2015

Marginal structure model to estimate the causal effect of cognitive decline on body mass index changes in elderly people

Can Meng

ABSTRACT

MARGINAL STRUCTURE MODEL TO ESTIMATE THE CAUSAL EFFECT OF COGNITIVE DECLINE ON BODY MASS INDEX CHANGES IN ELDERLY PEOPLE

Can Meng, M.S.
Division of Statistics
Northern Illinois University, 2015
Dr. Duchwan Ryu, Director

Marginal structure models (MSM) are causal models proposed to estimate the causal effect of time-varying treatments by controlling time-dependent confounders in observational studies. The objective of this thesis is to illustrate the fitting of an MSM for the causal effect of cognitive decline on BMI changes among old Mexican Americans participating in the Hispanic EPESE cohort study. Based on MSM analysis, we reported that those who suffered cognitive decline in later life will receive significant reduction in BMI by 0.1809 kg/m^2 on average (P-value=0.0305). We also conducted several sensitivity studies to explore different impacts on our MSM.

NORTHERN ILLINOIS UNIVERSITY
DEKALB, ILLINOIS

DECEMBER 2015

MARGINAL STRUCTURE MODEL TO ESTIMATE THE CAUSAL EFFECT OF COGNITIVE DECLINE ON BODY MASS INDEX CHANGES IN ELDERLY PEOPLE

BY

CAN MENG

A THESIS SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE

MASTER OF SCIENCE

DIVISION OF STATISTICS

Thesis Director:
Duchwan Ryu

TABLE OF CONTENTS

Page

LIST OF TABLES iv

LIST OF FIGURES v

CHAPTER1: INTRODUCTION 1

1.1 Background 1

1.2 Problems Description 1

1.3 Methods Consideration 3

1.4 Outline 4

CHAPTER 2: MARGINAL STRUCTURE MODEL 5

2.1 Introduction to Marginal Structure Model 5

2.2 Counterfactuals and MSM 6

2.3 Inverse Probability of Treatment Weight and MSM 7

2.4 Assumptions of MSM 10

2.5 Applications of IPTW and MSM 11

CHAPTER 3: ANALYSIS OF THE DATA 13

3.1 Description of the Dataset 13

3.2 Analysis of the Data 16

3.2.1 Crude Analysis 17

3.2.2 Models for IPTW 19

3.2.3 The Structure Model for MSM 23

3.3 The Structure Model Selection 24

CHAPTER 4: SENSITIVITY STUDIES 27

4.1 Positivity Checking and Sensitivity to Different Types of Variables 27

4.1.1 Positivity Assumption Checking 27

4.1.2 Sensitivity of IPTW Estimator to Different Types of Variables 28

4.2 Sensitivity Study to Unmeasured Confounders 29

4.3 Sensitivity Study to Weight Truncation 33

CHAPTER 5: DISCUSSION 35

REFERENCES 37

APPENDIX: SAS-CODE 42

LIST OF TABLES

Table Page

3.1. Descriptive statistics of variables at baseline (wave 1) 14

3.2. Selected notation used in data analysis..... 16

3.3. Unweighted correlation of BMI, cognitive decline, and CES-D scores 19

3.4. Comparison of three unweighted models 19

3.5. Weighted correlation of BMI, cognitive decline, and CES-D scores 22

3.6. Comparison weighted and unweighted CES-D scores by cognitive decline..... 22

3.7. Descriptive statistics of the stabilized weights $sw_{j,t}^{F}$, $sw_{j,t}^{E}$, and $sw_{j,t}^{C}$ 23

3.8. Covariance matrix selection for the MSM..... 25

3.9. Coefficients of final weighted structured model..... 26

4.1. Proportions of cognitive decline by categories of time-dependent confounder 28

4.2. Comparison of effect of the exposure under different categorization

of the confounder 29

4.3. Sensitivity of the MSM analysis to unmeasured confounding 33

4.4. Truncation of inverse probability weights..... 34

LIST OF FIGURES

Figure Page

1.1. The relationships between BMI, cognitive decline (Cog), and depressive symptoms (Dep) 3

3.1. Boxplots of the distribution of CES-D by cognitive decline 18

3.2. The histogram of the stabilized weights $sw_{j,t}^{F}$ 22

CHAPTER 1: INTRODUCTION

1.1 Background

Recent studies have documented the association between body weight and cognitive performance, but those reports were not consistent. Poor cognitive performance or dementia associated with both low body mass index (BMI) and obesity were found in previous studies (Cronk et al., 2010; Ng et al., 2008; Whitmer et al., 2005); while long-term abnormal body weights (both obesity and underweight) were found to have an association with lower cognitive scores in late midlife (Sabia et al., 2009). In the latest study, the decline in cognitive function can predict the loss of body weight in later life, which implied that preclinical dementia may result in BMI declines (Suemoto et al., 2015). This thesis aims to investigate the causal association between cognitive function decline or cognitive decline (measured by Mini-Mental State Examination [MMSE] and defined as MMSE score $\leq$ 23) (Tombaugh & McIntyre, 1992) and BMI in old Mexican Americans.

1.2 Problems Description

In observational study, potential confounders are not controlled due to non-randomization. To investigate the causal effect on an outcome of interest from certain treatments or risk factors, one should well adjust those potential confounders. The conventional methods to control confounders include stratification, multivariate linear models, and survival analysis. These methods may be sufficient when confounders are time-independent variables. However, in the

presence of time-dependent confounders, those techniques may lead to biased estimations in longitudinal studies with repeated measurements over time (Hernan et al, 2000, 2001; Robins et al., 2000).

The data obtained here was from a cohort study named Hispanic EPESE (Hispanic Established Populations for the Epidemiologic Studies of the Elderly, Waves I-IV, 1993-2001). In this cohort study, participants were 3050 noninstitutionalized Mexican Americans aged 65 or older. They were tracked and re-surveyed in order to record their current health conditions every 2 to 3 years. Since this is an observational study without randomization, the observed association between cognitive performance and BMI may be confounded by other factors, i.e., depressive symptoms. Recent studies demonstrated the different associations between depressive symptom and BMI: positive association (Dragan & Akhtar-Danesh, 2007; Johnston et al., 2004), or U-shaped one (De Wit et al., 2009); Kim et al. (2014) reported that depression can result in weight loss in middle-aged and elderly Asian population. The relationship between cognitive function and depression also remains controversial: depression predicting subsequent cognitive decline (Comijs et al., 2001; Chodosh et al., 2007; Geerlings et al., 2000; Kobler et al., 2010; Raji et al., 2007; Wilson et al., 2002; Wilson et al., 2004), or cognitive decline preceding depressive symptoms (Perrino, et al., 2008; Vinkers, et al., 2004). The relationship between cognitive function and depressive symptoms may indicate that they mutually influence each other over time, i.e., previous depressive symptoms affect the cognitive decline, whereas the cognitive decline will influence subsequent changes in depressive symptoms. In this case, depressive symptoms are not only a confounder but also an intermediate between cognitive decline and BMI. The relationships between BMI, cognitive decline, and depressive symptoms are shown in Figure

1.1. The conventional statistical techniques may lead to biased estimations of causal effect from cognitive decline and BMI.

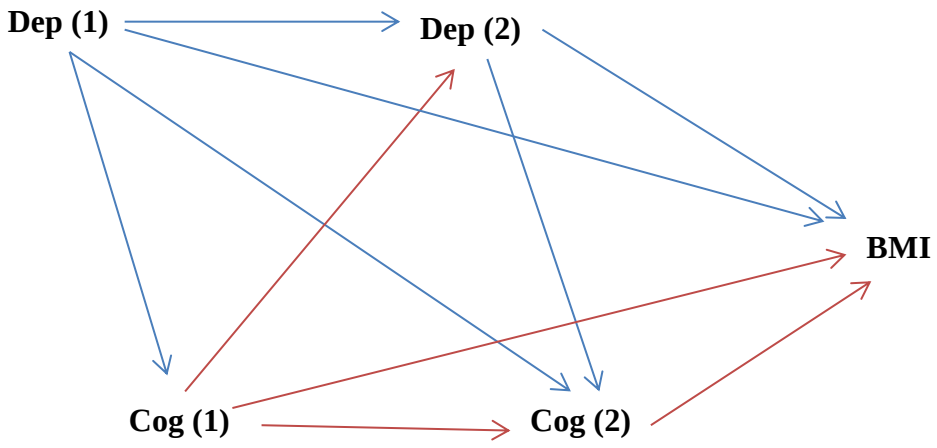

Figure 1.1. The relationships between BMI, cognitive decline (Cog), and depressive symptoms (Dep).

1.3 Methods Consideration

Marginal structural models (MSM), proposed by Robins (1998, 1999) and Robins et al. (2000), are a new class of causal models to estimate the causal effect of a time-varying exposure in the presence of time-dependent confounders. The parameters of MSMs can be consistently estimated by using the inverse probability of treatment weights (IPTW). Specifically, the IPTW obtained from the observed population can be used to create a pseudo-population in which the imbalance of treatment assignments inherent in the observed population will be adjusted, and the treatment will be unassociated with those confounders. Then one can fit a structural model to the pseudo-population in order to investigate the marginal association of the exposure and outcome

of interest. In this thesis, the outcome of interest is the BMI; the time-varying exposure is cognitive decline, and the potential time-dependent confounder is depressive symptoms.

1.4 Outline

In the second chapter, we are going to introduce MSM in detail, including relevant concepts, assumptions, and estimation of IPTW. The goal of this chapter is to understand what MSM is and how MSM and IPTW work. In the third chapter, we applied MSM to analyze the data of the Hispanic EPESE cohort study and compare the results of MSM with those of conventional regression models. In the fourth chapter, we conduct several sensitivity studies to explore: 1) the impacts of transforming the continuous confounder into a categorical variable on the IPTW estimation, 2) the sensitivity of our MSM to unmeasured confounders, and 3) the bias-variance tradeoff. And finally, we will discuss our findings and pros and cons of our model in Chapter 5.

CHAPTER 2: MARGINAL STRUCTURE MODEL

2.1 Introduction to Marginal Structure Model

A causal relationship of a treatment and its outcomes is equivocal when there are confounders involved. When some factors for the outcome are associated with the treatment, the treatment effect is confounded by those factors. Investigators employ randomization in experimental studies to eliminate or minimize such confounding impacts so that they can measure the true causal effect of the treatment; whereas in observational studies, randomization cannot be applied. Conventional methods to control confounders include some statistical techniques such as stratification and multivariate regression analysis. Those statistical techniques may be sufficient to adjust potential confounders in a point treatment study where the treatment is only administered once. However, such techniques may lead to biased estimation in longitudinal studies with repeated treatments over time if: 1) there exists a time-varying covariate that is a risk factor for, or an independent predictor of, the outcome of interest, and also predicts subsequent treatment, and 2) past treatment history predicts the subsequent covariate (Hernan et al., 2000, 2001; Robins et al., 2000). A covariate or risk factor satisfying condition 1) is called time-dependent confounders. Robins (1998, 1999) and Robins et al. (2000) proposed a marginal structural model (MSM) to estimate the causal effects of time-varying treatment to outcomes in the presence of time-dependent covariates that may be both confounders (condition 1) and intermediate variables (condition 2).

Technically, MSM employs a weighting method, known as the inverse probability of treatment weights (IPTW), to create a pseudo-population where the treatment is no longer confounded. In other words, the pseudo-population is used to simulate data of an experimental study. Then the causal effect of the treatment can be estimated by fitting a structure model (e.g., a regression model) to the data of the pseudo-population without caring for the time-dependent confounders. Since MSM uses a two-step modeling strategy to detach confounder control from the structure model, we need two kinds of models for MSM: 1) models to estimate IPTW and 2) a structure model to estimate the causal effect. In this thesis, cognitive decline is a risk factor rather than a treatment; thus, we call the models to estimate IPTW the exposure models.

