Received: 14 February 2020 | Revised: 23 December 2020 | Accepted: 3 February 2021

DOI: 10.1002/sim.8921

RESEARCH ARTICLE

Statistics in Medicine WILEY

Accounting for selection bias due to death in estimating the effect of wealth shock on cognition for the Health and Retirement Study

Yaoyuan Vincent Tan[1] | **Carol A. C. Flannagan**[2] | **Lindsay R. Pool**[3] | **Michael R. Elliott**[4]

[1]Biometrics Division, Vertex Pharmaceuticals, Boston, Massachusetts, USA

[2]Transportation Research Institute, University of Michigan, Ann Arbor, Michigan, USA

[3]Preventive Medicine (Epidemiology), Feinberg School of Medicine, Northwestern University, Chicago, Illinois, USA

[4]Biostatistics Department, School of Public Health, University of Michigan, Ann Arbor, Michigan, USA

Correspondence
Yaoyuan Vincent Tan, Biometrics Division, Vertex Pharmaceuticals, 50 Northern Avenue, Boston, Suite 11063I, Boston, MA 02210.
Email: vincent_tan@vrtx.com

The Health and Retirement Study (HRS) is a longitudinal study of U.S. adults enrolled at age 50 and older. We were interested in investigating the effect of a sudden large decline in wealth on the cognitive ability of subjects measured using a dataset provided composite score. However, our analysis was complicated by the lack of randomization, time-dependent confounding, and a substantial fraction of the sample and population will die during follow-up leading to some of our outcomes being censored. The common method to handle this type of problem is marginal structural models (MSM). Although MSM produces valid estimates, this may not be the most appropriate method to reflect a useful real-world situation because MSM upweights subjects who are more likely to die to obtain a hypothetical population that over time, resembles that would have been obtained in the absence of death. A more refined and practical framework, principal stratification (PS), would be to restrict analysis to the strata of the population that would survive regardless of negative wealth shock experience. In this work, we propose a new algorithm for the estimation of the treatment effect under PS by imputing the counterfactual survival status and outcomes. Simulation studies suggest that our algorithm works well in various scenarios. We found no evidence that a negative wealth shock experience would affect the cognitive score of HRS subjects.

KEYWORDS

Bayesian additive regression trees, causal inference, longitudinal study, missing data, penalized spline of propensity methods in treatment comparisons, time-dependent confounding

1 | INTRODUCTION

Late middle age adults commonly experience chronic health conditions as well as declining cognitive abilities. Factors known to be associated with accelerated decrease in cognitive abilities include smoking, high alcohol consumption, physical inactivity, high dietary intake of sodium and saturated fats, low dietary intake of fruits and vegetables;[1] hypertension, elevated serum cholesterol, diabetes, obesity, cerebrovascular and cardiovascular disease;[2] depression, lower socioeconomic status, and exposure to acute stressful life events and chronic perceived stress.[3] In particular, the acute stress of a sudden decrease in wealth—"a negative wealth shock"—may have a negative impact on the cognitive ability of late middle aged adults. Because income typically exceeds consumption at this stage in life, there are fewer remaining years left to

replenish the lost wealth.[4] The stress of losing substantial wealth during the savings period of the life cycle coupled with the pressure to replenish the lost wealth can lead to stress-related health conditions which in turn reduces the cognitive ability of an individual.[5]

To investigate the effect of a negative wealth shock on the cognitive ability of late middle age U.S. adults, we used the Health and Retirement Study (HRS). The HRS is a biennial survey that began in 1992 and collects detailed financial status and health information from U.S. adults age 50 and older.[6] Three issues arose when we tried to estimate the causal effect of a negative wealth shock on cognitive ability using this dataset. First, there was a lack of randomization. Negative wealth shocks are not randomly distributed in the population, but rather are confounded by factors such as gender and socio-economic status. Second, due to time-dependent confounding, the risk of the wealth shock at any point in time may depend on the prior cognitive ability up to the point. Finally, a sufficiently large fraction of our population will die during follow-up, leading to "censoring by death" implying that their cognitive ability measure would be missing. This "missingness" is different from the measure of cognition being "missing" due to dropout, where the cognitive ability measure exists but is unobserved. Note further that subjects observed to have survived in the absence of a negative wealth shock include: (i) subjects that would survive regardless of whether they received a negative wealth shock and (ii) subjects that would survive only in the absence of a negative wealth shock. Because death may be positively associated with variables that increase the risk of a negative wealth shock, increased cognitive ability decline, and the experience of a negative wealth shock, the censored cognitive ability outcome may be confounded by death.

Methods dealing with these three types of barriers to causal inference have been developed requiring varying assumption. Under the assumption that, conditional on available covariates, negative wealth shocks would truly be random, conditioning on the probability of receiving a negative wealth shock as a function of these covariates—the propensity scores[7]—can be used to remove the effect of confounding, either by regression, matching, or weighting.[8] For the second issue—time-dependent confounding—marginal structural models MSM,[9] and more recently, penalized spline of propensity methods in treatment comparisons PENCOMP,[10] have been used to account for confounding by the time-dependence association of the cognitive measures. For censoring by death, MSMs have typically been extended by multiplying the treatment assignment weights with the inverse of the predicted probability of death.[11] The issue with this approach—perhaps under appreciated—is that the resulting pseudo-population is not only balanced with respect to exposure "assignment," but also "immortal," in the sense that those more likely to die are upweighted so that the population over time resembles that would have been obtained in the absence of death up till time t.[12] This is arguably not a sensible population for inference, at least from a policy and public health perspective. We shall elaborate on this point using a synthetic example in Section 1.1.

A more refined approach would be to compare the effect of negative wealth shock on cognitive ability among subjects who would have survived whether they experienced a negative wealth shock or not. This approach uses the potential outcomes definition of Neyman[13] and Rubin,[14] which defines causal effects as the within-subject difference of an outcome at a particular time under different exposure or treatment regimen, averaged over the population. The result will be the recognition that the treatment effect for some subjects will be nonexistent and hence should not be included in the calculation of the average treatment effect. This idea is not new[15] and can be viewed as a specific example of the principal stratification (PS) method discussed in Frangakis and Rubin.[16] Our innovation here is to embed this in a longitudinal setting where time-dependent confounding is present. We view this as a large missing data problem where survival status and, among survivors, unobserved outcomes under a given exposure pattern, are imputed. We shall illustrate the necessity for this refinement using a short synthetic example.

