This product is part of the Pardee RAND Graduate School (PRGS) dissertation series. PRGS dissertations are produced by graduate fellows of the Pardee RAND Graduate School, the world's leading producer of Ph.D.'s in policy analysis. The dissertation has been supervised, reviewed, and approved by the graduate fellow's faculty committee.

DISSERTATION

# Economic Determinants of Obesity Among Older Americans

Yuhui Zheng

This document was submitted as a dissertation in September 2008 in partial fulfillment of the requirements of the doctoral degree in public policy analysis at the Pardee RAND Graduate School. The faculty committee that supervised and approved the dissertation consisted of Dana P. Goldman (Chair), Darius N. Lakdawalla, and Pierre-Carl Michaud.

RAND PARDEE RAND GRADUATE SCHOOL

The Pardee RAND Graduate School dissertation series reproduces dissertations that have been approved by the student's dissertation committee.

The RAND Corporation is a nonprofit research organization providing objective analysis and effective solutions that address the challenges facing the public and private sectors around the world. RAND's publications do not necessarily reflect the opinions of its research clients and sponsors.





Published 2008 by the RAND Corporation
1776 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138
1200 South Hayes Street, Arlington, VA 22202-5050
4570 Fifth Avenue, Suite 600, Pittsburgh, PA 15213
RAND URL: http://www.rand.org
To order RAND documents or to obtain additional information, contact
Distribution Services: Telephone: (310) 451-7002;
Fax: (310) 451-6915; Email: order@rand.org

# Abstract

During the past thirty years the prevalence of obesity among the U.S. population has doubled. Obesity, characterized by excess body fat, is associated with worse health, higher medical costs, and lost productivity. A series of studies have explored social economic factors behind the obesity epidemic: It is likely a by-product of technological changes that have made food cheaper, more accessible, while physical activity more expensive. Longitudinal data shows that nowadays individuals continue gaining weight until their sixties, and obesity rate among the near-elderly population is the highest. Therefore identifying effective interventions to lower obesity rate among this population has the potential of generating immediate health and financial benefits. Government agencies, employers, and health professionals are actively involved in preventing and treating obesity. However, empirical evidence about effectiveness of various interventions is inadequate. In this thesis I address two issues requiring further investigation: one is the long-term effect of physical activity on body weight, the other is the effect of food price on body weight. The study population is Americans between age 50 and age 73.

In the first part I examine the effect on body weight of retiring from strenuous and sedentary jobs, and the attendant effects on health. Retirement frees up time for exercise, but at the mean time reduces on-the-job physical activity. As a result retirees from strenuous jobs experience a larger drop in total physical activity, relative to retirees from sedentary jobs. Retirees from strenuous jobs gain 0.6 to 0.7 units of BMI, and exhibit declines in exercise participation. In contrast, retirees from sedentary jobs gain just 0.1 units of BMI over an 8-year period. Most significantly, retirement seems to cause an increase of about 1 percentage point in the incidence of diabetes among strenuous job-leavers, but no such increase among their counterparts in non-strenuous jobs. There is also evidence of relative increases in hypertension, although in this case it is harder to rule out non-causal mechanisms.

In the second part of the dissertation I analyze the effects of various food prices on body weight among older Americans. Using a longitudinal dataset and temporal and regional price data, I found evidence that price of calories, price of at-home food, and price of margarine are negatively associated with body mass index (BMI), while fast food price and price of produce are not significantly associated with BMI. Gasoline price is also negatively associated with BMI but the mechanisms are unclear. The negative association between price of margarine and BMI is the most robust result. Some of my results are in line with findings from other studies, but discrepancies remain to explore.

# Acknowledgement

I would like to thank my committee members, Dana Goldman, Darius Lakdawalla, and Pierre-Carl Michaud, for their guidance and encouragement. Without their support, my dissertation is not possible. I also want to thank them for leading me to the area of health economics and teaching me on how to do research. My gratitude extends to other RAND researchers, PRGS faculty and students who have helped me during my Ph.D. study, and the names are too many to list. I'd like to acknowledge funding from the National Institute on Aging. Finally, I want to thank my family, for always being there with me.

## Table of Contents

## List of Figures



## List of Appendix Figures

## List of Tables

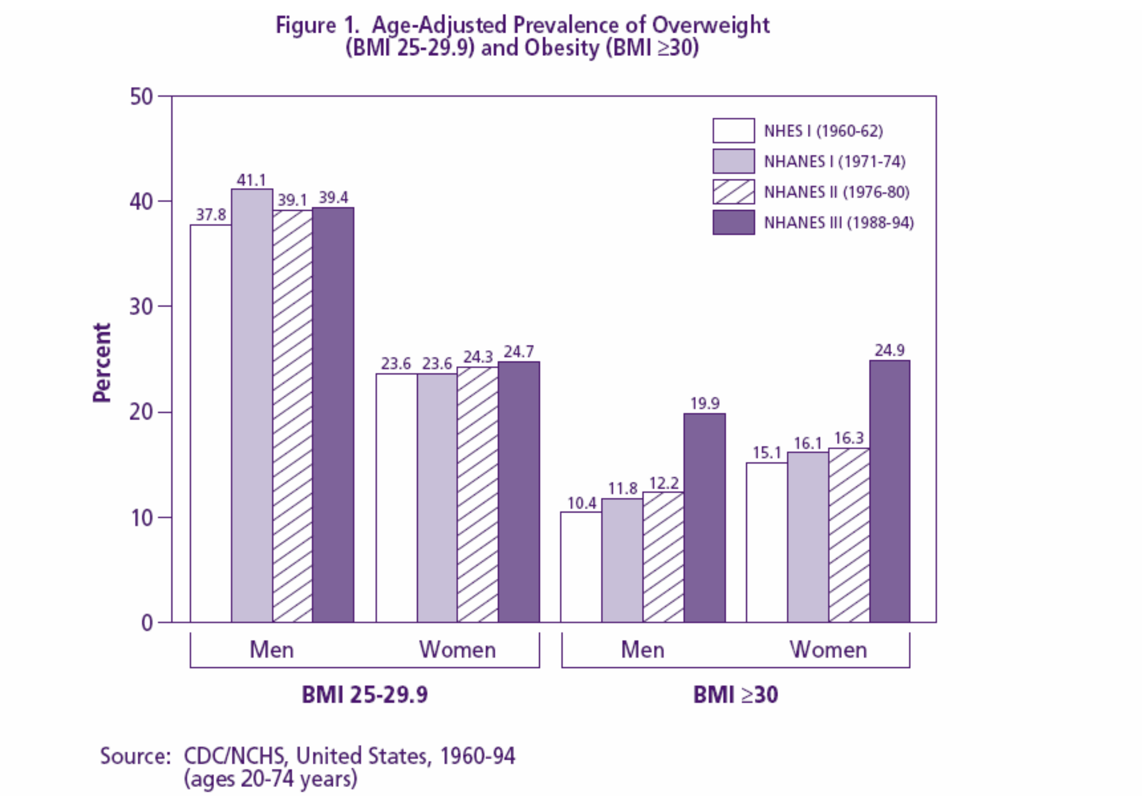
**List of Appendix Tables**

# Chapter 1 Introduction

## *Trends in obesity*

The prevalence of obesity in the United States has been rising rapidly since 1980s. As Figure 1-1 shows, from year 1960 to 1980, the prevalence of overweight and obesity among American adults only increased slightly. But from 1980 to 1994, male obesity rate nearly doubled, while female obesity rate increased from 15% to 25%. During year 2005-2006, more than one third of the adult American population (aged 20 and over) is obese (Ogden , Carroll  et al. 2007). More Americans are obese than smoke, use illegal drugs, or suffer from ailments unrelated to obesity (Wolf and Colditz 1998; Tuomilehto, Lindstrom et al. 2001).

**Figure 1-1. Age-adjusted Prevalence of Overweight (BMI 25-29.9) and Obesity (BMI≥30)**




Source: NIH 1998

Obesity is characterized by excess adipose tissue. The most accurate measures are to weigh a person underwater or in a chamber that uses air displacement to measure body volume, or to use an X-ray test called Dual Energy X-ray Absorptiometry, also known as DEXA (NIDDK 2006). But these methods are costly to implement in epidemiological and social science surveys. Thus, large-scale population surveys often approximate body fat using measures of body weight adjusted for stature (Gray and Fujioka 1991). The most widely used index is body mass index (BMI), defined as weight in kilograms divided by height in meters squared ($kg/m^2$). Despite not being able to distinguish between lean body mass and fat, a BMI exceeding $30kg/m^2$ is widely recognized as a definition of obesity, and it has the advantage of correlating strongly with body fat and weakly with height for both men and women (Revicki and Israel 1986). Other three approximate measures of body fatness include waist circumference, skinfold thickness, and bioimpedance (Kopelman 2000). But each has its own drawbacks and usually cannot be self-reported.

Throughout this thesis, unless otherwise specified, "overweight" and "obese" are measured by BMI, based on measured or self-reported weight and height. "Overweight" could indicate either $BMI \geq 25kg/m^2$, or $25 \leq BMI < 30kg/m^2$. "Obese" refers to $BMI \geq 30kg/m^2$.

### *Health consequences of obesity*

Numerous studies have found that obesity significantly increases the risk of hypertension, dyslipidemia, type 2 diabetes, coronary heart disease, stroke, gallbladder disease, osteoarthritis, respiratory disorders including sleep apnea and asthma, and some types of cancer (National Institutes of Health, 1998). Overweight accounts for about 400,000 deaths per year, second only to tobacco (Mokdad 2000).

Obesity, especially excessive abdominal fat, is associated with insulin resistance and glucose intolerance through elevated release of free fatty acids (FFAs) (Kopelman 2000). Therefore obesity is a significant risk factor of type II diabetes. In the Nurses Health study – a prospective cohort study of females ages 30 to 55 years old, Colditz et. al. (1995) found that gaining only 7 to 10.9 kg after the age of 18 corresponds to a twofold increase in the risk for diabetes mellitus (Colditz, Willett et al. 1995). Moreover, women with a body mass index (BMI) between 23-25 $kg/m^2$ are 4 times more likely to develop diabetes than women with a BMI less than 22 $kg/m^2$. Another prospective study of males aged 40 to 75 years also found a strong positive association between overall obesity and risk of diabetes (Chan, Rimm et al. 1994). Men with a BMI of > or = 35 $kg/m^2$ had relative risk of 42.1 (95% confidence interval [CI] 22.0-80.6) compared with men with a BMI < 23.0 $kg/m^2$. BMI at age 21 and absolute weight gain throughout adulthood were also significant independent risk factors for diabetes. In a third longitudinal study with a national cohort of 8,545 US adults (Ford, Williamson et al. 1997), adults who gained 20 kg or more during a 10-year period had relative risk of 3.85 (95% CI 2.04-7.22) compared with participants whose weights remained relatively stable.

Multiple cross-sectional studies have found the positive association between BMI and hypertension (Stamler, Stamler et al. 1978; Havlik, Hubert et al. 1983; MacMahon, Blacket et al. 1984; Dyer and Elliott 1989; Brown, Higgins et al. 2000).Using nationally representative data from NHANES III, Brown et. al. (2000) find that prevalence of high blood pressure (where high blood pressure is defined as mean systolic BP ≥ 140mm Hg or mean diastolic BP ≥ 90mm Hg) among adult men with BMI<25 was 15% compared to 42% for men with BMI ≥ 30. For women, the prevalence of high blood pressure for the two ranges of BMI was 15% and 38%, respectively. The effect of BMI on blood pressure was greatest for men and women ages 20 to 59 (Brown, Higgins et al. 2000). In a prospective study of U.S. female health professionals 30-55 years of age, Huang et. al. (1998) found that BMI at 18 years of age and midlife were positively associated with occurrence of hypertension. A weight gain of 5.0 to 9.9 kg since age 18 had a relative risk of 1.74, compared with women with weight gain 0.0 to 2.0kg. Weight loss since age 18 was related to significant lower risk of hypertension. And the benefit of weight loss was greater among women in the top tertile of BMI at age 18 years (Huang, Willett et al. 1998). Another prospective study – Framingham Heart Study – also found that weight gain was associated with increased risk of hypertension, while weight loss was associated with lower risk of hypertension(Higgins, Kannel et al. 1988).

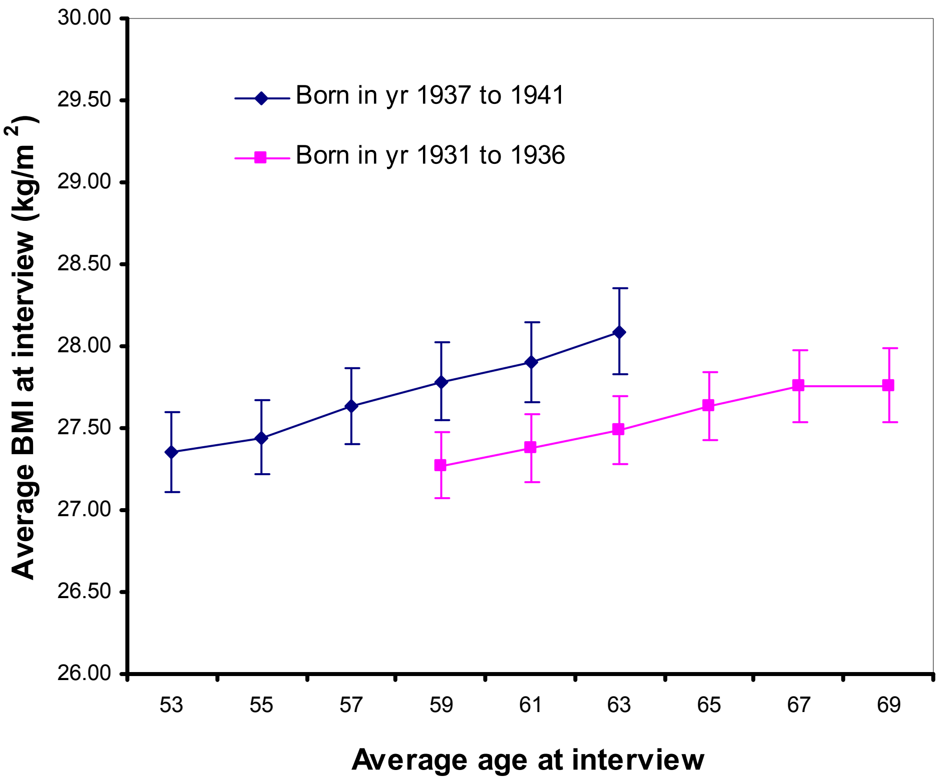
Obesity is directly related to cardiovascular risk factors, including high levels of total cholesterol, LDL cholesterol, triglycerides, blood pressure, fibrinogen and insulin, and low levels of HDL-cholesterol (NIH 1998). Using data from the Nurses Health Study, Willett et. al. (1995) identified obesity as a major contributor to coronary heart disease (CHD). After controlling for age, smoking, menopausal status, hormone use, and parental history of CHD, they found that women with BMIs from 25 to 28.9 faced risks for CHD that were two times higher than women with BMI<21. The relative risk was three times as high for women with a BMI ≥ 29 (Willett, Manson et al. 1995). BM≥ 25 is also found to be a risk factor of coronary heart disease among US men aged 40–75 years (Rimm, Stampfer et al. 1995). The association was much stronger among men younger than 65, compared with that among elderly men. Another longitudinal study evaluated risk of late life coronary heart disease associated with being overweight in late middle or old age (Harris, Launer et al. 1997). Among older men and women aged 70-86 and free of coronary heart disease, being overweight 10 years ago was associated with an increased risk of coronary heart disease in the next four years, after controlling for sex, age, and cigarette smoking. Current BMI of 27 or more was also associated with increased risk once excluding those with weight loss of 10% or more.

In a 26-year follow up study of the original Framingham cohort (Hubert, Feinleib et al. 1983), obesity, measured by Metropolitan Relative Weight, was found to be a significant independent predictor of incidences of coronary disease and congestive heart failure, and of coronary death, after controlling for age, cholesterol, systolic blood pressure, cigarettes, left ventricular hypertrophy and glucose intolerance.

### *Obesity among older Americans*

Obesity among older Americans is especially alarming, given its immediate impact. Cross-sectional data in year 2001-2004 shows that the prevalence of obesity peaks at age group 55-64, with the rate of 36.0% in male and 39.0% in female (NCHS 2007). Analysis of the longitudinal data of Health and Retirement Study (Figure 1-2) also demonstrates that among American males 50 and over, younger birth cohorts are heavier than older birth cohorts, and people continued gaining weight until their late sixties, and might also be gaining weight more rapidly.

**Figure 1-2. Trajectory of average body weight index for American males by two birth cohorts**




Note: Vertical bars show 1.96*standard errors of the sample mean.
Data source: Health and Retirement Study 1992-2004 and author's analysis

Evaluating health consequences of obesity among the elderly is more difficult. Studies find that underweight elderly individuals have higher mortality rates, while there is no consensus on mortality differentials between the normal and the overweight or obese. Studying the mortality risks of overweight or obesity among older population suffers from two methodological problems. First is "reverse causation", meaning that those who have lower weights suffer from illness which is either unobserved, or not adequately considered in analysis. And such illness leads to faster death. Second is the inadequate control of confounding factors, most importantly smoking status and smoking history, which could distort the relationship between body mass status and mortality. But missing confounding factors, like SES, could lead to an over-estimation of the mortality hazard. A closer examination of the studies reveals some systematic pattern. Studies that have a dataset with large sample size allow for regression analysis among only healthy, non-smokers. And these studies found that both overweight and obesity are associated with higher mortality rates across all age groups (up to age 75) (Stevens 1998, Calle 1999, Adams 2006). Studies using datasets with small sample sizes usually didn't find that overweight is associated with higher mortality. But these studies might lack enough power to detect the effect (Visscher 2000, Flegal 2007). One study examined mortality by BMI status among 70 years and older (Grabowski 2001), they found that thin people (BMI < 19.4) were more likely to die than normal-weighted people, while over-weight (BMI > 28.5) were less likely to die than normal-weighted. This article didn't control for smoking status. Although it controls for comorbidities, it is unknown whether the Cox proportional hazard assumption holds.

As for morbidity, studies have repeatedly shown that obesity is associated with increased cardiovascular risk factors, function limitations, and morbidity impairments in older adults (National Institute of Health, 1998, Willett 1999).

## *Economic costs of obesity*

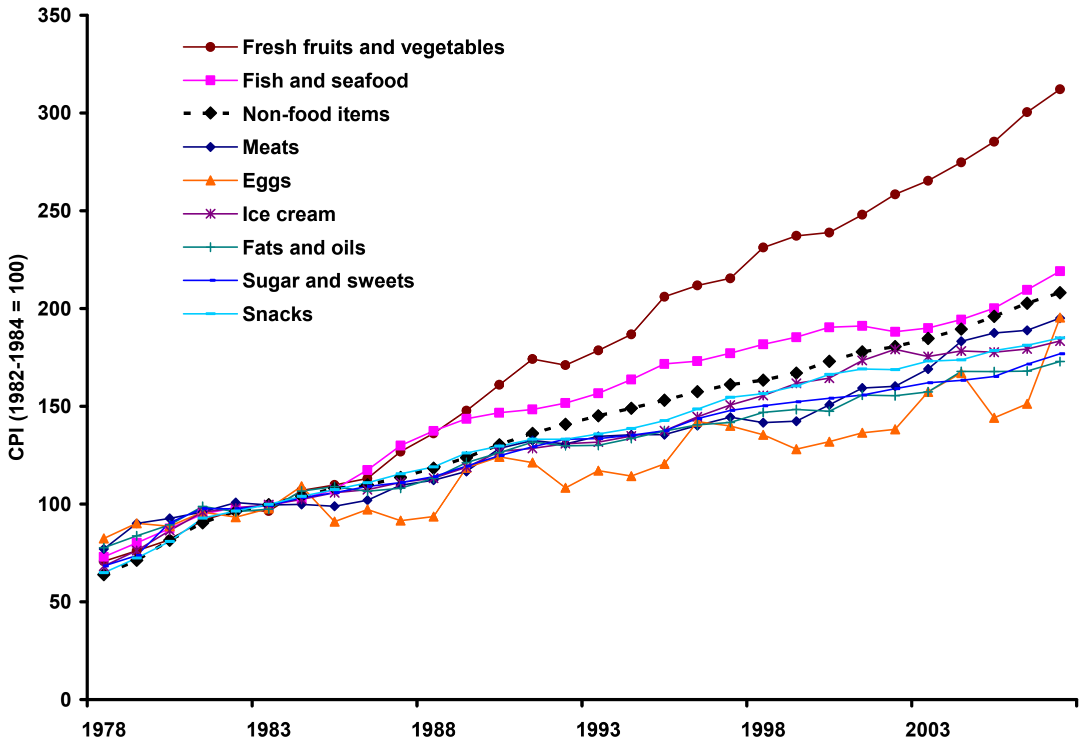
According to a study of national costs attributed to both overweight (BMI 25–29.9) and obesity (BMI greater than 30), medical expenses accounted for 9.1 percent of total U.S. medical expenditures in 1998 and may have reached as high as $78.5 billion ($92.6 billion in 2002 dollars) (Finkelstein, Fiebelkorn et al. 2003). Approximately half of these costs were paid by Medicaid and Medicare. In addition to higher medical costs, obesity is also associated with lost productivity (Wolf and Colditz 1998), Studies also examined the effect of obesity on labor market outcomes but the results are mixed. One study examined the associations of body weight with family income among men and women. They found negative association among women, but not among men. The lower family income among obese women was mainly due to the disadvantage in the marriage market, not the labor market (Averett and Korenman 1996). Another study found that obesity is associated with lower wages among white female, but not among other demographic groups (Cawley 2004). Using genetic information as instrumental variables for BMI, One study found no effect of obesity on either employment or wages, among men and women in their mid twenties (Norton and Han 2007).

## *Economic causes of rising obesity*

The rapid increase of obesity in the last three decades implies that either energy intake increased, energy expenditure decreased, or both. Genetic factors contribute to obesity but they are unlikely to undergo sharp changes in a short period of time. So it is very likely that the changing social economic factors have caused this energy imbalance. Lakdawalla and Philipson (2002) argues that technological change has induced weight growth by making home- and market production more sedentary and by lowering food prices through agricultural innovation (Lakdawalla and Philipson 2002). Their findings show that about forty percent of the recent growth in weight seems to be due to agricultural innovation that has lowered food prices, while sixty percent may be due to demand factors such as declining physical activity from technological changes in home and market production. Other studies discussed the relative contribution of the changes in energy expenditure and energy intake in the United States. Culter (2003) argues that it is the increased calorie intake, and not the calorie consumption, that has made Americans more obese in the past several decades. The mass production of food and improved food storage has made food cheaper and there are more varieties. They also found that the increased calorie intake is largely a result of consuming more meals rather than more Calories per meal. And these additional meals are often snacks (Cutler, Glaeser et al. 2003). Four other studies also found increased Calorie consumptions (Finkelstein, Ruhm et al. 2005).

Food price has decreased slightly overtime relative to non-food prices. From year 1978 the annual rate of increase was 3.7% for food items, while 4.3% for non-food items, the difference is small: 0.6% per year. But if we look into specific foods, we would find that the changes of food prices overtime were not uniform across food groups. Prices of processed foods or energy-dense foods – snacks, sweets – has decreased relative to non-food items, while the prices of low-calorie and nutritious foods – fresh vegetables and fish – has increased more rapidly than non-food items (Figure 1-3). As a result, processed foods or energy-dense foods become increasingly cheaper. In addition, these foods usually require little time to prepare – making them even more cost-attractive. Drewnowski and Specter (2004) found that there is an inverse relationship between energy density and energy cost, such that energy-dense foods represent the cheapest options to the consumer. Low-income groups are more likely to choose energy-dense food, have low vegetable consumption, and poor quality diets.

**Figure 1-3. Consumer price index for varieties of foods, and for non-food items**




Data source: Bureau of Labor Statistics.

Other explanations of the rising obesity apply to certain demographic groups. Studies have found that rising female employment might have contributed to female obesity, as well as child obesity (Anderson, Butcher et al. 2003; Cutler, Glaeser et al. 2003). Another explanation is the decrease in smoking. But the results are mixed (Chou, Grossman et al. 2004; Gruber and Frakes 2006).

Other causes of obesity put forward – food availability, built environment, urban sprawl (Lopez 2004; Booth, Pinkston et al. 2005; Powell, Slater et al. 2007)– consist of alternative ways of saying that the social economic environment has lowered the price of energy intake while raising the price of energy expenditure.

A last phenomenon related to the obesity epidemic is how food companies have used advertisements to affect people's – especially children' and adolescents' - preference to high-calorie non-nutritious food. American children are exposed to approximately 40,000 food advertisements annually, majority of which are candy, cereal, and fast food. Reviews found that these advertisements shape children's product preferences and eating habits. (Mello, Studdert et al. 2006)

In conclusion, researchers pointed out that the rising obesity in the past three decades is the unintended consequence of technology improvement – to make lifestyle more sedentary, and to make food cheaper and widely available. And food companies have deliberately channeled children and adolescents' preference toward energy-dense, and non-nutritious food.

## *Social economic policies to obesity prevention and intervention*

Although diet is a matter of personal choice, there are several rationales for the government to intervene. The major reason is is the "negative externality"related to obesity. Obesity is related to worse health status, and more health care utilization. In the United States, Increases in the proportion of and spending on obese people relative to people of normal weight account for 27 percent of the rise in inflation-adjusted per capita spending between 1987 and 2001(Thorpe, Florence et al. 2004). It is estimated that obesity contributes to $75 billion (in 2003 dollars) of medical spending(Finkelstein, Fiebelkorn et al. 2003). Using micro-simulation methods, three studies both found that obese older population has higher lifetime medical costs than non-obese older individuals. One study shows that obese seventy-year-olds will live about as long as those of normal weight, but will spend more than $39,000 more on health care(Lakdawalla, Goldman et al. 2005). The other study shows that elderly men who were overweight or obese at age 65 had 6–13 percent more lifetime health care expenditures than the same age cohort within normal weight range at age 65. Elderly women who were overweight or obese at age 65 spent 11–17 percent more than those in a normal weight range and those who are obese live about 1-2 years shorter(Yang and Hall 2008). A third study estimated that among those aged 55-64, higher BMI is associated with lower life expectancy and higher lifetime medical costs. Relative to those with initial BMI of

22.5kg/m$^2$, those with initial BMI of 37.5kg/m$^2$ live about 1 year shorter, and have lifetime medical care costs of $16,000 higher(Thompson, Edelsberg et al. 1999).

If an individual bear the full extra medical costs caused by obesity, then there will be no external costs. But this is not the case. In the United States both public and private health insurance programs don't adjust premium contribution according to obesity status. Therefore non-obese insured individuals subsidized obese insured individuals. In addition, high obesity rates among the poor raises costs of governmental transfer programs, like Medicaid. It is estimated that among the $75 billion obesity-related medical spending, Medicare and Medicaid paid half of them. Obesity also causes moral hazard problems in the context of health insurance – people have reduced incentives to control weight because they don't bear the full costs of excessive weight. Empirical studies about the magnitude of the moral hazard problem have generated mixed results (Newhouse and The 1993; Bhattacharya and Sood 2005; Rashad and Markowitz 2007).

Few studies have examined whether obesity is associated with the outlays of other government programs or group insurance programs, like Disability Insurance (DI), Old Age and Survivor Insurance (OASI), and group private pension and life insurance programs. An increasing disability trend has been reported among the young, which could be due to higher obesity rate (Lakdawalla, Bhattacharya et al. 2004; Sturm, Ringel et al. 2004). Unlike smoking, which has a significant effect on mortality and therefore imposes a "positive externality" on public and private pensions, the effect of obesity on life expectancy is less clear, and is not as dramatic as that of smoking.

Other market failures regarding food consumption and obesity may also justify government interventions. There is information asymmetry regarding food quality. The consumers know much less than the food manufactures about the quality of the food he or she is purchasing. So the government should reduce such information asymmetry by promoting information disclosure, like food labeling, and to prevent deceptive advertising practices. In addition, the government can encourage physical activity through better urban planning, improving public infrastructure, like parks, public transportation, and walkways.

Employers are increasing actively involved in combating obesity problems among their employees. One reason is that obese employees increase health care costs for non-obese employees, because of the pooling health insurance. The other reason is that obesity intervention might be a cost-effective way to improve productivity.

Although both the government and the employers have good rationale to prevent obesity, there remain a lot of questions about what strategies are effective. The purpose of this thesis is to examine whether economic incentives have significant impact on people's body weight, through either increasing physical activity or reducing food consumption. We focus the study population among the older population – aged 50 years and older. Since the prevalence of obesity peaks at age 60-70, and since the older population is the major health care users, reducing obesity rates among the older population could have immediate impact on the national health expenditures.

The rest of the thesis is organized as two essays. The first essay examines the relationship between retirement and weight gain. The second essay estimates the effects of food prices on weight.

# Chapter 2 Retirement and Weight[1]

## *Introduction*

Over the past thirty years, the rate of obesity in the US has more than doubled, and height-adjusted weight has increased for nearly every demographic group (Lakdawalla and Philipson 2006). Few doubt the negative implications of these trends for health. The only real debate is about the magnitude of the damage.[2]

The dramatic secular growth has put an increasing number of Americans at risk for obesity-related disease. More than half of Americans are overweight. Roughly one-quarter are obese. For all of these individuals, weight gain poses elevated risks of acquiring diseases like diabetes, hypertension, and heart disease, and weight loss is of particular benefit to health. A large body of observational studies demonstrates that health risks are higher for the overweight, and higher still for the obese(Adams, Schatzkin et al. 2006).[3] Randomized trials echo these qualitative findings: for instance, one study(Knowler, Barrett-Connor et al. 2002) show that diet and exercise regimens producing a three-year weight loss of 5-7% in overweight subjects reduce Type II diabetes risk by 58%.

The Knowler study finds an extremely large effect on health, but also induces a particularly sharp change in body weight. While weight changes of this size rarely occur so rapidly outside randomized settings, substantial changes do occur when individuals undergo significant transitions in diet or physical activity. For example, Lakdawalla and

[1] This is a co-authored paper with Dana Goldman and Darius Lakdawalla.

[2] In 2004, CDC researchers published estimates in *JAMA* indicating that poor diet and exercise were responsible for 400,000 deaths annually, suggesting that obesity would soon supplant smoking as the number one cause of death in the United States (Mokdad 2004). One year later, a different set of CDC researchers would publish estimates in the same journal locating the number of deaths at about 110,000 annually.

[3] To be sure, there is some controversy about the magnitude of health risks for the *elderly*, as some authors have found that overweight status is protective, relative to normal weight (Grabowski and Ellis 2001)

Philipson (Lakdawalla and Philipson 2007) show that entry into the labor force causes young workers in sedentary jobs to gain approximately 2.5% more body weight than their peers in strenuous jobs, over a three-year period. This raises the question of whether commonly observed life-cycle growth in weight is large enough to harm health. The secular growth in weight makes this question even more important, since more Americans find themselves closer to (or even above) the threshold of unhealthy weight.

Retirement provides a particularly salient example of a common but substantial life-cycle change in body weight. Departure from the labor force produces dramatic changes in patterns of physical activity, depending on the nature of one's job. A construction worker suddenly finds he is no longer paid to exercise, while an accountant finally finds time away from his desk to spend on recreational exercise. Moreover, as Figure 2-1 indicates, secular growth in body weight among the retirement-age population has made them ever-more vulnerable to obesity-related disease. The figure depicts the growth in Body Mass Index across five adjacent birth cohorts of elderly males in the Health and Retirement Study (HRS). In the space of just ten years, age-adjusted Body Mass Index (BMI) rose between 0.75 and 1.4 units.

Retirement may be particularly damaging to the health of retirees from strenuous jobs. This represents a departure from historical patterns in which the end of backbreaking labor may have relieved stress on the body (Case 2003). That tendency is likely to be present today as well, but – for obesity-related diseases at least – leaving a "backbreaking" strenuous job may end up posing significant risks to health.

In this paper, we explore this hypothesis by quantifying the effect on body weight of retiring from strenuous and sedentary jobs, and the attendant effects on health. Consistent with the simple incentives involved, retirement appears to function as a weight control device for workers in sedentary jobs, but a cause of significant body weight increase for others. Retirees from strenuous jobs gain 0.7 units of BMI, and exhibit declines in exercise participation. In contrast, retirees from sedentary jobs gain just 0.1 units of BMI over an 8-year period.

Most significantly, retirement seems to cause an increase in the incidence of diabetes among strenuous job-leavers, but no such increase among their counterparts. There is also evidence of relative increases in hypertension, although in this case it is harder to rule out non-causal mechanisms. To put these results in context, it is useful to note that contracting diabetes reduces life-expectancy by 4 years, disability-free life-expectancy by 5 years, and raises old-age health care costs by 55,000 dollars.

In subsequent sections we lay out the conceptual framework, describe data and methods, report the effect of retirement on weight, analyze the mechanisms for the retirement effect, and also examine the health effects of weight changes at retirement. The last part is conclusion.

## *Conceptual Framework*

Retirement has at least three effects on weight: (1) Direct reduction in job-related exercise; (2) Change in the individual's leisure-time endowment, which leads to a behavioral response of more exercise; and (3) A variety of possible income effects, but only if consumption-smoothing fails to be perfect. Each of these can vary across people in different types of occupations and at different levels of weight.

Consider an individual who works for one period and is retired for another. In addition to consumption $(C)$ and leisure $(L)$, she values three additional goods: food intake $(F)$, weight $(W)$, and exercise $(E)$. All else equal, people always prefer more food to less and more leisure to less. Utility is initially rising in weight, but eventually decreasing, and is maximized at some subjectively ideal weight level. We are agnostic about whether people enjoy exercising or prefer to avoid it.

Weight is determined by food intake and time spent exercising. The latter includes time spent exercising on the job, and at home. We take on-the-job exercise as given, so that weight can be expressed as $W(F, E_j + E_h)$, where $E_j$ is pre-determined time spent exercising at work, and $E_h$ is (chosen) time spent exercising at home. Time spent exercising reduces weight, but at a decreasing rate.

In period $t$, the individual has income $I_t$. This evolves over time according to the individual's consumption-smoothing decisions, which we do not model. Finally, the individual has time $T$ available for leisure and exercise; retired individuals have more time available than workers.[4] In each period, the individual faces a decision problem of the form:

$$\begin{aligned} &\max_{c_t, F_t, L_t, E_{ht}} U(c_t, F_t, L_t, W(F_t, E_{ht} + E_{jt}), E_{ht} + E_{jt}) \\ &s.t.\ c_t + pF_t \leq I_t \\ &\quad E_{ht} + L_t \leq T_t \end{aligned} \tag{1}$$

Optimal behavior is characterized by the following first-order conditions:

$$\begin{aligned} &U_c = \lambda_t \\ &U_F + U_W W_F = \lambda_t p \\ &U_L = \gamma_t \\ &U_W W_E + U_E = \gamma_t \end{aligned} \tag{2}$$

$\lambda$ is the marginal utility of income, and $\gamma$ is the marginal utility of time. The three effects of retirement operate on the first-order conditions as follows: the activity effect is a decrease in $E_{jt}$, which induces a first-order increase in $W$ and in $W_E$; the time-endowment effect is a decrease in $\gamma_t$, the marginal value of time; and the income effect (if permanent-income falls at retirement) corresponds to an increase in $\lambda_t$, the marginal utility of income.

Direct Effects. Job-related exercise pays the individual for activity. Therefore, when job-related exercise falls, it becomes more expensive to exercise, and less of it gets done. The individual may compensate by increasing leisure-time exercise but never by enough to increase total activity, as shown in the appendix. This decrease in total activity has two effects, also proven in the appendix: (1) The individual gains weight, because it is more expensive to be thin; and (2) He eats less, because the reduction in activity makes a

[4] We do not explicitly model health, which may affect time endowments and could reverse this result. We choose this approach, because our empirical strategy employs a variety of techniques to control for health explicitly.

given amount of food intake costlier in terms of weight. Note that food intake falls, but this is a second-order effect, not enough to offset the reduction in physical activity.

Time-Endowment Effects. Retirement intrinsically increases the amount of time an individual has for leisure and exercise. This increase in $T_t$ has the predictable effect of increasing time spent on leisure and on exercise. The effect on weight depends in part on whether the increase in time affects the marginal utility of food intake. However, as long as these cross-effects are not dominant, the increase in time endowment lowers weight by increasing physical activity.

Income Effects. If people smooth consumption perfectly (and if subjective discount factors are equal to financial discount factors), the marginal utility of income does not change at retirement, and there are no income effects. However, if subjective discount factors are smaller than financial ones, or if people fail to save optimally for any number of reasons, people would be "poorer" in retirement than in their working years. These income effects *per se* are ambiguous. Decreases in income may raise or lower food consumption, and they also reduce the time that can be spent on weight-control activities. The net effects are unclear.

Overall Effects of Retirement. Abstracting from the ambiguous income effects that may or may not operate at retirement, the simple model has clear predictions. For retirees from sedentary jobs, retirement strictly reduces weight, at least holding constant the standard biological effects of aging on weight. For these retirees, there is little if any job-related exercise to lose, but retirement boosts their time available for leisure-time exercise. On the other hand, retirement may raise or lower weight for retirees from active jobs, depending on the gain in leisure time relative to the reduction in job-related exercise. Clearly, workers who spend their entire work-day exercising will gain weight at retirement, because it will not be optimal to spend all their leisure-time exercising. Workers who spend part of their day exercising fall into an intermediate case. However, it is clear that, among those who work the same number of hours, retirement will tend to increase weight more for those in strenuous jobs.