2.2 Counterfactuals and MSM

If the treatment is not confounded by covariates (no matter if it is time dependent or not), we can infer the causal effect of a treatment on the outcome *Y*. Ideally, causal effects are comparisons of different potential outcomes under the treatment ($Y_{A=1}$) and no treatment ($Y_{A=0}$) for the same subjects. The causal effects within the population are the average of the individual causal effect. However, the outcomes are not observable simultaneously for the same subject since every subject is either treated or untreated. Then a concept of counterfactuals is involved. Literally, counterfactual means contrary to the fact. For example, if a person has received a treatment and died, we may ask whether this person's death occurred due to the potential side effect of the treatment. To answer this question, we want to know the probability that the person would have died had he been untreated. This probability, contrary to what we actually observed, is a counterfactual variable. Counterfactuals are sets of random variables representing all possible outcomes for each subject (Mortimer et al., 2005; Robins, 1999). The possible outcomes

comprise the actually observed ones as well as that would occur if, contrary to facts, the subject were exposed to each treatment history (Mortimer et al., 2005; Robins et al., 2000). In an observational study, each subject only receives one exposure at a given time, and the corresponding outcomes are typically observed. Mortimer et al. (2005) considered counterfactual outcomes that would have been observed had the subject received the treatment other than those actually received as a missing-data problem because not all possible treatment/outcome combinations are actually observed.

The concept of counterfactuals is fundamental for MSM. A marginal structural model is a structural model for marginal expectation/densities of a counterfactual outcome given a specified treatment plan (Chakraborty & Moodie, 2013). These models are said to be *marginal* because they model the marginal distribution of counterfactual variables (like $Y_{A=1}$ and $Y_{A=0}$); they are also called *structural* models because they model the counterfactual variables and describe the causal effects rather than associational ones (Robins et al., 2000). The average causal effect of a treatment on potential (i.e., counterfactual) outcomes can be estimated via comparison of the distribution of $Y_{A=1}$ and $Y_{A=0}$ on the aggregate or group defined by pretreatment covariates (Joffe et al., 2004). MSMs provide a way to estimate these aggregate causal effects.

2.3 Inverse Probability of Treatment Weight and MSM

The standard method to estimate the parameters of MSM is inverse probability of treatment weights (IPTW). A model (such as a logistic regression) for the probability of treatment conditional on treatment and confounder histories is used to calculate the weights that equal to the inverse probability of observed treatment given treatment and confounder histories (Cole & Hernan, 2008; Platt et al., 2013; Robins et al., 2000). The weights are then applied to the

observed data to create a pseudo-population (weighted data) where the treatment is unassociated with the confounder. To create such a pseudo-population, each participant is assigned a weight which is proportional to the participant's conditional probability of receiving his own treatment given his past treatment history and covariate history. The IPTW, denoted as w_j for subject *j*, has such a form:

$$w_j = \prod_{k=0}^{t} \frac{1}{Pr(A_{j,k} = a_{j,k} | \overline{A}_{j,k-1}, \overline{L}_{j,k})} \tag{2.1}$$

where $\overline{A}_{j,k-1}$ is a vector of the treatment history through time *k-1* for subject *j*, and $\overline{L}_{j,k}$ is the vector of covariate history through time *k*; $\overline{A}_{j,-1}$ is defined as $A_{j,0}$; $a_{j,k} = 0,1$. The denominator of w_j is informally the propensity score (Rosenbaum & Rubin, 1983) for the subject *j*. However, when w_j has extreme values (either large or small), the IPTW performs inefficiently in estimation. Robins et al. (2000) used the probability of the treatment conditional on the past treatment history and baseline covariates instead of 1 as the numerator in order to stabilize the distribution of w_j. The stabilized weight for the treatment, denoted as sw_j^T, is:

$$sw_j^T = \prod_{k=0}^{t} \frac{Pr(A_{j,k} = a_{j,k} | \overline{A}_{j,k-1},\ L_{j,0})}{Pr(A_{j,k} = a_{j,k} | \overline{A}_{j,k-1},\ \overline{L}_{j,k})} \tag{2.2}$$

where $L_{j,0}$, which represents the baseline covariates, is a subset of $\overline{L}_{j,k}$. It is noted that if there is no confounder, which means $A_{j,k}$ is unassociated with $\overline{L}_{j,k}$, then $sw_j = 1$. In sw_j^T, the difference between the numerator and the denominator only reflects the confounding caused by the time-dependent covariates $\overline{L}_{j,k}$ that the standard regression or stratification fails to adjust for. Even

though, compared to w_j, sw_j^T is less variable and skewed, it still tends to be skewed due to extreme values (Li et al., 2010; Robins et al., 2000).

In a longitudinal study, loss to follow-up or attrition is one important source leading to biased estimation. Robins et al. (2000) and Hernan et al. (2000) considered censoring (right censoring) as another time-dependent treatment and derive censoring weights to adjust for the loss to follow-up. Let $C_{j,k} = 0$ denote the subject *j* was not lost to follow-up at time *k* and $\overline{C}_{j,k-1} = 0$ denote the history of no censoring through time *k-1*. The censoring weight is:

$$sw_j^C = \prod_{k=0}^{t} \frac{Pr(C_{i,k} = 0 | \overline{C}_{i,k-1} = 0,\ \overline{A}_{i,k-1},\ L_{i,0})}{Pr(C_{i,k} = 0 | \overline{C}_{i,k-1} = 0,\ \overline{A}_{i,k-1},\ \overline{L}_{i,k})} \tag{2.3}$$

Then the final weights for subject *j* is:

$$sw_j^F = sw_j^T \times sw_j^C \tag{2.4}$$

where sw_j^F is used to create the pseudo-population where association of treatment and measured confounders will be removed by replicating sw_j^F copies of each subject (Hernan et al., 2000; Robins et al., 2000).

A specified structure model is to fit the pseudo-population in order to obtain the estimation of the parameters. To fit such a model in pseudo-population is equivalent to fitting a weighted model in the actual population because the true causal relationship will remain the same in pseudo-population as that in the actual population (Robins et al, 2000). If IPTW and MSM are correctly specified and assumptions of MSM are met (see Section 2.4), the parameters have causal interpretation as the effect of treatment on the outcome of interest.

2.4 Assumptions of MSM

As models for observational study, strong assumptions are required for MSM to obtain the causal interpretation of model parameters, which include consistency, no unmeasured confounding (exchangeability or sequential randomization), and positivity (experimental treatment assignment assumption) (Cole & Hernan, 2008; Moodie et al, 2013; Moore et al., 2012; Mortimer, et al., 2005; Robins et al., 2000). Consistency states that a person's counterfactual outcome under his actual course of exposure is exactly his observed outcome. In other words, if the $Y_{A=1}$ is the counterfactual outcome under the treatment $(A = 1)$ that would have been observed if a person had received the treatment, then $Y_{A=1} = Y(A = 1)$, where $Y(A = 1)$ is the actual outcome observed when the person actually received the treatment.

The fundamental assumption of MSM is no unmeasured confounders. Robins (1999) showed that under the assumption of no unmeasured confounders, IPTW estimation can consistently and unbiasedly estimate the causal effect of a time-dependent treatment. Although this assumption is not testable directly from an observational study (Hernan et al., 2001), a sensitivity analysis can be used to explore the effect of the potential unmeasured confounders on the causal parameter estimates (Brumback et al., 2004; Cole et al., 2005; Robins, 1999). The assumption of no unmeasured confounders may be equivalent to exchangeability or sequential randomization (Cole & Hernan, 2008; Robins, 1999).

Another important assumption of MSM is positivity, which states the probability of receiving each of the treatment levels for every combination of treatment and confounder histories is nonzero (Cole & Hernan, 2008). Positivity is also known as the experimental treatment assumption because under positivity assumption, individuals are assumed to be

assigned to each level of the treatment group in an observational study (Cole & Hernan, 2008). If some individuals cannot possibly expose to certain levels of treatment or confounders, then positivity is violated due to a structural zero probability of receiving the exposure. An example of the violation of positivity assumption is to estimate effect of exposure to certain intervention in an occupational epidemiology study. In such a study, being at work is usually considered as a proxy for health status, but also a potential confounder. However, there is a structural zero probability of exposure to the intervention among individuals off work. If the structural zero probability of receiving a time-dependent confounder occurs, then restriction or censoring may result in bias, regardless of using weighting or other methods (Hernan et al., 2004). For such studies, structural nested model is suggested to use (Robins et al., 2000). Even though there is neither structural zeros nor random zeros, Mortimer et al. (2005) showed that extreme low probabilities of receiving treatment may lead to bias in the IPTW estimator in practice.

2.5 Applications of IPTW and MSM

Since the MSM was introduced by Robins (1998) and Robins et al. (2000), it has been applied to address the issue of time-dependent confounders in different epidemiological and medical studies: antiretroviral therapy on HIV infection (Cole et al., 2005; Hernan et al., 2000), aspirin treatment on cardiovascular mortality (Cook et al., 2002), heparin use to predict the outcome of arteriovenous fistula surgery (Joffe et al., 2004), and iron supplement on anemia at delivery (Bodnar et al., 2004). The IPTW alone has been used to address the issue of missing data (Robins et al., 2000; Hernan et al., 2000) and adjust survival curves (Cole & Hernan, 2004).

In the next chapter, we will apply MSM to investigate the causal effect of cognitive decline on BMI. We use logistic regression model to estimate IPTW and then use a mixed model as the structural model to infer the causal effect.

CHAPTER 3: ANALYSIS OF THE DATA

3.1 Description of the Dataset

The Hispanic EPESE dataset comprises four waves including baseline (1993-1994) and three follow-ups (1995-1996, 1998-1999 and 2000-2001). The data collected on baseline is a representative sample of 3050 community-dwelling Mexican Americans aged 65 years and older from Arizona, California, Colorado, New Mexico, and Texas. The 3050 participants have a mean age of 73.6; 42.4% of them are males and 57.6% are females, which reflects the gender ratio of the old Mexican Americans in the community. Starting from the first follow-up, some of the original participants either died or were lost to follow-up, and about 55% of them were interviewed at the forth wave. Every subject contributes four rows to the data.

Variables used in analysis include several baseline covariates and potential time-varying covariates. Baseline covariates are age, gender, years of education, annual household income, marital status, smoking status, and previous drug use; time-varying covariates consist of cognitive decline and depressive symptoms. Cognitive decline is the time-varying exposure, while depressive symptoms are considered as a time-dependent confounder in our study. The years of education consists of three categories: no more than 8^{th} grade education, 9^{th}-11^{th} grade education, and 12^{th} grade or above. The income variables have five categories: less than $5,000, $5,000-$9,999, $10,000-$14,999, $15,000-$19,999, and $20,000 and over. Marital status is dichotomized into married or single (including separated, widowed, divorced, or unmarried). The

previous drug use is defined as the use of any medicine prescribed by a doctor during the past two weeks (1=yes, 0=no). Of those 3050 participants, only 10% completed 12 or more years of schooling, about 72% had annual household income less than $15,000, 44.5% were single, and 12% were current smokers (see Table 3.1).

Table 3.1. Descriptive statistics of variables at baseline (wave 1)

Variables	Mean ± SD or Proportion
Age (years)	73.1±6.8
Gender	
Male	42.4%
Female	57.6%
Years of Education	
≤8	83.8%
9-11	6.4%
≥12	9.8%
Annual Household Income	
<$5,000	14.2%
$5,000-$9,999	37.2%
$10,000-$14,999	21.4%
$15,000-$19,999	9.9%
≥$20,000	17.3%
Marital Status	
Married	55.5%
Single	44.5%
BMI	27.8±5.3
Cognitive Decline	
Yes	35.5%
No	58.0%
CES-D (Depressive Symptoms)	9.9±9.6
Smoking Status	
Current Smokers	12.5%
Current Non-smokers	87.2%
Pre-drug use	
Yes	61.9%
No	30.9%

The outcome of interest is BMI. Even though BMI is questioned if it is a good indicator of body fatness, BMI may be well reflective of the body fatness of the elderly because old people usually do not have much proportion of muscles.