1.1 | MSM approach vs principal strata approach

Suppose we have an exposure $Z = \{0, 1\}$, negative wealth shock in our context, with 0 representing no wealth shock and 1 representing wealth shock. We use $S(Z) = \{0, 1\}$ to denote the potential outcome of whether the subject survives with 0 representing the subject died and 1 representing the subject survived. The refinement PS introduces is the recognition that there are four possible strata based on the combination of $(S(0), S(1))$: $(S(0) = 0, S(1) = 0)$—subject dies regardless of whether they are shocked or not, $(S(0) = 0, S(1) = 1)$—subject survives only when they are shocked, $(S(0) = 1, S(1) = 0)$—subject survives only when they are NOT shocked, and $(S(0) = 1, S(1) = 1)$—subject survives regardless. Note that stratum $(S(0) = 0, S(1) = 1)$ is arguably not sensible and we assume this group of subjects do not exist (This is the monoticity assumption which we shall elaborate in Section 3.). This refined grouping gives rise to a slightly

TABLE 1 Synthetic dataset with censoring by death to illustrate the necessity of the principal stratification framework

Subject	$S(0)$	$S(1)$	$Y(0,(S(0),S(1)))$	$Y(1,(S(0),S(1)))$	Z	S	Y
1	1	1	−1	−6	0	1	−1
2	1	1	−2	−7	0	1	−2
3	1	1	−3	−8	1	1	−8
4	1	1	−4	−9	1	1	−9
5	1	1	−5	−10	1	1	−10
6	1	0	−11	−6	0	1	−11
7	1	0	−12	−7	0	1	−12
8	1	0	−13	−8	1	0	?
9	1	0	−14	−9	1	0	?
10	1	0	−15	−10	1	0	?
11	0	0	−11	−6	0	0	?
12	0	0	−12	−7	0	0	?
13	0	0	−13	−8	1	0	?
14	0	0	−14	−9	1	0	?
15	0	0	−15	−10	1	0	?

more elaborate potential outcomes notation, $Y(Z,(S(0),S(1)))$, the outcome under different exposure and different PS. The idea of MSM still fits into this framework since PS is just the refinement of subjects in the population into three different strata (in our context). Table 1 shows a synthetic example of the potential outcomes under all the three possible strata.

Subjects 1 to 5 belong to the always survivors group, that is, $(S(0)=1,S(1)=1)$, Subjects 6 to 10 belong to the helped group, that is, $(S(0)=1,S(1)=0)$, and finally Subjects 11 to 15 belong to the doomed group, that is, $(S(0)=0,S(1)=0)$. Columns 2 to 5 of Table 1 show the potential outcomes while columns 6 to 8 show the observed outcomes. When running a study, the researcher only observes columns 6 to 8 and is unable to observe columns 2 to 5 although these columns are precisely the information the researcher needs in order to obtain valid and practical results. Suppose the researcher takes the naïve approach and ignores the missing data in column 8 and then calculates $E[Y|Z=1]-E[Y|Z=0]$. This yields an average treatment effect of −2.5. This is misleading because if we looked at the potential outcomes for this population, $E[Y(1,(S(0),S(1)))]-E[Y(0,(S(0),S(1)))]=1.67$, which suggests a completely opposite effect. In fact, 1.67 will be the estimated effect MSM produces assuming that MSM is able to accurately recover all the counterfactuals, that is, the $Y(0,(S(0),S(1)))$s and $Y(1,(S(0),S(1)))$s that we did not observe due to the exposure allocation as well as death. To understand the result produced by MSM a little better, let us interpret it in context: if we hypothetically gave all subjects the exposure of a negative wealth shock as well as somehow kept them alive and then take the average of their cognitive score; and then we compare this average score with the hypothetical situation where we prevented negative wealth shock from all subjects, kept them alive, and then measured the average cognitive score produced, we will find that allocating subjects to a negative wealth shock produces a cognitive score which is 1.67 higher on average compared to when we prevented negative wealth shocks in the population.

One may claim this to be valid, which it is, be content and report the result to policy makers. But it is clear that there is a huge issue with this interpretation. Unless researchers are able to somehow manipulate death, the potential outcomes $Y(0,(S(0),S(1)))$ and $Y(1,(S(0),S(1)))$ for subjects 11 to 15 can never be realized in a real-life setting because these subjects will die regardless of allocation to negative wealth shock or the prevention of one. Similarly, the potential outcome $Y(1,(S(0),S(1)))$ for subjects 6 to 10 can never be realized in the real world because they will die under such a setting. So although the solution produced by MSM is valid, it is not useful for the policy maker because it describes a situation which is not replicable in the real world and worse, might produce the wrong conclusion because if we used $E[Y(1,(S(0)=1,S(1)=1))]-E[Y(0,(S(0)=1,S(1)=1))]$ instead, we can see that it is −5. This estimand can be interpreted as: given the hypothetical situation where we allocated all subjects to a negative wealth shock and then compare

this with the hypothetical situation where we prevented all subjects from receiving a wealth shock, keeping only subjects who survived under these two hypothetical situations, and then take the difference in the average cognitive score of the two groups, we will find that population average cognitive ability score of subjects having exposed to a negative wealth shock will be −5 lower compared to the cognitive ability average score of the population where we hypothetically prevented all negative wealth shocks; which is clearly much more sensible, negative wealth shock reduces cognitive ability, and a better representation of the real world effect of the policy on the population compared to the naïve or MSM approach.

We organize our paper as follows. In Section 2, we set up the framework for our problem, and provide a brief review of the methods we require in our proposed algorithm. We develop our proposed method in Section 3. Section 4 explores some of the empirical properties of our proposed method compared to a naïve method and MSM using a simulation study. We then discuss how negative wealth shock and cognitive ability were determined as well as present the results of our analysis in Section 5. Section 6 discusses the implication of our results and identifies future work.

2 | REVIEW OF RELEVANT METHODS

2.1 | Notation

Let V be the set of baseline covariates, $Z(t)$ be our exposure at time $t = 1, \ldots, T$ where $Z(t) = 1$ indicates a subject receiving a negative wealth shock at t and $Z(t) = 0$ indicates no negative wealth shock, and $W(t)$ be the covariates that may vary with time, but are unaffected by the exposure. For any variable U, we use $\bar{U}(t) = \{U(1), \ldots, U(t)\}$, $\bar{U} = \{U(1), \ldots, U(T)\}$, and $\underline{U}(t) = \{U(t+1), \ldots, U(T)\}$ to denote the profile of the variable at time t, the full profile until time T, and the profile from time $t+1$ to T, respectively. Let $\bar{z}$ be a possible realization from $\bar{Z}$, $\bar{X}_{\bar{z}}$ be the covariates that are affected by the exposure profile $\bar{z}$, and $\bar{Y}_{\bar{z}}$ be the potential cognitive ability measure (cognitive score) profile under exposure profile $\bar{z}$. Finally, we let $\bar{S}_{\bar{z}(T-1)}$ be the potential survival profile for our dataset. We also assumed that our variables occur in the following sequence, $v, w(1), z(1), x(1), y(1), s(1), \ldots, s(T-1), w(T), z(T), x(T), y(T)$. Similar to Poll et al,[17] we assume that a negative wealth shock is an "absorbing state" so that once a subject receives a negative wealth shock at time t ($Z(t) = 1$), the subject would be "forever" shocked, that is $\underline{Z}(t) = 1$. In addition, subjects who die at time t will have structurally missing data for the future time-varying outcomes and covariates, that is, $S(t) = 0$ implies $\underline{Z}(t) = \underline{W}(t) = \underline{X}_{\bar{z}(t)}(t) = \underline{Y}_{\bar{z}(t)}(t) = NA$, where "NA" indicates a structurally missing observation.