## *Empirical Analysis*

We begin by studying the effect of retirement on body weight, by occupation. To better understand the mechanisms behind this relationship, we then turn to the study of how retirement affects exercise, food intake, hours worked, and income. Finally, we estimate the effect of retirement on the incidence of obesity-related disease.

## *Data*

We use data from the Health and Retirement Study (HRS), which is a nationally representative panel of individuals aged 51 and over.[5] The HRS contains longitudinal data on demographics, health status and health behaviors, financial and housing wealth, income, retirement plans and employment history. The original HRS birth cohort – those who were born between 1931 and 1941 – were first interviewed in 1992. Since then, data have been collected biennially. Seven waves of data are currently available. The War Babies birth cohort, born between 1942 and 1947, was first interviewed in 1998, and followed up biennially since then. Due to relatively low rates of labor force participation among females of this cohort, we focus on how retirement affects HRS males.

Our final sample for analysis consists of 3,936 males. 5,639 males born between 1931 and 1947 were interviewed in at least one wave. We drop the 1,200 males who were retired at baseline, because we cannot identify the date of retirement or how long they have been retired. For similar reasons, we dropped an additional 116 who reported in at least one of the interviews that they were neither working nor retired. An additional 184 were dropped due to missing information on longest-tenured occupation. We further dropped 5 observations with missing values for body mass index, socioeconomic status and health status. Finally, we dropped 198 individuals with only one wave of valid data. These refinements yield our analytical sample of 3,936 males.

[5] We use a publicly available version of the HRS data (RAND HRS Data file Version G) that was cleaned and processed by the RAND Corporation with support from the Social Security Administration.

We use the HRS to construct data on: body mass index, retirement status, socioeconomic status, health status, primary occupation and job strenuousness, vigorous activity, and expenditures on dining out. Details of variable construction are presented in a data appendix, but we provide the major highlights here.

Body mass index is defined as weight in kilograms divided by height in meters-squared. The HRS data on height and weight are self-reported. Since self-reported weight has been shown to be systematically biased - women under-report their weight, light men over-report, and heavy men under-report - we employ procedures developed by Cawley (2004) to correct for these errors (Cawley 2004). In the year 2006, HRS randomly selected half of the households and measured their weight and height. The self-reported weight and height is also available. As a result we can model the relationship between actual and measured weight and height. We restrict the estimation sample among those males aged 52 and 73. 52 is the lowest age for an eligible HRS respondent (the lowest age for our estimation sample is 50). While 73 is the highest age for our estimation sample. We carry out regressions by regressing actual weight on reported weight, reported weight squared. And we did the regressions separately for eight sub-groups: White male non-Hispanic, black male non-Hispanic, Hispanic male, and other males. The regression results are shown in Table B 2-1. We run regressions with and without constant. And we found that regressions without constants fit the data much better. So we will use the estimates without constant to predict weight and height for individuals in our estimation sample, based on their race/ethnicity, self-reported weight and height. And we re-calculate BMI based on corrected weight and height. All descriptive and regression analysis afterwards use BMI corrected for reporting error.

Data on retirement are taken from the "labor force status" variable and measure whether the individual currently reports being retired (as opposed to "working full-time", "working part-time", or "unemployed"). Measures of health status are all self-reported. The occurrence of particular diseases is measured based on whether the respondent reports that a doctor has diagnosed him with the condition. We also use self-reported health (whether the individual describes his overall health as excellent, very good, good, fair, or poor), and disability. Disability is measured as the number of limitations to

Activities of Daily Living (ADL's), and Instrumental Activities of Daily Living (IADL's). The HRS collects data on limitations to five ADL's: bathing or showering, dressing, eating, getting out of bed, and walking across a room. A "limitation" is the presence of any difficulty with that activity due to a health or memory problem. There are three IADL's: making phone calls, managing money, and taking medication.[6]

All these data are summarized for the HRS sample in Table 2-1. Observe first that the mean HRS male is significantly overweight, defined as being at or above a BMI of 25. Therefore, weight loss is likely to be valuable, and vice-versa. About half of the males are retired by the end of the observation window.

The table breaks down all the variables according to type of occupation. We differentiate between individuals retiring from physically demanding occupations and those retiring from sedentary occupations. To measure the physical demands of jobs, we begin with restricted-access HRS data on the 3-digit 1980 Census occupation category of the respondent, and use this to link the HRS to data from the US Department of Labor's Dictionary of Occupation Titles (DOT). The DOT is, literally, a dictionary of all occupations in the US. First published in 1939, the DOT has been updated over time. We use the 4th edition, first published in 1977, the release closest to the peak labor force attachment period of the HRS cohort. Each entry in the dictionary lists a title for the occupation, as well as a description of the occupation's skill requirements and demands. Among these listed demands are the job's physical demands, which are reported for each occupation.[7] Specifically, the DOT reports on the presence of up to 3 "physical demands" in each occupation: climbing or balancing; stooping, kneeling, crouching, or crawling; and reaching, handling, fingering, or feeling. We take the number of demands (0-3) as an index of an occupation's physical demands. Clearly, this is not a complete set of all possible physical tasks on the job, but previous research has shown it to be

[6] At baseline, the HRS asked about different IADL's: using a map, using a calculator, and using a microwave.

[7] We use England and Kilbourne's cross-walk of the DOT data to the 1980 Census 3-digit occupation coding scheme (England and Kilbourne 1989)

reasonably well-correlated with individuals' reports of how physically taxing their jobs are (Lakdawalla and Philipson 2006).

Table 2-2 lists a set of representative occupations throughout the physical demands distribution, and how we classify strenuous versus non-strenuous jobs. Note that 1980 Census occupation codes are associated with physical demands measures that are not necessarily integers. This occurs because the DOT data are based on a very fine occupational coding scheme that does not always coincide with the Census occupation coding. The measure for each Census occupation code represents a mean across all original DOT codes that are encompassed by the Census occupation(England and Kilbourne 1989). Based on this index of physical demands, we divide the HRS sample into two approximately equal halves; empirically, the cut-off turns out to be just above 1.0 physical demands.

Returning to Table 2-1, workers in strenuous jobs tend to be poorer, less educated, heavier, and sicker. This represents one of the key empirical challenges of the paper. If these differences are fixed, we can use the longitudinal structure of the data to net out the fixed differences. However, it is possible that the weight, health, and activity *trajectories* differ across types of occupation, in which case fixed-effects are insufficient for identification. To address this problem, we analyze whether or not differential trends in weight and other outcomes pre-date retirement. We find that trends are similar before retirement, but diverge thereafter. Moreover, our findings are also robust to strategies that instrument for retirement using the ages of eligibility for Social Security and Medicare.

## *The Effect of Retirement on Weight*

### Trends in the Unadjusted Data

Figure 2-2 reveals the life-cycle change in male BMI that coincides with retirement. The figure depicts unadjusted changes in BMI relative to date of retirement. It follows two halves of the HRS male population, where one half engages in more strenuous work than the other.[8] Each point on the graph corresponds to a biennial wave of the HRS. The interesting feature of this figure is the divergence in BMI that seems to occur immediately after retirement, but does not precede it. After ten years (or five waves) of retirement, men who retired from sedentary jobs are about 0.3 BMI units lighter than they were at retirement. However, after the same length of time, those retiring from strenuous jobs are more than 0.6 67 BMI units heavier. During the 8 years prior to retirement, growth in BMI is nearly identical for both groups.

The table below the figure presents mean differences and sample sizes corresponding to each wave of data. The asterisks reflect statistically significant differences between the two occupational groups. There are no significant differences in any year prior to retirement, but significant differences in the two waves after the observation of retirement. Subsequent waves show no significant difference, likely due to the declining sample sizes, which are also shown.

## Regression Analysis

Figure 2-2 provides suggestive evidence that retirement causes weight gain among strenuous workers, but weight loss among sedentary workers. We can test this interpretation more formally using the following regression model:

$$BMI_{it} = \beta_0 + \beta_1 retired_{it} + \beta_2 strenuous_i + \beta_3 retired_{it} * strenuous_i + \beta_4 Z_{it} + \beta_5 X_i + u_{it} \quad (3)$$

$BMI_{it}$ is individual $i's$ body mass index at wave $t$. The variable $retired_{it}$ is either a dummy for whether retired at wave $t$ or the number of waves since retirement at wave $t$. The variable $strenuous_i$ is a binary measure of whether individual $i$ retired from a strenuous or a non-strenuous job, while $retired_{it} * strenuous_i$ is the interaction between the measure of retirement (either the dummy variable or number of waves since retirement) and job strenuousness. $X_i$ is a vector of time-invariant characteristics for individual $i$: race, ethnicity, and education. $Z_{it}$ represents a vector of time-varying characteristics of individual $i$ at time $t$: age, age squared, log of household income (in

[8] Later, we discuss the detailed definition of “strenuous” versus “sedentary” work.

1998 dollars), a dummy for non-positive household wealth, and log of household wealth (in 1998 dollars). As sensitivity analyses of whether our effects are driven by health changes, we also include in $Z_{it}$ the following time-varying measures of health: self-rated general health, IADL limitations, ADL limitations, and physician-diagnosed illnesses (cancer, diabetes, heart diseases, hypertension, lung diseases, and stroke).

Table 2-3 presents the results of this model. The first two columns of the table display results from the OLS regression model in equation 3. The second column adds the time-varying controls for individual health. The next two columns repeat these analyses, but measure the per-wave effect of being retired, rather than a single combined effect of retirement. Finally, the last four columns repeat the OLS analysis, but with the addition of individual fixed-effects. Note that all specifications net out age-specific (quadratic) trends in weight. Both the OLS and fixed-effects results imply that strenuous-job retirees gain more weight than sedentary-job retirees.

The inclusion or exclusion of health conditions has virtually no quantitative impact on the interaction between retirement and occupation, particularly in the fixed-effects specifications. This suggests that the interaction appears not to be influenced by time-varying health observables. This is some evidence that it is similarly unrelated to time-varying health unobservables.

The differences between the OLS and fixed-effects coefficients suggest that BMI levels are higher for strenuous workers at retirement than sedentary ones.[9] Since these baseline differences are unrelated to retirement, we would like to exclude them. Accordingly, we focus on the fixed-effects results for the balance of the paper.

Focusing on the fixed-effects results implies that retirement lowers BMI by about 0.32 units for retirees from sedentary jobs, but raises it by the same amount for their counterparts retiring from strenuous jobs. This result is robust to the inclusion of controls

[9] This is consistent with the finding that strenuous jobs also impose strength demands that tend to increase muscle mass and body weight Lakdawalla, D. and T. Philipson (2007). "Labor Supply and Weight." Journal of Human Resources **42**(1): 85-116.

for time-varying health characteristics, which do not affect the estimated interaction between retirement and occupation. This provides some evidence that unobservable health shocks do not produce the differential effect of retirement across occupational groups.

### Results of instrumental variables models

There is no definitive evidence that BMI changes cause retirement, or that they are correlated with a third factor that does. Moreover, we also failed to find evidence that third factors like health shocks are related to the estimated interaction effects. Nonetheless, it is virtually impossible to rule out with certainty the possibility of reverse causation or an unobserved common causal factor. As a result, we explore models that allow for the possible endogeneity of the retirement decision. The more general IV specification reveals the same differential effect of retirement on the weight of sedentary and strenuous workers. However, it decreases our power to detect absolute weight loss among the sedentary retirees. This injects some uncertainty into our discussion of that result.

To model retirement as an endogenous process, we estimate a fixed-effects model that instruments for retirement using the ages of Social Security and Medicare eligibility (ages 62 and 65, respectively). We expect a discontinuous change in the incentive to retire at these ages. Note that a second-degree polynomial in age is present in the second-stage regression; therefore, the instruments pick up only the discrete break in retirement exactly at ages 62 and 65. The first-stages of the fixed-effects IV regression have the following form:

$$\begin{aligned} retired_{it} &= \gamma_0 + \gamma_1 age62_{it} + \gamma_2 age65_{it} + \gamma_3 age62_{it} * strenuous_i \\ &+ \gamma_4 age65_{it} * strenuous_i + \gamma_5 Z_{it} + \alpha_i + u_{it} \end{aligned} \tag{4}$$

$$\begin{aligned} retired_{it} * strenuous_i &= \delta_0 + \delta_1 age62_{it} + \delta_2 age65_{it} + \\ &\delta_3 age62_{it} * strenuous_i + \delta_4 age65_{it} * strenuous_i + \delta_5 Z_{it} + \alpha_i + u_{it} \end{aligned} \tag{5}$$

The variables $age62_{it}$ and $age65_{it}$ represent whether the individual is at or above age 62 or 65, respectively. The other variables are as defined previously. Since retirement is interacted with occupational strenuousness, we also interact strenuousness with the identifying instruments. One can think of this approach as running two IV models, separately by occupation, but imposing the restriction that the covariates $z_{it}$ have the same effects on both types of occupation.

The first-stage results of the IV model are shown in Table 2-4.[10] Turning age 62 and 65 have the expected effects on retirement: becoming eligible for Social Security (turning 62) increases the probability of retirement by about 15% for sedentary workers and 24% for strenuous workers; Medicare eligibility increases the probability by about half that much. The instruments are quite strong. The F-statistics easily pass the "cookbook" cut-off of 10.0, and the single instrument of age 62 has a t-statistic over 10.

The second-stage results are given in Table 2-5. The IV strategy increases standard errors relative to the fixed-effects model, and this eliminates the significance of the retirement effect for sedentary workers. However, retirement continues to increase weight relatively more for strenuous workers, by 0.62 BMI units over the sample window or 0.17 per HRS wave. These results confirm the fixed-effects finding that workers in strenuous jobs gain more weight after retirement.

### Effect of Retirement Over the Life-Course

However, while the fixed-effects results found *ceteris paribus* weight loss among sedentary retirees and weight gain among strenuous retirees, the instrumental variables results find the opposite. From a broader perspective, however, these discrepancies are of little consequence, because weight grows over time for all workers. Therefore, if one computes the total predicted change in weight, rather than the ceteris paribus change, both sets of results confirm substantial BMI gains for strenuous retirees, and modest BMI gains for sedentary retirees.

This point is made by Figure 2-3, which plots the predicted change in BMI from retirement onwards, for the mean sedentary and strenuous workers. Taking into account

[10] To save space, we suppress the sensitivity analyses with health controls. These are virtually identical to the models reported.

the age effects, a sedentary worker experiencing the average change in health and wealth will first gain a little weight, then slowly flat off. A strenuous worker, however, will experience a weight gain of 0.87 BMI (about 5-6 pounds) in a 10-year period.

## *Mechanisms for the Retirement Effect*

There are two aspects of the data in need of explanation. First, retirees from sedentary jobs gain more weight in retirement than those from strenuous jobs. Second, according to the fixed-effects estimates, retirees from sedentary jobs lose weight in absolute terms.[11]

### Differences Across Occupation

The divergence in trend at retirement is probably best explained by the direct effects discussed in the theory; these are the direct reductions in exercise faced by retirees from strenuous jobs. However, we need to rule out two alternative possibilities: (1) Retirees from sedentary jobs gain more time at retirement, and they spend more of this time on exercise; or (2) Different income effects across occupation create differences in food intake that drive the results.

To test the first hypothesis, we analyzed whether the reduction in hours worked differs across occupation. The results, shown in Appendix Table A6, imply that there is no differential change by occupation. On average, retirement leads to a 36 hour reduction in weekly work time. However, there is no significant difference across occupations. The difference in the hours reduction for strenuous occupations is estimated to be -0.36 hours, with a 95% confidence interval of [-1.2, 0.5]. The bottom of the confidence interval would imply that sedentary workers gain an extra 1.2 hours per week at retirement, but even this represents just 3% of the total effect of retirement on time available. It thus seems unlikely that differential time reduction explains patterns in weight.

[11] The OLS regression shows that this within-individual change in weight is offset by the between-individual variation in weight across retirees and non-retirees.

The presence of income effects is more difficult to test directly, because we do not observe consumption in the HRS waves.[12] However, including controls for wealth and income, which are likely correlated with the degree of consumption-smoothing, has no impact on the estimated interaction between retirement and occupation (Table A2, and Table A3 in the appendix). Moreover, it is difficult from a theoretical perspective to explain weight loss for sedentary retirees using income effects, unless these individuals become richer at retirement.[13] Appendix Table A5 explicitly analyzes the change in log household income at retirement, by occupation. The table demonstrates that there is no statistically significant increase in income for sedentary retirees, relative to strenuous retirees.

## Weight-Loss for Sedentary Retirees

The second effect, of weight loss for sedentary retirees, requires an additional mechanism. The theory predicts a behavioral response by workers who have more time to spend after retirement. Some of that time is likely to be spent on exercise. Workers who spent most of their working time in sedentary pursuits will be likely to experience an absolute decline in weight. In the next section, we present direct evidence that retirees from sedentary jobs increase their vigorous physical activity. The only alternative explanation would be a differential change in food consumption by retirees from different kinds of jobs. We explore that hypothesis in a later section, and find little evidence — both in our data, and in the previous literature — of changes in food consumption. These two analyses suggest that retiring from a sedentary job induces greater investments of leisure-time into physical activity.

---

[12] Consumption measures are available in the last two waves (2001 and 2003). We found, for models estimated only over these two waves, that including consumption had no impact on the interaction between retirement and occupation. Clearly, however, statistical power is a concern in this analysis.

[13] For those who are overweight, weight loss is likely to be a normal good.

## The Effect of Retirement on Physical Activity

At an immediate level, changes in weight are always caused by changes in physical activity, changes in food intake, or some combination of the two. From its third wave (in 1996) onwards, the HRS asks respondents about their level of physical activity in the following survey question: "On average over the last 12 months have you participated in vigorous physical activity or exercise three times a week or more? By vigorous physical activity, we mean: things like sports, heavy housework, or a job that involves physical labor." The question elicits either a "yes" or "no" answer. We use the responses to this question in waves 3 through 6 of the HRS. While this is a crude and imperfect measure of exercise, it sheds some much-needed light on how retirement affects exercise.

Table 2-6: Proportion of HRS males engaging in vigorous physical activity 3 or more times weekly suggests that changes in exercise may be playing an important role in driving the impact of retirement on weight changes. The table demonstrates that the frequency of thrice-weekly vigorous physical activity decreases for workers retiring from strenuous jobs, but increases for retirees from sedentary jobs. Figure 2-4 breaks this apart further by time until retirement (only HRS males who retire during the survey window are included in the figure). From the wave before retirement onwards, there is a steady decline in the vigorous physical activity of workers in strenuous jobs. However, retirement actually causes a slight increase in vigorous activity for workers in sedentary jobs.

We can test these patterns formally by repeating the OLS, fixed-effects, and fixed-effects instrumental variables analysis. The only difference from the previous analysis is the use of vigorous physical activity as the dependent variable, instead of BMI. To economize on space, we report results only for models with the binary measure of retirement. The results from these analyses are entirely consistent with the patterns in weight we documented earlier.

Table 2-7 reports the results for all the analyses of exercise. In this case, the OLS and fixed-effects results are quite similar, indicating that baseline differences in exercise do not play a significant role in the retirement-activity relationship. The proportion of retirees from sedentary jobs engaging in vigorous physical activity increases by 5 to 6 percentage points. In contrast, the proportion falls by 14 to 16 percentage points for retirees from strenuous jobs. Just as in the BMI analysis, instrumenting for retirement increases the standard errors and renders the effect of retirement insignificant for sedentary workers. However, retirees from strenuous jobs decrease their activity significantly more than others, even in the instrumental variables specification. The first-stage results associated with the instrumental variables regression appear in Table 2-8. These are substantially similar to the earlier first-stage findings.

### The Effect of Retirement on Food Intake and Preparation

Changes in exercise patterns appear to align with changes in weight; this suggests that exercise can explain the changes in weight. It remains to show whether differential changes in food intake (due, for example, to different income effects of retirement) play a role. While the HRS measures of food intake are not as powerful as the exercise measures, they do suggest that weight changes are not well-explained by changes in food intake. This finding is consistent with the previous literature on this subject.

The economics literature suggests that retirement does not affect food consumption. Aguiar and Hurst (2005) show that neither quality nor food intake change at retirement(Aguiar and Hurst 2005). Unfortunately, they cannot directly examine whether the effect of retirement on food intake differs by occupation, due to limitations in their data. However, their analysis demonstrates that — for the average worker — food intake does not change. The only way food intake could explain our findings, while remaining consistent with Aguiar and Hurst, would be if retirees from sedentary jobs cut their food intake by as much as retirees from strenuous jobs raised it.

It is difficult to rule this out, because the HRS data on food intake is rather limited, but the Consumption and Activities Mail Survey (CAMS) supplements to the HRS ask subsamples of respondents to self-report their spending on food and drink consumed at home, as well as away from home. 980 males in our analytic sample can be linked to CAMS. And among those 980, 487 respondents have only one wave of data, 122 of them have two waves of data, and 371 of them have three waves of data. The small samples, and particularly the small number of men with repeated observations, make fixed-effects models too imprecise for inference. Table A4 in the appendix reports results of OLS regressions that test differences in food expenditures among retirees from different occupations. We find no differential effects by occupation.

It is difficult to draw firm conclusions, given the small samples, and the lack of direct data on calorie intake. Therefore, we supplement this analysis with direct investigation of whether income changes differentially at retirement. As discussed previously, Appendix Table A5 demonstrates that household income changes uniformly at retirement for retirees from strenuous and sedentary occupations. This further cuts against the hypothesis that changes in calorie intake (and food expenditure) explain much if any of the differential change in weight.

## *Health Effects of Weight Changes at Retirement*

As discussed in the introduction, changes in body mass index could have impacts on health, particularly since the average US male in this age group is already overweight. To assess this hypothesis, we examine the impact of retirement — by occupation — on several obesity-related diseases: diabetes, hypertension, and heart disease (Chan, Rimm et al. 1994; Colditz, Willett et al. 1995; Willett, Manson et al. 1995; Huang, Willett et al. 1998). All are consistent with differentially worsening health for those retiring from strenuous jobs. In interpreting these results, a natural concern is that of time-varying unobservables across occupations. For example, if workers in strenuous jobs face more rapid deterioration in health, these patterns may be unrelated to retirement. It is rather difficult to rule out this hypothesis for hypertension and heart disease, but there is clearer evidence for diabetes.

### Trends in the Unadjusted Data

Figure 2-5 depicts the changes over time in diabetes, by occupation. Prior to retirement, the prevalence of diabetes is growing at roughly the same rate for both types of workers, with a possibly faster rate of growth among the sedentary workers. However, after retirement, the growth in diabetes prevalence is more rapid among the strenuous retirees, who are also gaining greater body weight. In the unadjusted data, the differences after retirement are not statistically significant, due to the relatively small numbers of sample respondents with the condition.

Figure 2-5 depicts the changes overtime in hypertension, by occupation. At retirement, strenuous workers had significantly higher increase of hypertension prevalence, relative to sedentary workers. 4 years previous to retirement, the prevalence of hypertension is growth at almost the same rate for both types of workers, but from 8 years to 4 years before retirement, hypertension rate is growing faster among strenuous workers, relative to sedentary workers. This pre-existing divergence in prevalence growth put into doubt whether the divergence after retirement is due to retirement, or it's just a continuation of the before-retirement trend.

Figure 2-7 depicts the changes overtime in heart disease, by occupation. The trends are very similar for the two types of workers.

### Regression Analysis

Since there is a pre-retirement divergence in hypertension growth, and the heart disease trends are almost parallel by occupation, we only carry out regression analysis for the outcome of diabetes. Table 2-9 shows the fixed effects estimates, using the whole sample, and using the balanced sample only. According to the whole sample estimation controlling for occupation-specific age trend, strenuous workers experience higher diabetes incidence at retirement. The difference is 1.5%, which is more than 20 percent of the diabetes prevalence at baseline (6.6%). Using balanced panel estimation, the magnitude is even larger: 2.7%. Hausman specification test rejects the hypothesis that the estimates from the whole sample are the same as those from the balanced sample. To correct for attrition bias, we adopt the method proposed by Wooldridge(Wooldridge 2002). We first run Probit models of whether an individual appears in the next wave as a function of this wave's conditions, for each of wave 2 to wave 7. Predictors include this wave's demographics, age, household income, household wealth, BMI, self-rated health, and chronic conditions. We then calculate inverse-mills ratios for those who were interviewed in the next wave. Finally we first-difference the data, and run an OLS regression of first-differenced BMI against first-differenced regressors, and also the inverse mills ratios. The results are shown in Table 2-11. The upper panel shows the first-difference results not corrected for attrition bias. The lower panel shows the first-difference results corrected for attrition bias. The two sets of results are very similar, which reveals that attrition bias is not the problem in the first-difference estimation. The magnitudes of the retirement * strenuous occupation are similar to those in Table 2-9 and are all statistically significant.

The regression results confirm the pattern of the unobserved trends shown in Figure 2-5. Strenuous workers experienced higher rates of diabetes incidence at retirement, relative to sedentary workers.

**Simulation Results of Increasing Diabetes Rates at Age 51/52**

Regression analysis shows that retirement imposes an extra 2% incidence rate of diabetes among strenuous workers. We use the Future Elderly Model(Goldman, Shang et al. 2005) to evaluate the lifetime impact of a 2% increase in diabetes prevalence among Americans aged 51/52 olds. As Table 2-12 shows, for those who are made diabetic, the average life-expectancy reduction is 4 years, and disability-free life-expectancy is reduced by 5 years, by the mean time their lifetime health care costs by 55,000 dollars.

## *The Effect of Retirement on Weight for Female*

Although we've focused our analysis on males, it is helpful to check how working females' weight change at retirement. At the baseline interview, among women aged 50-61, the labor participation rate is 61%, while the rate among men is 79%. Among 6,072 females interviewed, 2,999 females meet the criteria (the same as the criteria for males) to be included in analytic sample. By the same cutoff point for job strenuousness, the proportion of women working in strenuous job is lower than that in men – only one third, while among men it's more than half.

Unadjusted trends of BMI by occupation for females are shown in Figure C 2-1. The trends are similar by occupation. Both types of workers continue gaining weight until 4 years after retirement. After that the rates of increase flatten off. Figure C 2-2 shows the unadjusted trends of taking 3+ vigorous exercise per week, for females and by occupation. From 4 years before retirement to 4 years after retirement, the trends are very similar by occupation. After 4 years of retirement, strenuous workers seem to experience a larger drop in physical activity.

Next we turn to regression analysis. The individual fixed effects estimation shows that sedentary workers, the effect of retirement on BMI is insignificant. Strenuous workers have higher weight gain at retirement but the difference is insignificant. The individual fixed effects with IV estimation (Table C 2-3) shows even smaller difference in weight gain by retirement, by occupation.

As for the effect of retirement on physical activity, as Table C 2-4 shows, in both Individual fixed effects and individual fixed effects with IV models, the effects of retirement on physical activity both for sedentary and strenuous workers are insignificant. The strenuous workers experience more drop in physical activity relative to sedentary workers, but the difference is insignificant neither.

## *Conclusions*

We have presented evidence that retirement increases weight for those in physically active jobs, but not for those in sedentary jobs, where it may even lower weight. The data suggest that the difference among retirees is the result of differential declines in job-related exercise. Moreover, the weight-loss among retirees from sedentary jobs provides some evidence that retirement frees up time, some of which is spent on weight-controlling exercise. Indeed, the data suggest that the patterns in weight change are best explained by direct and indirect changes in physical exercise, rather than food intake.

The results in this paper suggest that labor force participation contributes to weight control for those in physically demanding jobs, but reduces weight control investments for those in sedentary jobs. Moreover, it also provides a glimpse into how labor supply affects weight, both through direct effects of work on exercise, and through reductions in the availability of leisure-time for exercise.

The data used in this paper have permitted more analysis of food intake and exercise than is typically possible with survey data, but it is clearly limited. As we have discussed, the food intake data are relatively sparse, and we were only able to examine a particular, crude measure of exercise. Better data are needed to quantify more precisely the causes of weight gain.

## *Data Appendix*

### Body Mass Index and Physical Activity

We use self-reported weight and height to calculate body mass index (BMI) ($kg/m^2$). Weight is reported in each wave. Height is usually reported only at the first two waves of interview and is regarded as constant in later waves. HRS asks questions about participating in light and heavy physical activities. Light physical activity was asked only in the first two waves. Vigorous physical activity was asked in all waves but the question in the first two waves is not comparable with the question from wave 3 to wave 6. As a result, we define physical activity according to the vigorous physical activity question in wave 3 to 6. It is a binary variable measuring participation in three or more vigorous physical activities three times a week or more. It takes the value of one if the answer is yes, and zero otherwise. Only wave 3-6 HRS data is used for the analysis on physical activity.

### Retirement Status

Retirement status is derived from the constructed variable of "labor force status" in the RAND-HRS data. Those respondents coded as being employed full-time, employed part-time, or unemployed, are considered to be 'working.' Those coded as being retired, partly retired, or disabled are considered to be 'retired.' We consider those respondents coded as not in the labor force to be neither working nor retired. The labor force status variable is constructed from several sources in the HRS questionnaire, addressing work, retirement, and disability. See St. Clair (2005) for a more detailed description of its construction.

### Primary Occupation and Job Strenuousness

We obtained access to the restricted RAND HRS data containing detailed 3-digit occupational codes for the jobs with the longest reported tenure at each interview. The occupation codes follow the 1980 Standard Occupational Classification system. We define primary occupation for an individual as the longest-held occupation at the wave of retirement. For those who are not retired until the end of the study, primary occupation is defined as the first available occupation for the longest tenure. After identifying primary occupation for each individual, we use the occupational measures file for the 1980 census detailed occupations (EK file) (England and Kilbourne 1989) to obtain job strenuousness measure. The EK file includes a physical demand score for each occupation. The score ranges from 0-3. The higher is the score, the more strenuous the occupation. The job strenuousness variable is defined as a binary one. It takes the value of 1 if the physical demand score for the primary occupation is greater than or equal to 1, and 0 otherwise. About half of the males in our sample have strenuous jobs.

## Consumption and Activities Mail Survey (CAMS)

The Consumption and Activities Mail Survey (CAMS) is a supplement to the core HRS survey. In the fall of 2001, questionnaires were mailed out to a subsample of 5,000 households interviewed in the 2000 HRS survey. Questions were asked on individual activities and household patterns of consumption. If the household has two panel members, one member was randomly selected to receive the survey. 3,866 respondents completed the interview in 2001. In 2003 and 2005, the same households were contacted for the survey. 3,254 completed the survey in 2003. In 2005 there is a slight change of sample. If a household has two eligible members, the respondent – contacted for the 2001 or 2003 CAMS – was mailed the "full" survey which included a complete list of questions. The spouse or partner of that respondent received a "partial" questionnaire containing only questions on activities. In 2005, 3,880 completed the "full" version and 1,935 completed the "partial" version.

CAMS asks two questions on household food expenditure, one is about spending on food and drinks (including alcoholic) that people buy in grocery or other stores. The other is about spending on dining/drinking out: items in restaurants, cafes, diners, and take-out restaurants. Respondents can report weekly, monthly or yearly spending. We convert these into annual spending and into year 2005 dollars.

CAMS also asks questions on time use in various activities during last week. We examined two activities that are related to food production. These activities include: meal preparation time, clean-up time, and time spent shopping or running errands.

## Attrition bias

The original HRS cohort and the War Babies cohort were first interviewed in year 1992 and 1998, respectively. After initial interview they were then interviewed biennially up to year 2004. However, attrition could happen due to various reasons: the most obvious reason is mortality; there were also non-response among individuals alive; finally, some individuals were not interviewed yet their mortality status was unknown. In addition to the three types of attrition, we introduced two other types of attrition in our analytic sample: first, if at a certain wave an individual reported not being retired after the first wave of reporting being retired, the individual's waves of observations since that wave were dropped. Second, an observation is dropped if there is any missing value for one of the dependent or independent variables.

The analytic sample includes 3,936 males with 20,913 observations. 3,169 of them belong to the original HRS cohort while 767 of them belong to the War Babies cohort. If there were no attrition then the number of observations would be 25,251. Therefore there are 4,338 cases of attrition. Table Appendix A7 shows the break down of the 4,338 cases into five types, as mentioned above. The attrition might be non-absorbing if it is not due to mortality.

If attrition patterns are endogenous then our estimation will be inconsistent. To check whether attrition bias exists, we used a Hausman type specification test proposed by Nijman and Verbeek (Verbeek and Nijman). The key idea is to test whether the estimation using the whole (both balanced and unbalanced) sample is the same as the estimation using only the balanced sample. And we use bootstrap to calculate the standard errors of the difference between two estimations. Appendix Table A-8 shows the fixed effects estimations using the whole sample and the balanced sample only, under two model specifications. For Hausman specification test, the two p-values for the corresponding chi-square statistics are very high and we cannot reject the null hypothesis that the estimation from the balanced sample and the estimation from the unbalanced sample are the same. Therefore we don't find evidence of attrition bias.

## *Theoretical Appendix*

In this appendix, we prove that when $E_j$ falls, total exercise ($E_h$+$E_j$ ) falls, food intake (F) falls, and weight (W) rises. By displacement, we can rewrite the problem in 1 as a function of only food intake and exercise at home.

$$\max_{F_t, E_{ht}} U(I_t - pF_t, F_t, T_t - E_{ht}, W(F_t, E_{ht} + E_{jt}), E_{ht} + E_{jt})$$

Observe that the first order conditions associated with the displaced problem can be written as:

$$\begin{aligned} -pU_c + U_F + U_W W_F &= 0 \\ -U_L + U_W W_E + U_E &= 0 \end{aligned} \tag{6}$$

To simplify the analysis, define this displaced objective function as $V(F, E_h; E_j, T_t, p, I_t)$. $E_h$ and $E_j$ enter the analysis nearly symmetrically, except that increases in $E_h$ have the additional effect of reducing time available for labor. In particular, we have:

$$\begin{aligned} V_{FE_h} &= V_{FE_j} - \frac{\partial}{\partial T_t} V_F \\ V_{E_h E_h} &= V_{E_h E_j} - \frac{\partial}{\partial T_t} V_{E_h} \end{aligned} \tag{7}$$

We assume that reductions in time available raise the marginal utility of food and exercise, because income effects dominate, so that $\frac{\partial}{\partial T_t} V_F > 0$ and $\frac{\partial}{\partial T_t} V_{E_h}$.

To ensure the existence and uniqueness of this optimum, we assume that $V(F, E)$ is jointly concave. We assume further that $V_{FE_h} > 0$, $V_{FE_j} > 0$, and $V_{FE} > 0$, where $E$ is total exercise. Therefore, whenever total exercise rises — and no matter how it rises —

food intake is more valuable. Finally, we assume that $V_{E_hE_j} < 0$, so that exercise at home and at work are substitutes. Taken together, these assumptions imply the following key conditions:

$$\begin{aligned} &V_{FE_j} > V_{FE_h} > 0 \\ &V_{E_hE_h} < V_{E_hE_j} < 0 \end{aligned} \tag{8}$$

Performing comparative statics on the objective function $V$ yields:

$$\begin{bmatrix} V_{FF} & V_{FE_h} \\ V_{FE_h} & V_{E_hE_h} \end{bmatrix} \begin{bmatrix} \dfrac{\partial F}{\partial E_j} \\ \dfrac{\partial E_h}{\partial E_j} \end{bmatrix} = \begin{bmatrix} -V_{FE_j} \\ -V_{E_hE_j} \end{bmatrix} \tag{9}$$

$$\frac{\partial E_h}{\partial E_j} = -\frac{V_{FF}V_{E_hE_j} - V_{FE_h}V_{FE_j}}{V_{FF}V_{E_hE_h} - V_{FE_h}^2} \tag{10}$$

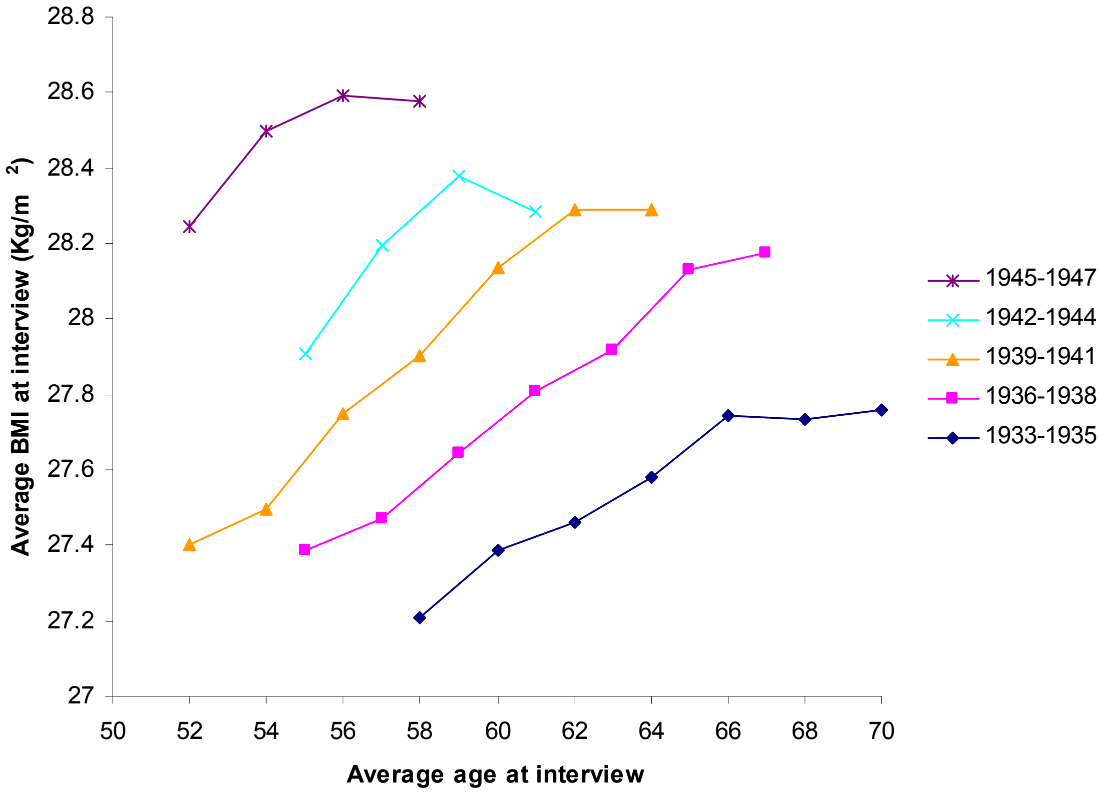
The concavity of the problem and assumptions about cross-partials imply that $\dfrac{\partial E_h}{\partial E_j} < 0$.