Cognitive performance was evaluated by MMSE scores, which ranged from 0 to 30 points. A score of 23 or lower indicates impaired cognitive functioning (Kurlowicz & Wallance, 1999; Tombaugh & McIntyre, 1992). Thus, if a subject's MMSE score was lower than or equal to 23, it was coded as "1" to denote his cognitive decline; otherwise, the variable was coded as "0" to indicate normal cognitive functioning. At baseline (wave 1), the mean MMSE score of all subjects was 24.7 with a standard deviation (SD) of 4.6.

Depressive symptom is assessed by the Center for Epidemiologic Studies Depression Scale (CES-D Scale) (Radloff, 1977). Twenty CES-D items selected from a pool of items which are symptoms associated with depression are measured on a four-point frequency scale (e.g., 0=rarely/none; 1=some/little; 2=occasionally/moderate; 3=most/all of the time) (Markides, 2009). Therefore, the possible range of CES-D score is 0 to 60, and the larger scores indicate higher levels of depressive symptoms. The mean and SD of CES-D scores at baseline were 9.9 and 9.6 respectively.

Table 3.2 lists the selected notations and their explanations we are going to use in this chapter.

Table 3.2. Selected notation used in data analysis

$Y_{i,j,t}$	The BMI scores for subject *j* under exposure A_i at wave *t* (*t*=1, 2, 3, or 4)
$A_{i,t}$	Indicator variable of cognitive decline at wave *t*, *i*=1, 0 (1=decline, 0=not decline)
$\bar{A}_{i,t}$	The over-bar of the time-varying exposure denotes the history of it, i.e., $\bar{A}_{i,t}=\{A_{i,1},\ldots, A_{i,t}\}$
$L_{j,t}$	Time-dependent confounder CES-D scores for subject *j* at wave *t*
$\bar{L}_{j,t}$	The over-bar of the time-dependent confounder denotes the history of it for subject *j*, i.e., $\bar{L}_{j,t}=\{L_{j,1},\ldots, L_{j,t}\}$
$\bar{B}$	The vector of baseline covariates*, i.e., gender, annual income, years of education, smoking status, marital status, and pre-drug use.
$C_{j,t}$	Indicator variable of censoring status for subject *j* at wave *t* (1=censored, 0=not censored)
$T_{j,t}$	The current age** for subject *j* at wave *t*
μ	The overall intercept in mixed model
$S_{j,t}$	The subject effect at wave *t*, $S_{j,t}\sim N(0,\sigma_S^2)$
$\varepsilon_{i,j,t}$	The error term, $\varepsilon_{i,j,t}\sim N(0,\sigma^2)$

*All the covariates in $\bar{B}$ are categorical variables.
**For each wave, the current age was used for each observation. E.g., if a subject was 65 years old in wave1, and two years later he was interviewed, then the age for the person used at wave 2 was 67.

3.2 Analysis of the Data

The ideal method to investigate the causal effect of cognitive decline on BMI changes is to set up a randomized trial in which let half of the participants develop cognitive decline. But it is impossible because, first of all, it is not ethical; second, the procedure of developing cognitive decline still remains debated. Therefore, to draw a valid conclusion from observational studies is crucial for us to understand cognitive dysfunctions and dementia.

In this cohort study, cognitive decline is a time-varying variable because subjects' cognitive performances change along with the time. And the effects of cognitive decline on BMI changes may be confounded by a time-dependent confounder, such as depression. As mentioned above, those conventional approaches may have biased estimations of the causal effects,

particularly when the time-dependent confounder is also intermediate variables. In this case, we employed MSM to assess the causal effect of cognitive decline on BMI by controlling depressive symptoms through IPTW.

3.2.1 Crude Analysis

First, it is indispensable to do a crude analysis of the data to investigate the impact of depressive symptoms as both a confounder and an intermediate variable. To see whether the effect of cognitive performance is likely to be confounded by depressive symptoms, the distribution of depressive symptoms measured by CES-D scores at different levels of cognitive performances (decline or not) was examined. Figure 3.1 shows the boxplot of CES-D scores among subjects suffering and not suffering cognitive decline. From the boxplot, individuals who suffered from cognitive decline are more likely to have higher CES-D scores (which indicate more depressive symptoms). Table 3.3 shows that CES-D scores are positively associated with both BMI and cognitive decline. Therefore, three models are fitted to control: 1) no other covariates, 2) only the baseline covariates, and 3) both baseline covariates and time-dependent confounder.

$$Y_{i,j,t} = \mu + A_{i,t} + S_{j,t} + \varepsilon_{i,j,t} \tag{3.1}$$

$$Y_{i,j,t} = \mu + A_{i,t} + S_{j,t} + \bar{B} + \beta_1 \cdot T_{j,t} + \varepsilon_{i,j,t} \tag{3.2}$$

$$Y_{i,j,t} = \mu + A_{i,t} + S_{j,t} + \bar{B} + \beta_1 \cdot T_{j,t} + \beta_2 \cdot L_{j,t} + \varepsilon_{i,j,t} \tag{3.3}$$

The estimated effects of the three models were obtained by using Proc mixed procedure with repeated options and specifying an unstructured working covariance matrix in SAS software. Table 3.4 shows the estimated effects of the time-varying exposures $A_{i,t}$ in these three

unweighted models. The effect is -0.2961 for model (3.1), which means, without any adjustment, persons with cognitive decline will have BMI scores on average 0.2961 lower than those without it. When baseline covariates and time-dependent confounder were adjusted, it turns out that the effects of cognitive decline do not differ much. The effect is -0.1363 for model (3.2) controlling only baseline covariates and -0.1269 for model (3.3) controlling both baseline covariates and time-dependent confounder. Note that when baseline covariates and time-varying confounder were adjusted, the effect of cognitive decline on BMI is not significant. However, when comparing model (3.2) and model (3.3), the small difference of the effect for the time-varying exposure may not indicate that depressive symptoms have a small impact on cognitive performance because, as mentioned above, depressive symptoms are time-dependent confounders as well as intermediates.

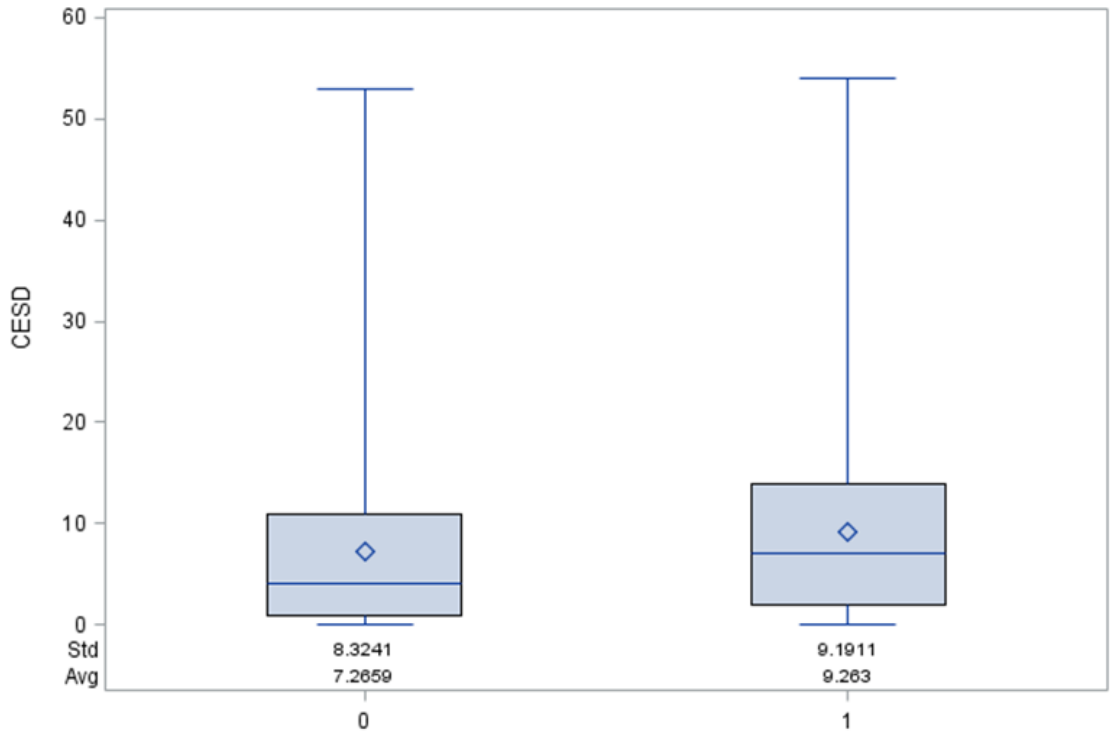

Figure 3.1. Boxplots of the distribution of CES-D by cognitive decline.

Table 3.3. Unweighted correlation of BMI, cognitive decline, and CES-D scores

	Pearson correlation coefficient	P-value
Cognitive decline vs. CES-D	0.1131	<0.0001
Cognitive decline vs. BMI	-0.0565	<0.0001
CES-D vs. BMI	0.0286	0.0121

Table 3.4. Comparison of three unweighted models

	Covariates controlled	Effects	SE	P-value
Model (3.1)	No other covariates	-0.2981	0.0822	0.0003
Model (3.2)	Only the baseline covariates	-0.1345	0.0852	0.1146
Model (3.3)	Both baseline covariates and time-dependent confounder	-0.1252	0.0856	0.1434

3.2.2 Models for IPTW

As introduced above, IPTW can create a pseudo-population by assigning each observation a weight in which the cognitive decline is no longer dependent on the depressive symptoms. Since each subject has multiple observations, each observation of the subject needs to be weighted separately. In this study, there are two kinds of exposure models: the exposure models for cognitive decline to balance the exposure assignment and the censoring models to adjust potential selection bias (Bodnar et al., 2004; Cole & Hernan, 2008).

Throughout the analysis, the stabilized weights are used. The stabilized weight of exposure for subject *j* at wave *t*, denoted as $sw_{j,t}^{E}$, is

$$sw_{j,t}^{E} = \prod_{k=1}^{t} \frac{Pr(A_{j,k} = a_{j,k} | A_{j,k-1}, \bar{B})}{Pr(A_{j,k} = a_{j,k} | A_{j,k-1}, \bar{L}_{j,k}, \bar{B})} \tag{3.4}$$

where, $a_{j,k}$=1 or 0. Similarly, the stabilized weight of censoring for subject *j* at wave *t*, denoted as $sw_{j,t}^{C}$, is

$$sw_{j,t}^{C} = \prod_{k=1}^{t} \frac{Pr(C_{j,k} = 0 | C_{j,k-1} = 0, A_{j,k-1}, \bar{B})}{Pr(C_{j,k} = 0 | C_{j,k-1} = 0, A_{j,k-1}, \overline{L}_{j,k}, \bar{B})} \quad (3.5)$$

Then the final weight for subject *j* at wave *t*, denoted as $sw_{j,t}^{F}$, is the product of $sw_{j,t}^{E}$ and $sw_{j,t}^{C}$:

$$sw_{j,t}^{F} = sw_{j,t}^{E} \times sw_{j,t}^{C} \quad (3.6)$$

Note that when *k*=1, $A_{j,0}$ and $C_{j,0}$ are defined as 0. For the sake of simplicity, the probabilities in both numerator and denominator are conditional on previous exposure $A_{j,t-1}$ rather than the whole history of exposure $\bar{A}_{j,t-1}$=$\{A_1,\ldots, A_{t-1}\}$; the history of time-dependent confounder $\overline{L}_{j,t}$ only consists of the values of time-dependent confounder at wave *t* and those at wave *t-1*, i.e., $\overline{L}_{j,t}$=$\{L_{j,t}, L_{j,t-1}\}$, assuming Markov property is held.

