs for health investments and consumption goods respectively.

Individuals choose their optimal time paths for H(t) and Z(t) given in 2.1 subject to 2.2 -2.4. The resulting first order condition for health capital is (see Appendix II for derivation)

$$\left\{ \frac{\partial U(t)/\delta s(t)))}{\partial \pi(0)} e^{-(\rho - r)t)} + \frac{\partial Y(t)}{\partial s(t)} \right\} \frac{\partial s(t)}{\partial H(t)} = \left\{ r + \delta(t) - \frac{\dot{\pi}^H(t)}{\pi^H(t)} \pi^H(t) \right\} \tag{2.5}$$

where $\pi(0)$ is the shadow price of initial assets and $\frac{\dot{\pi}^H(t)}{\pi^H(t)}\pi^H(t)$ is the change in marginal cost of health investments. The first term on the LHS of 2.5 represents the marginal benefit due to increased consumption of health from a reduction in sick days, while the second term denotes the change in production as a result of health investments that reduce sick days. The RHS of 2.5 illustrates the marginal cost of health capital and is composed of interest foregone from pecuniary investments in health, depreciation, and the change in price of health investments.

The Grossman Model is characterized by two simultaneous demand equations: the demand for health and the demands for health inputs. To get estimable equations for the demand for health and health inputs we simplify the model such that the remainder of our analysis will refer to the pure investment model.[16] The pure investment model assumes that $\frac{\partial U(t)}{\partial s(t))} = 0$. Therefore, we derive the demand for health by simplifying 2.5 and taking logs to get the following

$$ln \frac{\partial s(t)}{\partial H(t)} - lnw(t) = ln\delta(t) + ln\pi^h(t) - ln\Psi(t) \tag{2.6}$$

where w(t) is the market wage and $\Psi(t) = \frac{\delta(t)}{r+\delta(t)-\dot{\pi}^H(t)/\pi^H(t)}$. To get an estimable form of 2.6 we used disclosed functional forms for s(.), $\delta(.)$ and $\pi(.)$ presented by Grossman (1972b), Wagstaff (1986) and Cropper (1981). We assume that s(.) takes on the form

$$s(t) = \alpha_1 H(t)^{-\alpha_2} \tag{2.7}$$

and α_1, α_2 >0. The depreciation function is defined to be

$$ln\delta(t) = \delta_0 + \alpha_3 t + \alpha_4 Z(t). \tag{2.8}$$

We assume that the investment function is Cobb-Douglas with constant returns to scale and is produced by time and dietary intake,

$$I(t) = D(t)^{\alpha_5}(t^p)^{1-\gamma-\alpha_5}E(t)^{\alpha_6} \tag{2.9}$$

[16]Following Grossman (1972b), Cropper (1981), Wagstaff (1986), Wagstaff (1993) and Nocera and Zweifel (1998), we estimate the pure investment sub-model. The main appeal with this approach is that is avoids non-linear estimation.

where E(t) is education and is an environmental variable and α_5 + 1 - γ - α_5 + $\alpha_6 \equiv 1$ and $0 < \gamma \leq 1$. Constant returns to scale gives rise to the marginal cost of investment given by,

$$ln\pi^H(t) = (1 - \gamma - \alpha_5)lnw(t) + \alpha_5 lnP^D(t) + \alpha_6 E(t). \tag{2.10}$$

We assume that the cost for health investment does not change over time such that $\dot{\pi}^H(t) = 0$. After simplifying, $\Psi(t) = \frac{\delta(t)}{r} + \delta(t)$ which we assume to be increasing with t so Ψ is defined as

$$\Psi(t) = \alpha_7 t \tag{2.11}$$

Substituting 2.7 and 2.11 into 2.6, the demand for health function is as follows

$$\begin{aligned} lnH(t) = \ & \frac{1}{1+\alpha_2}(A_1 + (\alpha_5 + \gamma)lnw(t) - \alpha_5 lnP - (\alpha_3 - \alpha_7)t \\ & + \alpha_6 E(t) - \alpha_4 Z(t) + u_1(t)) \end{aligned} \tag{2.12}$$

where $A_1 = ln(\alpha_1\alpha_2)$ and $u_1(t) = -ln\delta_0$ and α_3 -$\alpha_7 > 0$.

The demand for health inputs must also be estimated. Using 2.2, 2.9 and the cost minimizing condition for health investment we get the demand for health inputs

$$\begin{aligned} lnD(t) = & A_2 + H(t) + (1-\alpha_5)lnw(t) - (1-\alpha_5)lnP^D(t) \\ & + \alpha_3(t) + \alpha_4(t)Z + \alpha_6 E(t) + u_2(t) \end{aligned} \tag{2.13}$$

where $A_2 = -(1-\alpha_5)ln\left[\frac{1-\alpha_5}{\alpha_5}\right]$, P^D is a vector of prices of macronutrients and $u_2(t) = ln\delta_0 + ln\left(1 + \frac{H(t)}{\dot{H}(t)}/\delta(t)\right)$. If we substitute 2.12 in to 2.13 we get the following reduced form function of health inputs,

$$
\begin{aligned}
lnD(t) =& A_2 + \frac{A_1}{1-\alpha_2} + \left(1 - \alpha_5 + \frac{\alpha_5}{1-\alpha_2}\right) lnw(t) \\
& - \left(1 - \alpha_5 - \frac{\alpha_5}{1-\alpha_2}\right) lnP^D(t) + \left(\alpha_3 + \frac{\alpha_9 - \alpha_3}{1-\alpha_2}\right) t \\
& + \left(\alpha_4 - \frac{\alpha_4}{1-\alpha_2}\right) Z + \alpha_6 + \left(\frac{\alpha_6}{1-\alpha_2}\right) E(t) + \frac{u_1(t)}{1-\alpha_2} + u_2(t). \qquad (2.14)
\end{aligned}
$$

To make 2.12 dynamic, a lagged health term must be included on the RHS side. Moving to discrete time, 2.2 is analogous to $H(t) - H(t-1) = I(t-1) - \delta(t)H(t-1)$. In addition, we deduce from 2.12 that

$$H(t) = F(X(t), W(t)) \qquad (2.15)$$

where X and W are variables related to the procurement and production of health respectively. In order to be able to include lagged levels of health stock as regressors, a relaxation of the Grossman model allowing partial adjustment of health stock over time is needed. Thus $H(t) - H(t-1) = \kappa(\tilde{H}(t) - H(t))$ where $\tilde{H}(t)$ represents the ideal level of health stock at t and $0 \leq \kappa \leq 1$ and is the fractional health adjustment. In Grossman's original formulation $\kappa \equiv 1$, which assumes instantaneous health adjustment. Since in reality this is surely not the case, substituting the partial adjusted health stock condition into 2.15 yields

$$H(t) = G(H(t-1), X(t), W(t)) \qquad (2.16)$$

2.12 then becomes

$$
\begin{aligned}
lnH(t) = & \ \kappa H(t-1) + \frac{1}{1+\alpha_2}(A_1 + (\alpha_5 + \gamma) lnw(t) - \alpha_5 lnP - (\alpha_3 - \alpha_7)t \\
& + \alpha_6 E(t) - \alpha_4 Z(t) + u_1(t)) \qquad (2.17)
\end{aligned}
$$

In this paper we estimate the structural demand for health 2.17 and both reduced form and structural estimates for the demand for health inputs 2.13 and 2.14.

2.3 The Econometric Model

To estimate our model we employ a dynamic panel data estimation pioneered by Holtz-Eakin, Newey, and Rosen (1988)and later popularized by Arellano and Bond (1991). Blundell and Bond (1998) further refine this work to extend the estimator to a system GMM. The estimator uses a set of moment conditions of both the levels equation and the differenced equation. Arellano and Bover (1995) and Blundell and Bond (1998) later found that lagged levels performed poorly for the transformed equations particularly if variables behaved closely to random walks. As a result, the estimation was expanded to include both lagged levels and lagged differences as instruments. In our estimation we use the expanded system GMM estimation for our analysis.

The Grossman Model is a multi-level estimation. We start by estimating the demand for health stock thus giving rise to the following AR(1)

$$B_{i,t} = \beta B_{i,t-1} + \Theta Y_{i,t} + \mathbf{Z'_{i,t}}\zeta + \mathbf{P'_{i,t}}\varphi + \boldsymbol{T'}\xi + \omega_i + \varepsilon_{i,t} \tag{2.18}$$

where $B_{i,t}$ represents the current level of health proxied by BMI for individual i at t; $B_{i,t-1}$ is the lagged value of health thus making the estimation dynamic and $|\beta| < 1$; $Y_{i,t}$ is total current income; $Z_{i,t}$ is a vector of consumption and environmental variables which includes education; $P_{i,t}$ is a vector of food prices; T is time; ω_i is the time invariant fixed effects and $\epsilon_{i,t}$ is the disturbance term.[17]

[17]This is also supported in the scientific literature, which has found lagged measures of BMI to be a significant determinant of current BMI (Block et al. 2013). In addition, previous empirical

To remove fixed effects the traditional approach is to use the first difference (FD) transformation. However, in the case of unbalanced panels the disadvantage of using the FD transformation is that it amplifies gaps in the data. As an alternative, Arellano and Bover (1995) suggest to instead use forward orthogonal deviations (FOD). In contrast with the FD transformation which subtracts the lagged values from the current values, the FOD transformation subtracts the average of all available future periods from the current value. Thus, the FOD transformation is computable for each individual even in the presence of data gaps. Our transformed equation is the following:

$$B_{i,t}^{FOD} = \beta B_{i,t-1}^{FOD} + \Theta Y_{i,t}^{FOD} + Z_{i,t}^{FOD'}\zeta + P_{i,t}^{FOD'}\varphi + \varepsilon_{i,t}^{FOD}\text{[18]} \tag{2.19}$$

The demands for inputs are similarly estimated using system GMM. The dependent variables are the three macronutrients: protein, fat and carbohydrates. Following Grossman, structural estimation of inputs is a function of the predicted values of health stock, $\hat{B}_{i,t}$ calculated in 2.19. As a result, the levels model for each macronutrient is

$$N_{i,t} = \beta N_{i,t-1} + \gamma \hat{B}_{i,t} + \Theta Y_{i,t} + Z_{i,t}^{'}\zeta + P_{i,t}^{'}\varphi + \omega_{i,t} + \varepsilon_{i,t}\text{[19]} \tag{2.20}$$

analysis from Wagstaff (1993), Goldman, Lakdawalla, and Zheng (2009) and Ng et al. (2012) also specify that the demand for health stock to be an AR(P) process.

[18]For example, the FOD of the dependent variable is: $B_{i,t}^{FOD} = B_{i,t} - \frac{\sum_{t=1}^{T} B_{i,t+1}}{T}$.

[19]While we only specify the structural form for the demand for health inputs the reduced form equations can easily be deduced. Our estimation follows Nocera and Zweifel (1998) analysis and we also estimate reduced form estimates for comparative purposes.

$N_{i,t}$ is a 3×1 vector indicating the macronutrient profile of person i at t and is a function of previous consumption levels of N. Similarly, we use FOD to transform 2.20 to get the following

$$N_{i,t}^{FOD} = \beta N_{i,t-1}^{FOD} + \gamma \hat{B}_{i,t}^{FOD} + \Theta Y_{i,t}^{FOD} + Z_{i,t}^{FOD'} \zeta + P_{i,t}^{FOD'} \varphi + \varepsilon_{i,t}^{FOD} \tag{2.21}$$

Obesity studies are plagued with many econometric hurdles. In particular, weight outcomes are a result of many endogenous processes. For example, excessive carbohydrate intake may lead to weight gain, but a higher weight may also induce increased carbohydrate consumption. Additionally, unobservable characteristics that determine income may also contribute to a person's weigh status. For example, a highly dedicated individual might receive higher wages because of more investments he or she is willing to put forth toward education. This attribute may also make this individual more apt to invest in health. As a result, he or she may be more likely to be of normal weight. Measures of obesity such as Body Mass Index (BMI) are subject to high amounts of inertia since intertemporal changes in BMI maybe small in consecutive periods.[20] However, small perturbations in weight related behavior may eventually lead to large changes in weight outcomes in the long-run.[21]

We choose to estimate our model using system GMM to take advantage of several analytical conveniences. The use of instruments to control endogenity is best practice, but consistent estimation is dependent upon finding strictly exogenous instruments (Cameron and Trivedi 2005). System GMM allows us to exploit the longitudinal nature of our dataset and use lagged values as instruments for endogenous variables.

[20] BMI is equal to weight (kg)/height2(cm)

[21] Cutler, Glaeser, and Shapiro (2003) found that an average daily caloric excess of 100 calories explains the spread of obesity in the United States observed from the last 30 years.

Additionally, we are able to capture the dynamic behavior for health stock and health inputs with the autoregressive specification of the model.

2.3.1 Data

This paper uses data from the China Health and Nutrition Survey (CHNS), a longitudinal dataset conducted across nine Chinese provinces.[22] In our estimation we use the four most recent complete waves (2009, 2006, 2004, 2000).[23] We limit our analysis to adults 18 years and older as health markers may not be appropriate for children who usually cannot make individual decisions about their health. CHNS contains extensive individual and household-specific information on health outcomes, dietary intake, demographic and economic variables. Additionally, pricing data for foods came from the National Bureau of Statistics of China annual data. Pricing information is made available from 1996 onward in the form of annual price indices for each province.

The original data set contains 60,332 observations with 17,653 individuals. Upon restricting our sample to adults with information on variables included in the estimation we are left with 6,665 observations for 4064 groups.[24] The large amount of sample attrition is similar to attrition rates found in Ng et al. (2012). This is attributable

[22] The provinces surveyed are: Liaoning, Heilongjiang, Jiangsu, Shandong, Henan, Hubei, Hunan, Guangxi and Guizhou.

[23] The latest round of data was collected in 2011; however nutritional information is not yet available and thus was excluded from this study.

[24] At first glance it may appear that there are less than two waves of data on average for each individual. However, the number of observations is a bit misleading. Recall that the minimum number of waves required to execute system GMM is three. Differencing reduces the number of observations to two and using the lagged dependent variable further reduces to one. Therefore, individuals with one reported observation have at least three waves of data.

to missing information from non-responsiveness of individuals as well as exclusion of new entrants since there were not at least two rounds of data for these individuals.

Previous empirical applications of the Grossman Model have typically used number of healthy days to measure health stock. In this paper, we use the log of BMI to measure obesity which was collected during physical exams. Standard BMI thresholds established by the World Health Organization (WHO) have been proven to underestimate adiposity and health risk among Asian populations (Zhou 2002; Misra 2003). For this reason adjusted BMI categorizes determined by Wu (2006) are used instead. Table 2.1 shows the differences between WHO standards and adjusted standards.

Table 2.1: BMI Classification

WHO Standardized		Asian Adjusted (Wu 2003)	
BMI Range	Obesity Classification	BMI Range	Obesity Classification
< 18.5	Underweight	< 18.5	Underweight
$[18.5, 25)$	Normal	$[18.5, 23.9)$	Normal
$[25, 30)$	Overweight	$[24, 27.9)$	Overweight
$[30, 34.9)$	Obese I	≥ 28	Obese
$[34.9, 39.9)$	Obese II		
≥ 39.9	Obese III		

We restrict our data sample to include only individuals with a BMI of least 18.5 or the minimum normal adjusted BMI value the first time entering the survey. We limit our analysis to only these individuals because deviations from the ideal BMI could signify worsening of health for those who were normal, but dropped weight in subsequent periods such that they were underweight. Table 2.2. shows a steady increase in average BMI from 1997 to 2009. Logged values of BMI show an increase from 3.14 to 3.16, which roughly translates to geometric BMI levels of 22.6 and 23.6 respectively and is just within the normal adjusted BMI range. This constitutes roughly about a 7 pound increase in weight on average.

Table 2.2: Summary Statistics

	Means			
Variable	2000	2004	2006	2009
$log(BMI)$	3.14	3.14	3.15	3.16
$log(CarbohydrateIntake)(g)$	5.76	5.72	5.68	5.64
$log(FatIntake)(g)$	4.18	4.15	4.07	4.15
$log(ProteinIntake)(g)$	4.15	4.13	4.13	4.13
$log(Income)$	9.83	9.55	9.74	10.23
$Urban$	0.33	0.35	0.34	0.32
$Female$	0.52	0.52	0.52	0.52
Age	46.38	49.12	50.48	53.19
Age^2	2363.45	2623.28	2754.77	3018.74
$Education$	6.78	7.44	7.48	7.2
$Job\ Sedentary$	0.18	0.22	0.2	0.19
$Alcohol$	0.36	0.33	0.32	0.34
$Household\ Chores(days)$	0.05	0.04	0.04	0.04
$Sedentary\ time\ per\ day(days)$		0.2	0.19	0.2
$log(Price_{grain}$	4.75	4.89	4.92	5.08
$log(Price_{oil\ \&\ fat})$	4.64	4.69	4.55	5.03
$log(Price_{meat\ \&\ poultry})$	4.74	4.95	4.98	5.43
$log(Price_{eggs})$	4.5	4.58	4.63	4.83
$log(Price_{aquatic\ products})$	4.6	4.71	4.8	5
$log(Price_{vegetables})$	5.02	5.19	5.38	5.66
$log(Price_{fruit})$	4.94	4.94	5.06	5.35
$log(Price_{alcohol})$	4.75	4.75	4.75	4.87

Log(Carbohydrates), log(Fat) and log(Protein) compose the vector of dependent variables for the demand for health inputs estimation. Carbohydrate, fat and protein consumption are represented in average daily intake in grams and should be thought of a proxies for consumption of carbohydrate, fat and protein rich foods respectively. To collect these variables each household member was observed and evaluated by a surveyor over the course of three consecutive days. Once their dietary consumption was collected, macronutrient information was compiled from these records and then averaged over three days. From table 2.2. we see that there is a clear decrease in consumption of carbohydrate and protein intake over the 12 years. Average fat intake decreased up till 2006, but quickly rises again in 2009.

2.3.2 Independent Variables

Many studies of developed countries have found that in fact low income is associated with higher incidence of obesity (e.g. Sobal and Stunkard (1989), McLaren (2007)). Conversely, Cawley, Han, and Norton (2009) note past correlational evidence to suggest that income and weight are in fact positively correlated in developing countries. We hypothesize that in developing countries such as China increases in income will in fact promote weight gain because cultural perceptions of weight may interpret weigh gain as a sign of affluence (Sobal and Stunkard 1989; Monteiro et al. 2004). We treat income as endogenous because it is a reflection of labor market choices made by the individual which may be influenced by weight. For example, overweight individuals may select jobs with better medical benefits and forego income because of future expectations of illness. Table 2.2. shows that log real income increased consistently throughout the 12 years and is a reflection of new economic prosperity experienced in China.

We include dummy variables for living in an urban area, gender and alcohol consumption. Age and age^2 is also included as demographic variables. Table 2.2. indicates that between 32%-35% of the population lived in urban areas across all years. Approximately 52% of the sample is female and the percentage of sedentary jobs stays relatively constant over the years from approximately 18% to 19%. Again, the increase in sedentary jobs maybe a reflection of economic growth and employment shifts toward skilled labor.

Education is also included in our estimation. Much like income, we assume that schooling is also endogenous both in the demand for health stock and demand for health inputs estimation. Li (1995) found that obesity adversely affected cognitive skills of a sample of Chinese elementary-aged children. Likewise, Campos et al. (1996) studied the intellectual disparities between Brazilian children and found that obese children had lower IQs and Mikkilä et al. (2003), found similar results, using a sample of Finnish students. Obesity may therefore deter educational attainment in adulthood (Kristjánsson, Sigfúsdóttir, and Allegrante 2010).

Education is treated as endogenous in the demand for health inputs estimation because it is likely that cognitive abilities may influence the demand for health inputs (eg. decisions about consuming healthy foods), while health inputs may also encourage cognitive ability through physiological and metabolic processes (Smith et al. 2011). Table 2.2. shows that the average number of years of education increased over the 12 year span, though decreasing slightly in 2006. The overall increasing trend is a reflection of economic growth as the demand for skilled labor may encourage the demand of more schooling.

To capture time allocation behavior we also include daily averages of time spent doing household chores and sedentary pastimes.[25] Over the 12 year span time devoted to household chores decreases. When converted to minutes time devoted to sedentary activities went from 72 minutes per day in 2000 to 57.6 minutes in 2004 and beyond. The amount of sedentary time stayed relatively constant and accounts for 283.7 minutes per day.

In our estimation we use the logged prices of grain, vegetables, fruit, alcohol, meat and poultry, eggs, aquatic products and oil and fats. From table 2.2 we see that all prices increased during the 12 year period. Prices of meat and poultry experienced the most rapid price increase over the 12 year span. The price of vegetables was the highest overall for all years while the prices for fruit were second highest until 2009 when the relative price of meat superseded fruits. Grain, aquatic products and eggs exhibited similar price growth while the price of oil experienced a decrease in price in 2006 followed by a dramatic increase in price in 2009.

2.4 Results

2.4.1 Demand for Health

Table 2.3 shows that lagged BMI is significant and positive. The short-run income elasticity is positive and significant and is consistent with previous empirical studies, which find higher socioeconomic status is associated with weigh gain (McLaren 2007; Nube, Asenso-Okyere, and Van den Boom 1998; Monteiro, Conde, and Popkin 2004; Mendez et al. 2004; Shah et al. 2004; Chee et al. 2004; Ulijaszek 2003; Reddy 1998; Townsend et al. 2001). Income elasticity was slightly higher (0.03) in the long-run. It

[25] Sedentary pastime includes time allocated to watching television, playing video games, using the internet and reading.