2.2 | Penalized spline of propensity methods for treatment comparison

To tackle the time-dependent confounding on observational longitudinal studies, Zhou, Elliott, and Little[10] proposed the Penalized Spline of Propensity Methods for Treatment Comparison (PENCOMP). In brief, PENCOMP views the time-dependent confounding issue as a big missing data problem and then employs identifying assumptions to construct valid imputation models for the counterfactuals. In order to enhance the robustness of the imputed counterfactuals, Zhou, Elliott, and Little[10] proposed to use the penalized splines of propensity prediction (PSPP) proposed by Zhang and Little[18] as the imputation method. PSPP is a doubly robust method that imposes a penalized spline on the propensity score of missingness (treatment allocation in the context of causal inference) and then combines this with predictors of the outcome in a regression model. We describe a particular implementation of PENCOMP in Appendix S1 A.2.

3 | CENSORING BY DEATH

3.1 | Identifiability assumptions

Before we introduce our proposed algorithm proper, we require a few identifying assumptions in order to allow us to construct valid imputation models. First, we require

$$Pr(S_{Z(t-1)} = 1 | \bar{z}(t-1), \bar{y}_{\bar{z}(t-1)}, \bar{x}_{\bar{z}(t-1)}, \bar{w}(t-1), v) > 0, \tag{1}$$

and $Pr(Z(t)=1|\bar{z}(t-1),\bar{y}_{\bar{z}(t-1)}=0,\bar{x}_{\bar{z}(t-1)},\bar{w}(t-1),v)>0$ for any $Z(t)$, that is, the probability of survival under wealth shock profile $\bar{z}(t-1)$ and the probability of wealth shock at time t are bounded away from 0. This is an extension of the standard positivity assumption to allow that at least some subjects will be observed to have survived under all exposure profiles we are interested in, and that wealth shocks are possible if they have not already occurred. Second, we require no interference between subjects, $Y_{\bar{Z}}=Y_{\bar{z}}$, that is, the potential outcome of subject i is independent of whatever exposure subject j is allocated to $i\neq j$ as well as $X_{\bar{Z}}=X_{\bar{z}}$. Third, we need the no unmeasured confounding and sequential randomization condition: $Y_{\bar{Z}(t)}\perp Z(t)|\bar{z}(t-1),\bar{y}_{\bar{z}(t-1)},\bar{x}_{\bar{z}(t-1)},\bar{w}(t-1),v$. In addition, we also require that

$$Y_{\bar{Z}(t)}\perp X_{\bar{Z}(t)}|\bar{x}_{\bar{z}(t-1)},\bar{y}_{\bar{z}(t-1)},\bar{z}(t),v, \tag{2}$$

and

$$X_{\bar{Z}(t)}\perp Z(t)|\bar{x}_{\bar{z}(t-1)},\bar{y}_{\bar{z}(t-1)},\bar{z}(t-1),v. \tag{3}$$

Finally, we assume "monotonicity":

$$\text{If } Z(1)\leq Z'(1),Z(2)\leq Z'(2),\ \dots\ ,Z(t-1)\leq Z'(t-1) \text{ then } S_{\bar{Z}(t-1)}\geq S_{\bar{Z}'(t-1)}, \tag{4}$$

$Z(i)\neq Z'(i)$ for any i. That is, for subjects observed to have survived a negative wealth shock, we assume they would also survive in the absence of that wealth shock.

3.2 | Determining the principal strata

Table 2 shows the set of potential outcomes and structural missing values for $t=3$ and no time-varying covariates. In this table, "x" indicates an observed value, "?" represents a missing observation which needs to be imputed, and "NA" indicates a structurally missing observation. The goal of our analysis is to provide inference about contrasts where both potential outcomes exist, that is $E[\overline{Y}_{\bar{z}(t)}-\overline{Y}_{\bar{z}'(t)}|S_{\bar{z}(t-1)}=S_{\bar{z}'(t-1)}=1]$, where $z(l)\neq z'(l)$ for at least one l with $l=1,\dots,t$, that is, we condition on subjects who would potentially survive under two different exposure profiles $\bar{z}(t)$ and $\bar{z}'(t)$. Thus, the distribution of $(S_{\bar{z}(t-1)},S_{\bar{z}'(t-1)})$ form our principal strata. For example, if we want to estimate the effect for a negative wealth shock at $t=2$ vs no negative wealth shock by $t=2$, that is $E[Y_{01}-Y_{00}|S_0=1]$, we restrict to subjects who survive if they did not receive a negative wealth shock at $t=1$, that is, subjects with $S_0=1$ (Subjects 1-12 in Table 2). Note that the definition, $E[\overline{Y}_{\bar{z}(t)}-\overline{Y}_{\bar{z}'(t)}|S_{\bar{z}(t-1)}=S_{\bar{z}'(t-1)}=1]$, is different from the parameter MSM estimates which is $E[\overline{Y}_{\bar{z}(t)}-\overline{Y}_{\bar{z}'(t)}]$.

3.3 | Proposed method

Our proposed method estimates $E[\overline{Y}_{\bar{z}(t)}-\overline{Y}_{\bar{z}'(t)}|S_{\bar{z}(t-1)}=S_{\bar{z}'(t-1)}=1]$ by imputing the survival status of each subject at the current time t and then, together with the observed survival status, estimate which principal stratum a subject would belong to. To improve the robustness of this method for our data analysis, we then use a Bayesian additive regression tree or BART-modified PENCOMP approach to impute the counterfactual outcomes among the potentially surviving subjects to account for the bias due to time-dependent confounding. This approach is doubly robust and reduces the burden of model specification by the researcher.[19,20] Subsequently, the average difference in the exposure effect within the desired principal strata is calculated. Variance is estimated using Rubin's combine rule to account for the imputation uncertainty.[21] Detailed steps for our method are given below.