To obtain the estimated $sw_{j,t}^{F}$, four models are needed for each subject at each wave, in particular, two models for the two numerators and the other two for the two denominators. For instance, $Pr(A_{j,t} = a_{j,t} | A_{j,t-1}, \overline{L}_{j,t}, \bar{B})$ in denominator of $sw_{j,t}^{E}$ is the probability of exposure to $a_{i,t}$ (cognitive decline or not) conditioning on exposure history, history of depressive symptoms and baseline covariates for the subject *j* who are not censored at wave *t*. Therefore, an ordinal logistic regression is used to model this probability, and the predicted probability is obtained by using Proc logistic procedure in SAS. The logistic regression model is written as

$$logit(\Pr(A_t = a_t | A_{t-1}, \overline{L}_t, \bar{B})) = \beta_0 + \beta_1 \cdot A_{t-1} + \beta_2 \cdot T_t + \beta_3 \cdot L_t + \beta_4 \cdot L_{t-1} + \bar{B} \quad (3.7)$$

As the stabilized weight, probabilities in the numerator are only conditional on preview exposure A_{t-1} and baseline covariates $\bar{B}$. Therefore, a reduced logistic regression model of the

denominator excluding time-dependent confounder is used to estimate the probability in numerator:

$$logit(\Pr(A_t = a_t | A_{t-1}, \bar{B})) = \beta_0 + \beta_1 \cdot A_{t-1} + \beta_2 \cdot T_t + \bar{B} \quad (3.8)$$

The probabilities of not being censored in both numerator and denominator of $sw_{j,t}^{C}$ can be modeled in a similar way:

$$logit(\Pr(C_t = 0 | C_{t-1} = 0, A_{t-1}, \bar{B})) = \beta_0 + \beta_1 \cdot A_{t-1} + \beta_2 \cdot T_t + \bar{B} \quad (3.9)$$

$$logit(\Pr(C_t = 0 | C_{t-1} = 0, A_{t-1}, \bar{L}_t, \bar{B})) = \beta_0 + \beta_1 \cdot A_{t-1} + \beta_2 \cdot T_t + \beta_3 \cdot L_t + \beta_4 \cdot L_{t-1} + \bar{B} \quad (3.10)$$

In order to improve the prediction of models (3.7)—(3.10), five-knot restricted cubic splines (RCS) to age (T_t) and three-knot RCS to CES-D scores (L_t and L_{t-1}) are added into models (3.7)—(3.10) using SAS macro by Professor Frank Harrell (Harrell, 2009).

By weighting the original population with $sw_{j,t}^{F}$, the cognitive decline is no longer confounded by the depressive symptoms. Table 3.5 shows that even though the correlation coefficient of cognitive decline and CES-D is 0.0194, they are unassociated because the corresponding P-value for testing null hypothesis of correlation=0 is greater than 0.05. Table 3.6 shows that after weighting, the mean CES-D among those who had cognitive decline decreased to 8.6563 from 9.2630 (unweighted mean) while the mean CES-D for those without cognitive decline increased to 8.3320 from 7.2659 (unweighted mean). It indicates that the distribution of the exposure A_t has been balanced in the pseudo-population by IPTW. Figure 3.2 shows the distribution of the stabilized weights $sw_{j,t}^{F}$. The mean of $sw_{j,t}^{F}$ in this study is 0.9947 with

standard deviation of 0.2140 (see Table 3.7). Note that the closer the mean of weights is to 1, the less likely the models are misspecified (Cole & Hernan, 2008; Hernan & Robins, 2006).

Table 3.5. Weighted correlation of BMI, cognitive decline, and CES-D scores

	Pearson correlation coefficient	P-value
Cognitive decline vs. CES-D	0.0194	0.0869
Cognitive decline vs. BMI	-0.0613	<0.0001
CES-D vs. BMI	0.0365	0.0019

Table 3.6. Comparison weighted and unweighted CES-D scores by cognitive decline

	CES-D scores			
	Weighted		Unweighted	
Cognitive decline	Mean	Std. Dev.	Mean	Std. Dev.
Yes	8.6595	8.9696	9.2630	9.1911
No	8.3107	8.9258	7.2659	8.3240

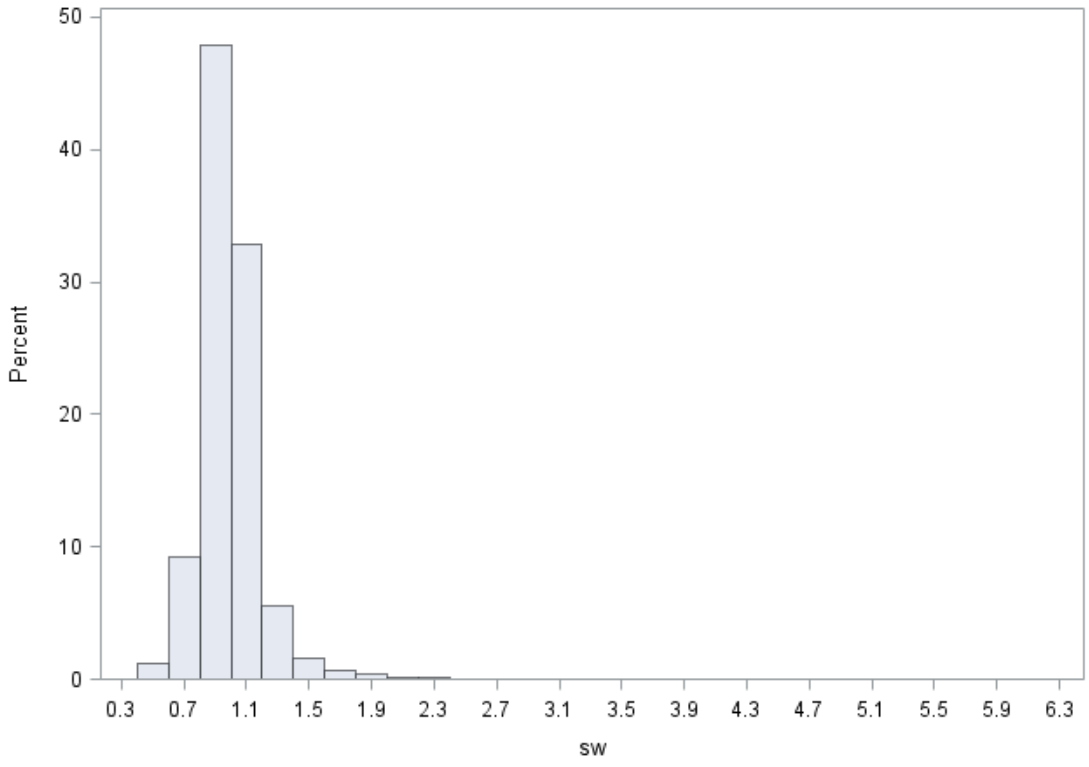

Figure 3.2. The histogram of the stabilized weights $sw_{j,t}^{F}$.

Table 3.7. Descriptive statistics of the stabilized weights $sw_{j,t}^{F}$, $sw_{j,t}^{E}$, and $sw_{j,t}^{C}$

		Mean	SD	Minimum	Maximum
$sw_{j,t}^{F}$	Pooled	0.9947	0.2140	0.2785	6.2352
	Wave 1	1.0000	0.0799	0.7240	1.2877
	Wave 2	0.9974	0.1767	0.5242	3.1817
	Wave 3	0.9978	0.2250	0.4116	3.1817
	Wave 4	0.9848	0.2864	0.2785	6.2352
$sw_{j,t}^{E}$	Pooled	1.0004	0.1741	0.3230	4.2575
	Wave 1	1.0000	0.0841	0.7240	1.2877
	Wave 2	1.0005	0.1616	0.5018	2.4155
	Wave 3	1.0004	0.1949	0.3443	2.4155
	Wave 4	1.0008	0.2235	0.3230	4.2575
$sw_{j,t}^{C}$	Pooled	0.9898	0.1286	0.4586	2.3573
	Wave 2	0.9952	0.0851	0.7306	1.5002
	Wave 3	0.9967	0.1213	0.7209	2.0394
	Wave 4	0.9775	0.1654	0.4586	2.3573

3.2.3 The Structure Model for MSM

After weighting the original population with $sw_{j,t}^{F}$, a structure model can be fitted in the pseudo-population, while in practice we can fit the weighted structure model into the original data. The structure model in our analysis is

$$Y_{i,jt} = \mu + A_{i,t} + S_{j,t} + \bar{B} + \beta_1 \cdot T_{j,t} + \beta_2 \cdot L_{j,1} + \varepsilon_{i,j,t} \quad (3.11)$$

where $L_{j,1}$ is the CES-D score at baseline (wave 1). To obtain the estimated effect of time-varying exposure $A_{i,t}$, model (3.11) is specified in SAS using Proc Mixed statement with repeated and weight options and assuming the covariance matrix is unstructured. The resulting effect of time-varying exposure in MSM is -0.1863 (95% confidence interval [-0.3549,-0.01757]). Compared to the result of the crude analysis (see Table 3.4), the estimation of MSM

shows a significant effect of cognitive decline on BMI changes (P-value=0.0305); the unweighted models indeed underestimate the effect of the exposure.

3.3 The Structure Model Selection

The weighted structural model (3.11) used above is a simple model including all baseline covariates and CES-D score at baseline, but no interactions between those covariates, and assuming the covariance matrix is unstructured. Now we want to select a best fit weighted structural model based on Akaike's information criterion (AIC) and Schwarz's Bayesian criterion (BIC). In the first place, we find the covariance structure that best fits; we then remove terms when that would improve AIC and BIC, but always keep the time-varying exposure A_t.

Thirteen different working covariance structures are selected, including: unstructured (UN), variance component (VC), compound symmetric (CS), heterogeneous compound symmetric (CSH), autoregressive (1) (AR [1]), heterogeneous autoregressive (1) (ARH [1]), autoregressive with random effect for subjects (AR RE), Toeplitz (TOEP), heterogeneous Toeplitz (TOEPH), Huynh-Feldt (HF), first-order autoregressive moving-average (ARMA [1,1]), first-order factor analytic (FA [1]) and first-order ante-dependence (ANTE [1]). The AIC, BIC, coefficients of time-varying exposure, and the corresponding standard errors from the MSM are shown in Table 3.8. The covariance structure of AR RE with less parameters, the lowest BIC and fairly low AIC is preferred. When the covariance structure of AR RE is selected, the next step is to decide which covariates are kept in the structure model. We use backward model selection by removing terms when that would improve AIC as well as BIC. Finally, the marital status and years of education are removed from the weighted structure model. The coefficients, standard error, and the corresponding P-values for the time-varying exposure and each covariate are

shown in Table 3.9. The causal effect of cognitive decline on BMI is -0.1809, which means, for persons with the same history of exposure, the person suffering cognitive decline will have, on average, 0.1809 loss in BMI.