Table 2.3: Demand for Health Stock

VARIABLES		VARIABLES contd.	
$log(BMI_{t-1})$	0.36***	2000	0.10***
	(0.05)		(0.03)
$log(Income)$	0.02*	*Job Sedentary*	0.04**
	(0.01)		(0.02)
$log(Price_{grain})$	0.09**	*Female*	-0.03
	(0.04)		(0.05)
$log(Price_{oil\ \&\ fat})$	-0.01	*Age*	0.01***
	(0.03)		(0.00)
$log(Price_{meat\ \&\ poultry})$	0.09***	Age^2	-0.00***
	(0.03)		0.00
$log(Price_{eggs})$	-0.15***	*Urban*	0.01
	(0.03)		(0.01)
$log(Price_{aquatic\ products})$	0.08***	*Education*	-0.01**
	(0.03)		(0.01)
$log(Price_{vegetables})$	0.03*	*Alcohol*	-0.00
	(0.02)		(0.09)
$log(Price_{fruit})$	0.04*	*Sedentary Activities (days)*	0.01
	(0.02)		(0.01)
$log(Price_{alcohol})$	-0.07*	*Household Chores(days)*	0.04**
	(0.04)		(0.02)
2006	0.05*	*Constant*	1.19***
	(0.03)		(0.41)
2004	0.08***		
	(0.02)		
Observations			6,665
Groups			4,064
Instruments			27
AR(2) Test			0.975
Sargan test			0.376
Hansen Test			0.549

Standard errors in parentheses
$***p < 0.01, **p < 0.05, *p < 0.1$

suggests that over longer periods of time, income impacts weight outcomes more so. This makes intuitive sense as individuals have more opportunities to consume calorie dense food and perhaps are more likely to form habits conducive to weight gain. For example, persistently higher income households may choose to eat more food away-from-home since it is time-saving. This behavior may eventually become habitual if households continue to consistently eat away-from-home. Since these meals are typically more calorie dense, it may lead to weight gain. The prices of meat and poultry and aquatic products are significant and positive while the price of eggs is negative. It might be expected that the sign of the coefficients should all be the same since all three foods are sources of animal proteins. However, the negative sign on eggs may illustrate less substitution toward other cheaper weight gaining foods or that individuals are substituting to other cheap proteins like legumes and vegetable-based proteins, which may not lead to weight gain. As expected, having a sedentary job contributed positively to a higher BMI. Education was surprisingly a slightly negative factor for weight gain and differs from previous literature. On possible explanation for this is that education may enforce healthier eating habits and thus mitigates weight gain. The negative coefficient on education may signify increased awareness of weight promoting behavior and more educated individuals maybe more likely to make healthful choices.

2.4.2 Demand for Health Inputs

The context and interpretation of the dependent variables is imperative here. Wagstaff (1986) considers the derived demand hypothesis, which leads us to expect a positive coefficient from the predicted health stock values in the structural estimation. Table 4. shows that for significant estimates of log BMI our coefficient is in fact negative. But, unlike Wagstaff (1986) and even Nocera and Zweifel (1998) our dependent variable interprets higher deviations of BMI (health capital) as decreases in health stock and as such when health stock is lower than ideally we should expect a decrease in consumption of food inputs since it contributes to weight gain.

Our first observation should be that both reduced form and structural estimates across all three inputs only vary slightly and are comparable in sign. In fact the variables which do have opposite sign prove to be all insignificant in the corresponding estimation. Like Nocera and Zweifel (1998), the use of panel data to capture dynamics of health ameliorates the estimation inconsistencies faced by Wagstaff (1986), in which his estimates found opposing signs in the coefficients. More explicitly, to address estimation concerns mentioned by Wagstaff (1986), for the most part log BMI is negative in the structural estimation and education and income have expected signs.

It should also be noted that Table 4. shows that we fail to reject the null hypothesis of the Hansen test for the reduced form estimation of protein. Since the structural estimation does pass the Hansen test, our discussion for the demand of protein is restricted to the structural estimates.

2.4.3 Carbohydrates

A closer look at the estimation for the demand for carbohydrates shows that lagged carbohydrate consumption is significant and positive in determining current carbohydrate consumption. This suggests that long-run carbohydrate consumption has some explanatory power on current consumption. We found income to be significant in the reduced form estimation with the expected sign since we expect increases in socioeconomic status in developing countries to induce a shift in diet away from cheap grains and toward more expensive, calorie density animal-products and fat (Popkin 1998;

Popkin 2001). In the long-run, the income elasticity (-0.10) was found to decrease carbohydrate consumption only slightly. Living in an urban area was significant and is shown to decrease the demand for carbohydrates. As expected, women are shown to demand less carbohydrates then men. In terms of the price variables, the logged price of vegetables was significant and negative, which is expected since vegetables are carbohydrate rich foods. In the long-run the price elasticity for vegetables was -0.24 indicating that individuals progressively consume less carbohydrates when high prices of vegetables persist. The logged price of oil and fat was found to be significant and negative which suggests that carbohydrates and fat are complements of each other and the price of meat and poultry was significant and positive for the reduced form estimates and reflects substitutive behavior between carbohydrates and protein. Lastly, logged BMI was not significant in the structural estimation, but was of the expected sign as we should expect a high BMI to induce a decrease in food consumption.

2.4.4 Protein

The structural estimation shows that income is highly significant and positive in determining the demand for protein. This effect is even more pronounced in the long run since income elasticity is 0.19. Again, these findings support the theory of nutritional transition discussed by Popkin (1998) and Popkin (2001). As expected we should see a rise in protein demand when individuals become richer and can afford to purchase more expensive meat-based products. The demand for protein as a health input may originate from cultural perceptions of health. Protein and more specifically, animal protein maybe believed to be a weight gaining food. In a country which historically suffered from food scarcity, higher weight may signify better health and so we see a positive relationship with income and demand for protein. Contrastingly, education is found to be a negative and significant determinant of the demand for protein, which suggests that education may mitigate the consumption of protein because of increased awareness of the impact on health from meat consumption. Moreover, age was shown to be a positive influence on the demand for protein, but was only very slightly quadratic while time allocated to household chores was also positive and significant. Similar to the demand for health stock results, the price elasticities for aquatic products and meat and poultry are significant and positive while eggs is negative and these effects increase in magnitude by 1% in the long-run. The price of grain also shows a negative impact on the demand for protein. Finally, logged BMI was as expected, very significant and negative.

2.4.5 Fat

Income was found to be a positive and significant factor to determining demand for fat in both the structural and reduced form estimations. In the long-run, income has a larger effect of the demand for fat as the elasticities were 0.19 and 0.25 for the structural and reduced form estimations respectively. Again, this may be due to cultural perceptions of weight and thus could induce more consumption of fat. Surprisingly, the price of oil and fat had an increasing effect on demand for fat and may signify that individuals substitute toward other sources of fat like that found in meat products when the price of oil increases. Again, this effect is far more pronounced in the long run as the price elasticities for oil and fat indicate 0.52 and 0.67 for the reduced and structural estimations, a 5 and 6% increase from the short run. Additionally, the price of meat and poultry is negative and suggest that fat consumption is a complement to meat consumption. Finally, logged BMI was shown to be a positive contributor to demand for fat. This is somewhat surprising as we would expect individuals with high BMI's to demand less fat. However, the positive sign of the variable maybe attributable to higher calorie demands of heavier individuals.

Table 2.4: Demand for Health Inputs

VARIABLES	**Carbohydrates**		**Protein**		**Fat**	
	Structural Form	Reduced Form	Structural Form	Reduced Form	Structural Form	Reduced Form
$log(CarbohydrateIntake_{t-1})$	0.09***	0.09***				
	(0.03)	(0.03)				
$log(ProteinIntake_{t-1})$			0.03	0.04		
			(0.03)	(0.03)		
$log(FatIntake_{t-1})$					0.10***	0.09**
					(0.03)	(0.04)
$log(BMI)$	-0.17		-1.99***		1.17***	
	(1.85)		(0.63)		(0.41)	
$log(Income)$	-0.08	-0.09*	0.18***	0.12***	0.17***	0.23***
	(0.08)	(0.05)	(0.03)	(0.03)	(0.05)	(0.06)
2009	0.17	0.18	-0.26***	-0.07	-0.03	
	(0.29)	(0.12)	(0.08)	(0.06)	(0.12)	
2006	0	0.01	-0.10***	-0.02	-0.02	0.17
	(0.15)	(0.06)	(0.04)	(0.03)	(0.05)	(0.16)
2004	0.02	0.02	-0.04	0	-0.01	0.18
	(0.08)	(0.04)	(0.03)	(0.03)	(0.04)	(0.14)
2000						0.2
						(0.18)
$Education$	-0.01	-0.01	-0.05***	-0.03*		
	(0.04)	(0.03)	(0.02)	(0.02)		
$JobSedentary$	0.01	0	0.09	0.04	-0.08**	-0.13**
	(0.13)	(0.08)	(0.06)	(0.06)	(0.04)	(0.06)
$Female$	-0.39**	-0.38*	-0.05	-0.11	0.23	0.75**
	(0.16)	(0.19)	(0.11)	(0.11)	(0.22)	(0.38)
Age	0.02	0.02**	0.02***	0	-0.01	-0.01
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
Age^2	0	-0.00***	-0.00***	0	0	0
	0.00	0.00	0.00	0.00	0.00	0.00
$Urban$	-0.05*	-0.05*	0.01	0.02	0.01	-0.07
	(0.03)	(0.03)	(0.03)	(0.02)	(0.04)	(0.06)
$Alcohol$	-0.38	-0.37	0.38*	0.17	0.58	1.55**
	(0.24)	(0.38)	(0.20)	(0.18)	(0.40)	(0.68)
$HouseholdChores(days)$	0.06	0.05	0.29***	0.20***	0.17	0.30*
	(0.09)	(0.07)	(0.07)	(0.06)	(0.11)	(0.16)
$Sedentarytime(days)$	0.06	0.05	0.04	0.06	0	-0.07
	(0.07)	(0.05)	(0.05)	(0.05)	(0.07)	(0.11)

Continued on next page

Table 2.4 – continued from previous page

$log(Price_{vegetables})$	-0.22**	-0.22***				
	(0.10)	(0.08)				
$log(Price_{fruit})$	-0.06	-0.07				
	(0.13)	(0.08)				
$log(Price_{alcohol})$	-0.02	0.01				
	(0.26)	(0.13)				
$log(Price_{aquatic\ products})$	-0.27	-0.29***	0.39***	0.11		
	(0.20)	(0.10)	(0.13)	(0.08)		
$log(Price_{eggs})$	0.07	0.11	-0.34***	0.01		
	(0.45)	(0.11)	(0.13)	(0.07)		
$log(Price_{oil\ \&\ fat})$	-0.36***	-0.35***			0.47***	0.61**
	(0.11)	(0.13)			(0.14)	(0.24)
$log(Price_{grain})$	-0.09	-0.09	-0.28***	-0.27***	-0.23	-0.17
	(0.31)	(0.15)	(0.11)	(0.10)	(0.15)	(0.24)
$log(Price_{meat\ \&\ poultry})$	0.37	0.34***	0.27**	0	-0.49***	-0.42***
	(0.29)	(0.10)	(0.12)	(0.09)	(0.08)	(0.11)
$Constant$	9.38**	8.84***	8.24***	3.87***	-0.35	0.74
	(4.21)	(1.31)	(1.52)	(0.54)	(1.57)	(2.26)
Observations	6,958	6,958	6,958	6,958	6,958	6,958
Groups	4,231	4,231	4,231	4,231	4,231	4,231
Instruments	29	27	29	27	27	25
AR(2) Test	0.201	0.352	0.394	0.331	0.437	0.256
Sargan test	0.155	0.133	0.105	0	0.623	0.458
Hansen Test	0.137	0.122	0.121	0	0.593	0.748

Standard errors in parentheses

$***p < 0.01, **p < 0.05, *p < 0.1$

2.5 Discussion

2.5.1 Limitations of Research

While there are many estimation advantages associated with panel analysis, specifically its ability to control for individual heterogeneity and variable persistence, a looming problem is its high vulnerability to attrition bias. This is of particular concern in our study since only 18% of the observations were retained for the estimations. Sources of attrition can be attributable to i) death of the individual in the survey, ii) non-response to questions in the survey and iii) non-participation of individuals in subsequent rounds. Because of this, it brings question to the accuracy and generality of our findings. Previous analysis using the CHNS have had similar attrition rates (Ng et al. 2012). Popkin et al. (2010) further discuss the collection problems with CHNS. In particular, there are issues with non-participation of individuals in later waves as they found difficulties in finding previous participants either because of migration, natural disaster and participant refusal to be subjected to the exams.

For this analysis we apply Lee Bounds (Lee 2005) to limit the the possible values of the variable effects. Since our analysis is concerned with socioeconomic status we calculate these bounds for income and education. Lee Bounds require that a treatment variable and a binary selection indicator variable be defined. In this case, the treatment variable is income or education in which we create a dummy variable indicating whether the individual is in the top quartile of the relevant distribution. The selection indicator was created from the post-estimation of the Arellano-Bond regressions and indicates whether the observation was included in the estimation or not. The 95% confidence interval for income and education are $[-0.1303,\ 0.1684]$ and

$[-0.0925,\ 0.0802]$ respectively. This is of concern since both intervals imply that income and education are not significantly different from zero. However, this result may not necessarily be unexpected. McLaren (2007) provide a tabulation of all previous associations between measures of socioeconomic status and obesity in medium and low developed countries. Her findings show that there are relatively equal numbers of articles, which show positive or non-significant associations. The Lee Bounds may then be a reflection of previous mixed results. For future research, techniques such as pseudo-panel analysis (utilized in chapter 3) can be used to mitigate these effects.

2.5.2 Conclusion

Obesity is a relatively new problem for developing countries, however the rate of obesity in China is spreading at an alarming rate. Developing countries are particularly vulnerable to the damaging economic effects obesity has as coexistence of malnutrition and obesity put a double burden on the medical systems of developing countries. Therefore, informative policy should be implemented to combat this phenomenon. In this paper we have estimated production functions of health stock and for health inputs. Our results show that income and education both have a direct impact on the demand for health stock. We find that income promotes increased BMI. Contrastingly, education was found to be negative and suggests that increased education mitigates weight gain. When analyzing results from the input demands we find evidence which shows that income contributes negatively toward the demand for carbohydrates, but is positive for both protein and fat demands. In the case of developing countries cultural perceptions of health and in particular weight maybe we different than those held in developed countries. As a result, health seeking behavior may illustrate itself differently in developing countries. Our results show evidence of this behavior as measures of socio-economic status (income) where shown to increase the consumption of protein and fat; macronutrients which are more likely to be perceived as weight gaining.

Chapter 3: A Pseudo-Panel Analysis of Obesity Prevalence using NHANES Data

3.1 Introduction

At present more than two thirds of adults are overweight and over a third are considered obese in the United States (Flegal et al. 2012). While this might seem alarming, current projections indicate that the incidence of obesity will rise to 42% by the year 2030 (Finkelstein et al. 2012), but could be as high as 50% (Levi et al. 2012). Obesity is related to increased morbidity from weight related illnesses such as diabetes, heart disease and stroke. It is estimated that $190 billion was spent on health care in 2005 related to obesity illnesses (Cawley and Meyerhoefer 2012). And, spending is only expected to rise as more obese individuals become sick (Wang et al. 2011). Indirect costs are additionally incurred through higher absenteeism due to obesity related illness (Cawley 2004), higher insurance premium payments (Bhattacharya and Bundorf 2009) and possible workforce discrimination (Conley and Glauber 2005). As such, considerable public attention is given to the obesity epidemic because of the economic burden it poses and the damaging effects on individual welfare.

The increased concern from obesity has spurred an extensive literature on its causes and incidence. The determinants of obesity are plenty and somewhat nebulous making it difficult to both treat and understand (Finegood 2011; Hammond 2009). Current social science literature has placed focus on the socioeconomic and behavioral factors of obesity. However,the vast majority of these studies are relegated to cross-sectional analysis despite having several explanatory limitations. Specifically, they cannot give insight into time dependent effects nor can they resolve issues related to variable persistence. Furthermore, when endogeneity is suspected it is a much more strenuous task to find viable instruments using cross-sectional data. When put in the context of obesity studies, which is inherently dynamic (Auld and Grootendorst 2011; Block et al. 2013) cross-sectional analysis does not lend itself particularly well to explaining the incidence of obesity.

McLaren (2007) updates Sobal and Stunkard (1989) seminal review on the literature and reports 333 journal articles all of which utilize cross-sectional data, while Ball and Crawford (2005) conducts a review of only longitudinal studies and found just 34 articles where less than half used US data all of which pre-date 2000. The predominance of cross-section analysis is due in part to a greater availability of publicly available cross-sectional data. Among studies which control for endogeneity in the socioeconomic variable (most often measured in income), Cawley, Moran, and Simon (2010) conduct a study on the effect of income on weight among the elderly. To control for endogeneity they exploit a natural experiment in which retirees receive different Social Security payments based on age. Their findings show that among the elderly there was no significant effect from income on weight. Similarly, Schmeiser (2009) takes advantage of inter-state variation in Earned Income Tax Credit (ETIC) allotted to young (25-43) and poor individuals. He found no significant effect of income on obesity among men, but found an inverse relationship with women.

This paper is consistent with the work by Cawley, Moran, and Simon (2010) and Schmeiser (2009) in addressing endogeneity in income, but extends their work to include dynamics in the model. To do so pseudo-panel methods are used to conduct a panel study. Of course, the best recourse when genuine panel data (GPD) is available is to use more typical panel methods. However, in the absence of GPD, as is in this case, pseudo-panel analysis is an alternative that first does not rely on GPD, but still allows for dynamic estimation. Pseudo-panel data are constructed by exploiting repeated cross-sectional data (RCS) in which at each cross-section individuals are pooled together to form statistically representative cohorts based on time invariant criteria[26]. Panel estimation can then be done at the cohort-level in which each time-stable cohort is represented in consecutive rounds of data despite the fact that individual membership to each cohort is different in each year. This paper applies pseudo-panel techniques to the Continuous National Health and Nutrition Examination Survey (NHANES) and aims to fill the deficiency of panel estimation in the literature. The estimation is framed by the Grossman Model (Grossman 1972b; Grossman 1972a), the standard dynamic human health capital model. As such, the system first estimates the demand for health stock, represented by body mass index (BMI) and persistence of overweight and obesity catergories, which is then used as a determinant for the estimation of demand for health inputs represented by food intake.

Previous applications of the Grossman Model have traditionally looked at the impact of medical services usage on health outcomes. Furthermore, these papers assume instantaneous health adjustment, precluding the use of past realizations of

[26] Alternatively, pseudo panel data is also known as synthetic panels

health as a contributing factor to current demands for health (Nocera and Zweifel 1998; Wagstaff 1986). Furthermore, Wagstaff (1993) states that exclusion of the lagged health stock "fails to take into account the inherently dynamic character of the health investment process". Goldman, Lakdawalla, and Zheng (2009) present a dynamic model of health, but their empirical analysis is not generalizable to the general public as they are confined to data on individuals of retirement age, who may not show considerable weight variation (Block et al. 2013).

This paper contributes to the literature in several ways. As discussed previously, there is a major research paucity in panel estimation on the incidence of obesity. To my knowledge this is the first paper to apply pseudo panel techniques in investigating the determinants of health status using US data. In addition, pseudo-panel techniques allow for the opportunity to exploit the rich nutritional data collected by NHANES, which is surprisingly absent in concurrent longitudinal studies. Finally, previous articles that have applied Grossman's Model typically have estimated static models (Wagstaff 1986; Nocera and Zweifel 1998). This paper extends the Grossman Model such that a dynamic model is estimated.

Results show that consistent with past cross-sectional studies, women show an inverse relationship with socioeconomic status and weight. Additionally, marriage was also a decreasing factor for women's weight outcomes. There was no apparent relationship between socioeconomic status and men (also consistent with prior literature), but they seemed to show moderate amounts of weight gain with marriage. Interestingly, past realizations of weight were positive and significant for women only as the results for men were largely variable and insignificant. This would indicate that men are resistant to a *"weight legacy"* and are more able to transform their current weight despite previous weight status, which does not seem to be the case for women. In terms of food consumption, there was a general trend showing that the youngest cohorts eat less animal protein and more added sugar. Female demand for added sugar was shown to be positively determined by socioeconomic factors, marital status and elevated weight status. This behavior may explain the differences in obesity incidence between gender, women being more likely to be overweight and obese.

The paper is organized in the following manner. Section 3.2 is a literature review. Section 3.3 presents a brief summary of the the theoretical and empirical framework, essentially laid out by Wagstaff (1993). Section 3.4 provides descriptive statistics and results and section 3.5 provides conclusions.

3.2 Literature Review

The obesity epidemic is a widespread problem particularly pronounced in developed countries like the United States. The current social science literature has focused on the impact of socioeconomic factors on weight gain, the results of which have been largely mixed. In McLaren (2007)'s literature review of cross-sectional studies, she finds that the prevalence of studies finding negative associations between income and obesity among women are approximately equal in number to studies which found no particular significance. With men this disparity is much more pronounced as there were more than 4 times the number of insignificant associations as there were ones that found a negative correlation. Amongst studies that pooled both genders the number of articles finding no significance was slightly more common then those that found a negative significance. And, for all categories only a small minority of studies found income and obesity to be positively related.

In what predates the pseudo–panel literature by Moffitt (1993) and Verbeek and Vella (2005), Flegal, Harlan, and Landis (1988a) and Flegal, Harlan, and Landis (1988b) present a model which compares the average BMI of sub-populations determined by race, gender, age, education and income across time. Their findings show varied associations between BMI and education and a positive relationship to income for men. Women were found to have a distinctively negative relationship between education, income and average BMI. Racial disparities where only detected among women where average BMI for black women was significantly higher than white women. Alternatively, Zhang and Wang (2004) found no significance between incidence of obesity and socioeconomic inequality among women. However, this association was positive and significant for male incidence of overweight, but was more pronounced among white males then in minorities. The findings on women were more consistent and were shown to be negatively correlated with socioeconomic inequality.

Behavioral factors are also important in determining obesity and are not considered in Flegal, Harlan, and Landis (1988a), Flegal, Harlan, and Landis (1988b) and Zhang and Wang (2004). Of the articles that incorporate food intake, Jeffery and French (1996) found that fat consumption was much higher among women with lower socioeconomic (SES) standing and efforts to control weight were less likely among low SES women. Alternatively, Lin, Huang, and French (2004) found that food away from home, which is typically more fat dense was a more pronounced positive factor for wealthy women, while vegetarianism elicited lower BMI's among the poorest women.

This paper makes two refinements from the previous work described above. First, while Flegal, Harlan, and Landis (1988a) and Flegal, Harlan, and Landis (1988b) may not have intended to do so, they lay the ground work for pseudo-panel analysis by essentially constructing cohort-level means and variance trends. However, stratifying the cross-sections by age, education and income which are time variant does not guarantee cohort stability and thus cannot be used to make causal inference using pseudo-panel analysis. Second, Jeffery and French (1996) and Lin, Huang, and French (2004) incorporate behavioral variables as determinants for weight. Alternatively, following the Grossman Model this paper estimates the demand for food intake and uses measures of average BMI as one of the determining factors. Like Cawley, Moran, and Simon (2010) and Schmeiser (2009) this paper also controls for endogeneity, but extends their work to include dynamics in the model. The findings from this study will provide evidence of causal relationships between income and obesity and food intake behavior and obesity.

