The Importance of Objective Health Measures in Predicting Early Receipt of Social Security Benefits: The Case of Fatness

Richard V. Burkhauser and John Cawley
Cornell University

Prepared for the 8th Annual Joint Conference of the Retirement Research Consortium
"Pathways to a Secure Retirement"
August 10-11, 2006
Washington, DC

The research reported herein was pursuant to a grant from the U.S. Social Security Administration (SSA) funded as part of the Retirement Research Consortium (RRC). The findings and conclusions expressed are solely those of the authors and do not represent the views of SSA, any agency of the Federal Government, or the RRC. We thank Maki Ueyama for research assistance and Krisy Gashler for administrative assistance. For helpful comments we thank John Bound, Charlie Brown, Dean Lillard, Bob Schoeni, Dan Silverman, and conference participants at the University of Michigan Retirement Research Center.

Abstract

Theoretical models argue that poor health will contribute to early exit from the labor market and the decision to take early Social Security retirement benefits (Old-Age or OA benefits). However, most empirical estimates of the causal importance of health on the decision to take early OA benefits have been forced to rely on global measures such as self-rated work limitations or self-rated health. We contribute to the empirical literature by using a more objective measure of health, fatness, to predict early receipt of OA benefits. We do so by estimating the causal impact of fatness within an empirical model using the method of instrumental variables, and testing the robustness of our findings using the most common measure of fatness in the social science literature – body mass index – with what is a more theoretically appropriate measure of fatness – total body fat and percent body fat. Overall, our conclusion is that fatness and obesity are strong predictors of early receipt of OA benefits.

Introduction

A substantial literature has established that health is a critical factor in the decision of workers to apply for early Social Security Old-Age (OA) benefits (see Lumsdaine and Mitchell, 1999, for a review of this literature and Currie and Madrian, 1999 for a broader review of the econometric problems of estimating the effect of health in labor supply models). The vast majority of studies in this literature have used subjective measures of health such as self-rated work limitations or self-rated health. However as Bound, Steinbricker and Waidmann (2006) recently point out, there are three potential problems with such survey measures: they are discrete, whereas the construct researchers are interested in measuring is presumably continuous; they are presumably error ridden, since not everyone will use the same scale when responding to survey questions; and they are likely to be endogenous to retirement decisions, since it seems plausible that responses to these global questions will be related to labor market status.

In this paper we use data from various years of the Panel Study of Income Dynamics (PSID) to utilize more objective measures of health to model early receipt of OA benefits in a manner that addresses each of the three objections outlined by Bound et al. (2006) above. Early receipt of OA benefits, not early retirement, is our outcome of interest both because the timing of OA benefits is well-defined whereas there are many competing definitions of retirement, which complicates the study of that outcome and because the timing of OA acceptance is of policy concern in and of itself.

In 1986, the PSID first obtained self reported weight and height information from its respondents and in many cases the weight and height of their biological relatives. We take advantage of these variables to create alternative measures of fatness, which the medical literature has consistently found to have adverse effects on health. We estimate the causal impact of various measures of fatness and obesity on early receipt of OA benefits. That is, we model the decision of males to first take OA benefits at their earliest possible age of eligibility—age 62.

Due to the rising prevalence of obesity, it is important to determine the role that fatness plays in the decision to take early OA benefits, since recent changes in Social Security retirement rules, which are gradually pushing back the normal retirement age for OA benefits to 67 and lowering the actuarial value of taking Social Security benefits at

age 62, were in part passed to discourage further cohorts of older workers from exiting the labor force and taking OA benefits at age 62.

A recent literature has linked fatness or obesity to a variety of social science outcomes such as wages (Cawley, 2004), disability (Lakdawalla, Bhattacharya, and Goldman, 2004; Ferraro et al., 2002; Cawley, 2000; Narbro et al., 1996), and the transition from welfare to work (Cawley and Danziger, 2005). However, this is the first paper to examine the link between fatness or obesity and early receipt of OA benefits. Olshansky et al. (2005) projected the implications for the recent rise in obesity for the Social Security program, but they focus on how obesity impacts mortality and not on how obesity might affect the timing of OA benefit receipt. This paper fills that gap by estimating the correlation and causal impact of fatness and obesity on whether males start taking OA benefits when they first become available at age 62. To preview our results, we find that fatness and obesity are strong predictors of early receipt of OA benefits. For example, obesity defined using body mass index is associated with a roughly 25 percentage point higher probability of taking OA benefits at age 62.

Fatness and Obesity: Definitions, Trends, Implications, and Measurement

A wide variety of social science outcomes are affected by health (Culyer and Newhouse, 2000), and one important dimension of health is fatness. Fatness is a concept that refers to the abundance of adipose tissue, in which energy is stored in the form of fat cells (Bjorntorp, 2002). Fatness is a risk factor for ischemic heart disease, congestive heart failure, stroke, cancer, respiratory disease, diabetes, hyperlipidemia, hypertension, asthma, sleep apnea, arthritis, degenerative joint disease, gastric reflux, and depression (Pi-Sunyer, 2002; U.S. D.H.H.S., 2001; NIH, 1998). The current view of fatness in the medical literature is that fat collectively constitutes an endocrine organ that secretes leptin, which damages the cardiovascular system, and resistin, which causes insulin resistance and Type II diabetes (Trayhurn and Beattie, 2001). The link between fatness and morbidity has been confirmed in hundreds of randomized controlled trials in the medical literature showing the correlation between weight loss and improvement in obesity-related comorbidities (NIH, 1998; U.S. DHHS, 2001).

The health impact of fatness is particularly troubling because the average weight of Americans has risen considerably in the past few decades (Hedley et al., 2004; Ogden et al., 2006). Given the link between fatness and morbidity and mortality, excessive fatness is now recognized as one of the most serious public health challenges facing the U.S. (U.S. DHHS, 2001) and other industrialized countries (International Obesity Task Force, 2005).

In this paper, we investigate the best way to measure fatness for the purpose of better predicting early receipt of OA benefits. To date, fatness has almost universally been measured in the social science literature using body mass index (BMI), which is weight in kilograms divided by height in meters squared (U.S. DHHS, 2001; NIH, 1998).[2] The advantage of BMI is that the information required to calculate it (weight and height) is easy to collect and relatively common in social science datasets such as the National Longitudinal Surveys of Youth (NLSY), the Panel Study of Income Dynamics (PSID), the Health and Retirement Study (HRS), the Behavioral Risk Factor Surveillance System (BRFSS), the National Health Interview Survey (NHIS), and the National Longitudinal Survey of Adolescent Health (Add Health).

Despite the widespread use of BMI among social scientists, within the medical literature BMI is considered to be a very limited measure of fatness and obesity because it does not distinguish body composition (McCarthy et al., 2006; Yusuf et al 2005; Gallagher et al 1996; Smalley et al 1990; Garn et al 1986). For example, it overestimates fatness among those who are muscular (U.S. DHHS, 2001; Prentice and Jebb, 2001). Gallagher et al. (1996) calculated that BMI alone accounts for just 25 percent of between-individual differences in percent body fat.

Obesity is a concept that refers to excessive fatness (Bjorntorp, 2002; Bray, Bouchard, and James, 1998). The most common definition of obesity used in the social science literature is based on BMI: a BMI greater than or equal to 30. However, there are a variety of definitions of obesity, corresponding to the various measures of fatness, and the strengths and weaknesses of each definition of obesity depend on the strengths and weaknesses of the fatness definition on which it is based. For this reason, the clinical

[2] For example, EconLit lists 55 articles with the words "body mass index" or "BMI" in the abstract or among the keywords, but zero articles with the more accurate measures of fatness in the abstract or among the keywords.

weight classification of obesity that is based on BMI suffers the same limitation as BMI: it ignores body composition. Smalley et al (1990) tested the accuracy of BMI-based definitions of obesity at identifying those determined to be obese through measurements of body fat and found that BMI correctly identified only 44.3 percent of obese men and 55.4 percent of obese women. Put another way, a majority of truly obese men are misclassified as non-obese if one uses the BMI-based definition of obesity. Subsequent studies in the medical literature have confirmed this finding (Wellens et al 1996).