1. Generate a bootstrap sample b from the data by sampling the units with replacement.
2. Estimate the model $X_{z^{(b)}(1)}|z^{(b)}(1),w^{(b)}(1),v^{(b)}$. Use this model to compute the counterfactual of $X_{z^{(b)}(1)}$ for bootstrap sample b.
3. Estimate the distribution of $Z^{(b)}(1)\,|w^{(b)}(1),v^{(b)}$. Use this model to estimate the propensity of exposure $Z^{(b)}(1)=z^{(b)}(1)$ as $P^*_{z^{(b)}(1)}=Pr(Z^{(b)}(1)=z^{(b)}(1)|w^{(b)}(1),v^{(b)})$. Note that we did not perform a logit transformation to obtain $P^*_{z^{(b)}(1)}$ (See PENCOMP, Appendix S1 A.2, Steps 2 and 5). This is because by using PENCOMP modified

TABLE 2 Sample example of a censoring by death dataset until $t = 3$ where $Z_t = 1$ indicates a subject having experienced a negative wealth shock and $Z_t = 0$ indicates a subject have not experienced any negative wealth shock till time t

	V	$Z(1)$	Y_1	Y_0	S_1	S_0	$Z(2)$	Y_{00}	Y_{01}	Y_{11}	S_{00}	S_{01}	S_{11}	$Z(3)$	Y_{000}	Y_{001}	Y_{011}	Y_{111}
Subject 1	x	1	x	?	1	1	1	?	?	x	1	1	1	1	?	?	?	x
Subject 2	x	0	?	x	1	1	1	?	x	?	1	1	1	1	?	?	x	?
Subject 3	x	1	x	?	1	1	1	?	?	x	1	1	0	NA	?	?	?	NA
Subject 4	x	0	?	x	1	1	1	?	x	?	1	1	0	1	?	?	x	NA
Subject 5	x	0	?	x	1	1	0	x	?	?	1	0	1	0	x	?	NA	?
Subject 6	x	0	?	x	1	1	0	x	?	?	0	1	1	NA	NA	NA	?	?
Subject 7	x	0	?	x	1	1	0	x	?	?	0	1	1	NA	NA	NA	?	?
Subject 8	x	0	?	x	1	1	0	x	?	?	1	0	0	0	x	?	NA	NA
Subject 9	x	1	x	?	0	1	NA	?	?	NA	1	1	0	NA	?	?	?	NA
Subject 10	x	1	x	?	0	1	NA	?	?	NA	0	1	0	NA	NA	NA	?	NA
Subject 11	x	0	?	x	0	1	1	?	x	NA	0	1	0	1	NA	NA	x	NA
Subject 12	x	0	?	x	0	1	0	x	?	NA	0	1	0	NA	NA	NA	?	NA
Subject 13	x	1	x	?	1	0	1	NA	NA	x	0	0	1	1	NA	NA	NA	x
Subject 14	x	0	?	x	1	0	NA	NA	NA	?	0	0	1	NA	NA	NA	NA	?

with BART to predict the outcomes, the nonlinear effect of the propensity of exposure will be handled automatically.

4. Estimate the model $Y_{z^{(b)}(1)}|P^*_{z^{(b)}(1)}, z^{(b)}(1), X_{z^{(b)}(1)}, w^{(b)}(1), v^{(b)}$ using BART to take care of any linear or nonlinear main effects as well as linear or nonlinear interactions. PENCOMP (See Appendix S1 A.2) is constructed using a nonlinear spline specification on the propensity of exposure combined with possible linear interactions between the propensity of exposure and remaining covariates. This fits well with the type of estimation problem that BART was designed to solve. We then use the model produced by BART-modified PENCOMP to compute the counterfactual of $Y_{z^{(b)}(1)}$ for bootstrap sample b.
5. Estimate the distribution for $S_{z^{(b)}(1)}|z^{(b)}(1), Y_{z^{(b)}(1)}, X_{z^{(b)}(1)}, w^{(b)}(1), v^{(b)}$. Use this model to generate a survival status for the counterfactual of $S_{z^{(b)}(1)}$ taking into account the assumption of monotonicity in Equation (4), that is, if S_0 is observed and $S_0 = 0$ then $S_1 = 0$. Similarly, if S_1 is observed and $S_1 = 1$ then $S_0 = 1$.
6. Estimate the model $X_{\overline{z}^{(b)}(2)}|\overline{z}^{(b)}(2), Y_{z^{(b)}(1)}, X_{z^{(b)}(1)}, \overline{w}^{(b)}(2), v^{(b)}$. Use the respective models to impute the counterfactual of $X_{\overline{z}^{(b)}(2)}$, using any previously imputed values for the unobserved exposure profiles and restricting to the subjects that are observed and predicted to survive under the given exposure profile of interest at $t = 1$.
7. Estimate the distribution of $Z^{(b)}(2)|z^{(b)}(1), y^{(b)}(1), x^{(b)}(1), \overline{w}^{(b)}(2), v^{(b)}$. Use this model to estimate the propensity to be assigned exposure $Z^{(b)}(2) = z^{(b)}(2)$ as $P_{z^{(b)}(2)} = Pr(Z^{(b)}(2) = z^{(b)}(2)|X_{z^{(b)}(1)}, z^{(b)}(1), w^{(b)}(1), v^{(b)})$. The probability of exposure $\overline{z}^{(b)}(2)$ is denoted as $P^*_{z^{(b)}(2)} = P_{z^{(b)}(2)}P^*_{z^{(b)}(1)}$.
8. Estimate the model $Y_{\overline{z}^{(b)}(2)}|P^*_{z^{(b)}(2)}, \overline{z}^{(b)}(2), Y_{z^{(b)}(1)}, \overline{X}_{\overline{z}^{(b)}(2)}, \overline{w}^{(b)}(2), v^{(b)}$ again restricting to subjects that are observed and predicted to survive under the exposure profiles of interest at $t = 2$. Use the respective models to impute the counterfactual of $Y_{\overline{z}^{(b)}(2)}$.
9. Using a similar procedure for steps 5-8, with the restriction determined by $S_{\overline{z}^{(b)}(t-1)} = S_{\overline{z}'^{(b)}(t-1)} = 1$ for time t where at least one $z^{(b)}(t) \neq z'^{(b)}(t)$, extend the estimation until the desired time point $t = T$.
10. Repeat Steps 1-9 to obtain B bootstrap values for $\hat{\Delta}_{\overline{z}^{(b)}(t),\overline{z}'^{(b)}(t)} = E[Y_{\overline{z}^{(b)}(t)} - Y_{\overline{z}'^{(b)}(t)}|S_{\overline{z}^{(b)}(t)} = S_{\overline{z}'^{(b)}(t)} = 1]$ with associated pooled variance $Q_{\overline{z}^{(b)}(t),\overline{z}'^{(b)}(t)}$.
11. The estimate of $\Delta_{\overline{Z}(t),\overline{Z}'(t)} = E[Y_{\overline{Z}(t)} - Y_{\overline{Z}'(t)}|S_{\overline{Z}(t-1)} = S_{\overline{Z}'(t-1)} = 1]$ is then $\overline{\Delta}_{\overline{z}^{(b)}(t),\overline{z}'^{(b)}(t),B} = \sum_{b=1}^{B} \frac{\hat{\Delta}_{\overline{z}^{(b)}(t),\overline{z}'^{(b)}(t)}}{B}$, and the estimate of the variance of $\overline{\Delta}_{\overline{z}^{(b)}(t),\overline{z}'^{(b)}(t),B}$ is $T_B = \overline{Q}_{\overline{z}^{(b)}(t),\overline{z}'^{(b)}(t),B} + (1 + 1/B)D_{\overline{z}^{(b)}(t),\overline{z}'^{(b)}(t),B}$, where $\overline{Q}_{\overline{z}^{(b)}(t),\overline{z}'^{(b)}(t),B} = \sum_{b=1}^{B} \frac{Q_{\overline{z}^{(b)}(t),\overline{z}'^{(b)}(t)}}{B}$ and