Table 3.8. Covariance matrix selection for the MSM

Covariance structure	AIC	BIC	Effect	SE
Unstructured (UN)	39412	39471	-0.1863	0.0860
Variance component (VC)	44012	44018	-0.2070	0.1293
Compound symmetric (CS)	39926	39938	-0.2202	0.0925
Heterogeneous compound symmetric (CSH)	39897	39927	-0.2149	0.0919
Autoregressive (1) (AR (1))	39541	39553	-0.1844	0.0868
Heterogeneous autoregressive (1) (ARH (1))	39523	39553	-0.1815	0.0863
Autoregressive with random effect for subjects (AR RE)	39428	39444	-0.1813	0.0876
Toeplitz (TOEP)	39485	39509	-0.1868	0.0876
Heterogeneous Toeplitz (TOEPH)	39468	39510	-0.1843	0.0872
Huynh-Feldt (HF)	39917	39947	-0.2183	0.0925
First-order autoregressive moving-average (ARMA (1,1))	39484	39502	-0.1841	0.0876
First-order factor analytic (FA (1))	39824	39871	-0.1941	0.0901
First-order ante-dependence (ANTE (1))	39470	39511	-0.1810	0.0853

Table 3.9. Coefficients of final weighted structured model

	Coefficient ($\hat{\beta}_1$)	SE	P-value
Intercept	36.0927	0.8990	<0.0001
Cognitive decline			
Yes	-0.1809	0.0873	0.0383
No	0 (reference)	—	—
CES-D at baseline	-0.0027	0.0103	0.7960
Age	-0.1204	0.0112	<0.0001
Gender			
Male	-1.0970	0.1984	<0.0001
Female	0 (reference)	—	—
Annul income			
$0-$4,999	0.7405	0.3482	0.0335
$5,000-$9,999	0.6300	0.2810	0.0250
$10,000-$14,999	0.6091	0.3111	0.0503
$15,000-$19,999	0.2988	0.3812	0.4332
$20,000 and more	0 (reference)	—	—
Smoking status			
Current Non-smokers	1.3205	0.2838	<0.0001
Current Smokers	0 (reference)	—	—
Pre-drug use			
Yes	1.1680	0.2026	<0.0001
No	0 (reference)	—	—

CHAPTER 4: SENSITIVITY STUDY

4.1 Positivity Checking and Sensitivity to Different Types of Variables

4.1.1 Positivity Assumption Checking

Positivity means that the probability of receiving each level of exposure A_t for every combination of exposure and confounder histories A_t and L_t is non-zero. If an individual cannot possibly be exposed to at least one level of the confounders, then the probability of receiving the exposure is always zero, which is called a structural zero probability. Therefore, positivity is violated in this case due to the presence of the structural zero probability. We can roughly check the positivity assumption by looking at possible combinations of values of the exposure (cognitive decline) and the confounder (depressive symptoms). Table 4.1 shows the number and proportion of those who suffer cognitive decline under seven categories of CES-D scores. Recall that higher CES-D score indicates more depressive symptoms. In the table, the highest level of CES-D score only contains six observations, with four of them having cognitive decline. Such low frequency indicates that there must be zero probability of receiving some specific CES-D scores under the category of 51-60. Actually, in this dataset, there is no person having CES-D score above 56. However, we do not believe this is an issue of structural zero probability because the participants of this cohort study are sampled from the general population where less people have severe depression (i.e., high CES-D scores). Thus, it is more likely a random zeros issue, also known as practical violation of the positivity (Wang et al., 2006), due to the lack of depression patients in the general population. Cole and Hernan (2008) pointed out that zero

proportions may occur when there are highly stratified or continuous covariates. To fix this issue here, we treat the CES-D as a continuous variable and add three-knot RCS to it in the exposure models (3.7)—(3.10) so that we can borrow the information from persons with similar histories to smooth over the random zeros.

Table 4.1. Proportions of cognitive decline by categories of time-dependent confounder

CES-D scores	No. of cognitive decline	No. of person-waves	Proportion
0	619	1537	0.40
1-10	2116	4267	0.50
11-20	935	1676	0.56
21-30	348	586	0.59
31-40	120	190	0.63
41-50	34	56	0.61
51-60	4	6	0.67

4.1.2 Sensitivity of IPTW Estimator to Different Types of Variables

An alternative to avoid random zeros problem is to transforming the continuous confounder into a categorical variable so that those random zeros can be hidden in the categories. However, the number of categories may impact on the estimated effect of MSM through IPTW. To explore this impact, we transform the CES-D score into a categorical variable and split it into different numbers of categories. The range of the value of CES-D is 0 to 60. Note that CES-D score=0 means there is no any depressive symptom, and higher CES-D indicates more depressive symptoms. Thus, we always set CES-D=0 as an individual category. We split CES-D into: 1) two levels with CES-D=0 and CES-D≠0; 2) three levels with CES-D=0, CES-D=1~30 and CES-D=31~60; 3) five levels with CES-D=0, CES-D=1~15, CES-D=16~30, CES-D=31~45 and CES-D=46~60; and 4) seven levels as shown in Table 4.1. The estimated weights of IPTW and estimated effect of the time-varying exposure from MSM are reported in Table 4.2. From the results we can see that when the number of categories increases, the mean of estimated weights

is closer to 1, which indicates more precision in estimating the causal effect (Cole & Hernan, 2008) because the effects under small number of categories are close to those from the unweighted models. Compared with the confounder treated as a continuous variable, even though the mean of the weight based on the confounder treated as a categorical variable with seven levels is closer to 1 (0.9949 vs. 0.9947), the corresponding effect is underestimated because it is closer to the unweighted result (-0.1757 vs. -0.1809). A possible explanation may be that the categorical variable is less informative than the continuous one (Royston et al., 2006); in our case, when the CES-D is categorized, the severity of depression indicated by the number of depressive symptoms is neglected.

Table 4.2. Comparison of effect of the exposure under different categorization of the confounder

No. of categories	Estimated weights		Estimated effect from MSM	
	Mean (SD)	Min/Max	Effect	SE
2	0.9935 (0.1093)	0.3112/2.1361	-0.1540	0.0868
3	0.9936 (0.1532)	0.2426/3.1105	-0.1638	0.0868
5	0.9942 (0.1670)	0.2624/4.6682	-0.1647	0.0868
7	0.9949 (0.2022)	0.2655/4.8357	-0.1757	0.0870
Continuous	0.9947 (0.2140)	0.2785/6.3252	-0.1809	0.0873

4.2 Sensitivity Study to Unmeasured Confounders

MSM requires that there is no unmeasured confounder. Robins (1999) argued that all the related confounders should be adjusted in order to consistently estimate the causal effect of a time-varying exposure from the IPTW estimation. Hernan et al. (2001) also argued that it is almost impossible to directly test the assumption of no unmeasured confounders in an observational study. However, one can examine the effect caused by unmeasured confounders on the estimated parameter of the time-varying exposure by conducting a sensitivity study (Robins, 1999; Robins et al., 1999). Brumback et al. (2004) proposed a sensitivity analysis to unmeasured

confounders for MSM by using augmented IPTW estimation and specific sensitivity functions. They first specified several sensitivity functions with bias parameters γ which quantified the magnitude and direction of unmeasured confounders. And then they used the sensitivity function to compute bias-adjusted outcomes which were used, instead of the initial outcomes, to re-fit the MSM, and obtained the IPTW estimator of the coefficient for the time-varying exposure. Finally, they made comparisons between these augmented IPTW estimations and the original IPTW estimation.

The sensitivity function f defined by Brumback et al. (2004) represents the average difference in counterfactual outcomes Y_A between those exposed to $A_t = a$ and those exposed to $A_t = 1 - a$ $(a = 0,1)$, conditional on the covariates L_t and $\bar{B}$.

$$f = E[Y_a|A_t = a, L_t, \bar{B}] - E[Y_a|A_t = 1 - a, L_t, \bar{B}] \quad (4.1)$$

When $f \equiv 0$, there is no unmeasured confounders for the mean outcomes, given the measured covariates and confounders.

In this paper, we use two sensitivity functions, denoted as f_1 and f_2, proposed by Brumback et al. (2004) to represent two different types of unmeasured confounders. The forms of the sensitivity functions are

$$f_1 = \gamma \cdot (2A_t - 1) \quad (4.2)$$

$$f_2 = \gamma \cdot A_t \quad (4.3)$$

where γ is the bias parameter; $A_t = 1 - a$ $(a = 0,1)$ indicates the exposure or no exposure to the cognitive decline at wave *t*.

For the first scenario $f_1 = \gamma \cdot (2A_t - 1)$, when $a = 1$, by model (4.1):

$$E[Y_1|A_t = 1, L_t, \bar{B}] - E[Y_1|A_t = 0, L_t, \bar{B}] = \gamma \quad (4.4)$$

and when $a = 0$,

$$E[Y_0|A_t = 1, L_t, \bar{B}] - E[Y_0|A_t = 0, L_t, \bar{B}] = \gamma \quad (4.5)$$

We can obtain model (4.6) by subtracting model (4.5) from model (4.4):

$$E[Y_1 - Y_0|A_t = 1, L_t, \bar{B}\,] = E[Y_1 - Y_0|A_t = 0, L_t, \bar{B}\,] \quad (4.6)$$

Therefore, the causal effect is identical among the exposed and unexposed groups. When the bias parameter $\gamma > 0$, the mean counterfactual outcomes to the exposure and to no exposure tend to be higher in the exposed group than those in the unexposed group, whereas, when parameter $\gamma < 0$, the situation is reversed.

For the second scenario $f_2 = \gamma \cdot A_t$, when $a = 1$, by model (4.1):

$$E[Y_1|A_t = 1, L_t, \bar{B}] - E[Y_1|A_t = 0, L_t, \bar{B}] = \gamma \quad (4.7)$$

and when $a = 0$,

$$E[Y_0|A_t = 1, L_t, \bar{B}] - E[Y_0|A_t = 0, L_t, \bar{B}] = 0 \quad (4.8)$$

From model (4.7), the mean Y_1 is higher in the exposed than in the unexposed by γ unit, whereas, in model (4.8), the mean Y_0 is identical in both groups. Moreover, the causal effect in the exposed is different from that in the unexposed by γ, i.e.,

$$E[Y_1 - Y_0|A_t = 1, L_t, \bar{B}\,] = E[Y_1 - Y_0|A_t = 0, L_t, \bar{B}\,] + \gamma \quad (4.9)$$

This type of unmeasured confounders has no impact on the counterfactual outcomes to the unexposed.

The augmented IPTW estimation employs the bias-adjusted outcome $Y_{i,j,t}^{\gamma}$to replace the observed outcome $Y_{i,j,t}$ and re-fit the MSM using $Y_{i,j,t}^{\gamma}$ (Brumback et al., 2004; Cole et al., 2005). The bias-adjusted outcome is defined as

$$Y_{i,j,t}^{\gamma} = Y_{i,j,t} - \sum_{k=0}^{t-1} f \times \Pr[1 - A_k | \bar{A}_{k-1}, \bar{L}_k] \tag{4.10}$$

Note that, when the subject is exposed to cognitive decline, his $Y_{i,j,t}^{\gamma}$ equals to the value subtracts the product of sensitivity function f and the conditional probability of being unexposed from his observed BMI $Y_{i,j,t}$; however, when the subject is unexposed, his $Y_{i,j,t}^{\gamma}$ is the difference between observed one and the product of sensitivity function f and the conditional probability of being exposed.

Table 4.3 shows the results of the sensitivity study. The first row consists of the results of the original MSM assuming no unmeasured confounders for the sake of comparison. The columns in the table include augmented IPTW estimates of the coefficients for the time-varying exposure, the standard error, and nominal effect directions (-,+,0) which are based on a two-sided hypothesis test assuming augmented IPTW estimation has no difference with the original IPTW estimation. For both functions f_1 and f_2, we examine how much the magnitudes of γ can switch the direction of exposure effect. From the Table 4.3, f_1 can switch the direction when γ is very small (-0.5, 0.5), which means our MSM model is very sensitive to this type of unmeasured confounders; on the contrary, the MSM is somewhat insensitive to f_2, especially on the negative

direction, because when γ equals to 1, -1, and -2, the augmented IPTW estimates of exposure effect show no significant difference from -0.1809.