The use of BMI to classify people as obese also results in false positives because people who are muscular but not fat have a higher BMI; these false positives totaled 9.9 percent of non-obese men and 1.8 percent of non-obese women (Smalley et al, 1990). (Women are less likely to be inaccurately classified as obese on the basis of BMI because they are less likely to be heavily muscular.) Based on all of these findings, the researchers concluded that the ability of BMI in particular, and weight-height indices in general, to identify obesity defined using direct measures of fatness is "poor" (Ibid, p. 408). Moreover, the inferiority of BMI at predicting health outcomes relative to more accurate measures of fatness led a 2005 editorial in the British medical journal *The Lancet* to conclude "…current practice with body-mass index as the measure of obesity is obsolete, and results in considerable underestimation of the grave consequences of the overweight epidemic" (Kragelund and Omland, 2005, p. 1590).

Despite the skepticism in the medical literature toward BMI as a measure of either fatness or obesity, virtually no tests of the robustness of social science-based findings using these more accurate measures of fatness have been undertaken on social science-based outcomes.[3] Here we conduct such a test with respect to one important outcome: early receipt of Social Security benefits.

More Accurate Measures of Fatness

While there is consensus in the medical literature that BMI is a poor measure of fatness (McCarthy et al., 2006; Yusuf et al 2005; Gallagher et al 1996; Smalley et al

[3] To our knowledge, the only such study is Wada (2005), which uses the NHANES III to estimate total body fat and fat-free mass in the NLSY but is focused more narrowly to explore the differences across race and gender in the correlation between BMI and wages found in Cawley (2004) and Averett and Korenman (1999).

1990; Garn et al 1986), there is no consensus on which of the more accurate measures of fatness is best (Freedman and Perry, 2000). Candidates include: total body fat (TBF), percent body fat (PBF) which is total body fat divided by total mass, waist circumference (WC), and waist-to-hip ratio (WHR).

Total body fat (TBF) and percent body fat (PBF) are appealing measures of fatness because the medical literature suggests that it is fat that causes morbidity and mortality (Pi-Sunyer, 2002; U.S. DHHS, 2001). For example, Trayhurn and Beattie (2001) argue that fat directly causes Type II diabetes and cardiovascular disease by secreting resistin and leptin; these findings suggest that TBF may be the most relevant measure of fatness for predicting social science outcomes affected by health because the sheer volume of fat may determine the amount of leptin and resistin released; on the other hand PBF may be a better measure if additional fat-free-mass can dilute the health impacts of those secretions.

Findings from the medical literature also suggest that it is not just the amount of fat that matters, but also the location or distribution of that fat. In particular, abdominal visceral fat (i.e. that located around the internal organs) is associated with an elevated risk of morbidity (Bray, Bouchard, and James, 1998). The amount of abdominal visceral fat can be assessed using laboratory methods like dual-energy X-ray absorptiometry, but in practice it is frequently measured using either waist circumference or waist-to-hip ratio; comparisons have found that these two are highly correlated with abdominal fat (Snijder et al., 2002). Yusuf et al. (2005) conclude that by a variety of standards, waist-to-hip ratio (WHR) and, to a lesser extent, waist circumference better predict heart attack than does BMI; an accompanying comment in *The Lancet* entitled "A Farewell to Body-Mass Index?" concluded that these findings represent "…the final nail in the casket for body-mass index as an independent cardiovascular risk factor…" (Kragelund and Omland, 2005, pp. 1589, 1590).

While it is generally accepted that central adiposity (abdominal fat) is associated with greater risk of morbidity and mortality, it is not clear that waist-to-hip ratio is the best way to measure it. For example, a 1998 NIH report entitled *Clinical Guidelines on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults*

recommends the use of waist circumference rather than waist to hip ratio to measure central adiposity (NIH, 1998, p. xxiv).

Until recently the costs of collecting both measurements of fatness and in-depth social science outcomes in one dataset were prohibitive. As a result, medical datasets that include direct measures of fatness contain few outcomes of interest to social scientists, and social science datasets do not contain accurate measures of fatness such as TBF, PBF, WC, or WHR. We expect this to change as social science researchers become more aware of the benefits of more accurate measures of fatness and as the costs of collecting them continue to decline. Until then however, TBF and FFM can be estimated in social science datasets using a two-step method. We use the NHANES III to generate prediction equations for these more accurate measures of fatness; this method is described in detail in our recent NBER working paper (Cawley and Burkhauser, 2006).

Data: The Panel Study of Income Dynamics (PSID)

The PSID is administered by the Survey Research Center at the University of Michigan. The PSID began in 1968 with a sample of 5,000 families. All current PSID families contain at least one member who was either part of the original 5,000 families or was born to a member of one of those families. Although the original sampling scheme disproportionately selected individuals from low-income families, a representative sample of the U.S. population can be obtained by excluding the original oversample from the data or by utilizing the sample weights included with the data. Starting in 1997 the PSID began to administer its survey every other year and some members of the low-income sample were dropped from the study because resources were not available to continue to follow all members of the ever expanding sample.

The PSID collected information on each respondent's weight and height in 1986, 1999, 2001, 2003, and 2005 (which is not yet publicly available). A unique strength of the PSID for our research question is that it contains the weight and height of biological relatives for many respondents. This is available two ways. First, data on biological relatives is available as a natural byproduct of the PSID design, which follows families that spin off from the original 1968 families. As a result, the children of PSID respondents who become themselves heads of families remain in the survey, and their

self-reported weight and height is recorded. The Family Identification Mapping System (FIMS) is used to merge data on adult children of respondents. The FIMS provides identification codes for each of a respondent's family members by type of relationship (e.g. biological parent, non-biological parent, biological grandparent, full sibling, half sibling). The FIMS ensures that our linking of fathers to adult children is straightforward and accurate.

Second, the PSID collected data on the height and weight of children in the Child Development Supplements (CDS) I (conducted in 1997) and II (conducted in 2002-2003). The CDS data are not useful for our project because respondents who reach age 62 by 2003 and have reported their height and weight in 1987 prior to reaching age 62 are generally too old to have had minor children in the household in 1997.

A weakness of the PSID is that it records individual receipt of Social Security benefits and the type of benefit (old age, survivor, disability, Supplemental Security Income, etc.) only in certain years. Hence for male heads that turn age 62 in 1993 or earlier we know whether or not they take OA benefits for the first time at age 62. After 1993, questions are asked regarding whether Social Security benefits are received by any member of the family but not about whom in the family receives these benefits or the type of Social Security benefit they receive.

Thus, after 1993, we can not directly identify who in the family is receiving Social Security benefits. Hence for those male heads that turn age 62 after 1993 we use these more limited data in the following way to determine whether the family head first began receiving OA benefits at age 62. We restrict our sample to male family heads who turn 62 before 2003, and whose family is not receiving Social Security benefits when he is aged 61. This conservative method of allowing male heads into our sample will reduce sample size but will also reduce the risk of identifying a male head as first receiving OA benefits at age 62 when in fact those benefits are coming from an alternative source. To be in our sample, no family member can be receiving benefits when the male head is 61. Our presumption is that the appearance of Social Security benefits in the family for the first time in the year when the male head turns age 62 can plausibly be assumed to be caused by his taking early OA benefits. Of course, the tradeoff is that we will inappropriately exclude some male heads from our sample because at age 61 someone

else in the family is receiving some alternative form of Social Security benefits. We will also exclude all male heads who were receiving Disability Insurance (DI) benefits at age 61.

To clarify, when possible, we use the PSID questions that are specific to the head to determine receipt of OA benefits at age 62. When individual-specific questions are not available (i.e. after 1993), we impute receipt of such benefits in the manner described above.

We limit our sample to male heads of households because in this cohort, the males are considerably more likely to be the first to reach age 62 (i.e. to be older than their wives) and so we can observe families transitioning from not receiving Social Security benefits when the male head is 61 to receiving Social Security benefits when the male head turns 62. This is strong evidence that the male head began receiving OA benefits at age 62. In contrast, if the husband is older than the wife, the household is likely to already be receiving Social Security benefits by the time the wife turns 62 and therefore it is harder to determine whether the wife begins receiving OA benefits at 62. Summary statistics for our PSID sample are provided in Appendix Table 1.