$$D_{\bar{z}^{(b)}(t),\bar{z}'^{(b)}(t),B} = \sum_{b=1}^{B} \frac{(\hat{\Delta}_{\bar{z}^{(b)}(t),\bar{z}'^{(b)}(t)} - \overline{\Delta}_{\bar{z}^{(b)}(t),\bar{z}'^{(b)}(t),B})^2}{B-1} \text{ with } \frac{\hat{\Delta}_{\bar{z}^{(b)}(t),\bar{z}'^{(b)}(t)} - \overline{\Delta}_{\bar{z}^{(b)}(t),\bar{z}'^{(b)}(t),B}}{\sqrt{T_B}} \sim t_\nu,\ \nu = (B-1)\left(1 + \frac{\overline{Q}_{\bar{z}^{(b)}(t),\bar{z}'^{(b)}(t),B}}{D_{\bar{z}^{(b)}(t),\bar{z}'^{(b)}(t),B}(B+1)}\right)^2.$$

4 | SIMULATION

We conducted a simulation study to investigate how well our method would perform under three scenarios: (1) weak association between exposure and confounder as well as exposure and survival status; (2) strong association between exposure and confounder as well as exposure and survival status; and finally (3) strong association between exposure and confounder, exposure and survival status, and an interaction between exposure, confounder, and survival status. We compared our method with the naïve method and MSM (See Appendix S1 A.1 for details on how we implemented MSM). We expect all three methods to perform well in the first scenario because there is little to no confounding. For the second scenario, we expect MSM and our proposed method to perform well because there is no difference in the exposure effect between the principal strata, and other stratification groups. The naïve method should not perform well due to the strong association between exposure and confounder as well as exposure and survival status. Finally, for Scenario 3, we expect only our proposed method to perform well because an association between the exposure effect and principal strata, $S_{\overline{Z}(t-1)} = S_{\overline{Z}'(t-1)} = 1$, is induced by the stronger interaction effect between exposure, confounder, and survival status.

To focus our investigation on these scenarios, we did not include any dropouts or complicated nonlinear and multiway interaction effects in our simulation. Although these simulation setups would provide us with more realistic results, we feel that including dropouts or complicated nonlinear and multiway interaction effects in our simulation setup will divert attention away from the censoring by death issue we are trying to tackle. Hence, out estimation methods replaced BART with linear and logistic regression models; we reserve the use of BART for the application. Details of our simulation setup can be found in Appendix S1 A.3.

Using $E[\overline{Y}_{\bar{z}(t)} - \overline{Y}_{\bar{z}'(t)} | S_{\bar{z}(t-1)} = S_{\bar{z}'(t-1)} = 1]$ as the true average treatment effect, we measured performance using the empirical bias, root mean squared error (RMSE), 95% coverage, and the average 95% Confidence Interval (CI) length (AIL). 1000 simulations were used to estimate these quantities. Under each simulation, a simple random sample of 4000 (Appendix S1 A.4) or 8000 (sample size of eligible HRS data we used was 7106) subjects was drawn from the target population data. All methods were then implemented on the sampled data to obtain the effect estimates. We assumed correct model specification for MSM and our proposed method. Codes to implement our simulation scenarios can be found in Appendix S1.

4.1 | Results

Table 3 shows the results with the sample size set to 8000, approximately the sample size in our application. All three methods under Scenario 1 were relatively unbiased and achieved similar RMSE. MSM and our proposed method reported slightly greater than nominal coverage due to the wider AIL for time point 3 estimates. Under Scenario 2, the absolute bias and RMSE of the naïve method was always larger than MSM and our proposed method and coverage was often far below the nominal 95% value. Finally, under Scenario 3, the naïve method was clearly biased with poor RMSE and coverage. MSM performed slightly better compared to the naïve method but absolute bias clearly increased compared to Scenario 2. Coverage for some exposure effects were poor as well. Our proposed method remained unbiased, produced a lower RMSE compared to the other two methods, and reached nominal coverage under Scenario 3. Comparing these results with sample size 4000 (See Appendix S1 A.4 for results), we found that bias relationship for the three methods remained relatively the same when the sample size decreased, though reduction in sample size reduces the impact of bias and results in better (though still below normal) coverage.

5 | DETERMINING THE EFFECT OF A NEGATIVE WEALTH SHOCK ON COGNITIVE SCORE FOR HRS SUBJECTS

For our analysis, we used HRS data collected from 1996 to 2002 because consistent collection of a subject's cognitive ability measure began in 1996. We treated the variables collected in 1996 as the baseline for our analysis and excluded subjects