Moreover, the inequalities in equation 8 imply that: $\left|\dfrac{\partial E_h}{\partial E_j}\right| < 1$. This implies that increases in on-the-job exercise lead to partially compensating reductions in exercise at-home, and vice-versa. This implies that reductions in $E_j$ lead to reductions in $E_h + E_j$. QED

Define $F^o$ and $E^o$ as the original levels of food consumption and at-home exercise, and $F^n$ and $E^n$ as the new levels. To prove that $F$ falls, assume that $F^n \geq F^o$. Note that $V_F(F^o, E_h^0; E_j^o) = 0$. Moreover, in the new equilibrium, total exercise is lower. Therefore, $V_F(F^o, E_h^n; E_j^n) < 0$. By concavity, it must then be true that $V_F(F^n, E_h^n; E_j^n) \leq V_F(F^o, E_h^n; E_j^n) < 0$. This is a contradiction.

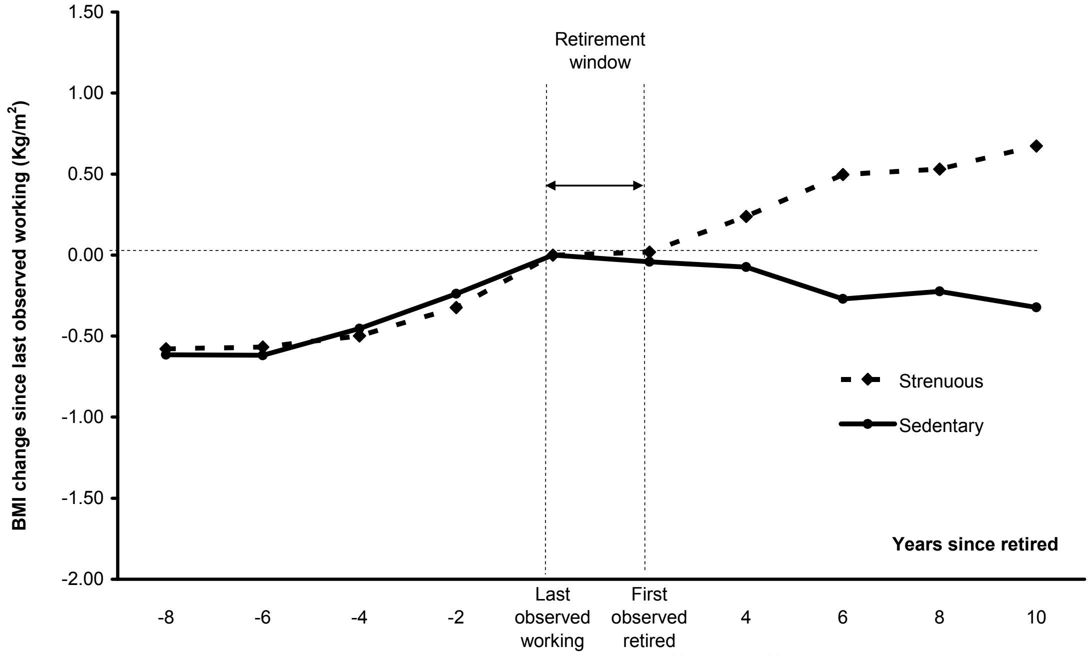
Finally, we prove that $W^n > W^o$. Since weight is a function of exercise and food intake, we can rewrite our objective function as a function of only exercise and weight. Define this objective function as $Z(W, E_h; E_j)$, and note that $Z_{WE} > 0$, based on the properties of $V$. Suppose that $W^n \leq W^o$. Clearly, $Z_{E_h}(W^o, E_h^o; E_j^o) = 0$. Moreover, by concavity, $Z_E(W^n, E_h^o; E_j^o) \geq Z_E(W^o, E_h^o; E_j^o) = 0$. Since $E^n < E^o$, concavity implies $Z_E(W^n, E_h^n; E_j^n) > Z_E(W^n, E_h^o; E_j^o) \geq 0$. This is a contradiction.

**Figure 2-1: Growth in BMI for five elderly birth cohorts.**




Source: Health and Retirement Study 1992 to 2004
Notes: This figure depicts BMI trajectories of five birth cohorts born between 1933 and 1947. Those males with non-missing BMI in all eligible waves are included. As a result, sample sizes are identical across years for a given cohort, and are as follows: n = 462 (1945-1947), n = 481(1942-1944), n = 780(1939-1941) n = 737(1936-1938), and n = 653(1933-1935).

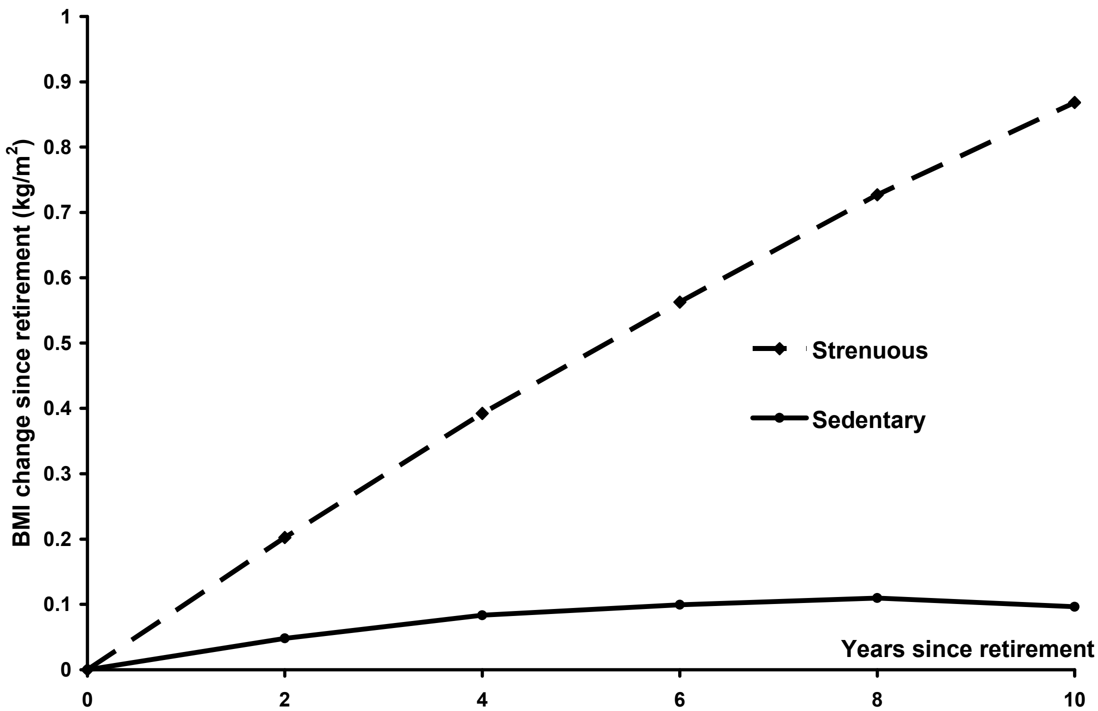
**Figure 2-2: Unadjusted change in BMI since last observed working for males by occupation**




| | -8 | -6 | -4 | -2 | Last observed working | First observed retired | 4 | 6 | 8 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Strenuous | -0.58 | -0.57 | -0.50 | -0.32 | 0.00 | 0.02** | 0.24** | 0.50 | 0.53 | 0.67 |
| N | 306 | 511 | 724 | 984 | 1262 | 943 | 643 | 423 | 268 | |
| Sedentary | -0.62 | -0.62 | -0.45 | -0.24 | 0.00 | -0.04** | -0.07** | -0.27 | -0.22 | -0.32 |
| N | 243 | 384 | 566 | 749 | 957 | 712 | 477 | 304 | 191 | |

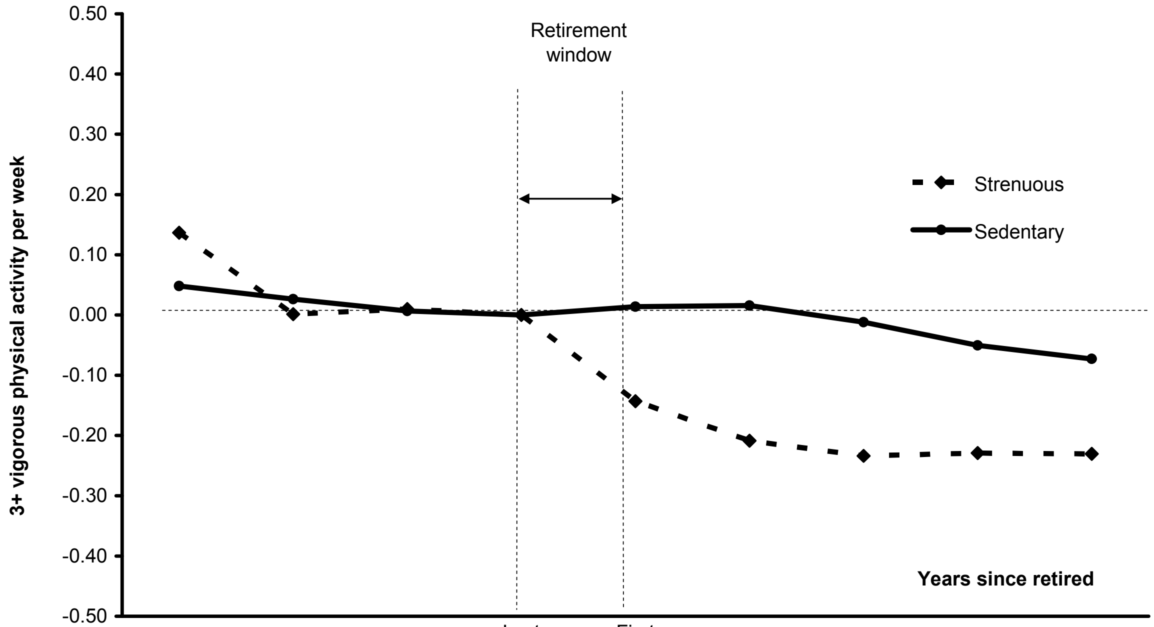
Source: Health and Retirement Study, 1992 to 2004.
Notes: The figure is based on the men in our sample who retire at some point during the HRS observation window. Since we can define a retirement date for each such person, we can also calculate time until or since retirement. This figure calculates cumulative changes in mean BMI since last observed working, as a function of time from retirement, for males in different occupations. The table at the bottom of the figure shows the point estimates for cumulative changes in BMI from the retirement date, as well as the relevant sample sizes. "**" indicates that the change of BMI for males in strenuous occupation at a certain time since retirement is significantly different from that in sedentary occupation, at the 5% significance level.

**Figure 2-3: Simulated changes in BMI since retirement for males, by occupation**




Source: Health and Retirement Study, 1992 to 2004.
Notes: this graph is the simulated BMI changes since retirement by occupation, based on the fixed-effects estimation results, as shown in the last column of Table 3. We take the average age last observed working in HRS as the age at retirement year 0, which is 59. We then increase age (and years since retirement) by two years a step, until 10 years later. We set the values of other covariates in the BMI fixed effects regression - *ln(income), non-positive wealth, ln(wealth), self-rated health is fair or poor, number of IADL limitations, number of ADL limitations, diagnosed with cancer, diagnosed with lung disease, diagnosed with heart disease, diagnosed with a stroke, diagnosed with hypertension* - at their mean values of each projected age (59,61,63,65,67,69). We set the occupation to be sedentary and simulate BMI changes since retirement for those with sedentary occupation. We then set the occupation to be strenuous and simulate BMI changes since retirement for those with strenuous occupation. As a result, the difference in the two BMI trajectories is completely due to the differential retirement effect by occupation, as measured by the interaction term of "waves since retire * strenuous job" in the last column of Table 3.

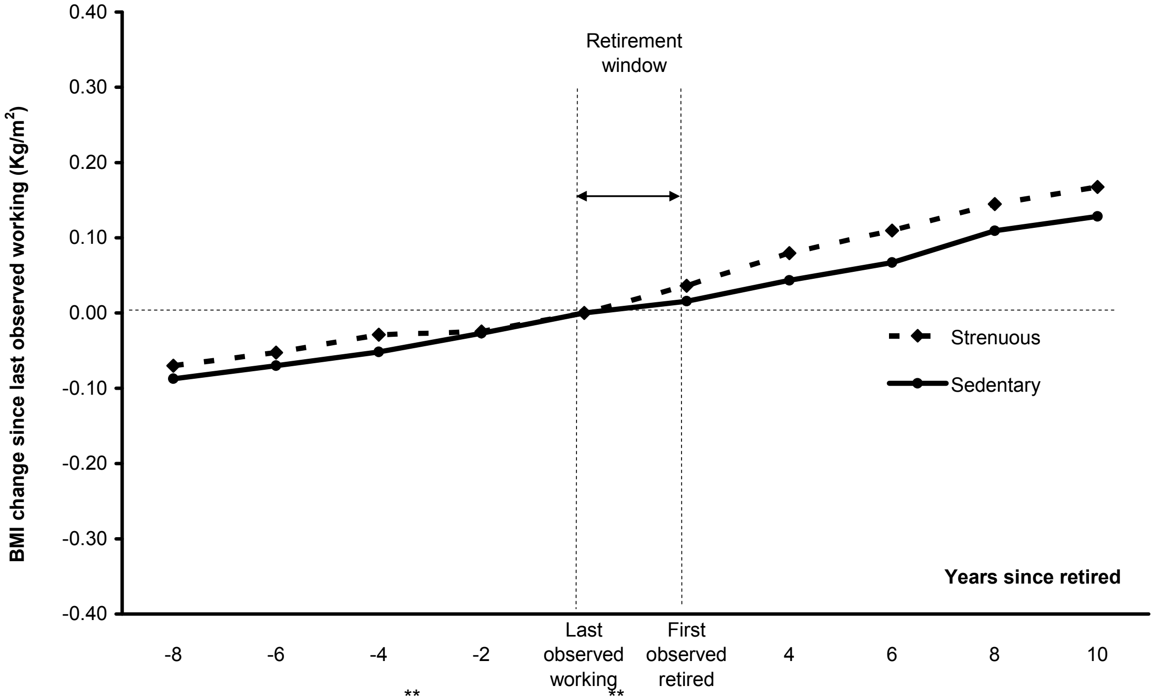
**Figure 2-4 Unadjusted change in percentage taking vigorous exercise since last observed working for males by occupation**




| | -6 | -4 | -2 | Last observed working | First observed retired | 4 | 6 | 8 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| Strenuous | 0.14 | 0.00 | 0.01 | 0.00** | -0.14** | -0.21 | -0.23 | -0.23 | -0.23 |
| N | 133 | 326 | 540 | 613 | 579 | 483 | 294 | 115 | |
| Sedentary | 0.05 | 0.03 | 0.01 | 0.00** | 0.01** | 0.02 | -0.01 | -0.05 | -0.07 |
| N | 91 | 280 | 430 | 473 | 413 | 351 | 208 | 96 | |

Source: Health and Retirement Study, 1992 to 2004.
Notes: The figure is based on the men in our sample who retire at some point during the HRS observation window. Since we can define a retirement date for each such man, we can also determine time until or since retirement. This figure calculates cumulative changes in percentage of individuals taking vigorous exercise since last observed working, as a function of time, for males in different occupations. The table at the bottom of the figure shows the cumulative changes, and the number of persons that are interviewed both at a certain wave since retire and at the next adjacent wave, for males in different occupation. "**" indicates that the change of percentage points from last wave in strenuous occupation is significantly different from that in sedentary occupation, at the 5% significance level.

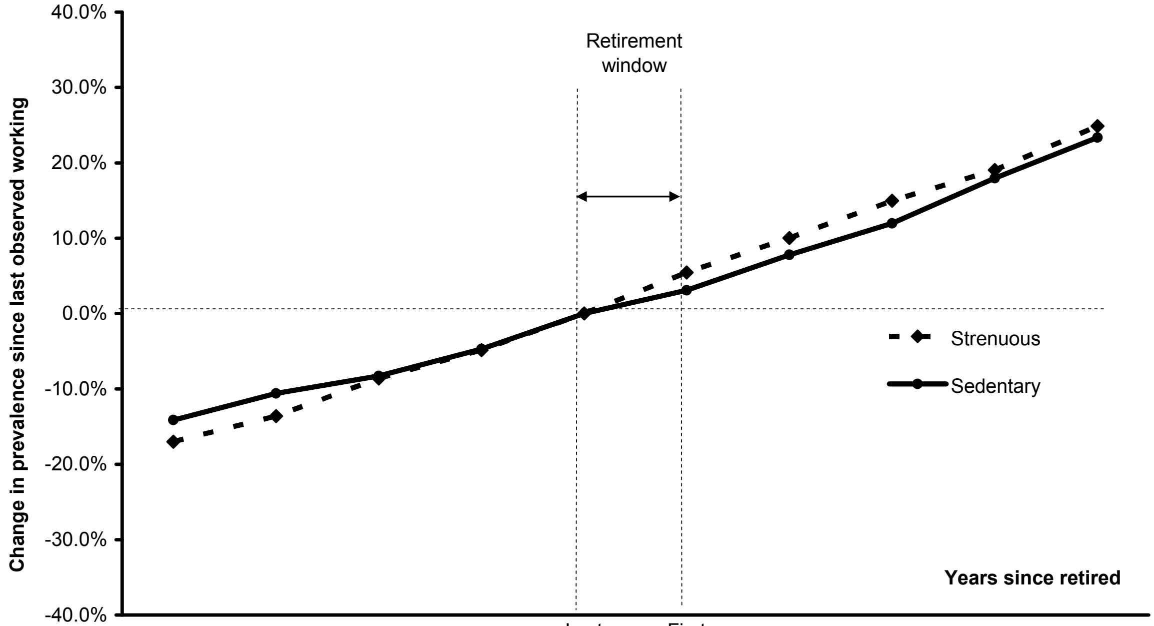
**Figure 2-5 Unadjusted change in percentage ever diagnosed with diabetes since last observed working for males by occupation**




| | -8 | -6 | -4 | -2 | Last observed working | First observed retired | 4 | 6 | 8 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Strenuous | -0.07 | -0.05 | -0.03** | -0.02 | 0.00** | 0.04 | 0.08 | 0.11 | 0.14 | 0.17 |
| N | 306 | 511 | 724** | 984 | 1,262** | 943 | 643 | 423 | 268 | |
| Sedentary | -0.09 | -0.07 | -0.05 | -0.03 | 0.00 | 0.02 | 0.04 | 0.07 | 0.11 | 0.13 |
| N | 243 | 384 | 566 | 749 | 957 | 712 | 477 | 304 | 191 | |

Source: Health and Retirement Study, 1992 to 2004.
Notes: The figure is based on the men in our sample who retire at some point during the HRS observation window. Since we can define a retirement date for each such man, we can also determine time until or since retirement. This figure calculates cumulative changes in the prevalence of diabetes since last observed working, as a functional of time, for males in different occupations. The table at the bottom of the figure shows the cumulative changes, and the number of persons that are interviewed both at a certain time since retire and at the next adjacent wave. "**" indicates that the incidence of diabetes for males in strenuous occupation at a certain time since retirement is significantly different from that in sedentary occupation, at the 5% significance level.

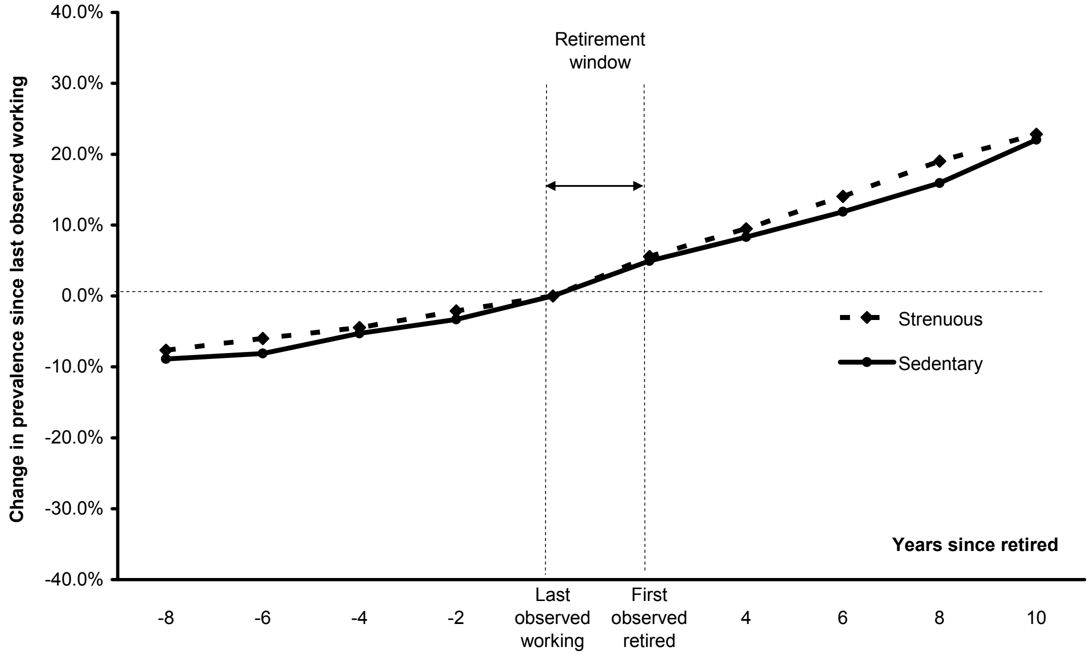
**Figure 2-6 Unadjusted change in percentage ever diagnosed with hypertension since last observed working for males by occupation**




| | -8 | -6 | -4 | -2 | Last observed working | First observed retired | 4 | 6 | 8 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Strenuous | -17.0% | -13.6%** | -8.6% | -4.9% | 0.0%** | 5.5% | 10.0% | 15.0% | 19.1% | 24.9% |
| N | 306 | 511 | 724 | 984 | 1,262 | 943 | 643 | 423 | 268 | |
| Sedentary | -14.1% | -10.6%** | -8.3% | -4.7% | 0.0%** | 3.1% | 7.8% | 12.0% | 18.0% | 23.3% |
| N | 243 | 384 | 566 | 749 | 957 | 712 | 477 | 304 | 191 | |

Source: Health and Retirement Study, 1992 to 2004.
Notes: The figure is based on the men in our sample who retire at some point during the HRS observation window. Since we can define a retirement date for each such man, we can also determine time until or since retirement. This figure calculates cumulative changes in the prevalence of hypertension since last observed working, as a functional of time, for males in different occupations. The table at the bottom of the figure shows the cumulative changes, and the number of persons that are interviewed both at a certain time since retire and at the next adjacent wave. "**" indicates that the incidence of hypertension for males in strenuous occupation at a certain time since retirement is significantly different from that in sedentary occupation, at the 5% significance level.

**Figure 2-7 Unadjusted change in percentage ever diagnosed with heart disease since last observed working for males by occupation**




| | -8 | -6 | -4 | -2 | Last observed working | First observed retired | 4 | 6 | 8 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Strenuous | -7.7% | -6.0% | -4.5% | -2.1% | 0.0% | 5.6% | 9.5% | 14.1% | 19.0% | 22.8% |
| N | 306 | 511 | 724 | 984 | 1,262 | 943 | 643 | 423 | 268 | |
| Sedentary | -8.9% | -8.1% | -5.3% | -3.3% | 0.0% | 4.9% | 8.3% | 11.9% | 15.9% | 22.0% |
| N | 243 | 384 | 566 | 749 | 957 | 712 | 477 | 304 | 191 | |

Source: Health and Retirement Study, 1992 to 2004.
Notes: The figure is based on the men in our sample who retire at some point during the HRS observation window. Since we can define a retirement date for each such man, we can also determine time until or since retirement. This figure calculates cumulative changes in the prevalence of heart disease since last observed working, as a functional of time, for males in different occupations. The table at the bottom of the figure shows the cumulative changes, and the number of persons that are interviewed both at a certain time since retire and at the next adjacent wave. "**" indicates that the incidence of heart disease for males in strenuous occupation at a certain time since retirement is significantly different from that in sedentary occupation, at the 5% significance level.

**Table 2-1: HRS Summary Statistics.**

| | Type of occupation | | | | | |
|---|---|---|---|---|---|---|
| | All (n= 3,936) | | Non-strenuous (n= 1,806) | | Strenuous (n= 2,130) | |
| Characteristic | Mean | Std. Dev. | Mean | Std. Dev. | Mean | Std. Dev. |
| **Body mass (kg/m$^2$):** | | | | | | |
| Baseline | 28.88 | 5.48 | 28.75 | 5.28 | 29.01 | 5.67 |
| End of study* | 29.44 | 5.85 | 29.16 | 5.57 | 29.72 | 6.10 |
| **Retirement:** | | | | | | |
| Retired by end of study | 0.521 | 0.500 | 0.484 | 0.500 | 0.557 | 0.497 |
| Age when retirement first reported | 61.1 | 3.9 | 61.1 | 4.0 | 61.0 | 3.8 |
| **Socioeconomic status (at baseline):** | | | | | | |
| Age | 54.3 | 2.9 | 54.2 | 2.9 | 54.4 | 2.9 |
| White | 0.887 | 0.317 | 0.926 | 0.263 | 0.850 | 0.358 |
| Hispanic | 0.063 | 0.243 | 0.029 | 0.168 | 0.096 | 0.294 |
| Education | | | | | | |
| Less than high-school degree | 0.151 | 0.358 | 0.047 | 0.212 | 0.252 | 0.434 |
| High school degree | 0.347 | 0.476 | 0.211 | 0.408 | 0.479 | 0.500 |
| Some college (no degree) | 0.210 | 0.408 | 0.233 | 0.423 | 0.188 | 0.391 |
| College degree | 0.292 | 0.455 | 0.509 | 0.500 | 0.081 | 0.273 |
| Household income ($1998) | 79,477 | 85,473 | 103,040 | 105,492 | 56,561 | 50,260 |
| Household wealth ($1998) | 308,616 | 605,175 | 430,258 | 768,810 | 190,317 | 345,526 |
| **Health status (baseline):** | | | | | | |
| Self-rated health is fair or poor | 0.113 | 0.317 | 0.071 | 0.257 | 0.154 | 0.361 |
| Number of IADL limitations | 0.069 | 0.312 | 0.032 | 0.204 | 0.106 | 0.386 |
| Number of ADL limitations | 0.023 | 0.184 | 0.015 | 0.151 | 0.030 | 0.211 |
| Cancer | 0.022 | 0.146 | 0.028 | 0.166 | 0.015 | 0.122 |
| Lung disease | 0.031 | 0.172 | 0.026 | 0.159 | 0.035 | 0.184 |
| Heart disease | 0.088 | 0.283 | 0.083 | 0.275 | 0.092 | 0.290 |
| Stroke | 0.013 | 0.111 | 0.016 | 0.125 | 0.009 | 0.096 |
| Diabetes | 0.066 | 0.248 | 0.065 | 0.247 | 0.066 | 0.249 |
| Hypertension | 0.297 | 0.457 | 0.286 | 0.452 | 0.307 | 0.461 |

Notes: Baseline BMI for each respondent is measured in 1992 for the original HRS cohort (n = 3,169), and in 1998 for the War Babies cohort (n = 767). "End of study" BMI is measured in the respondent's final HRS wave. The average elapsed time between baseline and the end of study is 10 years for the original HRS cohort, and 5.5 years for the War Babies cohort. Instrumental activities of daily living (IADL) include three activities: using the phone, managing money, and taking medications. Activities of daily living (ADL) include five activities: bathing, eating, dressing, getting in and out of bed, and walking across a room.  The existence of disease is determined by asking the respondent, “Has a doctor ever told you that you had [disease]?”

**Table 2-2: Examples of Strenuous and Non-Strenuous Occupations.**

| Occupation (3-digit 1980 Census Occupation Code) | Physical Demands |
|---|---|
| Non-strenuous | |
| Social work teachers (146) | 0 |
| Teachers, elementary school (156) | 0.07 |
| Legislators (003) | 0.43 |
| Managers and administrators (019) | 0.63 |
| Supervisors and preprietors, sales occupation (243) | 0.67 |
| Biological and life scientists (078) | 0.97 |
| | |
| Strenuous | |
| Mail carriers, postal services (355) | 1.03 |
| File clerks (335) | 1.35 |
| Automobile mechanics (505) | 1.92 |
| Janitors and cleaners (453) | 2.45 |
| Farm workers (479) | 2.57 |
| Electricians (575) | 2.72 |
| Carpenters (567) | 2.87 |

Notes: The table shows a sampling of strenuous and sedentary occupations in our data. Each occupation is a 1980 Census 3-digit occupation. The relevant 3-digit code is shown in parentheses next to the occupation title. The scores are derived from the England-Kilbourne file described in the text. They represent the number of physical demands present in the following list: climbing or balancing; stooping, kneeling, crouching, or crawling; reaching, handling, fingering, or feeling.

Table 2-3: OLS[a] and fixed effects regression results for effect of retirement on weight, by occupation

| Dependent variable: BMI | Specification | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | OLS | | | | Individual fixed effects | | | |
| Retired | -0.097 | -0.369 | | | -0.315*** | -0.301*** | | |
| | (0.242) | (0.234) | | | (0.062) | (0.062) | | |
| Retired*strenuous occupation | 0.966*** | 0.901*** | | | 0.324*** | 0.321*** | | |
| | (0.289) | (0.277) | | | (0.074) | (0.074) | | |
| Waves since retire | | | -0.003 | -0.101 | | | -0.150*** | -0.139*** |
| | | | (0.094) | (0.092) | | | (0.026) | (0.026) |
| Waves since retire*strenuous job | | | 0.340*** | 0.331*** | | | 0.155*** | 0.154*** |
| | | | (0.103) | (0.099) | | | (0.027) | (0.027) |
| Age | 0.310 | 0.194 | 0.508** | 0.293 | 0.352*** | 0.329*** | 0.278*** | 0.269*** |
| | (0.215) | (0.213) | (0.248) | (0.249) | (0.060) | (0.060) | (0.067) | (0.067) |
| Age squared | -0.003 | -0.002 | -0.005** | -0.003 | -0.002*** | -0.002*** | -0.002*** | -0.001** |
| | (0.002) | (0.002) | (0.002) | (0.002) | (0.001) | (0.001) | (0.001) | (0.001) |
| ln(income) | 0.163*** | 0.172*** | 0.160*** | 0.174*** | 0.040*** | 0.041*** | 0.040*** | 0.041*** |
| | (0.057) | (0.056) | (0.057) | (0.056) | (0.015) | (0.015) | (0.015) | (0.015) |
| Non-positive wealth | -0.423 | -0.124 | -0.406 | -0.131 | 0.114 | 0.104 | 0.099 | 0.092 |
| | (0.618) | (0.588) | (0.620) | (0.590) | (0.170) | (0.170) | (0.170) | (0.170) |
| ln(wealth) | -0.060 | -0.011 | -0.059 | -0.013 | -0.003 | -0.002 | -0.004 | -0.004 |
| | (0.049) | (0.046) | (0.049) | (0.046) | (0.015) | (0.015) | (0.015) | (0.015) |
| Control for health-related variables[b] | No | Yes | No | Yes | No | Yes | No | Yes |
| No. individuals | 3,936 | 3,936 | 3,936 | 3,936 | 3,936 | 3,936 | 3936 | 3,936 |
| No. of observations | 20,913 | 20,913 | 20,913 | 20,913 | 20,913 | 20,913 | 20,913 | 20,913 |
| Retired + Retired * strenuous occupation | 0.868 | 0.531 | | | 0.009 | 0.020 | | |
| P value of the Wald test of Retired + Retired * strenuous occupation = 0 | 0.000 | 0.020 | | | 0.882 | 0.738 | | |
| Waves since retire + Waves since retire*strenuous job | | | 0.337 | 0.231 | | | 0.006 | 0.015 |
| P value of the Wald test of Waves since retire + Waves since retire*strenuous job = 0 | | | 0.000 | 0.013 | | | 0.824 | 0.539 |

* p<0.10, ** p<0.05, *** p<0.01

Notes:
[a] OLS denotes ordinary least squares. all OLS regressions also include dummy variables for: strenuous occupation, white, Hispanic, less than high school degree, high school degree, and some college without degree, census region is Northeast, Middle west, and West.
[b] Health-related variables include number of IADL limitations, number of ADL limitations, self-rated health is fair/poor, ever being diagnosed with the following six chronic conditions: cancer, diabetes, heart disease, hypertension, lung disease, and stroke.

**Table 2-4: First-stage IV estimates[a] on retirement of turning age 62 and 65**

| | Specification | | | |
|---|---|---|---|---|
| | (1) | | (2) | |
| | **retired** | **retired* strenuous** | **waves of retirement** | **waves of retirement* strenuous** |
| Age | 0.005 | 0.009 | -1.319*** | -0.730*** |
| | (0.012) | (0.009) | (0.034) | (0.034) |
| Age squared | 0.000*** | 0.000 | 0.013*** | 0.007*** |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| Age >= 62 | 0.151*** | -0.108*** | -0.071*** | -0.463*** |
| | (0.013) | (0.004) | (0.027) | (0.014) |
| Age >= 65 | 0.091*** | -0.083*** | 0.002 | -0.704*** |
| | (0.015) | (0.006) | (0.037) | (0.023) |
| (Age >= 62) * Strenuous occupation | 0.092*** | 0.462*** | 0.105*** | 0.905*** |
| | (0.015) | (0.010) | (0.034) | (0.024) |
| (Age >= 65) * Strenuous occupation | 0.002 | 0.264*** | 0.254*** | 1.546*** |
| | (0.018) | (0.012) | (0.043) | (0.032) |
| ln(income) | -0.044*** | -0.023*** | -0.071*** | -0.030*** |
| | (0.003) | (0.003) | (0.007) | (0.005) |
| Non-positive wealth | 0.036 | -0.012 | -0.166** | -0.143*** |
| | (0.028) | (0.023) | (0.065) | (0.056) |
| ln(wealth) | -0.000 | -0.003 | -0.022*** | -0.017*** |
| | (0.002) | (0.002) | (0.006) | (0.005) |
| Control for health-related variables(b) | NO | NO | NO | NO |
| No. of observations | 3,936 | 3,936 | 3,936 | 3,936 |
| No. individuals | 20,913 | 20,913 | 20,913 | 20,913 |
| Joint F-statistic of age>=62, age >= 65, and interactions with strenuous occupation | 146 | 1501 | 22.7 | 1595 |
| P value of over-identification test | 0.360 | | 0.669 | |

* p<0.10, ** p<0.05, *** p<0.01

Notes:

[a] Models include individual fixed-effects. Age>=62, Age>=65, and interactions are instruments for retirement and its interaction.

[b] Health-related variables include number of IADL limitations, number of ADL limitations, self-rated health is fair/poor, ever being diagnosed with the following six chronic conditions: cancer, diabetes, heart disease, hypertension, lung disease, and stroke.

**Table 2-5: Second-stage IV estimates[(a)] of retirement on weight, by occupation**

| Dependent variable: BMI | Specification (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Retired | -0.199 | | -0.180 | |
| | (0.297) | | (0.295) | |
| Retired*strenuous occupation | 0.631*** | | 0.621*** | |
| | (0.140) | | (0.141) | |
| Waves since retire | | 0.148 | | 0.209 |
| | | (0.553) | | (0.552) |
| Waves since retire*strenuous occupation | | 0.178** | | 0.166* |
| | | (0.090) | | (0.091) |
| Age | 0.382*** | 0.717 | 0.358*** | 0.763 |
| | (0.074) | (0.719) | (0.074) | (0.708) |
| Age squared | -0.003*** | -0.006 | -0.002*** | -0.006 |
| | (0.001) | (0.007) | (0.001) | (0.007) |
| ln(income) | 0.052** | 0.062 | 0.052** | 0.065 |
| | (0.023) | (0.042) | (0.023) | (0.041) |
| Non-positive wealth | 0.117 | 0.160 | 0.103 | 0.158 |
| | (0.216) | (0.231) | (0.214) | (0.230) |
| ln(wealth) | -0.001 | 0.003 | -0.002 | 0.004 |
| | (0.019) | (0.022) | (0.018) | (0.021) |
| Control for health-related variables[(b)] | NO | NO | YES | YES |
| No. individuals | 3,936 | 3,936 | 3,936 | 3,936 |
| No. observations | 20,913 | 20,913 | 20,913 | 20,913 |
| Retired + Retired * strenuous occupation | 0.433 | | 0.440 | |
| P value of the Wald test of | | | | |
| Retired + Retired * strenuous occupation = 0 | 0.082 | | 0.074 | |
| Waves since retire + Waves since retire*strenuous job | | 0.326 | | 0.375 |
| P value of the Wald test of | | | | |
| Waves since retire + Waves since retire*strenuous job = 0 | | 0.491 | | 0.427 |

* p<0.10, ** p<0.05, *** p<0.01

Notes:
[(a)] Models include individual fixed-effects. Age>=62, Age>=65, and interactions are instruments for retirement and its interaction.
[(b)] Health-related variables include number of IADL limitations, number of ADL limitations, self-rated health is fair/poor, ever being diagnosed with the following six chronic conditions: cancer, diabetes, heart disease, hypertension, lung disease, and stroke

**Table 2-6: Proportion of HRS males engaging in vigorous physical activity 3 or more times weekly**

| Retirement status | Non-strenuous | | | Strenuous | | |
|---|---|---|---|---|---|---|
| | Mean | Std. Dev | N | Mean | Std. Dev | N |
| Not retired | 0.54 | 0.50 | 3,687 | 0.64 | 0.50 | 3,826 |
| Retired | 0.57 | 0.50 | 1,931 | 0.51 | 0.48 | 2,668 |

Source: Health and Retirement Study 1996 to 2002

Notes: Vigorous physical activity includes activity on-the-job and off-the-job.