An alternative dataset that could be used to study early receipt of OA benefits is the Health and Retirement Study (HRS). The HRS offers many advantages over the PSID: it is a larger sample, it more clearly assigns type of Social Security benefit received to specific individuals, it has linked Social Security administrative records on benefit timing and receipt. The HRS, like the PSID, also includes data on the weight and height of respondents. However, the PSID offers one key advantage over the HRS that is of critical importance for our research agenda: the PSID includes the weight and height of biological relatives, which allows us to estimate IV models and make statements about causality.

Data: National Health and Nutrition Examination Survey III (NHANES III)

The NHANES III is a nationally representative cross-sectional survey conducted from 1988 to 1994. All respondents were asked to complete an extensive interview (during which they were asked to report their weight and height) and undergo a subsequent medical examination in a large mobile examination center (during which their

weight and height were measured). The NHANES III sample consists of 31,311 examined respondents. In this paper we use NHANES III adults aged 18-65.

The NHANES III is the "Rosetta Stone" for estimating more accurate measures of fatness, because it includes the data necessary to calculate many measures of fatness; it includes: self-reported weight and height, measured weight and height, measured waist circumference, measured waist-to-hip ratio, and Bioelectrical Impedance Analysis (BIA) readings that can be used to calculate fat-free mass and therefore total body fat and percent body fat (we explain this calculation in the next section).[4] This allows researchers to regress measured fatness on self-reported weight and height, transport the coefficients on the self-reported variables to any social science dataset that includes self-reported values, and to then construct estimates of measured fatness. We drop as implausible observations with self-reported height either under four feet or over seven feet (one observation of each) or with self-reported weight under 80 pounds (one observation).

As part of the examination for respondents 12 years of age and older, measurements from Bioelectrical Impedance Analysis (BIA) were recorded. NHANES examiners attached a pair of electrodes to the right wrist and ankle, and passed a very small (50 kHz) electrical current through the body and measured the resistance of the body to the current. These measurements can be used to calculate fat and fat-free mass because the resistance to an electric current is inversely related to the amount of fat-free mass in the body; the water in muscles conducts electricity while fat is an insulator.

BIA is a well-established method widely used to measure body fat, and considerable planning and training were devoted to maximizing the validity and

[4] There are a variety of ways of measuring TBF and therefore PBF, which range from methods that use very expensive equipment that can be used only in a lab setting and which require subject cooperation or exposure to radiation (e.g. magnetic resonance imaging or MRI, dual x-ray absorptiometry) to more portable (field-based) methods that are less expensive and rapid (e.g. Bioelectrical Impedance Analysis or BIA) (Freedman and Perry 2000). In this paper, we use measures of TBF and PBF that were estimated by BIA. Each method of measuring body composition has its pros and cons (Freedman and Perry 2000); for example, the BIA method of estimating fat-free mass is less accurate for the severely obese (NIH, 1996). However, despite this limitation, the NIH endorses BIA as a useful technique for measuring body fat and body composition generally (NIH 1996). Prentice and Jebb (2001) conclude: "Bioimpedance is probably the only technique that can meet the criteria of being simple, rapid, and free from operator variability" (page 146). However, it is not our position that BIA is preferable to all other methods of measuring body fat; we use it because it is one method endorsed by the NIH, it is the only method for which data is now available in a large, nationally representative U.S. dataset - the National Health and Nutrition Examination Surveys (U.S. DHHS, 1994).

reliability of the BIA measurements taken as part of the NHANES (NIH, 1996; US DHHS, 1994).[5] BIA readings are missing for many NHANES III respondents. We have complete data for 2,138 white females, 1,861 African American females, 1,911 white males, and 1,634 African American males. BIA readings are converted into measures of Fat Free Mass using the equations provided in Sun et al. (2003); for more detail on this procedure see Cawley and Burkhauser (2006). BIA predicts fat-free mass, but total body fat is easily determined using the identity below:

$$\text{Total body fat} = \text{weight} - \text{fat-free mass}$$

and:

$$\text{Percent body fat} = (\text{Total body fat} / \text{weight}) * 100$$

We cannot calculate FFM or TBF in NHANES III for Mexican Americans because the sample used to generate the equations provided in Sun et al. (2003) excluded Hispanics.

The NHANES III contains little data on labor market activity or other social science outcomes, so in isolation it is not of much value in conducting social science research. However, in conjunction with social science datasets such as the PSID it is very useful, as it allows us to combine its rich biomarker data with the rich data on employment, relationships, and other outcomes in social science datasets.

We take into account the complex survey design of the NHANES III by estimating our models using svy commands in Stata version 9.2 that account for the strata, primary sampling units, and sample weights of the NHANES III. Because we look separately at race-gender cells, and there are missing values of BIA readings for many respondents, there were instances in which strata had only one primary sampling unit, which violates the requirements of Stata's complex survey design commands. Rather than drop these observations and lose information, we merge the stratum with only one PSU with another stratum, and give the PSU a unique identifier within the new stratum.[6]

[5] Although BIA readings were also included in each of the annual NHANES surveys since 1999, the NHANES III is the most recent survey for which there are consensus prediction equations in the medical literature to convert these readings to Total body fat and Fat-free mass.

[6] This was not a problem for white females, but there was one such stratum (with three observations) for white males, six such strata (with 21 observations) for African American females, and five such strata (also with 21 observations) for African American males.

We use the WTPFEX6 sample weight, which is recommended when the medical examination variables (such as weight and BIA measurements) are studied.

Using the NHANES III data, we regress total body fat (TBF) and fat-free mass (FFM) on self-reported weight and its square, self-reported height and its square, and age and its square. The set of regressors in this NHANES III regression consists entirely of variables also available in social science datasets like the PSID, NLSY, and HRS. Since we found that the coefficients in these regressions differ significantly by race and sex, we estimated models separately for white females, white males, African American females, and African American males. (Because there were no Hispanics in the Sun et al. (2003) sample, we restrict our analysis to whites and African Americans.)

The coefficients from these NHANES III regressions can be transported to various social science datasets and multiplied by the values of the relevant variables to predict TBF and FFM.[7] We find that self-reported weight, self-reported height, and age are good predictors of total body fat; R-squared values range from .76 to .90. This indicates that the basic self-reported weight data available in several social science datasets are sufficient to accurately estimate TBF. The the fit is better for women (around 89 percent of variance explained) than for men (77 percent).

The results also indicate that self-reported weight, self-reported height, and age are good predictors of fat-free mass; the R-squares are in the narrow range of .81 to .82. (Unlike the TBF results, the goodness of fit in the FFM regressions is approximately equal for men and women.) This confirms that the basic self-reported weight data available in several social science datasets are also sufficient to accurately estimate FFM.

Researchers can utilize the full variation in PBF, or they can convert it into an indicator variable for obesity. The NIH classifies a man as obese if his PBF exceeds 25 percent and a woman as obese if her PBF exceeds 30 percent (NIDDK, 2006).

Methods

[7] Although TBF and FFM are in kilogram units, the transformation between pounds and kilograms is linear so we keep the self-report of weight in pounds since those are the units in which self-reported weight is recorded in most social science datasets.

We assume that individual i takes early OA benefits at time t if his health H_{it} falls below some critical limit H^*. Health is assumed to be a function of fatness F_{it} and other characteristics X_{it}. Specifically:

$$H_{it} = F_{it}\,\beta + X_{it}\,\delta + u_{it}$$

Health H is not observed, but we know whether an individual takes early OA benefits; we denote $SS_{it}=1$ if individual i is disabled at time t, and $SS_{it}=0$ otherwise. Formally, early OA receipt relates to latent health in the following way:

$$SS_{it} = 0 \text{ if } H_{it} \geq H^*$$

$$SS_{it} = 1 \text{ if } H_{it} < H^*$$

Normalizing H^* at $H=0$, the probability that one takes early OA benefits is equal to the following.

$$\begin{aligned} \Pr[SS_{it} = 1 \mid X_{it}] &= \Pr[H_{it} < 0] \\ &= \Pr[F_{it}\,\beta + X_{it}\,\delta + u_{it} \leq 0] \\ &= \Pr[u_{it} < -F_{it}\,\beta - X_{it}\,\delta] \end{aligned}$$

$$\Pr[SS_{it} = 1 \mid F_{it}, X_{it}] = \Pr[u_{it} < -F_{it}\beta - X_{it}\delta] \tag{0.1}$$

With certain assumptions about the distribution of the error term u, one can estimate the probability of taking OA as a function of fatness F and characteristics X using probit regression.