Table 4.3. Sensitivity of the MSM analysis to unmeasured confounding

Sensitivity function	Bias parameter γ	Effect estimate	SE	Nominal effect direction
—	0	-0.1809	0.0873	0
f_1	-2	-0.9283	0.0904	-
	-1	-0.5526	0.0882	-
	-0.5	-0.3636	0.0877	0
	0.5	0.01638	0.0881	0
	1	0.2060	0.0889	+
	2	0.5819	0.0917	+
f_2	-5	-0.8608	0.0952	-
	-2	-0.4410	0.0890	0
	-1	-0.3060	0.0880	0
	1	-0.0426	0.0879	0
	2	0.0848	0.0888	+
	5	0.4648	0.0947	+

4.3 Sensitivity Study to Weight Truncation

The denominator of IPTW is informally the propensity score of receiving the time-varying exposure, given the confounders and covariates history (Robins et al., 2000). If the frequency in one combination of certain levels of the exposure and confounders is very low, the corresponding probability of receiving the exposure, which is the denominator of the weights, should be small. In other words, the corresponding weight is fairly large. For example, there are less people with high CES-D scores in our dataset, which means the probability of no cognitive decline for a person with a high CES-D score (e.g., above 50) should be very small, and the weight for the person should be very large. This is the reason why the stabilized weights with smaller variance are preferred. Even though the best behaved IPTW are considered to have a

mean close to 1 and a small range, Cole and Hernan (2008) argued that such weights may fail to control for time-dependent confounders. This leads to a discussion of bias-variance tradeoff.

In this thesis, we explore this bias-variance tradeoff by progressively truncating the weights as Cole and Hernan (2008) proposed. We progressively truncate the weights by resetting the value of weights lower than the i^{th} percentile to the value of the i^{th} percentile and resetting the value of weights greater than the $(100\text{-}i)^{th}$ percentile to the value of the $(100\text{-}i)^{th}$ percentile. The first row of Table 4.4 shows the original MSM without truncation while the last row shows the truncated IPTW from the median. Note that the median of the stabilized weights $sw_{j,t}^{F}$ in this study is not 1 due to the skewed distribution of $sw_{j,t}^{F}$. From the Table 4.4 we can see the means of weights are gradually apart from 1 when the weights are progressively truncated, whereas the standard deviations are approaching to 0. When the weights are truncated no more than 5%, the corresponding estimations of the effects are consistent (-0.1812, -0.1811, and -0.1800), but the standard deviations of the truncated mean decrease up to 25% ([0.1620-0.2140]/0.2140). Therefore, the weights truncated up to 5% may be acceptable to optimize the bias-variance tradeoff in this study.

Table 4.4. Truncation of inverse probability weights

Truncation percentile	Estimated weights		Coefficient of time-varying exposure	
	Mean (SD)	Min/Max	Estimate	SE
0, 100	0.9947 (0.2140)	0.2785/6.2352	-0.1809	0.0873
0.5, 99.5	0.9920 (0.1882)	0.5287/1.9800	-0.1812	0.0873
1, 99	0.9899 (0.1768)	0.5736/1.6782	-0.1811	0.0874
2.5, 97.5	0.9875 (0.1620)	0.6479/1.4596	-0.1800	0.0874
5, 95	0.9839 (0.1417)	0.7166/1.2901	-0.1763	0.0875
10, 90	0.9818 (0.1175)	0.7931/1.1809	-0.1714	0.0875
25, 75	0.9803 (0.0663)	0.8995/1.0644	-0.1668	0.0874
50, 50	0.9786 (0)	0.9786/0.9786	-0.1485	0.0865

CHAPTER 5: DISCUSSION

The analysis showed that the cognitive decline among old Mexican Americans will lead to loss in BMI by 0.1809 kg/m^2 ($\beta = -0.1809$). Even though this effect of cognitive dysfunction seems too tiny to be useful in practice, it significantly differs from those estimated by conventional models (-0.1345 and -0.1252) because MSM efficiently controls the time-dependent confounder depressive symptoms through the IPTW estimator which constructs a pseudo-population to simulate the experimental data. Our results support previous researches that cognitive decline can predict BMI loss (Suemoto et al., 2015). Suemoto et al. (2015) reported that low memory scores predicted an annual BMI change of -0.158 kg/m^2, which is similar to our result: -0.1809 kg/m^2. Moreover, our findings confirm that there exists a causal association from cognitive decline to BMI loss in later life. Unfortunately, we didn't find any results from experimental studies that evaluate the effect of cognitive dysfunction on BMI changes, so we do not know how far our estimated effect is from the "truth." Most subjects involved in this cohort study were elderly Mexican Americans who were less educated and had low incomes, so it is probably inappropriate to generalize our finding to the general population.

MSM is based on several strong assumptions that we cannot directly test. We conducted a sensitivity study to explore the potential impact of unmeasured confounders on our MSM by manipulating the size and direction of this unmeasured impact. We found that our MSM is somewhat sensitive to the unmeasured confounders, which may imply that there exist unmeasured confounders that we did not include in our model, because if all relevant

confounders are controlled, the model should be fairly insensitive to those unmeasured. To further investigate the uncertainty due to confounding by unmeasured factors, one may want to conduct a Bayesian analysis, which is considered to be complementary to the sensitivity study (Brumback et al., 2004).

Besides unmeasured confounders, another limitation may be misspecification of the exposure model. Misspecification of the exposure can result in substantial bias in the structural model due to the use of those weights. We did not include any interaction term in models for IPTW in order to lower the risk of misspecification. Although the mean of the IPTW (0.9947) is very close to 1, which, empirically, is considered a necessary condition for correct model specification (Hernan & Robins, 2006), misspecification may still exist due to irrelevant covariates added into the models. One may want to employ the covariate balancing propensity score method proposed by Imai and Ratkovic (2014) that make the treatment/exposure model robust to misspecification.

REFERENCES

Bodnar, L.M., Davidian, M., Maria Siega-Riz, A., & Tsiatis, A.A. (2004). Marginal structural models for analyzing causal effects of time-dependent treatments: an application in perinatal epidemiology. *Am. J. Epidemiol.*, 159, 926-934.

Brumback, B.A., Hernan, M.A., Haneuse, S.J.P.A., & Robins, J.M. (2004). Sensitivity analyses for unmeasured confounding assuming a marginal structural model for repeated measures. *Statist. Med.*, 23, 749-767.

Chakraborty, B., & Moodie, E.E.M. (2013). *Statistical methods for dynamic treatment regimes*. New York: springer, p11.

Chodosh, J., Kado, D., Seeman, T.E., & Karlamangla, A.S. (2007). Depressive symptoms as a predictor of cognitive decline: MacArthur studies of successful aging. *Am. J. Geriatr. Psychiatry*, 15(5), 406-415.

Cole., S.R., & Hernan, M.A. (2004). Adjusted survival curves with inverse probability weights. *Computer Methods and Programs in Biomedicine*, 75, 45-49.

Cole, S.R., & Hernan, M.A. (2008). Constructing inverse probability weights for marginal structural models. *Am. J. Epidemiol*., 168, 656-664.

Cole, S.R., Hernan, M.A., Margolick, J.B., Cohen, M.H., & Robins, J.M. (2005). Marginal structural models for estimating the effect of highly active antiretroviral therapy initiation on CD4 cell count. *Am. J. Epidemiol*., 162, 471-478.

Comijs, H., Jonker, C., Beekman, A.T.F., & Deeg, D.J.H. (2001). The association between depressive symptoms and cognitive decline in community-dwelling elderly persons. *Int. J. Geriatr. Psychiatry*, 16, 361-367.

Cook, N., Cole, S., & Hennekens, C. (2002). Use of a marginal structural model to determine the effect of aspirin on cardiovascular mortality in the physicians' health study. *Am. J. Epidemiol.*, 155, 1045-1053.

Cronk, B.B., Johnson, D.K., & Burns, J.M. (2010). Body mass index and cognitive decline in mild cognitive impairment. *Alzheimer Dis. Assoc. Disord.*, 24, 126-130.

De Wit, L.M., van Straten, A., van Herten, M., Penninx, B., & Cuijpers, P. (2009). Depression and body mass index, a u-shaped association. *BMC Public Health*, 9:14.

Dragan, A., & Akhtar-Danesh, N. (2007). Relation between body mass index and depression: a structural equation modeling approach. *BMC Medical Research Methodology*, 7:17.

Geerlings, M.I., Schoevers, R.A., Beekman, A.T.F., Jonker, C., Deeg, D.J.H., Schmand, B., Ader, H.J., Bouter, L.M., & Van Tilburg, W. (2000). Depression and risk of cognitive decline and Alzheimer's disease; results of two prospective community-based studies in the Netherland. *British Journal of Psychiatry*, 176, 568-575.

Harrell, F. (2009). Survival analysis: logistic regression with splines. Available at: https://math.uc.edu/~deddens/yr2009/stat572/rcspline.txt

Hernan, M.A., Brumback, B., & Robins, J.M. (2000). Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology*, 11(5), 561-570.

Hernan, M.A., Brumback, B., & Robins, J.M. (2001). Marginal structural models to estimate the joint causal effect of nonrandomized treatments. *Journal of the American Statistical Association*, 96 (454), 440-448.

Hernan, M.A., Hernandez-Diaz, S., & Robins, J.M. (2004). A structural approach to selection bias. *Epidemiology*, 15, 615-625.

Hernan, M.A., & Robins, J.M. (2006). Estimating causal effects from epidemiological data. *J Epidemiol Community Health*, 60, 578-586.

Imai, K., & Ratkovic, M. (2014). Robust estimation of inverse probability weights for marginal structural models. Journal of the American Statistical Association. Available at: https://www.princeton.edu/~ratkovic/public/MSM.pdf

Joffe, M.M., Have, T.R.T., Feldman, H.I., & Kimmel, S.E. (2004). Model selection, confounder control, and marginal structural models: review and new applications. *The American Statistician*, 58(4), 272-279.

Johnston, E., Johnson S., McLeod, P., & Johnson, M. (2004). The relation of body mass index to depressive symptoms. *Can. J. Public Health*, 95, 179-183.

Kim, J. Noh, J., Park, J., & Kwon, Y.D. (2014). Body mass index and depressive symptoms in older adults: a cross-lagged panel analysis. *PLoS ONE*, 9(12).

Kobler, S., Van Boxtel, M.P.J., Van Os, J., Thomas, A.J., O'Brien, J.T., Jolles, J., Verhey, F.R.J., & Allardyce, J. (2010). Depressive symptoms and cognitive decline in community-dwelling older adults. *J. Am. Geriatr. Soc.*, 58, 873-879.

Kurlowicz, L., & Wallance, M. (1999). The mini mental state examination (MMSE). Try this: Best Practices in Nursing Care to Older Adults, 3(1), 1-2. Available at: https://www.mountsinai.on.ca/care/psych/on-call-resources/on-call-resources/mmse.pdf

Li, L., Evans, E., & Hser, Y.I. (2010). A marginal structural modeling approach to assess the cumulative effect of drug treatment on the later drug use abstinence. *J. Drug Issues*, 40(1), 221-240.

Markides, K. S. (2009). Hispanic Established Populations for the Epidemiologic Studies of the Elderly, 1993-1994: [Arizona, California, Colorado, New Mexico, and Texas]. ICPSR02851-v2. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2009-12-14.

Moodie, E.E.M., Stephens, D.A., & Klein, M.B. (2013). A marginal structural model for multiple-outcome survival data: assessing the impact of injection drug use on several causes of death in the Canadian co-infection cohort. *Statistics in Medicine*, 33, 1409-1425.

Moore, K.L., Neugebauer, R.S., van der Laan, M.J., & Tager, I.B. (2012). Causal inference in epidemiological studies with strong confounding. *Statistics in Medicine*, 31, 1380-1404.