TABLE 3 Simulation results for sample size 8000

Scenario 1		Naïve				MSM				Proposed			
Parameter	**True value**	**Bias**	**RMSE**	**95% Coverage**	**AIL**	**Bias**	**RMSE**	**95% Coverage**	**AIL**	**Bias**	**RMSE**	**95% Coverage**	**AIL**
$\Delta_{1,0}$	−1.497	−0.001	0.023	94.2	0.087	0.0003	0.023	94.0	0.087	0.0003	0.023	96.2	0.100
$\Delta_{01,00}$	−1.499	−0.002	0.036	95.3	0.143	−0.001	0.036	94.8	0.143	−0.001	0.037	95.5	0.151
$\Delta_{11,00}$	−1.005	−0.002	0.034	94.8	0.134	−0.0007	0.034	95.1	0.134	−0.001	0.035	98.7	0.183
$\Delta_{11,01}$	0.493	0.001	0.034	94.6	0.134	0.002	0.034	94.4	0.134	0.002	0.035	98.6	0.184
$\Delta_{001,000}$	−1.502	0.005	0.057	95.0	0.222	0.005	0.057	99.0	0.289	0.005	0.058	95.4	0.235
$\Delta_{011,000}$	−1.008	0.004	0.051	94.5	0.201	0.004	0.051	98.4	0.260	0.004	0.052	97.1	0.246
$\Delta_{111,000}$	−0.504	0.005	0.051	95.1	0.201	0.007	0.052	98.3	0.261	0.006	0.053	99.7	0.369
$\Delta_{011,001}$	0.495	−0.002	0.051	94.6	0.200	−0.002	0.051	99.1	0.260	−0.002	0.052	97.8	0.247
$\Delta_{111,001}$	1.000	−0.001	0.052	94.2	0.201	0.0001	0.052	98.7	0.261	−0.0005	0.054	99.8	0.369
$\Delta_{111,011}$	0.502	0.003	0.046	93.8	0.177	0.004	0.047	98.4	0.229	0.004	0.048	99.8	0.308
Scenario 2		**Naïve**				**MSM**				**Proposed**			
Parameter	**True value**	**Bias**	**RMSE**	**95% Coverage**	**AIL**	**Bias**	**RMSE**	**95% Coverage**	**AIL**	**Bias**	**RMSE**	**95% Coverage**	**AIL**
$\Delta_{1,0}$	−3.367	−0.047	0.055	59.6	0.109	0.002	0.029	94.0	0.113	0.003	0.029	95.9	0.125
$\Delta_{01,00}$	−1.727	−0.036	0.045	73.4	0.105	−0.031	0.041	78.8	0.106	−0.001	0.026	96.1	0.113
$\Delta_{11,00}$	−1.199	−0.134	0.139	4.0	0.142	−0.018	0.041	92.4	0.144	−0.001	0.036	96.9	0.161
$\Delta_{11,01}$	0.528	−0.098	0.105	21.9	0.140	0.013	0.038	94.0	0.142	−0.001	0.036	97.4	0.158
$\Delta_{001,000}$	−1.727	−0.029	0.049	87.9	0.156	−0.024	0.047	93.2	0.170	0.0001	0.038	96.2	0.160
$\Delta_{011,000}$	−1.183	−0.066	0.075	54.3	0.141	−0.048	0.060	77.7	0.153	−0.001	0.036	97.7	0.156
$\Delta_{111,000}$	−1.169	−0.166	0.173	7.8	0.193	−0.040	0.065	90.4	0.215	−0.003	0.049	98.9	0.246
$\Delta_{011,001}$	0.544	−0.038	0.050	81.5	0.131	−0.025	0.042	92.1	0.142	−0.002	0.032	97.2	0.145
$\Delta_{111,001}$	0.558	−0.137	0.145	17.3	0.186	−0.017	0.050	96.7	0.208	−0.005	0.046	98.7	0.233
$\Delta_{111,011}$	0.013	−0.098	0.108	38.0	0.174	0.010	0.046	96.4	0.194	0.0004	0.044	98.2	0.210

(Continues)

TABLE 3 (Continued)

Scenario 3		Naïve				MSM				Proposed			
Parameter	**True value**	**Bias**	**RMSE**	**95% Coverage**	**AIL**	**Bias**	**RMSE**	**95% Coverage**	**AIL**	**Bias**	**RMSE**	**95% Coverage**	**AIL**
$\Delta_{1,0}$	−2.347	−0.123	0.126	1.5	0.113	0.002	0.029	94.5	0.113	0.003	0.029	95.7	0.124
$\Delta_{01,00}$	−2.559	−0.114	0.118	3.4	0.117	−0.060	0.067	49.3	0.113	−0.001	0.028	96.1	0.115
$\Delta_{11,00}$	−3.062	−0.230	0.234	0.1	0.164	−0.032	0.052	86.5	0.159	−0.004	0.040	97.3	0.183
$\Delta_{11,01}$	−0.502	−0.118	0.125	19.4	0.164	0.026	0.048	91.0	0.160	−0.003	0.040	98.2	0.184
$\Delta_{001,000}$	−2.820	−0.125	0.133	16.6	0.171	−0.063	0.076	78.3	0.192	−0.002	0.039	95.6	0.157
$\Delta_{011,000}$	−3.605	−0.143	0.147	2.2	0.140	−0.087	0.093	40.8	0.159	−0.007	0.032	96.7	0.142
$\Delta_{111,000}$	−4.032	−0.290	0.296	0.1	0.225	−0.081	0.099	81.5	0.265	−0.010	0.057	98.7	0.282
$\Delta_{011,001}$	−0.785	−0.018	0.044	93.4	0.159	−0.024	0.047	94.7	0.181	−0.005	0.037	97.0	0.161
$\Delta_{111,001}$	−1.217	−0.160	0.171	22.4	0.238	−0.013	0.062	97.1	0.278	−0.008	0.059	98.7	0.311
$\Delta_{111,011}$	−0.432	−0.142	0.152	26.4	0.216	0.011	0.056	97.9	0.255	−0.006	0.052	98.7	0.264

Abbreviations: AIL, average 95% Confidence Interval (CI) length; MSM, marginal structural models; RMSE, root mean square error.

who did not have longitudinal measurements for net worth in 1992, because we were unable to distinguish whether they have already experienced a negative wealth shock; subjects with zero or negative net worth at baseline, because we did not know if these subjects have lifelong asset poverty or experienced a negative wealth shock prior to study entry; and subjects who experienced a negative wealth shock and death between 1992 to 1996, because they were no longer at risk for a negative wealth shock or death. Finally 7106 participants (72.9%) were eligible for our analysis. Details and descriptive statistics for this dataset can be found in Appendix S1 A.5.

To determine whether a subject experienced a negative wealth shock from the previous follow-up period to the current follow-up period, we first obtained data from the module assessing net worth administered at every wave of HRS. Measured assets include housing value, net value of businesses, individual retirement accounts, checking/savings accounts, certificates of deposits and savings bonds, investment holdings, net value of vehicles, and the value of any other substantial assets. From this asset total, debts were subtracted, including home mortgages, other home equity loans, and unsecured debt values, like credit card balances, student loans, and medical debts. Missing values for wealth were imputed at the level of each asset or debt, using an unfolding bracket imputation method.[22] Wealth data were not imputed for those who do not participate in a given wave. Negative wealth shock was measured and then dichotomized (yes or no) for each time point. Loss of 75% or more of total wealth between two consecutive waves was used as the cut-point for negative wealth shock.[23] Subjects were considered at risk for negative wealth shock until they have experienced a negative wealth shock or reached age 65. With this definition of a negative wealth shock, about 7% of subjects in our analysis ever experienced a negative wealth shock.