**Table 2-7: Effect of retirement on vigorous physical activity of HRS males**

| | Specification | | |
|---|---|---|---|
| | **OLS** | **Individual fixed effects** | **Individual fixed effects with IV** |
| Retired | 0.049** | 0.063*** | 0.146 |
| | (0.020) | (0.022) | (0.114) |
| Retired*strenuous occupation | -0.139*** | -0.164*** | -0.150** |
| | (0.025) | (0.028) | (0.068) |
| Age | -0.035 | -0.024 | -0.014 |
| | (0.028) | (0.026) | (0.032) |
| Age squared | 0.000 | 0.000 | -0.000 |
| | (0.000) | (0.000) | (0.000) |
| ln(income) | 0.010** | 0.010** | 0.013** |
| | (0.005) | (0.005) | (0.006) |
| Non-positive wealth | 0.171*** | 0.033 | 0.033 |
| | (0.053) | (0.053) | (0.056) |
| ln(wealth) | 0.017*** | 0.003 | 0.004 |
| | (0.004) | (0.005) | (0.005) |
| Self-rated health is fair or poor | -0.104*** | -0.028* | -0.031 |
| | (0.018) | (0.016) | (0.020) |
| Number of IADL limitations | -0.032 | -0.012 | -0.015 |
| | (0.021) | (0.023) | (0.027) |
| Number of ADL limitations | -0.077*** | -0.045*** | -0.048*** |
| | (0.011) | (0.014) | (0.016) |
| Diagnosed with cancer | -0.056** | -0.057* | -0.059 |
| | (0.025) | (0.033) | (0.038) |
| Diagnosed with lung disease | -0.061** | -0.049 | -0.057 |
| | (0.031) | (0.050) | (0.061) |
| Diagnosed with heart disease | -0.001 | 0.026 | 0.021 |
| | (0.018) | (0.029) | (0.037) |
| Diagnosed with stroke | -0.074** | -0.148*** | -0.154*** |
| | (0.033) | (0.052) | (0.055) |
| Diagnosed with diabetes | -0.050** | 0.048 | 0.048 |
| | (0.021) | (0.031) | (0.038) |
| Diagnosed with hypertension | -0.049*** | 0.058** | 0.059** |
| | (0.014) | (0.023) | (0.025) |
| No. individuals | 3,712 | 3,712 | 3,417 |
| No. of observations | 12,112 | 12,112 | 11,797 |
| Retired + Retired * strenuous occupation | -0.090 | -0.101 | -0.004 |
| P value of the Wald test of Retired + Retired * strenuous occupation = 0 | <0.001 | <0.001 | 0.962 |

* p<0.10, ** p<0.05, *** p<0.01

Source: Health and Retirement Study 1996-2002
Notes: OLS regressions also include dummy variables for: strenuous occupation, white, Hispanic, less than high school degree, high school degree, and some college without degree. The model of “individual fixed effects with IV” adds the following instruments for retirement: Age>=62 and Age>=65

**Table 2-8: First-stage IV effects on retirement of turning age 62 and 65**

| | retired | retired* strenuous |
|---|---|---|
| Age | 0.028** | 0.030* |
| | (0.022) | (0.016) |
| Age squared | 0.000** | -0.000** |
| | (0.000) | (0.000) |
| Age >= 62 | 0.108*** | -0.069*** |
| | (0.017) | (0.005) |
| Age >= 65 | 0.090*** | -0.061*** |
| | (0.019) | (0.006) |
| (Age >= 62) * Strenuous occupation | 0.113*** | 0.370*** |
| | (0.022) | (0.015) |
| (Age >= 65) * Strenuous occupation | -0.027** | 0.203*** |
| | (0.023) | (0.016) |
| ln(income) | -0.031*** | -0.012*** |
| | (0.004) | (0.003) |
| Non-positive wealth | 0.024** | -0.023** |
| | (0.036) | (0.028) |
| ln(wealth) | -0.002** | -0.004** |
| | (0.003) | (0.002) |
| Self-rated health is fair or poor | 0.030** | 0.024** |
| | (0.013) | (0.011) |
| Number of IADL limitations | 0.042** | 0.025* |
| | (0.019) | (0.015) |
| Number of ADL limitations | 0.040*** | 0.028*** |
| | (0.012) | (0.010) |
| Diagnosed with cancer | 0.039** | 0.033** |
| | (0.029) | (0.023) |
| Diagnosed with lung disease | 0.072* | 0.031** |
| | (0.042) | (0.034) |
| Diagnosed with heart disease | 0.063*** | 0.039** |
| | (0.024) | (0.018) |
| Diagnosed with a stroke | 0.064** | 0.071* |
| | (0.046) | (0.037) |
| Diagnosed with diabetes | 0.010** | 0.036* |
| | (0.023) | (0.019) |
| Diagnosed with hypertension | -0.015** | 0.006** |
| | (0.017) | (0.013) |
| | | |
| Joint F-statistic of age>=62, age >= 65, and interactions with strenuous occupation | 55.1 | 275.8 |
| P value of over-identification test | 0.153 | |

* p<0.10, ** p<0.05, *** p<0.01

Data Source: Health and Retirement Study 1996-2002
Notes: OLS regressions also include dummy variables for: strenuous occupation, white, Hispanic, less than high school degree, high school degree, and some college without

degree. The model of "individual fixed effects with IV" adds the following instruments for retirement: Age>=62 and Age>=65

**Table 2-9: Fixed effects regression results for the effect of retirement on ever diagnosed with diabetes**

| Covariates | Specification (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Panel A: whole sample | | | | |
| Retired | -0.005 | -0.004 | | |
| | (0.005) | (0.006) | | |
| Retired*strenuous occupation | 0.017*** | 0.015* | | |
| | (0.007) | (0.008) | | |
| Waves since retire | | | 0.001 | 0.001 |
| | | | (0.002) | (0.003) |
| Waves since retire*strenuous occupation | | | 0.003 | 0.003 |
| | | | (0.002) | (0.004) |
| Age | 0.003 | -0.006 | 0.007 | -0.004 |
| | (0.005) | (0.008) | (0.006) | (0.008) |
| Age squared | 0.000* | 0.000** | 0.000 | 0.000* |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| Age*strenuous occupation | | 0.019* | | 0.022* |
| | | (0.011) | | (0.012) |
| Age squared*strenuous occupation | | -0.000* | | -0.000* |
| | | (0.000) | | (0.000) |
| Number of individuals | 20,913 | 20,913 | 20,913 | 20,913 |
| Retired + Retired * strenuous occupation | 0.012 | 0.011 | | |
| P value of the Wald test of Retired + Retired * strenuous occupation = 0 | 0.021 | 0.064 | | |
| Waves since retire + Waves since retire*strenuous job | | | 0.004 | 0.004 |
| P value of the Wald test of Waves since retire + Waves since retire*strenuous job = 0 | | | 0.044 | 0.121 |
| Panel B: balanced sample only | | | | |
| Retired | -0.018*** | -0.020*** | | |
| | (0.007) | (0.007) | | |
| Retired*strenuous occupation | 0.023*** | 0.027*** | | |
| | (0.008) | (0.010) | | |
| Waves since retire | | | -0.005** | -0.007** |
| | | | (0.003) | (0.003) |
| Waves since retire*strenuous occupation | | | 0.005* | 0.008* |
| | | | (0.003) | (0.004) |
| Age | 0.001 | -0.007 | -0.002 | -0.014 |
| | (0.006) | (0.009) | (0.007) | (0.010) |
| Age squared | 0.000* | 0.000** | 0.000** | 0.000*** |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| Age*strenuous occupation | | 0.018 | | 0.025* |
| | | (0.012) | | (0.014) |
| Age squared*strenuous occupation | | -0.000 | | -0.000* |
| | | (0.000) | | (0.000) |
| Number of individuals | 14,141 | 14,141 | 14,141 | 14,141 |
| Retired + Retired * strenuous occupation | 0.005 | 0.007 | | |
| P value of the Wald test of Retired + Retired * strenuous occupation = 0 | 0.43 | 0.82 | | |
| Waves since retire + Waves since retire*strenuous job | | | 0.000 | 0.001 |
| P value of the Wald test of | | | | |

| | | | | |
|---|---|---|---|---|
| Waves since retire + Waves since retire*strenuous job = 0 | | | 0.32 | 0.760 |
| Hausman specification test of attrition bias | | | | |
| F-statistic | 20.4 | 28.6 | 26.7 | 33.7 |
| P-value | 0.005 | 0.000 | 0.002 | 0.000 |
| * p<0.10, ** p<0.05, *** p<0.01 | | | | |

Data Source: Health and Retirement Study 1992-2004
Note: All models control for household income and household wealth

**Table 2-10: Fixed effects with IV regression results for the effect of retirement on ever diagnosed with diabetes**

| Covariates | Specification (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Panel A: whole sample | | | | |
| Retired | -0.005 | -0.008 | | |
| | (0.028) | (0.048) | | |
| Retired*strenuous occupation | 0.019 | 0.028 | | |
| | (0.014) | (0.054) | | |
| Waves since retire | | | 0.058 | -0.057 |
| | | | (0.055) | (0.109) |
| Waves since retire*strenuous occupation | | | -0.002 | 0.180 |
| | | | (0.009) | (0.126) |
| Age | 0.003 | -0.006 | 0.082 | -0.079 |
| | (0.007) | (0.011) | (0.072) | (0.140) |
| Age squared | 0.000 | 0.000 | -0.001 | 0.001 |
| | (0.000) | (0.000) | (0.001) | (0.001) |
| Age*strenuous occupation | | 0.019 | | 0.279 |
| | | (0.015) | | (0.170) |
| Age squared*strenuous occupation | | -0.000 | | -0.003 |
| | | (0.000) | | (0.002) |
| Number of individuals | 20,913 | 20,913 | 20,913 | 20,913 |
| Retired + Retired * strenuous occupation | 0.014 | 0.020 | | |
| P value of the Wald test of Retired + Retired * strenuous occupation = 0 | 0.529 | 0.452 | | |
| Waves since retire + Waves since retire*strenuous job | | | 0.056 | 0.123 |
| P value of the Wald test of Waves since retire + Waves since retire*strenuous job = 0 | | | 0.237 | 0.040 |
| Panel B: balanced sample only | | | | |
| Retired | 0.010 | -0.023 | | |
| | (0.034) | (0.055) | | |
| Retired*strenuous occupation | 0.020 | 0.077 | | |
| | (0.016) | (0.065) | | |
| Waves since retire | | | 0.072 | -0.013 |
| | | | (0.068) | (0.092) |
| Waves since retire*strenuous occupation | | | -0.002 | 0.198 |
| | | | (0.010) | (0.134) |
| Age | 0.004 | -0.008 | 0.105 | -0.024 |
| | (0.008) | (0.013) | (0.093) | (0.127) |
| Age squared | 0.000 | 0.000 | -0.001 | 0.000 |
| | (0.000) | (0.000) | (0.001) | (0.001) |
| Age*strenuous occupation | | 0.023 | | 0.318 |
| | | (0.017) | | (0.199) |
| Age squared*strenuous occupation | | -0.000 | | -0.003 |
| | | (0.000) | | (0.002) |
| Number of individuals | 14,141 | 14,141 | 14,141 | 14,141 |
| Retired + Retired * strenuous occupation | 0.030 | 0.054 | | |
| P value of the Wald test of Retired + Retired * strenuous occupation = 0 | 0.307 | 0.119 | | |

| | | | | |
|---|---|---|---|---|
| Waves since retire + Waves since retire*strenuous job | | | 0.070 | 0.185 |
| P value of the Wald test of Waves since retire + Waves since retire*strenuous job = 0 | | | 0.242 | 0.054 |
| Hausman specification test of attrition bias | | | | |
| F-statistic | 1.207 | 2.712 | 0.278 | 0.901 |
| P-value | 0.547 | 0.258 | 0.870 | 0.637 |

* p<0.10, ** p<0.05, *** p<0.01

Data Source: Health and Retirement Study 1992-2004
Note: All models control for household income and household wealth.

**Table 2-11: First-difference regression results for the effect of retirement on ever diagnosed with diabetes**

| Covariates | Specification (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Panel A: no correction for attrition bias | | | | |
| Retired | -0.009** | -0.009* | | |
| | (0.004) | (0.005) | | |
| Retired*strenuous occupation | 0.019*** | 0.018** | | |
| | (0.007) | (0.008) | | |
| Waves since retire | | | -0.001 | -0.005 |
| | | | (0.004) | (0.005) |
| Waves since retire*strenuous occupation | | | 0.012*** | 0.020*** |
| | | | (0.005) | (0.007) |
| Age | -0.021** | -0.026** | -0.013 | -0.031** |
| | (0.010) | (0.013) | (0.012) | (0.015) |
| Age squared | 0.000*** | 0.000** | 0.000 | 0.000** |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| Age*strenuous occupation | | 0.010 | | 0.036* |
| | | (0.018) | | (0.021) |
| Age squared*strenuous occupation | | -0.000 | | -0.000* |
| | | (0.000) | | (0.000) |
| Number of individuals | 16,412 | 16,412 | 16,412 | 16,412 |
| Retired + Retired * strenuous occupation | 0.010 | 0.009 | | |
| P value of the Wald test of | | | | |
| Retired + Retired * strenuous occupation = 0 | 0.000 | 0.000 | | |
| Waves since retire + Waves since retire*strenuous job | | | 0.011 | 0.015 |
| P value of the Wald test of | | | | |
| Waves since retire + Waves since retire*strenuous job = 0 | | | 0.000 | 0.000 |
| Panel B: with correction for attrition bias | | | | |
| Retired | -0.009** | -0.008* | | |
| | (0.004) | (0.004) | | |
| Retired*strenuous occupation | 0.020*** | 0.018** | | |
| | (0.007) | (0.008) | | |
| Waves since retire | | | -0.002 | -0.005 |
| | | | (0.004) | (0.005) |
| Waves since retire*strenuous occupation | | | 0.013*** | 0.019*** |
| | | | (0.005) | (0.007) |
| Age | -0.017 | -0.022 | -0.009 | -0.027* |
| | (0.011) | (0.014) | (0.013) | (0.015) |
| Age squared | 0.000 | 0.000 | 0.000 | 0.000* |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| Age*strenuous occupation | | 0.011 | | 0.037* |
| | | (0.018) | | (0.021) |
| Age squared*strenuous occupation | | -0.000 | | -0.000* |
| | | (0.000) | | (0.000) |
| Number of individuals | 16,412 | 16,412 | 16,412 | 16,412 |
| Retired + Retired * strenuous occupation | 0.011 | 0.010 | | |
| P value of the Wald test of | | | | |
| Retired + Retired * strenuous occupation = 0 | 0.000 | 0.000 | | |
| Waves since retire + Waves since retire*strenuous job | | | 0.011 | 0.014 |

| | | | | |
|---|---|---|---|---|
| P value of the Wald test of<br>Waves since retire + Waves since retire*strenuous job = 0 | | | 0.000 | 0.000 |

* p<0.10, ** p<0.05, *** p<0.01

Data Source: Health and Retirement Study 1992-2004
Note: All models control for household income and household wealth

**Table 2-12 Lifetime outcomes of aged 51/52 in the United States in year 2004**

| | Years lived since age 51 | Healthy years lived since age 51 | Lifetime medical spending* |
|---|---|---|---|
| **For the whol sample** | | | |
| Status_quo | 31.1 | 28.3 | 245,276 |
| Increased diabetes prevalence** | 31.0 | 28.3 | 246,205 |
| Relative change due to increased diabetes prevalence | -0.2% | -0.3% | 0.4% |
| | | | |
| **For those whose are made diabetic in year 2004** | | | |
| Status quo | 30.2 | 28.8 | 275,386 |
| Increased diabetes prevalence | 26.3 | 23.8 | 330,500 |
| Relative change due to increased diabetes prevalence | -13.1% | -17.2% | 20.0% |

Notes:
* Dollars are in 2004 values and are discounted by 3% per year.
** We randomly select those who are obese and non-diabetic in year 2004 and make them diabetic, so that the total prevalence of diabetes in year 2004 increases by 2%.

## *Appendix A. More Regression Results*

**Table A 2-1. Sensitivity analysis of estimating the effect of retirement on weight – Part I**

| Dependent variable: BMI | OLS | | Individual fixed effects | |
|---|---|---|---|---|
| | (1) | (2) | (1) | (2) |
| Retired | -0.369 | 0.060 | -0.301*** | -0.204* |
| | (0.234) | (0.457) | (0.062) | (0.123) |
| Retired*strenuous occupation | 0.901*** | 0.880*** | 0.321*** | 0.317*** |
| | (0.277) | (0.278) | (0.074) | (0.074) |
| Hours worked per week | | 0.021*** | | |
| | | (0.007) | | |
| Retired*Hours worked per week | | -0.008 | | -0.002 |
| | | (0.009) | | (0.002) |
| Age | 0.194 | 0.192 | 0.329*** | 0.330*** |
| | (0.213) | (0.214) | (0.060) | (0.060) |
| Age squared | -0.002 | -0.002 | -0.002*** | -0.002*** |
| | (0.002) | (0.002) | (0.001) | (0.001) |
| ln(income) | 0.172*** | 0.155*** | 0.041*** | 0.041*** |
| | (0.056) | (0.056) | (0.015) | (0.015) |
| Non-positive wealth | -0.124 | -0.274 | 0.104 | 0.104 |
| | (0.588) | (0.592) | (0.170) | (0.170) |
| ln(wealth) | -0.011 | -0.025 | -0.002 | -0.002 |
| | (0.046) | (0.047) | (0.015) | (0.015) |
| Number of IADL limitations | -0.774*** | -0.742*** | -0.152** | -0.151** |
| | (0.189) | (0.188) | (0.062) | (0.062) |
| Number of ADL limitations | 0.554** | 0.555** | 0.033 | 0.031 |
| | (0.223) | (0.224) | (0.043) | (0.043) |
| Diagnosed with cancer | 0.311 | 0.295 | -0.221** | -0.222** |
| | (0.308) | (0.309) | (0.093) | (0.093) |
| Diagnosed with lung disease | -0.392 | -0.395 | -0.150 | -0.150 |
| | (0.402) | (0.401) | (0.127) | (0.127) |
| Diagnosed with heart disease | 0.464* | 0.475** | 0.006 | 0.006 |
| | (0.243) | (0.242) | (0.079) | (0.079) |
| Diagnosed with a stroke | -0.270 | -0.318 | -0.448*** | -0.446*** |
| | (0.462) | (0.459) | (0.143) | (0.143) |
| Diagnosed with diabetes | 2.706*** | 2.717*** | -0.538*** | -0.538*** |
| | (0.299) | (0.299) | (0.086) | (0.086) |
| Diagnosed with hypertension | 2.106*** | 2.106*** | 0.200*** | 0.199*** |
| | (0.174) | (0.173) | (0.066) | (0.066) |
| Self-rated health is fair or poor | 0.194 | 0.225 | 0.097* | 0.098* |
| | (0.209) | (0.209) | (0.052) | (0.052) |
| N | 20,913 | 20,913 | 20,913 | 20,913 |

* p<0.10, ** p<0.05, *** p<0.01

Data source: Health and Retirement Study 1992-2004

Note:
Hours worked per week are the sum of hours worked for the main job and the secondary jobs. For respondents who are retired by the end of the observation window, we take hours worked per week at the wave before they first report retirement. For respondents who are not retired by the end of the observation window, we take the average of hours worked per week over all waves in which they are interviewed.
OLS denotes ordinary least squares. All OLS regressions also include dummy variables for: strenuous occupation, white, hispanic, less than high school degree, high school degree, and some college without degree.

**Table A 2-2. Sensitivity analysis of estimating effect of retirement on weight - Part II**

| Dependent variable: BMI | OLS | | Fixed effects | | Fixed effects with IV | |
|---|---|---|---|---|---|---|
| | **(1)** | **(2)** | **(1)** | **(2)** | **(1)** | **(2)** |
| Retired | -0.369 | -0.463** | -0.301*** | -0.318*** | -0.180 | -0.190 |
| | (0.234) | (0.235) | (0.062) | (0.062) | (0.295) | (0.294) |
| Retired*strenuous occupation | 0.901*** | 0.907*** | 0.321*** | 0.321*** | 0.621*** | 0.620*** |
| | (0.277) | (0.277) | (0.074) | (0.074) | (0.141) | (0.141) |
| ln(income) | 0.172*** | | 0.041*** | | 0.052** | |
| | (0.056) | | (0.015) | | (0.023) | |
| Non-positive wealth | -0.124 | | 0.104 | | 0.103 | |
| | (0.588) | | (0.170) | | (0.214) | |
| ln(wealth) | -0.011 | | -0.002 | | -0.002 | |
| | (0.046) | | (0.015) | | (0.018) | |
| Age | 0.194 | 0.188 | 0.329*** | 0.326*** | 0.358*** | 0.357*** |
| | (0.213) | (0.214) | (0.060) | (0.060) | (0.074) | (0.074) |
| Age squared | -0.002 | -0.002 | -0.002*** | -0.002*** | -0.002*** | -0.002*** |
| | (0.002) | (0.002) | (0.001) | (0.001) | (0.001) | (0.001) |
| Number of IADL limitations | -0.774*** | -0.794*** | -0.152** | -0.150** | -0.146* | -0.145* |
| | (0.189) | (0.189) | (0.062) | (0.062) | (0.081) | (0.082) |
| Number of ADL limitations | 0.554** | 0.540** | 0.033 | 0.034 | 0.017 | 0.018 |
| | (0.223) | (0.225) | (0.043) | (0.043) | (0.084) | (0.084) |
| Diagnosed with cancer | 0.311 | 0.327 | -0.221** | -0.220** | -0.223* | -0.222* |
| | (0.308) | (0.309) | (0.093) | (0.093) | (0.132) | (0.132) |
| Diagnosed with lung disease | -0.392 | -0.395 | -0.150 | -0.150 | -0.196 | -0.197 |
| | (0.402) | (0.402) | (0.127) | (0.127) | (0.171) | (0.171) |
| Diagnosed with heart disease | 0.464* | 0.465* | 0.006 | 0.007 | -0.014 | -0.014 |
| | (0.243) | (0.243) | (0.079) | (0.079) | (0.098) | (0.098) |
| Diagnosed with stroke | -0.270 | -0.261 | -0.448*** | -0.453*** | -0.500** | -0.507** |
| | (0.462) | (0.463) | (0.143) | (0.143) | (0.207) | (0.208) |
| Diagnosed with diabetes | 2.706*** | 2.696*** | -0.538*** | -0.541*** | -0.551*** | -0.554*** |
| | (0.299) | (0.301) | (0.086) | (0.086) | (0.121) | (0.121) |
| Diagnosed with hypertension | 2.106*** | 2.111*** | 0.200*** | 0.197*** | 0.202*** | 0.199*** |
| | (0.174) | (0.174) | (0.066) | (0.066) | (0.071) | (0.071) |
| Self-rated health is fair or poor | 0.194 | 0.151 | 0.097* | 0.096* | 0.079 | 0.077 |
| | (0.209) | (0.210) | (0.052) | (0.052) | (0.075) | (0.075) |
| N | 20,913 | 20,913 | 20,913 | 20,913 | 20,913 | 20,913 |

* p<0.10, ** p<0.05, *** p<0.01

Data source: Health and Retirement Study 1992-2004
Notes: OLS regressions also include dummy variables for: strenuous occupation, white, Hispanic, less than high school degree, high school degree, and some college without degree. The model of "individual fixed effects with IV" adds the following instruments for retirement: Age>=62 and Age>=65

**Table A 2-3. Sensitivity analysis of estimating effect of retirement on weight - Part III**

| Dependent variable: BMI | OLS | | Fixed effects | | Fixed effects with IV | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (1) | (2) | (1) | (2) |
| Waves since retire | -0.101 | -0.134 | -0.139*** | -0.145*** | 0.209 | 0.244 |
| | (0.092) | (0.092) | (0.026) | (0.026) | (0.552) | (0.577) |
| Waves since retire* strenuous occupation | 0.331*** | 0.336*** | 0.154*** | 0.155*** | 0.166* | 0.160* |
| | (0.099) | (0.099) | (0.027) | (0.027) | (0.091) | (0.095) |
| ln(income) | 0.174*** | | 0.041*** | | 0.065 | |
| | (0.056) | | (0.015) | | (0.041) | |
| Non-positive wealth | -0.131 | | 0.092 | | 0.158 | |
| | (0.590) | | (0.170) | | (0.230) | |
| ln(wealth) | -0.013 | | -0.004 | | 0.004 | |
| | (0.046) | | (0.015) | | (0.021) | |
| Age | 0.293 | 0.258 | 0.269*** | 0.260*** | 0.763 | 0.809 |
| | (0.249) | (0.250) | (0.067) | (0.067) | (0.708) | (0.742) |
| Age squared | -0.003 | -0.003 | -0.001** | -0.001** | -0.006 | -0.007 |
| | (0.002) | (0.002) | (0.001) | (0.001) | (0.007) | (0.007) |
| Number of IADL limitations | -0.783*** | -0.802*** | -0.150** | -0.148** | -0.173* | -0.175* |
| | (0.189) | (0.189) | (0.062) | (0.062) | (0.091) | (0.093) |
| Number of ADL limitations | 0.551** | 0.535** | 0.034 | 0.035 | -0.003 | -0.008 |
| | (0.225) | (0.227) | (0.043) | (0.043) | (0.098) | (0.100) |
| Diagnosed with cancer | 0.302 | 0.318 | -0.211** | -0.209** | -0.271 | -0.276 |
| | (0.307) | (0.308) | (0.093) | (0.093) | (0.167) | (0.169) |
| Diagnosed with lung disease | -0.402 | -0.405 | -0.158 | -0.157 | -0.267 | -0.277 |
| | (0.402) | (0.402) | (0.127) | (0.127) | (0.216) | (0.221) |
| Diagnosed with heart disease | 0.463* | 0.463* | 0.010 | 0.011 | -0.050 | -0.055 |
| | (0.243) | (0.244) | (0.079) | (0.079) | (0.128) | (0.130) |
| Diagnosed with stroke | -0.283 | -0.274 | -0.447*** | -0.453*** | -0.558** | -0.578** |
| | (0.463) | (0.464) | (0.143) | (0.143) | (0.251) | (0.258) |
| Diagnosed with diabetes | 2.705*** | 2.695*** | -0.535*** | -0.538*** | -0.546*** | -0.548*** |
| | (0.299) | (0.301) | (0.086) | (0.086) | (0.121) | (0.121) |
| Diagnosed with hypertension | 2.106*** | 2.110*** | 0.199*** | 0.196*** | 0.210*** | 0.208*** |
| | (0.174) | (0.174) | (0.066) | (0.066) | (0.075) | (0.075) |
| Self-rated health is fair or poor | 0.197 | 0.151 | 0.095* | 0.094* | 0.092 | 0.087 |
| | (0.209) | (0.210) | (0.052) | (0.052) | (0.075) | (0.075) |
| N | 20,913 | 20,913 | 20,913 | 20,913 | 20,913 | 20,913 |

* p<0.10, ** p<0.05, *** p<0.01

Data saource: Health and Retirement Study 1992-2004
Notes: OLS regressions also include dummy variables for: strenuous occupation, white, Hispanic, less than high school degree, high school degree, and some college without degree. The model of "individual fixed effects with IV" adds the following instruments for retirement: Age>=62 and Age>=65

**Table A 2-4: OLS regressions of retirement effect on food expenditure and time use**

| | Annual spending on grocery ($) | Annual spending on dining-out ($) | Annual spending on food ($) | Hours per week preparing meals | Hours per week shopping |
|---|---|---|---|---|---|
| Retired | -737.904** | -219.834 | -970.631** | 1.582*** | 1.298*** |
| | (300.390) | (312.049) | (489.842) | (0.556) | (0.388) |
| Retired*strenuous occupation | 348.719 | 302.762 | 640.260 | -0.079 | -0.452 |
| | (368.369) | (347.789) | (580.913) | (0.934) | (0.602) |
| Age | -986.571* | -430.815 | -1446.866* | -0.432 | 0.039 |
| | (556.447) | (384.933) | (776.535) | (0.742) | (0.495) |
| Age squared | 7.750* | 3.605 | 11.604* | 0.003 | -0.000 |
| | (4.535) | (3.085) | (6.299) | (0.006) | (0.004) |
| Married/Partnered | 1637.012*** | 42.599 | 1702.340*** | -1.146** | 0.653* |
| | (271.871) | (231.998) | (426.960) | (0.499) | (0.390) |
| ln(income) | 275.715** | 303.582*** | 570.413*** | 0.017 | -0.245 |
| | (127.739) | (96.191) | (193.783) | (0.198) | (0.182) |
| Non-positive wealth | 85.343 | 1837.421* | 2019.039 | 2.362 | -1.382 |
| | (1044.843) | (1025.409) | (1733.408) | (2.318) | (1.122) |
| ln(wealth) | -4.173 | 153.997* | 152.852 | 0.239 | -0.074 |
| | (76.110) | (85.327) | (138.257) | (0.206) | (0.093) |
| Self-rated health is fair or poor | 320.407 | -350.816 | -40.602 | 0.222 | -0.540* |
| | (244.160) | (249.897) | (401.848) | (0.996) | (0.294) |
| Number of IADL limitations | -23.322 | 223.486 | 252.898 | 5.367 | 1.101 |
| | (235.918) | (324.229) | (356.737) | (4.358) | (1.097) |
| Number of ADL limitations | -98.226 | 104.559 | -5.320 | -0.265 | 0.009 |
| | (125.557) | (167.347) | (225.658) | (0.446) | (0.332) |
| Diagnosed with cancer | 934.412* | 168.024 | 1115.365 | -0.838 | -0.320 |
| | (489.756) | (360.963) | (748.107) | (0.536) | (0.308) |
| Diagnosed with lung disease | 636.111 | 25.465 | 679.078 | 1.062 | 0.047 |
| | (394.984) | (374.673) | (665.415) | (0.899) | (0.377) |
| Diagnosed with heart disease | -174.543 | 378.495* | 222.468 | 0.074 | 0.136 |
| | (270.167) | (220.980) | (372.229) | (0.463) | (0.293) |
| Diagnosed with a stroke | -492.192 | -238.660 | -751.471 | -1.594 | -0.087 |
| | (385.886) | (378.518) | (610.054) | (1.209) | (0.620) |
| Diagnosed with diabetes | 440.060 | 327.885 | 818.826** | 0.304 | 0.813 |
| | (273.627) | (234.465) | (379.400) | (0.945) | (0.549) |
| Diagnosed with hypertension | -91.565 | 21.472 | -56.896 | 0.334 | 0.130 |
| | (201.909) | (178.001) | (293.242) | (0.483) | (0.292) |
| Adjusted R square | 0.080 | 0.086 | 0.108 | 0.050 | 0.022 |
| N | 1,782 | 1,791 | 1,776 | 1,820 | 1,816 |

* p<0.10, ** p<0.05, *** p<0.01

Data source: HRS in 2000, 2002 and 2004, and CAMS (Consumption and Activities Mail Survey) supplements in 2001, 2003, 2005. Spending is reported at the household level. Hours spent on various activities are reported at the individual level.

Spending on food is the sum of spending on grocery and dining-out.

Note: all regressions also include dummy variables for: strenuous occupation, white, Hispanic, less than high school degree, high school degree, and some college without degree.

**Table A 2-5. Individual fixed effects models of retirement effects on household income**

| | Log (household income + 1) | Log(household income)* |
|---|---|---|
| Retired | -0.428*** | -0.357*** |
| | (0.031) | (0.020) |
| Retired*strenuous occupation | -0.008 | -0.026 |
| | (0.037) | (0.023) |
| Age | -0.014 | 0.034* |
| | (0.030) | (0.019) |
| Age squared | 0.000 | -0.000* |
| | (0.000) | (0.000) |
| Number of IADL limitations | -0.007 | -0.027 |
| | (0.031) | (0.020) |
| Number of ADL limitations | -0.008 | -0.032** |
| | (0.022) | (0.014) |
| Diagnosed with cancer | 0.009 | -0.021 |
| | (0.046) | (0.030) |
| Diagnosed with lung disease | -0.016 | 0.037 |
| | (0.063) | (0.040) |
| Diagnosed with heart disease | 0.020 | -0.023 |
| | (0.039) | (0.025) |
| Diagnosed with a stroke | -0.044 | -0.017 |
| | (0.071) | (0.046) |
| Diagnosed with diabetes | 0.003 | 0.014 |
| | (0.043) | (0.027) |
| Diagnosed with hypertension | -0.062* | -0.011 |
| | (0.033) | (0.021) |
| Self-rated health is fair or poor | -0.041 | -0.014 |
| | (0.026) | (0.017) |
| Constant | 11.267*** | 9.984*** |
| | (0.896) | (0.573) |
| N | 20,913 | 20,799 |

* p<0.10, ** p<0.05, *** p<0.01

Data source: HRS 1992-2004
* Among those with positive household income

**Table A 2-6. Individual fixed effects model of retirement effect on hours worked per week**

| | **Hours worked per week** |
|---|---|
| Retired | -35.821*** |
| | (0.369) |
| Retired*strenuous occupation | -0.365 |
| | (0.440) |
| Age | -0.212 |
| | (0.358) |
| Age squared | -0.001 |
| | (0.003) |
| Number of IADL limitations | -0.759** |
| | (0.368) |
| Number of ADL limitations | -0.202 |
| | (0.258) |
| Diagnosed with cancer | -0.685 |
| | (0.552) |
| Diagnosed with lung disease | 0.343 |
| | (0.754) |
| Diagnosed with heart disease | -0.240 |
| | (0.470) |
| Diagnosed with a stroke | -1.153 |
| | (0.853) |
| Diagnosed with diabetes | -0.611 |
| | (0.514) |
| Diagnosed with hypertension | -0.042 |
| | (0.391) |
| Self-rated health is fair or poor | -1.290*** |
| | (0.309) |
| Constant | 63.394*** |
| | (10.720) |
| N | 20,631 |

* p<0.10, ** p<0.05, *** p<0.01

Data source: HRS 1992-2004.
Notes: "Hours worked per week" includes both hours worked per week on the main job and hours worked per week on the 2nd job.

