Does Higher Body Mass Index Lower Cognitive Functioning at Old-Age?

Bo MacInnis
Institute for Social Research
University of Michigan

September 20, 2009

Abstract

Cognitive performance declines with age, and impaired cognition is a major health problem and imposes substantial economic costs. Old-age obesity is prevalent and increasing. While the physical health effects of obesity have been well-documented and consistently found to be adverse, the consequences of obesity on cognitive health at old-age are less well understood with widely varied results and conflicting evidences, ranging from obesity being harmful to, protective of, to no relationship to cognition. The inconclusiveness of the literature may result from differences in study samples, cognitive domains, and methodology. We contribute to the literature by using a nationally representative sample—Health and Retirement Study—and measures of multiple cognitive domains, and importantly, an econometric analysis that models permanent individual heterogeneity and the potential endogoneity of obesity. We demonstrate the importance of these econometric modeling issues and how the differences in estimation methods yield to different (opposite direction) results that could characterize the existing literature.

JEL Classification: I12, I18, I21, I28.

Keywords: Obesity, overweight, body weight, old-age cognitive health, memory.

Correspondence: Bo MacInnis, University of Michigan, Population Studies Center, Institute for Social Research, 426 Thompson Street, Ann Arbor, MI 48106. Phone: 734 647 9274; E-mail: macinnis@umich.edu

Acknowledgement: The author is grateful to the insightful comments by Elena Gouskova, Jack Iwashyna, Kandice Kapinos, Ken Langa, Helen Levy, Bob Schoeni, and David Weir.

1. Introduction

Age-related cognitive impairment is a major health problem that exerts substantial burdens on individuals, their families, health care resources, and national safety net programs. Cognitive impairment is a disorder with marked deficits in many areas of thought and action, such as language, attention, reasoning, judgment, reading and writing, and particularly memory. The prevalence of cognitive impairment in the elderly population is high: in 2002, among the age 71+ population, 22.2%—5.4 million Americans—were cognitively impaired but not-demented, and additionally 13.9%—3.4 million Americans—were demented (Plassman et al. 2008, 2007). Because age is the strongest risk factor for cognitive impairment and dementia (Aronson et al. 1991), and the U.S. population is rapidly aging, the prevalence of cognitive impairment is projected to increase in this country (Brookmeyer, Gray, and Kawas 1998). The investigation for new prevention and treatment strategies to maintain higher cognitive functioning throughout life, particularly at old-age, is of substantial medical and economic importance.

Obesity among older Americans has grown over the past decades: the prevalence of obese adults rose from about 15% in 1991 to over 25% in 2001 among the age 60-69 population, and from about 12% in 1991 to 27% in 2001 among the age 70+ population. Arteburn et al. (2004) estimated that the prevalence of obesity in adults aged 60+ will increase to 37.4% in 2010, and the number of obese adults aged 60+ will increase 20.9 million. Obesity is strongly associated with chronic diseases of Type II diabetes, cardiovascular disease, certain types of kidney, colon and breast cancers, and musculoskeletal disorders. This association is particularly strong among older adults. For example, among the age 70+ population, in 2000, 27% had diabetes among the obese compared to 14% among the non-obese; 66% had hypertension among the obese and 53 among the non-obese, 77% had arthritis among the obese and 62 among the non-obese, and 34% had heart condition among the obese and 30% among the non-obese.

While the physical health effects of obesity are well-documented and consistently found to be negative and large, the cognitive health consequences of obesity are less well understood with widely varied results and conflicting evidences, ranging from obesity being harmful to, protective of, to no relationship to, cognition. First, on the body weight and cognitive functioning relationship, Bagger et al. (2004) find that elders with higher baseline Body Mass Index (BMI) performed better in follow-up cognitive tests than those with lower BMI, while Elias et al. (2003) report the evidence for the opposite findings at least among men: higher baseline BMI is associated with lower cognitive functioning at follow-ups. Second, on the body weight and cognitive decline relationship, Struman et al. (2008) find that elders with lower baseline BMI had less severe decline in their cognitive functioning over time than those with higher baseline BMI; however, Kanaya et al. (2009) report the opposite findings among men—higher baseline BMI leads to accelerated cognitive declines. Third, on the body weight and dementia relationship, several recent studies have led to a consistent finding that greater body weight is protective of cognitive health and reduces the risk of dementia (Fitzpatrick et al. 2009, West and Haan 2008, Dahl et al. 2008, Buchman et al. 2005).

The inconclusiveness of the obesity and cognition literature may result from differences in study samples, cognitive domains, and methodology. The relationship between obesity and old-age cognition may be heterogeneous between genders; for example, the adverse effect of higher BMI on cognition level and cognitive aging is only found among men not women (Elias et al. 2008, Kanaya et al. 2009). The obesity-cognition relationship may vary by life-course—mid-life obesity is a strong and consistent predictor for lower cognition at late-life, whereas the relationship between obesity and cognition at late-life is inconclusive. The old-age obesity and cognition relationship can differ across study populations—small specific communities or selective subpopulations such as healthy individuals without chronic diseases or without disability, rather than the nationally representative general population. The obesity-cognition link may be reflected differently in different measures of cognitive domains such as executive functioning, and various forms of memory and learning. We address these potential differences by using the Health and Retirement Study, a nationally representative longitudinal survey of middle-aged and elder Americans, and by using measures of multiple cognitive domains.