3.3 Theoretical and Empirical Framework

3.3.1 The Grossman Model

The Grossman Model characterizes the optimizing individual as one who maximizes utility, subject to a time-dependent process of health stock, 3.1.[27]

$$H_t - H_{t-1} = I_{t-1} - \delta_{t-1} H_{t-1} \tag{3.1}$$

Individuals are endowed with H_0 while succeeding periods of health are determined by 3.1. H_t is current health stock, I_{t-1} is gross investment on health in the previous period and δ_{t-1} is the depreciation rate of health. The individual must weigh the

[27]For a thorough derivation of the pure investment model refer to section 1 of chapter 2 and appendix C.

benefits of good health with the costs of procuring and maintaining it over time. As derived in chapter 2 (equation 2.5) , the resulting optimal condition is

$$i_t + c_t = (r + \delta_t - \tilde{\pi}_{t-1})\pi_t \tag{3.2}$$

where i_t is the monetary benefit, c_t is the non-monetary benefit from good health, r is the interest rate, π_t is the marginal cost of investment and $\tilde{\pi}_{t-1}$ is the percent change.

To make the Grossman Model estimable, simplifications are made to 3.2 such that either i_t or c_t are only estimated.[28] Grossman's Model is implemented first by estimating the demand for health stock, which is then used as a determinant for the structural estimation of health inputs. To make such computations possible, functional forms for i_t(or c_t), δ_t and π_t are assumed.[29] As derived in equation 2.12, the demand for health stock for the pure investment model is

$$lnH_t = \beta_1 lnw_t - \beta_2 P_t^D + \beta_3 t + \beta_4 E_t + \beta_5 Z_t \tag{3.3}$$

where w_t is wage, P^D is the price of inputs , E_t is education and Z_t is a vector of behavioral variables, which effect the production of health.[30]

[28]The simplified models are aptly named the pure investment model and the pure consumption model (Grossman 1972a; Grossman 1972b).

[29]See appendix C for derivations

[30]Examples of Z_t are marital status, smoking, alcohol consumption, marital status etc $\cdots$

To derive the demand for health inputs the cost minimization condition for health investment, the Cobb-Douglas investment production function, $I_t = f(F_t, t, E_t)$ and the log of 3.1 are used. The structural equation is the following

$$lnF_t = lnH_t + (1-\beta_1)lnw_t - (1-\beta_2)lnP_t^D + \beta_3 + \beta_4 E_t + \beta_5 Z_t \qquad (3.4)$$

where F_t is a vector of health inputs (food intake).

3.3 is notably missing a lagged health stock term. This is because the typical derivation of the Grossman Model assumes instantaneous health adjustment. Specifically, under instantaneous adjustment 3.3 states that

$$H_t = \bar{H}_t = f(w_t, P_t^D, t, E_t, Z_t). \qquad (3.5)$$

where $\bar{H}_t$ is the ideal level of health stock. In the case of partial adjustment of health stock, let $\bar{H}_t$ be the ideal level of health stock such that

$$H_t \leq \bar{H}_t = f(w_t, P_t^D, t, E_t, Z_t). \qquad (3.6)$$

and

$$H_t - H_{t-1} = \alpha(\bar{H}_t - H_{t-1}). \qquad (3.7)$$

The special case of instantaneous adjustment is when $\alpha \equiv 1$ and the ideal health stock corresponds to the actual level of health. If $0 \leq \alpha \leq 1$ then substituting 3.6 into 3.7 results in

$$H_t = g(H_{t-1}, w_t, P_t^D, t, E_t, Z_t). \qquad (3.8)$$

Then 3.3 can be written as

$$lnH_t = \alpha lnH_{t-1} + \beta_1 lnw_t - \beta_2 P_t^D + \beta_3 t + \beta_4 E_t + \beta_5 Z_t \tag{3.9}$$

and is the full dynamic estimable model.

3.3.2 Pseudo–Panel Data Sets

First proposed by Deaton (1985), when repeated cross-sectional (RCS) data is available, cohorts can be constructed where membership is based on relevant time-stable criteria.[31] The result being that the constructed data can be treated as if it were GPD at the cohort-level. There are several advantages to using pseudo panels. As Deaton (1985) shows, pseudo-panels can be constructed per cohort as long as the same time invariant variables are collected each year. This flexibility affords potential to combine datasets which collect the same information. Second, the construction of pseudo panels has a mitigating effect on data attrition where the averaged values smooth over instances of non-response in each cohort. In fact, even in cases where GPD is available, when T is large, pseudo-panels maybe still have an upper hand in answering long-run questions since it is less subject to attrition bias.

Suppose we wish to estimate the following dynamic regression model:

$$y_{it} = \gamma y_{i,t-1} + x_{i,t}^{'} \beta + \alpha_i + u_{i,t}. \tag{3.10}$$

In the context of RCS data it is obvious that $y_{i,t-1}$ is unobservable. In order to estimate 3.10 the data can be pooled together to form synthetic panels and unlike, GPD each cross-section contains a new set of individuals. Cohorts are then created

[31] For example, Deaton (1985) creates cohorts based on birth year intervals

such that member i of cohort C will always belong to cohort C for any given T. For example, a non-Hispanic white male born in 1970 will belong to the cohort of white, male individuals born between 1970-1974. And, if it were possible to survey that individual in a later cross-section he would still belong to the same cohort. The results are statistically representative cohorts, which can be tracked over time despite the fact that membership in each cohort is different from year-to-year. In lieu of 3.10, 3.11 is estimated such that

$$\bar{y}_{ct} = \gamma \bar{y}_{c,t-1} + \bar{x}'_{ct}\beta + \bar{u}_{ct} \tag{3.11}$$

with error

$$\bar{u}_{ct} = \alpha_c + \gamma(\bar{y}^*_{ct} - \bar{y}_{ct}) - e_{ct} \tag{3.12}$$

where $\bar{y}_{c,t-1}$ is the predicted lagged dependent variable equal to the sample average lagged dependent variable, $\bar{x}'_{ct}$ and $\bar{u}_{ct}$ are the average cohort independent regressors and $\bar{y}^*_{ct}$ is the cohort population mean. As Moffitt (1993) proved, averaging the data by cohort is analogous to instrumenting the individual data with a vector of cohort dummies interacted with time and ensures that the regressors are exogenous assuming that the cohort dummies pass the usual criteria for instruments. However, Verbeek and Vella (2005) find that these criteria drastically limit the number of viable instruments. To relax the instrument criteria they exploit several asymptotic assumptions of the data. Particularly, they assume that as $N \to \infty$, C, the number of cohorts is assumed to be fixed such that the number of members in each cohort, n_c tends to infinity. In this case, time invariant cohort effects can be isolated in much the same manner as with traditional fixed effects by including the cohort dummies in the main equation. This gives way to the augmented IV estimator (Verbeek and Vella 2005)

$$\bar{y}_{ct} = \gamma \bar{y}_{c,t-1} + \bar{x}'_{ct}\beta + \alpha_c + \bar{u}_{ct}. \tag{3.13}$$

Under the illustrated asymptotic assumptions, 3.13 is assumed to be a one-way error component model where $\bar{u}_{ct}$ is expected to contain time varying cohort effects, but since $n_c \to \infty$ it is assumed that the time varying unobservable component is zero in expectation such that the regressors and the error is uncorrelated to the instruments (cohort dummies).

Collado (1997) provide an alternative model. Essentially, she proposes a two-way error component model where the time varying cohort effects are not assumed to be zero in expectation. To estimate such a model, Collado (1997) uses the Arellano-Bond estimator. While there is a gain in generality it comes at the expense of efficiency since second lags are used as instruments for the differenced equation.

In equation 3.13, $\bar{y}_{ct}$ and $\bar{y}_{c,t-1}$ are the current and lagged dependent variables. The log of BMI and cohort percentages of overweight and obese categories were used as the dependent variable. $\bar{x}'_{ct}$ is composed of a vector of variables effecting the production of health including: age, age^2, indicator variables for marital status and food security and the log of income. The use of these variables are consistent with previous studies (Cawley, Moran, and Simon 2010; Schmeiser 2009; Griffiths et al. 2013; Laraia, Siega-Riz, and Evenson 2004; CDC et al. 2003; Dinour, Bergen, and Yeh 2007). Year indicators where also included to control for time trends. α_c is a vector of indicator variables for each time stable cohort. The cohort dummies interact race, gender and birth generation.

3.3.3 Data

This paper uses RCS data collected by the National Health and Nutrition Examination Survey (NHANES). NHANES is designed to evaluate health and nutritional status of adults and children in the United States. The NHANES program has been an ongoing survey since the 1960s. However, in 1999 the survey became continuous, collecting information on people's health on a yearly basis. Each year, 15 counties are chosen to be surveyed from which a group of approximately 5000 nationally representative individuals are sampled. The NHANES data is an extensive repository for information on demographic, economic, health outcomes and particularly dietary information. To gather this information, the respondents are subjected to a medical exam, laboratory work, an interview and are required to self-report food consumption.

The Continuous NHANES survey currently has six rounds of completed data starting in the 1999-2000 survey year and ending in 2009-2010. Cohorts were formed using race, gender and birth year ranges. Deaton (1985) was the first to use birth year ranges for cohort formation. Like, Russell and Fraas (2005) birth year, race and gender are used to form cohorts. Cohorts were formed using such variables because of evidence indicating weight differentials by age (Flegal et al. 1998; Cook and Daponte 2008; Baum II and Ruhm 2009), gender (Flegal et al. 2010) and race (Hedley et al. 2004; Flegal et al. 2010). There were three categories for race: non-Hispanic white, non-Hispanic black and Mexican American/other Hispanic/multi-racial. Gender includes both males and females while there are 16 birth year ranges composed of 5 year spans starting in 1915 in the earliest round through to 1992 in the last wave of data. The cohort formation variables produce 96 potential cohorts. However, as shown by Verbeek and Vella (2005), a necessary condition for identification of the model is, at a minimum, 3 rounds of data are needed to insure time variation and avoid multicolinearity. Eliminating cohorts with less than 3 rounds of data resulted in elimination of the oldest and youngest groups leaving 84 unique cohorts to be estimated.

Table 3.1 shows the member participation for each cohort by year where each number represents the number of individuals in each cohort. From the tables it shows that there are more non-hispanic white men and women relative to non-Hispanic Blacks and Mexican American/ Other Hispanic/Multi-race. The missing values in each table are a reflection of the oldest cohorts aging out of the survey and being replaced by younger ones.

Table 3.1: Cohort Size

	Women						Men					
Non-Hispanic White												
	2000	2002	2004	2006	2008	2010	2000	2002	2004	2006	2008	2010
Born 1920 – 1924	69	89	170	124			71	87	139	109		
Born 1925 – 1929	83	92	103	73	180		84	117	87	100	169	
Born 1930 – 1934	76	83	96	77	106	226	87	78	108	84	123	227
Born 1935 – 1939	85	94	84	78	100	125	83	87	94	73	101	111
Born 1940 – 1944	50	87	99	82	108	88	72	84	87	96	97	99
Born 1945 – 1949	75	115	85	79	112	97	76	133	88	66	116	113
Born 1950 – 1954	71	102	96	76	88	106	63	101	84	97	84	105
Born 1955 – 1959	80	100	88	87	103	104	77	112	93	103	126	114
Born 1960 – 1964	85	115	92	102	121	127	86	95	105	107	111	114
Born 1965 – 1969	109	130	96	88	107	130	76	89	86	91	105	116
Born 1970 – 1974	110	138	132	101	107	144	70	82	91	84	109	124
Born 1975 – 1979	97	128	115	125	98	109	70	83	80	78	97	113
Born 1980 – 1984	57	82	113	139	70	124	62	95	90	87	92	98
Born 1985 – 1989			127	114	81	117			137	113	87	99
Mean	81	104	107	96	106	125	75	96	98	92	109	119
Non-Hispanic Black												
	2000	2002	2004	2006	2008	2010	2000	2002	2004	2006	2008	2010
Born 1920 – 1924	26	18	19	19			17	18	13	10		
Born 1925 – 1929	29	29	17	17	39		29	22	16	13	29	
Born 1930 – 1934	30	33	21	21	26	46	38	47	15	30	11	35
Born 1935 – 1939	50	52	24	40	42	33	40	49	29	37	46	29
Born 1940 – 1944	25	28	40	38	50	38	21	28	35	52	52	40

Continued on next page

Table 3.1 – continued from previous page

Born 1945 – 1949	28	32	32	40	71	47	31	34	37	44	81	54
Born 1950 – 1954	44	49	40	43	61	42	36	55	35	41	34	55
Born 1955 – 1959	49	57	45	58	45	58	50	54	37	51	56	71
Born 1960 – 1964	51	54	62	53	58	52	33	43	48	43	46	49
Born 1965 – 1969	59	44	44	52	48	45	42	35	39	50	44	41
Born 1970 – 1974	40	47	41	47	52	46	25	32	42	48	57	38
Born 1975 – 1979	37	62	42	62	55	49	35	46	46	50	46	29
Born 1980 – 1984	60	84	68	61	50	47	66	92	54	54	45	52
Born 1985 – 1989			123	134	60	55			124	133	65	50
Mean	41	45	44	49	51	47	36	43	41	47	47	45

Hispanic/Mexican/Other

	2000	2002	2004	2006	2008	2010	2000	2002	2004	2006	2008	2010
Born 1920 – 1924	32	18	24	16			42	18	17	11		
Born 1925 – 1929	53	43	30	10	46		59	39	29	9	33	
Born 1930 – 1934	78	49	39	16	40	58	63	36	39	17	35	44
Born 1935 – 1939	89	70	51	30	58	54	94	62	61	36	37	43
Born 1940 – 1944	49	27	64	46	63	63	34	29	56	47	65	71
Born 1945 – 1949	72	43	50	46	112	113	50	49	40	38	79	82
Born 1950 – 1954	79	63	35	34	68	85	59	72	33	32	74	75
Born 1955 – 1959	82	85	55	40	99	80	80	73	47	49	107	89
Born 1960 – 1964	73	77	43	69	82	104	76	63	53	50	75	100
Born 1965 – 1969	78	88	66	61	76	126	60	71	59	82	90	85
Born 1970 – 1974	97	104	58	93	105	101	74	84	59	70	96	106
Born 1975 – 1979	133	119	80	110	100	97	78	92	78	67	75	90
Born 1980 – 1984	176	110	60	125	83	112	156	129	65	91	94	88
Born 1985 – 1989			155	162	122	102			121	132	107	106
Mean	84	69	58	61	81	91	71	63	54	52	74	82

3.3.4 Descriptive Statistics

The estimations presented in this paper use several different measures of health stock. Table 3.2. illustrates that the continuous variable ln(BMI) was used as well as a proportional variable indicating the prevalence of obesity. The average cohort BMI is 28, which is in the overweight category. Furthermore, on average 76% of individuals were at least overweight by cohort, 41% were at least classified as obese 1, 21% were obese 2 and 13% were obese 3. This weight stratification is consistent with national averages for overweight and obesity rates (NIDDK 2012).

By design there were equal numbers of male and female cohorts and equal number of cohorts that were white, black and Mexican American/ Other Hispanic/ Multi-racial. The average cohort age was 49.8 years. Upon running a fractional polynomial analysis it was found that the quadratic transformation of age fit much better relative to the untransformed variable. For this reason age^2 was included. Over 50% of individuals were married in each cohort. Food security is an indicator variable where 1 represents food secure.

Table 3.2 shows that on average each cohort was 84% food secure. Income is a continuous variable and is in real 1990 USD. When converted to levels, the average cohort income was just over $37,000.

The time variables indicate that the number of observations is roughly equal over all the years with fewer observations in the first and last round of data. This is to be expected as the youngest and oldest cohorts were eliminated due to having less then 3 rounds of data.

Table 3.2: Descriptive Statistics

Variable	μ	σ
Health Stock Variables		
$log(BMI)$	3.334	0.066
$log(\%\ overweight)$	-0.271	0.201
$log(\%\ obese1)$	-0.881	0.314
$log(\%\ obese2)$	-1.549	0.451
$log(\%\ obese3)$	-2.072	0.547
Demographic Variables		
sex	0.5	0.501
$white$	0.333	0.472
$black$	0.333	0.472
$hisp$	0.333	0.472
age	49.83	19.067
age^2	2845.789	1935.707
$married$	0.561	0.211
Time Variables		
2000	0.165	0.371
2002	0.165	0.371
2004	0.177	0.382
2006	0.177	0.382
2008	0.165	0.371
2010	0.152	0.359
Socioeconomic Variables		
$log(income)$	10.53	0.209
$food\ security$	.839	.091
Cohort Indicator Variables		
$Born$ $1920 - 1924$	0.051	0.219
$Born$ $1925 - 1929$	0.063	0.244
$Born$ $1930 - 1934$	0.076	0.265
$Born$ $1935 - 1939$	0.076	0.265
$Born$ $1940 - 1944$	0.076	0.265
$Born$ $1945 - 1949$	0.076	0.265
$Born$ $1950 - 1954$	0.076	0.265

Continued on next page

Table 3.2 – continued from previous page		
Born 1955 – 1959	0.076	0.265
Born 1960 – 1964	0.076	0.265
Born 1965 – 1969	0.076	0.265
Born 1970 – 1974	0.076	0.265
Born 1975 – 1979	0.076	0.265
Born 1980 – 1984	0.076	0.265
Born 1985 – 1989	0.051	0.219
Health Input Variables		
$log(meat)$	3.648	0.277
$log(added\ sugar)$	4.117	0.643
$log(fat)$	4.11	0.258

Finally, to make use of NHANES' extensive nutritional data, consumption of meat, added sugar and fat was first calculated per individual using data from a two-day dietary interview. Individuals recorded and self-reported quantities of food consumed on two days at least three days apart and the nutritional values of their diets were imputed by surveyors. The totals were averaged to get mean individual daily consumption levels. These values were then averaged again over cohort and then logged. Table 3.2. shows the logged values of the consumption of each food item in grams. The average cohort daily consumption of meat, added sugar and fat were 38, 61, and 60.9 grams respectively. This indicates that on average individuals eat over 1.5 times as much added sugar and fat as they do meat.

3.4 Results

3.4.1 Demand for Health Stock

Tables 3.3 and 3.4 show the results from the demand for health stock estimation for women and men respectively. While the cohort indicator variables were only significant for women, the general trend across both genders is that the oldest cohorts tend to be the heaviest and this disparity between oldest to youngest progressively widens with increasing weight class. This is not surprising since the highest prevalence of obesity is among 60-69 year olds (Flegal et al. 2002).

For women, marital status had a decreasing effect on average BMI and incidence of obesity for all categories. Sobal, Rauschenbach, and Jr. (1992) also find a negative effect but, it was not statistically significant in their study. Since women typically bare more household responsibility for food preparation, marriage may impose increased nutritional accountability for other members of the household including children, which in turn may make women more prone to eat healthy. Estimates show that being food secure spurs weight loss for women, but was only significant for overweight and obese 2 categories. This negative effect is consistent with prior literature, which found that in developed countries women in particular exhibit an inverse relationship with weight and food insecurity (Dinour, Bergen, and Yeh 2007). Elevated weight may actually be demonstrative of rational behavior by food insecure individuals as it may induce over-consumption during periods of relative food abundance to compensate for future periods of food scarcity, which may lead to weight gain (Dinour, Bergen, and Yeh 2007; Townsend et al. 2001). Furthermore, food insecurity may induce consumption of highly processed, non-perishable foods and fast foods which are conducive to rapid weight gain. Income was found to be a negative determinant for the continuous BMI variable, overweight and obese 1, though was only significant for overweight. For obese 2 and obese 3, the coefficients were both positive, but only significant for obese 2. One possible explanation for the positive coefficient in the obese 2(3) estimation is that for individuals with the highest BMIs, a positive income effect may facilitate excess calorie consumption. This would be particularly true if the most obese individuals have formed eating habits which promote overconsumption. These findings differ from Cawley, Moran, and Simon (2010) and Schmeiser (2009)'s analysis which find no significant effect from income, but is consistent with many other cross-sectional studies (Sobal and Stunkard 1989; McLaren 2007). The lagged dependent variable was also significant and positive for women in all estimations except in the case of obese 3. This highly suggests that female weight outcomes is a dynamic process for all but the the heaviest individuals.

For men, marriage is a positive effect for overweight incidence, but is shown to be a decreasing factor for higher obesity categories. The positive coefficient in the overweight estimation maybe a reflection of behavioral changes induced by marriage. As a result, moderate male weight gain may be caused by more consistent consumption of meals and a less active lifestyle brought upon by the advent of marriage (Sobal, Rauschenbach, and Jr. 1992). For higher obese categories, the negative coefficients may reflect heightened spousal concern for health when body weight is very high and signify increased spousal support for weight loss efforts. Food security is not a significant determinant for men's weight outcomes of which the coefficients were generally positive. Unlike women, income was not significant for men and was positive. Surprisingly, the lagged dependent variable for men was insignificant for all but the continuous BMI variable. Moreover, the sign of each coefficient was inconsistent across estimations as the obese 1 and obese 3 lagged dependent variables were negative. This suggests that for men, past weight outcomes is not a factor for current weight. Men maybe less hindered by a weight *"legacy"* as they may be more adaptable to weight loss (or weight gain) then women.