**Table A 2-7. Cases of attrition by reason**

| Reasons for attrition | Number of cases |
|---|---|
| Interviewed but with non-absorbing retirement status | 953 |
| Interviewed but with missing values in dependent or independent variables | 172 |
| Not interviewed but alive | 1,659 |
| Died | 1,066 |
| Not interviewed and mortality status unknown | 488 |
| Total | 4,338 |

Data source: Health and Retirement Study 1992-2004

**Table A 2-8. Compare fixed effects estimation using the whole sample versus the balanced sample**

| Dependent variable: BMI | Model specification I | | Model specification II | |
|---|---|---|---|---|
| | Whole sample | Balanced only | Whole sample | Balanced only |
| Retired | -0.301*** | -0.328*** | | |
| | (0.062) | (0.075) | | |
| Retired*strenuous occupation | 0.321*** | 0.335*** | | |
| | (0.074) | (0.090) | | |
| Waves since retire | | | -0.139*** | -0.118*** |
| | | | (0.026) | (0.030) |
| Waves since retire*strenuous job | | | 0.154*** | 0.133*** |
| | | | (0.027) | (0.031) |
| Age | 0.329*** | 0.298*** | 0.269*** | 0.248*** |
| | (0.060) | (0.069) | (0.067) | (0.077) |
| Age squared | -0.002*** | -0.002*** | -0.001** | -0.001* |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| ln(income) | 0.041*** | 0.050*** | 0.041*** | 0.052*** |
| | (0.015) | (0.018) | (0.015) | (0.018) |
| Non-positive wealth | 0.104 | 0.111 | 0.092 | 0.092 |
| | (0.170) | (0.223) | (0.170) | (0.223) |
| ln(wealth) | -0.002 | 0.003 | -0.004 | 0.001 |
| | (0.015) | (0.019) | (0.015) | (0.019) |
| Self-rated health is fair or poor | 0.097* | 0.109* | 0.095* | 0.108* |
| | (0.052) | (0.063) | (0.052) | (0.063) |
| Number of IADL limitations | -0.152** | -0.217*** | -0.150** | -0.217*** |
| | (0.062) | (0.082) | (0.062) | (0.082) |
| Number of ADL limitations | 0.033 | 0.083 | 0.034 | 0.083 |
| | (0.043) | (0.055) | (0.043) | (0.055) |
| Diagnosed with cancer | -0.221** | -0.172 | -0.211** | -0.168 |
| | (0.093) | (0.109) | (0.093) | (0.109) |
| Diagnosed with lung disease | -0.150 | -0.139 | -0.158 | -0.150 |
| | (0.127) | (0.152) | (0.127) | (0.153) |
| Diagnosed with heart disease | 0.006 | -0.014 | 0.010 | -0.013 |
| | (0.079) | (0.095) | (0.079) | (0.095) |
| Diagnosed with a stroke | -0.448*** | -0.336* | -0.447*** | -0.335* |
| | (0.143) | (0.174) | (0.143) | (0.173) |
| Diagnosed with diabetes | -0.538*** | -0.547*** | -0.535*** | -0.543*** |
| | (0.086) | (0.103) | (0.086) | (0.103) |
| Diagnosed with hypertension | 0.200*** | 0.204*** | 0.199*** | 0.206*** |
| | (0.066) | (0.077) | (0.066) | (0.077) |
| Number of individuals | 3,936 | 2,288 | 3,936 | 2,288 |
| Number of observations | 20,913 | 14,141 | 20,913 | 14,141 |
| Hausman specification test for attrition bias | | | | |
| F statistic | 0.40 | | 2.25 | |
| P-value | 0.82 | | 0.32 | |

* p<0.10, ** p<0.05, *** p<0.01

Data source: Health and Retirement Study 1992-2004

**Table A 2-9. First stage of the IV regressions for retirement effects on BMI, for individuals who were obese at initial interview**

| | retired | retired* strenuous | waves of retirement | waves of retirement* strenuous |
|---|---|---|---|---|
| Age | 0.023 | 0.030* | -1.202*** | -0.744*** |
| | (0.023) | (0.017) | (0.062) | (0.061) |
| Age squared | 0.000 | -0.000 | 0.012*** | 0.007*** |
| | (0.000) | (0.000) | (0.001) | (0.001) |
| Age 62+ | 0.157*** | -0.120*** | -0.022 | -0.512*** |
| | (0.023) | (0.008) | (0.049) | (0.025) |
| Age 65+ | 0.063** | -0.090*** | -0.016 | -0.772*** |
| | (0.029) | (0.011) | (0.069) | (0.044) |
| (Age 62+)*strenuous occupation | 0.041 | 0.433*** | 0.051 | 0.920*** |
| | (0.027) | (0.018) | (0.059) | (0.041) |
| (Age 65+)*strenuous occupation | 0.039 | 0.281*** | 0.293*** | 1.542*** |
| | (0.032) | (0.021) | (0.079) | (0.058) |
| ln(income) | -0.048*** | -0.026*** | -0.071*** | -0.034*** |
| | (0.006) | (0.005) | (0.012) | (0.010) |
| Non-positive wealth | -0.047 | -0.077** | -0.308*** | -0.301*** |
| | (0.048) | (0.039) | (0.102) | (0.090) |
| ln(wealth) | -0.004 | -0.006* | -0.032*** | -0.029*** |
| | (0.004) | (0.003) | (0.009) | (0.008) |
| Self-rated health is fair or poor | 0.032** | 0.023** | -0.046 | -0.027 |
| | (0.015) | (0.012) | (0.032) | (0.026) |
| Number of IADL limitations | 0.045** | 0.023 | 0.100** | 0.036 |
| | (0.019) | (0.018) | (0.040) | (0.037) |
| Number of ADL limitations | 0.046*** | 0.030*** | 0.066** | 0.037 |
| | (0.013) | (0.011) | (0.027) | (0.026) |
| Diagnosed with cancer | 0.025 | 0.015 | 0.203*** | 0.102 |
| | (0.030) | (0.022) | (0.074) | (0.063) |
| Diagnosed with lung disease | 0.083** | 0.073** | 0.207** | 0.175** |
| | (0.036) | (0.034) | (0.095) | (0.085) |
| Diagnosed with heart disease | 0.073*** | 0.002 | 0.226*** | 0.054 |
| | (0.024) | (0.017) | (0.055) | (0.044) |
| Diagnosed with a stroke | 0.086** | 0.094*** | 0.257** | 0.237*** |
| | (0.039) | (0.036) | (0.105) | (0.091) |
| Diagnosed with diabetes | -0.017 | -0.005 | 0.037 | -0.042 |
| | (0.020) | (0.016) | (0.051) | (0.040) |
| Diagnosed with hypertension | 0.010 | 0.041** | -0.062 | 0.005 |
| | (0.022) | (0.017) | (0.045) | (0.034) |
| Number of individuals | 1,305 | 1,305 | 1,305 | 1,305 |
| Number of observations | 6,870 | 6,870 | 6,870 | 6,870 |
| Joint F-statistic of age>=62, age >= 65, and interactions with strenuous occupation | 35.439 | 468.320 | 6.670 | 510.728 |
| P value of over-identification test | 0.899 | | 0.082 | |

* p<0.10, ** p<0.05, *** p<0.01

Data source: Health and Retirement Study 1992-2004

**Table A 2-10. Second stage of the IV regressions for retirement effects on BMI, for individuals who were obese at initial interview**

| | Model I | Model II |
|---|---|---|
| Retired | 0.902 | |
| | (0.860) | |
| Retired*strenuous occupation | 1.232*** | |
| | (0.362) | |
| Waves since retire | | -0.047 |
| | | (1.156) |
| Waves since retire*strenuous occupation | | 0.398 |
| | | (0.259) |
| Age | 0.063 | 0.244 |
| | (0.185) | (1.281) |
| Age squared | -0.001 | -0.002 |
| | (0.002) | (0.012) |
| ln(income) | 0.124** | 0.061 |
| | (0.058) | (0.077) |
| Non-positive wealth | -0.491 | -0.622 |
| | (0.523) | (0.589) |
| ln(wealth) | -0.068 | -0.077 |
| | (0.045) | (0.056) |
| Self-rated health is fair or poor | 0.161 | 0.223 |
| | (0.156) | (0.151) |
| Number of ADL limitations | 0.115 | 0.178 |
| | (0.166) | (0.184) |
| Diagnosed with cancer | -0.820** | -0.875** |
| | (0.324) | (0.410) |
| Diagnosed with lung disease | 0.167 | 0.260 |
| | (0.389) | (0.376) |
| Diagnosed with heart disease | -0.052 | 0.007 |
| | (0.205) | (0.304) |
| Diagnosed with a stroke | -0.944** | -0.778 |
| | (0.412) | (0.506) |
| Diagnosed with diabetes | -0.609*** | -0.599*** |
| | (0.176) | (0.181) |
| Diagnosed with hypertension | -0.014 | -0.027 |
| | (0.174) | (0.217) |
| Number of observations | 4,650 | 4,650 |
| Number of individuals | 892 | 892 |

* p<0.10, ** p<0.05, *** p<0.01

Data source: Health and Retirement Study 1992-2004

**Table A 2-11: Specification test of retirement effects on BMI and 3+ vigorous physical activities per week**

| | BMI | | Vigorous physical activity | |
|---|---|---|---|---|
| | (1) | (2) | (1) | (2) |
| Waves before retire(negative only) | 0.103 | | -0.016 | |
| | (0.109) | | (0.035) | |
| Waves before retire* strenuous occupation | 0.036 | | -0.020 | |
| | (0.033) | | (0.016) | |
| Waves after retire(positive only) | 0.012 | | -0.014 | |
| | (0.110) | | (0.035) | |
| Waves after retire(positive only)* strenuous occupation | 0.147*** | | -0.038*** | |
| | (0.027) | | (0.012) | |
| 12 dummies of waves to retire: -5 to 6 | Output omitted | | Output omitted | |
| (waves to retire=-5)*strenuous occupation | | -0.395 | | |
| | | (0.298) | | |
| (waves to retire=-4)*strenuous occupation | | -0.046 | | |
| | | (0.196) | | |
| (waves to retire=-3)*strenuous occupation | | -0.055 | | 0.037 |
| | | (0.160) | | (0.068) |
| (waves to retire=-2)*strenuous occupation | | -0.183 | | -0.027 |
| | | (0.133) | | (0.040) |
| (waves to retire=-1)*strenuous occupation | | -0.047 | | 0.004 |
| | | (0.118) | | (0.032) |
| (waves to retire=0)*strenuous occupation | | ref. | | ref. |
| (waves to retire=1)*strenuous occupation | | 0.077 | | -0.150*** |
| | | (0.107) | | (0.031) |
| (waves to retire=2)*strenuous occupation | | 0.257** | | -0.201*** |
| | | (0.122) | | (0.037) |
| (waves to retire=3)*strenuous occupation | | 0.693*** | | -0.177*** |
| | | (0.141) | | (0.045) |
| (waves to retire=4)*strenuous occupation | | 0.457*** | | -0.137** |
| | | (0.170) | | (0.055) |
| (waves to retire=5)*strenuous occupation | | 0.750*** | | -0.141* |
| | | (0.205) | | (0.075) |
| (waves to retire=6)*strenuous occupation | | 0.425 | | (dropped) |
| | | (0.247) | | |
| Joint F statistic for strenuous * pre-retirement trend (dummies) | 1.2(p=0.27) | 0.66 (p=0.65) | 1.56 (p=0.21) | 0.38 (p=0.77) |
| Joint F statistic for strenuous * post-retirement trend (dummies) | 28.8(p<0.01) | 5.70(p<0.01) | 10.78(p<0.01) | 7.28 (p<0.01) |
| Number of individuals | 2437 | 2437 | 2340 | 2340 |
| Number of observations | 13,849 | 13,849 | 7,881 | 7,881 |

* p<0.10, ** p<0.05, *** p<0.01

Data source: Health and Retirement Study 1992-2004
Note: all regressions include individual fixed-effects. All regressions also include the following covariates: age, age squared, ln(income), non-positive wealth, ln(wealth), self-rated health is fair or poor, number of IADL limitations, number of ADL limitations, diagnosed with cancer, diagnosed with lung disease, diagnosed with heart disease, diagnosed with stroke, diagnosed with diabetes, and diagnosed with hypertension.

**Table A 2-12: First-difference with IV regression results for the effect of retirement on ever diagnosed with diabetes**

| Covariates | Specification (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Panel A: no correction for attrition bias | | | | |
| Retired | -0.017 | 0.002 | | |
| | (0.030) | (0.042) | | |
| Retired*strenuous occupation | 0.006 | -0.022 | | |
| | (0.024) | (0.048) | | |
| Waves since retire | | | 0.040 | 0.029 |
| | | | (0.054) | (0.090) |
| Waves since retire*strenuous occupation | | | -0.002 | 0.011 |
| | | | (0.013) | (0.101) |
| Age | -0.023** | -0.026* | 0.037 | 0.016 |
| | (0.011) | (0.014) | (0.074) | (0.121) |
| Age squared | 0.000*** | 0.000** | -0.000 | -0.000 |
| | (0.000) | (0.000) | (0.001) | (0.001) |
| Age*strenuous occupation | | 0.008 | | 0.030 |
| | | (0.018) | | (0.133) |
| Age squared*strenuous occupation | | -0.000 | | -0.000 |
| | | (0.000) | | (0.001) |
| Number of individuals | 16,412 | 16,412 | 16,412 | 16,412 |
| Panel B: with correction for attrition bias | | | | |
| Retired | -0.010 | 0.012 | | |
| | (0.031) | (0.043) | | |
| Retired*strenuous occupation | 0.007 | -0.026 | | |
| | (0.024) | (0.048) | | |
| Waves since retire | | | 0.046 | 0.039 |
| | | | (0.053) | (0.090) |
| Waves since retire*strenuous occupation | | | -0.002 | 0.004 |
| | | | (0.013) | (0.101) |
| Age | -0.017 | -0.020 | 0.050 | 0.034 |
| | (0.012) | (0.015) | (0.072) | (0.119) |
| Age squared | 0.000 | 0.000 | -0.000 | -0.000 |
| | (0.000) | (0.000) | (0.001) | (0.001) |
| Age*strenuous occupation | | 0.008 | | 0.023 |
| | | (0.018) | | (0.133) |
| Age squared*strenuous occupation | | -0.000 | | -0.000 |
| | | (0.000) | | (0.001) |
| Number of individuals | 16,412 | 16,412 | 16,412 | 16,412 |
| * p<0.10, ** p<0.05, *** p<0.01 | | | | |

Data source: Health and Retirement Study 1992-2004
Note: All models control for household income and household wealth

**Table A 2-13: Fixed effects regression results for the effect of retirement on ever diagnosed with hypertension**

| Covariates | Specification (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Panel A: whole sample | | | | |
| Retired | -0.028*** | -0.024*** | | |
| | (0.007) | (0.008) | | |
| Retired*strenuous occupation | 0.036*** | 0.028** | | |
| | (0.009) | (0.011) | | |
| Waves since retire | | | -0.010*** | -0.007** |
| | | | (0.003) | (0.004) |
| Waves since retire*strenuous occupation | | | 0.012*** | 0.007 |
| | | | (0.003) | (0.005) |
| Age | 0.032*** | 0.038*** | 0.028*** | 0.031*** |
| | (0.007) | (0.010) | (0.008) | (0.011) |
| Age squared | -0.000 | -0.000 | -0.000 | -0.000 |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| Age*strenuous occupation | | -0.012 | | -0.007 |
| | | (0.014) | | (0.016) |
| Age squared*strenuous occupation | | 0.000 | | 0.000 |
| | | (0.000) | | (0.000) |
| Number of individuals | 20,913 | 20,913 | 20,913 | 20,913 |
| Panel B: balanced sample only | | | | |
| Retired | -0.036*** | -0.028*** | | |
| | (0.009) | (0.010) | | |
| Retired*strenuous occupation | 0.044*** | 0.029** | | |
| | (0.011) | (0.014) | | |
| Waves since retire | | | -0.011*** | -0.006 |
| | | | (0.004) | (0.004) |
| Waves since retire*strenuous occupation | | | 0.013*** | 0.003 |
| | | | (0.004) | (0.006) |
| Age | 0.033*** | 0.039*** | 0.029*** | 0.033*** |
| | (0.008) | (0.012) | (0.009) | (0.013) |
| Age squared | -0.000 | -0.000 | -0.000 | -0.000 |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| Age*strenuous occupation | | -0.011 | | -0.012 |
| | | (0.016) | | (0.019) |
| Age squared*strenuous occupation | | 0.000 | | 0.000 |
| | | (0.000) | | (0.000) |
| Number of individuals | 14,141 | 14,141 | 14,141 | 14,141 |
| Hausman specification test of attrition bias | | | | |
| F-statistic | 6.4 | 4.9 | 8.3 | 7.8 |
| P-value | 0.499 | 0.667 | 0.503 | 0.556 |

* p<0.10, ** p<0.05, *** p<0.01

Data source: Health and Retirement Study 1992-2004

Note: All models control for household income and household wealth

**Table A 2-14: Fixed effects with IV regression results for the effect of retirement on ever diagnosed with hypertension**

| Covariates | Specification (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Panel A: whole sample | | | | |
| Retired | -0.056 | -0.050 | | |
| | (0.036) | (0.059) | | |
| Retired*strenuous occupation | 0.052*** | 0.045 | | |
| | (0.017) | (0.069) | | |
| Waves since retire | | | 0.021 | 0.083 |
| | | | (0.070) | (0.141) |
| Waves since retire*strenuous occupation | | | 0.011 | -0.061 |
| | | | (0.011) | (0.160) |
| Age | 0.029*** | 0.035** | 0.071 | 0.147 |
| | (0.010) | (0.015) | (0.092) | (0.182) |
| Age squared | -0.000 | -0.000 | -0.000 | -0.001 |
| | (0.000) | (0.000) | (0.001) | (0.002) |
| Age*strenuous occupation | | -0.010 | | -0.088 |
| | | (0.020) | | (0.217) |
| Age squared*strenuous occupation | | 0.000 | | 0.001 |
| | | (0.000) | | (0.002) |
| Number of individuals | 20,913 | 20,913 | 20,913 | 20,913 |
| Panel B: balanced sample only | | | | |
| Retired | -0.085* | -0.050 | | |
| | (0.045) | (0.068) | | |
| Retired*strenuous occupation | 0.065*** | 0.008 | | |
| | (0.020) | (0.084) | | |
| Waves since retire | | | 0.039 | 0.103 |
| | | | (0.086) | (0.130) |
| Waves since retire*strenuous occupation | | | 0.011 | -0.122 |
| | | | (0.012) | (0.167) |
| Age | 0.028** | 0.035** | 0.100 | 0.185 |
| | (0.011) | (0.017) | (0.118) | (0.181) |
| Age squared | -0.000 | -0.000 | -0.001 | -0.002 |
| | (0.000) | (0.000) | (0.001) | (0.002) |
| Age*strenuous occupation | | -0.012 | | -0.188 |
| | | (0.023) | | (0.246) |
| Age squared*strenuous occupation | | 0.000 | | 0.002 |
| | | (0.000) | | (0.002) |
| Number of individuals | 14,141 | 14,141 | 14,141 | 14,141 |
| Hausman specification test of attrition bias | | | | |
| F-statistic | 6.3 | 3.8 | 9.6 | 2.8 |
| P-value | 0.505 | 0.802 | 0.388 | 0.971 |

* p<0.10, ** p<0.05, *** p<0.01

Data source: Health and Retirement Study 1992-2004

Note: All models control for household income and household wealth

**Table A 2-15: First-difference regression results for the effect of retirement on ever diagnosed with hypertension**

| Covariates | Specification (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Panel A: no correction for attrition bias | | | | |
| Retired | -0.018*** | -0.018*** | | |
| | (0.007) | (0.007) | | |
| Retired*strenuous occupation | 0.025** | 0.023** | | |
| | (0.010) | (0.011) | | |
| Waves since retire | | | -0.007 | -0.005 |
| | | | (0.005) | (0.006) |
| Waves since retire*strenuous occupation | | | 0.008 | 0.005 |
| | | | (0.006) | (0.009) |
| Age | 0.001 | 0.004 | -0.002 | -0.001 |
| | (0.014) | (0.020) | (0.016) | (0.022) |
| Age squared | 0.000 | 0.000 | 0.000 | 0.000 |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| Age*strenuous occupation | | -0.006 | | -0.002 |
| | | (0.026) | | (0.029) |
| Age squared*strenuous occupation | | 0.000 | | 0.000 |
| | | (0.000) | | (0.000) |
| Number of individuals | 16,412 | 16,412 | 16,412 | 16,412 |
| Panel B: with correction for attrition bias | | | | |
| Retired | -0.017*** | -0.016** | | |
| | (0.006) | (0.007) | | |
| Retired*strenuous occupation | 0.026*** | 0.023** | | |
| | (0.010) | (0.011) | | |
| Waves since retire | | | -0.008 | -0.006 |
| | | | (0.005) | (0.006) |
| Waves since retire*strenuous occupation | | | 0.009 | 0.005 |
| | | | (0.006) | (0.009) |
| Age | 0.018 | 0.021 | 0.014 | 0.014 |
| | (0.015) | (0.021) | (0.017) | (0.022) |
| Age squared | -0.000 | -0.000 | -0.000 | -0.000 |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| Age*strenuous occupation | | -0.005 | | -0.001 |
| | | (0.026) | | (0.028) |
| Age squared*strenuous occupation | | 0.000 | | 0.000 |
| | | (0.000) | | (0.000) |
| Number of individuals | 16,412 | 16,412 | 16,412 | 16,412 |

* p<0.10, ** p<0.05, *** p<0.01

Data source: Health and Retirement Study 1992-2004

Note: All models control for household income and household wealth

**Table A 2-16: First-difference with IV regression results for the effect of retirement on ever diagnosed with hypertension**

| Covariates | Specification (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Panel A: no correction for attrition bias | | | | |
| Retired | -0.101*** | -0.110** | | |
| | (0.036) | (0.051) | | |
| Retired*strenuous occupation | 0.056* | 0.071 | | |
| | (0.029) | (0.060) | | |
| Waves since retire | | | -0.081 | -0.122 |
| | | | (0.066) | (0.114) |
| Waves since retire*strenuous occupation | | | 0.027* | 0.093 |
| | | | (0.015) | (0.128) |
| Age | -0.008 | -0.007 | -0.100 | -0.160 |
| | (0.014) | (0.021) | (0.093) | (0.154) |
| Age squared | 0.000 | 0.000 | 0.001 | 0.001 |
| | (0.000) | (0.000) | (0.001) | (0.001) |
| Age*strenuous occupation | | -0.001 | | 0.102 |
| | | (0.026) | | (0.168) |
| Age squared*strenuous occupation | | -0.000 | | -0.001 |
| | | (0.000) | | (0.002) |
| Number of individuals | 16,412 | 16,412 | 16,412 | 16,412 |
| Panel B: with correction for attrition bias | | | | |
| Retired | -0.088** | -0.093* | | |
| | (0.037) | (0.051) | | |
| Retired*strenuous occupation | 0.058** | 0.065 | | |
| | (0.028) | (0.060) | | |
| Waves since retire | | | -0.074 | -0.112 |
| | | | (0.065) | (0.111) |
| Waves since retire*strenuous occupation | | | 0.028* | 0.090 |
| | | | (0.015) | (0.126) |
| Age | 0.009 | 0.009 | -0.068 | -0.127 |
| | (0.016) | (0.022) | (0.089) | (0.149) |
| Age squared | -0.000 | -0.000 | 0.001 | 0.001 |
| | (0.000) | (0.000) | (0.001) | (0.001) |
| Age*strenuous occupation | | -0.001 | | 0.102 |
| | | (0.026) | | (0.167) |
| Age squared*strenuous occupation | | 0.000 | | -0.001 |
| | | (0.000) | | (0.002) |
| Number of individuals | 16,412 | 16,412 | 16,412 | 16,412 |

* p<0.10, ** p<0.05, *** p<0.01

Data source: Health and Retirement Study 1992-2004

Note: All models control for household income and household wealth

## ***Appendix B.*** *OLS Regressions of Measured Weight/height against Reported Weight/height*

**Table B 2-1. OLS regressions of measured weight/height against self-reported weight/height, among males aged 52-73**

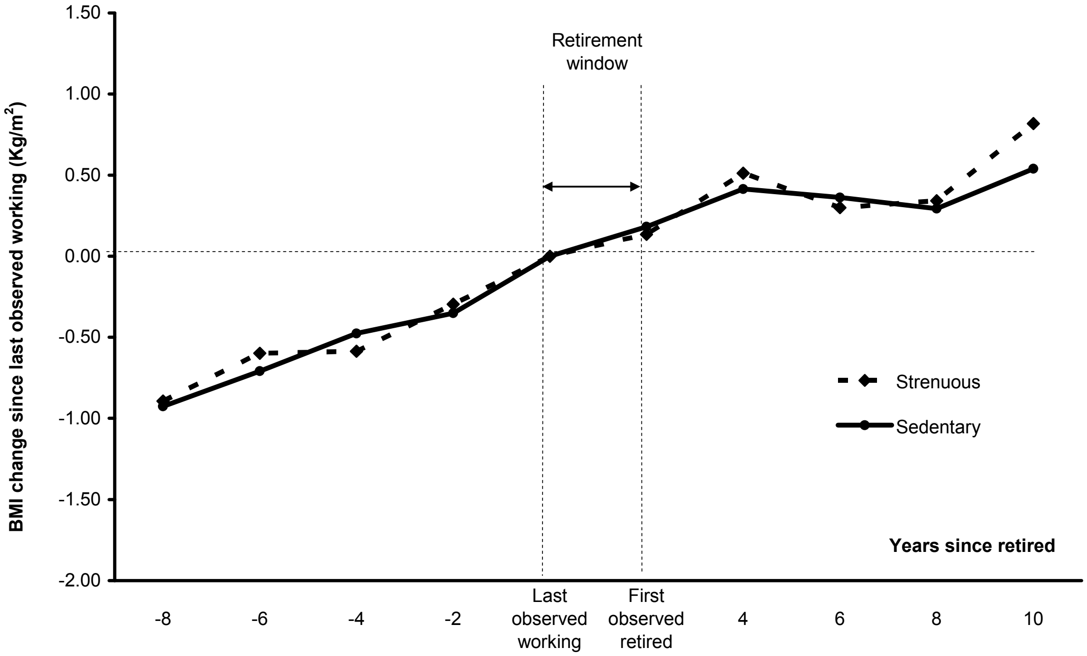
| | Non-hispanic white | | Non-hispanic black | | Hispanic | | Other | |
|---|---|---|---|---|---|---|---|---|
| | no constant | with constant | no constant | with constant | no constant | with constant | no constant | with constant |
| **Weight (Kgs)** | | | | | | | | |
| Self-reported weight | 0.85458*** | -0.65228** | 1.00388*** | 0.95523*** | 1.05346*** | 1.53617*** | 0.98195*** | 0.24352 |
| | (0.02890) | (0.26129) | (0.02621) | (0.25067) | (0.02517) | (0.21503) | (0.03419) | (0.35586) |
| Self-reported weight squared | 0.00187*** | 0.00979*** | -0.00000 | 0.00026 | -0.00047* | -0.00319** | 0.00028 | 0.00448** |
| | (0.00030) | (0.00140) | (0.00027) | (0.00138) | (0.00028) | (0.00124) | (0.00039) | (0.00205) |
| Constant | | 69.69147*** | | 2.18538 | | -20.85333** | | 31.36070** |
| | | (12.01198) | | (11.19848) | | (9.22691) | | (15.04796) |
| Adjusted R square | 0.968 | 0.529 | 0.996 | 0.873 | 0.998 | 0.912 | 0.998 | 0.953 |
| N | 1,565 | 1,565 | 240 | 240 | 161 | 161 | 48 | 48 |
| **Height (Meters)** | | | | | | | | |
| Self-reported height | 1.05773*** | -1.59108*** | 1.14252*** | 1.46816 | 1.06206*** | 5.45421* | 1.56881*** | 8.57784*** |
| | (0.01593) | (0.55868) | (0.03468) | (1.11040) | (0.07841) | (2.96718) | (0.09220) | (1.98111) |
| Self-reported height squared | -0.03956*** | 0.70656*** | -0.08773*** | -0.17962 | -0.04310 | -1.30689 | -0.33635*** | -2.27914*** |
| | (0.00893) | (0.15756) | (0.01942) | (0.31377) | (0.04556) | (0.85468) | (0.05226) | (0.55064) |
| Constant | | 2.34744*** | | -0.28795 | | -3.80976 | | -6.30324*** |
| | | (0.49492) | | (0.98140) | | (2.57285) | | (1.78007) |
| Adjusted R square | 0.999 | 0.686 | 0.999 | 0.712 | 0.998 | 0.453 | 0.999 | 0.428 |
| N | 1,567 | 1,567 | 240 | 240 | 162 | 162 | 48 | 48 |

* p<0.10, ** p<0.05, *** p<0.01

Data source: Health and Retirement Study, 2006

## ***Appendix C.*** *Figures and Regression Results for Female*

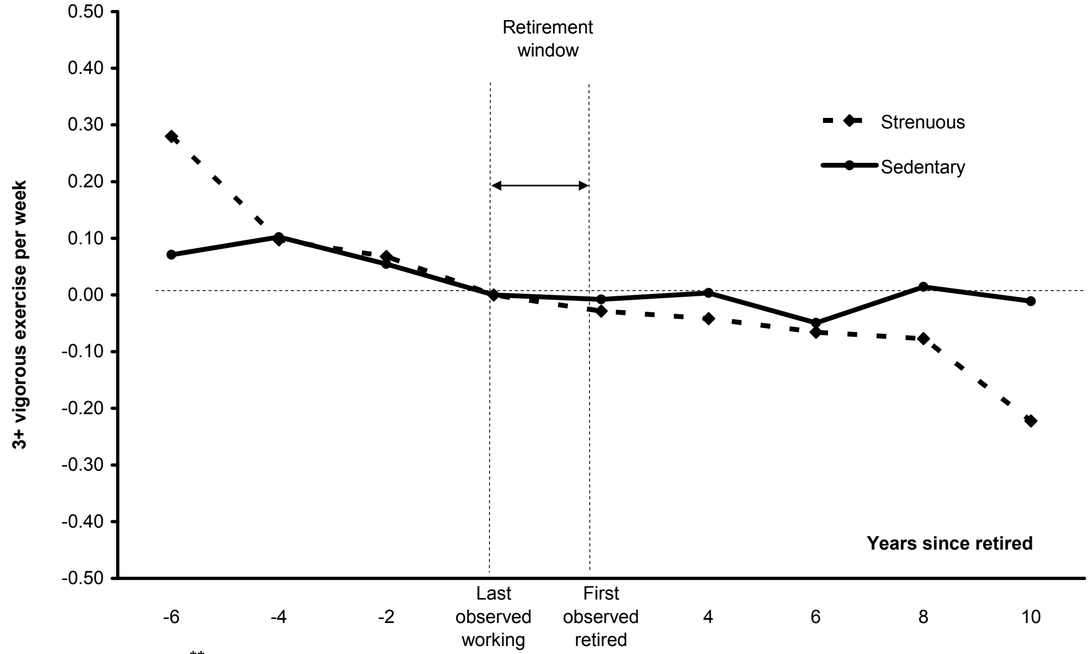
**Figure C 2-1: Unadjusted change in BMI since last observed working for females by occupation**




| | -8 | -6 | -4 | -2 | Last observed working | First observed retired | 4 | 6 | 8 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Strenuous | -0.89 | -0.60 | -0.59 | -0.30 | 0.00 | 0.13 | 0.51 | 0.30 | 0.34 | 0.82 |
| N | 164 | 252 | 371 | 500 | 650 | 483 | 353 | 239 | 152 | |
| Sedentary | -0.93 | -0.71 | -0.48 | -0.35 | 0.00 | 0.18 | 0.41 | 0.36 | 0.29 | 0.54 |
| N | 271 | 433 | 618 | 851 | 1065 | 795 | 539 | 358 | 215 | |

Source: Health and Retirement Study, 1992 to 2004.
Notes: The figure is based on the women in our sample who retire at some point during the HRS observation window. Since we can define a retirement date for each such person, we can also calculate time until or since retirement. This figure calculates cumulative changes in mean BMI since last observed working, as a function of time from retirement, for females in different occupations. The table at the bottom of the figure shows the point estimates for cumulative changes in BMI from the retirement date, as well as the relevant sample sizes. "**" indicates that the change of BMI for females in strenuous occupation at a certain time since retirement is significantly different from that in sedentary occupation, at the 5% significance level.

**Figure C 2-2: Unadjusted change in percentage taking vigorous exercise since last observed working for females by occupation**




| | -6 | -4 | -2 | Last observed working | First observed retired | 4 | 6 | 8 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| Strenuous | 0.28** | 0.10 | 0.07 | 0.00 | -0.03 | -0.04 | -0.07 | -0.08 | -0.22 |
| N | 63 | 155 | 249 | 321 | 305 | 260 | 163 | 63 | |
| Sedentary | 0.07** | 0.10 | 0.05 | 0.00 | -0.01 | 0.00 | -0.05 | 0.01 | -0.01 |
| N | 113 | 288 | 480 | 528 | 485 | 396 | 236 | 96 | |

Source: Health and Retirement Study, 1992 to 2004.
Notes: The figure is based on the women in our sample who retire at some point during the HRS observation window. Since we can define a retirement date for each such man, we can also determine time until or since retirement. This figure calculates cumulative changes in percentage of individuals taking vigorous exercise since last observed working, as a function of time, for females in different occupations. The table at the bottom of the figure shows the cumulative changes, and the number of persons that are interviewed both at a certain wave since retire and at the next adjacent wave, for females in different occupation. "**" indicates that the change of percentage points from last wave in strenuous occupation is significantly different from that in sedentary occupation, at the 5% significance level.

**Table C 2-1: OLS[(a)] and fixed effects regression results for effect of retirement on weight for females, by occupation**

| Dependent variable: BMI | Specification | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | OLS | | | | Individual fixed effects | | | |
| Retired | 0.789*** | 0.124 | | | 0.027 | 0.023 | | |
| | (0.305) | (0.254) | | | (0.072) | (0.072) | | |
| Retired*strenuous occupation | 0.434 | 0.060 | | | 0.121 | 0.136 | | |
| | (0.417) | (0.398) | | | (0.103) | (0.103) | | |
| Waves since retire | | | 0.282** | 0.034 | | | -0.020 | -0.010 |
| | | | (0.115) | (0.101) | | | (0.032) | (0.032) |
| Waves since retire*strenuous job | | | 0.197 | 0.059 | | | 0.038 | 0.049 |
| | | | (0.138) | (0.128) | | | (0.036) | (0.036) |
| Age | 0.572** | 0.372 | 0.950*** | 0.433 | 0.593*** | 0.552*** | 0.574*** | 0.554*** |
| | (0.273) | (0.255) | (0.319) | (0.293) | (0.079) | (0.079) | (0.090) | (0.090) |
| Age squared | -0.005** | -0.004* | -0.008*** | -0.004* | -0.004*** | -0.004*** | -0.004*** | -0.004*** |
| | (0.002) | (0.002) | (0.003) | (0.002) | (0.001) | (0.001) | (0.001) | (0.001) |
| ln(income) | -0.174* | -0.054 | -0.188** | -0.056 | 0.020 | 0.022 | 0.016 | 0.018 |
| | (0.091) | (0.077) | (0.090) | (0.076) | (0.021) | (0.021) | (0.021) | (0.021) |
| Non-positive wealth | -2.737*** | -2.155*** | -2.685*** | -2.147*** | 0.065 | 0.038 | 0.066 | 0.044 |
| | (0.781) | (0.702) | (0.778) | (0.701) | (0.189) | (0.189) | (0.189) | (0.189) |
| ln(wealth) | -0.373*** | -0.242*** | -0.367*** | -0.241*** | -0.008 | -0.009 | -0.008 | -0.008 |
| | (0.065) | (0.057) | (0.064) | (0.057) | (0.017) | (0.017) | (0.017) | (0.017) |
| Control for health-related variables[(b)] | No | Yes | No | Yes | No | Yes | No | Yes |
| No. individuals | 2,999 | 2,999 | 2,999 | 2,999 | 2,999 | 2,999 | 2,999 | 2,999 |
| No. of observations | 16,121 | 16,121 | 16,121 | 16,121 | 16,121 | 16,121 | 16,121 | 16,121 |
| Retired + Retired * strenuous occupation | 1.223 | 0.183 | | | 0.148 | 0.159 | | |
| P value of the Wald test of Retired + Retired * strenuous occupation = 0 | 0.002 | 0.626 | | | 0.116 | 0.094 | | |
| Waves since retire + Waves since retire*strenuous job | | | 0.480 | 0.093 | | | 0.018 | 0.038 |
| P value of the Wald test of Waves since retire + Waves since retire*strenuous job = 0 | | | 0.001 | 0.480 | | | 0.629 | 0.308 |

* p<0.10, ** p<0.05, *** p<0.01

Notes:

[(a)] OLS denotes ordinary least squares. all OLS regressions also include dummy variables for: strenuous occupation, white, Hispanic, less than high school degree, high school degree, and some college without degree, census region is Northeast, Middle west, and West.

[(b)] Health-related variables include number of IADL limitations, number of ADL limitations, self-rated health is fair/poor, ever being diagnosed with the following six chronic conditions: cancer, diabetes, heart disease, hypertension, lung disease, and stroke.

**Table C 2-2: First-stage IV estimates[(a)] on retirement for famales of turning age 62 and 65**

| | Specification | | | |
|---|---|---|---|---|
| | (1) | | (2) | |
| | **retired** | **retired* strenuous** | **waves of retirement** | **waves of retirement* strenuous** |
| Age | -0.005 | 0.000 | -1.420*** | -0.524*** |
| | (0.014) | (0.008) | (0.036) | (0.034) |
| Age squared | 0.000*** | 0.000 | 0.013*** | 0.005*** |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| Age >= 62 | 0.197*** | -0.064*** | -0.035 | -0.310*** |
| | (0.013) | (0.004) | (0.026) | (0.013) |
| Age >= 65 | 0.102*** | -0.049*** | 0.105*** | -0.478*** |
| | (0.016) | (0.005) | (0.036) | (0.023) |
| (Age >= 62) * Strenuous occupation | 0.011 | 0.419*** | 0.034 | 0.878*** |
| | (0.018) | (0.014) | (0.040) | (0.034) |
| (Age >= 65) * Strenuous occupation | 0.022 | 0.296*** | 0.096** | 1.544*** |
| | (0.020) | (0.016) | (0.048) | (0.043) |
| ln(income) | -0.046*** | -0.019*** | -0.067*** | -0.017*** |
| | (0.004) | (0.003) | (0.008) | (0.005) |
| Non-positive wealth | 0.044 | -0.001 | -0.164** | -0.106** |
| | (0.028) | (0.022) | (0.064) | (0.052) |
| ln(wealth) | 0.003 | -0.001 | -0.025*** | -0.018*** |
| | (0.003) | (0.002) | (0.006) | (0.005) |
| Control for health-related variables(b) | NO | NO | NO | NO |
| No. of observations | 2,999 | 2,999 | 2,999 | 2,999 |
| No. individuals | 16,121 | 16,121 | 16,121 | 16,121 |
| Joint F-statistic of age>=62, age >= 65, and interactions with strenuous occupation | 110.326 | 787.303 | 7.558 | 839.227 |
| P value of over-identification test | 0.963 | | 0.257 | |

* p<0.10, ** p<0.05, *** p<0.01

Notes:

[(a)] Models include individual fixed-effects. Age>=62, Age>=65, and interactions are instruments for retirement and its interaction.

[(b)] Health-related variables include number of IADL limitations, number of ADL limitations, self-rated health is fair/poor, ever being diagnosed with the following six chronic conditions: cancer, diabetes, heart disease, hypertension, lung disease, and stroke.