Additionally and importantly, none of the existing studies have addressed the issue of individual heterogeneity beyond the control for the baseline BMI and demographics. There are many unobserved individual factors that affect both body weight and cognitive functioning beyond the observable characteristics that are recorded on the survey and controlled for in the regression analysis. some of these unobserved individual factors are permanent and time-invariant, such as time preference and personal taste for health. Estimates of the obesity-cognition relationship are biased because of the presence of the permanent individual factors that influence both obesity and cognition outcomes, and the magnitude of the bias may be large in part depending on the strength of the correlation between these permanent individual factors and obesity. In our analysis, we purge out the individual heterogeneity using a first-difference method afforded by the longitudinal nature of the data, and by the virtue of first-difference estimation, we essentially ask: what is the effect of the change in body weight on the change in cognitive functioning?

While the first-difference estimation purges out the individual factors that are permanent, constant and time-invariant, there might be time-varying individual factors that are correlated with both obesity and cognitive functioning outcomes, leading to biased estimates of obesity-cognition relationship. The existing studies in the literature have treated obesity status as an exogenous variable, making it difficult to interpret the estimates of the obesity and cognition relationship beyond mere associations. The body weight or obesity status may be endogenous because these time-varying individual factors that affect both obesity and cognition are omitted from the analytical analysis. An example of these omitted variables that make obesity an endogenous variable is the utilization of prescription drugs that target obesity-related conditions such as diabetes, and such drugs make affect body weight as well as cognitive health. We use an instrumental variable method to account for the potential endogenous nature of body weight, and the instrument variable is twice-lagged body weight. The relevance criteria of valid instrument variable is that they are correlated with the endogenous variable, and this can be easily established because individual body weight tends to be serially correlated across adjacent time periods, and we find body weight at current period is strongly correlated with once or twice-lagged body weight but not three-lagged body weight. The exclusion criteria of instrument variables are that they are uncorrelated with the error or residual term in determining the outcome. While this assumption is un-testable empirically, we suppose that it is reasonable to assume that the two-lagged body weight is uncorrelated with the error term in determining the relationship between the cognitive outcome and body weight at current period conditional on the observable.

Using two econometric approaches of the instrument variable method to address the potential endogenous nature of body weight and the individual fixed effect method to address permanent individual heterogeneity, we have three findings that advance the literature. First, being heavy is not good to the brain—a greater BMI exerts an adverse effect on the cognitive functioning. Second, the effect of BMI on cognition varies by gender and race subgroups, and the use of instrument variable in addressing the endogenous nature of body weight is found to be valid among white men and women, and among the whites, the effect of BMI on cognition is strongest among white women. Third, the effect of BMI on cognition also varies by cognitive domains as well as gender and race difference, and BMI's negative effect is strongest on verbal memory—immediate, delayed, or combined word recall, and not on mental status and series 7 working memory. Importantly, our econometric modeling is verified by statistical specification tests, and our findings are insensitive to several robustness checks.

2. Methods

2.1 Empirical Specifications

Our empirical framework in examining the effect of body weight on cognitive functioning is to relate the cognitive health outcomes to body mass index. Let y_{it} and bmi_{it} represent the cognitive health measure and body mass index at time *t* for individual *i*, respectively. Suppose that cognitive health measure is determined by a relation in the form:

$$y_{it} = \alpha + \beta bmi_{it} + \phi x_{it} + \gamma z_i + \tau t_t + \omega_i + \varepsilon_{it} \tag{1}$$

Where z_i represents time-varying personal traits (gender, race, ethnicity, educational attainment, and veteran status), x_{it} represents a set of measured personal traits that varies over time (age, marital status, residential region, as well as chronic health condition indictors and economic resource variables), t_t represents a set of time-related controls (dummy variables for survey years and number of waves since enrollment to control for the "practice effect"), ω_i represents a permanent individual-specific effect, and ε_{it} random error term. Among the coefficients $\alpha, \beta, \phi, \gamma, \tau$ to be estimated, β is the coefficient of interest.

The consistent estimation of the effect of body mass index on cognitive health outcome depends on the modeling assumptions of the individual-specific effect, ω_i, and the variable of interest, bmi_{it}, and four modeling assumptions and the corresponding four specifications are considered and analyzed below. The first specification is to model the individual effect to be uncorrelated with all the covariates including the error term (2a) and the body weight to be exogenous (2b):

$$\omega_i \perp \left(bmi_{it}, x_{it}, z_{it}, t_t, \varepsilon_{it} \right) \tag{2a}$$

$$bmi_{it} \perp \varepsilon_{it} \tag{2b}$$

The specification (1) under assumption (2a)-(2b) (thereafter referred to as specification (2) for brevity) is considered cross-sectional individual random effect and exogenous body weight specification. Under this specification, the coefficient β is estimated using a random effect generalized least square (GLS) estimation, and interpreted as the association between body weight and cognitive outcome. A positive coefficient β indicates that higher body mass index is associated with better cognitive functioning level.

To relax the assumption (2a), we allow the individual effect, ω_i, to be arbitrarily correlated with any or all of the covariates including the error term, but maintain the exogeneity assumption on body weight:

$$\omega_i \not\perp \left(bmi_{it}, x_{it}, z_{it}, t_t, \varepsilon_{it}\right) \tag{3a}$$

$$bmi_{it} \perp \varepsilon_{it} \tag{3b}$$

The specification (1) under assumption (3a)-(3b) (thereafter referred to as specification (3) for brevity) is considered longitudinal individual fixed effect and exogenous body weight specification. Under this specification, the coefficient β is estimated using a fixed effect within (for first-difference) estimation, and interpreted as the association between the change in body weight and the change in cognitive outcome. A positive coefficient β indicates that an increase in body mass index is associated with an improvement on cognitive functioning level.