Mortimer, K., Neugebauer, R., van der Laan, M., & Tager, I. (2005). An application of model-fitting procedures for marginal structural models. *American Journal of Epidemiology*, 162(4), 382-388.

Ng, T., Feng, L., Niti, M., & Yap,. K.B. (2008). Albumin, haemoglobin, BMI and cognitive performance in older adults. *Age and Ageing*, 37, 423-429.

Perrino, T., Mason, C.A., Brown, S.C., Spokane, A., & Szapocznik, J. (2008). Longitudinal relationships between cognitive functioning and depressive symptoms among Hispanic older adults. *J. Gerontal. B. Psychol. Sci. Soc. Sci.*, 63, 309-317.

Platt, R.W., Brookhart, M.A., Cole, S.R., Westreich, D., & Schisterman, E.F. (2013). An information criterion for marginal structural models. *Statistics in Medicine*, 32(8), 1383-1393.

Radloff, L.S. (1977). The CES-D scale: a new self-report depression scale for research in the general population. *Appl. Psych. Meas.*, 1, 385-401.

Raji, M.A., Reyes-Ortiz, C.A., Kuo, Y.F., Markides, K.S., & Ottenbacher, K.J. (2007). Depressive symptoms and cognitive change in older Mexican Americans. *J. Geriatr. Psychiatry Neurol.*, 20(3), 145-152.

Robins, J.M. (1998). Marginal structural models. In: 1997 proceedings of the section on Bayesian statistical science, Alexandria, VA: American Statistical Association, 1-10.

Robins, J.M. (1999). Association, causation, and marginal structural models. *Synthese*, 121(1-2), 151-179.

Robins, J.M., Greenland, S., & Hu, F. (1999). Estimation of the causal effect of a time-varying exposure on the marginal mean of a repeated binary outcome. *Journal of the American Statistical Association*, 94(447), 687-700.

Robins, J.M., Hernan, M.A., & Brumback, B. (2000). Marginal structural models and causal inference in epidemiology. *Epidemiology*, 11(5), 550-560.

Rosenbaum, P.R., & Rubin, D.B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70, 41-55.

Royston, P., Altman, D. G., & Sauerbrei, W. (2006), "Dichotomizing continuous predictors in multiple regression: a bad idea," *Statistics in Medicine*, 25, 127–141.

Sabia, S., Kivimaki, M., Shipley, M.J., Marmot, M.G., & Signh-Manoux, A. (2009). Body mass index over the adult life course and cognition in late midlife: the Whitehall II Cohort study. *Am. J. Clin. Nutr*., 89, 601-607.

Suemoto, C.K., Gilsanz, P., Mayeda, E.R., & Glymour, M.M. (2015). Body mass index and cognitive function: the potential for reverse causation. *International Journal of Obesity*, 1-7.

Tombaugh, T.N., & McIntyre, N.J. (1992). The mini-mental state examination: A comprehensive review. *JAGS*, 40, 922-953.

Vinkers, D.J., Gussekloo, J., Stek, M.L., Westendorp, R.G.J., & Van der Mast, R.C. (2004). Temporal relation between depression and cognitive impairment in old age: prospective population based study. *BMJ*, 329:881.

Wang, Y., Petersen, M.L., Bangberg, D., & van der Laan, M.J. (2006). Diagnosing bias in the inverse probability of treatment weighted estimator resulting from violation of experimental treatment assignment. Berkeley, CA: Division of Biostatistics, University of California. Available at: http://biostats.bepress.com/ucbbiostat/paper211.

Whitmer, R.A., Gunderson, E.P., Barrett-Connor, E., Quesenberry, C.P., & Yaffe, K. (2005). Obesity in middle age and future risk of dementia: a 27 year longitudinal population based study. *BMJ*, doi: 10.1136.

Wilson, R.S., Barnes, L.L., Mendes de Leon, C.F., Aggarwal, N.T., Schneider, J.S., Bach, J., Pilat, J., Beckett, L.A., Arnold, S.E., Evans, D.A., & Bennett, D.A. (2002). Depressive symptoms, cognitive decline, and risk of AD in older persons. *Neurology*, 59, 364-370.

Wilson, R.S., Mendes de Leon, C.F., Bennett, D.A., Bienias, J.L., & Evans, D.A. (2004). Depressive symptoms and cognitive decline in a community population of older persons. *J. Neurol Neurosurg. Psychiatry*, 75, 126-129.