The cognitive ability of a subject is assessed in HRS using the Telephone Interview for Cognitive Status. Unfortunately, the full HRS cognitive battery is not available for participants under 65. Hence, we used an abbreviated measure that included questions about episodic memory (Immediate Word recall [10 points] and Delayed Word recall [10 points]) and mental status (Serial 7's [5 points], backwards counting from 20 [2 points]).[24] All responses were combined to create a composite score ranging from 0 to 27, with a higher score indicating higher cognitive ability. We treated this measure as continuous and normally distributed.

5.1 | Analysis

We were interested in how a negative wealth shock would affect the cognitive ability of late middle-aged adults in the HRS during the 6 years of follow-up as well as how the duration of a negative wealth shock affects cognitive ability. We employed four different methods to make inference about this effect: (i) the naïve method, where all subjects who died under their observed negative wealth shock status were removed from analysis; (ii) the baseline adjusted method, where all subjects who died were removed from analysis but the mean cognitive score was adjusted using a model that included all baseline covariates; (iii) the MSM method, where negative wealth shock allocation, missingness, and censoring by death were accounted for by inverse probability weighting; and (iv) our proposed method including the PENCOMP modification described in Section 3.2. We assumed that depression was the time-varying covariate that depends on the negative wealth shock status ($\overline{X}_{\overline{Z}}$ in Section 2 and the rest of the time-varying covariates are: self-reported health status, whether subject was insured, labor force status of subject, income, level of alcohol consumption, current smoking status, and number of health conditions [$\overline{W}$ in Section 3]). We also assumed that the cognitive score is missing at random given the baseline variables, past negative wealth shock status, time-varying covariates, and cognitive score. For MSM, we accounted for this missingness by modeling the propensity of response while for our proposed method, we imputed the missing cognitive score by using the modified version of PENCOMP discussed in Section 3.2. All our models (baseline adjusted, MSM, and our proposed method) were specified using BART. For the naïve, baseline adjusted, and MSM method, we employed 1000 bootstrap samples to calculate the mean and the 95% CI. The 95% CI was determined by taking the 2.5 and 97.5 percentile. For our proposed method, we estimated the effect and accounted for our uncertainty using our algorithm described in Section 3.2. Codes to implement all data analysis can be found in Appendix S1.

5.2 | Results

Table 4 shows the adjusted effect estimate of a negative wealth shock on cognitive score depending on the duration of the shock for late middle aged adults in the original HRS cohort from 1996 to 2002. In general, the naïve and baseline adjusted

TABLE 4 Effect estimate of negative wealth shock on cognitive score for late middle aged adults in original Health Retirment Study cohort from 1996 to 2002

	Naïve		Baseline adjusted[a]		MSM[b]		Proposed[b]	
	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI
2 years vs no shock	−0.51	(−1.45, 0.35)	−0.51	(−1.37, 0.3)	−0.01	(−1.18, 1.07)	−0.13	(−0.83, 0.58)
4 years vs no shock	−0.69	(−1.45, 0.05)	−0.7	(−1.4, 0.03)	−0.31	(−1.23, 0.58)	0.18	(−0.73, 1.09)
6 years vs no shock	−1.95	(−2.62, −1.25)	−1.94	(−2.6, −1.26)	−0.12	(−1.12, 0.89)	−0.18	(−0.87, 0.51)
4 years vs 2 years	−0.18	(−1.33, 1.04)	−0.19	(−1.26, 0.94)	−0.3	(−1.78, 1.15)	0.31	(−0.58, 1.20)
6 years vs 2 years	−1.45	(−2.54, −0.38)	−1.43	(−2.46, −0.4)	−0.1	(−1.61, 1.36)	−0.03	(−0.83, 0.78)
6 years vs 4 years	−1.26	(−2.27, −0.2)	−1.24	(−2.2, −0.24)	0.19	(−1.11, 1.61)	−0.38	(−1.36, 0.61)

[a]Adjusted by gender, education category, race, cognitive score, BMI, self-reported health status, alcohol consumption, insurance status, depression status, income, labor force status, marital status, age, smoking status, diabetes status, heart condition, HBP status, psychological problem status, and stroke status at baseline.

[b] Adjusted by gender, education category, race, cognitive score, BMI, self-reported health status, alcohol consumption, insurance status, depression status, income, labor force status, marital status, age, smoking status, diabetes status, heart condition, HBP status, psychological problem status, and stroke status at baseline as well as time-varying self-reported health status, alcohol consumption, insurance status, income, labor force status, smoking status, number of health conditions, and depression.

method suggests that experiencing a negative wealth shock has a much larger negative effect on the cognitive score of subjects in our sample compared to the adjusted estimates reported by MSM and our proposed method. In addition, we observed that MSM produced slightly wider CIs compared to our proposed method. The naïve and baseline adjusted method produced very similar results suggesting low association between cognitive score and the baseline covariates. The effect for subjects who experienced a negative wealth shock within the first 2 years of follow-up vs no shock (6 years vs no shock), subjects who experienced a negative wealth shock within the first 2 years of follow-up vs subjects who experienced a negative wealth shock between the second and fourth year of follow-up (6 years vs 2 years), and subjects who experienced a negative wealth shock within the first 2 years of follow-up vs subjects who experienced a negative wealth shock between the fourth and sixth year of follow up (6 years vs no shock), were significantly larger than 0 under the naïve and baseline adjusted method. For MSM and our proposed method all effects were reported to be not significant. This suggests that there is a substantial amount of "confounding by exposure" in the HRS data, where health status at earlier times influences risk of wealth shock, but little effect of "censoring by death," with the potential outcomes among the survivors not differing greatly from the potential outcomes among those who would die under alternative exposure regimes once confounding by exposure is accounted for.

6 | DISCUSSION

In this paper, we were interested in how a negative wealth shock affects the cognitive ability of late middle-aged Americans participating in the HRS from 1996 to 2002. The main difficulty we faced was the presence of death in some subjects causing their cognitive score to be censored. Because it is very possible that subjects with lower cognitive ability and/or experience of a negative wealth shock would have a higher risk of death, accounting for censoring by death is required. This is because, even with randomization or its conditional equivalent, subjects who did not experience a wealth shock and were observed to have survived are a mix of those who would survive regardless of whether they experienced a negative wealth shock and those who would survive only if they did not experience a negative wealth shock. As a result, if we remove only the subjects we observed to die from our analysis, the effect of the negative wealth shock on cognitive ability that we measure would be confounded by death. Although MSM is commonly employed to weight the subjects who survived, this approach is arguably not sensible and would likely produce biased estimates when the effect depends on the principal strata as well as when adjustments on the weights have to be employed in order to stabilize the MSM estimate. To overcome these issues, we propose a new method to estimate the effect by imputing the counterfactual survival status of each subject in order to compare outcomes among individuals who would survive regardless of whether they experience a wealth shock. Our method remained unbiased for all the simulation scenarios we tried and produced

reasonable coverage. When applied to the HRS dataset, our method suggested that the effect of a negative wealth shock on the cognitive ability is close to null.