The next two specifications focus on the modeling assumption of body weight. To relax the assumption (2b), we allow body weight, bmi_{it}, to be endogenous, that is to be correlated with the error term; by maintaining the individual random-effect assumption (2a), we have the following specification:

$$\omega_i \perp \left(bmi_{it}, x_{it}, z_{it}, t_t, \varepsilon_{it}\right) \tag{4a}$$

$$bmi_{it} \not\perp \varepsilon_{it}, bmi_{it} = a_0 + a_1 bmi_{it-2} + a_3 x_{it} + a_4 z_i + a_t t_t + e_{it}, \varepsilon_{it} \perp e_{it} \tag{4b}$$

The specification (1) under assumption (4a)-(4b) (thereafter referred to as specification (4) for brevity) is considered cross-sectional individual random effect and endogenous body weight specification. Under this specification, we use the twice-lagged body mass index, bmi_{it-2}, as an instrument variable to bmi_{it}. By the definition of an instrument variable, we assume that the error terms, ε_{it} and e_{it}, are not correlated, that is, the instrument variable (bmi_{it-2}) affects the outcome of interest (y_{it}) only through its effect on the endogenous variable (bmi_{it}). This is not a testable assumption. We think it is a reasonable assumption that the body weight four-years prior does not affect the factors that determine the cognitive performance other than its relationship to current period body weight. The coefficient β is estimated using Generalized Two-Stage Least Square (G2LS) instrument variable (IV) estimation, and interpreted as the impact of body weight on the cognitive outcome. A positive (negative) coefficient β indicates that higher body mass index increases (decreases) cognitive functioning.

Lastly, we relax the assumption (4a) to allow individual effect to be arbitrarily correlated with all the covariates and the error term, leading to the most general specification:

$$\omega_i \not\perp \left(bmi_{it}, x_{it}, z_{it}, t_t, \varepsilon_{it} \right) \tag{5a}$$

$$bmi_{it} \not\perp \varepsilon_{it}, bmi_{it} = a_0 + a_1 bmi_{it-2} + a_3 x_{it} + a_4 z_i + a_t t_t + e_{it}, \varepsilon_{it} \perp e_{it} \tag{5b}$$

The specification (1) under assumption (5a)-(5b) (thereafter referred to as specification (5) for brevity) is considered longitudinal individual fixed effect and endogenous body weight specification. The coefficient β is estimated using fixed-effect (within or first-difference) IV estimation, and interpreted as the impact of body weight on the cognitive outcome. A positive (negative) coefficient β indicates that an increase in body mass index increases (decreases) cognitive functioning.

Among the four models, model (2) is the most restrictive while model (5) the least restrictive, and model (3) is more general than model (2), model (4) is more general than model (2), and model (5) is more general than model (3). Using Hausman test, we can compare each of the three pairs of models listed above, from which the choice of the best model can be made.

2.2 Variables

We use data from the Health and Retirement Study (HRS), a biennial, longitudinal, nationally representative population-based survey of U.S adults above age 50. Study details of the HRS are provided elsewhere (Juster and Suzman 1995; Heeringa and Connor 1995; Wallace, Herzog, et al 2005). Our study samples are individuals aged 65-90 who had performed some or all domains of the HRS cognitive tests. We exclude the over-90 population because of the possible selection issue involving survival. Our main results are for the 65-90 population, with those aged 90+ included in robustness checks.

The HRS provides a battery of tests to reflect a range of cognitive abilities, including memory, language, orientation, and attention. Four cognitive tests were given to all respondents aged 65+ at all waves (Herzog and Wallace 1997): mental status, serial 7's, immediate word recall, and delayed word recall. Mental status refers to the Telephone Interview of Cognitive Status, which comprises 10 items relating to knowledge, language and attention; specific tasks include: naming the current president and vice president, identifying the current date and day of the week, naming a word for a tool used to cut paper, naming a word of the name of a prickly desert plant, and counting backward from 20. Mental status scores range from 0 to 10. Serial 7's refers to a subtraction test in which respondents are asked to count backward by 7s from 100 five times. The Serial 7's score ranges from 0 to 5 for number of correct subtractions. In the immediate word recall test, respondents were asked to recall a list of 10 words immediately after it was read to them; in the delayed word recall test, respondents were asked to recall the same list of 10 words approximately five minutes after it was first read to them. For the word recall tests, respondents were advised to listen carefully and told that they would be asked to recall as many words as possible.[1] The word recall score, the total number of words correctly recalled in both tests, ranges from 0 to 20. The series 7's variable will hereafter be referred to as "working memory"; word recall and verbal memory will be used interchangeably in the remainder of the text. (TODO: insert a distribution graph of recall & cogall scores)

The HRS provides self-reported body weight and height at each survey wave, from which Body Mass Index (BMI), defined as body weight in kilograms divided by the square of body height in meters, is calculated. We use BMI as a continuous variable in most of our analysis, and to capture the nonlinear effect of body weight we also use weight status as a categorical variable based on BMI: underweight (BMI < 18.5), normal weight (18.5 ≤ BMI < 25), overweight (18.5 ≤ BMI < 25), and obese (BMI ≥ 30). (TODO: insert a distribution graph of BMI)

We account for a standard set of demographic variables in our analysis, including time-unvarying demographic characteristics such as gender, race (white, and non-white), ethnicity (Hispanic origin), educational attainment in years of completed schooling, and military service status; as well as time-varying variables such as age, residential region, and marital status (married, and non-married). We also control for an extended list of economic resources that include annual household incomes and total household assets, as well as indicators for the

[1] The questionnaire wording is follows: "I'll read a set of 10 words and ask you to recall as many as you can. We have purposely made the list long so that it will be difficult for anyone to recall all the words—most people recall just a few. Please listen carefully as I read the set of words because I cannot repeat them. When I finish, I will ask you to recall aloud as many of the words as you can, in any order. Is this clear?"

presence of common chronic conditions that include hypertension, diabetes, cancer, lung disease, heart disease, arthritis, stroke, and psychological problems.