Table 3.3: Demand for Health Stock: Women

VARIABLES	(1) $log(BMI_t)$	(2) $log(\%\ overweight_t)$	(3) $log(\%\ obese1_t)$	(4) $log(\%\ obese2_t)$	(5) $log(\%\ obese3_t)$
$log(BMI_{t-1})$	0.37***				
	(0.067)				
$log(\%\ overweight_{t-1})$		0.18***			
		(0.066)			
$log(\%\ obese1_{t-1})$			0.17**		
			(0.071)		
$log(\%\ obese2_{t-1})$				0.15**	
				(0.069)	
$log(\%\ obese3_{t-1})$					0.02
					(0.079)
age	-0.00	-0.06**	-0.06	-0.06	-0.08
	(0.010)	(0.029)	(0.055)	(0.077)	(0.118)
age^2	-0.00***	-0.00***	-0.00***	-0.00***	-0.00***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$married$	-0.15***	-0.28***	-0.74***	-1.26***	-1.72***
	(0.027)	(0.068)	(0.133)	(0.183)	(0.274)
$food\ secure$	-0.06	-0.36***	-0.32	-0.61*	-0.09
	(0.045)	(0.136)	(0.247)	(0.345)	(0.526)
$log(income)$	-0.05	-0.23**	-0.20	0.57**	0.42
	(0.032)	(0.093)	(0.174)	(0.244)	(0.370)
2004	0.15**	0.64***	0.78**	0.80	0.91
	(0.064)	(0.186)	(0.350)	(0.491)	(0.748)
2006	0.14**	0.63***	0.80**	0.74	0.74
	(0.064)	(0.185)	(0.348)	(0.488)	(0.743)
2008	0.17**	0.84***	1.03**	0.95	1.01
	(0.081)	(0.234)	(0.441)	(0.618)	(0.940)
2010	0.22**	1.04***	1.26**	1.18	1.29
	(0.101)	(0.291)	(0.548)	(0.769)	(1.168)

Continued on next page

Table 3.3 – continued from previous page

Born 1920 – 1924	1.15*	6.54***	7.97**	9.47**	11.21
	(0.603)	(1.736)	(3.271)	(4.589)	(6.982)
Born 1925 – 1929	1.05*	5.95***	7.18**	8.49**	10.11
	(0.555)	(1.597)	(3.009)	(4.220)	(6.422)
Born 1930 – 1934	0.97*	5.42***	6.51**	7.63*	8.93
	(0.514)	(1.477)	(2.784)	(3.903)	(5.940)
Born 1935 – 1939	0.90*	4.95***	5.99**	7.00*	8.11
	(0.468)	(1.345)	(2.534)	(3.552)	(5.406)
Born 1940 – 1944	0.79*	4.40***	5.32**	6.10*	7.19
	(0.421)	(1.212)	(2.284)	(3.201)	(4.873)
Born 1945 – 1949	0.71*	3.97***	4.77**	5.37*	6.25
	(0.378)	(1.088)	(2.050)	(2.873)	(4.373)
Born 1950 – 1954	0.62*	3.43***	4.16**	4.65*	5.57
	(0.331)	(0.951)	(1.793)	(2.513)	(3.826)
Born 1955 – 1959	0.51*	2.95***	3.53**	3.95*	4.63
	(0.285)	(0.819)	(1.542)	(2.163)	(3.292)
Born 1960 – 1964	0.43*	2.46***	2.92**	3.22*	3.80
	(0.238)	(0.685)	(1.291)	(1.809)	(2.754)
Born 1965 – 1969	0.34*	1.98***	2.38**	2.62*	3.16
	(0.192)	(0.553)	(1.042)	(1.460)	(2.222)
Born 1970 – 1974	0.29**	1.60***	1.97**	2.13*	2.51
	(0.146)	(0.419)	(0.789)	(1.106)	(1.682)
Born 1975 – 1979	0.20**	1.06***	1.31**	1.38*	1.70
	(0.098)	(0.282)	(0.530)	(0.742)	(1.129)
Born 1980 – 1984	0.10*	0.54***	0.64**	0.72*	0.80
	(0.053)	(0.153)	(0.287)	(0.403)	(0.613)
Constant	2.62***	2.99***	2.18	-5.80**	-4.88
	(0.415)	(0.889)	(1.669)	(2.332)	(3.548)
Observations	195	195	195	195	195
R^2	0.747	0.710	0.549	0.469	0.417

Standard errors in parentheses

$***p < 0.01, **p < 0.05, *p < 0.1$

Table 3.4: Demand for Health Stock: Men

VARIABLES	(1) $log(BMI_t)$	(2) $log(\%\ overweight_t)$	(3) $log(\%\ obese1_t)$	(4) $log(\%\ obese2_t)$	(5) $log(\%\ obese3_t)$
$log(BMI_{t-1})$	0.19***				
	(0.072)				
$log(\%\ overweight_{t-1})$		0.12			
		(0.074)			
$log(\%\ obese1_{t-1})$			-0.04		
			(0.071)		
$log(\%\ obese2_{t-1})$				0.01	
				(0.077)	
$log(\%\ obese3_{t-1})$					-0.04
					(0.077)
age	0.01	0.03	0.04	0.04	0.00
	(0.008)	(0.035)	(0.067)	(0.111)	(0.132)
age^2	-0.00***	-0.00***	-0.00***	-0.00***	-0.00*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$married$	0.02	0.32***	-0.27	-0.70**	-0.83**
	(0.024)	(0.101)	(0.188)	(0.319)	(0.371)
$food\ secure$	0.03	0.04	0.10	0.19	-0.55
	(0.041)	(0.169)	(0.325)	(0.543)	(0.636)
$log(income)$	0.05	0.09	0.37	0.32	0.60
	(0.030)	(0.126)	(0.242)	(0.411)	(0.482)
2004	0.04	0.06	0.31	0.21	0.21
	(0.053)	(0.224)	(0.428)	(0.709)	(0.843)
2006	0.04	0.05	0.34	0.23	0.20
	(0.053)	(0.221)	(0.423)	(0.699)	(0.835)
2008	0.05	0.04	0.42	0.37	0.24
	(0.066)	(0.278)	(0.532)	(0.881)	(1.054)
2010	0.06	0.06	0.51	0.44	0.21
	(0.082)	(0.347)	(0.663)	(1.098)	(1.312)

Continued on next page

Table 3.4 – continued from previous page

Born 1920 – 1924	0.25	0.62	3.39	4.36	5.48
	(0.493)	(2.066)	(3.953)	(6.560)	(7.804)
Born 1925 – 1929	0.25	0.62	3.14	3.93	4.94
	(0.455)	(1.909)	(3.651)	(6.055)	(7.220)
Born 1930 – 1934	0.23	0.53	2.85	3.46	4.36
	(0.422)	(1.769)	(3.383)	(5.615)	(6.681)
Born 1935 – 1939	0.20	0.37	2.46	2.95	3.54
	(0.383)	(1.609)	(3.077)	(5.107)	(6.078)
Born 1940 – 1944	0.18	0.32	2.18	2.68	3.26
	(0.345)	(1.450)	(2.772)	(4.600)	(5.477)
Born 1945 – 1949	0.14	0.22	1.85	2.24	2.80
	(0.309)	(1.298)	(2.482)	(4.119)	(4.905)
Born 1950 – 1954	0.11	0.16	1.46	1.75	2.28
	(0.270)	(1.135)	(2.171)	(3.599)	(4.287)
Born 1955 – 1959	0.09	0.10	1.13	1.43	1.82
	(0.233)	(0.978)	(1.871)	(3.103)	(3.697)
Born 1960 – 1964	0.08	0.05	0.94	1.08	1.45
	(0.195)	(0.819)	(1.567)	(2.600)	(3.098)
Born 1965 – 1969	0.07	0.10	0.93	1.27	1.50
	(0.157)	(0.657)	(1.257)	(2.087)	(2.489)
Born 1970 – 1974	0.07	0.11	0.73	1.01	1.37
	(0.118)	(0.495)	(0.949)	(1.576)	(1.879)
Born 1975 – 1979	0.04	0.10	0.58	0.79	1.05
	(0.080)	(0.336)	(0.644)	(1.072)	(1.278)
Born 1980 – 1984	0.03	0.06	0.34	0.42	0.47
	(0.045)	(0.187)	(0.358)	(0.595)	(0.708)
Constant	1.93***	-2.23*	-6.37***	-6.38	-8.55*
	(0.339)	(1.267)	(2.427)	(4.079)	(4.810)
Observations	195	195	195	193	191
R^2	0.701	0.624	0.399	0.206	0.247

Standard errors in parentheses

$***p < 0.01, **p < 0.05, *p < 0.1$

3.4.2 Demand for Health Inputs

Determinants for the consumption of meat, added sugar and fat were estimated by gender using the appropriate predicted values from the demand of health stock estimations. The estimates show that food security induced more meat consumption from women. The only other significant factor for women was lagged meat consumption. This suggests that women's consumption of meat is highly driven by habitual behaviors and the magnitude of each coefficient was relatively consistent over all weight categories.

For men, when marital status was significant, marriage decreased the demand for meat. This behavior could be caused by two effects. First, a perceived health assumption linking consumption of saturated fat, often found in meat, to decreasing health may induce familial accountability. Specifically, men maybe more motivated to make perceived health conscious decisions like reducing meat consumption to prevent premature death and familial abandonment as a result. Second, this behavior could also be a reflection of spousal peer effects. Particularly, if women are predominantly in charge of food preparation lower consumption of meat by women may also influence male consumption. Being food secure seemed to induce more meat consumption, but income is found to be a negative determinant of meat consumption. These findings are somewhat counterintuitive as one would expect a consistent sign between both variables. A possible explanation is that the positive coefficient on food security may reflect capital intensity relevant to food preparation in food secure households. Specifically, for meals-at-home, fresh meat consumption generally requires adequate refrigeration and preparation tools, which may not necessarily be available in food insecure households. Additionally, as an effort to hedge for unknown future food scarcity, food insecure households may be more inclined to purchase non-perishable food items, which are typically more carbohydrate intense, require little preparation and have low risk of spoilage. Contrastingly, the negative coefficient on income maybe reflective of health investments by richer households in which they substitute away from animal protein toward perceived health foods. Somewhat eluded in the descriptive statistics, the birth cohort variables show a decrease in meat consumption from the oldest to the youngest cohorts. Again, this could reflect generational changes in in eating habits where mitigating meat consumption is believed to be health promoting. This trend is consistent with an overall decline in national meat consumption (Johnson 2012).

Table 3.6 shows that marriage increased added sugar consumption with increasing magnitude by weight class for women. In particular women who were the most obese (obese 3) show that marriage has a tremendous effect on sugar consumption as the coefficient was over 8.5 times larger then the second largest coefficient from the remaining weight categories. Food security and income also induced more added sugar consumption for women. A conceivable reason for this is again a reflection of changing food habits over time. Increase in consumption of fat-free or low-fat foods as they are perceived to be health boosting may have inadvertently increased added sugar consumption since these food often contain more sugar to increase palatability. Weight status was positive, significant and increasing in magnitude by weight classification. This may suggest that some physiological aspect of obesity maybe driving the demand for added sugar consumption. Added sugar consumption may also be determined by habit formations as lagged sugar consumption was also significant for all female estimations and is relatively consistent in magnitude across all BMI categories.

Weight status also seemed to be a positive factor in sugar demand and appears to increase in magnitude with higher weight status. It suggests that some physiological factor may drive the demand for sugar in women.

Income was a positive factor in demand for added sugar for men and increased in magnitude by weight status. Similar to the women, lagged sugar consumption was consistent across all BMI categories and was positive and significant. For both genders it appears that habitual sugar consumption plays a significant role in current added sugar consumption. As for cohort effects there was an obvious increasing tread of added sugar consumption among women. The results for men where however, not easily interpretable across categories. For obese 1 and obese 3, there is an increase in sugar consumption from oldest to youngest while obese 2 showed an decreasing trend. Obesity is more prevalent in women across all obese categories (Flegal et al. 2010) and the cohort results suggest that increased added sugar in women's diets maybe attributing to higher preponderance of obesity in women relative to men.

Men and women exhibit comparable behavior for fat consumption. Food security and lagged fat consumption were both positive, significant determinants for food consumption. For both genders, fat consumption was also influenced by habit formation as the lagged consumption for both genders is positive and significant across all weight categories.

Table 3.5: Demand for Meat: Women

(1) VARIABLES	(2) $log(BMI_t)$[32]	(3) $log(\%\ overweight_t)$	(4) $log(\%\ obese1_t)$	(5) $log(\%\ obese2_t)$	$log(\%\ obese3_t)$
age	-0.03	0.03	0.01	0.00	0.06
	(0.032)	(0.041)	(0.035)	(0.035)	(0.126)
age^2	0.00	0.00	0.00	0.00	0.00
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
$married$	0.04	0.03	0.35	0.38	1.48
	(0.151)	(0.150)	(0.227)	(0.277)	(2.457)
$food\ secure$	0.14	0.47**	0.29*	0.30*	0.12
	(0.150)	(0.238)	(0.168)	(0.179)	(0.186)
$log(income)$	-0.03	0.10	0.03	-0.39**	-0.56
	(0.111)	(0.144)	(0.116)	(0.150)	(0.613)
2004	0.14	-0.24	-0.22	-0.04	-0.55
	(0.228)	(0.340)	(0.291)	(0.270)	(1.329)
2006	0.07	-0.29	-0.29	-0.07	-0.44
	(0.230)	(0.337)	(0.294)	(0.260)	(1.091)
2008	0.11	-0.40	-0.37	-0.09	-0.62
	(0.288)	(0.439)	(0.374)	(0.331)	(1.481)
2010	0.14	-0.47	-0.44	-0.10	-0.79
	(0.360)	(0.542)	(0.464)	(0.411)	(1.889)
$Born\ 1920-1924$	0.92	-3.42	-3.21	-1.98	-8.47
	(2.093)	(3.387)	(2.864)	(2.789)	(16.202)
$Born\ 1925-1929$	0.79	-3.15	-2.91	-1.78	-7.66
	(1.922)	(3.091)	(2.605)	(2.532)	(14.612)
$Born\ 1930-1934$	0.76	-2.79	-2.54	-1.49	-6.62
	(1.780)	(2.827)	(2.382)	(2.306)	(12.922)
$Born\ 1935-1939$	0.65	-2.57	-2.38	-1.39	-6.02
	(1.627)	(2.582)	(2.182)	(2.107)	(11.735)
$Born\ 1940-1944$	0.67	-2.21	-2.03	-1.10	-5.25
	(1.460)	(2.304)	(1.949)	(1.867)	(10.401)
$Born\ 1945-1949$	0.57	-2.04	-1.86	-0.97	-4.54
	(1.310)	(2.077)	(1.750)	(1.662)	(9.050)
$Born\ 1950-1954$	0.52	-1.72	-1.59	-0.81	-4.05
	(1.147)	(1.804)	(1.529)	(1.448)	(8.063)
$Born\ 1955-1959$	0.46	-1.50	-1.35	-0.69	-3.34
	(0.982)	(1.553)	(1.307)	(1.241)	(6.710)
$Born\ 1960-1964$	0.38	-1.26	-1.11	-0.53	-2.73
	(0.820)	(1.295)	(1.085)	(1.023)	(5.506)
$Born\ 1965-1969$	0.36	-0.97	-0.88	-0.40	-2.26
	(0.661)	(1.045)	(0.883)	(0.831)	(4.583)
$Born\ 1970-1974$	0.22	-0.84	-0.79	-0.36	-1.82
	(0.513)	(0.835)	(0.705)	(0.650)	(3.638)
$Born\ 1975-1979$	0.13	-0.55	-0.52	-0.22	-1.23
	(0.348)	(0.557)	(0.472)	(0.429)	(2.462)
$Born\ 1980-1984$	0.08	-0.26	-0.22	-0.10	-0.55
	(0.186)	(0.287)	(0.238)	(0.226)	(1.160)
$log(meat_{t-1})$	0.29***	0.31***	0.29***	0.29***	0.34***
	(0.071)	(0.069)	(0.069)	(0.070)	(0.069)
$log(\%\ overweight_t)$		0.94**			
		(0.412)			
$log(\%\ obese1_t)$			0.74***		
			(0.258)		
$log(\%\ obese2_t)$				0.47**	
				(0.195)	
$log(\%\ obese3_t)$					1.01

Continued on next page

[32]The BMI variables designated on the top of each column does not represent the dependent variable but rather indicates the predicted BMI variable used in the food intake regression

Table 3.5 – continued from previous page

					(1.418)
Constant	-1.51	1.24	2.69**	6.95***	8.97
Observations	195	195	195	195	195
R^2	0.650	0.650	0.656	0.651	0.640

Standard errors in parentheses
$***p < 0.01, **p < 0.05, *p < 0.1$

Table 3.6: Demand for Added Sugar: Women

VARIABLES	(6) $log(BMI_t)$	(7) $log(\%\ overweight_t)$	(8) $log(\%\ obese1_t)$	(9) $log(\%\ obese2_t)$	(10) $log(\%\ obese3_t)$
age	-0.05	0.03	-0.01	-0.02	0.23*
	(0.035)	(0.044)	(0.038)	(0.038)	(0.132)
age^2	0.00***	0.00***	0.00***	0.00***	0.00**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
married	0.38**	0.38**	0.64***	0.67**	5.78**
	(0.159)	(0.160)	(0.240)	(0.291)	(2.568)
food secure	0.27	0.72***	0.42**	0.41**	0.42**
	(0.168)	(0.253)	(0.186)	(0.195)	(0.205)
$log(income)$	0.25**	0.43***	0.29**	-0.15	-1.31**
	(0.118)	(0.154)	(0.124)	(0.149)	(0.634)
2004	-1.41***	-1.93***	-1.73***	-1.56***	-4.18***
	(0.247)	(0.362)	(0.312)	(0.292)	(1.391)
2006	-1.14***	-1.64***	-1.48***	-1.23***	-3.35***
	(0.253)	(0.362)	(0.321)	(0.283)	(1.148)
2008	-1.22***	-1.92***	-1.67***	-1.36***	-4.27***
	(0.311)	(0.467)	(0.403)	(0.356)	(1.554)
2010	-1.27***	-2.10***	-1.80***	-1.43***	-5.16**
	(0.386)	(0.574)	(0.497)	(0.440)	(1.981)
Born 1920 − 1924	-1.17	-7.05*	-5.03	-3.78	-36.37**
	(2.252)	(3.591)	(3.050)	(2.988)	(16.942)
Born 1925 − 1929	-0.97	-6.31*	-4.42	-3.28	-32.68**
	(2.069)	(3.278)	(2.776)	(2.715)	(15.280)
Born 1930 − 1934	-0.90	-5.71*	-3.97	-2.89	-28.80**
	(1.916)	(3.000)	(2.540)	(2.476)	(13.515)
Born 1935 − 1939	-0.80	-5.16*	-3.60	-2.59	-26.08**
	(1.752)	(2.741)	(2.328)	(2.263)	(12.274)
Born 1940 − 1944	-0.69	-4.58*	-3.19	-2.24	-23.11**
	(1.572)	(2.447)	(2.080)	(2.007)	(10.880)
Born 1945 − 1949	-0.56	-4.09*	-2.81	-1.91	-20.01**
	(1.411)	(2.206)	(1.868)	(1.788)	(9.469)
Born 1950 − 1954	-0.48	-3.51*	-2.43	-1.63	-17.82**
	(1.236)	(1.917)	(1.633)	(1.560)	(8.436)
Born 1955 − 1959	-0.33	-2.97*	-2.01	-1.33	-14.75**
	(1.059)	(1.651)	(1.398)	(1.337)	(7.022)
Born 1960 − 1964	-0.24	-2.45*	-1.62	-1.04	-12.06**
	(0.885)	(1.378)	(1.161)	(1.104)	(5.763)
Born 1965 − 1969	-0.19	-1.97*	-1.33	-0.85	-10.06**
	(0.715)	(1.113)	(0.946)	(0.899)	(4.797)
Born 1970 − 1974	-0.16	-1.59*	-1.09	-0.66	-7.95**
	(0.555)	(0.888)	(0.754)	(0.703)	(3.808)
Born 1975 − 1979	-0.11	-1.04*	-0.71	-0.40	-5.36**
	(0.376)	(0.593)	(0.506)	(0.465)	(2.577)
Born 1980 − 1984	-0.03	-0.49	-0.30	-0.18	-2.48**
	(0.201)	(0.306)	(0.255)	(0.245)	(1.215)
$log(BMI_t)$	1.82***				
	(0.617)				

Continued on next page

Table 3.6 – continued from previous page

$log(\%\ overweight_t)$		1.26***			
		(0.426)			
$log(\%\ obese1_t)$			0.78***		
			(0.264)		
$log(\%\ obese2_t)$				0.50**	
				(0.199)	
$log(\%\ obese3_t)$					3.35**
					(1.480)
$log(addedsugar_{t-1})$	0.26***	0.26***	0.25***	0.27***	0.24***
	(0.081)	(0.081)	(0.081)	(0.082)	(0.082)
$Constant$	-4.35	-0.85	1.43	5.88***	19.61***
	(2.827)	(1.805)	(1.303)	(1.543)	(7.303)
Observations	195	195	195	195	195
R^2	0.954	0.954	0.954	0.954	0.953

Standard errors in parentheses
$***p < 0.01, **p < 0.05, *p < 0.1$

Table 3.7: Demand for Fat: Women

	(11)	(12)	(13)	(14)	(15)
VARIABLES	$log(BMI_t)$	$log(\%\ overweight_t)$	$log(\%\ obese1_t)$	$log(\%\ obese2_t)$	$log(\%\ obese3_t)$
age	-0.02	-0.02	-0.00	-0.00	-0.01
	(0.027)	(0.034)	(0.029)	(0.029)	(0.105)
age^2	-0.00	-0.00	0.00	0.00	-0.00
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
$married$	-0.18	-0.21*	-0.01	0.14	-0.15
	(0.125)	(0.126)	(0.188)	(0.226)	(2.059)
$food\ secure$	0.48***	0.43**	0.55***	0.61***	0.48***
	(0.136)	(0.204)	(0.150)	(0.158)	(0.169)
$log(income)$	-0.08	-0.11	-0.03	-0.20*	-0.09
	(0.093)	(0.121)	(0.098)	(0.116)	(0.509)
2004	0.13	0.20	-0.04	-0.08	0.12
	(0.192)	(0.283)	(0.244)	(0.224)	(1.112)
2006	0.09	0.16	-0.08	-0.10	0.08
	(0.193)	(0.281)	(0.246)	(0.216)	(0.913)
2008	0.12	0.21	-0.10	-0.13	0.10
	(0.242)	(0.365)	(0.313)	(0.275)	(1.240)
2010	0.19	0.30	-0.08	-0.12	0.17
	(0.302)	(0.451)	(0.388)	(0.341)	(1.582)
$Born\ 1920-1924$	1.00	1.68	-0.73	-1.44	0.79
	(1.757)	(2.813)	(2.388)	(2.300)	(13.575)
$Born\ 1925-1929$	0.93	1.54	-0.64	-1.27	0.74
	(1.614)	(2.568)	(2.173)	(2.089)	(12.243)
$Born\ 1930-1934$	0.91	1.47	-0.51	-1.06	0.74
	(1.495)	(2.350)	(1.988)	(1.904)	(10.826)
$Born\ 1935-1939$	0.82	1.33	-0.48	-0.99	0.67
	(1.366)	(2.146)	(1.821)	(1.740)	(9.832)
$Born\ 1940-1944$	0.79	1.25	-0.36	-0.79	0.66
	(1.226)	(1.916)	(1.628)	(1.543)	(8.715)
$Born\ 1945-1949$	0.71	1.12	-0.33	-0.69	0.59
	(1.100)	(1.727)	(1.462)	(1.375)	(7.584)
$Born\ 1950-1954$	0.65	1.01	-0.26	-0.56	0.55
	(0.964)	(1.501)	(1.278)	(1.199)	(6.756)
$Born\ 1955-1959$	0.59	0.90	-0.19	-0.45	0.50
	(0.825)	(1.292)	(1.093)	(1.027)	(5.622)
$Born\ 1960-1964$	0.50	0.75	-0.14	-0.34	0.43
	(0.689)	(1.078)	(0.908)	(0.848)	(4.613)
$Born\ 1965-1969$	0.40	0.61	-0.12	-0.29	0.34