We estimate model (0.1) for the following measures of fatness: total body fat, percent body fat, body mass index, and clinical weight classifications based on BMI. While the PSID does not have measures of TBF or PBF, we estimate them in the PSID by transporting coefficients from regressions of TBF and PBF on self-reported weight in the NHANES III. The NHANES III coefficients are multiplied by the values of the same variables in the PSID to estimate TBF and PBF for PSID respondents.

Probit models will be used to estimate the correlation of each of these binary outcomes with TBF, PBF, FFM, BMI, or clinical weight classification based on BMI. Goodness-of-fit measures (e.g. percent of correct predictions) will be used to determine which of the anthropometrics best predicts our outcomes of interest.

The primary limitation of the probit models is that the correlations reflected in the probit coefficients cannot be interpreted as causal. To classify the sources of potential endogeneity in weight, the residual probability of early receipt u_{it} in equation (0.1) can be

decomposed as having a genetic component G^{SS}, a non-genetic component NG^{SS}, and a residual ν that is i.i.d. over individuals and time.

$$u_{it} = G_{it}^{SS} + NG_{it}^{SS} + \nu_{it} \tag{0.2}$$

Research in behavioral genetics suggests that roughly 25-40 percent of the variation in weight is due to genetics while twin studies suggest that the correct figure is between 50-80 percent (Bouchard et al., 2003).

Equation (0.1) indicates how early receipt of OA benefits may be correlated with fatness. Fatness, in turn, may be affected by early receipt of OA and personal characteristics.

$$F_{it} = X_{it}\gamma + SS_{it}\alpha + Z_{it}\phi + G_{it}^{F} + NG_{it}^{F} + \xi_{it} \tag{0.3}$$

In equation (0.3), X is the same vector of variables that affect early receipt of OA benefits in equation (0.1), SS is an indicator variable that equals one if the respondent receives early OA benefits, Z is a vector of variables that affect fatness but do not directly affect the probability of taking early OA benefits, G^F represents the influence of genetics on fatness, NG^F represents the influence of non-genetic factors (such as an individual's choices, upbringing, and culture) on fatness. Residual fatness is represented by ξ.

The variables on the right-hand-side of equation (0.3) illustrate the potential pitfalls of a probit estimation of equation (0.1). First, early receipt of OA benefits may affect fatness (if $\alpha \neq 0$). For example, early receipt may be associated with retirement, and presumably a decrease in caloric expenditure. If the decrease in exercise is not offset with a decrease in caloric intake, a gain in weight and fatness will result. A second potential pitfall of a probit estimation of equation (0.1) is that genetic factors that influence fatness (G^F) may be correlated with genetic factors that affect early receipt of OA benefits (G^{SS}). A third potential pitfall is that non-genetic factors that influence fatness (NG^F) may be correlated with non-genetic factors that affect early receipt of OA benefits (NG^{SS}). For example, people may differ in the extent to which they think about and plan for distant future outcomes. People who are farsighted and value outcomes in the future are likely to have more successful careers (that they are less likely to retire early from) and also to stay in shape. In contrast, those who are myopic and do not care

about future consequences may be more likely to retire early and more likely to become overweight. Each of these scenarios implies that the assumption that F is uncorrelated with ε in equation (0.1) is violated and that a probit estimate of β is biased.

Method of Instrumental Variables

Perhaps the most convincing research design that would allow one to determine whether fatness or obesity causally affected early OA receipt would be a randomized controlled trial. As a thought experiment, imagine that a large number of randomly selected subjects were randomly assigned between treatment and control groups. The treatment group would then be "treated" with additional fatness or weight and the control group would not. One could measure the causal impact of fatness or weight on early receipt of OA benefits by comparing changes in outcomes for the treatment group to changes in outcomes for the control group.

While that research design would be convincing, it is neither ethical nor feasible to endow a treatment group with additional weight. Because of this, randomized controlled trials can never be used to answer research questions like the ones we pose. Researchers are forced to look for alternatives to randomized controlled trials. One viable alternative is to find a natural experiment—some variation in weight that was not chosen. Our strategy for identifying the causal impact of fatness or weight on early receipt of OA benefits is to exploit one such natural experiment: the variation in weight due to genetics. Each person is endowed before their birth with a set of genes that imply a certain predisposition to fatness or body weight. We explain below how we will exploit the genetic variation in weight to answer our research question.

To eliminate these influences and thereby generate a consistent estimate of the impact of fatness on early receipt of OA benefits, we estimate a model of instrumental variables (IV). This becomes possible using IV one can identify a variable or set of variables Z from equation (0.3) that are correlated with body weight but not u_{it}, the error term in the early OA receipt regression (equation (0.1)).

We will use an instrument correlated with the genetic variation in fatness (G^F): the fatness of an adult biological child. In the PSID, children are tracked and surveyed after they form their own families. The Family Information Mapping System of the PSID ensures that we are matching biological (as opposed to adopted or step-) children to

respondents. The fatness of the adult biological child could be correlated with the father's personal characteristics, early receipt of Social Security benefits, and genes:

$$F_{Ct} = X_{Ct}\gamma + D_{Ct}\alpha + G^F_{Ct} + NG^F_{Ct} + \xi_{Ct}$$

The subscript C indicates the biological child of the respondent *i*.

The identifying assumption in our method of instrumental variables has two parts. The first is that the fatness of a biological child is strongly correlated with the fatness of the respondent. A parent and child are expected to share half of their genes, ensuring a high correlation between these relatives' genetic variation in fatness G^F_C and G^F_i. An extensive review of the genetics literature concluded that roughly 25-40 percent of the variation in body fat (and weight) is due to genetics (Bouchard et al., 1998). These findings imply a strong correlation between the fatness of the respondent F_i and the fatness of a biological child F_C that is consistent with the first part of our identifying assumption.

The second part of the identifying assumption is that the fatness of a biological child is uncorrelated with u_{it}, the respondent's residual probability of early receipt of Social Security benefits. One might be concerned that the nongenetic variation in the biological child's fatness NG^F_C is correlated with the respondent's early receipt of benefits through the nongenetic variation in the respondent's probability of early receipt NG^{SS} if both are, in part, determined by habits learned in the parents' family.

However, studies have been unable to detect any effect of common family environment on body weight (Hewitt 1997; Grilo and Pogue-Geile, 1991). (To clarify, individual environment has been found to have a significant influence on weight and obesity, but the environment common to members of a family has not.) Adoption studies have consistently found that the correlation in BMI between a child and his biological parents is the same regardless of whether the child grew up in the home of the biological parents or with unrelated adults (i.e. was adopted); i.e. all of the correlation in weight between biological relatives can be attributed to shared genes with no detectable effect of shared family environment. This has been found for BMI (Vogler et al., 1995), weight class (Stunkard et al., 1986), and even body silhouette (Sorensen and Stunkard, 1993).

Consistent with these findings, studies have been unable to reject the hypotheses that the correlations in weight, weight for height, and skinfold measures between unrelated adopted siblings are equal to zero (Grilo and Pogue-Geile, 1991). Studies of twins reared apart also find no effect of a shared family environment on BMI; there is no significant difference between the correlation in weight of twins reared together and those reared apart, nor is the correlation affected by age at separation or the similarity of separate rearing environments (Maes et al., 1997; Price and Gottesman, 1991). Other studies have found that genes explain large fractions of variance in diet and eating behaviors and no detectable impact of shared environment (Tholin et al 2005; Hur et al 1998). A comprehensive review of studies of the genetic and environmental influences on weight and obesity concludes that "...only environmental experiences that are not shared among family members appear to be important. In contrast, experiences that are shared among family members appear largely irrelevant in determining individual differences in weight and obesity" (Grilo and Pogue-Geile, 1991, p. 520). Similarly, Hewitt (1997) refers to "the impotence of the shared family environment" for obesity (p. 353). Finally, we use as an instrument the weight of an adult child (who in our sample average 29 years old in 1986); as a result, they have likely not been living at the same address of the respondent for many years, further decreasing the likelihood that common environment affects both adult child weight and respondent OA receipt. Absent randomized controlled studies, it is difficult to directly test the null hypothesis of no effect of shared environment on body weight. Consequently, the repeated failure to reject the null hypothesis is the strongest evidence that will ever be available.