In addition, we include a set of dummy variables to indicate each HRS survey wave for two purpose. One purpose is that we address the potential "practice effect" in cognitive testing (Rabbitt et al., 2004, Alley et al., 2007) because the HRS respondents participated in the cognitive tests in multiple survey waves. For example, the mental status questions and serial 7's test for working memory are virtually identical across all survey waves and all respondents and thus may be subject to the potential great practice effect. The other purpose is that although the HRS has high cross-wave consistency on survey instruments and questionnaires, it is possible that the interview process in general and cognitive tests in particular might have some wave-specific features that could possibly influence respondents' cognitive test scores. The inclusion of a set of dummy variables indicating each interview wave will also account for this potential wave-specific effect.

There are 55,776 person-wave observations aged 65+ that are nationally representative with valid sampling weights available from HRS 1992-2006 including the Assets and Health Dynamics among the Oldest Old 1993-1995; and among those, 54,345 observations were aged between 65 and 90, and 1,431 were aged 91+ (age 91-102). Among those aged 65-90, there are 618 observations with missing data on body mass index, leading to a maximum of 53,727 observations for analysis. The participation rate in the cognitive test is considerably high, but varies by cognitive domain since the HRS asked and obtained the consent of the respondent to take each domain of the cognitive test separately. There are two sources of non-participation in cognitive test: one is that not self-respondents whereby a proxy cognition survey was administered to a respondent in the same household, and the other is some self-respondents refused to take the cognitive test. For the cognitive measure of immediate verbal recall, there are 48,335 useable observations for the final analysis; for delayed verbal recall, 47,638 observations; for mental status, 48,290 observations; for series 7 of working memory, 39,903 observations. The sample size is slightly reduced when we include the data on income, wealth and chronic health conditions in the analysis.

2.3 Descriptive Summary

Table 1 presents the summary statistics of our study sample of individuals aged 65-90. About 89.2% are white, 42.4% are male, 28.2% are veterans, 4.7% are of Hispanic origin, and 55.7% are married. Our sample has an average age of 75.1, and completed an average of 11.8 years of completed schooling. The average height is 1.68 meters, the average body weight is nearly 74 kilograms, and the average body mass index is a little over 26 and considered overweight. The prevalence of chronic diseases is high: about 53.1% for hypertension, 16.0% diabetes, 15.5% cancer, 10.2% lung diseases, 30.3% heart diseases, 10.3% stroke, 11.1% psychological problems, and 56.6% arthritis. The average household income is 10.1 in log, and the average total household wealth is 11.8 in log among those with positive wealth accumulation.

Without adjusting age and other demographics, our study sample performed well on several measures of cognitive tests: they obtained an average score of 9.15 out of 10 in mental status test, and were able to recall an average of a little more than half of the 10 words in immediate word recall. Our sample did less well in delayed verbal recall than in immediate recall, and this is to be expected because it is more cognitive challenging in performing delayed recall than immediate recall, our sample was able to recall about 3.86 out of 10 in delayed recall. Among those who participated both immediate and delayed recall test, the average score is 8.46 out of 20. Even with the relatively high refusal rate, our sample scored an average of 3.03 out of 5 in series 7 working memory test; those who refused to take the series 7 subtraction test do less well on other domains of the cognitive test conditional on their willingness to participate. Among those who participated in the test of all cognitive domains, they achieved an average score of 21.50 out of 35.

3. Results

3.1 Main results

We investigate the BMI-cognition relationship using four econometric specifications in how we model the individual effects and BMI. Specifically, we estimate four specifications: BMI is modeled as exogenous with individual random and fixed effect, respectively; BMI is modeled as endogenous with individual random and fixed effect, respectively. Results are presented in Table 2 for the entire study sample. Shown in the columns (1)-(2), assumed to be exogenous, BMI is positively and significantly correlated with all measures of cognitive domain outcomes regardless the modeling assumption of the individual effect, and correlation is small in magnitude but larger in the individual fixed effect model than in the individual random effect model. With individual random effects, every additional unit is associated with about one-tenth more word in either immediate or delayed recall, 0.006 and 0.010 higher in series 7 and mental status test scores, respectively, and 0.028 greater in the all cognition test score. With individual fixed effects, every additional unit is associated with about two-tenths more word in either immediate or delayed recall, 0.024 and 0.031 higher in series 7 and mental status test scores, respectively, and 0.085 greater in the all cognition test score.

We proceed to model BMI as an endogenous variable, and use the two-stage least square estimation to implement the just-identified instrumental variable method; results of first- and second-stage estimation are presented in columns (3)-(4) in Table 2 for individual random and fixed effect, respectively. In the first stage estimating, twice-lagged BMI is negatively correlated with current BMI in the individual random effect model, and positively correlated in the individual fixed effect model; the correlation is highly significant in both models. In assessing the strength of the instrument variables, Stock and Yogo (2002) provides empirical tests for weak instruments in linear instrumental variable regression, and one major requirement for ruling out weak instruments is that the *F*-statistics in the first stage be greater than 10. We find that this requirement is met (*F*-statistics = 41.56).