**Table C 2-3: Second-stage IV estimates[(a)] of retirement on weight for females, by occupation**

| Dependent variable: BMI | Specification | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Retired | 0.519 | | 0.446 | |
| | (0.340) | | (0.332) | |
| Retired*strenuous occupation | 0.046 | | 0.091 | |
| | (0.175) | | (0.175) | |
| Waves since retire | | 0.176 | | 0.303 |
| | | (0.635) | | (0.634) |
| Waves since retire*strenuous occupation | | 0.007 | | 0.016 |
| | | (0.063) | | (0.060) |
| Age | 0.648*** | 0.865 | 0.599*** | 1.013 |
| | (0.102) | (0.953) | (0.102) | (0.947) |
| Age squared | -0.005*** | -0.007 | -0.004*** | -0.008 |
| | (0.001) | (0.009) | (0.001) | (0.009) |
| ln(income) | 0.042 | 0.029 | 0.040 | 0.038 |
| | (0.030) | (0.050) | (0.029) | (0.048) |
| Non-positive wealth | 0.044 | 0.141 | 0.017 | 0.126 |
| | (0.284) | (0.300) | (0.284) | (0.297) |
| ln(wealth) | -0.009 | -0.000 | -0.011 | 0.001 |
| | (0.024) | (0.029) | (0.024) | (0.028) |
| Control for health-related variables[(b)] | NO | NO | YES | YES |
| No. individuals | 2,999 | 2,999 | 2,999 | 2,999 |
| No. observations | 16,121 | 16,121 | 16,121 | 16,121 |
| Retired + Retired * strenuous occupation | 0.565 | | 0.537 | |
| P value of the Wald test of Retired + Retired * strenuous occupation = 0 | 0.095 | | 0.109 | |
| Waves since retire + Waves since retire*strenuous job | | 0.183 | | 0.319 |
| P value of the Wald test of Waves since retire + Waves since retire*strenuous job = 0 | | 0.762 | | 0.602 |

* p<0.10, ** p<0.05, *** p<0.01

Notes:
[(a)] Models include individual fixed-effects. Age>=62, Age>=65, and interactions are instruments for retirement and its interaction.
[(b)] Health-related variables include number of IADL limitations, number of ADL limitations, self-rated health is fair/poor, ever being diagnosed with the following six chronic conditions: cancer, diabetes, heart disease, hypertension, lung disease, and stroke.

**Table C 2-4: Regression results for effect of retirement on exercise for females, by occupation**

| Dependent variable: 3+ vigorous exercise per wk | Specification | | | | | |
|---|---|---|---|---|---|---|
| | OLS[a] | | Individual fixed effects | | Individual fixed effects with IV | |
| Retired | 0.013 | 0.049** | 0.014 | 0.018 | 0.079 | 0.054 |
| | (0.020) | (0.020) | (0.021) | (0.021) | (0.113) | (0.110) |
| Retired*strenuous occupation | -0.124*** | -0.098*** | -0.033 | -0.019 | -0.094 | -0.087 |
| | (0.031) | (0.030) | (0.034) | (0.034) | (0.072) | (0.072) |
| Age | -0.042 | -0.032 | -0.004 | -0.003 | -0.011 | -0.013 |
| | (0.034) | (0.033) | (0.030) | (0.030) | (0.038) | (0.038) |
| Age squared | 0.000 | 0.000 | -0.000 | -0.000 | -0.000 | 0.000 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| ln(income) | 0.014** | 0.007 | 0.004 | 0.003 | 0.005 | 0.003 |
| | (0.007) | (0.007) | (0.006) | (0.006) | (0.008) | (0.008) |
| Non-positive wealth | 0.163*** | 0.131** | 0.114** | 0.101** | 0.108* | 0.096 |
| | (0.054) | (0.051) | (0.051) | (0.051) | (0.063) | (0.062) |
| ln(wealth) | 0.019*** | 0.012*** | 0.005 | 0.004 | 0.005 | 0.004 |
| | (0.004) | (0.004) | (0.005) | (0.005) | (0.006) | (0.006) |
| Control for health-related variables[b] | No | Yes | No | Yes | No | Yes |
| No. individuals | 2,592 | 2,592 | 2,592 | 2,592 | 2,592 | 2,592 |
| No. of observations | 9,265 | 9,265 | 9,265 | 9,265 | 9,265 | 9,265 |
| Retired + Retired * strenuous occupation | -0.110 | -0.049 | -0.019 | -0.001 | -0.015 | -0.033 |
| P value of the Wald test of Retired + Retired * strenuous occupation = 0 | 0.000 | 0.054 | 0.525 | 0.966 | 0.891 | 0.755 |

* p<0.10, ** p<0.05, *** p<0.01

Notes:
[a] OLS denotes ordinary least squares. all OLS regressions also include dummy variables for: strenuous occupation, white, Hispanic, less than high school degree, high school degree, and some college without degree, census region is Northeast, Middle west, and West.
[b] Health-related variables include number of IADL limitations, number of ADL limitations, self-rated health is fair/poor, ever been diagnosed with the following six chronic conditions: cancer, diabetes, heart disease, hypertension, lung disease, and stroke.

# Chapter 3 Food price and body weight among older Americans[14]

## *Background*

In the United States, the obesity rate has doubled among adults between 1980 and 2004. During year 2005-2006, more than one third of the adult population (aged 20 and over) is obese. The Center for Disease Control declared that obesity is one of its top public health priorities (Nestle and Jacobson 2000).

Obesity indicates excessive energy intake relative to energy expenditure. Among the various government interventions proposed to deal with the obesity epidemic, one initiative is to impose a food tax or "fat tax" on selected foods that are high in energy, fat, but are low in other nutrients (Jacobson and Brownell 2000; Nestle and Jacobson 2000).

Until year 2000 there were 19 states and cities in the United States that imposed taxes on less nutritious foods, like soft drinks, sweets, or snack foods (Jacobson and Brownell 2000). The major purpose is to raise revenue rather than to improve health. Usually the revenue goes to general funds. Some is ear-marked to specific purposes, like violence prevention (Washington), funding Medicaid (Arkansas), or funding medical schools (West Virginia). Such taxes were strongly opposed by soft drink and food industries and 12 cities, counties, or states have reduced or repealed such taxes in recent years. Despite arguments for or against food taxes, existing literature examining the association between food price and obesity is limited.

The most rigorous evidence would come from randomized control trials. Imagine we take a representative sample of the population of interest, assign them with different food prices, record their food consumption and body weight, and follow them up for a sufficiently long period of time. Then we examine how food consumption and body weight vary by food prices. So far there is no such experiment done, and the existing experiments regarding food price and food consumption are very few, narrowly focused and didn't use a randomized design. One study conducted a non-randomized experiment in a University looking at relative food prices of low-fat snacks in vending machines (French 2003). The author found that sales of low-fat snacks from vending machines

[14] This paper is co-authored with Dana Goldman and Darius Lakdawalla

increased significantly when prices were lowered, while the overall sales volume didn't change, meaning that the sales of fatty snacks must have decreased. However, the experiment cannot track whether students' total food consumption or weight changed.

There are a series of empirical study examining what have contributed to the obesity epidemic in the United States. One study examined the relationship between relative food price and weight among young or middle-aged Americans (Lakdawalla and Philipson 2002). The authors used the National Longitudinal Survey of Youth (NLSY), inter-city prices from the American Chamber of Commerce Researchers Association (ACCRA), as well as data from the Bureau of Labor Statistics (BLS) on price variation over time, within particular cities. They used the state-level difference in food-sales tax and non-food sales tax as the instrument for relative food price. They found that 1 percent increase in relative price leads to at most 0.6 percent decrease in BMI, meaning a price elasticity of 0.6. Using repeated cross-sectional data of the U.S. population, and regional price data from ACCRA and other data sources, Chou et. al. found that the real fast-food restaurant price, the real food at home price and the real full-service restaurant price were negatively associated with weight (Chou, Grossman et al. 2004). A review by the World Health Organization(Goodman and Anise 2006) claimed that there is no direct evidence of a causal relationship between policy-related economic instruments (taxes, prices, etc) and consumption of saturated fats and energy- dense foods. In another study of U.S. children from kindergarten to the third grade, the authors examined the association between food price and weight gain (Sturm and Datar 2005). They found that lower real food prices of fruits and vegetables are significantly associated with lower weight gain for children from kindergarten to third grade. They did not find significant associations for diary or fast-food prices and weight. A fourth study examines the effects of raising the relative price of non-healthy food versus healthy food. Using data from the National Health Interview Survey (NHIS) for the period of 1982-1996, and the BLS statistics on prices for different items at the census region level, the authors found very small effects: their results indicate that 100 percent tax on unhealthful foods could reduce average BMI by about 1 percent, and the same tax could reduce the incidence of being overweight and the incidence of obesity by 2 percent and 1 percent respectively (Gelbach, Klick et al. 2007). Chouinard et. al. found that imposing a tax on diary products in proportional to fat contents is ineffective in reducing fat consumption. The elasticity is lower than 0.1. Such a tax will mainly raise revenues but is highly regressive (Chouinard, Davis et al. 2007). Finally, Sheroeter et. al. showed that a tax on food away from home could actually lead to an increase in fat intake. This is because there are plenty of at-home foods that are Calorie-dense and are close substitutes of food away from home (Schroeter, Lusk et al. 2007).

In conclusion, studies on the effect of food price and body weight are limited and the conclusions are inconsistent with each other. And none of the study population is for the near-elderly and the elderly in the United States. In this study, we will examine the associations between various food prices and body weight in adults age 50 and over in the United States. This is an important population to examine because the prevalence of obesity and overweight among this population is high, and the associated health costs are enormous. We will use a longitudinal and national-representative survey of older Americans, along with data on geographical and temporal variations in food price inside the United States to identify the associations. The results will provide valuable information on the potential effects of food tax/subsidy on curtailing the obesity epidemic in the U.S.

The rest of the paper is organized as the following: Section 2 outlines the conceptual framework of analyzing the effects of food prices on body weight. Section 3 describes data and methods. Section 4 reports the results. Section 5 is the discussion.

## *Conceptual framework*

Obesity is caused by excessive energy intake, relative to energy consumption. Energy intake means eating food, while energy consumption could be achieved through various ways: work, leisure exercise, dieting, etc. Following the conceptual frameworks proposed by several previous studies (Philipson and Posner 1999; Lakdawalla and Philipson 2002; Schroeter, Lusk et al. 2007), I assume that an individual maximizing the utility by choosing hours of work (H), exercise (E), two types of food, one high-calorie food (Fh) and low calorie food (FL). She also values non-food consumption (C) and leisure time (L).

$$\max_{C,H,F_1,F_2,E,L} U(W(F_1, F_2, H, E), C, F_1, F_2, E, L) \quad \textbf{(1)}$$

Budget constraint:

$$C + p_1 F_1 + p_2 F_2 + qE \leq wH + I \quad \textbf{(2)}$$

Time constraint:

$$H + L + E \leq T \quad \textbf{(3)}$$

$p_1$, $p_2$ and q are the resource prices for $F_1, F_2$ and $E$. $w$ is the wage rate; $T$ is the total time available.

Since we're addressing the problem of "obesity", we are assuming that higher weight lowers utility, $U_W < 0$; weight is a byproduct of other goals that are valuable(Chou, Grossman et al. 2004). Taking food as an example, people value food for multiple reasons other than energy intake: nutrition, taste, palatability, etc. No food is good at all aspects – nutritious, healthy, tasty, palatable, and not too energy dense. In reality, some food provides high values in nutrition, but lacks palatability, like vegetables. Other food provides taste and palatability, but is not nutritious, like sweets and fatty food. To simplify the model, we assume that as a whole, more food intake increases utility, though at a decreasing rate:

$U_{F_1} > 0$, $U_{F_1F_1} < 0$; $U_{F_2} > 0$, $U_{F_2F_2} < 0$;

We further assume that the two types of food are supplements: $U_{F_1F_2} > 0$; $U_{F_2F_1} > 0$

And the high-calorie food has a larger impact on weight:

$W_{F_1} > W_{F_2} > 0$

Working reduces weight through work-related exercise. $W_H < 0$.

But as job becomes more sedentary, the impact is smaller: $|W_H|$ would decrease.

Exercise (or strictly speaking, time devoted to exercise), will reduce weight: $W_E < 0$.

The optimal weight: $W^*$, is a function of the optimal food consumption, work and exercise.

$$W^* = W(F_1^*, F_2^*, H^*, E^*) \quad \textbf{(4)}$$

What we are interested is how the optimal $W^*$ will shift with the change in food prices: $P_{F_1}$, and $P_{F_2}$.

Take the derivative of both sides in equation (4) against $P_{F_1}$, we obtain:

$$\frac{\partial W^*}{\partial P_{F_1}} = \frac{\partial W^*}{\partial F_1^*}\frac{\partial F_1^*}{\partial P_{F_1}} + \frac{\partial W^*}{\partial F_2^*}\frac{\partial F_2^*}{\partial P_{F_1}} + \frac{\partial W^*}{\partial H^*}\frac{\partial H^*}{\partial P_{F_1}} + \frac{\partial W^*}{\partial E^*}\frac{\partial E^*}{\partial P_{F_1}} \quad \textbf{(5)}$$

Multiply both sides by $\frac{P_{F_1}}{W^*}$, we could rewrite the above equation using elasticities:

$$\varepsilon_{W,P_{F_1}} = \varepsilon_{W,F_1}\varepsilon_{F_1F_1} + \varepsilon_{W,F_2}\varepsilon_{F_2F_1} + \varepsilon_{W,H}\varepsilon_{HF_1} + \varepsilon_{W,E}\varepsilon_{EF_1} \quad \textbf{(6)}$$

$\varepsilon_{W,P_{F_1}}$ is the percentage change in weight resulting from 1 percentage change in the price of high-calorie food. $\varepsilon_{W,F_1}$, $\varepsilon_{W,F_2}$, $\varepsilon_{W,H}$, and $\varepsilon_{W,E}$ represent percentage changes in weight resulting from 1 percentage change in $F_1^*, F_2^*, H^*, E^*$. $\varepsilon_{F_1F_1}$ is the own price elasticity of high-calorie food, while $\varepsilon_{F_2F_1}, \varepsilon_{HF_1}$ and $\varepsilon_{EF_1}$ are cross-price elasticities of $F_2^*, H^*, E^*$ versus $F_1^*$.

Similarly, we could derive the following equation:

$$\varepsilon_{W,P_{F2}} = \varepsilon_{W,F_2}\varepsilon_{F_2F_2} + \varepsilon_{W,F_1}\varepsilon_{F_1F_2} + \varepsilon_{W,H}\varepsilon_{HF_2} + \varepsilon_{W,E}\varepsilon_{EF_2} \quad (7)$$

The strategy of taxing high-calorie food or subsidizing low-calorie food will be effective only if $\varepsilon_{W,P_{F_1}} < 0$ or $\varepsilon_{W,P_{F_2}} > 0$. But as Schroeter (2007) pointed out, it is not always the case. From Equation (5), $\varepsilon_{W,P_{F_1}} < 0$ is more likely to hold as $\left|\varepsilon_{F_1F_1}\right|$ increases, $\varepsilon_{W,H}\varepsilon_{HF_1} + \varepsilon_{W,E}\varepsilon_{EF_1}$ decreases, $\varepsilon_{W,F_2} / \varepsilon_{W,F_1}$ increases, or as $\varepsilon_{F_2F_1}$ decreases. This means that raising the price of high-calorie food is more likely to lower weight if $F_1$ is price-elastic, if $F_1$ is much more calorie-dense than $F_2$, if the change in exercise and the change in hours worked doesn't induce much weight gain, and if the two foods are not close substitutes.

On the other hand, based on Equation (6), $\varepsilon_{W,P_{F_2}} > 0$ is more likely to hold if $F_2$ is price inelastic, the two foods are close-substitutes, and the change in exercise and the change in hours worked doesn't induce much weight gain.

To sum up, the effect of a food tax or subsidy will depend on own- and cross-price elasticities: in a two-food example, if the low-calorie food is price-elastic, and the two foods are not close substitutes, then a tax on high-calorie food might work. If the low-calorie food is price inelastic, and the two foods are close subsitutes, then a subsidy on low-Calorie food is more likely to work. Another caution of imposing tax on high-calorie food is to ensure that the high-Calorie food is not a "Giffen food". In reality, there are huge varieties of foods, therefore determine what food to tax and what food to subsidy is very important. Otherwise there might be unexpected consequences.

In this study we will examine the effects of certain types of foods on obesity. We will examine the effects of fast food prices and at-home food prices. These are two broad categories of foods. We will also analyze the effects of price for calorie-dense and non-nutritious food (margarine) and price for low-calorie and nutritious food (produce).

Instead of taxing certain type of food, a more reasonable approach is to impose tax in proportion to the "unhealthy" attributes of the food (Chouinard, Davis et al. 2007). In order to lower obesity and discourage energy over-intake, a tax can be imposed according to the calorie density of food. In order to know how calorie intake is sensitive to price of Calories, we will also examine the effect of price of calories on body weight.

## *Data and methods*

### Health and Retirement Study

We employed the Health and Retirement Study (HRS), a biennial survey of the population over age 50, to carry out the analysis. The original HRS cohort – first interviewed in 1992 – was a national representative sample of approximately 7,600 households (n = 12,654 individuals) with at least one member who was born between 1931-1941. In year 1998, the HRS cohort was merged with another study with similar focus but looking at the older birth cohort: the AHEAD cohort including households with at least one member who was born before year 1923. The AHEAD cohort was first interviewed in 1993 and then in 1995. Meanwhile, two new birth cohorts were added to the study in 1998: the War Baby cohort (households with at last one member born between 1942 and 1947) and the CODA cohort (households with at least one member born between 1923 and 1930). Therefore in year 1998 the HRS includes a representative sample of population born in 1947 or earlier and surveys more than 22,000 Americans. Subsequent interviews were made for every two years and data is now publicly available until year 2004. In my main analysis, the analytic sample is the original HRS cohort, since they have the most waves of data, and their age is not too old so that voluntary weight loss could be beneficial.

In every wave of interview HRS asked detailed questions about demographics, employment, occupation, income and wealth, and health insurance. Questions were also asked about self-reported general health status, prevalence and incidence of chronic conditions, functional status and disability, exercise, and self-reported body height and weight. County residence is available as restricted data and therefore we can link geographical information on food prices and taxes to HRS respondents.

Body mass index in HRS is constructed based on self-reported weight and height. Since there might be measurement errors in self-reported variables, for example, heavier persons are more likely to underreport their weight, we employ procedures developed by Cawley (1999) to correct for these errors. In the year 2006, HRS randomly selected half of the households and measured their weight and height. The self-reported weight and height is also available. As a result we can model the relationship between actual and measured weight and height. We restrict the estimation sample among those aged 52 and 73. 52 is the lowest age for an eligible HRS respondent (the lowest age for our estimation sample is 50). While 73 is the highest age for our estimation sample. We carry out regressions by regressing actual weight on reported weight, reported weight squared. And we did the regressions separately for eight sub-groups: White male non-Hispanic, white female non-Hispanic, black male non-Hispanic, black female non-Hispanic, Hispanic male, Hispanic female, other male, and other female. The regression results are shown in Table A 3-1 and Table A 3-2. We run regressions with and without constant. And we found that regressions without constants fit the data much better [IS'NT THAT USUALLY THE CASE?]. So we will use the estimates without constant to predict weight and height for individuals in our estimation sample, based on their gender, race/ethnicity, self-reported weight and height. And we re-calculate BMI based on corrected weight and height.

The means of adjusted BMI and obesity rate are both higher than the unadjusted means. For example, at baseline, without adjustment, the mean BMI is 26.78kg/m2, and the prevalence of obesity is 20.4%. After adjustment, the mean BMI is 28.03kg/m2, and the prevalence of obesity is 28.9%.

## ACCRA Price Data

To construct food indexes for food and other goods, I will use the ACCRA cost of living index, published quarterly by the American Chamber of Commerce Researchers Association (ACCRA), for more than 200 cities. The data covers year 1992 to 2003.

ACCRA collected prices for 57 different items and reports the weight of each item in the typical budget of household whose head holds a mid-management position. A list of the 57 items is shown in Table B 3-1. Since the ACCRA data was reported quarterly, we average prices over the four quarters in a given year to get annual prices. ACCRA reports prices at the level of metropolitan areas. We will use the population-weighted averages of the city prices to constructed prices at the county level. We will inflate prices in different years using the consumer price indexes for all goods, released by the Bureau of Labor Statistics (BLS).

We use the basket of “grocery goods” excluding laundry detergent, facial tissue and cigarettes to construct at-home food price. Price of fast food is constructed using the prices a MacDonald’s Quarter-Pounder with cheese, a thin crusted cheese Pizza at Pizza Hut or Pizza Inn, and a fried chicken at Kentucky Fried Chicken or Church’s. The remained items are used to construct “non-food” items, excluding cigarettes and gasoline prices. We also explore the effects of sub-groups of food: fruits and vegetables, margarine. Finally, we will construct a price index of calories. Information on Calories per unit of purchase is obtained through the USDA website of "What's In The Foods You Eat *Search Tool*"[15]. The formula is shown in equation (7).

$$price\_percal_{at} = \frac{1}{\sum_{j=1}^{27} \exp\_share_j \times \frac{Calories\_per\_purchase\_unit_j}{price_{atj}}} \qquad \textbf{(8)}$$

a : area

t : year

j : food item

$\exp\_share_j$ : share of item j in all expenditure

$Calories\_per\_purchased\_unit_j$ : Calories per unit of purchase for food j

$price_{atj}$ : Price per unit of purchase for food j at area a and time t

There are two basic types of food price indices. Laspeyres index measures the changes in the cost of a fixed basket of goods from a base period. It assumes no substitution due to relative price changes and therefore overestimates the true CPI. Paasche index weights prices by current consumption pattern, and thus likely overstates substitution and underestimates the true CPI. The two indices can be viewed as the "upper bound" and "lower bound" of the price change.

People also use one of three price indexes that stands in-between the Laspeyres index and the Paasche index. These are called superlative indices and there is no conclusion about which one is better (Hill 2006). A detailed description of the three superlative indices is in Appendix C. In this study we will construct both Laspeyres-type price indices and Paasche-type price indices. Estimates for the two sets of price indices will provide upper and lower bounds of the true price effects.

## Merge HRS data with ACCRA

We begin with 9,733 individuals who were born between 1931 and 1941, first interviewed in year 1992 and with positive sampling weight. If they didn't die or drop out of the sample, they were followed biennially until year 2004. For the analytic sample, we

[15] http://www.ars.usda.gov/Services/docs.htm?docid=17032

have to exclude a large part of the sample, due to the following reasons: first, we identify the price effects by examining how individual weight changes correlate with price changes at the county level. To take into account both the individual-level and county-level time-invariant unobservables, we only include individuals who didn't move during the period of 1992 to 2004. Those who moved to another county during this period are excluded from the analysis. Second, ACCRA does not cover all cities in the United States, and the aggregated data at the county-level don't include all counties either. As a result, when we merge the HRS data with the ACCRA data through the county identifier and year, we can only match less than half of the HRS respondents. Third, we removed observations with missing values for the independent variables and control variables that we will use in the regression analysis. Finally we exclude observations with non adjacent waves of data. For example, if an individual was interviewed in wave 1, wave 3, and wave 4, we will exclude the wave 1 data.

The detailed sample selection process is shown in Figure 3-1. 3,178 individuals are included in the analytic sample.

## Correction for sample attrition bias

Since we include only 32.7% of the initial sample (3,178 out of 9,733) in our analysis, we are concerned that the selected sample might not be representative of the study population, and that the selection is non-random in terms of the relationship between food price change and weight change. To reduce sample selection bias, we first model whether an individual in an initial sample will be included in analytic sample. We run a Probit model and regressors include demographics and health and economic conditions at the 1992 interview. We then predict the probability of sample inclusion for those in the analytic sample, and multiply the sampling weight by the inverse of the predicted probability. All descriptive and regression analysis will be carried out using the modified sampling weight. Using this procedure, we take care of selection based on observables, but not the selection based on unobservables (Wooldridge 2002).

The selection model is presented in Table 3-1. It shows that residing in rural areas greatly decreases the probability of being included in the analytic sample. This is because the

ACCRA price data is only collected in cities. Even though we aggregate data at the county level, those in rural counties are excluded. Individuals with diabetes at baseline interview are also less likely to appear in the analytic sample. In addition, non-Hispanic blacks, Hispanic, and those with less than high school education are more likely to be included. Finally, those with higher household wealth are also more likely to appear in the analytic sample.

The original sampling weight has a mean of 2,340, standard deviation of 1,048, a minimum of 563 and a maximum of 7,710. After multiplied by the inverse of the probability, the mean is raised to 5,813, the standard deviation is 3,560, the minimum is 921, while the maximum is 26,687. Descriptive statistics of the analytic sample is shown in Table 3-2.

## Econometric analysis

The effect of food price on body weight could be shown as the following:

$$W_{igt} = \beta_0 + P_{gt}\beta_1 + Z_{gt}\beta_2 + X_{igt}\beta_3 + \alpha_i + \tau_t + \varepsilon_{igt} \qquad \textbf{(9)}$$

In equation (8) $i$ indicates the ith individual, g indicates the gth geographic region (at the county level); $t$ indicates time. The dependent variable, $W_{igt}$, is body weight for individual $i$ in region g at time $t$; $P_{gt}$ represents a vector of food-price related variables, including relative price of food, and relative price of certain types of food, $X_{igt}$ is a vector of individual-level time-variant characteristics, including age, income, wealth, health status, smoking behavior, and exercise. $Z_{gt}$ is a vector of variables representing regional contextual variables. $\alpha_i$ is the individual fixed effects, and $\tau_t$ is time fixed effects. $\varepsilon_{igt}$ represents the idiosyncratic error term.

A model allows for lag effect of body weight is shown in equation (9).

$$W_{igt} = \gamma W_{ig,t-1} + \beta_0 + P_{gt}\beta_1 + Z_{gt}\beta_2 + X_{igt}\beta_3 + \alpha_i + \tau_t + \varepsilon_{igt} \qquad \textbf{(10)}$$

Equation (9) is a dynamic linear model. Fixed effects or first-difference estimation will generate inconsistent estimates .So I will use the method proposed by Arellano and Bond (Arellano and Bond 1991): first-difference equation (10), obtain Equation (11), and then use $(W_{igt1},...,W_{igt-2},\Delta Z_{gt},\Delta X_{igt},\Delta P_{gt})$ as instruments.

$$\Delta W_{igt} = \gamma \Delta W_{ig,t-1} + \beta_1 \Delta P_{gt} + \beta_2 \Delta Z_{gt} + \beta_3 \Delta X_{igt} + \Delta \tau_t + \Delta \varepsilon_{igt} \quad \textbf{(11)}$$

In using first-difference as the estimation strategy, we have to assume that $(P_{gt}, Z_{gt}, X_{igt})$ are weakly exogenous. But this might not hold.

First, the simultaneity between supply and demand for food implies that price at time t ($P_{gt}$) might be impacted by body weight at time t ($W_{igt}$). But an exogenous increase in body weight in the short period is more likely to raise food prices rather than lowering them. So we are likely to under-estimate the effect of food price on body weight. One solution is to find good instruments for food prices. But it is hard to find one instrument for food price, and even harder to find several instruments for prices of different kinds of food. Assume we need only one instrument, for the average food price. Then fuel price is one possibility. But as we will mention later, it will impact body weight not only through food prices, but through other channels, like physical activity. Another possibility is local extreme weather, like draught. But how food price fluctuates with local weather change is not clear. And extreme weather itself might affect human body's metabolism. Historical weather data is also hard to get. As a result, we will not try using instruments for food prices.

Second, there could always be some unobserved factors that impact both food prices and body weight. We will include the price of cigarettes, and the price of gasoline, and the price of non-food items (except cigarettes and gasoline) as control variables in our regression analysis.

Measurement error is another issue. Since the price data is collected at the city level, it only partially reflects the price at the local community level. The solution is to find some variables that are strongly correlated with local food price, but have measurement errors uncorrelated with the current price measure. We could not find such a variable. So our estimation will be biased towards zero, thus giving a lower bound on the real price effect. In addition, since the data we use only report self-reported height and weight, there is also measurement error in the dependent variable – body mass index (BMI). But we have found a dataset with both actual and reported height and weight, so we can use the strategy as described in Cawley (2004) to correct for this measurement error.

### Alternative Model Specification

Individual first-difference estimation only makes use of within-individual variations. But as Table 3-2 shows, the between-individual variations of price variables are of the same magnitude, or even larger than within-individual variations. Ignoring individual fixed effects, we estimate the following OLS regressions, making use of the total variation in key independent variables:

$$W_{igt} = \beta_0 + P_{gt}\beta_1 + Z_{gt}\beta_2 + X_{igt}\beta_3 + \tau_t + \mu_{igt} \qquad \textbf{(12)}$$

It is very likely that $E(\mu_{igt}P_{gt}) \neq 0$, which means that the OLS estimation will be biased. But the direction of bias is unknown. If we can find some instrumental variables – variables that are strongly correlated with $P_{gt}$ but are not correlated with $\mu_{igt}$ - then the problem will be solved. Such good instruments are difficult to find. There are a number of candidates – transportation costs, gasoline prices, weather – that might be correlated with food prices but it is hard to argue that they affect body weight only through food prices. Since the first-difference estimation is less prone to unobserved heterogeneity problem, it is the preferred model specification. The OLS estimation is used as a robustness check.

## *Results*

The key outcome variable is either BMI or the natural logarithm of BMI. We examined different sets of food price specifications: (1) log of price per 1,000 Calories; (2) log of price of at home food, log of price of fast food; (3) log of price of margarine, log of price of produce, log of at-home food excluding margarine and produce.

## Unadjusted Associations

Figure 3-2 to Figure 3-6 show the average 2-year change in BMI at the individual level, against the terciles of average 2-year declines in the log of relative prices (real price divided by the price for all goods in that county), for price per 1,000 Calories, price of at-home food, price of fast food, price of produce, and price of margarine. We observe positive correlations between BMI change and price decline in at-home food and margarine. The cheaper food is that home, the more likely individuals are to gain weight. There is a weak, non-linear positive correlation between BMI change and the decline in price per 1,000 Calories. There is no obvious relationship between BMI change and the price declines of fast food and produce. In all the graphs, the 95% confidence intervals are wide.

## First-difference estimation results

We next turn to the regression analysis. All regressions account for county/year clustering. Except for the food price vectors mentioned above, all regressions include the following variables: log of price of cigarettes, log of price of gasoline, log of price of non-food goods excluding cigarettes and gasoline, self-reported diagnosis of chronic conditions, self-rated health, current smoking status, marital status, whether working for pay, total household income, and total household wealth.

In order to examine whether price responses differ by social economic status, we added interaction terms of price variables and one of the two measures of social economic status: whether the individual is at the bottom tercile of the initial household wealth, and whether the individual is a high-school drop-out.

As shown in Table 3-3, price per 1,000 Calories is negatively correlated with BMI change. Using Laspeyres type of price indices, 100 percent increase in the price per 1,000 Calories is associated a reduction of 0.94 unit of BMI, or 2.7 percent. Both effects are statistically significant at the 0.05 significance level. As we expected, the price effects using Paasche type of price indices are slightly smaller, and are only significant at the 0.10 significance level. Another noticeable effect is the price of gasoline. 100 percent increase in gasoline price is associated with 0.65 unit decline in BMI, or 2.1 percent. We didn't find a significant effect of cigarette price on weight.

The middle part of Table 3-3 reports regression estimates with interactions of price variables and the indicator of at the bottom tercile of the initial household wealth. We find the price of calories only affects BMI of those at the bottom tercile of the initial household wealth. Using Laspeyres type of price indices, 100 percent increase in the price per 1,000 Calories is associated a reduction of 2.51 units of BMI, or 7.5 percent.

The bottom part of Table 3-3 reports regression estimates if we add interactions of price variables and the indicator of less than high school education. We don't find that the price of Calories differentially affect BMI of those with less than high-school education.

Table 3-4 shows the regression results using the second set of food price vectors: log of price of at home food, and log of price of fast food. Using Laspeyres type of price indices, 100 percent increase in the price per 1,000 Calories is associated a reduction of 1.00 unit of BMI, or 3.1 percent. Both effects are statistically significant at the 0.10 significance level. However, if we use Paasche type of price indices, the estimates reduce to 0.73 unit of BMI, or 2.3 percent, and are no longer statistically significant. We don't find a significant association between fast food price and BMI. Similar to the price of Calories, the price of at-home food only affects those at the bottom tercile of the initial household wealth. The bottom of Table 3-4 shows that none of the interactions of price variables and the indicator of less than high school education is statistically significant.

Table 3-5 shows the effects of two specific types of foods – fat and produce. We use margarine to represent fat, which is very calorie-dense, high in cholesterol and low in other nutrients. It is a potential target of "food tax" or "fat tax". Three items were included by ACCRA as produce: potatoes, bananas and iceberg lettuce. Produce is the type of low-calorie, and nutritious food and a potential food category to subsidize. From Table 3-5 we learn that the price of margarine is negatively associated with BMI. Using Laspeyres type of price indices, 100 percent increase in the price of margarine is associated a reduction of 0.32 unit of BMI, or 1.0 percent. The estimates using Paasche type of price indices are very similar, which is not surprising given that the price of margarine is constructed by a single item. The price of produce is not statistically associated with BMI. Excluding margarine and produce, the effect of price index of at-home food is negative but no longer statistically significant.

We are also interested in the effects of other control variables –self-rated health, chronic conditions, working status, marital status, household income and household wealth. Columns "II" in Table 3-7 shows the full regression estimates except for the effects of year dummies.

Questions for chronic conditions are in the format of: "Has a doctor ever told you that you have……" Therefore these chronic conditions are absorbing. Therefore the first difference regression reveals the correlation between BMI change and getting a certain condition.

We observe that variables of getting diabetes, stroke, and cancer are negatively associated with BMI. This is not likely to be the effect of BMI change on disease incidence, since previous studies have established that more obese individuals have increased risks for diabetes, stroke and certain types of cancers(NIH 1998). It is more likely due to the effects of getting a certain condition on BMI. Widowhood is negatively correlated with BMI change, indicating the involuntary weight loss effect of losing a spouse. We don't find significant effects of income or wealth on BMI.

### Robustness of first-difference estimations

As a robustness check, we also examined how the inclusion or exclusion or demographic and economic variables affects the estimation of price effects. Columns "I" in Table 3-7 doesn't include the following variables that are included in columns "II": self-rated health, chronic conditions, working status, marital status, household income and household wealth. The estimates of price effects in Columns "I" and those in columns "II" are very similar. This added the confidence that regional price change is not strongly correlated with demand side shocks.

Another robustness check is to relax the assumption of dynamic completeness. Our major regression estimates are based on Equation (8). But we make the very strong assumption about the lag BMI effect. So we estimate first-difference regressions with lag BMI effects, as Equation (9) specifies. Table 3-6 shows the results. We found that the lag effects of BMI are insignificant in all specifications. The point estimates of the price effects are very similar to the previous estimates, but the standard errors are much larger. As a result none of the price effects are significant. Since the lagged BMI effect is insignificant, and by doing the sample size is reduced by 30%, we prefer to report regression results without including the lagged BMI effect.

In addition to the continuous outcome of BMI, we also examine the effects of food prices on the binary outcome of obesity (BMI >=30 kg/m$^2$). Using either linear panel models or conditional logit models, we don't find significant effects of food prices on obesity.

### Alternative model specifications

Table 3-8 shows the OLS regression results as specified in Equation (11). Ignoring individual-specific effects, we find that all price effects are insignificant except for the price of margarine. It is negatively associated with BMI and the magnitude is much higher than that in the first-difference results. 100 percent increase in the price of Margarine is associated with 0.80 unit of BMI, or 2.5 percent.

## *Discussion*

In this study we examined the relationships between food prices and body weight among older Americans. Using first-difference estimation, we found that 100 percent increase in price of calories, or price of at-home food, is associated with 2-3 percent reduction in body weight. Interacting price variables with household wealth, we found that the price effects concentrate among individuals at the bottom one third of household wealth. For this population, 100 percent increase in price of calories, or price of at-home food, is associated with 5-7 percent reduction in body weight.

We also examined the price effects of specific types of foods. 100 percent increase in price of margarine is associated with 1 percent reduction in body weight. We don't find significant effects for price of fast food and price of produce.

Two other price factors that are likely to affect body weight are price of cigarettes and price of gasoline. We don't find significant effects of price of cigarettes on body weight. However, we found significant and negative effects of gasoline prices on body weight.

We next compare the elasticities of BMI estimated in this study with other studies. For at-home food, Chou (2004) estimated a price elasticity of -0.04, while our estimates are -0.02 to -0.03. The highest elasticity is reported by Lakdawalla (2002), which is -0.6. For fast food, Chou (2004) estimated a price elasticity of -0.05; we don't find a significant effect of fast food price on BMI. Datar (2005) doesn't find a significant effect of fast food price on children weight gain neither. Furthermore, Schroeter (2007) shows that rising fast food price might instead raise body weight. For "unhealthy" food, we found that 100 percent increase in price of margarine is associated with 1 percent reduction in BMI, and Gelbach (2007) find that 100 percent increase in the price of "unhealthy" food is associated with about 1 percent reduction in BMI. In addition, Datar (2005) finds that higher price of vegetables is associated with higher weight gain among children. We don't find significant effect of price of produce on BMI.

This study adds to the literature about the effects of food prices on body weight/obesity. We find very small effects of food prices on body weight among older Americans. A food tax, either for a specific type of food, or a general category of foods, might not be effective in combating the obesity epidemic. Furthermore, such a tax might be regressive since the poor households are more sensitive to price variation.

Our analysis has several limitations. First, we only have price measures at the county-level, not at the community level. This introduces measurement errors that would bias our estimates toward zero. Our estimates can be regarded as a lower bound of the true effects. But to our knowledge, there is no other better data source for regional price data. Second, we assume that price change is exogenous, which might not be true. And we don't have a good idea about what has caused differential price changes across regions.