APPENDIX

SAS-CODE

```
/*Restrict Cubic Splines*/
%MACRO
 RCSPLINE(x,knot1,knot2,knot3,knot4,knot5,knot6,knot7,knot8,knot9,knot10,
 norm=2);
 %LOCAL j v7 k tk tk1 t k1 k2;
 %LET v7=&x; %IF %LENGTH(&v7)=8 %THEN %LET v7=%SUBSTR(&v7,1,7);
    %*Get no. knots, last knot, next to last knot;
     %DO k=1 %TO 10;
     %IF %QUOTE(&&knot&k)=  %THEN %GOTO nomorek;
     %END;
 %LET k=11;
 %nomorek: %LET k=%EVAL(&k-1); %LET k1=%EVAL(&k-1); %LET k2=%EVAL(&k-2);
 %IF &k<3 %THEN %PUT ERROR: <3 KNOTS GIVEN.  NO SPLINE VARIABLES
    CREATED.;
       %ELSE %DO;
       %LET tk=&&knot&k;
       %LET tk1=&&knot&k1;
       DROP _kd_; _kd_=
       %IF &norm=0 %THEN 1;
       %ELSE %IF &norm=1 %THEN &tk - &tk1;
       %ELSE (&tk - &knot1)**.666666666666; ;
            %DO j=1 %TO &k2;
             %LET t=&&knot&j;
 &v7&j=max((&x-&t)/_kd_,0)**3+((&tk1-&t)*max((&x-&tk)/_kd_,0)**3
               -(&tk-&t)*max((&x-&tk1)/_kd_,0)**3)/(&tk-&tk1)%STR(;);
            %END;
   %END;
 %MEND;
/*weight wave1*/
data one;
```

```
set project1;
%RCSPLINE(AGE1,65,68,71.5,77,86);
%RCSPLINE(CESD1,0,7,23);
run;
proc logistic data=one descending noprint;
class predrug1 income gender education marry1 smoke1 cogscore1;
model cogscore1=age1 age11-age13 predrug1 income gender education marry1 smoke1/rsq;
output out=a1 predicted=n1;
run;
data n1;
set a1;
if cogscore1=1 then wn1=n1;if cogscore1=0 then wn1=1-n1;
keep ID CogScore1 wn1;
run;
proc logistic data=one descending noprint;
class predrug1 income gender education marry1 smoke1 cogscore1;
model cogscore1=age1 age11-age13 CESD1 CESD11 predrug1 income gender education
    marry1 smoke1/rsq;
output out=a2 predicted=d1;
run;
data d1;
set a2;
if cogscore1=1 then wd1=d1; if cogscore1=0 then wd1=1-d1;
keep ID CogScore1 wd1;
run;
data w1;
merge n1 d1;
by ID;
w1=wn1/wd1;
if w1=. then w1=1;
```

```
keep ID w1;
run;
/*weight wave2*/
data two;
set project1;
%RCSPLINE(AGE2,67,70,73,79,87);
%RCSPLINE(CESD1,0,7,23);
%RCSPLINE(CESD2,0,4,18);
run;
proc logistic data=two descending noprint;
class predrug1 income gender education marry1 smoke1 cogscore1 cogscore2;
model cogscore2=CogScore1 age2 age21-age23 predrug1 income gender education marry1
    smoke1;
output out=b1 predicted=n2;
run;
data n2;
set b1;
if cogscore2=1 then wn2=n2;if cogscore2=0 then wn2=1-n2;
keep ID CogScore2 wn2;
run;
proc logistic data=two descending noprint;
class predrug1 income gender education marry1 smoke1 cogscore1 cogscore2;
model cogscore2=age2 age21-age23 CESD2 CESD21 CogScore1 CESD1 CESD11 predrug1
    income gender education marry1 smoke1/rsq;
output out=b2 predicted=d2;
run;
data d2;
set b2;
if cogscore2=1 then wd2=d2; if cogscore2=0 then wd2=1-d2;
keep ID CogScore2 wd2;
```

```
run;
data w2;
merge n2 d2;
by ID;
w2=wn2/wd2;
if w2=. then w2=1;
keep ID w2;
run;
/*weight wave3*/
data three;
set project1;
%RCSPLINE(AGE3,70,73,76,81,89);
%RCSPLINE(CESD2,0,4,18);
%RCSPLINE(CESD3,0,6,22);
run;
proc logistic data=three descending noprint;
class predrug1 income gender education marry1 smoke1 cogscore2 cogscore3;
model cogscore3=CogScore2 age3 age31-age33 predrug1 income gender education marry1
    smoke1;
output out=c1 predicted=n3;
run;
data n3;
set c1;
if cogscore3=1 then wn3=n3;if cogscore3=0 then wn3=1-n3;
keep ID CogScore3 wn3;
run;
proc logistic data=three descending noprint;
class predrug1 income gender education marry1 smoke1 cogscore2 cogscore3;
model cogscore3=age3 age31-age33 CESD2 CESD21 CogScore2 CESD3 CESD31 predrug1
    income gender education marry1 smoke1/rsq;
```

```
output out=c2 predicted=d3;
run;
data d3;
set c2;
if cogscore3=1 then wd3=d3; if cogscore3=0 then wd3=1-d3;
keep ID CogScore3 wd3;
run;
data w3;
merge n3 d3;
by ID;
w3=wn3/wd3;
if w3=. then w3=1;
keep ID w3;
run;
/*weight wave4*/
data four;
set project1;
%RCSPLINE(AGE4,72,75,78,82,90);
%RCSPLINE(CESD3,0,6,22);
%RCSPLINE(CESD4,0,5,18);
run;
proc logistic data=four descending noprint;
class predrug1 income gender education marry1 smoke1 cogscore3 cogscore4;
model cogscore4=CogScore3 age4 age41-age43 predrug1 income gender education marry1
    smoke1;
output out=e1 predicted=n4;
run;
data n4;
set e1;
if cogscore4=1 then wn4=n4;if cogscore4=0 then wn4=1-n4;
```

```
keep ID CogScore4 wn4;
run;
proc logistic data=four descending noprint;
class predrug1 income gender education marry1 smoke1 cogscore3 cogscore4;
model cogscore4=age4 age41-age43 CESD3 CESD31 CogScore3 CESD4 CESD41 predrug1
    income gender education marry1 smoke1/rsq;
output out=e2 predicted=d4;
run;
data d4;
set e2;
if cogscore4=1 then wd4=d4; if cogscore4=0 then wd4=1-d4;
keep ID CogScore4 wd4;
run;
data w4;
merge n4 d4;
by ID;
w4=wn4/wd4;
if w4=. then w4=1;
keep ID w4;
run;
/*censor wave2*/
proc logistic data=one descending noprint;
class predrug1 income gender education marry1 smoke1 cogscore1 censor2;
model censor2=cogscore1 age1 age11-age13 predrug1 income gender education marry1 smoke1;
output out=f1 predicted=n5;
run;
data n5;
set f1;
if censor2=0 then cn2=1-n5; if censor2=1 then cn2=1;
keep ID Censor2 cn2;
```

```
run;
proc logistic data=one descending noprint;
class predrug1 income gender education marry1 smoke1 cogscore1 censor2;
model censor2=cogscore1 age1 age11-age13 CESD1 CESD11 predrug1 income gender
    education marry1 smoke1;
output out=f2 predicted=d5;
run;
data d5;
set f2;
if censor2=0 then cd2=1-d5;if censor2=1 then cd2=1;
keep ID censor2 cd2;
run;
data w5;
merge n5 d5;
by ID;
cw2=cn2/cd2;
if cw2=. then cw2=1;
keep ID cw2;
run;
/*censor wave3*/
proc logistic data=two descending noprint;
class predrug1 income gender education marry1 smoke1 cogscore2 censor3;
model censor3=cogscore2 age2 age21-age23 predrug1 income gender education marry1 smoke1;
output out=g1 predicted=n6;
run;
data n6;
set g1;
if censor3=0 then cn3=1-n6; if censor3=1 then cn3=1;
keep ID Censor3 cn3;
run;
```

```
proc logistic data=two descending noprint;
class predrug1 income gender education marry1 smoke1 cogscore2 censor3;
model censor3=cogscore2 age2 age21-age23 CESD2 CESD21 predrug1 income gender
    education marry1 smoke1;
output out=g2 predicted=d6;
run;
data d6;
set g2;
if censor3=0 then cd3=1-d6;if censor3=1 then cd3=1;
keep ID censor3 cd3;
run;
data w6;
merge n6 d6;
by ID;
cw3=cn3/cd3;
if cw3=0 then cw3=1;
keep ID cw3;
run;
/*censor wave4*/
proc logistic data=three descending noprint;
class predrug1 income gender education marry1 smoke1 cogscore3 censor4;
model censor4=cogscore3 age3 age31-age33 predrug1 income gender education marry1 smoke1;
output out=h1 predicted=n7;
run;
data n7;
set h1;
if censor4=0 then cn4=1-n7; if censor4=1 then cn4=1;
keep ID Censor4 cn4;
run;
proc logistic data=three descending noprint;
```

```
class predrug1 income gender education marry1 smoke1 cogscore3 censor4;
model censor4=cogscore3 age3 age31-age33 CESD3 CESD31 predrug1 income gender
    education marry1 smoke1;
output out=h2 predicted=d7;
run;
data d7;
set h2;
if censor4=0 then cd4=1-d7;if censor4=1 then cd4=1;
keep ID censor4 cd4;
run;
data w7;
merge n7 d7;
by ID;
cw4=cn4/cd4;
if cw4=. then cw4=1;
keep ID cw4;
run;
data weight;
merge w1 w2 w3 w4 w5 w6 w7;
by ID;
sw1=w1;sw2=w1*w2*cw2;sw3=w1*w2*w3*cw2*cw3;sw4=w1*w2*w3*w4*cw2*cw3*cw4;
keep ID sw1 sw2 sw3 sw4;
run;
data project2;
merge weight project1;
by ID;
if sw1=1 then sw1=.;if sw2=1 then sw2=.;if sw3=1 then sw3=.;if sw4=1 then sw4=.;
run;
data project3;
set project2;
```

```
array age_all(4) age1-age4;
array Smoke_all(4) Smoke1-Smoke4;
array Marry_all(4) Marry1-Marry4;
array CogScore_all(4) CogScore1-CogScore4;
array Censor_all(4) Censor1-Censor4;
array sw_all(4) sw1-sw4;
array BMI_all(4) BMI1-BMI4;
array CESD_all(4) CESD1-CESD4;
do wave=1 to 4;
age=age_all{wave};
Smoke=Smoke_all{wave};
Marry=Marry_all{wave};
CogScore=CogScore_all{wave};
Censor=Censor_all{wave};
sw=sw_all{wave};
BMI=BMI_all{wave};
CESD=CESD_all{wave};
output;
end;
drop age1-age4 Smoke2-Smoke4 Marry2-Marry4 CogScore1-CogScore4 Censor1-Censor4 sw1-
    sw4 BMI1-BMI4 CESD2-CESD4;
run;
proc corr data=project3;
var cogscore CESD BMI;
weight sw;
run;
proc sort data=project3;
by cogscore CESD;
run;
proc boxplot data=project3;
```

```
plot CESD*cogscore;
insetgroup mean stddev;
run;
proc means data=project3;
by cogscore;
var CESD;
weight sw;
run;
proc means data=project3;
var sw;
run;
proc univariate data=project3;
var sw;
histogram;
run;
/*unweighted model*/
proc mixed data=project3;
class ID cogscore (ref='0');
model BMI=cogscore /solution;
repeated /type=un subject=ID;
run;
proc mixed data=project3;
class ID cogscore (ref='0') education income marry1 gender smoke1 predrug1;
model BMI=cogscore age education income marry1 gender smoke1 predrug1/solution;
repeated /type=un subject=ID;
run;
proc mixed data=project3;
class ID cogscore (ref='0') education income marry1 gender smoke1 predrug1;
model BMI=cogscore age education income marry1 gender smoke1 predrug1 CESD/solution;
repeated /type=un subject=ID;
```

```
run;
/*weighted model*/
proc mixed data=project3;
class ID cogscore (ref='0') education income marry1 gender smoke1 predrug1;
model BMI=cogscore age education income marry1 gender smoke1 predrug1 CESD1/solution;
repeated /type=un subject=ID;
weight sw;
run;
/*covariance selection*/
proc mixed data=project3;
class ID education income marry1 gender smoke1 predrug1 cogscore(ref='0');
model BMI=cogscore age education income marry1 gender smoke1 predrug1 CESD1/solution ;
repeated /type=vc subject=ID;
weight sw;
run;
proc mixed data=project3;
class ID education income marry1 gender smoke1 predrug1 cogscore(ref='0');
model BMI=cogscore age education income marry1 gender smoke1 predrug1 CESD1/solution ;
repeated /type=cs subject=ID;
weight sw;
run;
proc mixed data=project3;
class ID education income marry1 gender smoke1 predrug1 cogscore(ref='0');
model BMI=cogscore age education income marry1 gender smoke1 predrug1 CESD1/solution ;
repeated /type=csh subject=ID;
weight sw;
run;
proc mixed data=project3;
class ID education income marry1 gender smoke1 predrug1 cogscore(ref='0');
model BMI=cogscore age education income marry1 gender smoke1 predrug1 CESD1/solution ;
```

```
repeated /type=ar(1) subject=ID;
weight sw;
run;
proc mixed data=project3;
class ID education income marry1 gender smoke1 predrug1 cogscore(ref='0');
model BMI=cogscore age education income marry1 gender smoke1 predrug1 CESD1/solution ;
repeated /type=arh(1) subject=ID;
weight sw;
run;
proc mixed data=project3;
class ID education income marry1 gender smoke1 predrug1 cogscore(ref='0');
model BMI=cogscore age education income marry1 gender smoke1 predrug1 CESD1/solution ;
random int/sub=ID;
repeated /type=ar(1) subject=ID;
weight sw;
run;
proc mixed data=project3;
class ID education income marry1 gender smoke1 predrug1 cogscore(ref='0');
model BMI=cogscore age education income marry1 gender smoke1 predrug1 CESD1/solution ;
repeated /type=toep subject=ID;
weight sw;
run;
proc mixed data=project3;
class ID education income marry1 gender smoke1 predrug1 cogscore(ref='0');
model BMI=cogscore age education income marry1 gender smoke1 predrug1 CESD1/solution ;
repeated /type=toeph subject=ID;
weight sw;
run;
proc mixed data=project3;
class ID education income marry1 gender smoke1 predrug1 cogscore(ref='0');
```

```
model BMI=cogscore age education income marry1 gender smoke1 predrug1 CESD1/solution ;
repeated /type=hf subject=ID;
weight sw;
run;
proc mixed data=project3;
class ID education income marry1 gender smoke1 predrug1 cogscore(ref='0');
model BMI=cogscore age education income marry1 gender smoke1 predrug1 CESD1/solution ;
repeated /type=arma(1,1) subject=ID;
weight sw;
run;
proc mixed data=project3;
class ID education income marry1 gender smoke1 predrug1 cogscore(ref='0');
model BMI=cogscore age education income marry1 gender smoke1 predrug1 CESD1/solution ;
repeated /type=fa(1) subject=ID;
weight sw;
run;
proc mixed data=project3;
class ID education income marry1 gender smoke1 predrug1 cogscore(ref='0');
model BMI=cogscore age education income marry1 gender smoke1 predrug1 CESD1/solution ;
repeated /type=ante(1) subject=ID;
weight sw;
run;
/*Final MSM*/
proc mixed data=project3;
class ID education income marry1 gender smoke1 predrug1 cogscore(ref='0');
model BMI=cogscore age income  gender smoke1 predrug1 CESD1/solution ;
random int/sub=ID;
repeated /type=ar(1) subject=ID;
weight sw;
run;
```

```
/*truncate 0.5th and 99.5th percentile*/
data project3;
set project3;
if 0<sw<0.5287 then sw=0.5287;if 1.9800<sw then sw=1.9800;
run;
proc means data=project3;
var sw;
run;
proc mixed data=project3;
class ID education income marry1 gender smoke1 predrug1 cogscore(ref='0');
model BMI=cogscore age income  gender smoke1 predrug1 CESD1/solution ;
random int/sub=ID;
repeated /type=ar(1) subject=ID;
weight sw;
run;
/*truncate 1st and 99th percentile*/
data project3;
set project3;
if 0<sw<0.5736 then sw=0.5736;if 1.6782<sw then sw=1.6782;
run;
proc means data=project3;
var sw;
run;
proc mixed data=project3;
class ID education income marry1 gender smoke1 predrug1 cogscore(ref='0');
model BMI=cogscore age income  gender smoke1 predrug1 CESD1/solution ;
random int/sub=ID;
repeated /type=ar(1) subject=ID;
weight sw;
run;
```

```
/*truncate 2.5th and 97.5th percentile*/
data project3;
set project3;
if 0<sw<0.6479 then sw=0.6479;if 1.4596<sw then sw=1.4596;
run;
proc means data=project3;
var sw;
run;
proc mixed data=project3;
class ID education income marry1 gender smoke1 predrug1 cogscore(ref='0');
model BMI=cogscore age income  gender smoke1 predrug1 CESD1/solution ;
random int/sub=ID;
repeated /type=ar(1) subject=ID;
weight sw;
run;
/*truncate 5th and 95th percentile*/
data project3;
set project3;
if 0<sw<0.7166 then sw=0.7166;if 1.2901<sw then sw=1.2901;
run;
proc means data=project3;
var sw;
run;
proc mixed data=project3;
class ID education income marry1 gender smoke1 predrug1 cogscore(ref='0');
model BMI=cogscore age income  gender smoke1 predrug1 CESD1/solution ;
random int/sub=ID;
repeated /type=ar(1) subject=ID;
weight sw;
run;
```

```
proc mixed data=project3;
class ID ;
model BMI=cogscore age education income marry1 gender smoke1 predrug1/solution;
repeated /type=un subject=ID;
weight sw;
run;
/*sensitivity function */
data project1;
set project1;
drop ADL1-ADL4 health1-health4 chronic1-chronic4 predrug2 obesity1-obesity4;
a1=2*cogscore1-1;a2=2*cogscore2-1;a3=2*cogscore3-1;
b1=1-cogscore1;b2=1-cogscore2;b3=1-cogscore3;
r1=5;r2=-0.1;
run;
proc logistic data=project1 descending noprint;
model b1=CESD1;
output out=temp1 p=p1;
run;
proc logistic data=project1 descending noprint;
model b2=CESD2;
output out=temp2 p=p2;
run;
proc logistic data=project1 descending noprint;
model b3=CESD3;
output out=temp3 p=p3;
run;
data temp1;
set temp1;
keep ID p1;
run;
```

```
data temp2;
set temp2;
keep ID p2;
run;
data temp3;
set temp3;
keep ID p3;
run;
data adjust;
merge project1 temp1 temp2 temp3;
by ID;
run;
data adjust;
set adjust;
BMI_ad1=BMI1;
BMI_ad2=BMI2-r1*a1*p1;
BMI_ad3=BMI3-r1*a1*p1-r1*a2*p2;
BMI_ad4=BMI4-r1*a1*p1-r1*a2*p2-r1*a3*p3;
run;
```