Alternately, one might be concerned that the genetic variation in the biological child's fatness G_C^F is correlated with the respondent's residual probability of early receipt of OA through genetic variation in respondent's likelihood of early receipt G^{SS}. For this to be true, the genes that determine fatness and any genes that determine early receipt would have to be either the same or bundled in transmission. While it is impossible to prove the null hypothesis that relative's fatness is uncorrelated with the residual in the respondent's early receipt equation, it can be informative to examine whether relative's fatness or weight is correlated with observables that are believed to be related to unobserved factors that affect the residual probability of early receipt of OA benefits.

Ideally what we would like to measure is the average treatment effect (ATE)- the impact of an extra unit of fatness on the average person - but what we measure through our method of instrumental variables is the local average treatment effect (LATE)- the impact of an extra unit of genetically-determined fatness on the average person in our sample (Imbens and Angrist, 1994). If genetically-determined fatness has a different impact on early OA receipt than does fatness caused by discretionary behavior (i.e. not genes), the LATE will differ from the ATE. To our knowledge, no research literature has documented a different impact of genetically-determined fatness than environmentally-determined or behaviorally-determined fatness, but we acknowledge that our results may represent a LATE rather than an ATE.

Our specific instrument is the fatness or obesity of an adult child (we use the same measure of fatness for the child as we do for the respondent), controlling for the adult child's age and gender. Summary statistics for the adult children are provided in the bottom of Appendix Table 1. Probit IV models are estimated using STATA version 9.

How You Measure Fatness Matters: Misclassification of Obesity

We have shown that more accurate measures of fatness can be estimated using the NHANES III and one of many social science datasets. But is worth the trouble to do it? Obesity defined using BMI, though theoretically inferior, may be a reasonable proxy, generally correctly classifying who is obese and who is not obese.

In this section we show that it is worth doing because obesity defined using BMI (calculated using measured weight and height) is only weakly correlated with obesity defined using percent body fat, and that obesity defined using BMI results in substantial misclassification of individuals into weight classifications.

Although there are many measures of fatness, because of space constraints we focus on two measures of obesity: body mass index (based on measured weight and height) greater than or equal to 30, and a more accurate measure—percent body fat greater than 25% for men or greater than 30% for women (NIDDK, 2006). The correlation between the two is relatively weak: .45 for males and .38 for females.

Taking obesity status defined using PBF to be the true obesity status, we examine how well obesity defined using BMI correctly classifies people. Table 1 shows that the

accuracy of obesity defined using BMI varies by sex. For example, only 0.17 percent of all women classified as obese by BMI are actually not obese judging by percent body fat. In contrast, 16.45 percent of all positives are false for men. One important reason for this difference by sex is that men are more likely to have considerable muscle mass.

In general, false negatives are a much bigger problem than false positives for obesity defined using BMI. In other words, many more people are obese than are classified as such by BMI. The extent of false negatives in Table 1 also varies by sex: 62.85 percent for females compared to 28.66 percent for males. Consistent with previous studies (Smalley et al., 1990; Wellens et al., 1996), we find that obesity defined using BMI does a poor job of classifying people as obese or non-obese.

One important difference between obesity defined using BMI and that defined using PBF is that the PBF-defined obesity is a lower threshold; far larger percentages of people are classified as obese using the PBF threshold. Table 2 shows that while 23.3 percent of women are classified as obese according to BMI, three times as many (70.1 percent) are classified as obese by the PBF standard. The difference for men, though not as large, is still striking: 18.9 percent are obese according to BMI, but 43.3 percent are obese according to PBF. The fact that PBF results in a strikingly higher rate of obesity is not necessarily an indictment of that measure. Yusuf et al. (2005) found that waist-to-hip ratio was a far better predictor of heart attack than BMI and concluded that BMI greatly underestimates the number of people for whom fatness impacts health. So while it is true that more people are classified as obese by PBF than BMI, it is not clear that the prevalence of obesity we have become accustomed to, because BMI is usually used to define obesity, is the "right" prevalence of obesity.

To show how similarly BMI and PBF classify people as obese when controlling for this difference in threshold, we choose cutoffs for PBF that generate the same rates of obesity as are found when one uses BMI (roughly 23.3 percent for women and 18.9 percent for men). The analogous PBF cutoff is 41 percent body fat for women and 29 percent body fat for men. Using those PBF cutoffs, we find that the correlation in obesity classifications is 70.2 percent for women and 47.1 percent for men. Treating the PBF classification as correct, Table 3 shows that the BMI classification of obesity has a false positive rate of 27.74 percent for women and 44.79 percent for men, and a false negative

rate of 5.42 percent for women and 9.31percent for men. (Lowering the PBF cutoff that defines obesity resulted in an increase in the false positives rate and a decrease in the false negative for BMI, reflecting the tradeoff between Type I and Type II error.) Even controlling for the threshold effect, we find that BMI does a poor job of classifying people as obese or non-obese. As we show below, the major reason is that BMI does not distinguish between fat and fat free mass.

How You Measure Fatness Matters: Racial Disparities in Obesity

In the next two sections we provide examples of the value of using more accurate measures of fatness in social science research. Although there are many measures of fatness, because of space constraints we focus on a few. Our comparisons confirm that the measure of fatness chosen by the researcher has enormous consequences for the conclusions drawn about who is obese and how fatness affects social science outcomes.

As Tables 1, 2 and 3 demonstrate, BMI is a noisy measure of fatness since it does not distinguish between fat and fat free mass. This would be less of an issue if this noise were random across the population. However as we show in this section, this is not the case across race and gender. Hence the prevalence of obesity will vary across race and gender depending on how one defines obesity. We start by showing in Table 4 how the amount of fat-free mass and total body fat vary with race and gender. On average, African American females have 3.56 more kg of fat-free mass (such as muscle, bone, and fluid) than white females and African American males have 1.33 more kg of fat-free mass than white males. Both of these differences are statistically significant. African American women also have on average 3.16 more kg of total body fat, but their additional fat-free mass almost perfectly offsets that, so African American females' percent body fat is only 0.79 percentage points greater than that of white females, a difference that is not statistically significant.

African American men not only have more fat-free mass on average, they also have on average 2.33 fewer kg of total body fat compared to white men; as a result, their average percent body fat is 2.85 percentage points lower than that of white men. This is a statistically significant difference.

These differences across race in fat-free mass and total body fat are critical, because BMI does not account in any way for differences in fat-free mass, nor are there different BMI cutoffs for overweight or obesity for African Americans. Since African Americans have on average higher Fat-free mass such as muscle and bone, BMI mistakenly classifies them as more likely to be overweight and obese. Table 5 shows that when one defines obesity using BMI, the obesity rate among African American women is 11.40 percentage points higher than that among white women, a statistically significant difference. Again when one defines obesity using BMI, the obesity rate among African American men is 0.56 percentage points higher than that of white men, a difference that is not statistically significant.

The black-white gap in obesity rates changes dramatically when one classifies people as obese using the more accurate measure of percent body fat. The second column of Table 5 shows that, while the black-white gap in obesity rates among women continues to be statistically significant, it falls by more than half when one uses the more accurate measure of PBF. Even more dramatic is the change among men. Whereas African American and white men have statistically indistinguishable obesity rates when one uses BMI, when one defines obesity using PBF the difference is statistically significant with white men having an obesity rate that is 16.26 percentage points higher than that of African American men. In summary, the use of a more accurate measure of fatness generates obesity rates that challenge the conventional wisdom about who is obese and why.