Modeled as endogenous, BMI bears no relationship with any measures of cognitive domain outcomes when the individual effect is assumed to be random effect. When we model BMI as endogenous and individual effect as fixed effect, we find that BMI exerts a negative effect on all measures of cognition outcomes, and the effect is significant in immediate or delayed recall, and the total cognition test, and insignificant in mental status or series 7 test. Specifically, for every additional unit of BMI, there is a reduction of 0.317 word, and 0.308 word in delayed and immediate word recall, and a reduction of 0.555 word in both delayed and immediate word recall, as well as a decline of 0.778 point in the total cognition test.

To put these estimates in the context, we can evaluate whether the a adverse effects of BMI on cognition are large and clinically and economically important. Calculating at the mean of the study sample, a reduction of 0.317 word in delayed recall per unit increase in BMI amounts to about 8% reduction in delayed recall, and similarly, per unit increase in BMI brings an decrease of about 6% decline in immediate recall, and 4% reduction in total cognitive test score. On the relative scale, the negative effect of BMI on cognition is largest in magnitude in delayed recall, and this finding may be significant considering the prominent importance of delayed recall as a marker for cognitive functioning. (TODO insert references)

3.2 Specification checks

The estimation results of the four specifications in modeling BMI and individual effects represent a snapshot of the existing literature—the BMI and cognition relationship differs drastically from studies to studies, from no relationship, protective (positive relationship), to harmful (negative relationship), as described in the earlier section. A major contribution of this study is to use econometric analysis to assess which model captures the true relationship between BMI and cognition that is embodied in the data available. From the modeling perspective among the four models analyzed here, specification (5) requires the least assumption and is the most general model, while specification (1) requires the strongest assumption and is the most restricted model, and the other two models—specifications (3)-(4) fall in between. Since the true relationship between BMI and cognition is an empirical question, specification checks can be performed to evaluate which model represents the best fit for the data. We use the Hausman test to check the relative strength between the four specifications, and present the results in Table 3. We perform Hausman test for three pairs of specifications: individual random effect versus fixed effect with exogenous BMI, endogenous versus exogenous BMI with individual random effect, and endogenous versus exogenous BMI with individual fixed effect. The test statistics strongly suggest that (1) with exogenous assumption of BMI, individual random effect is rejected when it is compared to the fixed effect with; (2) with individual random effect, the exogenous BMI model is rejected when it is compared to the endogenous BMI model; and (3) with individual fixed effect, the exogenous BMI model is rejected when it is compared to the endogenous BMI model. We thus conclude that the specification (5)—endogenous BMI with individual fixed effect—is the best specification among the four analyzed and compared. In this preferred model (5), BMI exerts a negative effect on cognition, especially on verbal memory.

Next we check whether our results are sensitive to the inclusion of health conditions and economic resources. Table 4 presents the estimation results using verbal recall (immediate and delayed recall combined) as a measure of cognition, using the same set of specifications as in Table 2 plus the additional set of covariates such as indicators for eight common chronic conditions, and household income and wealth. With the inclusion of these health and economics variables, the estimates of the BMI-cognition relationship when BMI is assumed exogenous become larger in magnitude and continue to be highly significant regardless of the modeling assumption of individual effects, while the estimate of the cognition relationship when BMI is modeled as endogenous in individual random effect becomes larger in magnitude and becomes significant. It is important to note that the estimate of the BMI-cognition relationship in the preferred specification—endogenous BMI with individual fixed effect—remains largely unchanged. We prefer the specification (5) without the inclusion of these health and economics variables because these variables are also used in the first stage regression. Furthermore, we perform the same specification checks as in Table 3, and conclude with strong evidence that specification (5) is most preferred.

3.3 Subgroup analysis

Table 5 presents the two-stage least square estimation results of specification (5)—with endogenous BMI and individual fixed effect—separately for white and black, men and women subgroups. We present our main results and robustness checks without accounting for chronic health conditions, and household incomes and wealth, and when these variables are controlled, our findings carry over. The first stage results in Table 5 indicate that twice-lagged BMI is significantly and negatively correlated with current BMI in among all gender and race subgroups except black men among whom the correlation is negative but significant. We find that this requirement is met among white men (*F*-statistics = 12.04) and white women (*F*-statistics = 30.20), but not met among black men (*F*-statistics = 4.59) or black women (*F*-statistics = 2.45). Because of the numerous issues involved in weak instruments, we focus on our seconds-stage estimation results among white men and women.

The main results among our study sample is that an increase in BMI leads to a decline in cognitive performance, and this effect is evident among white women and in all cognitive domains except series 7 working memory test. Specially, one unit increase in BMI leads to a reduction of 0.538 word in delayed word recall, a decrease of 0.434 word in immediate word recall, and a decline of 0.892 word in immediate and delayed word recall. The impact of one unit increase in BMI also includes a reduction of 0.465 in the mental status test score. Taking mental status, series 7 and immediate and delayed word recall altogether, we find that the marginal effect of BMI is a reduction of 1.319 points. Among white men, we find that a higher BMI leads to a deterioration in all measures of cognitive functioning outcomes, however, the standard errors are large, and the estimates are imprecise and thus insignificant.

To put these estimates in the context, we can evaluate whether the a adverse effects of BMI on cognition are large and clinically and economically important. Calculating at the mean of the white women of the study sample, a reduction of 0.538 word in delayed recall per unit increase in BMI amounts to about 14% reduction in delayed recall, and similarly, per unit increase in BMI brings an decrease of about 9% decline in immediate recall, 5% decline in mental status, and 6% reduction in total cognitive test score.