**Figure 3-1 Flow chart of forming the analytic sample**

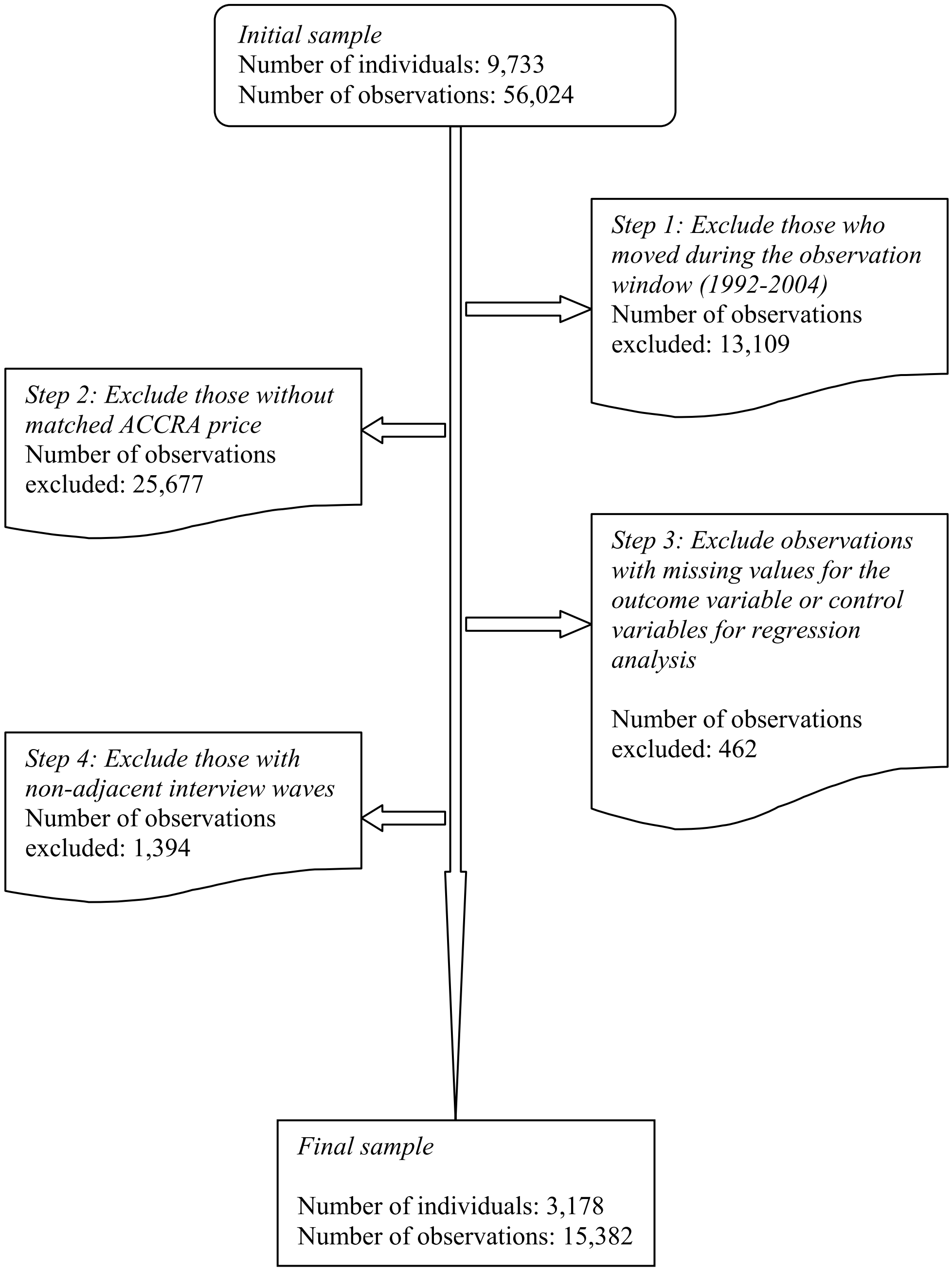


**Figure 3-2. BMI change and the decline of price per 1,000 Calories**

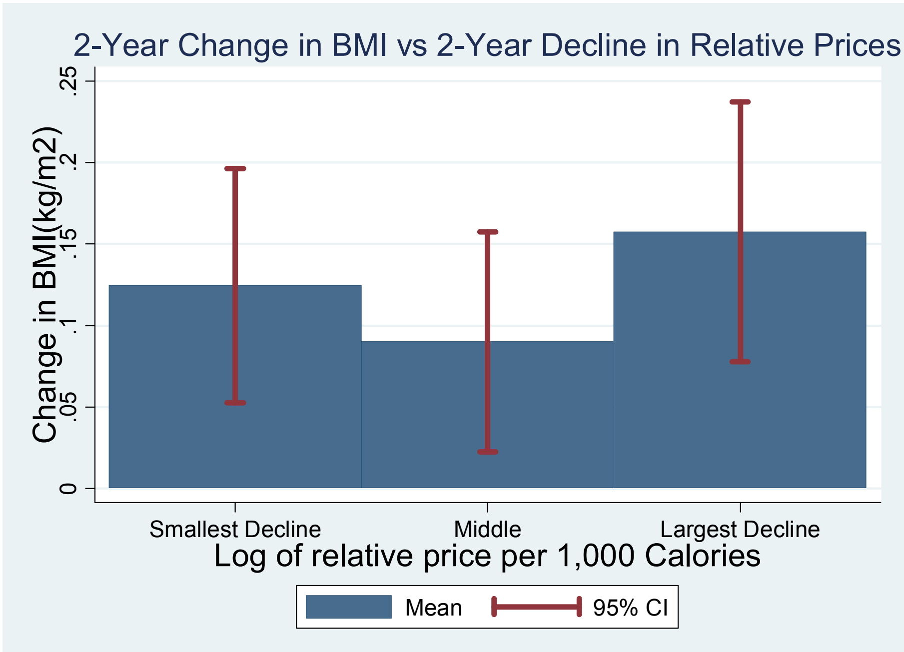



**Figure 3-3. BMI change and the decline of price of at-home food**

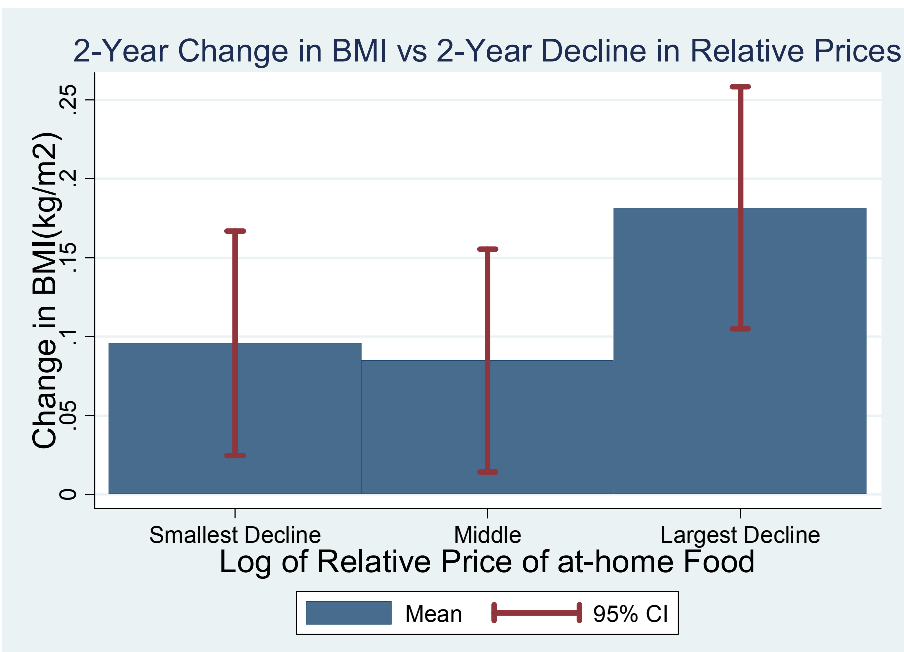


**Figure 3-4. BMI change and the decline of price of fast food**

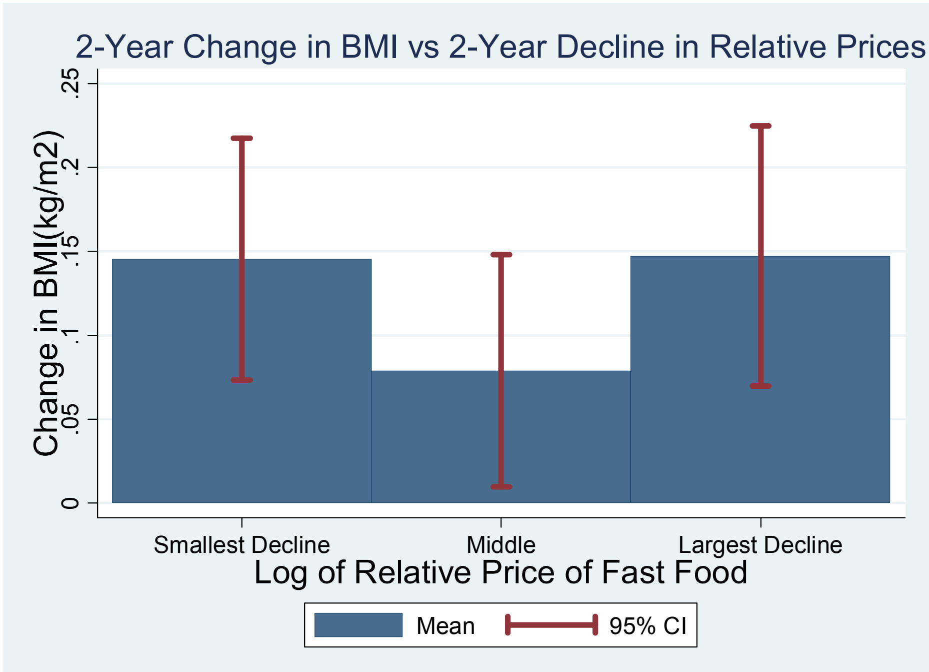



**Figure 3-5. BMI change and the decline of price of produce**

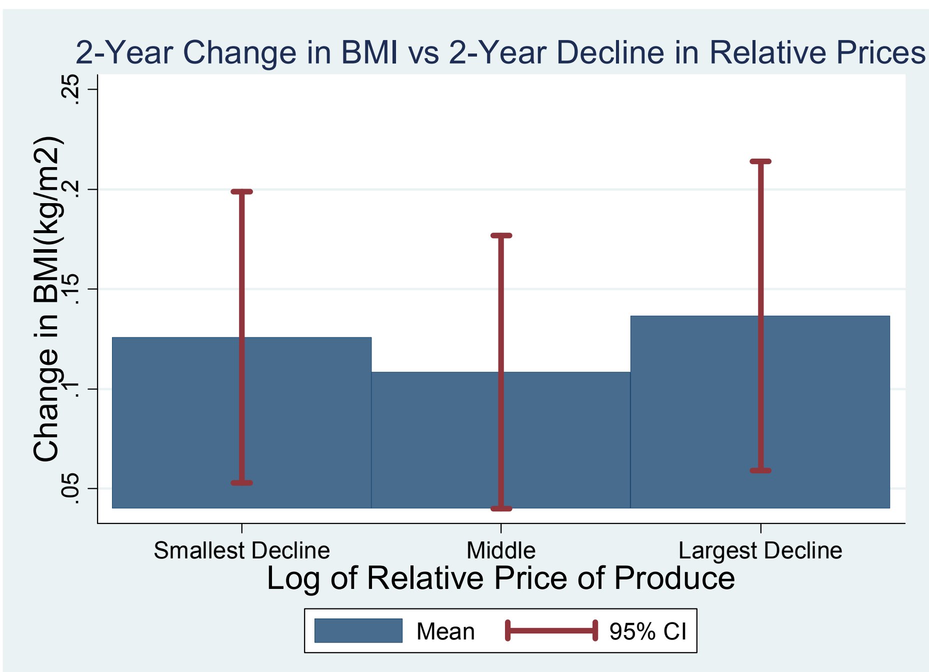


**Figure 3-6. BMI change and the decline of price of margarine**

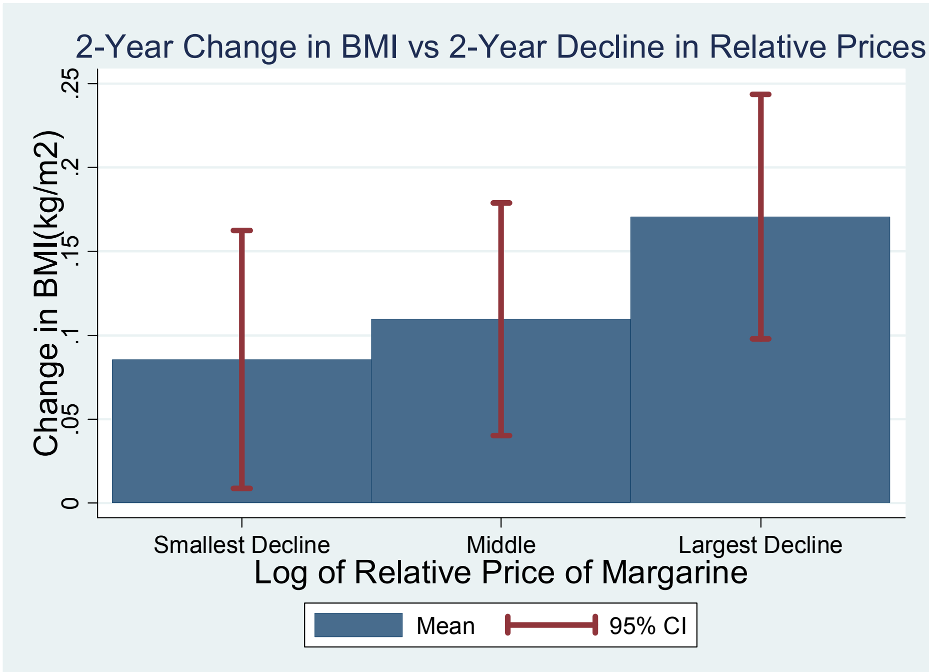


**Table 3-1. Probit Model of whether being included in the analytic sample**

| Covariate | Being included in the analytic sample |
|---|---|
| Male | -0.0213 |
| | (0.0283) |
| Non-Hispanic black | 0.2899*** |
| | (0.0388) |
| Hispanic | 0.3803*** |
| | (0.0500) |
| Less than high school | 0.0916** |
| | (0.0363) |
| Some college and above | 0.0389 |
| | (0.0319) |
| Suburban area | -0.0404 |
| | (0.0309) |
| Rural area | -0.7872*** |
| | (0.0353) |
| Initial Cancer | 0.0960 |
| | (0.0627) |
| Initial Diabetes | -0.0928* |
| | (0.0480) |
| Initial Heart disease | -0.0209 |
| | (0.0460) |
| Initial Hypertension | 0.0077 |
| | (0.0296) |
| Initial Lung disease | 0.0232 |
| | (0.0609) |
| Initial Stroke | -0.0998 |
| | (0.0876) |
| Initial Arthritis | 0.0285 |
| | (0.0295) |
| Initial Psyche problems | 0.0204 |
| | (0.0532) |
| Initial Current smoking | -0.0083 |
| | (0.0310) |
| Initial Self-rated Health is Fair/Poor | 0.0085 |
| | (0.0394) |
| Initial Age | 0.4489*** |
| | (0.1589) |
| Initial Age squared | -0.0040*** |
| | (0.0014) |
| Initial Log of household income | -0.0091 |
| | (0.0100) |
| Initial Non-positive household wealth | 0.0385 |
| | (0.1092) |
| Initial Log of household wealth | 0.0218** |
| | (0.0099) |
| Initial Widowed | -0.0034 |
| | (0.0582) |
| Initial Single | 0.0891** |

| | |
|---|---|
| | (0.0384) |
| Initial R working for pay | 0.0164 |
| | (0.0324) |
| N | 9715 |

* p<0.10, ** p<0.05, *** p<0.01

Data source: Health and Retirement Study 1992-2004, ACCRA 1992-2003

Note: A probit model is used to model the probability that a HRS respondent who is born between year 1931 and 1941 is included in the analytic sample.

**Table 3-2. HRS summary statistics for variables used in regression analysis**

Number of individuals: 3,178; Number of observations: 15,382

| **Variable** | **Mean** | **Standard deviation (total)** | **Standard deviation (between)** | **Standard deviation (within)** |
|---|---|---|---|---|
| Body mass index (kg/m2) | 29.02 | 6.24 | 6.00 | 1.89 |
| Prices (Laspeyres type of price index) | | | | |
| Price per 1,000 calories | 0.79 | 0.07 | 0.07 | 0.04 |
| Price of at-home food | 1.14 | 0.10 | 0.09 | 0.04 |
| Price of fast food | 2.74 | 0.18 | 0.14 | 0.12 |
| Price of produce | 0.70 | 0.11 | 0.08 | 0.07 |
| Price of margarine | 0.46 | 0.11 | 0.10 | 0.06 |
| Price of cigarettes | 15.27 | 5.00 | 3.67 | 3.95 |
| Price of gasoline | 0.78 | 0.09 | 0.06 | 0.07 |
| Price of non-food items | 220.00 | 99.98 | 93.06 | 40.05 |
| Chronic conditions (%) | | | | |
| Cancer | 9.0% | 28.6% | 25.5% | 13.6% |
| Diabetes | 15.7% | 36.3% | 33.9% | 15.2% |
| Heart disease | 16.2% | 36.9% | 34.5% | 15.9% |
| Hypertension | 45.8% | 49.8% | 46.1% | 20.7% |
| Lung disease | 7.2% | 25.8% | 24.7% | 11.1% |
| Stroke | 4.4% | 20.6% | 18.7% | 10.2% |
| Arthritis | 48.8% | 50.0% | 45.8% | 21.5% |
| Psyche problems | 11.4% | 31.7% | 30.6% | 12.3% |
| Self-rated Health is Fair/Poor (%) | 26.1% | 43.9% | 36.8% | 26.1% |
| Demographics | | | | |
| Age at interview | 61.28 | 4.91 | 3.85 | 3.41 |
| Widowed (%) | 10.6% | 30.8% | 27.7% | 14.8% |
| Single (%) | 19.3% | 39.5% | 37.7% | 13.0% |
| Economic conditions | | | | |
| R working for pay (%) | 51.4% | 50.0% | 41.3% | 30.2% |
| HH total income | 31,765 | 42,713 | 34,509 | 25,824 |
| HH wealth | 189,238 | 819,872 | 877,793 | 497,516 |

Data source: HRS 1992-2004, ACCRA 1992-2003.

**Table 3-3. First-difference estimations of BMI or Log(BMI), part 1**

| | Dependent variable: BMI | | Dependent variable: Log(BMI) | |
|---|---|---|---|---|
| | I | II | I | II |
| Log of price per 1,000 Calories | -0.938** | -0.817* | -0.027** | -0.023* |
| | (0.417) | (0.428) | (0.013) | (0.014) |
| Log of price of cigarettes | -0.183 | -0.194 | -0.002 | -0.002 |
| | (0.267) | (0.266) | (0.009) | (0.009) |
| Log of price of gasoline | -0.646* | -0.667* | -0.020* | -0.021* |
| | (0.363) | (0.370) | (0.012) | (0.012) |
| Log of price of non-food items excl. cigarettes and gasoline | 0.107 | 0.123 | 0.005 | 0.006 |
| | (0.185) | (0.179) | (0.006) | (0.006) |
| N | 11,871 | 11,871 | 11,871 | 11,871 |
| | | | | |
| Add interactions of price variables and the indicator of at the bottom tercile of the initial household wealth | | | | |
| Log of price per 1,000 Calories | -0.285 | -0.366 | -0.007 | -0.010 |
| | (0.472) | (0.457) | (0.015) | (0.015) |
| (Log of price per 1,000 Calories)*Bottom tercile of HH wealth | -2.221** | -1.502* | -0.068** | -0.042* |
| | (1.062) | (0.838) | (0.033) | (0.024) |
| Log of price of cigarettes | -0.150 | -0.155 | 0.000 | 0.000 |
| | (0.270) | (0.272) | (0.009) | (0.009) |
| (Log of price of cigaretttes)*Bottom tercile of HH wealth | -0.163 | -0.167 | -0.009 | -0.009 |
| | (0.270) | (0.264) | (0.009) | (0.009) |
| Log of price of gasoline | -0.714* | -0.739* | -0.024** | -0.025** |
| | (0.375) | (0.384) | (0.012) | (0.012) |
| (Log of price of gasoline)*Bottom tercile of HH wealth | 0.336 | 0.305 | 0.015 | 0.015 |
| | (0.417) | (0.436) | (0.014) | (0.015) |
| Log of price of non-food items | 0.100 | 0.097 | 0.004 | 0.004 |
| | (0.215) | (0.206) | (0.008) | (0.008) |
| (Log of price of non-food items excl. cigarettes and gasoline)* Bottom tercile of HH wealth | 0.004 | 0.054 | 0.004 | 0.005 |
| | (0.345) | (0.358) | (0.013) | (0.013) |
| N | 11,871 | 11,871 | 11,871 | 11,871 |
| | | | | |
| Add interactions of price variables and the indicator of less than high school education | | | | |
| Log of price per 1,000 Calories | -0.933** | -0.860* | -0.026* | -0.024* |
| | (0.434) | (0.442) | (0.014) | (0.014) |
| (Log of price per 1,000 Calories)*Less than high school | -0.124 | 0.109 | -0.010 | -0.001 |
| | (1.082) | (0.858) | (0.032) | (0.025) |
| Log of price of cigarettes | -0.086 | -0.093 | 0.001 | 0.001 |
| | (0.273) | (0.271) | (0.009) | (0.009) |
| (Log of price of cigaretttes)*Less than high school | -0.447 | -0.463 | -0.016 | -0.017 |
| | (0.444) | (0.445) | (0.013) | (0.014) |
| Log of price of gasoline | -0.728* | -0.754* | -0.023* | -0.024** |
| | (0.380) | (0.385) | (0.012) | (0.012) |
| (Log of price of gasoline)*Less than high school | 0.544 | 0.569 | 0.018 | 0.019 |
| | (0.397) | (0.393) | (0.013) | (0.013) |
| Log of price of non-food items | 0.206 | 0.213 | 0.008 | 0.008 |
| | (0.227) | (0.213) | (0.007) | (0.007) |
| (Log of price of non-food items excl. cigarettes and gasoline)* Less than high school | -0.478 | -0.446 | -0.014 | -0.013 |
| | (0.459) | (0.430) | (0.015) | (0.014) |

| | | | | |
|---|---|---|---|---|
| N | 11,871 | 11,871 | 11,871 | 11,871 |

* p<0.10, ** p<0.05, *** p<0.01

Data source: Health and Retirement Study, 1992-2004, ACCRA price data, 1992-2003
Notes:
Model I uses "Laspayres" type of price indices, while model II uses "Paasche" type of price indices.
All models also include the following variables: year dummies, self-rated health, chronic conditions, working status, marital status, household income and household wealth.

**Table 3-4. First-difference estimations of BMI or Log(BMI), part 2**

| | Dependent variable: BMI | | Dependent variable: Log(BMI) | |
|---|---|---|---|---|
| | I | II | I | II |
| Log of price of at-home food | -1.010* | -0.731 | -0.031* | -0.023 |
| | (0.545) | (0.554) | (0.017) | (0.017) |
| Log of price of fast food | 0.222 | 0.212 | 0.006 | 0.006 |
| | (0.424) | (0.430) | (0.014) | (0.014) |
| Log of price of cigarettes | -0.213 | -0.208 | -0.003 | -0.003 |
| | (0.259) | (0.259) | (0.008) | (0.008) |
| Log of price of gasoline | -0.649* | -0.703* | -0.020* | -0.022* |
| | (0.365) | (0.370) | (0.012) | (0.012) |
| Log of price of non-food items excl. cigarettes and gasoline | 0.123 | 0.141 | 0.006 | 0.006 |
| | (0.183) | (0.183) | (0.006) | (0.006) |
| N | 11,871 | 11,871 | 11,871 | 11,871 |
| | | | | |
| Add interactions of price variables and the indicator of at the bottom tercile of the initial household wealth | | | | |
| Log of price of at-home food | -0.130 | 0.103 | -0.006 | 0.003 |
| | (0.647) | (0.668) | (0.020) | (0.021) |
| (Log of price of at-home food )*Bottom tercile of HH wealth | -2.759** | -2.540** | -0.079* | -0.076** |
| | (1.382) | (1.165) | (0.043) | (0.036) |
| Log of price of fast food | 0.135 | 0.052 | 0.004 | 0.001 |
| | (0.620) | (0.639) | (0.020) | (0.021) |
| (Log of price of fast food)*Bottom tercile of HH wealth | 0.227 | 0.449 | 0.005 | 0.012 |
| | (1.057) | (1.118) | (0.032) | (0.034) |
| Log of price of cigarettes | -0.124 | -0.139 | 0.001 | 0.000 |
| | (0.268) | (0.268) | (0.009) | (0.009) |
| (Log of price of cigaretttes)*Bottom tercile of HH wealth | -0.276 | -0.216 | -0.013 | -0.011 |
| | (0.275) | (0.280) | (0.009) | (0.009) |
| Log of price of gasoline | -0.769** | -0.834** | -0.025** | -0.027** |
| | (0.385) | (0.389) | (0.012) | (0.012) |
| (Log of price of gasoline)*Bottom tercile of HH wealth | 0.399 | 0.458 | 0.017 | 0.019 |
| | (0.419) | (0.425) | (0.015) | (0.015) |
| Log of price of non-food items | 0.067 | 0.107 | 0.003 | 0.004 |
| | (0.226) | (0.216) | (0.008) | (0.008) |
| (Log of price of non-food items excl. cigarettes and gasoline)* Bottom tercile of HH wealth | 0.211 | 0.093 | 0.010 | 0.006 |
| | (0.390) | (0.375) | (0.014) | (0.014) |
| N | 11,871 | 11,871 | 11,871 | 11,871 |
| Add interactions of price variables and the indicator of less than high school education | | | | |
| Log of price of at-home food | -0.816 | -0.492 | -0.025 | -0.015 |
| | (0.629) | (0.604) | (0.019) | (0.019) |
| (Log of price of at-home food )*Less than high school | -0.843 | -1.065 | -0.027 | -0.035 |
| | (1.574) | (1.295) | (0.048) | (0.038) |
| Log of price of fast food | -0.001 | -0.045 | 0.000 | -0.001 |
| | (0.507) | (0.511) | (0.017) | (0.017) |
| (Log of price of fast food)*Less than high school | 0.615 | 0.728 | 0.014 | 0.018 |
| | (0.747) | (0.760) | (0.025) | (0.025) |
| Log of price of cigarettes | -0.120 | -0.122 | 0.001 | 0.001 |
| | (0.265) | (0.263) | (0.009) | (0.008) |
| (Log of price of cigaretttes)*Less than high school | -0.417 | -0.393 | -0.016 | -0.015 |

| | | | | |
|---|---|---|---|---|
| | (0.452) | (0.451) | (0.014) | (0.014) |
| Log of price of gasoline | -0.715* | -0.770** | -0.022* | -0.024** |
| | (0.382) | (0.386) | (0.012) | (0.012) |
| (Log of price of gasoline)*Less than high school | 0.485 | 0.515 | 0.017 | 0.018 |
| | (0.410) | (0.412) | (0.013) | (0.013) |
| Log of price of non-food items | 0.220 | 0.242 | 0.009 | 0.009 |
| | (0.226) | (0.222) | (0.007) | (0.007) |
| (Log of price of non-food items excl. cigarettes and gasoline)* Less than high school | -0.420 | -0.440 | -0.013 | -0.013 |
| | (0.451) | (0.426) | (0.014) | (0.014) |
| N | 11,871 | 11,871 | 11,871 | 11,871 |

* p<0.10, ** p<0.05, *** p<0.01

Data source: Health and Retirement Study, 1992-2004, ACCRA price data, 1992-2003
Notes:
Model I uses "Laspayres" type of price indices, while model II uses "Paasche" type of price indices.
All models also include the following variables: year dummies, self-rated health, chronic conditions, working status, marital status, household income and household wealth.

**Table 3-5. First-difference estimations of BMI or Log(BMI), part 3**

| | Dependent variable: BMI | | Dependent variable: Log(BMI) | |
|---|---|---|---|---|
| | I | II | I | II |
| Log of price of margarine | -0.324*** | -0.346*** | -0.010*** | -0.011*** |
| | (0.124) | (0.123) | (0.004) | (0.004) |
| Log of price of produce | -0.074 | -0.087 | -0.002 | -0.003 |
| | (0.312) | (0.311) | (0.011) | (0.010) |
| Log of price of at-home food excluding produce and margarine | -0.480 | -0.137 | -0.015 | -0.004 |
| | (0.522) | (0.540) | (0.016) | (0.017) |
| Log of price of fast food | 0.443 | 0.452 | 0.013 | 0.013 |
| | (0.409) | (0.411) | (0.014) | (0.014) |
| Log of price of cigarettes | -0.147 | -0.144 | -0.001 | -0.001 |
| | (0.268) | (0.267) | (0.009) | (0.009) |
| Log of price of gasoline | -0.755** | -0.815** | -0.023** | -0.025** |
| | (0.345) | (0.346) | (0.011) | (0.011) |
| Log of price of non-food items excl. cigarettes and gasoline | 0.139 | 0.153 | 0.006 | 0.007 |
| | (0.199) | (0.195) | (0.006) | (0.006) |
| N | 11,871 | 11,871 | 11,871 | 11,871 |
| Add interactions of price variables and the indicator of at the bottom tercile of the initial household wealth | | | | |
| Log of price of margarine | -0.247 | -0.279 | -0.008 | -0.009 |
| | (0.186) | (0.191) | (0.006) | (0.006) |
| (Log of price of margarine)*Bottom tercile of HH wealth | -0.230 | -0.198 | -0.007 | -0.006 |
| | (0.324) | (0.333) | (0.010) | (0.011) |
| Log of price of produce | -0.018 | -0.011 | 0.001 | 0.001 |
| | (0.391) | (0.392) | (0.014) | (0.014) |
| (Log of price of produce)*Bottom tercile of HH wealth | -0.120 | -0.195 | -0.007 | -0.009 |
| | (0.552) | (0.572) | (0.018) | (0.019) |
| Log of price of at-home food except produce and margarine | 0.224 | 0.521 | 0.005 | 0.016 |
| | (0.622) | (0.646) | (0.019) | (0.020) |
| (Log of price of at-home food except produce and margarine)*Bottom tercile of HH wealth | -2.245* | -2.027* | -0.062 | -0.061* |
| | (1.232) | (1.043) | (0.038) | (0.033) |
| Log of price of fast food | 0.327 | 0.276 | 0.010 | 0.009 |
| | (0.601) | (0.617) | (0.020) | (0.020) |
| (Log of price of fast food)*Bottom tercile of HH wealth | 0.290 | 0.458 | 0.006 | 0.011 |
| | (1.008) | (1.055) | (0.030) | (0.032) |
| Log of price of cigarettes | -0.065 | -0.084 | 0.003 | 0.002 |
| | (0.280) | (0.280) | (0.009) | (0.009) |
| (Log of price of cigaretttes)*Bottom tercile of HH wealth | -0.274 | -0.212 | -0.012 | -0.010 |
| | (0.282) | (0.287) | (0.010) | (0.010) |
| Log of price of gasoline | -0.854** | -0.919** | -0.028** | -0.030** |
| | (0.368) | (0.369) | (0.012) | (0.012) |
| (Log of price of gasoline)*Bottom tercile of HH wealth | 0.334 | 0.381 | 0.014 | 0.016 |
| | (0.504) | (0.505) | (0.017) | (0.017) |
| Log of price of non-food items | 0.083 | 0.114 | 0.003 | 0.004 |
| | (0.262) | (0.247) | (0.009) | (0.009) |

| | | | | |
|---|---|---|---|---|
| (Log of price of non-food items excl. cigarettes and gasoline)* Bottom tercile of HH wealth | 0.211 | 0.107 | 0.011 | 0.007 |
| | (0.454) | (0.437) | (0.016) | (0.016) |
| N | 11,871 | 11,871 | 11,871 | 11,871 |
| | | | | |
| Add interactions of price variables and the indicator of less than high school education | | | | |
| Log of price of margarine | -0.400*** | -0.432*** | -0.013*** | -0.014*** |
| | (0.137) | (0.135) | (0.004) | (0.004) |
| (Log of price of margarine)*Less than high school | 0.350 | 0.404 | 0.013 | 0.015 |
| | (0.423) | (0.426) | (0.013) | (0.013) |
| Log of price of produce | 0.008 | 0.006 | 0.001 | 0.001 |
| | (0.330) | (0.327) | (0.011) | (0.011) |
| (Log of price of produce)*Less than high school | -0.480 | -0.556 | -0.019 | -0.022 |
| | (0.491) | (0.511) | (0.016) | (0.016) |
| Log of price of at-home food except produce and margarine | -0.220 | 0.178 | -0.006 | 0.007 |
| | (0.623) | (0.606) | (0.019) | (0.018) |
| (Log of price of at-home food except produce and margarine)*Less than high school | -1.152 | -1.416 | -0.039 | -0.049 |
| | (1.523) | (1.231) | (0.046) | (0.036) |
| Log of price of fast food | 0.297 | 0.263 | 0.010 | 0.009 |
| | (0.483) | (0.481) | (0.016) | (0.016) |
| (Log of price of fast food)*Less than high school | 0.322 | 0.453 | 0.003 | 0.008 |
| | (0.820) | (0.817) | (0.027) | (0.027) |
| Log of price of cigarettes | -0.049 | -0.055 | 0.003 | 0.003 |
| | (0.279) | (0.276) | (0.009) | (0.009) |
| (Log of price of cigarettes)*Less than high school | -0.462 | -0.433 | -0.017 | -0.016 |
| | (0.475) | (0.472) | (0.014) | (0.014) |
| Log of price of gasoline | -0.823** | -0.882** | -0.026** | -0.028** |
| | (0.357) | (0.359) | (0.011) | (0.011) |
| (Log of price of gasoline)*Less than high school | 0.483 | 0.518 | 0.017 | 0.018 |
| | (0.422) | (0.416) | (0.013) | (0.013) |
| Log of price of non-food items | 0.211 | 0.226 | 0.008 | 0.009 |
| | (0.261) | (0.250) | (0.008) | (0.008) |
| (Log of price of non-food items excl. cigarettes and gasoline)* Less than high school | -0.293 | -0.310 | -0.008 | -0.008 |
| | (0.521) | (0.494) | (0.017) | (0.016) |
| N | 11,871 | 11,871 | 11,871 | 11,871 |

* p<0.10, ** p<0.05, *** p<0.01

Data source: Health and Retirement Study, 1992-2004, ACCRA price data, 1992-2003
Notes:
Model I uses "Laspayres" type of price indices, while model II uses "Paasche" type of price indices.
All models also include the following variables: year dummies, self-rated health, chronic conditions, working status, marital status, household income and household wealth.