Continued on next page

Table 3.7 – continued from previous page

	(0.556)	(0.871)	(0.739)	(0.689)	(3.841)
Born 1970 − 1974	0.33	0.50	-0.11	-0.23	0.28
	(0.431)	(0.695)	(0.589)	(0.539)	(3.050)
Born 1975 − 1979	0.23	0.34	-0.06	-0.14	0.20
	(0.293)	(0.464)	(0.395)	(0.356)	(2.064)
Born 1980 − 1984	0.11	0.17	-0.03	-0.07	0.10
	(0.156)	(0.239)	(0.200)	(0.188)	(0.973)
$log(BMI_{t-1})$	-0.01				
	(0.489)				
$log(\%\ overweight_t)$		-0.10			
		(0.336)			
$log(\%\ obese1_t)$			0.20		
			(0.209)		
$log(\%\ obese2_t)$				0.23	
				(0.155)	
$log(\%\ obese3_t)$					0.02
					(1.189)
$log(fat_{t-1})$	0.40***	0.40***	0.39***	0.39***	0.40***
	(0.071)	(0.071)	(0.071)	(0.071)	(0.073)
Constant	3.25	3.53**	2.76***	4.48***	3.30
	(2.171)	(1.369)	(0.972)	(1.216)	(5.911)
Observations	195	195	195	195	195
R^2	0.755	0.755	0.756	0.758	0.755

Standard errors in parentheses
$***p < 0.01, **p < 0.05, *p < 0.1$

Table 3.8: Demand for Meat: Men

VARIABLES	(1) $log(BMI_t)$	(2) $log(\%\ overweight_t)$	(3) $log(\%\ obese1_t)$	(4) $log(\%\ obese2_t)$	(5) $log(\%\ obese3_t)$
age	-0.07*	-0.10**	0.02	-0.19*	-0.03
	(0.034)	(0.039)	(0.042)	(0.104)	(0.030)
age^2	-0.00	0.00	-0.00***	0.00	-0.00***
	(0.000)	(0.000)	(0.001)	(0.002)	(0.000)
married	-0.34***	-0.65***	-0.83***	2.13	-1.07***
	(0.098)	(0.208)	(0.243)	(1.597)	(0.404)
food secure	0.35**	0.37**	0.57***	-0.28	-0.18
	(0.150)	(0.146)	(0.166)	(0.467)	(0.315)
$log(income)$	-0.30**	-0.33***	0.50	-1.33*	0.37
	(0.134)	(0.123)	(0.326)	(0.730)	(0.317)
2002	-0.43**	-0.47**	-1.14***	0.30	-0.78***
	(0.191)	(0.189)	(0.348)	(0.505)	(0.265)
2004					-0.20***
					(0.061)
2006	0.08**	0.10***	0.17***	0.05	-0.11*
	(0.031)	(0.029)	(0.044)	(0.046)	(0.058)
2008	0.18***	0.23***	0.44***	-0.33	
	(0.059)	(0.059)	(0.119)	(0.353)	
2010	0.36***	0.42***	0.83***	-0.41	0.14**
	(0.114)	(0.113)	(0.221)	(0.528)	(0.066)
Born 1920 − 1924	4.46**	4.59***	11.96***	-10.65	9.11***
	(1.764)	(1.749)	(3.615)	(9.980)	(3.187)
Born 1925 − 1929	4.11**	4.23***	11.10***	-9.44	8.34***
	(1.630)	(1.615)	(3.351)	(8.984)	(2.899)
Born 1930 − 1934	3.76**	3.92***	10.11***	-8.20	7.47***
	(1.512)	(1.497)	(3.056)	(7.934)	(2.603)
Born 1935 − 1939	3.48**	3.71***	8.98***	-6.74	6.43***
	(1.373)	(1.364)	(2.684)	(6.772)	(2.205)

Continued on next page

Table 3.8 – continued from previous page

Born 1940 − 1944	3.17**	3.40***	8.06***	-6.11	5.89***
	(1.237)	(1.229)	(2.392)	(6.154)	(2.010)
Born 1945 − 1949	2.85**	3.10***	6.99***	-4.95	5.15***
	(1.107)	(1.103)	(2.063)	(5.163)	(1.756)
Born 1950 − 1954	2.51**	2.74***	5.79***	-3.61	4.34***
	(0.967)	(0.966)	(1.692)	(4.049)	(1.474)
Born 1955 − 1959	2.26***	2.48***	4.84***	-2.76	3.69***
	(0.833)	(0.835)	(1.374)	(3.319)	(1.221)
Born 1960 − 1964	1.90***	2.12***	4.05***	-1.89	3.04***
	(0.698)	(0.702)	(1.145)	(2.521)	(0.997)
Born 1965 − 1969	1.58***	1.70***	3.66***	-2.85	2.82***
	(0.561)	(0.559)	(1.040)	(2.924)	(0.919)
Born 1970 − 1974	1.15***	1.24***	2.80***	-2.31	2.35***
	(0.426)	(0.420)	(0.802)	(2.308)	(0.785)
Born 1975 − 1979	0.77***	0.80***	2.04***	-1.95	1.69***
	(0.289)	(0.286)	(0.602)	(1.807)	(0.584)
Born 1980 − 1984	0.45***	0.47***	1.20***	-1.00	0.86***
	(0.162)	(0.159)	(0.347)	(0.971)	(0.276)
$log(meat_{t-1})$	0.00	0.02	0.01	0.03	0.02
	(0.074)	(0.072)	(0.071)	(0.072)	(0.072)
$log(BMI_t)$	1.44				
	(1.362)				
$log(\%\ overweight_t)$		0.99*			
		(0.525)			
$log(\%\ obese1_t)$			-2.00**		
			(0.854)		
$log(\%\ obese2_t)$				3.41	
				(2.265)	
$log(\%\ obese3_t)$					-0.98**
					(0.486)
Constant	3.47	9.58***	-4.63	28.10**	-1.24
	(3.404)	(1.836)	(5.037)	(14.143)	(4.194)
Observations	195	195	195	194	193
R^2	0.742	0.745	0.748	0.741	0.745

Table 3.9: Demand for Added Sugar: Men

VARIABLES	(6) $log(BMI_t)$	(7) $log(\%\ overweight_t)$	(8) $log(\%\ obese1_t)$	(9) $log(\%\ obese2_t)$	(10) $log(\%\ obese3_t)$
age	-0.08*	-0.04	0.04	-0.28**	-0.03
	(0.040)	(0.046)	(0.050)	(0.124)	(0.036)
age^2	0.00***	0.00	-0.00*	0.01**	-0.00
	(0.000)	(0.000)	(0.001)	(0.003)	(0.000)
married	-0.15	-0.14	-0.60**	3.91**	-1.05**
	(0.110)	(0.250)	(0.288)	(1.906)	(0.485)
food secure	-0.03	0.04	0.24	-1.05*	-0.66*
	(0.180)	(0.179)	(0.200)	(0.562)	(0.382)
log(income)	0.26*	0.43***	1.20***	-1.34	1.21***
	(0.156)	(0.149)	(0.388)	(0.872)	(0.381)
2002	1.38***	1.33***	0.63	3.43***	0.74*
	(0.225)	(0.227)	(0.415)	(1.221)	(0.384)
2004				0.91	-0.37**
				(0.629)	(0.162)
2006	0.34***	0.38***	0.46***	1.21**	0.01
	(0.105)	(0.107)	(0.113)	(0.589)	(0.134)
2008	0.31***	0.35***	0.60***	0.39*	0.01
	(0.117)	(0.122)	(0.174)	(0.215)	(0.079)
2010	0.32*	0.38**	0.84***		

Continued on next page

Table 3.9 – continued from previous page

	(0.160)	(0.165)	(0.281)		
$Born$ 1920 − 1924	0.48	0.62	8.30*	-23.95**	7.19*
	(2.069)	(2.096)	(4.312)	(11.897)	(3.829)
$Born$ 1925 − 1929	0.36	0.54	7.68*	-21.54**	6.49*
	(1.912)	(1.936)	(3.996)	(10.710)	(3.482)
$Born$ 1930 − 1934	0.32	0.54	7.01*	-18.96**	5.77*
	(1.773)	(1.794)	(3.645)	(9.459)	(3.127)
$Born$ 1935 − 1939	0.42	0.58	6.17*	-16.06**	4.80*
	(1.612)	(1.635)	(3.201)	(8.073)	(2.649)
$Born$ 1940 − 1944	0.35	0.51	5.47*	-14.61**	4.39*
	(1.452)	(1.473)	(2.852)	(7.336)	(2.414)
$Born$ 1945 − 1949	0.39	0.50	4.70*	-12.18**	3.81*
	(1.300)	(1.323)	(2.461)	(6.156)	(2.109)
$Born$ 1950 − 1954	0.48	0.51	3.86*	-9.40*	3.20*
	(1.136)	(1.158)	(2.018)	(4.828)	(1.771)
$Born$ 1955 − 1959	0.43	0.45	3.07*	-7.67*	2.59*
	(0.980)	(1.002)	(1.639)	(3.958)	(1.466)
$Born$ 1960 − 1964	0.37	0.42	2.59*	-5.72*	2.11*
	(0.821)	(0.842)	(1.366)	(3.007)	(1.198)
$Born$ 1965 − 1969	0.26	0.33	2.44*	-6.88*	2.10*
	(0.659)	(0.671)	(1.240)	(3.487)	(1.105)
$Born$ 1970 − 1974	0.13	0.25	1.90**	-5.43*	1.89**
	(0.500)	(0.505)	(0.956)	(2.753)	(0.944)
$Born$ 1975 − 1979	0.15	0.19	1.48**	-4.26**	1.46**
	(0.339)	(0.343)	(0.717)	(2.155)	(0.701)
$Born$ 1980 − 1984	0.04	0.10	0.85**	-2.29**	0.66**
	(0.190)	(0.191)	(0.414)	(1.158)	(0.332)
$log(BMI_t)$	3.41**				
	(1.563)				
$log(\%\ overweight_t)$		0.23			
		(0.636)			
$log(\%\ obese1_t)$			-2.07**		
			(1.018)		
$log(\%\ obese2_t)$				5.60**	
				(2.701)	
$log(\%\ obese3_t)$					-1.25**
					(0.585)
$log(addedsugar_{t-1})$	0.28***	0.29***	0.29***	0.30***	0.29***
	(0.071)	(0.073)	(0.071)	(0.071)	(0.071)
$Constant$	-10.07**	-1.31	-13.90**	31.98*	-11.86**
	(3.963)	(2.183)	(6.030)	(16.223)	(5.077)
$Observations$	195	195	195	194	193
R^2	0.952	0.950	0.951	0.952	0.952

Standard errors in parentheses
$***p < 0.01, **p < 0.05, *p < 0.1$

Table 3.10: Demand for Fat: Men

	(11)	(12)	(13)	(14)	(15)
VARIABLES	$log(BMI_t)$	$log(\%\ overweight_t)$	$log(\%\ obese1_t)$	$log(\%\ obese2_t)$	$log(\%\ obese3_t)$
age	-0.04	-0.00	0.03	-0.03	-0.01
	(0.030)	(0.035)	(0.038)	(0.094)	(0.027)
age^2	0.00	-0.00	-0.00	0.00	0.00
	(0.000)	(0.000)	(0.001)	(0.002)	(0.000)
$married$	-0.13	-0.01	-0.33	0.36	0.21
	(0.084)	(0.186)	(0.217)	(1.438)	(0.366)
$food\ secure$	0.32**	0.37***	0.46***	0.25	0.55*
	(0.135)	(0.135)	(0.153)	(0.420)	(0.288)

Continued on next page

Table 3.10 – continued from previous page

log(income)	0.04	0.18	0.53*	-0.04	-0.04
	(0.117)	(0.113)	(0.294)	(0.656)	(0.288)
2002	-0.04	-0.06	-0.42	0.08	0.04
	(0.168)	(0.169)	(0.312)	(0.454)	(0.240)
2004					0.00
					(0.055)
2006	0.02	0.04	0.08**	0.04	0.05
	(0.027)	(0.027)	(0.040)	(0.041)	(0.053)
2008	-0.00	0.01	0.14	-0.08	
	(0.052)	(0.054)	(0.107)	(0.318)	
2010	0.03	0.06	0.30	-0.07	0.05
	(0.100)	(0.101)	(0.198)	(0.475)	(0.059)
Born 1920 – 1924	1.13	1.23	5.03	-1.55	-0.89
	(1.555)	(1.568)	(3.246)	(8.979)	(2.893)
Born 1925 – 1929	0.91	1.05	4.57	-1.45	-0.87
	(1.438)	(1.448)	(3.009)	(8.083)	(2.631)
Born 1930 – 1934	0.81	0.96	4.16	-1.26	-0.76
	(1.333)	(1.343)	(2.744)	(7.139)	(2.363)
Born 1935 – 1939	0.79	0.87	3.67	-1.01	-0.53
	(1.211)	(1.223)	(2.410)	(6.092)	(2.002)
Born 1940 – 1944	0.68	0.76	3.25	-0.94	-0.52
	(1.092)	(1.103)	(2.148)	(5.537)	(1.825)
Born 1945 – 1949	0.63	0.66	2.79	-0.76	-0.44
	(0.977)	(0.990)	(1.853)	(4.645)	(1.594)
Born 1950 – 1954	0.62	0.59	2.30	-0.52	-0.31
	(0.853)	(0.867)	(1.519)	(3.643)	(1.338)
Born 1955 – 1959	0.57	0.53	1.88	-0.38	-0.20
	(0.735)	(0.750)	(1.234)	(2.986)	(1.108)
Born 1960 – 1964	0.44	0.42	1.55	-0.27	-0.16
	(0.616)	(0.630)	(1.028)	(2.268)	(0.905)
Born 1965 – 1969	0.37	0.39	1.46	-0.41	-0.19
	(0.495)	(0.502)	(0.934)	(2.631)	(0.835)
Born 1970 – 1974	0.20	0.27	1.10	-0.37	-0.25
	(0.376)	(0.377)	(0.720)	(2.077)	(0.713)
Born 1975 – 1979	0.14	0.18	0.81	-0.32	-0.22
	(0.255)	(0.257)	(0.540)	(1.626)	(0.530)
Born 1980 – 1984	0.08	0.12	0.49	-0.15	-0.06
	(0.143)	(0.142)	(0.312)	(0.874)	(0.251)
$log(BMI_t)$	2.14*				
	(1.195)				
$log(\%\ overweight_t)$		-0.16			
		(0.472)			
$log(\%\ obese1_t)$			-1.02		
			(0.767)		
$log(\%\ obese2_t)$				0.58	
				(2.039)	
$log(\%\ obese3_t)$					0.32
					(0.441)
$log(fat_{t-1})$	0.29***	0.31***	0.31***	0.32***	0.31***
	(0.070)	(0.070)	(0.069)	(0.069)	(0.070)
Constant	-3.82	0.84	-4.60	4.96	3.93
	(3.012)	(1.632)	(4.525)	(12.715)	(3.805)
Observations	195	195	195	194	193
R^2	0.807	0.803	0.805	0.800	0.799

Standard errors in parentheses
$***p < 0.01, **p < 0.05, *p < 0.1$

It is suspect that food security maybe over controlling for income. To investigate further, additional estimations excluding food security can be found in appendix D.

When comparing results from both models the estimates show that in general the coefficients are similar in magnitude and sign. In fact, in some instances the inclusive model yields more significant results for income.[33] The only major discrepancies in income are for male demand for health stock and demand for fat, which are significant in the controlled model, but not in the inclusive model.

3.5 Conclusions

Obesity continues to be a pervasive problem in the United States. Without intervention obesity has the potential to impose financial stress on the economy and decrease individual welfare. Despite this looming crisis, much of the social science literature studies the determinants of obesity as a static model. But, as can be attested, obesity is a time dependent problem and thus most social science obesity studies are merely correlational. To the best of my knowledge this is the first paper to exploit RCS to form a dynamic model using US data. In this paper I construct cohorts based on birth year, race and gender. The transformed data is then used to apply the Grossman Model for dynamic health human capital.

The results suggest that men experience moderate amounts of marital induced weight gain, while women were much more likely to be less heavy because of increased nutritional accountability for other household members. Like the previous literature, women exhibited an inverse relationship with socioeconomic factors (in this case food security and income). Men also did not show evidence that past realizations of weight effect current weight status. This was not the case for women in which the lagged regressor was positive and significant for most estimations. This implies that men do

[33]For example, tables D.3 and D.6 show less significance in income.

not suffer from a *"weight legacy"* and can change current weight outcomes despite previous weight status.

In terms of food consumption there were general trends suggesting that the youngest cohorts eat more sugar and increasingly less meat. It suggests that there are generational shifts in food consumption toward a less animal protein, higher added sugar diet. The preponderance of obesity is distinctly higher among women for overweight and all obese categories. The findings in this study found that women's demand for added sugar is driven by socioeconomic factors, marital status and weight status. This may suggest that elevated levels of obesity in women may at least partially be determined by added sugar consumption which may be amplified by physiological factors inducing increased sugar consumption as weight increases. Habit formation was also a recurring factor for both genders. These results imply that effective intervention will involve efforts to reduce sugar consumption, increase animal protein consumption and break habit formation.

Bibliography

Arellano, Manuel, and Stephen Bond. 1991. "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations." *Review of Economic Studies* 58 (2): 277–97 (April).

Arellano, Manuel, and Olympia Bover. 1995. "Another Look at the Instrumental Variable Estimation of Error-components Models." *Journal of Econometrics* 68 (1): 29–51.

Auchincloss, Amy H., Rick L. Riolo, Daniel G. Brown, Jeremy Cook, and Ana V. Diez Roux. 2011. "An Agent-Based Model of Income Inequalities in Diet in the Context of Residential Segregation." *American Journal of Preventive Medicine* 40 (3): 303 – 311.

Auld, Christopher M., and Paul Grootendorst. 2011. "Challenges for Causal Inference in Obesity Research." In *Oxford Handbook of the Social Science of Obesity*, edited by John Cawley. Oxford: Oxford University Press.

Ball, Kylie, and David Crawford. 2005. "Socioeconomic Status and Weight Change in Adults: A Review." *Social Science & Medicine* 60 (9): 1987 – 2010.

Baum II, Charles L, and Christopher J Ruhm. 2009. "Age, Socioeconomic Status and Obesity Growth." *Journal of Health Economics* 28 (3): 635–648.

Becker, Gary S., and Kevin M. Murphy. 1988. "A Theory of Rational Addiction." *Journal of Political Economy* 96 (4): pp. 675–700.

Berrington de Gonzalez, Amy, Patricia Hartge, James R. Cerhan, Alan J. Flint, Lindsay Hannan, Robert J. MacInnis, Steven C. Moore, Geoffrey S. Tobias, Hoda Anton-Culver, Laura Beane Freeman, and W. Lawrence Beeson. 2010. "Body-Mass Index and Mortality among 1.46 Million White Adults." *New England Journal of Medicine* 363 (23): 2211–2219. PMID: 21121834.

Bhattacharya, Jay, and M. Kate Bundorf. 2009. "The Incidence of the Healthcare Costs of Obesity." *Journal of Health Economics* 28 (3): 649 – 658.

Block, Jason P., S. V. Subramanian, Nicholas A. Christakis, and A. James OMalley. 2013. "Population Trends and Variation in Body Mass Index from 1971 to 2008 in the Framingham Heart Study Offspring Cohort." *PLoS ONE* 8 (5): e63217 (05).

Blundell, Richard, and Stephen Bond. 1998. "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models." *Journal of Econometrics* 87 (1): 115 – 143.

Bruzzone, Agostino, Vera Novak, and Francesca Madeo. 2012. ""Obesity Epidemics Modelling by Using Intelligent Agents"." *SCS M&S Magazine* 9 (3): 18–24 (February).

Cai, L, J He, Y Song, K Zhao, and W Cui. 2013. "Association of Obesity with Socio-economic Factors and Obesity-related Chronic Diseases in Rural Southwest China." *Public Health* 127 (3): 247–251.

Cameron, A Colin, and Pravin K Trivedi. 2005. *Microeconometrics: Methods and Applications*. Cambridge University Press.

Campos, Alba LR, Dirce M Sigulem, Denise EB Moraes, Arlete Escrivão, and Mauro Fisberg. 1996. "Intelligent Quotient of Obese Children and Adolescents by the Weschler Scale." *Revista de Saúde Pública* 30 (1): 85–90.

Cawley, John. 2004. "The Impact of Obesity on Wages." *The Journal of Human Resources* 39 (2): pp. 451–474.

———. 2007. "The Labor Market Impact of Obesity on Wages." Working paper E-FAN-04-004, Economic Research Service/USDA.

Cawley, John, Euna Han, and Edward C Norton. 2009. "Obesity and Labor Market Outcomes Among Legal Immigrants to the United States from Developing Countries." *Economics & Human Biology* 7 (2): 153–164.

Cawley, John, and Chad Meyerhoefer. 2012. "The Medical Care Costs of Obesity: An Instrumental Variables Approach." *Journal of Health Economics* 31 (1): 219 – 230.

Cawley, John, John Moran, and Kosali Simon. 2010. "The Impact of Income on the Weight of Elderly Americans." *Health Economics* 19 (8): 979–993.

CDC, Centers for Disease Control, Prevention (CDC, et al. 2003. "Self-reported Concern about Food Security Associated with Obesity–Washington, 1995-1999." *MMWR. Morbidity and mortality weekly report* 52 (35): 840.

Chee, Heng-Leng, Mirnalini Kandiah, Maimunah Khalid, Khadijah Shamsuddin, Jamilah Jamaluddin, NA Nordin, Rashidah Shuib, and Intan Osman. 2004. "Body Mass Index and Factors Related to Overweight among Women Workers in Electronic Factories in Peninsular Malaysia." *Asia Pac J Clin Nutr* 13 (3): 248–254.

Chen, Zhuo, and David Meltzer. 2008. "Beefing up with the Chans: Evidence for the Effects of Relative Income and Income Inequality on Health from the China Health and Nutrition Survey." *Social Science & Medicine* 66 (11): 2206–2217.

Collado, M. Dolores. 1997. "Estimating Dynamic Models from Time Series of Independent Cross-Sections." *Journal of Econometrics* 82 (1): 37 – 62.

Conley, Dalton, and Rebecca Glauber. 2005, May. "Gender, Body Mass and Economic Status." Working paper 11343, National Bureau of Economic Research.

———. 2007. "Gender, Body Mass, and Socioeconomic Status: New Evidence from the PSID." *Advances in Health Economics and Health Services Research* 17:253–275.