3.4 Robustness checks for subgroup longitudinal individual fixed-effect IV regression

We assess the robustness of our main results for the subgroups from several perspectives including excessive weight loss and stroke incidence. The first robustness check concerns the issue of excessive weight loss. Involuntary weight loss of large magnitude during a short period has been recognized as a marker for frailty, and associated with disability (Fried et al. 2004), poor health (Kahng, Dunkle, and Jackson 2004), and mortality (Calle et al. 1999; Wedick et al. 2002; Corrada et al. 2006). The fragility literature suggests that excessive involuntary weight loss may reflect an underlying cause that deteriorates health including brain health and cognitive functioning. Considering excessive weight loss an omitted variable, the inclusion of elders with excessive weight loss can bias our results.

To evaluate whether our main results are sensitive to this potential omitted variable, we conduct the same analysis but exclude individuals who lost at least 5 kilograms in weight during the previous 2-year period. The results are presented in columns (1)-(2) in Table 6. we focus on white men and women because of the support of the first stage results that the instrument variable is valid among whites not among blacks. Using the delayed recall as an example, there is about 9% reduction in number of observations among white men and women. The first stage results remain largely unchanged: the twice-lagged BMI is negatively and significantly correlated with current BMI, and the *F*-statistics are strong among white men and women.

The second stage estimations indicate two main results. One result is that as in the previous analysis, we find that a higher BMI leads to a worse cognitive functioning among white women, and the magnitude of this adverse effect of BMI is similar whether we include those with excessive weight loss, though the estimates become slightly less precise. The other result is that the estimates among white men become much more precise and significant when those with excessive weight loss are excluded, though the magnitude of the effect becomes smaller. We find that BMI adversely affects immediate recall, delayed recall, and cognition with all domains among white men, and these effects are sizable and comparable to those among white women. Specifically, the marginal effect of BMI is –0.399 word on delayed recall, -0.476 word on immediate recall, and –1.082 points on all cognitive tests. The conclusion is that our main results are insensitive to the inclusion or exclusion of those with excessive weight loss.

The second robustness check is related to the potential confounding issue of stroke incidence. Given the well-established damaging effect of stroke on brain functioning, the onset of stroke may exert a permanent change in the brain structure and cognitive reserve, and the relationship between body weight and cognition among stroke population than that among non-stroke population. To check whether our main results are sensitive to the potential difference on the BMI-cognition relationship induced by the onset of stroke, we repeat the analysis excluding individuals who ever had stroke, and report the results in the columns (3)-(4) in Table 6. The sample size—number of observations, using the delayed recall as an example, has about 10% reduction among white men, and 9% reduction among white women. The first stage results remain largely unchanged: the twice-lagged BMI is negatively and significantly correlated with current BMI, and the *F*-statistic remains strong among white women, and becomes less strong among white men.

The second stage estimations indicate that our main results among white women in the general population carry over in stroke patients are excluded from the analysis, with the magnitude of the adverse effect slightly attenuated. For example, the marginal effect of BMI on delayed recall declines to –0.458 word in non-stroke white women from –0.538 in white women in the general population, to –0.347 in mental status from –0.465, and to –1.052 in all cognitive test from –1.319. The conclusion from this exercise is that our main results are insensitive to the inclusion or exclusion of those with stroke.

4. Discussion

(TODO summary of findings). (TODO limitations of the study), (TODO mechanisms and mediators); (TODO policy implications)

Table 1: Demographics, BMI, and cognitive performance of the study sample

	Mean	Standard Error
Demographics		
White (0-1)	0.892	0.001
Male (0-1)	0.424	0.002
Years of schooling	11.753	0.015
Veteran (0-1)	0.282	0.002
Hispanic (0-1)	0.047	0.001
Age (years)	75.092	0.029
Married (0-1)	0.557	0.002
Household income ($ log)	10.095	0.005
Household wealth ($ log)	11.076	0.014
Body weight & health		
Height (meter)	1.681	0.000
Weight (kilogram)	73.995	0.076
BMI (kg/m_2)	26.093	0.023
Hypertension (0-1)	0.531	0.002
Diabetes (0-1)	0.160	0.002
Cancer (0-1)	0.155	0.002
Lung disease (0-1)	0.102	0.001
Heart disease (0-1)	0.303	0.002
Stroke (0-1)	0.103	0.001
Psychological problem (0-1)	0.111	0.001
Arthritis (0-1)	0.566	0.002
Cognitive performance		
Mental status (0-10)	9.150	0.007
Immediate word recall (0-10)	5.025	0.009
Delayed word recall (0-10)	3.860	0.010
Word recall (immediate and delayed) (0-20)	8.460	0.019
Series 7 subtraction (0-5)	3.028	0.010
Total cognition test score (0-35)	21.497	0.027

Notes: Data source is HRS 1992-2006 including AHEAD 1993-1995.Presented are population averages of demographics, economic resources, body weight, chronic conditions, and four domains of cognitive test scores in HRS of the study sample using sampling weight to adjust for complex survey designs. The study samples consists of individuals who were aged between 65 and 90 at time of the interview, and self-respondents and performed the cognitive tests. See the main text for the detailed construction of the study samples and data description.