**Table 3-6 Dynamic model of BMI or Log(BMI)**

| | Dependent variable: BMI | | | Dependent variable: Log of BMI | | |
|---|---|---|---|---|---|---|
| Lag of BMI(kg/m$^2$) | 0.006 | 0.006 | 0.006 | | | |
| | (0.078) | (0.078) | (0.078) | | | |
| Lag of log of BMI | | | | 0.071 | 0.072 | 0.072 |
| | | | | (0.062) | (0.062) | (0.062) |
| Log of price per 1000 Calories (Laspeyres index) | -0.878 | | | -0.028 | | |
| | (0.572) | | | (0.019) | | |
| Log of price of at-home food (Laspeyres index) | | -0.956 | | | -0.031 | |
| | | (0.702) | | | (0.024) | |
| Log of price of fast food (Laspeyres index) | | 0.033 | 0.058 | | -0.000 | 0.001 |
| | | (0.539) | (0.554) | | (0.018) | (0.019) |
| Log of price of margarine | | | -0.145 | | | -0.005 |
| | | | (0.232) | | | (0.008) |
| Log of price of produce (Laspeyres index) | | | -0.383 | | | -0.013 |
| | | | (0.324) | | | (0.011) |
| Log of price of at-home food except produce and margarine (Laspeyres index) | | | -0.547 | | | -0.018 |
| | | | (0.753) | | | (0.026) |
| Log of price of cigarettes | -0.042 | -0.046 | -0.089 | 0.004 | 0.004 | 0.003 |
| | (0.423) | (0.424) | (0.430) | (0.014) | (0.014) | (0.015) |
| Log of price of gasoline | -0.099 | -0.076 | -0.169 | -0.002 | -0.002 | -0.005 |
| | (0.581) | (0.586) | (0.597) | (0.020) | (0.020) | (0.020) |
| Log of price of non-food exclude cigarettes and gasoline (Laspeyres index) | 0.484 | 0.516 | 0.577* | 0.016 | 0.017 | 0.019 |
| | (0.334) | (0.338) | (0.340) | (0.012) | (0.012) | (0.012) |
| Age | 0.479** | 0.479** | 0.480** | 0.013* | 0.013* | 0.013* |
| | (0.219) | (0.219) | (0.219) | (0.007) | (0.007) | (0.007) |
| Age squared | -0.003** | -0.003** | -0.004** | -0.000 | -0.000 | -0.000 |
| | (0.002) | (0.002) | (0.002) | (0.000) | (0.000) | (0.000) |
| Log of household income | -0.010 | -0.010 | -0.010 | -0.000 | -0.000 | -0.000 |
| | (0.024) | (0.024) | (0.024) | (0.001) | (0.001) | (0.001) |
| Non-positive household wealth | -0.270 | -0.266 | -0.264 | -0.012 | -0.011 | -0.011 |
| | (0.293) | (0.293) | (0.293) | (0.010) | (0.010) | (0.010) |
| Log of household wealth | 0.007 | 0.007 | 0.007 | 0.000 | 0.000 | 0.000 |
| | (0.030) | (0.030) | (0.029) | (0.001) | (0.001) | (0.001) |
| Widowed | -0.418** | -0.413** | -0.416** | -0.015** | -0.015** | -0.015** |
| | (0.191) | (0.191) | (0.190) | (0.006) | (0.007) | (0.006) |
| Single | -0.168 | -0.166 | -0.173 | -0.006 | -0.006 | -0.007 |
| | (0.184) | (0.184) | (0.184) | (0.006) | (0.006) | (0.006) |
| R working for pay | 0.027 | 0.027 | 0.027 | 0.000 | 0.000 | 0.000 |
| | (0.071) | (0.071) | (0.070) | (0.002) | (0.002) | (0.002) |
| Cancer | -0.413* | -0.411* | -0.415* | -0.016** | -0.016** | -0.016** |
| | (0.228) | (0.227) | (0.227) | (0.008) | (0.008) | (0.008) |
| Diabetes | -1.010*** | -1.010*** | -1.011*** | -0.032*** | -0.032*** | -0.032*** |
| | (0.285) | (0.286) | (0.285) | (0.009) | (0.009) | (0.009) |
| Heart disease | -0.270 | -0.268 | -0.267 | -0.011* | -0.011* | -0.011* |
| | (0.189) | (0.190) | (0.190) | (0.006) | (0.006) | (0.006) |
| Hypertension | 0.080 | 0.076 | 0.078 | 0.003 | 0.003 | 0.003 |
| | (0.128) | (0.128) | (0.127) | (0.004) | (0.004) | (0.004) |
| Lung disease | -0.198 | -0.198 | -0.199 | -0.009 | -0.009 | -0.009 |
| | (0.243) | (0.244) | (0.243) | (0.009) | (0.009) | (0.009) |
| Stroke | -0.682** | -0.681** | -0.681** | -0.026*** | -0.026*** | -0.026*** |
| | (0.287) | (0.287) | (0.288) | (0.010) | (0.010) | (0.010) |
| Arthritis | 0.154 | 0.154 | 0.152 | 0.004 | 0.004 | 0.004 |
| | (0.103) | (0.102) | (0.102) | (0.004) | (0.004) | (0.004) |

| | | | | | | |
|---|---|---|---|---|---|---|
| Psyche problems | 0.095 | 0.096 | 0.095 | 0.003 | 0.003 | 0.003 |
| | (0.304) | (0.304) | (0.304) | (0.010) | (0.010) | (0.010) |
| Self-rated Health is Fair/Poor | 0.067 | 0.066 | 0.068 | -0.000 | -0.000 | -0.000 |
| | (0.092) | (0.092) | (0.092) | (0.003) | (0.003) | (0.003) |
| N | 8,360 | 8,360 | 8,360 | 8,360 | 8,360 | 8,360 |

* p<0.10, ** p<0.05, *** p<0.01

Data source: Health and Retirement Study, 1992-2004, ACCRA price data, 1992-2003
Note: All models include year dummies.

**Table 3-7. Compare different specifications of First-difference regressions of BMI or Log(BMI)**

| | Dependent variable: BMI | | Dependent variable: Log(BMI) | |
|---|---|---|---|---|
| | I | II | I | II |
| Log of price per 1,000 Calories (Laspeyres index) | -0.973** | -0.938** | -0.029** | -0.027** |
| | (0.409) | (0.417) | (0.013) | (0.013) |
| Log of price of gasoline | -0.615* | -0.646* | -0.020* | -0.020* |
| | (0.368) | (0.363) | (0.012) | (0.012) |
| Log of price of cigarettes | -0.208 | -0.183 | -0.003 | -0.002 |
| | (0.264) | (0.267) | (0.009) | (0.009) |
| Log of price of non-food exclude cigarette and gasoline (Laspeyres index) | 0.149 | 0.107 | 0.007 | 0.005 |
| | (0.190) | (0.185) | (0.006) | (0.006) |
| Age | 0.361 | 0.308 | 0.010 | 0.008 |
| | (0.239) | (0.242) | (0.008) | (0.008) |
| Age squared | -0.003 | -0.002 | -0.000 | -0.000 |
| | (0.002) | (0.002) | (0.000) | (0.000) |
| Self-rated Health is Fair/Poor | | 0.109 | | 0.001 |
| | | (0.080) | | (0.003) |
| Psyche problems | | 0.061 | | 0.001 |
| | | (0.193) | | (0.006) |
| Arthritis | | 0.138 | | 0.005 |
| | | (0.087) | | (0.003) |
| Stroke | | -0.901** | | -0.029*** |
| | | (0.367) | | (0.010) |
| Lung disease | | -0.109 | | -0.007 |
| | | (0.234) | | (0.009) |
| Hypertension | | 0.072 | | 0.004 |
| | | (0.127) | | (0.004) |
| Heart disease | | -0.357** | | -0.013** |
| | | (0.173) | | (0.005) |
| Diabetes | | -1.019*** | | -0.031*** |
| | | (0.260) | | (0.008) |
| Cancer | | -0.395** | | -0.016** |
| | | (0.201) | | (0.007) |
| R working for pay | | -0.021 | | -0.001 |
| | | (0.070) | | (0.002) |
| Single | | -0.144 | | -0.005 |
| | | (0.150) | | (0.005) |
| Widowed | | -0.358** | | -0.014*** |
| | | (0.142) | | (0.005) |
| Log of household wealth | | 0.008 | | 0.000 |
| | | (0.022) | | (0.001) |
| Non-positive household wealth | | -0.094 | | -0.004 |
| | | (0.258) | | (0.008) |
| Log of household income | | -0.006 | | -0.000 |
| | | (0.018) | | (0.001) |
| Constant | -0.164 | -0.154 | -0.004 | -0.003 |
| | (0.168) | (0.169) | (0.005) | (0.005) |
| N | 11,871 | 11,871 | 11,871 | 11,871 |

* p<0.10, ** p<0.05, *** p<0.01

Data source: Health and Retirement Study, 1992-2004, ACCRA price data, 1992-2003
Notes: All models include year dummies.
Model I doesn't include the following variables: self-rated health, smoking status, chronic conditions, working status, marital status, household income and household wealth.

**Table 3-8 OLS of BMI or Log(BMI)**

| | Dependent variable: BMI | | | Dependent variable: Log of BMI | | |
|---|---|---|---|---|---|---|
| Log of price per 1000 Calories (Laspeyres index) | -0.391 | | | -0.011 | | |
| | (0.983) | | | (0.032) | | |
| Log of price of at-home food (Laspeyres index) | | 0.109 | | | 0.006 | |
| | | (1.374) | | | (0.045) | |
| Log of price of fast food (Laspeyres index) | | 1.511 | 2.104 | | 0.062 | 0.080 |
| | | (1.624) | (1.507) | | (0.057) | (0.052) |
| Log of price of margarine | | | -0.802** | | | -0.025** |
| | | | (0.321) | | | (0.010) |
| Log of price of produce (Laspeyres index) | | | 0.490 | | | 0.015 |
| | | | (0.630) | | | (0.022) |
| Log of price of at-home food except produce and margarine (Laspeyres index) | | | 1.135 | | | 0.039 |
| | | | (1.469) | | | (0.050) |
| Log of price of cigarettes | 0.025 | -0.078 | 0.137 | 0.002 | -0.002 | 0.005 |
| | (0.657) | (0.681) | (0.658) | (0.023) | (0.024) | (0.024) |
| Log of price of gasoline | 0.152 | -0.194 | -0.385 | -0.001 | -0.014 | -0.021 |
| | (1.319) | (1.080) | (0.997) | (0.047) | (0.037) | (0.035) |
| Log of price of non-food exclude cigarettes and gasoline (Laspeyres index) | 0.042 | -0.146 | -0.236 | -0.001 | -0.008 | -0.011 |
| | (0.395) | (0.484) | (0.458) | (0.013) | (0.016) | (0.016) |
| Male | 0.265 | 0.264 | 0.262 | 0.016** | 0.016** | 0.016** |
| | (0.195) | (0.195) | (0.195) | (0.006) | (0.006) | (0.006) |
| Non-Hispanic black | 1.377*** | 1.380*** | 1.393*** | 0.045*** | 0.045*** | 0.046*** |
| | (0.291) | (0.290) | (0.291) | (0.009) | (0.009) | (0.009) |
| Hispanic | 0.713*** | 0.703*** | 0.756*** | 0.035*** | 0.034*** | 0.036*** |
| | (0.257) | (0.256) | (0.255) | (0.009) | (0.009) | (0.009) |
| Less than high school | 0.144 | 0.145 | 0.141 | 0.006 | 0.006 | 0.006 |
| | (0.353) | (0.351) | (0.349) | (0.011) | (0.011) | (0.011) |
| Some college and above | -0.228 | -0.231 | -0.224 | -0.009 | -0.009 | -0.009 |
| | (0.227) | (0.226) | (0.227) | (0.007) | (0.007) | (0.007) |
| Age | -0.082 | -0.088 | -0.098 | -0.003 | -0.004 | -0.004 |
| | (0.423) | (0.423) | (0.426) | (0.013) | (0.013) | (0.013) |
| Age squared | -0.000 | -0.000 | -0.000 | -0.000 | -0.000 | 0.000 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | (0.003) | (0.003) | (0.003) | (0.000) | (0.000) | (0.000) |
| Log of household income | 0.140*** | 0.141*** | 0.142*** | 0.005*** | 0.005*** | 0.005*** |
| | (0.051) | (0.051) | (0.051) | (0.002) | (0.002) | (0.002) |
| Non-positive household wealth | -1.763*** | -1.768*** | -1.754*** | -0.053*** | -0.053*** | -0.053*** |
| | (0.560) | (0.560) | (0.565) | (0.018) | (0.018) | (0.018) |
| Log of household wealth | -0.161** | -0.161** | -0.160** | -0.004* | -0.004* | -0.004* |
| | (0.067) | (0.067) | (0.067) | (0.002) | (0.002) | (0.002) |
| Widowed | -0.169 | -0.175 | -0.165 | -0.006 | -0.006 | -0.006 |
| | (0.359) | (0.358) | (0.356) | (0.012) | (0.012) | (0.012) |
| Single | -0.761** | -0.763** | -0.752** | -0.028** | -0.028** | -0.028** |
| | (0.317) | (0.318) | (0.318) | (0.011) | (0.011) | (0.011) |
| R working for pay | 0.261 | 0.263 | 0.261 | 0.011* | 0.011* | 0.011* |
| | (0.198) | (0.199) | (0.197) | (0.006) | (0.006) | (0.006) |
| Cancer | 0.159 | 0.165 | 0.166 | 0.004 | 0.004 | 0.004 |
| | (0.307) | (0.306) | (0.307) | (0.010) | (0.010) | (0.010) |
| Diabetes | 2.798*** | 2.801*** | 2.805*** | 0.093*** | 0.093*** | 0.093*** |
| | (0.383) | (0.382) | (0.382) | (0.012) | (0.012) | (0.012) |
| Heart disease | -0.062 | -0.062 | -0.065 | -0.001 | -0.001 | -0.001 |
| | (0.281) | (0.281) | (0.280) | (0.009) | (0.009) | (0.009) |
| Hypertension | 2.138*** | 2.136*** | 2.137*** | 0.073*** | 0.073*** | 0.073*** |
| | (0.204) | (0.204) | (0.203) | (0.006) | (0.006) | (0.006) |
| Lung disease | -1.514*** | -1.510*** | -1.508*** | -0.059*** | -0.059*** | -0.059*** |
| | (0.443) | (0.443) | (0.445) | (0.015) | (0.015) | (0.015) |
| Stroke | -0.446 | -0.443 | -0.454 | -0.018 | -0.018 | -0.018 |
| | (0.392) | (0.392) | (0.392) | (0.013) | (0.013) | (0.013) |
| Arthritis | 1.429*** | 1.431*** | 1.432*** | 0.047*** | 0.047*** | 0.047*** |
| | (0.161) | (0.160) | (0.160) | (0.006) | (0.006) | (0.006) |
| Psyche problems | 0.465 | 0.463 | 0.459 | 0.014 | 0.013 | 0.013 |
| | (0.383) | (0.382) | (0.383) | (0.012) | (0.012) | (0.012) |
| Self-rated Health is Fair/Poor | 0.398 | 0.394 | 0.399 | 0.009 | 0.009 | 0.009 |
| | (0.288) | (0.291) | (0.291) | (0.009) | (0.009) | (0.009) |
| N | 15,382 | 15,382 | 15,382 | 15,382 | 15,382 | 15,382 |

* p<0.10, ** p<0.05, *** p<0.01

Data source: Health and Retirement Study, 1992-2004, ACCRA price data, 1992-2003
Notes: All models include year dummies.
Model I doesn't include the following variables: self-rated health, smoking status, chronic conditions, working status, marital status, household income and household wealth.

## *Appendix A*

**Table A 3-1. OLS regressions of measured weight (Kgs) against self-reported weight, among aged 52-73**

| | Non-hispanic white | | Non-hispanic black | | Hispanic | | Other | |
|---|---|---|---|---|---|---|---|---|
| | no constant | with constant | no constant | with constant | no constant | with constant | no constant | with constant |
| **Female** | | | | | | | | |
| Self-reported weight | 0.94445*** | 0.35500*** | 0.84732*** | 0.33229 | 1.02330*** | 1.09120*** | 0.99669*** | 1.07097*** |
| | (0.01517) | (0.11216) | (0.09097) | (0.66429) | (0.01713) | (0.12537) | (0.02123) | (0.15173) |
| Self-reported weight squared | 0.00111*** | 0.00463*** | 0.00238** | 0.00530 | -0.00009 | -0.00049 | 0.00022 | -0.00023 |
| | (0.00018) | (0.00069) | (0.00100) | (0.00385) | (0.00021) | (0.00075) | (0.00024) | (0.00093) |
| Constant | | 23.51265*** | | 21.88187 | | -2.75817 | | -2.90781 |
| | | (4.43347) | | (27.95705) | | (5.04500) | | (5.88059) |
| Adjusted R square | 0.981 | 0.732 | 0.899 | 0.331 | 0.996 | 0.925 | 0.998 | 0.978 |
| N | 1,913 | 1,913 | 419 | 419 | 250 | 250 | 60 | 60 |
| **Male** | | | | | | | | |
| Self-reported weight | 0.85458*** | -0.65228** | 1.00388*** | 0.95523*** | 1.05346*** | 1.53617*** | 0.98195*** | 0.24352 |
| | (0.02890) | (0.26129) | (0.02621) | (0.25067) | (0.02517) | (0.21503) | (0.03419) | (0.35586) |
| Self-reported weight squared | 0.00187*** | 0.00979*** | -0.00000 | 0.00026 | -0.00047* | -0.00319** | 0.00028 | 0.00448** |
| | (0.00030) | (0.00140) | (0.00027) | (0.00138) | (0.00028) | (0.00124) | (0.00039) | (0.00205) |
| Constant | | 69.69147*** | | 2.18538 | | -20.85333** | | 31.36070** |
| | | (12.01198) | | (11.19848) | | (9.22691) | | (15.04796) |
| Adjusted R square | 0.968 | 0.529 | 0.996 | 0.873 | 0.998 | 0.912 | 0.998 | 0.953 |
| N | 1,565 | 1,565 | 240 | 240 | 161 | 161 | 48 | 48 |

* p<0.10, ** p<0.05, *** p<0.01

Data source: Health and Retirement Study, 2006

**Table A 3-2. OLS regressions of measured height (meters) against self-reported height, among aged 52-73**

| | Non-Hispanic white | | Non-Hispanic black | | Hispanic | | Other | |
|---|---|---|---|---|---|---|---|---|
| | no constant | with constant | no constant | with constant | no constant | with constant | no constant | with constant |
| **Female** | | | | | | | | |
| Self-reported height | 1.08784*** | 1.42432*** | 1.17019*** | 2.38085*** | 1.34023*** | -5.04342*** | 1.27161*** | -4.10097*** |
| | (0.01443) | (0.46389) | (0.03464) | (0.86759) | (0.05823) | (1.01126) | (0.07997) | (0.96188) |
| Self-reported height squared | -0.05978*** | -0.16260 | -0.11062*** | -0.48506* | -0.22523*** | 1.82659*** | -0.17711*** | 1.56443*** |
| | (0.00882) | (0.14196) | (0.02116) | (0.26895) | (0.03674) | (0.32636) | (0.05005) | (0.31372) |
| Constant | | -0.27484 | | -0.97676 | | 4.95291*** | | 4.12859*** |
| | | (0.37872) | | (0.69941) | | (0.78348) | | (0.73749) |
| Adjusted R square | 0.999 | 0.665 | 0.999 | 0.581 | 0.998 | 0.402 | 0.999 | 0.723 |
| N | 1,937 | 1,937 | 425 | 425 | 249 | 249 | 60 | 60 |
| **Male** | | | | | | | | |
| Self-reported height | 1.05773*** | -1.59108*** | 1.14252*** | 1.46816 | 1.06206*** | 5.45421* | 1.56881*** | 8.57784*** |
| | (0.01593) | (0.55868) | (0.03468) | (1.11040) | (0.07841) | (2.96718) | (0.09220) | (1.98111) |
| Self-reported height squared | -0.03956*** | 0.70656*** | -0.08773*** | -0.17962 | -0.04310 | -1.30689 | -0.33635*** | -2.27914*** |
| | (0.00893) | (0.15756) | (0.01942) | (0.31377) | (0.04556) | (0.85468) | (0.05226) | (0.55064) |
| Constant | | 2.34744*** | | -0.28795 | | -3.80976 | | -6.30324*** |
| | | (0.49492) | | (0.98140) | | (2.57285) | | (1.78007) |
| Adjusted R square | 0.999 | 0.686 | 0.999 | 0.712 | 0.998 | 0.453 | 0.999 | 0.428 |
| N | 1,567 | 1,567 | 240 | 240 | 162 | 162 | 48 | 48 |

* p<0.10, ** p<0.05, *** p<0.01

Data source: Health and Retirement Study, 2006

## *Appendix B*

**Table B 3-1. Items, descriptions, and expenditure shares for ACCRA data collected in year 1992**

| Expenditure share within each category | Item | Item Description |
|---|---|---|
| **Grocery (expenditure share 13%)** | | |
| 0.0527 | T-Bone Steak | Price per pound |
| 0.0527 | Ground Beef or hamburger | Price per pound, lowest price |
| 0.0492 | Sausage | Price per pound; Jimmy Dean 100% pork |
| 0.0371 | Frying Chicken | Price per pound, whole fryer |
| 0.0306 | Chunk Light Tuna | 6.125-6.5 oz can, Starkist or Chicken of the Sea, packed in oil |
| 0.0494 | Whole Milk | Half-Gallon carton |
| 0.009 | Eggs | One Dozen, Grade A, Large |
| 0.0376 | Margarine | One Pound, cubes, Blue Bonnet or Parkay |
| 0.0376 | Parmesan Cheese, Grated | 8 oz. Canister, Kraft Brand |
| 0.0228 | Potatoes | 10 pound sack, white or red |
| 0.0474 | Bananas | Price per pound |
| 0.0228 | Iceberg Lettuce | Head, approximately 1.25 pounds |
| 0.0818 | Bread, White | 24 oz. loaf, lowest price, or prorated 24-oz. equivalent, lowest price |
| 0.0748 | Cigarettes | Carton, Winston, king-size (85 mm.) |
| 0.0513 | Coffee, Vacuum-Packed | 13 oz. can, Maxwell House, Hills Brothers, or Foldgers |
| 0.0314 | Sugar | 5 pounds, Cane or Beet, lowest price |
| 0.0419 | Corn Flakes | 18 oz., Kellog's or Post Toasties |
| 0.0072 | Sweet Peas | 15-17 oz. can, Del Monte or Green Giant |
| 0.0072 | Tomatoes | 14-1/2 oz. can, Hunts or Del Monte |
| 0.0333 | Peaches | 29 oz. can, Hunt's, Del Monte, or Libby's, halves or slices |
| 0.0221 | Facial Tissues | 175-count box, Kleenex brand |
| 0.0417 | Washing Powder | 42 oz. ("Ultra"), Tide, Bold, or Cheer |
| 0.0184 | Shortening | 3 pound can, all-vegetable, Crisco brand |
| 0.0384 | Frozen Orange Juice | 12 oz. can, Minute Maid brand |
| 0.0072 | Frozen Corn | 10 oz., Whole Kernel, lowest price |
| 0.056 | Baby Food | 4-4.5 oz. jar, strained vegetables, lowest price |
| 0.0384 | Soft Drink | 2 liter Coca Cola, excluding any deposit |
| **Housing (expenditure share 28%)** | | |
| 0.2631 | Apartment, Monthly Rent | Two-Bedroom, unfurnished, excluding all utilities except water, 1-1/2 baths, approximately 950 sq.ft. |
| | Total Purchase Price | 1,800 sq.ft. living area new house, 8,000 sq.ft. lot, urban area with all utilities |
| | Mortgage rate | Effective rate, including points and origination fee, for 30-year conventional fixed- or adjustable-rate mortgage |
| 0.7369 | Monthly Payment | Principal and Interest, using mortgage rate from Item 29B and assuming 25% down payment |

**Utilities (expenditure share 9%)**

| | | |
|---|---|---|
| 0.9 | Total Home Energy Cost | Monthly Cost, at current rates, for average monthly consumption of all types of energy during the previous 12 months for the type of home specified in item 29A |
| | Electricity | Average monthly cost for all-electric homes is shown in column 30A; average monthly cost for homes using other types of energy as well is shown in column 30B |
| | Other Home Energy | Average monthly cost at current rates for natural gas, fuel oil, coal, wood and any other forms of energy except electricity |
| 0.1 | Telephone | Private residential line; Customer owns instruments. Price includes: basic monthly rate; additional local use charges, if any, incurred by a family of four; Touch Tone fee; all other mandatory monthly charges, such ass long distance access fee and 911 fee; and all taxes foregoing |

**Transportation (expenditure share 10%)**

| | | |
|---|---|---|
| 0.1 | Commuter Fare | One-way commuting fare, up to ten miles |
| 0.3541 | Auto Maintenance | Average price to computer- or spin balance- one front wheel |
| 0.5459 | Gasoline | One Gallon regular unleaded, national brand, including all taxes; cash price at self service pump if available |

**HealthCare (expenditure share 5%)**

| | | |
|---|---|---|
| 0.175 | Hospital room | Average cost per day for semi-private room |
| 0.3509 | Office Visit, Doctor | American Medical Association procedure 90050: general practitioner's routine examination of established patient |
| 0.3509 | Office Visit, Dentist | American Dental Association procedure 1110 (adult teeth cleaning) and 0120 (periodic oral examination) |
| 0.1232 | Aspirin | 100 tablet bottle, Bayer brand, 325-mg., tablets |

**Miscellaneous (expenditure share 35%)**

| | | |
|---|---|---|
| 0.095 | Hamburger Sandwich | 1/4 pound patty with cheese, pickle, onion, mustard, and catsup. McDonald's Quarter-Pounder with Cheese, where available |
| 0.095 | Pizza | 12"-13" thin crust cheese pizza. Pizza Hut or Pizza Inn, where available |
| 0.095 | Fried Chicken | Thigh and Drumstick, with or without extras, whichever is less expensive. Kentucky Fried Chicken or Church's, where available |
| 0.0174 | Haircut | Mans barber shop haircut, no styling |
| 0.0174 | Beauty Salon | Woman's shampoo, trim, and blow dry |
| 0.0174 | Toothpaste | 6 oz.-7oz. tube, crest or colgate |
| 0.0174 | Shampoo | 15 oz. Bottle, Alberto VO-5 |
| 0.0174 | Dry Cleaning | Man's two-piece suit |
| 0.115 | Man's Dress Shirt | Arrow, Enro, Van Huesen, or J.C Penny's Stafford. White, cotton/polyester blend (at least 55% cotton), long sleeves |
| 0.0523 | Boy's Underwear | Package of three briefs, size 10-14, cotton, lowest price |
| 0.115 | Man's Denim Jeans | Levi's Brand, 501s or 505s, rinsed washed or bleached, size 28/30-34/36 |

| | | |
|---|---|---|
| 0.0742 | Major Appliance repair | Home service call, clothes washing machine; minimum labor charge, excluding parts |
| 0.0271 | Newspaper Subscription | Daily and Sunday home delivery, large-city newspaper |
| 0.0459 | Movie | First-run, indoor, evening, no discount |
| 0.0459 | Bowling | Price per line (game), evening rate |
| 0.0654 | Tennis Balls | Can of three extra duty, yellow, Wilson or Penn Brand |
| 0.0384 | Board Game | Parket Brothers "Monopoly", No. 9 edition |
| 0.0163 | Liquor | J&B Scotch, 750-ml. bottle |
| 0.0162 | Beer | Budweiser or Miller Lite, 6-pack, 12 oz. containers, excluding any deposit |
| 0.0163 | Wine | Gallo chablis blanc, 1.5-liter bottle |

Data source: Council for Community and Economic Research (C2ER) - formerly known as ACCRA

## *Appendix C*

**Superlative price indices:**
People also use one of three price indexes that stands in-between the Laspeyres index and the Paasche index. These are called superlative indices and there is no conclusion about which one is better[16].
The three price indices are:

(1) Fisher Ideal index: the geometric mean of the Laspeyres and Paasche indexes (the square root of their product)

**(2) The Tornqvist index**[17]: **Tornqvist index** is a discrete approximation to a continuous Divisia index. A Divisia index is a weighted sum of the growth rates of the various components, where the weights are the component's shares in total value. When a Tornqvist index is used as an approximation to the continuous Divisia index, the growth rates are defined as the difference in natural logarithms of successive observations of the components and the weights are equal to the mean of the factor shares of the components in the corresponding pair of years. $D_t$ is the price index in year t, and $D_{t-1}$ is the price index in year t-1; $s_{i,t}$ and $p_{i,t}$ are budget share and price for component i at year t.

$$\log(D_t) - \log(D_{t-1}) = \sum_{i=1}^{n} \frac{1}{2}[s_{i,t} + s_{i,t-1}][\log(p_{i,t}) - \log(p_{i,t-1})],$$

*where*

$$s_{i,t} = \frac{q_{i,t} p_{i,t}}{\sum_{j=1}^{n} q_{j,t} p_{j,t}} \qquad (\mathrm{i} = 1,...\mathrm{n}).$$

**3) The Walsh index[18]: the formula is as equation XXX.**

$$\mathrm{D_t} = \frac{\sum_{i=1}^{n} (p_{i,t} / \sqrt{p_{i,0} p_{i,t}}) \sqrt{s_{i,0} s_{i,t}}}{\sum_{i=1}^{n} (p_{i,0} / \sqrt{p_{i,0} p_{i,t}}) \sqrt{s_{i,0} s_{i,t}}}$$

---

[16] Robert J. Hill, Superlative index numbers: not all of them are super, Journal of Econometrics,Volume 130, Issue 1, , January 2006, Pages 25-43.

[17] Hulten CR. Divisia Index numbers, 1973

[18] IMF, new Export and Import Price Index Manual. Chapter 16.
https://www.imf.org/external/np/sta/tegeipi/

## *Bibliography*

Adams, K. F., A. Schatzkin, et al. (2006). "Overweight, obesity, and mortality in a large prospective cohort of persons 50 to 71 years old." N Engl J Med **355**(8): 763-78.

Aguiar, M. and E. Hurst (2005). "Consumption versus Expenditure." Journal of Political Economy **113**(5): 919-48.

Anderson, P. M., K. F. Butcher, et al. (2003). "Maternal employment and overweight children." J Health Econ **22**(3): 477-504.

Arellano, M. and S. 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.

Averett, S. and S. Korenman (1996). "The Economic Reality of The Beauty Myth." Journal of Human Resources **31**(2): 304-30.

Bhattacharya, J. and N. Sood (2005). Health Insurance and the Obesity Externality. National Bureau of Economic Research, Inc, NBER Working Papers.

Booth, K. M., M. M. Pinkston, et al. (2005). "Obesity and the built environment." J Am Diet Assoc **105**(5 Suppl 1): S110-7.

Brown, C. D., M. Higgins, et al. (2000). "Body mass index and the prevalence of hypertension and dyslipidemia." Obes Res **8**(9): 605-19.

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

Chan, J. M., E. B. Rimm, et al. (1994). "Obesity, fat distribution, and weight gain as risk factors for clinical diabetes in men." Diabetes Care **17**(9): 961-9.

Chou, S.-Y., M. Grossman, et al. (2004). "An Economic Analysis of Adult Obesity: Results from the Behavioral Risk Factor Surveillance System." Journal of Health Economics **23**(3): 565-87.

Chouinard, H., D. Davis, et al. (2007). "Fat Taxes: Big Money for Small Change." *Forum for Health Economics & Policy*: **10**(2).

Colditz, G. A., W. C. Willett, et al. (1995). "Weight gain as a risk factor for clinical diabetes mellitus in women." Ann Intern Med **122**(7): 481-6.

Cutler, D. M., E. L. Glaeser, et al. (2003). "Why Have Americans Become More Obese?" Journal of Economic Perspectives **17**(3): 93-118.

Dyer, A. R. and P. Elliott (1989). "The INTERSALT study: relations of body mass index to blood pressure. INTERSALT Co-operative Research Group." J Hum Hypertens **3**(5): 299-308.

England, P. and B. Kilbourne. (1989). "Occupational Measures from the Dictionary of Occupational Titles for 1980 Census Detailed Occupations."

England, P. and B. Kilbourne (1989). "Occupational Measures from the Dictionary of Occupational Titles for 1980 Census Detailed Occupations
".

Finkelstein, E. A., I. C. Fiebelkorn, et al. (2003). "National medical spending attributable to overweight and obesity: how much, and who's paying?" Health Aff (Millwood) **Suppl Web Exclusives**: W3-219-26.

Finkelstein, E. A., C. J. Ruhm, et al. (2005). "Economic causes and consequences of obesity." Annu Rev Public Health **26**: 239-57.

Ford, E. S., D. F. Williamson, et al. (1997). "Weight change and diabetes incidence: findings from a national cohort of US adults." Am J Epidemiol **146**(3): 214-22.
French, S. A. (2003). "Pricing effects on food choices." J Nutr **133**(3): 841S-843S.
Gelbach, J. B., J. Klick, et al. (2007). Cheap Donuts and Expensive Broccoli: The Effect of Relative Prices on Obesity. Public Law Research Paper, FSU College of Law.
Goldman, D. P., B. Shang, et al. (2005). "Consequences of health trends and medical innovation for the future elderly." Health Aff (Millwood) **24 Suppl 2**: W5R5-17.
Goodman, C. and A. Anise (2006). What is known about the effectiveness of economic instruments to reduce consumption of foods high in saturated fats and other energy-dense foods for preventing and treating obesity? Copenhagen, WHO Regional Office for Europe
Gray, D. S. and K. Fujioka (1991). "Use of relative weight and Body Mass Index for the determination of adiposity." J Clin Epidemiol **44**(6): 545-50.
Gruber, J. and M. Frakes (2006). "Does falling smoking lead to rising obesity?" J Health Econ **25**(2): 183-97; discussion 389-93.
Harris, T. B., L. J. Launer, et al. (1997). "Cohort study of effect of being overweight and change in weight on risk of coronary heart disease in old age." BMJ **314**(7097): 1791-4.
Havlik, R. J., H. B. Hubert, et al. (1983). "Weight and hypertension." Ann Intern Med **98**(5 Pt 2): 855-9.
Higgins, M., W. Kannel, et al. (1988). "Hazards of obesity--the Framingham experience." Acta Med Scand Suppl **723**: 23-36.
Hill, R. J. (2006). "Superlative Index Numbers: Not All of Them Are Super." Journal of Econometrics **130**(1): 25-43.
Huang, Z., W. C. Willett, et al. (1998). "Body weight, weight change, and risk for hypertension in women." Ann Intern Med **128**(2): 81-8.
Hubert, H. B., M. Feinleib, et al. (1983). "Obesity as an independent risk factor for cardiovascular disease: a 26-year follow-up of participants in the Framingham Heart Study." Circulation **67**(5): 968-77.
Jacobson, M. F. and K. D. Brownell (2000). "Small taxes on soft drinks and snack foods to promote health." Am J Public Health **90**(6): 854-7.
Knowler, W. C., E. Barrett-Connor, et al. (2002). "Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin." N Engl J Med **346**(6): 393-403.
Kopelman, P. G. (2000). "Obesity as a medical problem." Nature **404**(6778): 635-43.
Lakdawalla, D. and T. Philipson (2002). The Growth of Obesity and Technological Change: A Theoretical and Empirical Examination, National Bureau of Economic Research, Inc, NBER Working Papers: 8946.
Lakdawalla, D. and T. Philipson (2007). "Labor Supply and Weight." Journal of Human Resources **42**(1): 85-116.
Lakdawalla, D. N., J. Bhattacharya, et al. (2004). "Are the young becoming more disabled?" Health Aff (Millwood) **23**(1): 168-76.
Lakdawalla, D. N., D. P. Goldman, et al. (2005). "The health and cost consequences of obesity among the future elderly." Health Aff (Millwood) **24 Suppl 2**: W5R30-41.

Lakdawalla, D. N. and T. J. Philipson (2006). Economics of Obesity. The Elgar Companion to Health Economics. A. M. Jones. Northampton, MA, Edward Elgar: 72-82.
Lakdawalla, D. N. and T. J. Philipson (2006). "Labor Supply and Weight." Journal of Human Resources **forthcoming**.
Lopez, R. (2004). "Urban sprawl and risk for being overweight or obese." Am J Public Health **94**(9): 1574-9.
MacMahon, S. W., R. B. Blacket, et al. (1984). "Obesity, alcohol consumption and blood pressure in Australian men and women. The National Heart Foundation of Australia Risk Factor Prevalence Study." J Hypertens **2**(1): 85-91.
Mello, M. M., D. M. Studdert, et al. (2006). "Obesity--the new frontier of public health law." N Engl J Med **354**(24): 2601-10.
NCHS (2007). Health, United States, 2007, With Chartbook on Trends in the Health of Americans, Hyattsville, MD.
Nestle, M. and M. F. Jacobson (2000). "Halting the obesity epidemic: a public health policy approach." Public Health Rep **115**(1): 12-24.
Newhouse, J. P. and I. E. G. The (1993). Free for all? Lessons from the RAND Health Insurance Experiment, A RAND Study. Cambridge and London: Harvard University Press.
NIDDK (2006). Understanding Adult Obesity, NIH.
NIH (1998). Clinical Guidelines on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults--The Evidence Report. National Institutes of Health, National Institute of Health.
Norton, E. and E. Han (2007). "Genetic Information, Obesity, and Labor Market Outcomes
" Health, Econometrics and Data Group (HEDG) Working Papers **HEDG**(07-15).
Ogden , C., M. Carroll , et al. (2007). Obesity among adults in the United States- no change since 2003-2004. NCHS data brief
Philipson, T. J. and R. A. Posner (1999). The Long-Run Growth in Obesity as a Function of Technological Change, National Bureau of Economic Research, Inc, NBER Working Papers: 7423.
Powell, L. M., S. Slater, et al. (2007). "Food store availability and neighborhood characteristics in the United States." Prev Med **44**(3): 189-95.
Rashad, I. and S. Markowitz (2007). Incentives in Obesity and Health Insurance, National Bureau of Economic Research, Inc, NBER Working Papers: 13113.
Revicki, D. A. and R. G. Israel (1986). "Relationship between body mass indices and measures of body adiposity." Am J Public Health **76**(8): 992-4.
Rimm, E. B., M. J. Stampfer, et al. (1995). "Body size and fat distribution as predictors of coronary heart disease among middle-aged and older US men." Am J Epidemiol **141**(12): 1117-27.
Schroeter, C., J. Lusk, et al. (2007). "Determining the impact of food price and income changes on body weight." J Health Econ.
St.Clair, P., et al (2005). RAND HRS Data Documentation, Version E. Santa Monica, CA, RAND Center for the Study of Aging.
Stamler, R., J. Stamler, et al. (1978). "Weight and blood pressure. Findings in hypertension screening of 1 million Americans." JAMA **240**(15): 1607-10.

Sturm, R. and A. Datar (2005). "Body mass index in elementary school children, metropolitan area food prices and food outlet density." Public Health **119**(12): 1059-68.

Sturm, R., J. S. Ringel, et al. (2004). "Increasing obesity rates and disability trends." Health Aff (Millwood) **23**(2): 199-205.

Thompson, D., J. Edelsberg, et al. (1999). "Lifetime Health and Economic Consequences of Obesity." Arch Intern Med **159**(18): 2177-2183.

Thorpe, K. E., C. S. Florence, et al. (2004). "The impact of obesity on rising medical spending." Health Aff (Millwood) **Suppl Web Exclusives**: W4-480-6.

Tuomilehto, J., J. Lindstrom, et al. (2001). "Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance." N Engl J Med **344**(18): 1343-50.

Verbeek, M. and T. Nijman Incomplete Panels and Selection Bias. L. Matyas and P. e. Sevestre.

Willett, W. C., J. E. Manson, et al. (1995). "Weight, weight change, and coronary heart disease in women. Risk within the 'normal' weight range." JAMA **273**(6): 461-5.

Wolf, A. M. and G. A. Colditz (1998). "Current estimates of the economic cost of obesity in the United States." Obes Res **6**(2): 97-106.

Wooldridge, J. M. (2002). Econometric analysis of cross section and panel data, Cambridge and London: MIT Press.

Yang, Z. and A. G. Hall (2008). "The financial burden of overweight and obesity among elderly Americans: the dynamics of weight, longevity, and health care cost." Health Serv Res **43**(3): 849-68.