Cook, Andrew, and Beth Daponte. 2008. "A Demographic Analysis of the Rise in the Prevalence of the US Population Overweight and/or Obese." *Population Research and Policy Re view* 27 (4): 403–426.

Cropper, M. L. 1981. "Measuring the Benefits from Reduced Morbidity." *The American Economic Review* 71 (2): pp. 235–240.

Cutler, David M., Edward L. Glaeser, and Jesse M. Shapiro. 2003. "Why Have Americans Become More Obese?" *Journal of Economic Perspectives* 17 (3): 93–118.

Deaton, Angus. 1985. "Panel Data from Time Series of Cross-sections." *Journal of Econometrics* 30 (1-2): 109–126.

Dinour, Lauren M., Dara Bergen, and Ming-Chin Yeh. 2007. "The Food Insecurity – Obesity Paradox: A Review of the Literature and the Role Food Stamps May Play." *Journal of the American Dietetic Association* 107 (11): 1952 – 1961.

Doak, Colleen M, Linda S Adair, Margaret Bentley, Carlos Monteiro, and Barry M Popkin. 2004. "The Dual Burden Household and the Nutrition Transition Paradox." *International Journal of Obesity* 29 (1): 129–136.

Du, Shufa, Tom A Mroz, Fengying Zhai, and Barry M Popkin. 2004. "Rapid Income Growth Adversely Affects Diet Quality In China–particularly For The Poor!" *Social Science & Medicine* 59 (7): 1505–1515.

Finegood, Diane. 2011. "The Complex Systems Science of Obesity." In *Oxford Handbook of the Social Science of Obesity*, edited by John Cawley. Oxford: Oxford University Press.

Finkelstein, Eric A., Olga A. Khavjou, Hope Thompson, Justin G. Trogdon, Liping Pan, Bettylou Sherry, and William Dietz. 2012. "Obesity and Severe Obesity Forecasts Through 2030." *American Journal of Preventive Medicine* 42 (6): 563 – 570.

Finkelstein, Eric A, Justin G Trogdon, Joel W Cohen, and William Dietz. 2009. "Annual Medical Spending Attributable to Obesity: Payer-and Service-specific Estimates." *Health Affairs* 28 (5): w822–w831.

Flegal, Katherine M, Margaret D Carroll, Cynthia L Ogden, and Clifford L Johnson. 2002. "Prevalence and Trends in Obesity Among US Adults, 1999-2000." *JAMA* 288 (14): 1723–1727.

Flegal, KM, MD Carroll, RJ Kuczmarski, and CL Johnson. 1998. "Overweight and Obesity in the United States: Prevalence and Trends, 1960-1994." *International Journal of Obesity* 22 (1): 39–47.

Flegal, KM, MD Carroll, CL Ogden, and LR Curtin. 2010. "Prevalence and Trends in Obesity Among US Adults, 1999-2008." *JAMA* 303 (3): 235–241.

Flegal, KM, WR Harlan, and JR Landis. 1988a. "Secular Trends in Body Mass Index and Skinfold Thickness With Socioeconomic Factors in Young Adult Men." *The American Journal of Clinical Nutrition* 48 (3): 544–551.

———. 1988b. "Secular Trends in Body Mass Index and Skinfold Thickness with Socioeconomic Factors in Young Adult Women." *The American journal of clinical nutrition* 48 (3): 535–543.

Flegal, KM, Carroll MD, Kit BK, and Ogden CL. 2012. "Prevalence of Obesity and Trends in the Distribution of Body Mass Index among US Adults, 1999-2010." *JAMA* 307 (5): 491–497.

Goldman, Dana, Darius Lakdawalla, and Yuhui Zheng. 2009, June. "Food Prices and the Dynamics of Body Weight." Working paper 15096, National Bureau of Economic Research.

Griffiths, C, P Gately, PR Marchant, and CB Cooke. 2013. "Area-level Deprivation and Adiposity in Children: Is the Relationship ?" *International Journal of Obesity* 37 (4): 486–492.

Grimm, V., U. Berger, D.L. DeAngelis, J.G. Polhill, J. Giske, and S.F. Railsback. 2010. "The ODD Protocol: A Review and First Update." *Ecological Modelling* 221 (23): 2760–2768. cited By (since 1996)224.

Grimm, Volker, Uta Berger, Finn Bastiansen, Sigrunn Eliassen, Vincent Ginot, Jarl Giske, John Goss-Custard, Tamara G, Simone K. Heinz, Geir Huse, Andreas Huth, Jane U. Jepsen, Christian Jrgensen, and Wolf M. Mooij. 2006. "A Standard Protocol for Describing Individual-based and Agent-based Models." *Ecological Modelling* 198 (12): 115 – 126.

Grossman, Michael. 1972a. *The Demand for Health: A Theoretical and Empirical, Investigation.* National Bureau of Economic Research, Inc.

———. 1972b. "On the Concept of Health Capital and the Demand for Health." *Journal of Political Economy* 80 (2): 223–255.

Hammond, Ross A. 2009. "Complex Systems Modeling for Obesity Research." *Preventing chronic disease* 6, no. 3.

Hedley, AA, Ogden CL, Johnson CL, Carroll MD, Curtin LR, and Flegal KM. 2004. "Prevalence of Overweight and Obesity among US Children, Adolescents, and Adults, 1999-2002." *JAMA* 291 (23): 2847–2850.

Holtz-Eakin, Douglas, Whitney Newey, and Harvey S Rosen. 1988. "Estimating Vector Autoregressions with Panel Data." *Econometrica: Journal of the Econometric Society*, pp. 1371–1395.

Iannaccone, Laurence R. 1986. "Addiction and satiation." *Economics Letters* 21 (1): 95 – 99.

Jacobson, Michael F, and Kelly D Brownell. 2000. "Small Taxes on Soft Drinks and Snack Foods to Promote Health." *American Journal of Public Health* 90 (6): 854.

Jeffery, Robert W, and Simone A French. 1996. "Socioeconomic Status and Weight Control Practices Among 20-to 45-year-old Women." *American Journal of Public Health* 86 (7): 1005–1010.

Johnson, Rachel J. 2012, November. "Livestock, Dairy, and Poultry Outlook." Technical Report, United States Department of Agriculture- Economic Research Service.

Kristjánsson, Álfgeir Logi, Inga Dóra Sigfúsdóttir, and John P Allegrante. 2010. "Health Behavior and Academic Achievement Among Adolescents: The Relative Contribution of Dietary Habits, Physical Activity, Body Mass Index, and Self-esteem." *Health Education & Behavior* 37 (1): 51–64.

Kuchler, Fred, Abebayehu Tegene, and James Michael Harris. 2004. "Taxing Snack Foods: What to Expect for Diet and Tax Revenues." Technical Report.

Laraia, Barbara A, Anna Maria Siega-Riz, and Kelly R Evenson. 2004. "Self-reported Overweight and Obesity are not Associated with Concern about Enough Food Among Adults in New York and Louisiana." *Preventive Medicine* 38 (2): 175–181.

Lee, David S. 2005, October. "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects." Working paper 11721, National Bureau of Economic Research.

Levi, Jeffrey, Laura M. Segal, Rebecca St. Laurent, Albert Lang, and Jack Rayburn. 2012, September. "F as in Fat: How Obesity Threatens Americas Future." Technical Report, Trust for Americas Health.

Li, Xaioming. 1995. “A Study of Intelligence and Personality in Children with Simple Obesity.” *International Journal of Obesity and Related Metabolic Disorders: Journal of the International Association for the Study of Obesity* 19 (5): 355–357.

Lin, BH, CL Huang, and SA French. 2004. “Factors Associated with Women’s and Children’s Body Mass Indices by Income Status.” *International Journal of Obesity* 28 (4): 536–542.

McLaren, Lindsay. 2007. “Socioeconomic Status and Obesity.” *Epidemiologic reviews* 29 (1): 29–48.

Mendez, MA, RS Cooper, A Luke, R Wilks, F Bennett, and T Forrester. 2004. “Higher Income is more Strongly Associated with Obesity than with Obesity-related Metabolic Disorders in Jamaican Adults.” *International journal of obesity* 28 (4): 543–550.

Mikkilä, Vera, Marjaana Lahti-Koski, Pirjo Pietinen, Suvi M Virtanen, and Matti Rimpelä. 2003. “Associates of Obesity and Weight Dissatisfaction Among Finnish Adolescents.” *Public Health Nutrition* 6 (01): 49–56.

Misra, A. 2003. “Revisions of Cutoffs of Body Mass Index to Define Overweight and Obesity are Needed for the Asian-ethnic Groups.” *International Journal of Obesity* 27 (11): 1294–1296.

Misra, Anoop, and Lokesh Khurana. 2008. “Obesity and the Metabolic Syndrome in Developing Countries.” *The Journal of Clinical Endocrinology & Metabolism* 93 (11_supplement_1): s9–s30.

Moffitt, Robert. 1993. "Identification and estimation of dynamic models with a time series of repeated cross-sections." *Journal of Econometrics* 59 (1-2): 99–123 (September).

Monteiro, Carlos A, Wolney L Conde, and Barry M Popkin. 2004. "The Burden of Disease from Undernutrition and Overnutrition in Countries Undergoing Rapid Nutrition Transition: A View from Brazil." *American Journal of Public Health* 94 (3): 433.

Monteiro, Carlos A, Erly C Moura, Wolney L Conde, and Barry M Popkin. 2004. "Socioeconomic Status and Obesity in Adult Populations of Developing Countries: A Review." *Bulletin of the World Health Organization* 82 (12): 940–946.

Ng, Shu Wen, Edward C. Norton, David K. Guilkey, and Barry M. Popkin. 2012. "Estimation of a Dynamic Model of Weight." *Empirical Economics* 42 (2): 413–443.

NIDDK. 2012. "Overweight and Obesity Statistics." Fact sheet 044158, U.S. Department of Health and Human Services National Institute of Diabetes and Digestive and Kidney Diseases.

Nocera, Sandra, and Peter Zweifel. 1998. "The Demand for Health: An Empirical Test of the Grossman Model using panel data." In *Health, The Medical Profession, and Regulation*, 35–49. Springer.

Nube, M, WK Asenso-Okyere, and GJ Van den Boom. 1998. "Body Mass Index as Indicator of Standard of Living in Developing Countries." *European Journal of Clinical Nutrition* 52 (2): 136–144.

Ogden, Cynthia L, National Center for Health Statistics (US), et al. 2010. *Obesity and Socioeconomic Status in Adults: United States, 2005-2008.* US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics.

Popkin, Barry M. 1998. "The Nutrition Transition and its Health Implications in Lower-income Countries." *Public Health Nutrition* 1 (01): 5–21.

———. 2001. "The Nutrition Transition and Obesity in the Developing World." *The Journal of Nutrition* 131 (3): 871S–873S.

Popkin, Barry M, Linda S Adair, and Shu Wen Ng. 2012. "Global Nutrition Transition and the Pandemic of Obesity in Developing Countries." *Nutrition Reviews* 70 (1): 3–21.

Popkin, Barry M, Shufa Du, Fengying Zhai, and Bing Zhang. 2010. "Cohort Profile: The China Health and Nutrition Surveymonitoring and Understanding Socioeconomic and Health Change in China, 1989–2011." *International Journal of Epidemiology* 39 (6): 1435–1440.

Powell, Lisa M. 2011. "Food Taxes and Subsidies: Evidence and Policies for Obesity Prevention." In *Oxford Handbook of the Social Science of Obesity*, edited by John Cawley. Oxford: Oxford University Press.

Prince, SA. 2009. "A Population Health Approach to Obesity in CanadaPutting the Problem Back into Context." *Transdisciplinary Studies in Population Health Series* 1 (1): 22–33.

Rand, William, and Roland T. Rust. 2011. "Agent-based Modeling in Marketing: Guidelines for Rigor." *International Journal of Research in Marketing* 28 (3):

181 – 193.

Rao, Mayuree, Ashkan Afshin, Gitanjali Singh, and Dariush Mozaffarian. 2013. “Do Healthier Foods and Diet Patterns Cost More than Less Healthy Options? A Systematic Review and Meta-analysis.” *BMJ Open* 3 (12): e004277.

Reddy, B Nirmala. 1998. “Body Mass Index and its Association with Socioeconomic and Behavioral Variables among Socioeconomically Heterogeneous Populations of Andhra Pradesh, India.” *Human Biology*, pp. 901–917.

Richards, Timothy J, and Paul M Patterson. 2006. “Native American Obesity: An Economic Model of the Thrifty Gene Theory.” *American Journal of Agricultural Economics* 88 (3): 542–560.

Roza, Allan M, and Harry M Shizgal. 1984. “The Harris Benedict Equation Reevaluated: Resting Energy Requirements and the Body Cell Mass.” *The American Journal of Clinical Nutrition* 40 (1): 168–182.

Russell, Jeffrey E, and John W Fraas. 2005. “An Application of Panel Regression to Pseudo Panel Data.” *Multiple Linear Regression Viewpoints* 31 (1): 1–15.

Sawaya, Ana L, Gerald Dallal, Gisela Solymos, Maria H Sousa, Maria L Ventura, Susan B Roberts, and Dirce M Sigulem. 1995. “Obesity and Malnutrition in a Shantytown Population in the City of São Paulo, Brazil.” *Obesity Research* 3 (S2): 107s–115s.

Schmeiser, Maximilian D. 2009. “Expanding Wallets and Waistlines: The Impact of Family Income on the BMI of Women and Men Eligible for the Earned Income Tax Credit.” *Health Economics* 18 (11): 1277–1294.

Shah, Syed M, Debra Nanan, Mohammad H Rahbar, Musa Rahim, and Gul Nowshad. 2004. "Assessing Obesity and Overweight in a High Mountain Pakistani Population." *Tropical Medicine & International Health* 9 (4): 526–532.

Shimokawa, Satoru, Hung-Hao Chang, and Per Pinstrup-Andersen. 2009. "Understanding the Differences in Obesity Among Working Adults Between Taiwan and China." *Asia Pacific Journal of Clinical Nutrition* 18 (1): 88.

Sichieri, R, KS Siqueira, and AS Moura. 2000. "Obesity and Abdominal Fatness Associated with Undernutrition Early in Life in a Survey in Rio de Janeiro." *International Journal of obesity and Related Metabolic Disorders: Journal of the International Association for the Study of Obesity* 24 (5): 614–618.

Smith, E, P Hay, L Campbell, and JN Trollor. 2011. "A Review of the Association Between Obesity and Cognitive Function Across the Lifespan: Implications for Novel Approaches to Prevention and Treatment." *Obesity Reviews* 12 (9): 740–755.

Sobal, Jeffery, and Albert J Stunkard. 1989. "Socioeconomic Status and Obesity: A Review of the Literature." *Psychological bulletin* 105 (2): 260.

Sobal, Jeffrey, Barbara S. Rauschenbach, and Edward A. Frongillo Jr. 1992. "Marital Status, Fatness and Obesity." *Social Science & Medicine* 35 (7): 915 – 923.

Stigler, George J, and Gary S Becker. 1977. "De gustibus non est Disputandum." *The American Economic Review*, pp. 76–90.

Tiebout, Charles M. 1956. "A Pure Theory of Local Expenditures." *Journal of Political Economy* 64:416.

Townsend, Marilyn S, Janet Peerson, Bradley Love, Cheryl Achterberg, and Suzanne P Murphy. 2001. “Food Insecurity is Positively Related to Overweight in Women.” *The Journal of Nutrition* 131 (6): 1738–1745.

Ulijaszek, SJ. 2003. “Socio-economic Factors Associated with Physique of Adults of the Purari Delta of the Gulf Province, Papua New Guinea.” *Annals of human biology* 30 (3): 316–328.

Verbeek, Marno, and Francis Vella. 2005. “Estimating Dynamic Models from Repeated Cross-Sections.” *Journal of Econometrics* 127 (1): 83 – 102.

Wagstaff, Adam. 1986. “The Demand for Health: Some New Empirical Evidence.” *Journal of Health Economics* 5 (3): 195 – 233.

———. 1993. “The Demand for Health: An Empirical Reformulation of the Grossman Model.” *Health Economics* 2 (2): 189–98.

Wang, Y Claire, Klim McPherson, Tim Marsh, Steven L Gortmaker, and Martin Brown. 2011. “Health and Economic Burden of the Projected Obesity Trends in the USA and the UK.” *The Lancet* 378 (9793): 815 – 825.

WHO. 1995. “Physical status: The use of and interpretation of anthropometry, Report of a WHO Expert Committee.”

Wu, Yangfeng. 2006. “Overweight and Obesity in China.” *BMJ* 333 (7564): 362–363.

Xiao, Yuanyuan, Naiqing Zhao, Hao Wang, Jie Zhang, Qingfang He, Danting Su, Ming Zhao, Lixin Wang, Xinwei Zhang, Weiwei Gong, et al. 2013. “”Association Between Socioeconomic Status and Obesity in a Chinese Adult Population.” *BMC Public Health* 13 (1): 355.

Xu, Fei, Xiao-Mei Yin, Min Zhang, Eva Leslie, Robert Ware, and Neville Owen. 2005. "Family Average Income And Body Mass Index Above The Healthy Weight Range Among Urban And Rural Residents In Regional Mainland China." *Public Health Nutrition* 8 (01): 47–51.

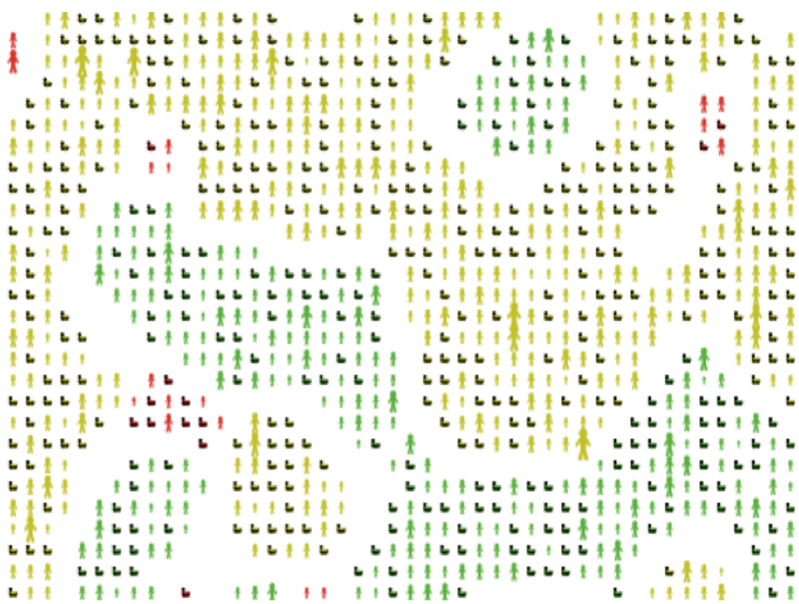
Zhang, Qi, and Youfa Wang. 2004. "Socioeconomic Inequality of Obesity in the United States: Do Gender, Age, and Ethnicity Matter?" *Social Science & Medicine* 58 (6): 1171 – 1180.

Zhou, BF. 2002. "Effect of Body Mass Index on All-cause Mortality and Incidence of Cardiovascular Diseases: Report for Meta-analysis of Prospective Studies on Optimal Cut-off Points of Body Mass Index in Chinese adults." *Biomedical and Environmental Sciences* 15 (3): 245–252.

Appendix A: Tiebout Sorting Model

Figure A.1 is an illustrative example of the Tiebout Sorting Model for one simulation. The sorting criterion is based on income class and is determined by teritiles. The first teritile or the poorest individuals are represented in red while the middle teritile agents are yellow and green agents represent the last teritile. There are two icons represented in figure A.1. The icons, which look like people, are the individual agents. Size corresponds to their current BMI and their initial coordinates represent the position of their household, which effects their local food environment. The other icon represents the restaurants and color and coordinates are similarly representative of income class and location.