Fatness and Early Receipt of OA Benefits

Table 6 contains the results of our probit regressions, in which the dependent variable is an indicator variable for whether the respondent started receiving OA benefits at age 62. It lists results for six regressions, one for each of our measures of fatness or obesity: body mass index, weight in kg controlling for height in cm, total body fat and fat free mass (both in kg), percent body fat, an indicator for obesity as defined by BMI, and an indicator for obesity as defined by PBF. Table 6 indicates that the correlation between each measure of fatness is statistically significant and positive; i.e. fatness is consistently associated with a higher probability of taking OA benefits at age 62. Moreover, the

correlation is large. Column 1 indicates that a one-unit increase in BMI is associated with a 1.7 percentage point higher probability of taking OA benefits at age 62. A unit of BMI is hard to grasp, as it varies with one's height, so it may be clearer to point out that a one standard deviation increase in BMI (4.3 units) is associated with a 7.3 percentage point higher probability of taking OA benefits at age 62. When weight in kg is the measure of fatness used, the results are similar. An increase in weight of 14.3 kilograms (one standard deviation) is associated with a 7.2 percentage point increase in the probability of taking OA benefits at age 62.

We next present results for measures of fatness that are based on body composition. In column 3, the results indicate that an extra kilogram of body fat is associated with a 2.8 percentage point higher probability of taking OA benefits at age 62. An extra standard deviation of total body fat is associated with a 19.3 percentage point higher probability of early receipt of OA benefits. Fat free mass is negative in sign but not statistically significant. These findings confirm that body composition is important to consider. While the BMI and weight in kg results imply that all body mass increases the probability of early OA receipt, the model that uses TBF and FFM indicates that it is only body fat, not fat-free mass like muscle, that predicts early receipt of OA.

Another measure of body composition is percent body fat (PBF). We find that an extra percentage point of one's body mass that is fat is associated with a two percentage point higher probability of taking OA benefits at age 62. A PBF that is one standard deviation from the mean is associated with a 7.6 percentage point higher probability of taking early benefits.

Finally, we present results for obesity. First, we use the most common measure of obesity, one based on BMI. Column 5 shows that this measure of obesity is associated with a whopping 25.8 percentage point higher probability of taking OA benefits at age 62. To put this in context, 45.6 percent of our sample took OA benefits at age 62. When we use the measure of obesity based on PBF, obesity is associated with a 11.7 percentage point increase in the probability of that outcome. (The magnitude of PBF-based obesity is lower than that for BMI-based obesity because PBF-based obesity has a lower threshold; that is, more people are classified as obese when one uses PBF so not surprisingly the correlation of PBF-based obesity with outcomes is less extreme.)

All of these results control for the strictly exogenous factors race and year in which the respondent turned age 62. When we add additional regressors that may also be affected by fatness and obesity (in particular, education, age of wife, and marital status), the results are very similar with the exception that the coefficient on PBF-based obesity goes from being just barely significant to not statistically significant (see Table 7).

These correlations are potentially unsatisfying because they may not accurately reflect the causal impact of fatness or obesity on early receipt of OA benefits. For example, there may be unobserved health differences between the obese and non-obese that leads to early receipt. To test this, we instrument for the fatness or obesity of the respondent using the same fatness or obesity measure for the respondent's adult biological child. Table 8 presents the IV probit results. There are two important conclusions. First, the IV procedure has considerably raised the standard errors, so no coefficients are statistically significant. More importantly, the Wald tests of exogeneity (test statistics not shown) indicate that we cannot reject the hypothesis that fatness and obesity are exogenous. The results of this test are important because they indicate that the benefit of the IV procedure (greater efficiency) is not worth the cost (higher standard errors); in other words, it indicates that our probit results are to be preferred to our probit IV results.

Conclusions

In this paper we have used more objective measures of health to model early receipt of OA benefits in a manner that addresses each of the three objections outlined by Bound et al. (2006). We find that fatness and obesity are strong predictors of early receipt of OA benefits. This is true even though two aspects of our PSID data work against finding a significant correlation: first, we use a measure of fatness or obesity from 1986, years before the respondents actually begin to receive OA benefits; and second, our measure of early receipt has an unknown degree of measurement error due to inaccuracy in our algorithm for imputing whether the head began receiving OA benefits at age 62; and third, we are working with a relatively small sample (N=233) although the magnitudes of the correlation are so large that our limited power was not an issue. Our results indicate that we cannot reject that fatness or obesity is exogenous, implying that

we can cautiously interpret the probit marginal effects as estimates of the causal impact of fatness and obesity on early receipt of Social Security.

Our results also confirm that the measure of fatness one uses can affect research findings; for example, we find that only body fat, not fat-free mass is associated with early OA receipt, although if one used BMI or weight in kg one might be led to believe that more body mass is always associated with early OA receipt. Moreover, we find that obesity defined using BMI is associated with a higher probability of early OA receipt than is obesity defined using percent body fat; the reason is that fewer people are classified as obese by BMI and so it is a more extreme form of obesity, more strongly correlated with early OA receipt.

Our findings have important implications for those who believe that recent changes in OA benefit payout rules (that will eventually raise the normal retirement age to 67 and decrease the actuarial value of first taking OA benefits at age 62) will lead to a smaller percentage of future old-age cohorts taking OA benefits at age 62. While these policy changes are likely to have the expected marginal impact on behavior, that impact may be overwhelmed by the potential marginal impact of increased obesity in future old-age cohorts on early OA receipt. The magnitude of our results suggest that the upward trend in obesity in the U.S. could actually result in a higher percentage of future old-age cohorts taking OA benefits at age 62 despite the reduced protection such benefits will provide.

References

Averett, Susan and Sanders Korenman. 1999. "Black-White Differences in Social and Economic Consequences of Obesity." *International Journal of Obesity*, 23: 166-173.

Beertema, Wieske; van Hezewijk, Marjan; Kester, Arnold; Forget, Pierre-Philippe; van Kreel, Bernard. 2000. "Measurement of Total Body Water in Children Using Bioelectrical Impedance: A Comparison of Several Prediction Equations." *Journal of Pediatric Gastroenterology & Nutrition*. 31(4):428-432.

Bjorntorp, Per. 2002. "Definition and Classification of Obesity." In Christopher G. Fairburn and Kelly D. Brownell (eds.) *Eating Disorders and Obesity: A Comprehensive Handbook, 2nd Edition*. (New York: Guilford Press).

Bound, John, Charles Brown, and Nancy Mathiowetz. 2002. "Measurement Error in Survey Data." in *Handbook of Econometrics*, volume 5, ed. James Heckman and Ed Leamer. (New York: Springer-Verlag).

Bound, John, Todd Stinebrickner and Tim Waidmann. 2006. *Health, Economic Resources and the Work Decision of Older Men*, Paper presented at the 2006 NBER Economics of Aging Workshop, July, 2006.

Brown, Peter J. and Vicki K. Bentley-Condit. 1998. "Culture, Evolution, and Obesity." In *Handbook of Obesity*, edited by George A. Bray, Claude Bouchard, and W.P.T. James. (Marcel Dekker: New York).

Bray, George A., Claude Bouchard, and W.P.T. James. 1998. "Definitions and Proposed Current Classification of Obesity." *Handbook of Obesity*, (Marcel Dekker: New York).

Burkhauser, Richard V., and John Cawley. 2005. "Obesity, Disability, and Movement onto the Disability Insurance Rolls." Paper presented at the International Health Economics Association 5th World Congress.

Burkhauser, Richard V., Andrew J. Houtenville, and David C. Wittenburg. 2003. "A User's Guide to Current Statistics on the Employment of People with Disabilities." In David C. Stapleton and Richard V. Burkhauser (eds.) *The Decline in the Employment of People with Disabilities: A Policy Puzzle*. (Kalamazoo, MI: W.E. Upjohn Institute for Employment Research), pp. 23-86.

Burkhauser, Richard V., Andrew J. Houtenville, and Ludmila Rovba. 2006. "Accounting for the Declining Fortunes of Working-Age People with Disabilities," Cornell University Working Paper (March).

Cawley, John. 2000. "An Instrumental Variables Approach to Measuring the Effect of Body Weight on Employment Disability." *Health Services Research*, 35(5, Part II): 1159-1179

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

Cawley, John, and Sheldon Danziger. 2005. "Morbid Obesity and the Transition From Welfare to Work." *Journal of Policy Analysis and Management*, 24(4): 727-743.