One shortcoming of our approach is our failure to incorporate the HRS sample design, in particular the sampling weights, in our inference. Given that a key use of weights in regression-type analysis is to reduce the effect of model misspecification,[25] we hope that our use of BART will minimize the degree of model misspecification. We leave the incorporation of such features in a general approach to future work. Another aspect of our method which could be improved is to allow our method to be applicable to studies where the follow-up time is not fixed. In such a situation, Cox-based survival models would have to be employed and time would have to be included as a covariate in the survival and outcome models. The difficulty in this extension would be how to develop a systematic way, applicable to all subjects, to determine the relation in time between exposure, measuring the outcome, and death.

Finally, we would like to point out an interesting question regarding the use of imputed potential outcomes as the covariates for predicting the potential outcomes at the next time point. At first glance, there may be the risk of bias or "aggregated" uncertainty at later time points. We argue that under most situations, when the imputed potential outcomes at the previous time point is within the "space" of the covariates in the potential outcomes prediction model, this risk would be negligible. However, if the imputed potential outcomes are based on covariates at the far edge of the covariate space, then there may be a risk that bias introduced at the current time point will aggregate to the later time points. We know of little work in this regards and believe it would constitute an interesting area for further research.

ACKNOWLEDGEMENTS

We would like to thank the associate editor for the constructive comments which have greatly improved this article.

DATA AVAILABILITY STATEMENT

The data that supports the findings of this study are available in Appendix S1 of this article.

ORCID


Yaoyuan Vincent Tan https://orcid.org/0000-0001-5950-9846

REFERENCES

1. Lee Y, Back JH, Kim J, et al. Systematic review of health behavioral risks and cognitive health in older adults. *Int Psychogeriatr*. 2010;22(2):174-187.
2. Plassman BL, Williams JW Jr, Burke JR, Holsinger T, Benjamin S. Systematic review: factors associated with risk for and possible prevention of cognitive decline in later life. *Ann Intern Med*. 2010;153(3):182-193.
3. Krieger N. Theories for social epidemiology in the 21st century: an eco-social perspective. *Int J Epidemiol*. 2001;30(4):668-677.
4. Butrica BA, Smith KE, Toder EJ. What the 2008 stock market crash means for retirement security. *J Aging Soc Policy*. 2010;22(4):339-359.
5. Shrira A, Palgi Y, Ben-Ezra M, Spalter T, Kavé G, Shmotkin D. For better and for worse: the relationship between future expectations and functioning in the second half of life. *J Gerontol B Psychol Sci Soc Sci*. 2011;66(2):195-203.
6. Sonnega A, Faul JD, Ofstedal MB, Langa KM, Phillips JW, Weir DR. Cohort profile: the health and retirement study (HRS). *Int J Epidemiol*. 2014;43(2):576-585.
7. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika*. 1983;70(1):41-55.
8. Imbens GW, Rubin DB. *Causal Inference for Statistical, Social, and Biomedical Sciences: An Introduction*. New York, NY: Cambridge University Press; 2015.
9. Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology*. 2000;11(5):550-560.
10. Zhou T, Elliott MR, Little RJA. Penalized spline of propensity methods for treatment comparison. *J Am Stat Assoc*. 2019;114(525):1-19.
11. Weuve J, Tchetgen Tchetgen EJ, Glymour MM, et al. Accounting for bias due to selective attrition: the example of smoking and cognitive decline. *Epidemiology*. 2012;23:119-128.
12. Chaix B, Evans D, Merlo J, Suzuki E. Commentary: weighing up the dead and missing: reflections on inverse probability weighting and principal stratification to address truncation by death. *Epidemiology*. 2012;23(1):129-131.
13. Neyman J. On the two different aspects of the representative method: the method of stratified sampling and the method of purposive selection. *J R Stat Soc*. 1934;97(4):558-625.
14. Rubin DB. Estimating causal effects of treatments in randomized and nonrandomized studies. *J Educ Psychol*. 1974;66(5):688-701.
15. Elliott MR, Joffe MM, Chen Z. A potential outcomes approach to developmental toxicity analyses. *Biometrics*. 2006;62(2):352-360.
16. Frangakis CE, Rubin DB. Principal stratification in causal inference. *Biometrics*. 2002;58(1):21-29.
17. Pool LR, Burgard SA, Needham BL, Elliott MR, Langa KM, Leon CFM. Association of a negative wealth shock with all-cause mortality in middle-aged and older adults in the united states. *J Am Med Assoc*. 2018;319(13):1341-1350.

18. Zhang G, Little RJA. Extensions of the penalized spline of propensity prediction method of imputation. *Biometrics*. 2009;65:911-918.
19. Chipman HA, George EI, McCulloch RE. BART: Bayesian additive regression trees. *Ann Appl Stat*. 2010;4(1):266-298.
20. Tan YV, Flannagan CAC, Elliott MR. "Robust-squared" imputation models using BART. *Journal of Survey Statistics and Methodology*. 2019;7(4):465-497.
21. Heitjan D, Little RJA. Multiple imputation for the fatal accident reporting system. *Appl Stat*. 1991;40:13-29.
22. Juster FT, Smith JP. Improving the quality of economic data: lessons from the HRS and AHEAD. *J Am Stat Assoc*. 1997;92(440):1268-1278.
23. Pool LR, Needham BL, Burgard SA, Elliott MR, Leon CFM. Negative wealth shock and short-term changes in depressive symptoms and medication adherence among late middle-aged adults. *J Epidemiol Community Health*. 2017;71:758-763.
24. Crimmins EM, Kim JK, Langa KM, Weir DR. Assessment of cognition using surveys and neuropsychological assessment: the health and retirement study and the aging, demographics and memory study. *J Gerontol Ser B Psychol Sci Soc Sci*. 2011;66(Suppl 1):162-171.
25. Korn EL, Graubard BI. Examples of differing weighted and unweighted estimates from a sample survey. *Am Stat*. 1995;49(3):291-295.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Tan YV, Flannagan CAC, Pool LR, Elliott MR. Accounting for selection bias due to death in estimating the effect of wealth shock on cognition for the Health and Retirement Study. *Statistics in Medicine*. 2021;40:2613–2625. https://doi.org/10.1002/sim.8921