Figure A.1: Depiction of Tiebout Sorting Model



Appendix B: Monthly Food Prices

Table B.1: U.S. City Average Price Data

Carbohydrate Rich Foods

Food item	Jan 1999	Feb 1999	Mar 1999	Apr 1999	May 1999	Jun 1999	Jul 1999	Aug 1999	Sep 1999	Oct 1999	Nov 1999	Dec 1999
Flour, white, all purpose, per lb. (453.6 gm)	0.297	0.293	0.289	0.283	0.306	0.303	0.308	0.311	0.31	0.288	0.272	0.267
Rice, white, long grain, uncooked, per lb. (453.6 gm)	0.551	0.54	0.544	0.548	0.551	0.553	0.55	0.557	0.547	0.511	0.488	0.502
Spaghetti and macaroni, per lb. (453.6 gm)	0.872	0.88	0.883	0.888	0.887	0.882	0.88	0.877	0.882	0.849	0.882	0.881
Bread, white, pan, per lb. (453.6 gm)	0.872	0.88	0.883	0.897	0.886	0.885	0.893	0.884	0.878	0.889	0.899	0.899
Bread, whole wheat, pan, per lb. (453.6 gm)	1.302	1.308	1.319	1.27	1.332	1.306	1.344	1.348	1.342	1.371	1.361	1.363
Cookies, chocolate chip, per lb. (453.6 gm)	2.61	2.594	2.607	2.564	2.573	2.54	2.609	2.602	2.583	2.584	2.627	2.673
Crackers, soda, salted, per lb. (453.6 gm)	1.59	1.559	1.52	1.615	1.595	1.696	1.714	1.643	1.612	1.535	1.608	1.65
Apples, Red Delicious, per lb. (453.6 gm)	0.86	0.87	0.852	0.87	0.881	0.893	0.905	0.921	0.972	0.919	0.902	0.918
Bananas, per lb. (453.6 gm)	0.489	0.509	0.506	0.482	0.492	0.502	0.494	0.49	0.481	0.471	0.48	0.494
Oranges, Navel, per lb. (453.6 gm)	0.83	0.889	0.869	0.944							0.884	0.641
Oranges, Valencia, per lb. (453.6 gm)					0.865	0.942	0.959	0.989	0.974	0.955		
Grapefruit, per lb. (453.6 gm)	0.543	0.545	0.546	0.556	0.606	0.712	0.778	0.803	0.762	0.71	0.631	0.582
Lemons, per lb. (453.6 gm)	1.402	1.274	1.167	1.188	1.159	1.183	1.282	1.397	1.463	1.535	1.538	1.414
Pears, Anjou, per lb. (453.6 gm)	0.923	0.925	0.942	0.953	0.96	0.913						1.034
Peaches, per lb. (453.6 gm)		1.856	1.941			1.413	1.16	1.098	1.1			
Strawberries, dry pint, per 12 oz. (340.2 gm)		2.102	1.96	1.751	1.419	1.49	1.375	1.557	1.679	1.664	1.948	
Grapes, Thompson Seedless, per lb. (453.6 gm)	2.341	1.663	1.613	2.262		1.864	1.678	1.522	1.453	1.557	1.897	2.403
Potatoes, white, per lb. (453.6 gm)	0.381	0.382	0.384	0.38	0.388	0.391	0.411	0.429	0.413	0.393	0.384	0.395
Lettuce, iceberg, per lb. (453.6 gm)	0.649	0.658	0.774	0.753	0.691	0.652	0.627	0.652	0.623	0.669	0.677	0.668
Tomatoes, field grown, per lb. (453.6 gm)	1.904	1.476	1.395	1.298	1.284	1.304	1.287	1.232	1.272	1.279	1.3	1.405
Cabbage, per lb. (453.6 gm)	0.425	0.412	0.396	0.406	0.421	0.42	0.402	0.414	0.429	0.435	0.445	0.424
Celery, per lb. (453.6 gm)	0.59	0.563	0.563	0.55	0.556	0.621	0.63	0.607	0.586	0.554	0.56	0.563
Carrots, short trimmed and topped, per lb. (453.6 gm)	0.552	0.575	0.578	0.578	0.604	0.587	0.574	0.572	0.521	0.527	0.544	0.523
Peppers, sweet, per lb. (453.6 gm)	1.429	1.311	1.334	1.456	1.622	1.308	1.429	1.25	1.321	1.308	1.617	1.53
Cucumbers, per lb. (453.6 gm)			1.07	0.91	0.833	0.764	0.709	0.726	0.798	0.955	0.976	0.841
Broccoli, per lb. (453.6 gm)	1.123	0.999	0.99	1.012	0.952	0.944	0.993	0.962	1.052	1.028	1.001	1.004
Orange juice, frozen concentrate, 12 oz. can, per 16 oz. (473.2 ml)	1.753	1.78	1.741	1.779	1.764	1.758	1.813	1.825	1.825	1.784	1.841	1.822
Potatoes, frozen, French fried, per lb. (453.6 gm)	1	1.022	0.969	0.995	1.03	1.003	0.955	1.042	1.036	1.076	1.032	1.039
Sugar, white, all sizes, per lb. (453.6 gm)	0.436	0.43	0.437	0.432	0.436	0.431	0.432	0.431	0.437	0.438	0.426	0.426
Sugar, white, 33-80 oz. pkg, per lb. (453.6 gm)	0.421	0.412	0.421	0.415	0.42	0.415	0.417	0.415	0.422	0.422	0.408	0.408
Cola, nondiet, per 2 liters (67.6 oz)	1.044	1.023	1.012	1.031	1.074	1.02	1.052	1.044	1.034	1.052	1.004	1.029
Potato chips, per 16 oz.	3.217	3.223	3.249	3.264	3.212	3.235	3.255	3.279	3.237	3.289	3.299	3.33

Continued on next page

Table B.1 – continued from previous page

Protein Rich Foods

Food Item	Jan 1999	Feb 1999	Mar 1999	Apr 1999	May 1999	Jun 1999	Jul 1999	Aug 1999	Sep 1999	Oct 1999	Nov 1999	Dec 1999
Ham, rump or shank half, bone-in, smoked,per lb. (453.6 gm)	2.03	1.961	1.974					2.118	2.152			
Ham, boneless, excluding canned, per lb. (453.6 gm)	2.718	2.813	2.837	2.721	2.793	2.866	3.007	3.032	3.025	2.947	2.838	2.766
Sausage, fresh, loose, per lb. (453.6 gm)	2.362	2.311	2.366	2.387	2.45	2.441	2.485	2.498	2.457	2.439	2.385	2.502
Bologna, all beef or mixed, per lb. (453.6 gm)	2.41	2.451	2.448	2.388	2.424	2.382	2.315	2.417	2.346	2.514	2.481	2.521
Chicken, fresh, whole, per lb. (453.6 gm)	1.072	1.064	1.057	1.057	1.026	1.041	1.045	1.043	1.08	1.055	1.078	1.053
Chicken breast, bone-in, per lb. (453.6 gm)	2.072	2.036	2.084	2.111	2.039	2.077	2.062	2.063	2.103	2.088	2.111	2.075
Chicken legs, bone-in, per lb. (453.6 gm)	1.274	1.28	1.281	1.253	1.237	1.283	1.316	1.284	1.269	1.274	1.268	1.168
Turkey, frozen, whole, per lb. (453.6 gm)	0.969	1.001	0.984	0.936	0.975	1.005	1.031	1.034	1.018	1.025	0.964	0.976
Tuna, light, chunk, per lb. (453.6 gm)	2.095	2.071	2.04	2.119	2.145	2.018	2.038	2.071	2.05	2.036	2.112	2.032
Beans, dried, any type, all sizes, per lb. (453.6 gm)	0.689	0.696	0.697	0.695	0.7	0.707	0.7	0.703	0.689	0.689	0.704	0.688
All uncooked ground beef, per lb. (453.6 gm)	1.849	1.887	1.87	1.884	1.869	1.889	1.861	1.877	1.916	1.925	1.93	1.933
All Uncooked Beef Roasts, per lb. (453.6 gm)	2.654	2.668	2.688	2.71	2.67	2.701	2.724	2.712	2.745	2.786	2.75	2.739
All Uncooked Beef Steaks, per lb. (453.6 gm)	3.69	3.676	3.669	3.697	3.746	3.783	3.78	3.772	3.782	3.822	3.862	3.882
All Uncooked Other Beef (Excluding Veal), per lb. (453.6 gm)	2.176	2.209	2.237	2.222	2.217	2.233	2.215	2.241	2.248	2.288	2.33	2.337
All Ham (Excluding Canned Ham and Luncheon Slices), per lb. (453.6 gm)	2.009	1.983	2.019	1.926	1.99	2.017	2.074	2.084	2.097	2.182	2.172	2.116
All Pork Chops, per lb. (453.6 gm)	2.866	2.99	2.937	2.951	3.007	3.026	3.043	3.066	3.112	3.057	3.056	3.061
All Other Pork (Excluding Canned Ham and Luncheon Slices), per lb. (453.6 gm)	1.632	1.646	1.647	1.651	1.67	1.666	1.654	1.675	1.71	1.654	1.713	1.673

Fat Rich Foods

Food Item	Jan 1999	Feb 1999	Mar 1999	Apr 1999	May 1999	Jun 1999	Jul 1999	Aug 1999	Sep 1999	Oct 1999	Nov 1999	Dec 1999
Butter, salted, grade AA, stick, per lb. (453.6 gm)	3.002	2.801	2.739	2.703	2.546	2.648	2.67	2.722	2.628	2.656	2.449	2.272
Shortening, vegetable oil blends, per lb. (453.6 gm)	1.053	1.053	1.055	1.05	1.053	1.046	1.052	1.053	1.041	1.027	0.996	1.037
Peanut butter, creamy, all sizes, per lb. (453.6 gm)	1.769	1.788	1.818	1.822	1.809	1.817	1.831	1.819	1.823	1.83	1.835	1.861

(Source: http://www.bls.gov/data/)

Appendix C: Derivation of Optimal Solution for Health Stock

Taking into account equations 1.1-1.5, the Hamiltonian is the following:

$$\Lambda = e^{\rho t} U[s(H(t)), Z(t)] + \lambda(t)[rA(t) + Y[s(H(t))] \\ - \pi(t)^H I(t) - \pi(t)^Z Z(t)] + \mu(t)[I(D(t), t^p) - \delta(t, R(t))\, H(t)] \tag{C.1}$$

$\lambda(t)$ and $\mu(t)$ are time dependent shadow prices for pecuniary assets and health stock respectively. $H(t)$ and $A(t)$ are state variables while $I(t)$ and $Z(t)$ are control variables. Taking the partial derivatives with respect to the control variables $I(t)$ and $Z(t)$ and setting to zero yields the following:

$$\frac{\partial \Lambda}{\partial I(t)} = \mu(t) - \lambda(t)\, \pi^H(t) \tag{C.2}$$

And

$$\frac{\partial \Lambda}{\partial Z(t)} = e^{-\rho t} \frac{\partial U(t)}{\partial Z(t)} - \lambda(t)\pi^Z(t) = 0 \tag{C.3}$$

Taking derivatives of C.2 with respect to t results in,

$$\dot{\mu}(t) = \dot{\lambda}(t)\pi^H(t) + \lambda(t)\dot{\pi}^H \tag{C.4}$$

The derivative of the Hamiltonian with respect to the state variables should be equal the negative of the derivative of the Lagrangian multiplier with respect to time and yields,

$$\frac{\partial \Lambda}{\partial H(t)} = -\dot{\mu}(t) = e^{-\rho t}\frac{\partial U(t)}{\partial s(t)}\frac{\partial s(t)}{\partial H(t)} + \lambda(t)\frac{\partial Y(t)}{\partial s(t)}\frac{\partial s(t)}{\partial H(t)} - \mu(t)\delta(t, R(t)) \quad \text{(C.5)}$$

And

$$\frac{\partial \Lambda}{\partial A(t)} = -\dot{\lambda}(t) = r\lambda(t) \quad \text{(C.6)}$$

We integrate C.6 with respect to t to get a value for $\lambda(t)$,

$$\lambda(t) = \lambda(0)e^{-rt} \quad \text{(C.7)}$$

Setting C.4 equal to C.5 and

$$-(\dot{\lambda}(t)\pi^H(t) + \lambda(t)\dot{\pi}^H) = e^{-\rho t}\frac{\partial U(t)}{\partial s(t)}\frac{\partial s(t)}{\partial H(t)} + \lambda(t)\frac{\partial Y(t)}{\partial s(t)}\frac{\partial s(t)}{\partial H(t)} - \mu(t)\delta(t, R(t)) \quad \text{(C.8)}$$

Substituting using C.2, C.4, C.6 and C.7,

$$\begin{aligned} r\lambda(0)e^{-rt}\pi^H(t) + \lambda(0)e^{-rt}\dot{\pi}^H = \\ & e^{-\rho t}\frac{\partial U(t)}{\partial s(t)}\frac{\partial s(t)}{\partial H(t)} + \lambda(0)e^{-rt}\frac{\partial Y(t)}{\partial s(t)}\frac{\partial s(t)}{\partial H(t)} \\ & - \lambda(0)e^{-rt}\pi^H(t)\delta(t, R(t)) \end{aligned} \quad \text{(C.9)}$$

Finally, we rearrange to get 2.5

$$\left\{\frac{\partial U(t)/\delta s(t))}{\partial \pi(0)}e^{-(\rho-r)t)} + \frac{\partial Y(t)}{\partial s(t)}\right\}\frac{\partial s(t)}{\partial H(t)} = \left\{r + \delta(t) - \frac{\dot{\pi}^H(t)}{\pi^H(t)}\pi^H(t)\right\}$$

Appendix D: Demand for Health Stock and Health Inputs Excluding Food Security

Table D.1: Demand for Health Stock: Women

VARIABLES	(1) $log(BMI_t)$	(2) $log(\%\ overweight_t)$	(3) $log(\%\ obese1_t)$	(4) $log(\%\ obese2_t)$	(5) $log(\%\ obese3_t)$
$log(BMI_{t-t})$	0.37***				
	(0.067)				
$log(\%\ overweight_{t-1})$		0.23***			
		(0.065)			
$log(\%\ obese1_{t-1})$			0.17**		
			(0.071)		
$log(\%\ obese2_{t-1})$				0.15**	
				(0.069)	
$log(\%\ obese3_{t-1})$					0.01
					(0.079)
age	-0.00	-0.06*	-0.05	-0.05	-0.08
	(0.010)	(0.030)	(0.055)	(0.077)	(0.117)
age^2	-0.00***	-0.00***	-0.00***	-0.00***	-0.00***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$married$	-0.15***	-0.26***	-0.73***	-1.27***	-1.72***
	(0.027)	(0.069)	(0.133)	(0.184)	(0.274)
$log(income)$	-0.07**	-0.36***	-0.33**	0.33	0.39
	(0.028)	(0.082)	(0.147)	(0.205)	(0.310)
2004	0.15**	0.64***	0.79**	0.83*	0.91
	(0.065)	(0.189)	(0.351)	(0.494)	(0.746)
2006	0.15**	0.64***	0.81**	0.78	0.75
	(0.064)	(0.188)	(0.349)	(0.491)	(0.740)
2008	0.18**	0.86***	1.05**	1.02	1.02
	(0.082)	(0.238)	(0.441)	(0.621)	(0.936)
2010	0.23**	1.08***	1.31**	1.28*	1.31
	(0.101)	(0.295)	(0.548)	(0.771)	(1.162)
$Born$ $1920-1924$	1.14*	6.40***	7.91**	9.46**	11.21
	(0.604)	(1.766)	(3.277)	(4.617)	(6.962)
$Born$ $1925-1929$	1.04*	5.81***	7.12**	8.48**	10.10
	(0.556)	(1.624)	(3.014)	(4.246)	(6.404)
$Born$ $1930-1934$	0.96*	5.29***	6.45**	7.60*	8.92
	(0.514)	(1.502)	(2.789)	(3.927)	(5.923)
$Born$ $1935-1939$	0.89*	4.81***	5.92**	6.95*	8.10
	(0.468)	(1.368)	(2.538)	(3.574)	(5.391)
$Born$ $1940-1944$	0.78*	4.29***	5.27**	6.07*	7.18
	(0.422)	(1.233)	(2.288)	(3.221)	(4.859)

Continued on next page

Table D.1 – continued from previous page

Born 1945 − 1949	0.70*	3.86***	4.72**	5.34*	6.24
	(0.379)	(1.106)	(2.053)	(2.890)	(4.360)
Born 1950 − 1954	0.61*	3.35***	4.13**	4.65*	5.56
	(0.331)	(0.968)	(1.796)	(2.529)	(3.815)
Born 1955 − 1959	0.51*	2.87***	3.50**	3.94*	4.62
	(0.285)	(0.833)	(1.545)	(2.176)	(3.282)
Born 1960 − 1964	0.42*	2.39***	2.89**	3.21*	3.79
	(0.238)	(0.697)	(1.293)	(1.820)	(2.746)
Born 1965 − 1969	0.33*	1.93***	2.37**	2.63*	3.16
	(0.193)	(0.562)	(1.044)	(1.469)	(2.216)
Born 1970 − 1974	0.28*	1.56***	1.96**	2.13*	2.51
	(0.146)	(0.426)	(0.791)	(1.113)	(1.678)
Born 1975 − 1979	0.20**	1.04***	1.31**	1.39*	1.70
	(0.098)	(0.286)	(0.531)	(0.746)	(1.126)
Born 1980 − 1984	0.10*	0.53***	0.63**	0.72*	0.80
	(0.053)	(0.155)	(0.288)	(0.405)	(0.611)
Constant	2.76***	3.90***	3.09**	-4.03*	-4.63
	(0.400)	(0.836)	(1.514)	(2.117)	(3.196)
Observations	195	195	195	195	195
R^2	0.744	0.698	0.545	0.459	0.416

Standard errors in parentheses
$***p < 0.01, **p < 0.05, *p < 0.1$

Table D.2: Demand for Health Stock: Men

VARIABLES	(1) $log(BMI_t)$	(2) $log(\%\ overweight_t)$	(3) $log(\%\ obese1_t)$	(4) $log(\%\ obese2_t)$	(5) $log(\%\ obese3_t)$
$log(BMI_t)$	0.19***				
	(0.072)				
$log(\%\ overweight_{t-1})$		0.12			
		(0.073)			
$log(\%\ obese1_{t-1})$			-0.04		
			(0.071)		
$log(\%\ obese2_{t-1})$				0.01	
				(0.076)	
$log(\%\ obese3_{t-1})$					-0.04
					(0.077)
age	0.01	0.03	0.04	0.04	0.01
	(0.008)	(0.035)	(0.067)	(0.110)	(0.132)
age^2	-0.00***	-0.00***	-0.00***	-0.00***	-0.00**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
married	0.02	0.32***	-0.27	-0.71**	-0.79**
	(0.024)	(0.100)	(0.187)	(0.316)	(0.369)
$log(income)$	0.06**	0.11	0.42**	0.41	0.34
	(0.023)	(0.096)	(0.186)	(0.317)	(0.370)
2004	0.04	0.06	0.31	0.23	0.17
	(0.053)	(0.223)	(0.426)	(0.705)	(0.842)
2006	0.04	0.05	0.35	0.23	0.20
	(0.052)	(0.220)	(0.421)	(0.697)	(0.834)
2008	0.05	0.04	0.42	0.37	0.25
	(0.066)	(0.277)	(0.530)	(0.879)	(1.053)
2010	0.06	0.06	0.51	0.44	0.23
	(0.082)	(0.346)	(0.661)	(1.096)	(1.311)
Born 1920 − 1924	0.27	0.65	3.47	4.55	5.09
	(0.491)	(2.057)	(3.932)	(6.523)	(7.785)
Born 1925 − 1929	0.27	0.65	3.23	4.11	4.56

Continued on next page

Table D.2 – continued from previous page

	(0.453)	(1.901)	(3.630)	(6.018)	(7.201)
Born 1930 − 1934	0.26	0.55	2.93	3.63	4.00
	(0.420)	(1.761)	(3.364)	(5.581)	(6.663)
Born 1935 − 1939	0.21	0.40	2.54	3.10	3.21
	(0.382)	(1.602)	(3.059)	(5.076)	(6.061)
Born 1940 − 1944	0.20	0.34	2.25	2.82	2.97
	(0.344)	(1.444)	(2.756)	(4.573)	(5.462)
Born 1945 − 1949	0.16	0.24	1.90	2.36	2.56
	(0.308)	(1.293)	(2.469)	(4.096)	(4.893)
Born 1950 − 1954	0.12	0.17	1.50	1.85	2.07
	(0.269)	(1.130)	(2.159)	(3.580)	(4.277)
Born 1955 − 1959	0.10	0.11	1.17	1.52	1.64
	(0.232)	(0.974)	(1.861)	(3.086)	(3.688)
Born 1960 − 1964	0.09	0.06	0.98	1.16	1.30
	(0.194)	(0.816)	(1.559)	(2.586)	(3.091)
Born 1965 − 1969	0.08	0.11	0.96	1.33	1.39
	(0.156)	(0.654)	(1.250)	(2.076)	(2.483)
Born 1970 − 1974	0.08	0.12	0.75	1.05	1.27
	(0.118)	(0.493)	(0.943)	(1.566)	(1.874)
Born 1975 − 1979	0.05	0.10	0.59	0.83	0.97
	(0.080)	(0.334)	(0.640)	(1.064)	(1.273)
Born 1980 − 1984	0.04	0.07	0.35	0.45	0.42
	(0.044)	(0.186)	(0.355)	(0.590)	(0.705)
Constant	1.82***	-2.39**	-6.75***	-7.16**	-6.29
	(0.299)	(1.051)	(2.039)	(3.425)	(4.029)
Observations	195	195	195	193	191
R^2	0.700	0.624	0.399	0.205	0.244

Standard errors in parentheses
$***p < 0.01, **p < 0.05, *p < 0.1$

Table D.3: Demand for Meat: Women

VARIABLES	(1) $log(BMI_t)$	(2) $log(\%\ overweight_t)$	(3) $log(\%\ obese1_t)$	(4) $log(\%\ obese2_t)$	(5) $log(\%\ obese3_t)$
age	-0.03	0.01	0.01	-0.00	0.06
	(0.032)	(0.036)	(0.034)	(0.034)	(0.129)
age^2	0.00	0.00	0.00*	0.00	0.00
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
married	0.03	-0.06	0.33	0.42	1.60
	(0.148)	(0.124)	(0.221)	(0.290)	(2.565)
log(income)	0.02	0.17	0.13	-0.28**	-0.54
	(0.104)	(0.158)	(0.121)	(0.112)	(0.590)
2004	0.15	-0.07	-0.21	-0.08	-0.61
	(0.228)	(0.293)	(0.288)	(0.278)	(1.389)
2006	0.07	-0.13	-0.28	-0.11	-0.50
	(0.229)	(0.294)	(0.292)	(0.269)	(1.144)
2008	0.11	-0.19	-0.37	-0.15	-0.70
	(0.287)	(0.383)	(0.374)	(0.344)	(1.555)
2010	0.13	-0.23	-0.45	-0.19	-0.90
	(0.360)	(0.477)	(0.465)	(0.430)	(1.988)
Born 1920 − 1924	1.01	-1.52	-2.96	-2.24	-9.20
	(2.078)	(2.842)	(2.805)	(2.855)	(16.889)
Born 1925 − 1929	0.88	-1.41	-2.69	-2.01	-8.31
	(1.908)	(2.595)	(2.551)	(2.588)	(15.229)
Born 1930 − 1934	0.84	-1.20	-2.33	-1.68	-7.19
	(1.767)	(2.376)	(2.332)	(2.354)	(13.464)

Continued on next page

Table D.3 – continued from previous page

Born 1935 − 1939	0.73	-1.10	-2.18	-1.56	-6.53
	(1.613)	(2.164)	(2.133)	(2.148)	(12.221)
Born 1940 − 1944	0.73	-0.91	-1.86	-1.25	-5.71
	(1.449)	(1.938)	(1.908)	(1.904)	(10.836)
Born 1945 − 1949	0.63	-0.87	-1.70	-1.10	-4.94
	(1.299)	(1.744)	(1.712)	(1.693)	(9.426)
Born 1950 − 1954	0.56	-0.72	-1.47	-0.94	-4.42
	(1.139)	(1.522)	(1.498)	(1.479)	(8.404)
Born 1955 − 1959	0.50	-0.63	-1.24	-0.79	-3.64
	(0.974)	(1.308)	(1.280)	(1.265)	(6.991)
Born 1960 − 1964	0.42	-0.53	-1.02	-0.62	-2.97
	(0.814)	(1.092)	(1.063)	(1.042)	(5.736)
Born 1965 − 1969	0.38	-0.39	-0.81	-0.48	-2.46
	(0.657)	(0.883)	(0.866)	(0.849)	(4.778)
Born 1970 − 1974	0.24	-0.36	-0.73	-0.42	-1.98
	(0.509)	(0.697)	(0.690)	(0.665)	(3.792)
Born 1975 − 1979	0.14	-0.24	-0.49	-0.26	-1.34
	(0.346)	(0.467)	(0.463)	(0.440)	(2.568)
Born 1980 − 1984	0.09	-0.10	-0.20	-0.12	-0.60
	(0.185)	(0.243)	(0.233)	(0.230)	(1.209)
$log(meat_{t-1})$	0.29***	0.31***	0.29***	0.29***	0.34***
	(0.071)	(0.068)	(0.069)	(0.070)	(0.068)
$log(BMI_t)$	1.38**				
	(0.602)				
$log(\%\ overweight_t)$		0.67**			
		(0.315)			
$log(\%\ obese1_t)$			0.71***		
			(0.250)		
$log(\%\ obese2_t)$				0.50**	
				(0.205)	
$log(\%\ obese3_t)$					1.07
					(1.481)
Constant	-1.73	1.08	1.91	6.19***	8.92
	(2.643)	(1.657)	(1.208)	(1.377)	(7.029)
Observations	195	195	195	195	195
R^2	0.650	0.648	0.655	0.651	0.640