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

Chou, Shin-Yi, Michael Grossman, and Henry Saffer. 2004. "An Economic Analysis of Adult Obesity: Results from the Behavioral Risk Factor Surveillance System." *Journal of Health Economics*, 23: 565-587.

Chumlea, W. Cameron, Shumei S. Guo, Robert J. Kuczmarski, Katherine M. Flegal, Clifford L. Johnson, Steven B. Heymsfield, Henry C. Lukaski, Karl Friedl, and Van S. Hubbard. 2002. "Body Composition Estimates from NHANES III Bioelectrical Impedance Data." *International Journal of Obesity*, 26: 1596-1609.

Culyer, Anthony J., and Joseph P. Newhouse. 2000. *Handbook of Health Economics, Volumes 1A and 1B.* (Elsevier: New York).

Currie, Janet and Brigitte C. Madrian. 1999. "Health, Health Insurance and the Labor Market," in Handbook of Labor Economics, ed. by Orley Ashenfelter and David Card, New York: Elsevier Science Publishing Co.: pp. 3309-3416.

Ferraro, Kenneth F., Ya-ping Su, Randall J. Gretebeck, David Black, and Stephen F. Badylak. 2002. "Body Mass Index and Disability in Adulthood: A 20-Year Panel Study," *American Journal of Public Health*, 92(5): 834-840.

Freedman, D.S. and G. Perry. 2000. "Body Composition and Health Status Among Children and Adolescents." *Preventive Medicine*, 31(2, Part II Supplement): S34-S53.

Gallagher, Dympna, Marjolein Visser, Dennis Sepulveda, et al. 1996. "How Useful is Body Mass Index for Comparison of Body Fatness Across Age, Sex, and Ethnic Groups?" *American Journal of Epidemiology*, 143(3): 228-39.

Garn, S.M., W.R. Leonard, V.M. Hawthorne. 1986. "Three Limitations of the Body Mass Index." *American Journal of Clinical Nutrition*, 44: 996-97.

Hedley, Allison A., Cynthia L.Ogden, Clifford L. Johnson, Margaret D. Carroll, Lester R. Curtin, and Katherine M. Flegal. 2004. "Overweight and Obesity Among U.S. Children, Adolescents, and Adults, 1999-2002." *JAMA*, 291(23): 2847-2850.

Hill, Martha S. (1992), *The Panel Study of Income Dynamics: A User's Guide*. Beverly Hills, CA: Sage Publications.

International Obesity Task Force. *EU Platform on Diet, Physical Activity, and Health, 2005*. Online at: http://www.iotf.org/media/euobesity3.pdf

Kragelund, Charlotte, and Tobjorn Omland. 2005. "A Farewell to Body-Mass Index?" *The Lancet*, 366:1589-1591.

Lakdawalla, Darius, and Tomas Philipson. 2002. "The Growth of Obesity and Technological Change: A Theoretical and Empirical Investigation." *NBER Working Paper* #8965.

Lakdawalla, Darius N., Jayanta Bhattacharya, and Dana P. Goldman. 2004. "Are the Young Becoming More Disabled?" *Health Affairs*, 23(1): 168-176.

Lee, Lung-fei and Jungsywan H. Sepanski. 1995. "Estimation of Linear and Nonlinear Errors-in-Variables Models Using Validation Data." *Journal of the American Statistical Association*, 90(429): 130-40.

Lumsdaine, Robin L. and Olivia S. Mitchell. 1999. "Retirement," in Handbook of Labor Economics, ed. by Orley Ashenfelter and David Card, New York: Elsevier Science Publishing Co.: pp. 3261-3308.

McCarthy, HD, TJ Cole, T Fry, SA Jebb, and AM Prentice. 2006. "Body Fat Reference Curves for Children." *International Journal of Obesity*, 30: 598-602.

Narbro, Kristina, Egon Jonsson, Barbro Larsson, Hans Waaler, Hans Wedel, and Lars Sjostrom. 1996. "Economic Consequences of Sick-Leave and Early Retirement in Obese Swedish Women." *International Journal of Obesity and Related Metabolic Disorders*, 20(10): 895-903.

National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institutes of Health. 2006. "Weight-Control Information Network." World wide web content. http://win.niddk.nih.gov/publications/understanding.htm#distribution Accessed May 19, 2006.

National Institutes of Health. 1996. "Bioelectrical Impedance Analysis in Body Composition Measurement: National Institutes of Health Technology Assessment Conference Statement." *American Journal of Clinical Nutrition*, 64(3S): 524S-532S.

National Institutes of Health. 1998. *Clinical Guidelines on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults.* NIH Publication 98-4083. (NIH: Washington, D.C.).

Nieto-Garcia, F. J., Bush, T. L. and Keyl, P. M. 1990. "Body Mass Definitions of Obesity: Sensitivity and Specificity Using Self-Reported Weight and Height." *Epidemiology*, 1(2), 146-152.

Ogden, Cynthia, Margaret D. Carroll, Lester R. Curtin, Margaret A. McDowell, Carolyn J. Tabak, and Katherine M. Flegal. 2006. "Prevalence of Overweight and Obesity in the United States, 1999-2004." *JAMA*, 295(13): 1549-1555.

Olshansky, S. Jay, Douglas J. Passaro, Ronald C. Hershow, et al. 2005. "A Potential Decline in Life Expectancy in the United States in the 21st Century." *New England Journal of Medicine*, 352(11): 1138-1145.

Pi-Sunyer, F. Xavier. 2002. "Medical Complications of Obesity in Adults." In Fairburn, Christopher G. and Kelly D. Brownell, *Eating Disorders and Obesity: A Comprehensive Handbook, 2nd Edition*. (Guilford Press: New York).

Plankey, MW, J Stevens, KM Flegal, and PF Rust. 1997. "Prediction Equations do not Eliminate Systematic Error in Self-Reported Body Mass Index." *Obesity Research*, 5: 308-314.

Prentice Andrew M. and Susan A. Jebb. 2001. "Beyond Body Mass Index." *Obesity Reviews*, 2(3): 141-147.

Rowland, M. L. 1989. "Reporting Bias in Height and Weight Data." *Statistical Bulletin*, 70(2), 2-11.

Smalley, K.J., A.N. Knerr, Z.V. Kendrick, J.A. Colliver, and O.E. Owen. 1990. "Reassessment of Body Mass Indices." *American Journal of Clinical Nutrition*, 52: 405-08.

Snijder, MG, M Visser, JM Dekker, et al. 2002. "The Prediction of Visceral Fat by Dual-Energy X-Ray Absorptiometry in the Elderly: A Comparison with Computed Tomography and Anthropometry." *International Journal of Obesity and Related Metabolic Disorders*, 26: 984-993.

Social Security Administration. 2006. "Social Security Beneficiary Statistics." Actuarial Publications. World wide web content. Accessed May 23, 2006. http://www.ssa.gov/OACT/STATS/OASDIbenies.html

Stapleton, David C. and Richard V. Burkhauser. (eds.) 2003. "The Decline in Employment of People with Disabilities, W.E. Upjohn Institute for Employment Research, Kalamazoo, MI

Sun, Shumei S., W. Cameron Chumlea, Steven B. Heymsfield, et al. 2003. "Development of Bioelectrical Impedance Analysis Prediction Equations for

Body Composition With the Use of a Multicomponent Model for Use in Epidemiologic Surveys." *American Journal of Clinical Nutrition*, 77: 331-40.

Trayhurn, Paul and John H. Beattie. 2001. "Physiological Role of Adipose tissue: White Adipose Tissue as an Endocrine and Secretory Organ." *Proceedings of the Nutrition Society*, 60:329-339.

U.S. Department of Health and Human Services. 1994. "Plan and Operation of the Third National Health and Nutrition Examination Survey, 1988-94." *Vital and Health Statistics*, Series 1: Programs and Collection Procedures, No. 32. (DHHS: Hyattsville, MD).

U.S. Department of Health and Human Services. 2001. *The Surgeon General's Call to Action to Prevent and Decrease Overweight and Obesity*. (Washington, DC: U. S. Government Printing Office).

Wada, Roy. 2005. "Obesity, Muscularity, and Body Composition: The Puzzle of Gender-Specific Penalty and Between-Ethnic Outcomes." Paper presented at 2005 Eastern Economic Association Meetings.