Table 2: Estimates of BMI-cognition relationship using multiple specifications

Specifications	(1)	(2)	(3)	(4)
Individual effect modeled as	Random effect	Fixed effect	Random effect	Fixed effect
BMI modeled as	Exogenous	Exogenous	Endogenous	Endogenous
First stage: Dep. variable = BMI				
			0.859 ***	-0.051 ***
Twice-lagged BMI			(0.003)	(0.008)
# of observations (# of groups)			29,748 (12,176)	2,355 (1,004)
Wald or *F*-statistics			69641	41.56
(*p*-value)			$p < 0.0000$	$p < 0.0000$
Second stage estimates				
Dep. variable = Delayed recall				
	0.009 ***	0.017 ***	-0.000	-0.317 *
BMI	(0.002)	(0.006)	(0.003)	(0.174)
Dep. variable = Immediate recall				
	0.010 ***	0.020 ***	0.001	-0.308 **
BMI	(0.002)	(0.005)	(0.003)	(0.144)
Dep. variable = Recall				
	0.019 ***	0.044 ***	-0.000	-0.555 **
BMI	(0.004)	(0.009)	(0.005)	(0.283)
Dep. variable = Series 7				
	0.006 ***	0.024 ***	0.000	-0.043
BMI	(0.002)	(0.006)	(0.003)	(0.166)
Dep. variable = Mental status				
	0.010 ***	0.031 ***	0.003	-0.116
BMI	(0.002)	(0.004)	(0.002)	(0.111)
Dep. variable = All cognitive tests				
	0.028 ***	0.085 ***	-0.001	-0.778 *
BMI	(0.006)	(0.014)	(0.008)	(0.411)

Notes: without health or economic conditions TODO HERE

Table 3: Specification checks of BMI-cognition relationship

Specifications	[1]	[2]	[3]	[4]
BMI modeled as	Exogenous	Exogenous	Endogenous	Endogenous
Individual effect modeled as	Random effect	Fixed effect	Random effect	Fixed effect
Dep. variable = Delayed recall				
	0.009 ***	0.017 ***	-0.000	-0.317 *
BMI	(0.002)	(0.006)	(0.003)	(0.174)
	[1] vs. [3]: test statistics = 443.89, $p < 0.0000$			
	[1] vs. [2]: test statistics = 477.52, $p < 0.0000$			
Hausman tests	[2] vs. [4]: test statistics = 117.15, $p < 0.0000$			
Dep. variable = Immediate recall				
	0.010 ***	0.020 ***	0.001	-0.308 **
BMI	(0.002)	(0.005)	(0.003)	(0.144)
	[1] vs. [3]: test statistics = 504.43, $p < 0.0000$			
	[1] vs. [2]: test statistics = 411.48, $p < 0.0000$			
Hausman tests	[2] vs. [4]: test statistics = 63.00, $p < 0.0000$			
Dep. variable = Recall				
	0.019 ***	0.044 ***	-0.000	-0.555 **
BMI	(0.004)	(0.009)	(0.005)	(0.283)
	[1] vs. [3]: test statistics = 236011.17, $p < 0.0000$			
	[1] vs. [2]: test statistics = 6961.67, $p < 0.0000$			
Hausman tests	[2] vs. [4]: test statistics = 147.47, $p < 0.0000$			
Dep. variable = Series 7				
	0.006 ***	0.024 ***	0.000	-0.043
BMI	(0.002)	(0.006)	(0.003)	(0.166)
	[1] vs. [3]: test statistics = 29.20, $p < 0.0226$			
	[1] vs. [2]: test statistics = 62.32, $p < 0.0000$			
Hausman tests	[2] vs. [4]: test statistics = 23.78, $p < 0.0137$			
Dep. variable = Mental status				
	0.010 ***	0.031 ***	0.003	-0.116
BMI	(0.002)	(0.004)	(0.002)	(0.111)
	[1] vs. [3]: test statistics = 219.59, $p < 0.0000$			
	[1] vs. [2]: test statistics = 324.26, $p < 0.0000$			
Hausman tests	[2] vs. [4]: test statistics = 124.65, $p < 0.0000$			
Dep. variable = All cognitive tests				
	0.028 ***	0.085 ***	-0.001	-0.778 *
BMI	(0.006)	(0.014)	(0.008)	(0.411)
	[1] vs. [3]: test statistics = 200.06, $p < 0.0000$			
	[1] vs. [2]: test statistics = 336.22, $p < 0.0000$			
Hausman tests	[2] vs. [4]: test statistics = 117.96, $p < 0.0000$			

Notes: without health or economic conditions TODO HERE

Table 4: Estimates of BMI-cognition (verbal recall) relationship accounting for health conditions and economic resources