Table D.4: Demand for Added Sugar: Women

	(6)	(7)	(8)	(9)	(10)
VARIABLES	$log(BMI_t)$	$log(\%\ overweight_t)$	$log(\%\ obese1_t)$	$log(\%\ obese2_t)$	$log(\%\ obese3_t)$
age	-0.05	-0.00	-0.01	-0.02	0.25*
	(0.035)	(0.040)	(0.037)	(0.037)	(0.135)
age^2	0.00***	0.00***	0.00***	0.00***	0.00**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
married	0.37**	0.26*	0.61***	0.73**	6.15**
	(0.157)	(0.133)	(0.235)	(0.304)	(2.687)
$log(income)$	0.34***	0.53***	0.43***	-0.01	-1.24**
	(0.112)	(0.169)	(0.132)	(0.112)	(0.612)
2004	-1.42***	-1.70***	-1.73***	-1.62***	-4.41***
	(0.246)	(0.319)	(0.310)	(0.299)	(1.457)
2006	-1.12***	-1.38***	-1.45***	-1.28***	-3.52***
	(0.252)	(0.316)	(0.319)	(0.293)	(1.206)
2008	-1.21***	-1.59***	-1.65***	-1.43***	-4.52***
	(0.311)	(0.409)	(0.403)	(0.370)	(1.636)
2010	-1.27***	-1.73***	-1.80***	-1.54***	-5.49***
	(0.387)	(0.509)	(0.499)	(0.461)	(2.090)

Continued on next page

Table D.4 – continued from previous page

Born 1920 − 1924	-1.14	-4.46	-4.80	-4.24	-38.83**
	(2.244)	(3.076)	(3.000)	(3.048)	(17.692)
Born 1925 − 1929	-0.93	-3.95	-4.21	-3.68	-34.89**
	(2.061)	(2.810)	(2.731)	(2.766)	(15.954)
Born 1930 − 1934	-0.86	-3.55	-3.77	-3.25	-30.75**
	(1.909)	(2.575)	(2.498)	(2.519)	(14.107)
Born 1935 − 1939	-0.75	-3.16	-3.40	-2.90	-27.83**
	(1.744)	(2.347)	(2.286)	(2.299)	(12.806)
Born 1940 − 1944	-0.66	-2.82	-3.03	-2.52	-24.67**
	(1.566)	(2.102)	(2.046)	(2.040)	(11.356)
Born 1945 − 1949	-0.53	-2.50	-2.66	-2.15	-21.36**
	(1.405)	(1.892)	(1.836)	(1.816)	(9.879)
Born 1950 − 1954	-0.46	-2.15	-2.32	-1.86	-19.05**
	(1.232)	(1.651)	(1.608)	(1.587)	(8.808)
Born 1955 − 1959	-0.31	-1.80	-1.91	-1.52	-15.76**
	(1.055)	(1.420)	(1.375)	(1.359)	(7.328)
Born 1960 − 1964	-0.23	-1.47	-1.54	-1.19	-12.89**
	(0.881)	(1.186)	(1.143)	(1.121)	(6.014)
Born 1965 − 1969	-0.18	-1.19	-1.28	-0.99	-10.76**
	(0.713)	(0.961)	(0.932)	(0.914)	(5.010)
Born 1970 − 1974	-0.16	-0.95	-1.04	-0.77	-8.51**
	(0.552)	(0.759)	(0.742)	(0.716)	(3.976)
Born 1975 − 1979	-0.11	-0.62	-0.68	-0.47	-5.75**
	(0.375)	(0.509)	(0.499)	(0.474)	(2.693)
Born 1980 − 1984	-0.03	-0.28	-0.28	-0.21	-2.66**
	(0.200)	(0.264)	(0.252)	(0.249)	(1.269)
$log(BMI_{t-1})$	1.76***				
	(0.606)				
$log(\%\ overweight_t)$		0.87***			
		(0.331)			
$log(\%\ obese1_t)$			0.75***		
			(0.257)		
$log(\%\ obese2_t)$				0.54**	
				(0.209)	
$log(\%\ obese3_t)$					3.56**
					(1.548)
$log(addedsugar_t)$	0.29***	0.30***	0.28***	0.28***	0.27***
	(0.076)	(0.077)	(0.076)	(0.077)	(0.077)
Constant	-4.98*	-1.35	0.21	4.77***	19.29***
	(2.843)	(1.837)	(1.314)	(1.253)	(7.248)
Observations	195	195	195	195	195
R^2	0.954	0.954	0.954	0.954	0.953

Standard errors in parentheses
$***p < 0.01, **p < 0.05, *p < 0.1$

Continued on next page

Table D.5: Demand for Fat: Women

VARIABLES	(11) $log(BMI_t)$	(12) $log(\%\ overweight_t)$	(13) $log(\%\ obese1_t)$	(14) $log(\%\ obese2_t)$	(15) $log(\%\ obese3_t)$
age	-0.02	-0.04	-0.01	-0.01	-0.03
	(0.028)	(0.031)	(0.030)	(0.030)	(0.112)
age^2	-0.00	-0.00	0.00	0.00	-0.00
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
married	-0.21	-0.27**	-0.04	0.19	-0.45
	(0.128)	(0.109)	(0.191)	(0.246)	(2.234)
$log(income)$	0.05	-0.07	0.13	0.00	0.14

Continued on next page

Table D.5 – continued from previous page

	(0.094)	(0.138)	(0.110)	(0.090)	(0.505)
2004	0.13	0.32	-0.02	-0.15	0.25
	(0.198)	(0.255)	(0.251)	(0.239)	(1.209)
2006	0.09	0.27	-0.07	-0.18	0.17
	(0.200)	(0.255)	(0.254)	(0.232)	(0.996)
2008	0.11	0.36	-0.10	-0.24	0.23
	(0.250)	(0.331)	(0.325)	(0.296)	(1.355)
2010	0.16	0.47	-0.10	-0.28	0.31
	(0.313)	(0.412)	(0.404)	(0.371)	(1.732)
Born 1920 − 1924	1.20	3.13	-0.27	-1.79	2.78
	(1.812)	(2.458)	(2.434)	(2.444)	(14.710)
Born 1925 − 1929	1.11	2.87	-0.21	-1.58	2.53
	(1.664)	(2.246)	(2.215)	(2.217)	(13.264)
Born 1930 − 1934	1.08	2.67	-0.12	-1.34	2.32
	(1.541)	(2.058)	(2.026)	(2.018)	(11.725)
Born 1935 − 1939	1.00	2.45	-0.11	-1.21	2.12
	(1.407)	(1.874)	(1.854)	(1.842)	(10.644)
Born 1940 − 1944	0.94	2.22	-0.05	-1.01	1.93
	(1.264)	(1.679)	(1.659)	(1.634)	(9.438)
Born 1945 − 1949	0.84	2.01	-0.04	-0.87	1.70
	(1.133)	(1.511)	(1.489)	(1.454)	(8.210)
Born 1950 − 1954	0.75	1.76	-0.03	-0.75	1.51
	(0.994)	(1.319)	(1.304)	(1.271)	(7.320)
Born 1955 − 1959	0.67	1.55	0.02	-0.60	1.31
	(0.851)	(1.134)	(1.114)	(1.087)	(6.089)
Born 1960 − 1964	0.56	1.29	0.02	-0.47	1.08
	(0.711)	(0.947)	(0.926)	(0.897)	(4.997)
Born 1965 − 1969	0.44	1.03	-0.00	-0.41	0.88
	(0.574)	(0.767)	(0.755)	(0.731)	(4.162)
Born 1970 − 1974	0.38	0.85	0.00	-0.32	0.72
	(0.444)	(0.605)	(0.601)	(0.572)	(3.304)
Born 1975 − 1979	0.25	0.57	-0.00	-0.21	0.48
	(0.302)	(0.406)	(0.404)	(0.379)	(2.238)
Born 1980 − 1984	0.13	0.29	0.01	-0.10	0.23
	(0.161)	(0.211)	(0.204)	(0.199)	(1.054)
$log(BMI_t)$	-0.22				
	(0.495)				
$log(\%\ overweight_t)$		-0.33			
		(0.265)			
$log(\%\ obese1_t)$			0.13		
			(0.210)		
$log(\%\ obese2_t)$				0.25	
				(0.170)	
$log(\%\ obese3_t)$					-0.17
					(1.290)
$log(fat_t)$	0.50***	0.49***	0.49***	0.48***	0.50***
	(0.068)	(0.067)	(0.068)	(0.068)	(0.070)
Constant	2.62	3.17**	1.22	2.63**	0.88
	(2.283)	(1.440)	(1.040)	(1.011)	(6.051)
Observations	195	195	195	195	195
R^2	0.736	0.738	0.736	0.739	0.736

Standard errors in parentheses
$***p < 0.01, **p < 0.05, *p < 0.1$

Table D.6: Demand for Meat: Men

VARIABLES	(1) $log(BMI_t)$	(2) $log(\%\ overweight_t)$	(3) $log(\%\ obese1_t)$	(4) $log(\%\ obese2_t)$	(5) $log(\%\ obese3_t)$

Continued on next page

age	-0.07**	-0.10**	0.01	-0.19**	-0.03
	(0.034)	(0.039)	(0.042)	(0.092)	(0.031)
age^2	-0.00	0.00	-0.00***	0.00	-0.00***
	(0.000)	(0.000)	(0.001)	(0.002)	(0.000)
$married$	-0.36***	-0.63***	-0.87***	2.07	-1.02***
	(0.098)	(0.213)	(0.253)	(1.525)	(0.334)
$log(income)$	-0.13	-0.14	0.77**	-1.42	0.27
	(0.127)	(0.106)	(0.363)	(0.880)	(0.162)
2004	0.47**	0.51***	1.19***	-0.29	0.56***
	(0.194)	(0.192)	(0.360)	(0.513)	(0.205)
2006	0.53***	0.58***	1.33***	-0.23	0.67***
	(0.194)	(0.191)	(0.385)	(0.520)	(0.209)
2008	0.62**	0.69***	1.58***	-0.59	0.78***
	(0.243)	(0.241)	(0.469)	(0.800)	(0.264)
2010	0.79***	0.89***	1.97***	-0.65	0.92***
	(0.303)	(0.301)	(0.575)	(0.964)	(0.316)
$Born$ 1920 − 1924	4.81***	4.94***	12.52***	-10.38	8.80***
	(1.790)	(1.778)	(3.749)	(9.749)	(2.759)
$Born$ 1925 − 1929	4.46***	4.57***	11.65***	-9.21	8.06***
	(1.654)	(1.641)	(3.478)	(8.805)	(2.506)
$Born$ 1930 − 1934	4.08***	4.24***	10.62***	-8.02	7.22***
	(1.535)	(1.521)	(3.173)	(7.796)	(2.255)
$Born$ 1935 − 1939	3.77***	3.99***	9.45***	-6.59	6.21***
	(1.394)	(1.385)	(2.788)	(6.680)	(1.924)
$Born$ 1940 − 1944	3.43***	3.64***	8.47***	-5.97	5.70***
	(1.256)	(1.249)	(2.484)	(6.059)	(1.755)
$Born$ 1945 − 1949	3.07***	3.29***	7.34***	-4.83	4.99***
	(1.123)	(1.121)	(2.141)	(5.088)	(1.544)
$Born$ 1950 − 1954	2.70***	2.90***	6.09***	-3.53	4.21***
	(0.981)	(0.982)	(1.756)	(4.010)	(1.305)
$Born$ 1955 − 1959	2.42***	2.62***	5.10***	-2.70	3.58***
	(0.845)	(0.849)	(1.427)	(3.299)	(1.087)
$Born$ 1960 − 1964	2.03***	2.23***	4.27***	-1.86	2.95***
	(0.708)	(0.713)	(1.189)	(2.522)	(0.892)
$Born$ 1965 − 1969	1.68***	1.79***	3.83***	-2.77	2.74***
	(0.569)	(0.568)	(1.078)	(2.859)	(0.808)
$Born$ 1970 − 1974	1.24***	1.33***	2.95***	-2.25	2.27***
	(0.433)	(0.427)	(0.833)	(2.265)	(0.675)
$Born$ 1975 − 1979	0.85***	0.87***	2.15***	-1.91	1.63***
	(0.293)	(0.290)	(0.627)	(1.775)	(0.492)
$Born$ 1980 − 1984	0.49***	0.51***	1.27***	-0.98	0.83***
	(0.165)	(0.161)	(0.362)	(0.956)	(0.235)
$log(meat_{t-1})$	0.02	0.03	0.03	0.04	0.03
	(0.075)	(0.073)	(0.072)	(0.073)	(0.073)
$log(BMI_t)$	1.28				
	(1.393)				
$log(\%\ overweight_t)$		0.87			
		(0.540)			
$log(\%\ obese1_t)$			-2.01**		
			(0.872)		
$log(\%\ obese2_t)$				3.30	
				(2.120)	
$log(\%\ obese3_t)$					-0.93**
					(0.417)
	1.93	7.23***	-8.17	28.69*	-0.99
$Constant$	(3.268)	(1.763)	(5.717)	(15.359)	(2.711)
$Observations$	195	195	195	194	193
R^2	0.731	0.734	0.738	0.730	0.736

Standard errors in parentheses
$***p < 0.01, **p < 0.05, *p < 0.1$

Table D.7: Demand for Added Sugar: Men

VARIABLES	(6) $log(BMI_t)$	(7) $log(\%\ overweight_t)$	(8) $log(\%\ obese1_t)$	(9) $log(\%\ obese2_t)$	(10) $log(\%\ obese3_t)$
age	-0.07*	-0.04	0.04	-0.24**	-0.02
	(0.040)	(0.046)	(0.049)	(0.107)	(0.036)
age^2	0.00***	0.00	-0.00*	0.01**	-0.00
	(0.000)	(0.000)	(0.001)	(0.002)	(0.000)
$married$	-0.15	-0.14	-0.62**	3.68**	-0.86**
	(0.108)	(0.250)	(0.294)	(1.782)	(0.394)
$log(income)$	0.25*	0.45***	1.32***	-1.67	0.83***
	(0.148)	(0.131)	(0.424)	(1.030)	(0.194)
2004	-1.39***	-1.33***	-0.61	-2.53***	-1.18***
	(0.224)	(0.226)	(0.421)	(0.600)	(0.242)
2006	-1.04***	-0.95***	-0.16	-2.16***	-0.76***
	(0.236)	(0.238)	(0.459)	(0.606)	(0.260)
2008	-1.07***	-0.98***	-0.02	-2.88***	-0.75**
	(0.289)	(0.294)	(0.557)	(0.931)	(0.322)
2010	-1.06***	-0.95***	0.22	-3.24***	-0.75*
	(0.355)	(0.361)	(0.679)	(1.122)	(0.380)
$Born\ 1920-1924$	0.44	0.65	8.55*	-23.04**	5.80*
	(2.061)	(2.086)	(4.382)	(11.380)	(3.259)
$Born\ 1925-1929$	0.32	0.57	7.92*	-20.79**	5.21*
	(1.905)	(1.927)	(4.065)	(10.279)	(2.960)
$Born\ 1930-1934$	0.28	0.56	7.23*	-18.34**	4.61*
	(1.767)	(1.786)	(3.709)	(9.101)	(2.663)
$Born\ 1935-1939$	0.38	0.60	6.37*	-15.59**	3.81*
	(1.605)	(1.626)	(3.259)	(7.799)	(2.273)
$Born\ 1940-1944$	0.32	0.53	5.65*	-14.16**	3.50*
	(1.447)	(1.466)	(2.903)	(7.074)	(2.073)
$Born\ 1945-1949$	0.37	0.52	4.85*	-11.81**	3.05*
	(1.295)	(1.316)	(2.503)	(5.941)	(1.823)
$Born\ 1950-1954$	0.45	0.53	3.98*	-9.16*	2.57*
	(1.131)	(1.153)	(2.053)	(4.683)	(1.542)
$Born\ 1955-1959$	0.41	0.46	3.18*	-7.50*	2.07
	(0.975)	(0.997)	(1.669)	(3.853)	(1.284)
$Born\ 1960-1964$	0.35	0.43	2.68*	-5.63*	1.69
	(0.818)	(0.838)	(1.391)	(2.946)	(1.054)
$Born\ 1965-1969$	0.25	0.33	2.51**	-6.62**	1.72*
	(0.657)	(0.668)	(1.260)	(3.339)	(0.955)
$Born\ 1970-1974$	0.12	0.26	1.96**	-5.25**	1.54*
	(0.499)	(0.502)	(0.974)	(2.646)	(0.798)
$Born\ 1975-1979$	0.14	0.20	1.53**	-4.12**	1.18**
	(0.338)	(0.341)	(0.732)	(2.073)	(0.582)
$Born\ 1980-1984$	0.04	0.10	0.88**	-2.22**	0.53*
	(0.190)	(0.189)	(0.423)	(1.116)	(0.278)
$log(BMI_t)$	3.41**				
	(1.568)				
$log(\%\ overweight_t)$		0.23			
		(0.641)			
$log(\%\ obese1_t)$			-2.08**		
			(1.018)		
$log(\%\ obese2_t)$				5.18**	
				(2.475)	
$log(\%\ obese3_t)$					-1.07**
					(0.493)
$log(addedsugar_{t-1})$	0.29***	0.30***	0.30***	0.30***	0.29***
	(0.069)	(0.071)	(0.069)	(0.069)	(0.070)
$Constant$	-8.57**	-0.18	-14.31**	36.66**	-7.25**
	(3.721)	(2.059)	(6.692)	(17.907)	(3.204)

Continued on next page

Table D.7 – continued from previous page

Observations	195	195	195	194	193
R^2	0.952	0.950	0.951	0.952	0.952

Standard errors in parentheses
$***p < 0.01, **p < 0.05, *p < 0.1$

Table D.8: Demand for Fat: Men

VARIABLES	(11) $log(BMI_t)$	(12) $log(\%\ overweight_t)$	(13) $log(\%\ obese1_t)$	(14) $log(\%\ obese2_t)$	(15) $log(\%\ obese3_t)$
age	-0.04	-0.00	0.02	-0.04	-0.01
	(0.030)	(0.035)	(0.038)	(0.083)	(0.028)
age^2	0.00	-0.00	-0.00	0.00	0.00
	(0.000)	(0.000)	(0.001)	(0.002)	(0.000)
married	-0.14	0.01	-0.36	0.39	0.06
	(0.084)	(0.190)	(0.226)	(1.375)	(0.304)
$log(income)$	0.19*	0.34***	0.72**	0.04	0.25
	(0.114)	(0.103)	(0.329)	(0.792)	(0.153)
2004	0.07	0.10	0.46	-0.07	0.03
	(0.172)	(0.172)	(0.324)	(0.462)	(0.186)
2006	0.08	0.12	0.52	-0.04	0.05
	(0.171)	(0.171)	(0.346)	(0.468)	(0.190)
2008	0.05	0.08	0.56	-0.17	-0.00
	(0.215)	(0.216)	(0.422)	(0.721)	(0.241)
2010	0.08	0.13	0.72	-0.17	0.04
	(0.268)	(0.269)	(0.517)	(0.868)	(0.288)
Born 1920 − 1924	1.47	1.59	5.46	-1.52	0.38
	(1.582)	(1.590)	(3.371)	(8.783)	(2.515)
Born 1925 − 1929	1.25	1.41	4.99	-1.40	0.31
	(1.463)	(1.468)	(3.127)	(7.932)	(2.284)
Born 1930 − 1934	1.13	1.28	4.55	-1.21	0.30
	(1.357)	(1.361)	(2.853)	(7.024)	(2.056)
Born 1935 − 1939	1.08	1.16	4.03	-0.95	0.37
	(1.232)	(1.240)	(2.507)	(6.018)	(1.754)
Born 1940 − 1944	0.94	1.01	3.56	-0.89	0.29
	(1.111)	(1.118)	(2.233)	(5.458)	(1.600)
Born 1945 − 1949	0.85	0.87	3.05	-0.72	0.26
	(0.994)	(1.004)	(1.925)	(4.584)	(1.407)
Born 1950 − 1954	0.80	0.77	2.53	-0.47	0.27
	(0.868)	(0.879)	(1.579)	(3.612)	(1.189)
Born 1955 − 1959	0.72	0.67	2.08	-0.34	0.28
	(0.748)	(0.760)	(1.284)	(2.972)	(0.991)
Born 1960 − 1964	0.57	0.54	1.72	-0.22	0.23
	(0.627)	(0.639)	(1.070)	(2.272)	(0.813)
Born 1965 − 1969	0.47	0.49	1.59	-0.40	0.16
	(0.504)	(0.509)	(0.969)	(2.576)	(0.737)
Born 1970 − 1974	0.29	0.36	1.21	-0.35	0.07
	(0.383)	(0.382)	(0.749)	(2.041)	(0.616)
Born 1975 − 1979	0.22	0.25	0.90	-0.31	0.03
	(0.259)	(0.260)	(0.563)	(1.599)	(0.449)
Born 1980 − 1984	0.12	0.17	0.54*	-0.14	0.06
	(0.146)	(0.144)	(0.325)	(0.861)	(0.215)
$log(BMI_t)$	1.91				
	(1.224)				
$log(\%\ overweight_t)$		-0.25			
		(0.486)			
$log(\%\ obese1_t)$			-1.01		
			(0.783)		

Continued on next page

Table D.8 – continued from previous page

$log(\%\ obese2_t)$				0.64	
				(1.911)	
$log(\%\ obese3_t)$					0.17
					(0.381)
$log(fat_{t-1})$	0.34***	0.36***	0.36***	0.36***	0.36***
	(0.069)	(0.069)	(0.068)	(0.069)	(0.070)
$Constant$	-4.58	-0.92	-6.79	4.45	0.81
	(2.915)	(1.579)	(5.147)	(13.825)	(2.475)
$Observations$	195	195	195	194	193
R^2	0.797	0.795	0.796	0.791	0.789

Standard errors in parentheses
$***p < 0.01, **p < 0.05, *p < 0.1$