Wellens, R.I., A.F. Roche, H.J. Khamis, A.S. Jackson, M.L. Pollock, R.M. Siervogel. 1996. "Relationships Between the Body Mass Index and Body Composition." *Obesity Research*, 4(1): 35-44.

Yusuf, Salim, Steven Hawken, Stephanie Ounpuu, et al. 2005. "Obesity and the Risk of Myocardial Infarction in 27000 Participants from 52 Countries: A Case-Control Study." *The Lancet*, 366: 1640-1649.

Table 1
False Positives and False Negatives
When Use BMI to Define Obesity, by Gender

	% Positives False	% Negatives False
Females	0.17	62.85
Males	16.45	28.66

Notes:
1) Data: NHANES III
2) Uses NIH-recommended cutoffs of PBF for obesity: 25% for men, 30% for women.

Table 2
Prevalence of Obesity Measured by BMI and PBF, by Gender

	Percent Obese Defined Using BMI	Percent Obese Defined Using PBF
Females	23.3	70.1
Males	18.9	43.3

Note: Data: NHANES III

Table 3
False Positives and False Negatives
When Use BMI to Define Obesity, by Gender

	% Positives False	% Negatives False
Females	27.74	5.42
Males	44.79	9.31

Notes:
1) Data: NHANES III
2) Controls for the threshold effect by using cutoffs of PBF for obesity that result in approximately the same rates of obesity as does the BMI cutoff for obesity (BMI>=30), which are: 29% body fat for men and 41% for women.

Table 4
Testing for Differences in Mean PBF, TBF, FFM by Race and Gender

	FFM (kg)	TBF (kg)	PBF (%)
African American Females	47.59	27.95	35.27
White Females	44.03	24.79	34.48
Black-White Difference	**3.56*** **	**3.16*** **	**0.79**
African American Males	64.01	18.28	21.19
White Males	62.68	20.61	24.04
Black-White Difference	**1.33*** **	**-2.33*** **	**-2.85*** **

Note:
1) Data: NHANES III
2) Asterisks indicate level of statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Table 5
Prevalence of Obesity Measured by BMI and PBF, by Race and Gender

	Percent Obese Defined Using BMI	Percent Obese Defined Using PBF
African American Females	33.11	74.56
White Females	21.71	69.33
Black-White Difference	**11.40*** **	**5.23*** **
African American Males	19.39	29.00
White Males	18.83	45.26
Black-White Difference	**0.56**	**-16.26*** **

Notes:
1) Data: NHANES III
2) BMI based on measured weight and height. PBF calculated using TBF and FFM generated from BIA readings.
3) Asterisks indicate level of statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Table 6:
Probit Regressions
Early Receipt of OA Benefits on Measures of Fatness

	(1)	(2)	(3)	(4)	(5)	(6)
Body Mass Index	0.04279*					
	(1.72)					
	ME=.017					
Weight in kg		0.01369*				
		(1.81)				
		ME=.005				
Height in cm		-0.02206				
		(-1.46)				
		ME= -.009				
Total Body Fat (kg)			0.06995*			
			(1.78)			
			ME=.028			
Fat Free Mass (kg)			-0.04049			
			(-1.32)			
			ME= -.016			
Percent Body Fat				0.05016**		
				(1.98)		
				ME=.020		
Obese (defined by BMI)					0.66158**	
					(2.39)	
					ME=.258	
Obese (defined by PBF)						0.29701*
						(1.65)
						ME=.117
Constant	-65.35452	-64.31430	-80.46653*	-72.82526	-56.18868	-69.71423
	(-1.42)	(-1.41)	(-1.68)	(-1.59)	(-1.21)	(-1.53)
Observations	233	233	233	233	233	233

Notes:
1) t statistics in parentheses
2) Marginal effects listed below t statistics
3) Statistical significance indicated with asterisks: *** $p<0.01$, ** $p<0.05$, * $p<0.1$
4) Other regressors include: the year respondent turned age 62, indicator for African-American.

Table 7:
Probit Regressions
Early Receipt of OA Benefits on Measures of Fatness
With Broader Set of Regressors

	(1)	(2)	(3)	(4)	(5)	(6)
Body Mass Index	0.04979**					
	(1.99)					
	ME= .020					
Weight in kg		0.01603**				
		(2.08)				
		ME=.006				
Height in cm		-0.02783*				
		(-1.82)				
		ME= -.011				
Total Body Fat (kg)			0.08456**			
			(2.13)			
			ME=.033			
Fat Free Mass (kg)			-0.05028			
			(-1.62)			
			ME= -.020			
Percent Body Fat				0.06065**		
				(2.36)		
				ME=.024		
Obese (defined by BMI)					0.63773**	
					(2.29)	
					ME= .250	
Obese (defined by PBF)						0.30208
						(1.62)
						ME=.119
Constant	-68.87333	-67.90309	-87.82960*	-78.07218*	-58.88961	-72.74588
	(-1.49)	(-1.48)	(-1.82)	(-1.70)	(-1.27)	(-1.58)
Observations	231	231	231	231	231	231

Notes:

1) t statistics in parentheses
2) Marginal effects listed below t statistics
3) Statistical significance indicated with asterisks: *** $p<0.01$, ** $p<0.05$, * $p<0.1$
4) Other regressors include: the year respondent turned age 62, highest grade completed, age of wife when head turned age 62, and indicator variables for African-American and marital status.

Table 8:
IV Probit Regressions
Early Receipt of OA Benefits on Measures of Fatness

	(1)	(2)	(3)	(4)	(5)	(6)
Body Mass Index	0.01174					
	(0.13)					
	ME=.005					
Weight in kg		0.00524				
		(0.21)				
		ME=.002				
Height in cm		-0.01794				
		(-0.38)				
		ME= -.007				
Total Body Fat (kg)						
Fat Free Mass (kg)						
Percent Body Fat				0.03541		
				(0.39)		
				ME=.014		
Obese (defined by BMI)					-0.56396	
					(-0.24)	
					ME= -.209	
Obese (defined by PBF)						0.72244
						(0.45)
						ME=.279
Constant	-64.51373	-65.35432		-70.64022	-67.48748	-74.74769
	(-1.42)	(-1.44)		(-1.49)	(-1.49)	(-1.60)
Observations	233	233		233	233	233

Notes:
1) t statistics in parentheses

2) Marginal effects listed below t statistics
3) Statistical significance indicated with asterisks: *** p<0.01, ** p<0.05, * p<0.1
4) Other regressors include: the year respondent turned age 62, indicator for African-American.
5) Instrument is the same measure of fatness for the respondent's adult biological child, controlling for the adult child's age and gender.
6) The probit IV model in which TBF and FFM were the measures of fatness failed to converge, so column 3 is left blank.

Appendix Table 1:
PSID Summary Statistics

Variable	Observations	Mean	Std. Dev.	Min	Max
Body Mass Index	233	26.50	4.34	15.78	46.59
Weight in kg	233	83.81	14.26	50	154.55
Height in cm	233	177.70	7.14	154.94	200.66
Total Body Fat	233	21.12	6.91	3.50	56.81
Fat-Free Mass	233	62.42	8.24	43.81	101.43
Percent Body Fat	233	24.72	3.84	6.79	35.98
Obesity defined by PBF	233	.46	.50	0	1
Obese defined by BMI	233	.16	.37	0	1
Black	233	.16	.37	0	1
Year Turned 62	233	1992.22	3.95	1987	2002
Married	233	.94	.23	0	1
Years of Education	231	12.32	3.37	2	17
Age	233	55.61	14.14	0	76
Child BMI	233	24.31	3.67	16.82	43.26
Child Weight in kg	233	164.15	31.87	90	260
Child Height in cm	233	68.72	4.05	51	77
Child Total Body Fat (kg)	233	18.58	6.40	7.14	52.42
Child Fat-Free Mass (kg)	233	55.97	10.88	33.22	83.00
Child Percent Body Fat	233	24.75	5.98	12.10	46.65
Child Obese defined by BMI	233	.086	.28	0	1
Child Obese defined by PBF	233	.28	.45	0	1
Child Age	233	29	4.11	20	39
Child is Male	233	.71	.45	0	1