Specifications	[1]	[2]	[3]	[4]
BMI modeled as	Exogenous	Exogenous	Endogenous	Endogenous
Individual effect modeled as	Random effect	Fixed effect	Random effect	Fixed effect
	0.038 ***	0.043 ***	0.013 **	-0.512 *
BMI	(0.011)	(0.009)	(0.006)	(0.272)
	-0.140 ***	0.025	-0.179 ***	-0.114
Age	(0.003)	(0.041)	(0.004)	(0.072)
	-0.101 **	-0.141 **	-0.207 **	0.049
Married (0-1)	(0.042)	(0.068)	(0.049)	(0.115)
	-1.200 ***	-	-1.300 ***	-
Male (0-1)	(0.004)		(0.066)	
	1.049 ***	-	1.046 ***	-
White (0-1)	(0.065)		(0.072)	
	0.249 ***	-	0.264 ***	-
Years of schooling	(0.007)		(0.008)	
	-0.054	-	-0.235 **	-
Hispanic (0-1)	(0.097)		(0.105)	
	0.295 ***	-	0.147 **	-
Veteran (0-1)	(0.065)		(0.071)	
Health & economic resources				
	-0.013	0.071	-0.038	0.145
Hypertension (0-1)	(0.038)	(0.067)	(0.044)	(0.101)
	-0.326 ***	-0.001	-0.354 ***	-0.117
Diabetes (0-1)	(0.051)	(0.095)	(0.057)	(0.155)
	0.165 ***	0.040	0.088	-0.221
Cancer (0-1)	(0.050)	(0.088)	(0.056)	(0.156)
	-0.008	0.030	0.063	0.196
Lung disease (0-1)	(0.061)	(0.107)	(0.068)	(0.157)
	-0.085 **	-0.100	-0.112 **	-0.125
Heart disease (0-1)	(0.041)	(0.073)	(0.047)	(0.106)
	-0.514 ***	-0.518 ***	-0.625 ***	-0.713 ***
Stroke (0-1)	(0.062)	(0.098)	(0.071)	(0.183)
	-0.391 ***	-0.219 **	-0.474 ***	-0.129
Psychological problem (0-1)	(0.060)	(0.104)	(0.066)	(0.164)
	0.523 ***	0.062	-0.054	0.113
Arthritis (0-1)	(0.037)	(0.055)	(0.044)	(0.107)
	0.243 ***	0.050 **	0.171 ***	0.044
Household income ($ log)	(0.021)	(0.025)	(0.024)	(0.040)
	0.073 ***	0.014	0.065 ***	0.017
Household wealth ($ log)	(0.006)	(0.009)	(0.008)	(0.012)
# of observations (# of groups)	45,427 (15,250)	45,427 (15,250)	29,544 (12,116)	29,544 (12,116)
R-squared (overall)	0.3237	0.0125	0.2667	0.1131

Notes:

Table 5: Gender- and race-specific BMI-cognition link with individual-specific fixed effects treating BMI as endogenous

	White males	Black males	White females	Black females
First stage: Dependent variable: BMI				
Twice-lagged BMI	-0.032 ** (0.014))	-0.061 (0.042)	-0.051 *** (0.012)	-0.087 *** (0.028)
Age	-0.154 *** (0.050)	-0.093 (0.160)	-0.168 *** (0.051)	-0.070 (0.167)
Married (0-1)	0.266 *** (0.097)	-0.170 (0.291)	0.239 *** (0.079)	-0.077 (0.311)
# of waves since enrollment	0.028 (0.530)	-0.035 (0.326)	0.274 *** (0.105)	0.253 (1.336)
# of observations (# of groups)	11,103 (4,568)	1,379 (619)	14,963 (6,013)	2,355 (1,004)
F-statistics (*p*-value)	12.04 (0.0000)	4.59 (0.0000)	30.20 (0.0000)	2.45 (0.0049)
Second stage				
Dep. variable = Delayed recall				
BMI	-0.510 (0.526)	0.312 (0.656)	-0.538 ** (0.273)	0.154 (0.257)
Dep. variable = Immediate recall				
BMI	-0.738 (0.492)	0.555 (0.697)	-0.434 ** (0.216)	0.114 (0.207)
Dep. variable = Recall				
BMI	-1.232 (0.948)	0.910 (1.152)	-0.892 ** (0.442)	0.332 (0.418)
Dep. variable = Series 7				
BMI	-0.451 (0.420)	0.870 * (0.521)	0.072 (0.252)	-0.144 (0.272)
Dep. variable = Mental status				
BMI	0.278 (0.293)	0.399 (0.736)	-0.465 ** (0.191)	0.185 (0.191)
Dep. variable = All cognitive tests				
BMI	-1.672 (1.313)	2.348 (1.452)	-1.319 ** (0.657)	0.345 (0.619)

Notes:

Table 8: BMI-cognition link with individual-specific fixed effects treating BMI as endogenous robustness checks: exclude those with excessive weight loss, or stroke among whites

	Exclude excessive weight loss		Exclude stroke	
	White males	White females	White males	White females
First stage: Dependent variable: BMI				
Twice-lagged BMI	-0.081 *** (0.016)	-0.053 *** (0.014)	-0.033 ** (0.014)	-0.059 *** (0.013)
Age	-0.120 *** (0.050)	-0.141 *** (0.050)	-0.177 *** (0.051)	-0.162 *** (0.052)
Married (0-1)	0.317 *** (0.100)	0.270 *** (0.079)	0.256 *** (0.101)	0.252 *** (0.082)
# of waves since enrollment	0.122 (0.499)	0.266 *** (0.103)	-0.017 (0.541)	0.260 ** (0.108)
# of observations (# of groups)	10,137 (4,440)	13,606 (5,831)	10,010 (4,145)	13,602 (5,521)
F-statistics (*p*-value)	11.17 (0.0000)	21.27 (0.0000)	9.06 (0.0000)	25.39 (0.0000)
Second stage				
Dep. variable = Delayed recall				
BMI	-0.399 * (0.244)	-0.453 (0.319)	-0.699 (0.570)	-0.458 * (0.242)
Dep. variable = Immediate recall				
BMI	-0.476 ** (0.211)	-0.508 * (0.265)	-0.766 (0.508)	-0.435 ** (0.196)
Dep. variable = Recall				
BMI	-0.930 ** (0.422)	-0.877 * (0.525)	-1.387 (0.986)	-0.842 ** (0.398)
Dep. variable = Series 7				
BMI	-0.271 (0.209)	-0.061 (0.327)	-0.326 (0.421)	0.193 (0.228)
Dep. variable = Mental status				
BMI	0.102 (0.141)	-0.377 * (0.206)	0.162 (0.271)	-0.347 ** (0.153)
Dep. variable = All cognitive tests				
BMI	-1.082 ** (0.536)	-1.609 * (0.887)	-1.730 (1.278)	-1.052 ** (0.537)

Notes: