Web Material

Marginal structural models for life-course theories and social epidemiology: Definitions, sources of bias, and simulated illustrations

Paola Gilsanz, Jessica G. Young, M. Maria Glymour, Eric J. Tchetgen Tchetgen, Chloe W. Eng, Karestan C. Koenen, and Laura D. Kubzansky

Web Appendices

- Web Appendix 1. Background: Data structure, identifying assumptions and MSM definitions
- Web Appendix 2. Data generating models
- Web Appendix 3. Bias and estimation
- Web Appendix 4. Simulation with time-varying confounding affected by past exposure and "collider bias"
- Web Appendix 5. SAS code

Web Tables

- Web Table 1. Measures of accuracy and precision of the estimated effects on the log odds scale under different theoretical life-course model when information regarding early-life exposure is missing
- Web Table 2. Measures of accuracy and precision of estimated effects on the log odds scale under different theoretical life-course models when information regarding early-life exposure is missing
- Web Table 3. Variance in estimated effects of depression on stroke on the log odds scale under an accumulation model comparing conventional regression and inverse probability weighting to estimate MSM parameters, in the presence of a late-study-start and time-varying confounders.

Web Appendix 1. Background: Data structure, identifying assumptions and MSM definitions

We consider a study where the following measurements are taken on each of n study subjects (assumed i.i.d): A_t is the subject's value of the exposure of interest (e.g. depression), L_t the value of a time-varying covariate (e.g. alcohol use disorder) occurring between A_{t-1} and A_t at each time $t = 0, 1, 2$ with $A_{-1} \equiv 0$ by convention. Let Y be a subject's value of the outcome of interest (e.g. stroke) at an end of follow-up time occurring after A_2. We denote the history of a random variable using overbars; for example, $\overline{A}_2 = (A_0, A_1, A_2)$. We make the simplifying assumptions that in this study Y is measured for all n subjects, there are no competing events, no loss to follow-up and no stroke events that occur prior to $t = 2$. We also make the simplifying assumption that L_0 is constant for all n subjects.

Let $Y^{\overline{a}_2}$ denote the counterfactual outcome for a subject had, contrary to fact, we intervened to set $\overline{A}_2$ to some $\overline{a}_2 = (a_0, a_1, a_2)$ and suppose the following identifying assumptions hold:

1. Consistency 1: If $\overline{A}_2 = \overline{a}_2$ then $Y = Y^{\overline{a}_2}$
2. Exchangeability 1: $Y^{\overline{a}_2} \coprod A_t | \overline{L}_t, \overline{A}_{t-1} = \overline{a}_{t-1}$, $t = 0, 1, 2$
3. Positivity 1: $f(\overline{a}_{t-1}, \overline{l}_t) \neq 0 \implies f(a_t | \overline{a}_{t-1}, \overline{l}_t) > 0$ w.p.1, $t = 0, 1, 2$, for $\overline{l}_t$ any possible realization of $\overline{L}_t$ in the study.

Exchangeability 1 holds under the causal diagram of Figure 2 in the main text by the absence of any unblocked backdoor paths between A_t and Y conditional on $\overline{A}_{t-1}$ and $\overline{L}_t$, $t = 0, 1, 2$ [1]. A stronger version of this assumption holds in Figure 1 of the main text where we can understand $\overline{L}_t = \varnothing$.

Remark on notation: In the main text, we used notation D_e, D_m, and D_l in place of A_0, A_1 and A_2, respectively; d_e, d_m and d_l in place of a_0, a_1 and a_2, respectively; and AUD_m and AUD_l in place of L_1 and L_2, respectively.

Following [2], under these identifying assumptions, we can write the counterfactual risk of stroke under intervention $\overline{a}_2$, or $\Pr[Y^{\overline{a}_2} = 1]$, as the g-formula characterized by $\overline{a}_2$, which in this simple setting can be written as the following function of the above observed variables:

$$h(a_2, a_1, a_0) = \sum_{l_2, l_1} \Pr[Y = 1 | A_2 = a_2, L_2 = l_2, A_1 = a_1, L_1 = l_1, A_0 = a_0] f(l_2 | a_1, l_1, a_0) f(l_1 | a_0) \quad (1)$$

where $f(l_2 | a_1, l_1, a_0) \equiv \Pr[L_2 = l_2 | A_1 = a_1, L_1 = l_1, A_0 = a_1]$ and $f(l_1 | a_0) = \Pr[L_1 = l_1 | A_0 = a_0]$.

As in the main text, the following is a marginal structural model assumption for $\Pr[Y^{\overline{a}_2} = 1]$

$$\Pr[Y^{\overline{a}_2} = 1] = \exp\{\psi_0 + \psi_1 a_2 + \psi_2 a_1 + \psi_3 a_0\} \tag{2}$$

which, given the identification results above, is equivalent to the following model assumption on the g-formula:

$$h(a_2, a_1, a_0) = \exp\{\psi_0 + \psi_1 a_2 + \psi_2 a_1 + \psi_3 a_0\} \tag{3}$$

We also consider a study identical to that above but with the complication that A_0 is not observed. In this case we might more naturally consider the alternative counterfactual risk $\Pr[Y^{a_2,a_1} = 1]$; i.e. the risk of stroke under an intervention where A_2 is set to a_2 and A_1 is set to a_1 (and no intervention is made on A_0). Consider the following alternative versions of consistency, exchangeability and positivity:

1. Consistency 2: If $(A_1, A_2) = (a_1, a_2)$ then $Y = Y^{a_2,a_1}$
2. Exchangeability 2: $Y^{a_2,a_1} \coprod A_1 | L_1$ and $Y^{a_2,a_1} \coprod A_2 | L_2, A_1 = a_1, L_1$
3. Positivity 2: $f(l_2, a_1, l_1) \neq 0 \implies f(a_2|l_2, a_1, l_1) > 0$ w.p.1 and $f(l_1) \neq 0 \implies f(a_1|, l_1) > 0$ w.p.1

Provided these assumptions hold, again, following Robins (1986), we can write the counterfactual risk $\Pr[Y^{a_2,a_1} = 1]$ as the following function of measured variables which does not depend on the unobserved A_0:

$$m(a_2, a_1) = \sum_{l_2, l_1} \Pr[Y = 1 | A_2 = a_2, L_2 = l_2, A_1 = a_1, L_1 = l_1] f(l_2|a_1, l_1) f(l_1) \tag{4}$$

We might analogously impose a marginal structural model on $\Pr[Y^{a_2,a_1} = 1]$ as a function of (a_2, a_1)

$$\Pr[Y^{a_2,a_1} = 1] = \exp\{\eta_0 + \eta_1 a_2 + \eta_2 a_1\} \tag{5}$$

which given our alternative identifying assumptions is equivalent to the observed data model assumption

$$m(a_2, a_1) = \exp\{\eta_0 + \eta_1 a_2 + \eta_2 a_1\} \tag{6}$$

Importantly, the alternative version of our identifying assumptions above equating $\Pr[Y^{a_2,a_1} = 1]$ to the observed data function $m(a_2, a_1)$, which ignores A_0, do not hold under either Figure 1 or Figure 2 by the fact that A_0 is a common cause of both Y and future exposure in these Figures. By this we say A_0 is an unmeasured confounder for effects of joint interventions on A_1 and A_2 on the outcome Y. Thus, even if the model (6) is correctly specified, the parameters η would not encode causal effects in this case – i.e. the models (5) and (6) would not simultaneously hold for some η.

However, we now show that, given both the MSMs (2) and (5) are correct and under certain additional conditions, we will have that ψ_j of the MSM (2) will equal η_j of the MSM (5), $j = 0, 1, 2$. By this, under conditions where we can compute a consistent estimator of η_j, it is also a consistent estimator of ψ_j.

Theorem 1: Suppose the following are true: (i) as in Figure 2, Exchageability 1 holds for $t = 0$; (ii) the following Consistency assumption holds

- Consistency 3: If $A_0 = a_0$ then $Y^{a_2,a_1} = Y^{a_2,a_1,a_0}$

and (iii) both the MSMs (2) and (5) are correctly specified. Then ψ_j of the MSM (2) equals η_j of the MSM (5), $j = 0, 1, 2$ Note Consistency 3 can be equivalently stated as follows: For an individual who is observed to have $A_0 = a_0$, we can equate that individual's outcome that would be observed in a world where there is no intervention on A_0 and only his/her values of A_1 and A_2 are intervened on and set to a_1 and a_2, respectively, to the outcome we would have observed for that individual had we also intervened to set his/her A_0 to a_0.

Proof: By probability laws, Consistency 3 and Exchangeability 1 we have

$$\begin{aligned}
\Pr[Y^{a_2,a_1} = 1] &= \sum_{a_0} \Pr[Y^{a_2,a_1} = 1 | A_0 = a_0] f(a_0) \\
&= \sum_{a_0} \Pr[Y^{a_2,a_1,a_0} = 1 | A_0 = a_0] f(a_0) \\
&= \sum_{a_0} \Pr[Y^{a_2,a_1,a_0} = 1] f(a_0) \qquad (7)
\end{aligned}$$

By the MSMs (2) and (5), along with (7) we have

$$\begin{aligned}
\exp\{\eta_0\} &= \Pr[Y^{a_2=0,a_1=0} = 1] \\
&= \sum_{a_0} \Pr[Y^{a_2=0,a_1=0,a_0} = 1] f(a_0) \\
&= \sum_{a_0} \exp\{\psi_0\} f(a_0) \\
&= \exp\{\psi_0\} \sum_{a_0} f(a_0) \\
&= \exp\{\psi_0\}
\end{aligned}$$

Also

$$\begin{aligned}
\exp\{\eta_1\} &= \frac{\Pr[Y^{a_2=1,a_1=0}=1]}{\Pr[Y^{a_2=0,a_1=0}=1]} \\
&= \frac{\sum_{a_0}\Pr[Y^{a_2=1,a_1=0,a_0}=1]f(a_0)}{\sum_{a_0}\Pr[Y^{a_2=0,a_1=0,a_0}=1]f(a_0)} \\
&= \frac{\sum_{a_0}\exp\{\psi_0+\psi_1\}f(a_0)}{\sum_{a_0}\exp\{\psi_0\}f(a_0)} \\
&= \exp\{\psi_1\}\sum_{a_0} f(a_0) \\
&= \exp\{\psi_1\}
\end{aligned}$$

and

$$\begin{aligned}
\exp\{\eta_2\} &= \frac{\Pr[Y^{a_2=0,a_1=1}=1]}{\Pr[Y^{a_2=0,a_1=0}=1]} \\
&= \frac{\sum_{a_0}\Pr[Y^{a_2=0,a_1=1,a_0}=1]f(a_0)}{\sum_{a_0}\Pr[Y^{a_2=0,a_1=0,a_0}=1]f(a_0)} \\
&= \frac{\sum_{a_0}\exp\{\psi_0+\psi_2\}f(a_0)}{\sum_{a_0}\exp\{\psi_0\}f(a_0)} \\
&= \exp\{\psi_2\}\sum_{a_0} f(a_0) \\
&= \exp\{\psi_2\}
\end{aligned}$$

End Proof.

Web Appendix 2. Data generating models

2.1 Simulation with time-varying confounding

The following parametric models were used to generate 10,000 datasets of $n = 100,000$ measurements of $(A_0, L_1, A_1, L_2, A_2, Y)$ such that time-varying confounding affected by past exposure is present:

- A_0: constant logistic regression model:

$$\text{logit}\{\Pr[A_0 = 1]\} = \alpha_0 \tag{8}$$

- L_1: logistic regression model, possibly depending on A_0:

$$\text{logit}\{\Pr[L_1 = 1|A_0]\} = \beta_0 + \beta_1 A_0 \tag{9}$$

- A_1: logistic regression model, possibly depending on A_0, L_1:

$$\text{logit}\{\Pr[A_1 = 1|L_1, A_0]\} = \alpha_0 + \alpha_1 L_1 + \alpha_2 A_0 \tag{10}$$

- L_2: logistic regression model, possibly depending on A_1:

$$\text{logit}\{\Pr[L_2 = 1|A_1, L_1, A_0]\} = \beta_0 + \beta_1 A_1 \tag{11}$$

- A_2: logistic regression model, possibly depending on A_1, L_2:

$$\text{logit}\{\Pr[A_2 = 1|L_2, A_1, L_1, A_0]\} = \alpha_0 + \alpha_1 L_2 + \alpha_2 A_1 \tag{12}$$

- Y: logistic regression model, possibly depending on A_0, L_1, A_1, L_2, A_2:

$$\text{logit}\{\Pr[Y = 1|A_2, L_2, A_1, L_1, A_0]\} = \theta_0 + \theta_1 A_2 + \theta_2 L_2 + \theta_3 A_1 + \theta_4 L_1 + \theta_5 A_0 \tag{13}$$

SAS code to implement this simulation along with estimators described in Sections 3.1 and 3.3 can be found in the Web Appendix 5.

2.2 Simulation without time-varying confounding

The following parametric models were used to generate 10,000 datasets of $n = 100,000$ measurements of (A_0, A_1, A_2, Y):

- A_0: constant logistic regression model:

$$\text{logit}\{\Pr[A_0 = 1]\} = \alpha_0 \tag{14}$$

- A_1: logistic regression model, possibly depending on A_0:

$$\text{logit}\{\Pr[A_1 = 1|A_0]\} = \alpha_0 + \alpha_1 A_0 \tag{15}$$

- A_2: logistic regression model, possibly depending on A_1, A_0:

$$\text{logit}\{\Pr[A_2 = 1|A_1, A_0]\} = \alpha_0 + \alpha_1 A_1 + \alpha_2 A_0 \tag{16}$$

- Y: logistic regression model, possibly depending on A_0, A_1, A_2:

$$\text{logit}\{\Pr[Y = 1|A_2, A_1, A_0]\} = \lambda_0 + \lambda_1 A_2 + \lambda_2 A_1 + \lambda_3 A_0 \tag{17}$$

SAS code for simulating data and implementing estimators described in Section 3.2 can be found in Web Appendix 5. Data generating models and parameters differed in this simulation from those of Section 2.1 by two key differences. These include the absence of any dependence of Y or A_t at any time on $\overline{L}_t$. In addition, we allow the complication that A_2 depends on A_0 conditional on A_1 in this case such that A_0 is an unmeasured confounder for the effect of A_2 on Y – by selecting $\alpha_2 \neq 0$ in the model (16). This complication was absent in Section 2.1. There, A_0 was not an unmeasured confounder for this effect as we did not allow A_2 to depend on A_0 in the model (12). The fact that we allow $\alpha_2 \neq 0$ in the model (16) here has implications for model misspecification in implementing IPW estimators when A_0 is not measured which we describe below.

Web Appendix 3. Bias and estimation

3.1 Bias due to inappropriate adjustment for time-varying confounding (no late study start/all variables observed)

In each of the $10,000$ data sets generated from models in Section 2.1, we applied two estimators of the MSM coefficient vector (ψ_1, ψ_2, ψ_3) in model (3): an inverse probability weighted (IPW) estimator under the model assumption (3) and a maximum likelihood estimator (MLE) of $(\theta_1, \theta_3, \theta_5)$ in the logistic regression model (43). IPW estimates in each of the datasets were constructed by fitting a logistic regression model with dependent variable Y and independent variables A_2, A_1, A_0 weighted by subject specific stabilized weights with denominator the product of the MLE of $f(A_t|\overline{A}_{t-1}, \overline{L}_t)$ over $t = 1, 2$ under the known data generating logistic regression models (10) and (12) and numerator a product of the MLEs of $f(A_t|\overline{A}_{t-1})$ over $t = 1, 2$ based on corresponding logistic regression models [3].

Bias is defined as the mean of an estimator of a given true parameter minus the true parameter value. In this case, the parameter(s) of interest are causal effects as encoded by contrasts in $\Pr[Y^{\overline{a}_2} = 1]$ for differing values of $\overline{a}_2$. As above, given our nonparametric identifying assumptions (consistency, exchangeability, positivity), our parameter(s) of interest are in turn encoded by contrasts in the g-formula $h(\overline{a}_2)$ for differing values of $\overline{a}_2$. By defining bias with respect to the true values of the coefficients ψ in the model (3) we are inherently assuming that this model is correct. For example, we would like to interpret ψ_2 as the causal effect of A_1 on Y had we intervened to set $A_2 = a_2$ and $A_0 = a_0$ for all subjects. However, this is not the correct interpretation (even if our causal identifying assumptions hold) if this effect is different depending on the value of a_2 and/or a_0. In this case, the model (3) should include interaction terms. Further, even if (3) is correctly specified under our data generating models, it is not immediately clear from these models what the corresponding true value of ψ is as it is by definition a complex function of the parameters of data generating model coefficients by the definition of $h(\overline{a}_2)$ in (1). Without knowledge of this true value, bias cannot be computed.

We now prove that the model (3) is correctly specified under the data generating models of Section 2.1 and give formulas for the true values of these coefficients as functions of the data generating coefficients (β, θ) under the assumption of a rare outcome (which may be accomplished by selecting θ_0 large negative in all scenarios). These formulas were used to compute the true values of ψ and corresponding bias in Table 1 and 3 of the main text. For further discussion, see [4].

Theorem. 2: Given our data generating models (8) through (43) and a rare outcome, the MSM (3) is correctly specified with true ψ defined as

$$\psi_0 = \log\left[\exp(\theta_0)\left\{\frac{\exp(\theta_2+\beta_0)}{1+\exp(\beta_0)}+\frac{1}{1+\exp(\beta_0)}\right\}\left\{\frac{\exp(\theta_4+\beta_0)}{1+\exp(\beta_0)}+\frac{1}{1+\exp(\beta_0)}\right\}\right]$$

$$\psi_1 = \theta_1$$

$$\psi_2 = \log\left[\exp(\theta_3)\frac{\left\{\frac{\exp(\theta_2+\beta_0+\beta_1)}{1+\exp(\beta_0+\beta_1)}+\frac{1}{1+\exp(\beta_0+\beta_1)}\right\}}{\left\{\frac{\exp(\theta_2+\beta_0)}{1+\exp(\beta_0)}+\frac{1}{1+\exp(\beta_0)}\right\}}\right] \tag{18}$$

$$\psi_3 = \log\left[\exp(\theta_5)\frac{\left\{\frac{\exp(\theta_4+\beta_0+\beta_1)}{1+\exp(\beta_0+\beta_1)}+\frac{1}{1+\exp(\beta_0+\beta_1)}\right\}}{\left\{\frac{\exp(\theta_4+\beta_0)}{1+\exp(\beta_0)}+\frac{1}{1+\exp(\beta_0)}\right\}}\right]$$

Proof: By definition, we can write:

$$h(\overline{a}_2) = \exp\{\psi_0+\psi_1 a_2+\psi_2 a_1+\psi_3 a_0+\psi_4 a_2 a_1+\psi_5 a_2 a_0+\psi_6 a_1 a_0+\psi_7 a_2 a_1 a_0\} \tag{19}$$

Further, given our data generating model (43) and rare outcome we have

$$\Pr[Y=1|A_2,L_2,A_1,L_1,A_0] \approx \exp\{\theta_0+\theta_1 A_2+\theta_2 L_2+\theta_3 A_1+\theta_4 L_1+\theta_5 A_0\} \tag{20}$$

By (20), our other data generating assumptions and (1) we also have

$$h(\overline{a}_2) = h(a_2,a_1,a_0) = \exp\{\theta_0+\theta_1 a_2+\theta_3 a_1+\theta_5 a_0\}h^*(a_1,a_0) \tag{21}$$

where

$$h^*(a_1,a_0) = \sum_{l_2,l_1}\exp\{\theta_2 l_2+\theta_4 l_1\}f(l_2|a_1;\beta)f(l_1|a_0;\beta) \tag{22}$$

with $f(l_2|a_1;\beta)$ and $f(l_1|a_0;\beta)$ defined as in models (11) and (9), respectively.

By (19) and (21) we then have:

- $\exp\{\psi_0\} = h(0,0,0) = \exp\{\theta_0\}h^*(0,0)$ where

$$\begin{aligned} h^*(0,0) &= \frac{1+\exp\{\theta_2+\beta_0\}+\exp\{\theta_4+\beta_0\}+\exp\{\theta_2+\theta_4+2\beta_0\}}{(1+\exp\{\beta_0\})^2} \\ &= \left(\frac{1+\exp\{\theta_2+\beta_0\}}{1+\exp\{\beta_0\}}\right)\left(\frac{1+\exp\{\theta_4+\beta_0\}}{1+\exp\{\beta_0\}}\right) \end{aligned}$$

- $\exp\{\psi_1\} = \frac{\exp\{\psi_0+\psi_1\}}{\exp\{\psi_0\}} = \frac{h(1,0,0)}{\exp\{\psi_0\}} = \frac{\exp\{\theta_0+\theta_1\}h^*(0,0)}{\exp\{\theta_0\}h^*(0,0)} = \exp\{\theta_1\}$
- $\exp\{\psi_2\} = \frac{\exp\{\psi_0+\psi_2\}}{\exp\{\psi_0\}} = \frac{h(0,1,0)}{\exp\{\psi_0\}} = \frac{\exp\{\theta_0+\theta_3\}h^*(1,0)}{\exp\{\theta_0\}h^*(0,0)} = \exp\{\theta_3\}\frac{h^*(1,0)}{h^*(0,0)}$ where

$$\begin{aligned} h^*(1,0) &= \frac{1+\exp\{\theta_2+\beta_0+\beta_1\}+\exp\{\theta_4+\beta_0\}+\exp\{\theta_2+\theta_4+2\beta_0+\beta_1\}}{(1+\exp\{\beta_0+\beta_1\})(1+\exp\{\beta_0\})} \\ &= \left(\frac{1+\exp\{\theta_2+\beta_0+\beta_1\}}{1+\exp\{\beta_0+\beta_1\}}\right)\left(\frac{1+\exp\{\theta_4+\beta_0\}}{1+\exp\{\beta_0\}}\right) \end{aligned}$$

  In turn

$$\frac{h^*(1,0)}{h^*(0,0)} = \frac{\left\{\frac{\exp(\theta_2+\beta_0+\beta_1)}{1+\exp(\beta_0+\beta_1)}+\frac{1}{1+\exp(\beta_0+\beta_1)}\right\}}{\left\{\frac{\exp(\theta_2+\beta_0)}{1+\exp(\beta_0)}+\frac{1}{1+\exp(\beta_0)}\right\}}$$

- $\exp\{\psi_3\} = \frac{\exp\{\psi_0+\psi_3\}}{\exp\{\psi_0\}} = \frac{h(0,0,1)}{\exp\{\psi_0\}} = \frac{\exp\{\theta_0+\theta_5\}h^*(0,1)}{\exp\{\theta_0\}h^*(0,0)} = \exp\{\theta_5\}\frac{h^*(0,1)}{h^*(0,0)}$ where

$$\begin{aligned} h^*(0,1) &= \frac{1+\exp\{\theta_2+\beta_0\}+\exp\{\theta_4+\beta_0+\beta_1\}+\exp\{\theta_2+\theta_4+2\beta_0+\beta_1\}}{(1+\exp\{\beta_0\})(1+\exp\{\beta_0+\beta_1\})} \\ &= \left(\frac{1+\exp\{\theta_2+\beta_0\}}{1+\exp\{\beta_0\}}\right)\left(\frac{1+\exp\{\theta_4+\beta_0+\beta_1\}}{1+\exp\{\beta_0+\beta_1\}}\right) \end{aligned}$$

  In turn

$$\frac{h^*(0,1)}{h^*(0,0)} = \frac{\left\{\frac{\exp(\theta_4+\beta_0+\beta_1)}{1+\exp(\beta_0+\beta_1)}+\frac{1}{1+\exp(\beta_0+\beta_1)}\right\}}{\left\{\frac{\exp(\theta_4+\beta_0)}{1+\exp(\beta_0)}+\frac{1}{1+\exp(\beta_0)}\right\}}$$

- $\exp\{\psi_4\} = \frac{\exp\{\psi_0+\psi_1+\psi_2+\psi_4\}}{\exp\{\psi_0+\psi_1+\psi_2\}} = \frac{h(1,1,0)}{\exp\{\psi_0+\psi_1+\psi_2\}} = \frac{\exp\{\theta_0+\theta_1+\theta_3\}h^*(1,0)h^*(0,0)}{\exp\{\theta_0\}h^*(0,0)\exp\{\theta_1+\theta_3\}h^*(1,0)} = 1$
- $\exp\{\psi_5\} = \frac{\exp\{\psi_0+\psi_1+\psi_3+\psi_5\}}{\exp\{\psi_0+\psi_1+\psi_3\}} = \frac{h(1,0,1)}{\exp\{\psi_0+\psi_1+\psi_3\}} = \frac{\exp\{\theta_0+\theta_1+\theta_5\}h^*(0,1)h^*(0,0)}{\exp\{\theta_0\}h^*(0,0)\exp\{\theta_1+\theta_5\}h^*(0,1)} = 1$
- $\exp\{\psi_6\} = \frac{\exp\{\psi_0+\psi_2+\psi_3+\psi_6\}}{\exp\{\psi_0+\psi_2+\psi_3\}} = \frac{h(0,1,1)}{\exp\{\psi_0+\psi_2+\psi_3\}} = \frac{\exp\{\theta_0+\theta_3+\theta_5\}h^*(1,1)h^*(0,0)h^*(0,0)}{\exp\{\theta_0\}h^*(0,0)\exp\{\theta_3+\theta_5\}h^*(0,1)h^*(1,0)} = \frac{h^*(1,1)h^*(0,0)}{h^*(0,1)h^*(1,0)}$ where

$$\begin{aligned} h^*(1,1) &= \frac{1+\exp\{\theta_2+\beta_0+\beta_1\}+\exp\{\theta_4+\beta_0+\beta_1\}+\exp\{\theta_2+\theta_4+2\beta_0+2\beta_1\}}{(1+\exp\{\beta_0+\beta_1\})(1+\exp\{\beta_0+\beta_1\})} \\ &= \left(\frac{1+\exp\{\theta_2+\beta_0+\beta_1\}}{1+\exp\{\beta_0+\beta_1\}}\right)\left(\frac{1+\exp\{\theta_4+\beta_0+\beta_1\}}{1+\exp\{\beta_0+\beta_1\}}\right) \end{aligned}$$

  By this and definitions of $h^*(1,0)$, $h^*(0,1)$ and $h^*(0,0)$ above, we have $\frac{h^*(1,1)h^*(0,0)}{h^*(0,1)h^*(1,0)} = 1$.

•

$$\begin{aligned}\exp\{\psi_7\} &= \frac{\exp\{\psi_0+\psi_1+\psi_2+\psi_3+\psi_4+\psi_5+\psi_6+\psi_7\}}{\exp\{\psi_0+\psi_1+\psi_2+\psi_3+\psi_4+\psi_5+\psi_6\}} \\ &= \frac{h(1,1,1)}{\exp\{\psi_0+\psi_1+\psi_2+\psi_3+\psi_4+\psi_5+\psi_6\}} \\ &= \frac{\exp\{\theta_0+\theta_1+\theta_3+\theta_5\}h^*(1,1)h^*(0,0)h^*(0,0)h^*(1,0)h^*(0,1)}{\exp\{\theta_0\}h^*(0,0)\exp\{\theta_1+\theta_3+\theta_5\}h^*(1,0)h^*(0,1)h^*(1,1)h^*(0,0)} \\ &= 1\end{aligned}$$

End proof.

Note the following special case of Theorem 2: When we set $\theta_2 = \theta_4 = 0$ (i.e. there is no confounding by L_1 or L_2) or $\beta_1 = 0$ (confounders are not affected by past exposure), then $\psi_2 = \theta_3$ and $\psi_3 = \theta_5$. Thus, Theorem 2 clarifies that a standard MLE regression estimator of θ in the model (43) will recover ψ only when there is no time-varying confounding affected by past exposure. By contrast, an IPW estimator under the MSM (3) will recover ψ in this case as it recovers $h(\overline{a}_2)$ when (3) is correctly specified [5].

3.2 Bias due to late study start (A_0 unobserved) in data without time-varying confounding

Note that, when there is no time-varying confounding ($\overline{L}_t = \varnothing$), the g-formula $h(a_2, a_1, a_0)$ in (1) reduces to $\Pr[Y = 1|A_2, A_1, A_0]$. It immediately follows that, under the data generating model (17) of Section 2.2 and a rare disease assumption:

$$\Pr[Y = 1|A_2, A_1, A_0] \approx \exp\{\lambda_0 + \lambda_1 A_2 + \lambda_2 A_1 + \lambda_3 A_0\} \tag{23}$$

the model (3) holds with $\lambda_j \approx \psi_j$, $j = 0, \ldots, 3$.

To each of the 10,000 data sets generated under the models in Section 2.2, we calculated the MLE of δ in the model

$$\text{logit}\{\Pr[Y = 1|A_2, A_1]\} = \delta_0 + \delta_1 A_2 + \delta_2 A_1 \tag{24}$$

which ignores A_0. We now show that this model is misspecified under these data generating models even when (23) holds.

Theorem 3. Given our data generating models (14) through (17) as well as rare disease (23), we have approximately

$$\Pr[Y = 1|A_2 = a_2, A_1 = a_1] = \exp\{\delta_0 + \delta_1 a_2 + \delta_2 a_1 + \delta_3 a_1 a_2\} \tag{25}$$

with

$$\delta_0 = \log\left[\exp(\lambda_0)\left(\frac{\exp\{\lambda_3+\alpha_0\}c_1+c_2}{\exp\{\alpha_0\}c_1+c_2}\right)\right]$$

$$\delta_1 = \log\left[\exp(\lambda_1)\left(\frac{\exp\{\lambda_3+\alpha_0+\alpha_2\}c_1+c_2}{\exp\{\alpha_0+\alpha_2\}c_1+c_2}\right)\left(\frac{\exp\{\alpha_0\}c_1+c_2}{\exp\{\lambda_3+\alpha_0\}c_1+c_2}\right)\right]$$

$$\delta_2 = \log\left[\exp(\lambda_2)\left(\frac{\exp\{\lambda_3+\alpha_0+\alpha_1\}c_3+c_4}{\exp\{\alpha_0+\alpha_1\}c_3+c_4}\right)\left(\frac{\exp\{\alpha_0\}c_1+c_2}{\exp\{\lambda_3+\alpha_0\}c_1+c_2}\right)\right]$$

$$\delta_3 = \log\left[\frac{\left(\frac{\exp\{\lambda_3+\alpha_0+\alpha_1+\alpha_2\}c_3+c_4}{\exp\{\alpha_0+\alpha_1+\alpha_2\}c_3+c_4}\right)\left(\frac{\exp\{\lambda_3+\alpha_0\}c_1+c_2}{\exp\{\alpha_0\}c_1+c_2}\right)}{\left(\frac{\exp\{\lambda_3+\alpha_0+\alpha_2\}c_1+c_2}{\exp\{\alpha_0+\alpha_2\}c_1+c_2}\right)\left(\frac{\exp\{\lambda_3+\alpha_0+\alpha_1\}c_3+c_4}{\exp\{\alpha_0+\alpha_1\}c_3+c_4}\right)}\right]$$

and

$$c_1 = (1+\exp\{\alpha_0\})^2$$

$$c_2 = (1+\exp\{\alpha_0+\alpha_2\})(1+\exp\{\alpha_0+\alpha_1\})$$

$$c_3 = (1+\exp\{\alpha_0\})(1+\exp\{\alpha_0+\alpha_1\})$$

$$c_4 = (1+\exp\{\alpha_0+\alpha_1+\alpha_2\})(1+\exp\{\alpha_0+\alpha_1\})$$

Proof: Let $q(a_2,a_1) \equiv \Pr[Y=1|A_2=a_2,A_1=a_1]$. By definition, we can write:

$$q(a_2,a_1) = \exp\{\delta_0+\delta_1 a_2+\delta_2 a_1+\delta_3 a_1 a_2\} \tag{26}$$

Also by definition

$$q(a_2,a_1) = \sum_{a_0}\Pr[Y=1|A_2=a_2,A_1=a_1,A_0=a_0]\frac{f(a_2|a_1,a_0)f(a_1|a_0)f(a_0)}{\sum_{a_0}f(a_2|a_1,a_0)f(a_1|a_0)f(a_0)} \tag{27}$$

which, by our data generating assumptions of this section and (23), gives approximately

$$q(a_2,a_1) = \exp\{\lambda_0+\lambda_1 a_2+\lambda_2 a_1\}q^*(a_2,a_1) \tag{28}$$

where

$$q^*(a_2,a_1) = \frac{\sum_{a_0}\exp\{\lambda_3 a_0\}f(a_2|a_1,a_0;\alpha)f(a_1|a_0;\alpha)f(a_0;\alpha)}{\sum_{a_0}f(a_2|a_1,a_0;\alpha)f(a_1|a_0;\alpha)f(a_0;\alpha)} \tag{29}$$

with $f(a_2|a_1,a_0;\alpha)$, $f(a_1|a_0;\alpha)$ and $f(a_0;\alpha)$ defined as in models (16), (15) and (14), respectively.

By (26) and (28) we then have:

- $\exp\{\delta_0\} = q(0,0) = \exp\{\lambda_0\}q^*(0,0)$
- $\exp\{\delta_1\} = \frac{\exp\{\delta_0+\delta_1\}}{\exp\{\delta_0\}} = \frac{q(1,0)}{\exp\{\delta_0\}} = \frac{\exp\{\lambda_0+\lambda_1\}q^*(1,0)}{\exp\{\lambda_0\}q^*(0,0)} = \exp\{\lambda_1\}\frac{q^*(1,0)}{q^*(0,0)}$

- $\exp\{\delta_2\} = \frac{\exp\{\delta_0+\delta_2\}}{\exp\{\delta_0\}} = \frac{q(0,1)}{\exp\{\delta_0\}} = \frac{\exp\{\lambda_0+\lambda_2\}q^*(0,1)}{\exp\{\lambda_0\}q^*(0,0)} = \exp\{\lambda_2\}\frac{q^*(0,1)}{q^*(0,0)}$
- $\exp\{\delta_3\} = \frac{\exp\{\delta_0+\delta_1+\delta_2+\delta_3\}}{\exp\{\delta_0+\delta_1+\delta_2\}} = \frac{q(1,1)}{\exp\{\delta_0+\delta_1+\delta_2\}} = \frac{\exp\{\lambda_0+\lambda_1+\lambda_2\}q^*(1,1)q^*(0,0)q^*(0,0)}{\exp\{\lambda_0+\lambda_1+\lambda_2\}q^*(0,0)q^*(1,0)q^*(0,1)} = \frac{q^*(1,1)q^*(0,0)}{q^*(1,0)q^*(0,1)}$

Our results follow by $q^*(0,0) = \frac{\exp\{\lambda_3+\alpha_0\}c_1+c_2}{\exp\{\alpha_0\}c_1+c_2}$, $q^*(1,0) = \frac{\exp\{\lambda_3+\alpha_0+\alpha_2\}c_1+c_2}{\exp\{\alpha_0+\alpha_2\}c_1+c_2}$, $q^*(0,1) = \frac{\exp\{\lambda_3+\alpha_0+\alpha_1\}c_3+c_4}{\exp\{\alpha_0+\alpha_1\}c_3+c_4}$ and $q^*(1,1) = \frac{\exp\{\lambda_3+\alpha_0+\alpha_1+\alpha_2\}c_3+c_4}{\exp\{\alpha_0+\alpha_1+\alpha_2\}c_3+c_4}$ under (23) and our data generating assumptions of Section 2.2. End Proof.

Note the following special cases of Theorem 3:

- If $\lambda_3 = 0$ (i.e. relative to Figure 1 in main text, A_0 is not a direct cause of Y) then $\delta_1 = \lambda_1$, $\delta_2 = \lambda_2$ and $\delta_3 = 0$.
- If $\alpha_2 = 0$ (i.e. relative to Figure 1 in main text, A_0 is not a direct cause of A_2) then $\delta_1 = \lambda_1$ and $\delta_3 = 0$.
- If $\alpha_1 = 0$ (i.e. relative to Figure 1 in main text, A_0 is not an indirect cause of A_2 through A_1) then $\delta_2 = \lambda_2$ and $\delta_3 = 0$.

As $\lambda_1 \approx \psi_1$ and $\lambda_2 \approx \psi_2$ under the data generating models of Section 2.2 given (23), this shows that bias in the MLE of δ_1 and δ_2 based on the misspecified model (24) for ψ_1 and ψ_2, respectively, is a function of λ_3, α_1 and α_2.

3.3 Bias due to late study start (A_0 unobserved) in data with time-varying confounding (IPW versus standard regression)

We consider here the case of late-study-start (here meaning a study where A_0 is not measured) such that time-varying confounding affected by past exposure may also be present.

In each of the 10,000 data sets generated under the models of Section 2.1, we applied two estimators of the MSM coefficient vector (ψ_0, ψ_1, ψ_2) of the model (3). The first is an IPW estimator of η under the model (6) with stabilized weights such that the denominator of the weight was defined as an estimate of $\Pr[A_1 = 1|L_1] \times \Pr[A_2 = 1|L_2, A_1]$ under the saturated model

$$\Pr[A_1 = 1|L_1] = \frac{\exp\{\phi_0 + \phi_1 L_1\}}{1 + \exp\{\phi_0 + \phi_1 L_1\}} \tag{30}$$

and the data generating model (12) for $\Pr[A_2 = 1|L_2, A_1]$.

The second estimator is an MLE of the regression parameters γ in the model

$$\Pr[Y = 1|A_2, L_2, A_1, L_1] = \exp\{\gamma_0 + \gamma_1 A_2 + \gamma_2 L_2 + \gamma_3 A_1 + \gamma_4 L_1\} \tag{31}$$

The model (30) is a saturated model and therefore correctly specified. Further, under our data generating model (12), $\Pr[A_2 = 1|L_2, A_1] = \Pr[A_2 = 1|L_2, A_1, L_1, A_0]$. Therefore, this model is correctly specified for $\Pr[A_2 = 1|L_2, A_1]$ under our data generating mechanism. In the following we further show that the model (6) is correctly specified under the data generating models of Section 2.1.

Theorem 4: Given our data generating models (8) through (43) and rare outcome (20), the model (6) is correctly specified with

$$\begin{aligned} \eta_0 &= \log\left[\exp(\theta_0)m^*(a_1 = 0)\right] \\ \eta_1 &= \theta_1 \\ \eta_2 &= \log\left[\exp(\theta_3)\frac{m^*(a_1 = 1)}{m^*(a_1 = 0)}\right] \end{aligned} \tag{32}$$

where

$$\begin{aligned} m^*(a_1) = &\sum_{l_2} \exp\{\theta_2 l_2\} f(l_2|a_1; \beta) \times \\ &\sum_{l_1} \exp\{\theta_4 l_1\} \frac{\sum_{a_0} \exp\{\theta_5 a_0\} f(a_1|l_1, a_0; \alpha) f(l_1|a_0; \beta) f(a_0; \alpha)}{\sum_{a_0} f(a_1|l_1, a_0; \alpha) f(l_1|a_0; \beta) f(a_0; \alpha)} \sum_{a_0} f(l_1|a_0; \beta) f(a_0; \alpha) \end{aligned} \tag{33}$$

Proof: By definition we have

$$m(a_2, a_1) = \exp\{\eta_0 + \eta_1 a_2 + \eta_2 a_1 + \eta_3 a_2 a_1\} \tag{34}$$

By (20), our other data generating models and (4) we also have

$$m(a_2, a_1) = \exp\{\theta_0 + \theta_1 a_2 + \theta_3 a_1\} m^*(a_1) \tag{35}$$

By (34) and (35) we then have

- $\exp\{\eta_0\} = m(0, 0) = \exp\{\theta_0\} m^*(0)$
- $\exp\{\eta_1\} = \frac{\exp\{\eta_0 + \eta_1\}}{\exp\{\eta_0\}} = \frac{m(1,0)}{m(0,0)} = \frac{\exp\{\theta_0 + \theta_1\} m^*(0)}{\exp\{\theta_0\} m^*(0)} = \exp\{\theta_1\}$
- $\exp\{\eta_2\} = \frac{\exp\{\eta_0 + \eta_2\}}{\exp\{\eta_0\}} = \frac{m(0,1)}{m(0,0)} = \frac{\exp\{\theta_0 + \theta_3\} m^*(1)}{\exp\{\theta_0\} m^*(0)} = \exp\{\theta_3\} \frac{m^*(1)}{m^*(0)}$
- $\exp\{\eta_3\} = \frac{\exp\{\eta_0 + \eta_1 + \eta_2 + \eta_3\}}{\exp\{\eta_0 + \eta_1 + \eta_2\}} = \frac{m(1,1)}{\exp\{\eta_0 + \eta_1 + \eta_2\}} = \frac{\exp\{\theta_0 + \theta_1 + \theta_3\} m^*(1) m^*(0)}{\exp\{\theta_0 + \theta_1 + \theta_3\} m^*(0) m^*(1)} = 1$

End Proof.

Corollary 4. If either $\theta_5 = 0$ or $\alpha_2 = 0$ (i.e. A_0 is not a confounder for the effect of A_1 on Y) then η_2 of (32) is equivalent to ψ_2 of (47).

By (32) and $f(l_2|a_1 = 1; \beta)$ chosen as the data generating model (11), if $\theta_5 = 0$ then

$$\begin{aligned}\frac{m^*(a_1 = 1)}{m^*(a_1 = 0)} &= \frac{\sum_{l_2} \exp\{\theta_2 l_2\} f(l_2|a_1 = 1; \beta) \sum_{l_1} \exp\{\theta_4 l_1\} \sum_{a_0} f(l_1|a_0; \beta) f(a_0; \alpha)}{\sum_{l_2} \exp\{\theta_2 l_2\} f(l_2|a_1 = 0; \beta) \sum_{l_1} \exp\{\theta_4 l_1\} \sum_{a_0} f(l_1|a_0; \beta) f(a_0; \alpha)} \\ &= \frac{\sum_{l_2} \exp\{\theta_2 l_2\} f(l_2|a_1 = 1; \beta)}{\sum_{l_2} \exp\{\theta_2 l_2\} f(l_2|a_1 = 0; \beta)} \\ &= \frac{\left\{ \frac{\exp(\theta_2 + \beta_0 + \beta_1)}{1 + \exp(\beta_0 + \beta_1)} + \frac{1}{1 + \exp(\beta_0 + \beta_1)} \right\}}{\left\{ \frac{\exp(\theta_2 + \beta_0)}{1 + \exp(\beta_0)} + \frac{1}{1 + \exp(\beta_0)} \right\}}\end{aligned}$$

Alternatively, if $\alpha_2 = 0$ then

$$\begin{aligned}\frac{m^*(a_1 = 1)}{m^*(a_1 = 0)} &= \frac{\sum_{l_2} \exp\{\theta_2 l_2\} f(l_2|a_1 = 1; \beta) \sum_{l_1} \exp\{\theta_4 l_1\} \sum_{a_0} \exp\{\theta_5 a_0\} f(l_1|a_0; \beta) f(a_0; \alpha)}{\sum_{l_2} \exp\{\theta_2 l_2\} f(l_2|a_1 = 0; \beta) \sum_{l_1} \exp\{\theta_4 l_1\} \sum_{a_0} \exp\{\theta_5 a_0\} f(l_1|a_0; \beta) f(a_0; \alpha)} \\ &= \frac{\sum_{l_2} \exp\{\theta_2 l_2\} f(l_2|a_1 = 1; \beta)}{\sum_{l_2} \exp\{\theta_2 l_2\} f(l_2|a_1 = 0; \beta)} \\ &= \frac{\left\{ \frac{\exp(\theta_2 + \beta_0 + \beta_1)}{1 + \exp(\beta_0 + \beta_1)} + \frac{1}{1 + \exp(\beta_0 + \beta_1)} \right\}}{\left\{ \frac{\exp(\theta_2 + \beta_0)}{1 + \exp(\beta_0)} + \frac{1}{1 + \exp(\beta_0)} \right\}}\end{aligned}$$

End Proof. Theorems 2 and 4 with Corollary 4 confirm that, if A_0 is not a confounder for the effect of future exposure on Y, then an IPW estimator can recover ψ_1 and ψ_2 of the model (3) even when A_0 is unobserved.

We now show that the model (31) is misspecified under our data generating models of Section 2.1 when A_0 is a confounder of the effect of future exposure on the outcome.

Theorem 5: Given our data generating models and rare outcome (20), we have approximately

$$\Pr[Y = 1|A_2, L_2, A_1, L_1] = \exp\{\gamma_0 + \gamma_1 A_2 + \gamma_2 L_2 + \gamma_3 A_1 + \gamma_4 L_1 + \gamma_5 A_1 L_1\} \qquad (36)$$

where

$$\begin{aligned}\gamma_0 &= \log\left[\exp(\theta_0) r^*(a_1 = 0, l_1 = 0)\right] \\ \gamma_1 &= \theta_1 \\ \gamma_2 &= \theta_2\end{aligned}$$

$$\gamma_3 = \log\left[\exp(\theta_3)\left(\frac{c_1+\exp\{\theta_5+\alpha_2+\alpha_0\}c_2}{c_1+\exp\{\alpha_2+\alpha_0\}c_2}\right)\left(\frac{c_1+\exp\{\alpha_0\}c_2}{c_1+\exp\{\theta_5+\alpha_0\}c_2}\right)\right]$$

$$\gamma_4 = \log\left[\exp(\theta_4)\left(\frac{c_3+\exp\{\theta_5+\beta_1+\alpha_0\}c_4}{c_3+\exp\{\beta_1+\alpha_0\}c_4}\right)\left(\frac{c_1+\exp\{\alpha_0\}c_2}{c_1+\exp\{\theta_5+\alpha_0\}c_2}\right)\right]$$

$$\gamma_5 = \log\left[\frac{\left(\frac{c_3+\exp\{\theta_5+\alpha_2+\beta_1+\alpha_0\}c_4}{c_3+\exp\{\alpha_2+\beta_1+\alpha_0\}c_4}\right)\left(\frac{c_1+\exp\{\theta_5+\alpha_0\}c_2}{c_1+\exp\{\alpha_0\}c_2}\right)}{\left(\frac{c_1+\exp\{\theta_5+\alpha_2+\alpha_0\}c_2}{c_1+\exp\{\alpha_2+\alpha_0\}c_2}\right)\left(\frac{c_3+\exp\{\theta_5+\beta_1+\alpha_0\}c_4}{c_3+\exp\{\beta_1+\alpha_0\}c_4}\right)}\right]$$

and

$$c_1 = (1+\exp\{\alpha_0+\alpha_2\})(1+\exp\{\beta_0+\beta_1\})$$
$$c_2 = (1+\exp\{\alpha_0\})(1+\exp\{\beta_0\})$$
$$c_3 = (1+\exp\{\alpha_0+\alpha_1+\alpha_2\})(1+\exp\{\beta_0+\beta_1\})$$
$$c_4 = (1+\exp\{\alpha_0+\alpha_1\})(1+\exp\{\beta_0\})$$

Proof: Let $r(a_2,l_2,a_1,l_1) = \Pr[Y=1|A_2,L_2,A_1,L_1]$. By definition we have

$$\begin{aligned}&r(a_2,l_2,a_1,l_1) = \\ &\exp\{\gamma_0+\gamma_1a_2+\gamma_2l_2+\gamma_3a_1+\gamma_4l_1+\gamma_5a_1l_1+\gamma_6a_2a_1+\gamma_7a_2l_1+\gamma_8l_2a_1+\gamma_9l_2l_1+\gamma_{10}a_2l_2+ \\ &\gamma_{11}a_2l_2a_1+\gamma_{12}a_2a_1l_1+\gamma_{13}a_2l_2l_1+\gamma_{14}l_2a_1l_1+\gamma_{15}a_2l_2a_1l_1\}\end{aligned} \tag{37}$$

Also by definition

$$r(a_2,l_2,a_1,l_1) = \sum_{a_0}\Pr[Y=1|A_2=a_2,L_2=l_2,A_1=a_1,L_1=l_1,A_0=a_0]\times \frac{f(a_1|l_1,a_0;\alpha)f(l_1|a_0;\beta)f(a_0;\alpha)}{\sum_{a_0}f(a_1|l_1,a_0;\alpha)f(l_1|a_0;\beta)f(a_0;\alpha)} \tag{38}$$

Then by (20) and our other data generating models

$$r(a_2,l_2,a_1,l_1) = \exp\{\theta_0+\theta_1a_2+\theta_2l_2+\theta_3a_1+\theta_4l_1\}r^*(a_1,l_1) \tag{39}$$

where

$$r^*(a_1,l_1) = \frac{\sum_{a_0}\exp\{\theta_5a_0\}f(a_1|l_1,a_0;\alpha)f(l_1|a_0;\beta)f(a_0;\alpha)}{\sum_{a_0}f(a_1|l_1,a_0;\alpha)f(l_1|a_0;\beta)f(a_0;\alpha)} \tag{40}$$

Equations (37) and (39) give

- $\exp\{\gamma_0\} = r(0,0,0,0) = \exp\{\theta_0\}r^*(0,0)$

- $\exp\{\gamma_1\} = \frac{\exp\{\gamma_0+\gamma_1\}}{\exp\{\gamma_0\}} = \frac{r(1,0,0,0)}{\exp\{\gamma_0\}} = \frac{\exp\{\theta_0+\theta_1\}r^*(0,0)}{\exp\{\theta_0\}r^*(0,0)} = \exp\{\theta_1\}$
- $\exp\{\gamma_2\} = \frac{\exp\{\gamma_0+\gamma_2\}}{\exp\{\gamma_0\}} = \frac{r(0,1,0,0)}{\exp\{\gamma_0\}} = \frac{\exp\{\theta_0+\theta_2\}r^*(0,0)}{\exp\{\theta_0\}r^*(0,0)} = \exp\{\theta_2\}$
- $\exp\{\gamma_3\} = \frac{\exp\{\gamma_0+\gamma_3\}}{\exp\{\gamma_0\}} = \frac{r(0,0,1,0)}{\exp\{\gamma_0\}} = \frac{\exp\{\theta_0+\theta_3\}r^*(1,0)}{\exp\{\theta_0\}r^*(0,0)} = \exp\{\theta_3\}\frac{r^*(1,0)}{r^*(0,0)}$
- $\exp\{\gamma_4\} = \frac{\exp\{\gamma_0+\gamma_4\}}{\exp\{\gamma_0\}} = \frac{r(0,0,0,1)}{\exp\{\gamma_0\}} = \frac{\exp\{\theta_0+\theta_4\}r^*(0,1)}{\exp\{\theta_0\}r^*(0,0)} = \exp\{\theta_4\}\frac{r^*(0,1)}{r^*(0,0)}$
- $\exp\{\gamma_5\} = \frac{\exp\{\gamma_0+\gamma_3+\gamma_4+\gamma_5\}}{\exp\{\gamma_0+\gamma_3+\gamma_4\}} = \frac{r(0,0,1,1)}{\exp\{\gamma_0+\gamma_3+\gamma_4\}} = \frac{\exp\{\theta_0+\theta_3+\theta_4\}r^*(1,1)r^*(0,0)r^*(0,0)}{\exp\{\theta_0+\theta_3+\theta_4\}r^*(0,0)r^*(1,0)r^*(0,1)} = \frac{r^*(1,1)r^*(0,0)}{r^*(1,0)r^*(0,1)}$
- $\exp\{\gamma_6\} = \frac{\exp\{\gamma_0+\gamma_1+\gamma_3+\gamma_6\}}{\exp\{\gamma_0+\gamma_1+\gamma_3\}} = \frac{r(1,0,1,0)}{\exp\{\gamma_0+\gamma_1+\gamma_3\}} = \frac{\exp\{\theta_0+\theta_1+\theta_3\}r^*(1,0)r^*(0,0)}{\exp\{\theta_0+\theta_1+\theta_3\}r^*(0,0)r^*(1,0)} = 1$
- $\exp\{\gamma_7\} = \frac{\exp\{\gamma_0+\gamma_1+\gamma_4+\gamma_7\}}{\exp\{\gamma_0+\gamma_1+\gamma_4\}} = \frac{r(1,0,0,1)}{\exp\{\gamma_0+\gamma_1+\gamma_4\}} = \frac{\exp\{\theta_0+\theta_1+\theta_4\}r^*(0,1)r^*(0,0)}{\exp\{\theta_0+\theta_1+\theta_4\}r^*(0,0)r^*(0,1)} = 1$
- $\exp\{\gamma_8\} = \frac{\exp\{\gamma_0+\gamma_2+\gamma_3+\gamma_8\}}{\exp\{\gamma_0+\gamma_2+\gamma_3\}} = \frac{r(0,1,1,0)}{\exp\{\gamma_0+\gamma_2+\gamma_3\}} = \frac{\exp\{\theta_0+\theta_2+\theta_3\}r^*(1,0)r^*(0,0)}{\exp\{\theta_0+\theta_2+\theta_3\}r^*(0,0)r^*(1,0)} = 1$
- $\exp\{\gamma_9\} = \frac{\exp\{\gamma_0+\gamma_2+\gamma_4+\gamma_9\}}{\exp\{\gamma_0+\gamma_2+\gamma_4\}} = \frac{r(0,1,0,1)}{\exp\{\gamma_0+\gamma_2+\gamma_4\}} = \frac{\exp\{\theta_0+\theta_2+\theta_4\}r^*(0,1)r^*(0,0)}{\exp\{\theta_0+\theta_2+\theta_4\}r^*(0,0)r^*(0,1)} = 1$
- $\exp\{\gamma_{10}\} = \frac{\exp\{\gamma_0+\gamma_1+\gamma_3+\gamma_10\}}{\exp\{\gamma_0+\gamma_1+\gamma_3\}} = \frac{r(1,1,0,0)}{\exp\{\gamma_0+\gamma_1+\gamma_3\}} = \frac{\exp\{\theta_0+\theta_1+\theta_2\}r^*(0,0)}{\exp\{\theta_0+\theta_1+\theta_2\}r^*(0,0)} = 1$
-

$$\begin{aligned}
\exp\{\gamma_{11}\} &= \frac{\exp\{\gamma_0+\gamma_1+\gamma_2+\gamma_3+\gamma_6+\gamma_8+\gamma_{10}+\gamma_{11}\}}{\exp\{\gamma_0+\gamma_1+\gamma_2+\gamma_3+\gamma_6+\gamma_8+\gamma_{10}\}} \\
&= \frac{r(1,1,1,0)}{\exp\{\gamma_0+\gamma_1+\gamma_2+\gamma_3+\gamma_5+\gamma_6+\gamma_8+\gamma_{10}\}} \\
&= \frac{\exp\{\theta_0+\theta_1+\theta_2+\theta_3\}r^*(1,0)r^*(0,0)}{\exp\{\theta_0+\theta_1+\theta_2+\theta_3\}r^*(0,0)r^*(1,0)} \\
&= 1
\end{aligned}$$

$$\begin{aligned}
\exp\{\gamma_{12}\} &= \frac{\exp\{\gamma_0+\gamma_1+\gamma_3+\gamma_4+\gamma_5+\gamma_6+\gamma_7+\gamma_{12}\}}{\exp\{\gamma_0+\gamma_1+\gamma_3+\gamma_4+\gamma_5+\gamma_6+\gamma_7\}} \\
&= \frac{r(1,0,1,1)}{\exp\{\gamma_0+\gamma_1+\gamma_3+\gamma_4+\gamma_5+\gamma_6+\gamma_7\}} \\
&= \frac{\exp\{\theta_0+\theta_1+\theta_3+\theta_4\}r^*(1,1)r^*(0,0)r^*(0,0)r^*(1,0)r^*(0,1)}{\exp\{\theta_0+\theta_1+\theta_3+\theta_4\}r^*(0,0)r^*(1,0)r^*(0,1)r^*(0,0)r^*(1,1)} \\
&= 1
\end{aligned}$$

$$\begin{aligned}
\exp\{\gamma_{13}\} &= \frac{\exp\{\gamma_0+\gamma_1+\gamma_2+\gamma_4+\gamma_7+\gamma_9+\gamma_{10}+\gamma_{13}\}}{\exp\{\gamma_0+\gamma_1+\gamma_2+\gamma_4+\gamma_7+\gamma_9+\gamma_{10}\}} \\
&= \frac{r(1,1,0,1)}{\exp\{\gamma_0+\gamma_1+\gamma_2+\gamma_4+\gamma_7+\gamma_9+\gamma_{10}\}} \\
&= \frac{\exp\{\theta_0+\theta_1+\theta_2+\theta_4\}r^*(0,1)r^*(0,0)}{\exp\{\theta_0+\theta_1+\theta_2+\theta_4\}r^*(0,0)r^*(0,1)} \\
&= 1
\end{aligned}$$

$$\begin{aligned}
\exp\{\gamma_{14}\} &= \frac{\exp\{\gamma_0+\gamma_2+\gamma_3+\gamma_4+\gamma_5+\gamma_8+\gamma_9+\gamma_{14}\}}{\exp\{\gamma_0+\gamma_1+\gamma_2+\gamma_4+\gamma_7+\gamma_9+\gamma_{10}\}} \\
&= \frac{r(0,1,1,1)}{\exp\{\gamma_0+\gamma_2+\gamma_3+\gamma_4+\gamma_5+\gamma_8+\gamma_9\}} \\
&= \frac{\exp\{\theta_0+\theta_2+\theta_3+\theta_4\}r^*(1,1)r^*(0,0)r^*(0,0)r^*(1,0)r^*(0,1)}{\exp\{\theta_0+\theta_2+\theta_3+\theta_4\}r^*(0,0)r^*(1,0)r^*(0,1)r^*(1,1)r^*(0,0)} \\
&= 1
\end{aligned}$$

$$\begin{aligned}
\exp\{\gamma_{15}\} &= \frac{\exp\{\gamma_0+\gamma_1+\gamma_2+\ldots+\gamma_{14}+\gamma_{15}\}}{\exp\{\gamma_0+\gamma_1+\gamma_2+\ldots+\gamma_{14}\}} \\
&= \frac{r(1,1,1,1)}{\exp\{\gamma_0+\gamma_1+\gamma_2+\ldots+\gamma_{14}\}} \\
&= \frac{\exp\{\theta_0+\theta_1+\theta_2+\theta_3+\theta_4\}r^*(1,1)r^*(0,0)r^*(0,0)r^*(1,0)r^*(0,1)}{\exp\{\theta_0+\theta_1+\theta_2+\theta_3+\theta_4\}r^*(0,0)r^*(1,0)r^*(0,1)r^*(1,1)r^*(0,0)} \\
&= 1
\end{aligned}$$

Our results follow by

$$\frac{r^*(1,0)}{r^*(0,0)} = \left(\frac{c_1+\exp\{\theta_5+\alpha_2+\alpha_0\}c_2}{c_1+\exp\{\alpha_2+\alpha_0\}c_2}\right)\left(\frac{c_1+\exp\{\alpha_0\}c_2}{c_1+\exp\{\theta_5+\alpha_0\}c_2}\right),$$

$$\frac{r^*(0,1)}{r^*(0,0)} = \left(\frac{c_3+\exp\{\theta_5+\beta_1+\alpha_0\}c_4}{c_3+\exp\{\beta_1+\alpha_0\}c_4}\right)\left(\frac{c_1+\exp\{\alpha_0\}c_2}{c_1+\exp\{\theta_5+\alpha_0\}c_2}\right)$$

and

$$\frac{r^*(1,1)r^*(0,0)}{r^*(1,0)r^*(0,1)} = \frac{\left(\frac{c_3+\exp\{\theta_5+\alpha_2+\beta_1+\alpha_0\}c_4}{c_3+\exp\{\alpha_2+\beta_1+\alpha_0\}c_4}\right)\left(\frac{c_1+\exp\{\theta_5+\alpha_0\}c_2}{c_1+\exp\{\alpha_0\}c_2}\right)}{\left(\frac{c_1+\exp\{\theta_5+\alpha_2+\alpha_0\}c_2}{c_1+\exp\{\alpha_2+\alpha_0\}c_2}\right)\left(\frac{c_3+\exp\{\theta_5+\beta_1+\alpha_0\}c_4}{c_3+\exp\{\beta_1+\alpha_0\}c_4}\right)}$$

given the data generating models (10), (9), (8) for $f(a_1|l_1,a_0;\alpha)$, $f(l_1|a_0;\beta)$, and $f(a_0;\alpha)$, respectively. End Proof.

Note the following special cases of Theorem 5:

- If $\theta_5 = 0$ (i.e. relative to Figure 2, if A_0 is not a direct cause of Y) then $\gamma_3 = \theta_3$, $\gamma_4 = \theta_4$ and $\gamma_5 = 0$.
- If $\alpha_2 = 0$ (i.e. relative to Figure 2, if A_0 is not a direct cause of A_1) then $\gamma_3 = \theta_3$ and $\gamma_5 = 0$.
- If $\alpha_1 = 0$ (i.e. relative to Figure 2, if A_0 is not an indirect cause of A_1 through L_1) then $\gamma_4 = \theta_4$ and $\gamma_5 = 0$.

Thus bias in a standard regression based on the model (31) for θ of the model (43) is a function of θ_5, α_1 and α_2. Following Theorem 2, bias in this estimator for ψ is further a function of θ_2, θ_4 and β_1.

Web Appendix 4. Simulation with time-varying confounding affected by past exposure and "collider bias"

We have shown above that, when time-varying confounders are affected by past exposure, an outcome regression will not recover the true parameters of the MSM, even when identifying assumptions hold and even in the complete absence of model misspecification. Generally, the reason for this failure is that, in this setting, the outcome regression quantifies a conditional association that is not equal to the causal effect encoded by this MSM: this is a total time-varying treatment effect in the study population, capturing all paths by which the time-varying treatment affects the outcome. However, under the particular data generating scenarios considered above, the outcome regression parameters may still be linked to *some* (unarticulated) direct effect due to the absence of unmeasured common causes of the measured confounders and the outcome. The assumption of no unmeasured common causes of the measured confounders and the outcome is generally a strong assumption. Violation of this assumption would not only prevent a standard outcome regression from recovering our articulated total effect encoded by the MSM but, in general, prevent such a regression from recovering any causal effect [2, 5].

To illustrate the more realistic setting where unmeasured common causes are present, we extended scenario to allow the presence of an unmeasured common cause W of the measured confounders and the outcome but restrict that W does not affect exposure at any time (ensuring exchangeability). The data generating models of Section 2.1 were used for this scenario but with the following modifications:

- W was generated according to a constant logistic regression model:

$$\text{logit}\{\Pr[W = 1]\} = \omega_0 \tag{41}$$

- L_1 was generated according to a logistic regression model, possibly depending on A_0 and W:

$$\text{logit}\{\Pr[L_1 = 1|A_0]\} = \beta_0 + \beta_1 A_0 + \beta_2 W + \beta_3 A_0 \times W \tag{42}$$

- Y was generated according to a logistic regression model, possibly depending on the measured past and W:

$$\text{logit}\{\Pr[Y = 1|A_2, L_2, A_1, L_1, A_0]\} = \theta_0 + \theta_1 A_2 + \theta_2 L_2 + \theta_3 A_1 + \theta_4 L_1 + \theta_5 A_0 + \theta_6 W \tag{43}$$

Theorem 6: Given the modified data generating models of this section and a rare outcome, the MSM (3) is correctly specified with true ψ defined as

$$\psi_0 = \log\left[\exp(\theta_0) \times (b_0 + c_0)\right], \tag{44}$$

where

$$b_0 = \frac{\{1+\exp(\theta_2+\beta_0)\}\{1+\exp(\theta_4+\beta_0)\}}{\{1+\exp(\beta_0)\}\{1+\exp(\beta_0)\}\{1+\exp(\omega_0)\}},$$

and

$$c_0 = \frac{\exp(\theta_6+\omega_0)\left[\{1+\exp(\theta_2+\beta_0)\}\{1+\exp(\theta_4+\beta_0+\beta_2)\}\right]}{\{1+\exp(\beta_0)\}\{1+\exp(\beta_0+\beta_2)\}\{1+\exp(\omega_0)\}}.$$

$$\psi_1 = \theta_1. \tag{45}$$

$$\psi_2 = \log\left[\exp(\theta_3) \times \frac{(d_0+e_0)}{(b_0+c_0)}\right], \tag{46}$$

where

$$d_0 = \frac{\{1+\exp(\theta_2+\beta_0+\beta_1)\}\{1+\exp(\theta_4+\beta_0)\}}{\{1+\exp(\beta_0+\beta_1)\}\{1+\exp(\beta_0)\}\{1+\exp(\omega_0)\}},$$

and

$$e_0 = \frac{\exp(\theta_6+\omega_0)\left[\{1+\exp(\theta_2+\beta_0+\beta_1)\}\{1+\exp(\theta_4+\beta_0+\beta_2)\}\right]}{\{1+\exp(\beta_0+\beta_1)\}\{1+\exp(\beta_0+\beta_2)\}(\{1+\exp(\omega_0)\}}.$$

$$\psi_3 = \log\left[\exp(\theta_3) \times \frac{(f_0+g_0)}{(b_0+c_0)}\right], \tag{47}$$

where

$$f_0 = \frac{\{1+\exp(\theta_2+\beta_0)\}\{1+\exp(\theta_4+\beta_0+\beta_1)\}}{\{1+\exp(\beta_0)\}\{1+\exp(\beta_0+\beta_1)\}\{1+\exp(\omega_0)\}},$$

and

$$g_0 = \frac{\exp(\theta_6+\omega_0)\left[\{1+\exp(\theta_2+\beta_0)\}\{1+\exp(\theta_4+\beta_0+\beta_1+\beta_2+\beta_3)\}\right]}{\{1+\exp(\beta_0)\}\{1+\exp(\beta_0+\beta_1+\beta_2+\beta_3)\}(\{1+\exp(\omega_0)\}}.$$

The proof follows analogous logic to that of Theorem 2. As expected, the expressions for ψ in Theorem 6 reduce to those of Theorem 2 when $\omega_0 = \theta_6 = \beta_2 = \beta_3 = 0$.

References

[1] J. Pearl, "Causal diagrams for empirical research," *Biometrika*, vol. 82, pp. 669–710, 1995.

[2] J. Robins, "A new approach to causal inference in mortality studies with a sustained exposure period: application to the healthy worker survivor effect," *Mathematical Modelling*, vol. 7, pp. 1393–1512, 1986. [Errata (1987) in Computers and Mathematics with Applications 14, 917-921. Addendum (1987) in Computers and Mathematics with Applications 14, 923-945. Errata (1987) to addendum in Computers and Mathematics with Applications 18, 477.].

[3] J. Robins, M. Hernán, and B. Brumback, "Marginal structural models and causal inference in epidemiology," *Epidemiology*, vol. 11, no. 5, pp. 550–560, 2000.

[4] J. Young and E. Tchetgen Tchetgen, "Simulation from a known Cox MSM using standard parametric models for the g-formula," *Statistics in Medicine*, vol. 10, pp. 1001–14, 2014.

[5] J. Robins and M. Hernán, "Estimation of the causal effects of time-varying exposures," in *Advances in Longitudinal Data Analysis* (G. Fitzmaurice, M. Davidian, G. Verbeke, and G. Molenberghs, eds.), pp. 553–599, Boca Raton, FL: Chapman and Hall/CRC Press, 2009.

Web Table 1. Measures of accuracy and precision of the estimated effects on the log odds scale under different theoretical life-course model when information regarding early-life exposure is missing

	Accuracy and Precision of								
	Estimated direct effect of depression in early-life (λ_e) on log odds of stroke			**Estimated direct effect of depression in mid-life (λ_m) on log odds of stroke**			**Estimated direct effect of depression in late-life (λ_l) on log odds of stroke**		
	Standard error	MSE	95% CI coverage	Standard error	MSE	95% CI coverage	Standard error	MSE	95% CI coverage
CONVENTIONAL REGRESSION									
Varying the magnitude of the effect of AUD on stroke									
1. No confounding	0.07	0.01	95%	0.08	0.01	95%	0.08	0.01	95%
2. Moderate positive confounding	0.06	0.02	53%	0.07	0.02	61%	0.07	<0.01	95%
3. Strong positive confounding	0.05	0.04	6%	0.06	0.04	12%	0.06	<0.01	95%
4. Negative confounding	<0.01	<0.01	0%	<0.01	<0.01	81%	<0.01	0.01	81%
5. Greater confounding earlier in life	<0.01	<0.01	0%	<0.01	<0.01	51%	<0.01	<0.01	58%
6. Greater confounding later in life	0.06	0.02	47%	0.06	0.04	17%	0.06	<0.01	95%
Varying the magnitude of the effect of depression on AUD									
7. No effect of depression on AUD	0.06	<0.01	95%	0.06	<0.01	95%	0.06	<0.01	95%
Unmeasured confounder of AUD and stroke									
8. Moderate positive unmeasured confounder of AUD and stroke	0.08	0.01	86%	0.09	0.02	74%	0.09	0.01	95%
IPW ESTIMATION									
Varying the magnitude of the effect of AUD on stroke									
1. No confounding	0.07	<0.01	95%	0.08	0.01	95%	0.08	0.01	95%

2. Moderate positive confounding	0.06	<0.01	95%	<0.01	<0.01	94%	<0.01	<0.01	95%
3. Strong positive confounding	0.05	<0.01	95%	0.08	<0.01	95%	0.08	<0.01	95%
4. Negative confounding	0.08	<0.01	0%	0.22	<0.01	69%	0.21	<0.01	69%
5. Greater confounding earlier in life	0.05	<0.01	0%	0.05	<0.01	69%	0.06	<0.01	69%
6. Greater confounding later in life	0.05	<0.01	95%	<0.01	<0.01	95%	<0.01	<0.01	94%
Varying the magnitude of the effect of depression on AUD									
7. No effect depression on AUD	0.06	<0.01	95%	<0.01	<0.01	95%	<0.01	<0.01	95%
Unmeasured confounder of AUD and stroke									
8. Moderate positive unmeasured confounder of AUD and stroke	0.08	0.01	95%	0.09	0.01	95%	0.09	0.01	94%

Note: Data was generated with 10,000 replications and 100,000 people in each sample. 95% CI coverage is the proportion of samples for which the estimated 95% confidence intervals contain the true effect and is influenced by the sample size used in the simulation.

Web Table 2. Measures of accuracy and precision of estimated effects on the log odds scale under different theoretical life-course models when information regarding early-life exposure is missing

Theoretical model under which data is generated	Accuracy and Precision of: Estimated direct effect of depression in mid-life (λ_m) on log odds of stroke			Estimated direct effect of depression in late-life (λ_l) on log odds of stroke			Estimated lifetime effects of depression (λ_m+λ_l) on log odds of stroke		
	Standard error	MSE	95% CI coverage	Standard error	MSE	95% CI coverage	Standard error	MSE	95% CI coverage
Accumulation model with equal effects across the life-course	0.08	0.01	87%	0.08	0.01	87%	0.11	0.33	0%
Early-life critical period	0.09	0.02	83%	0.09	0.02	80%	0.14	0.83	0%
Pathway model	0.09	0.01	95%	0.11	0.01	95%	0.14	0.02	2%
Accumulation model with increasing effect across life-course	0.08	0.01	92%	0.08	0.01	93%	0.11	0.12	0%
Accumulation model with decreasing effect across life-course	0.08	0.01	83%	0.08	0.01	80%	0.10	0.58	0%

Note: Data was generated with 10,000 replications and 100,000 people in each sample. 95% CI coverage is the proportion of samples for which the estimated 95% confidence intervals contain the true effect and is influenced by the sample size used in the simulation.

Web Table 3. Variance in estimated effects of depression on stroke on the log odds scale under an accumulation model comparing conventional regression and inverse probability weighting to estimate MSM parameters, in the presence of a late-study-start and time-varying confounders.

Analytic method	Variance in Estimate of							
	Direct effect of depression in mid-life (δ_m) on log odds of stroke			Direct effect of depression in late-life (δ_l) on log odds of stroke			Lifetime effects of depression ($\delta_m+\delta_l$) on log odds of stroke	
	Standard error	MSE	95% CI coverage	Standard error	MSE	95% CI coverage	MSE	95% CI coverage
Conventional regression	0.06	0.02	35%	0.06	<0.01	95%	0.69	<0.01%
IPW estimation of MSM	0.06	0.01	87%	0.06	<0.01	94%	0.43	<0.01%

Notes: Data was generated with 10,000 replications and 100,000 people in each sample. 95% CI coverage is the proportion of samples for which the estimated 95% confidence intervals contain the true effect and is influenced by the sample size used in the simulation. IPW=inverse probability weight; MSM=marginal structural model

Web Appendix 5. SAS code

```
*************************************************************************************************************************************************
```

SAS code for Table 1 and Supplemental Table 1. Bias in estimates of the causal direct effect of depression in early-, mid-, and late-life on stroke under the accumulation model using conventional regression versus IPW in the presence or absence of time-varying confounding by alcohol use disorder

Order of simulation code for Table 1

- Scenarios 1-6
- Scenario 7
- Scenario 8

MACRO ORDER

- %samples
- %simulate
- %estimate
- %calculate_truth
- %trueptrtrange

Table 1 key:

CONVENTIONAL REGRESSION			IPW		
LATE-LIFE	MID-LIFE	LATE-LIFE	LATE-LIFE	MID-LIFE	LATE-LIFE
lambda 3 (theta1)	lambda 2 (theta3)	lambda 1 (theta5)	psi 3	psi 2	psi 1

```
*************************************************************************************************************************************************

options nonotes;

%macro samples(log=,theta0=,theta1=,theta2=,theta3=,theta4=,theta5=,beta0=,beta1=,alpha0=,alpha1=,alpha2=,nsubjects=,nsamples=);
        *calculates the true value of psi based on input choices of theta and beta;
        %calculate_truth;
        %trueptrtrange;
        *empty data set that will hold the nsamples IPW estimates of psi;
        data psifinal;
        run;
        *empty data set that will hold the nsamples MLEs of theta;
        data thetafinal;
        run;
        %let seed = 20;
        %do sample = 1 %to &nsamples;
            %put sample = &sample;
             %let seed = %eval(&seed + 3);
             %put seed= &seed;
             %simulate;
             %estimate;
        %end;
```

```
/*      PSI FINAL ---------------------------------------------------------- */
        title "psihat means";
        proc means data=psifinal n mean std STDERR clm;
        var A2 A2_stderr A2_pct_bias A2_CI_coverage A1 A1_stderr A1_pct_bias A1_CI_coverage A0 A0_stderr A0_pct_bias A0_CI_coverage;
        run;
        %let true = psi1;
        %let pred = A2;
        %let dataset = psifinal;
        proc sql;
                create table psi_mse_A2 as
                select mean((&pred.-&true.)**2) as MSE_A2, sqrt(mean((&pred.-&true.)**2)) as RMSE_A2 from &dataset.;
                quit;
        %let true = psi2;
        %let pred = A1;
        %let dataset = psifinal;
        proc sql;
                create table psi_mse_A1 as
                select mean((&pred.-&true.)**2) as MSE_A1, sqrt(mean((&pred.-&true.)**2)) as RMSE_A1 from &dataset.;
                quit;
        %let true = psi3;
        %let pred = A0;
        %let dataset = psifinal;
        proc sql;
                create table psi_mse_A0 as
                select mean((&pred.-&true.)**2) as MSE_A0, sqrt(mean((&pred.-&true.)**2)) as RMSE_A0 from &dataset.;
                quit;
        title "psi hat MSE";
        proc print data=psi_mse_A0; run;
        proc print data=psi_mse_A1; run;
        proc print data=psi_mse_A2; run;

/*      THETA FINAL → lambda ----------------------------------------------------------- */
        title 'thetahat means';
        proc means data=thetafinal n mean std STDERR clm;
        var A2 A2_stderr A2_pct_bias A2_CI_coverage A1 A1_stderr A1_pct_bias A1_CI_coverage A0 A0_stderr A0_pct_bias A0_CI_coverage;
        run;
        %let true = psi1;
        %let pred = A2;
        %let dataset = thetafinal;
        proc sql;
                create table theta_mse_A2 as
                select mean((&pred.-&true.)**2) as MSE_A2, sqrt(mean((&pred.-&true.)**2)) as RMSE_A2 from &dataset.;
                quit;
        %let true = psi2;
        %let pred = A1;
        %let dataset = thetafinal;
        proc sql;
                create table theta_mse_A1 as
                select mean((&pred.-&true.)**2) as MSE_A1, sqrt(mean((&pred.-&true.)**2)) as RMSE_A1 from &dataset.;
                quit;
        %let true = psi3;
```

```
        %let pred = A0;
        %let dataset = thetafinal;
        proc sql;
                create table theta_mse_A0 as
                select mean((&pred.-&true.)**2) as MSE_A0, sqrt(mean((&pred.-&true.)**2)) as RMSE_A0 from &dataset.;
                quit;
        title "theta hat MSE";
        proc print data=theta_mse_A0; run;
        proc print data=theta_mse_A1; run;
        proc print data=theta_mse_A2; run;
        title;
DM "log; clear; ";
%mend samples;


%macro simulate;
data simsample;
 *      generating simulation dataset ;
        do id= 1 to %eval(&nsubjects);
        call streaminit(&seed);
                * A0 = Early-life depression ------------------------------------------------;
                logitA0 = &alpha0 ;
                 pA0 = 1/(1+exp(-logitA0)) ;
                 uA0=rand('uniform');
                 if uA0<=pA0 then A0=1;
                 if uA0>pA0>. then A0=0;
                * L1 = Early-life AUD -------------------------------------------------------;
                logitL1 = &beta0 + (&beta1 * A0) ;
                               pL1= 1/(1+exp(-logitL1)) ;
                               uL1=rand('uniform');
                               if uL1<=pL1  then L1=1;
                               if uL1>pL1>. then L1=0;
                * A1 = Mid-life depression ------------------------------------------------;
                logitA1 = &alpha0 + (&alpha1 * L1) + (&alpha2 * A0);
                 pA1 = 1/(1+exp(-logitA1)) ;
                 uA1=rand('uniform');
                 if uA1<=pA1 then A1=1;
                 if uA1>pA1>. then A1=0;
                * L2 = Mid-life AUD -------------------------------------------------------;
                logitL2 = &beta0 + (&beta1 * A1) ;
                               pL2= 1/(1+exp(-logitL2)) ;
                               uL2=rand('uniform');
                               if uL2<=pL2  then L2=1;
                               if uL2>pL2>. then L2=0;
                * A2 = Late-life depression ------------------------------------------------;
         logitA2 = &alpha0 + (&alpha1 * L2) + (&alpha2 * A1);
                 pA2 = 1/(1+exp(-logitA2)) ;
                 uA2=rand('uniform');
                 if uA2<=pA2 then A2=1;
                 if uA2>pA2>. then A2=0;
                * Y ------------------------------------------------------------------------;
```

```
               logitY = &theta0 + (&theta1 * A2) + (&theta2 * L2) + (&theta3 * A1)  + (&theta4 * L1)  + (&theta5 * A0)    ;
                              pY = 1/(1+exp(-logitY)) ;
                uY=rand('uniform');
                if uY<=pY  then Y=1;
                if uY>pY>. then Y=0;
                      output;
       end;
       keep id A0 L1 A1 L2 A2 Y ;
run;
%mend;

%macro estimate;
*numerator IPTW A1;
proc logistic NOPRINT descending data=simsample;
        model A1= A0;
        output out=modtn1 p=ptrt_n1;
        run;
*denominator IPTW A1;
proc logistic NOPRINT descending data=simsample;
        model A1=L1 A0 ;
        output out=modtd1 p=ptrt_d1;
        run;
*numerator IPTW A2;
proc logistic NOPRINT descending data=simsample;
        model A2= A1 ;
       output out=modtn2 p=ptrt_n2;
       run;
*denominator IPTW A2;
proc logistic NOPRINT descending data=simsample;
        model A2=L2 A1 ;
        output out=modtd2 p=ptrt_d2;
        run;
data new_sim;
        merge modtn1 modtd1 modtn2 modtd2;
        by id;
       /*calculate time 1 part of weight*/
       if A1=1 then do;
          pan1=ptrt_n1;
          pad1=ptrt_d1;
          end;
       else do;
          pan1=1-ptrt_n1;
          pad1=1-ptrt_d1;
          end;
       /*calculate time 2 part of weight*/
       if A2=1 then do;
          pan2=ptrt_n2;
          pad2=ptrt_d2;
          end;
        else do;
           pan2=1-ptrt_n2;
```

```
            pad2=1-ptrt_d2;
           end;
        /*calculate products and stabilized weight*/
        prod_n=pan1*pan2;
        prod_d=pad1*pad2;
        stabw=prod_n/prod_d;
        run;

/*      GET PSIHAT ------------------------------------------------------------------------------------------------*/
        ods exclude all;
        proc genmod data=new_sim descending;
               model Y = A2 A1 A0 / link=logit dist=binomial cl;
               weight stabw;
               ods output parameterestimates=psidata_out;
               run;
        ods exclude none;
        * transpose parameter estimates;
        proc transpose data=psidata_out out=psidata_wide;
               id Parameter;
               var Estimate StdErr LowerWaldCL UpperWaldCL;
               run;
        data estimate; set psidata_wide(drop=Intercept Scale); where _NAME_='Estimate'; run;
        data stderr;   set psidata_wide(drop=Intercept Scale); where _NAME_='StdErr'; run;
        data lcl;              set psidata_wide(drop=Intercept Scale); where _NAME_='LowerWaldCL'; run;
        data ucl;              set psidata_wide(drop=Intercept Scale); where _NAME_='UpperWaldCL'; run;
        data psidata;
               set estimate;
               if _N_=1 then set truepsi;
               if _N_=1 then set stderr(rename=(A0=A0_stderr A1=A1_stderr A2=A2_stderr));
               if _N_=1 then set lcl  (rename=(A0=A0_lcl A1=A1_lcl A2=A2_lcl));
               if _N_=1 then set ucl  (rename=(A0=A0_ucl A1=A1_ucl A2=A2_ucl));
               drop _NAME_ _LABEL_;
               run;
        data psidata;
               set psidata;
               sample=&sample;
               * CI coverage ;
               A0_CI_coverage=.; if A0_lcl<psi3 and A0_ucl>psi3 then A0_CI_coverage=1; else A0_CI_coverage=0;
               A1_CI_coverage=.; if A1_lcl<psi2 and A1_ucl>psi2 then A1_CI_coverage=1; else A1_CI_coverage=0;
               A2_CI_coverage=.; if A2_lcl<psi1 and A2_ucl>psi1 then A2_CI_coverage=1; else A2_CI_coverage=0;
               * pct bias ;
               format A1_pct_bias percentn. A2_pct_bias percentn. A0_pct_bias percentn.;
               A0_pct_bias=((A0-psi3)/psi3);
               A1_pct_bias=((A1-psi2)/psi2);
               A2_pct_bias=((A2-psi1)/psi1);
               run;
               data psifinal;
               set psifinal psidata;
               if sample ^= .;

        /*      GET THETAHAT ----------------------------------------------------------------------------------------*/
```

```
        ods exclude all;
        proc genmod data=new_sim descending;
                model Y = A2 L2 A1 L1 A0 / link=logit dist=binomial cl;
                ods output parameterestimates=simsample_out;
                run;
        ods exclude none;
        * transpose parameter estimates;
        proc transpose data=simsample_out out=simsample_wide;
                id Parameter;
                var Estimate StdErr LowerWaldCL UpperWaldCL;
                run;
        data estimate; set simsample_wide(drop=Intercept Scale); where _NAME_='Estimate'; run;
        data stderr;   set simsample_wide(drop=Intercept Scale); where _NAME_='StdErr'; run;
        data lcl;              set simsample_wide(drop=Intercept Scale); where _NAME_='LowerWaldCL'; run;
        data ucl;              set simsample_wide(drop=Intercept Scale); where _NAME_='UpperWaldCL'; run;
        data thetadata;
                set estimate;
                if _N_=1 then set truepsi;
                if _N_=1 then set stderr(rename=(A0=A0_stderr A1=A1_stderr A2=A2_stderr));
                if _N_=1 then set lcl  (rename=(A0=A0_lcl A1=A1_lcl A2=A2_lcl));
                if _N_=1 then set ucl  (rename=(A0=A0_ucl A1=A1_ucl A2=A2_ucl));
                drop _NAME_ _LABEL_;
                run;
        data thetadata;
                set thetadata;
                sample=&sample;
                * CI coverage ;
                A0_CI_coverage=.; if A0_lcl<psi3 and A0_ucl>psi3 then A0_CI_coverage=1; else A0_CI_coverage=0;
                A1_CI_coverage=.; if A1_lcl<psi2 and A1_ucl>psi2 then A1_CI_coverage=1; else A1_CI_coverage=0;
                A2_CI_coverage=.; if A2_lcl<psi1 and A2_ucl>psi1 then A2_CI_coverage=1; else A2_CI_coverage=0;
                * pct bias ;
                format A1_pct_bias percentn. A2_pct_bias percentn. A0_pct_bias percentn.;
                A0_pct_bias=((A0-psi3)/psi3);
                A1_pct_bias=((A1-psi2)/psi2);
                A2_pct_bias=((A2-psi1)/psi1);
                run;
                data thetafinal;
                set thetafinal thetadata;
                if sample ^= .;
                run;
%mend;

%macro calculate_truth;
data truepsi; run;
data truepsi;
set truepsi;
 *      formula for psi0 --intercept of the MSM;
        psi0fracnum1=exp(&theta2 + &beta0) ;
        psi0fracden1=1+exp(&beta0);
        psi0frac1=psi0fracnum1/psi0fracden1;
        psi0frac2=1/psi0fracden1;
```

```
       psi0sum1=psi0frac1+psi0frac2;
       psi0fracnum3=exp(&theta4+&beta0);
       psi0frac3=psi0fracnum3/psi0fracden1;
       psi0sum2=psi0frac3+psi0frac2;
       exppsi0=exp(&theta0)*psi0sum1*psi0sum2;
       psi0=log(exppsi0);
 *     formula for psi1 -- coefficient of late life exposure ;
       psi1=&theta1;
 *     formula for psi2 -- coefficient of mid life exposure;
       psi2numfrac1=exp(&theta2 + &beta0 + &beta1) ;
       psi2denfrac1=1+exp(&beta0+&beta1);
       psi2frac1=psi2numfrac1/psi2denfrac1;
       psi2frac2=1/psi2denfrac1;
       psi2sum1=psi2frac1+psi2frac2;
       psi2frac=psi2sum1/psi0sum1;
       exppsi2=exp(&theta3)*psi2frac;
       psi2=log(exppsi2);
 *     formula for psi3 -- coefficient of early life exposure;
       psi3numfrac1=exp(&theta4 + &beta0 +&beta1);
       psi3frac1=psi3numfrac1/psi2denfrac1;
       psi3sum1=psi3frac1+psi2frac2;
       psi3frac=psi3sum1/psi0sum2;
       exppsi3=exp(&theta5)*psi3frac;
       psi3=log(exppsi3);
keep /*psi0*/ psi1 psi2 psi3;
run;
proc print data=truepsi;
run;
%mend calculate_truth;


%macro trueptrtrange;
data pat;
run;
data pat;
set pat;
       pat00=exp(&alpha0)/(1+exp(&alpha0));
       pat10=exp(&alpha0+&alpha1)/(1+exp(&alpha0+&alpha1));
       pat01=exp(&alpha0+&alpha2)/(1+exp(&alpha0+&alpha2));
       pat11=exp(&alpha0+&alpha1+&alpha2)/(1+exp(&alpha0+&alpha1+&alpha2));
run;
proc print data=pat;
run;
%mend;

*Scenario 1 No Confounding;
%samples(log=M1,theta0=-6,theta1=0.69, theta2=0.00, theta3=0.69, theta4=0.00, theta5=0.69, beta0=-.3, beta1=1.2, alpha0=0.00,
alpha1=0.41, alpha2=0.41, nsubjects=100000, nsamples=10000);

*Scenario 2 Moderate Positive Confounding;
```

```
%samples(log=M2,theta0=-6,theta1=0.69, theta2=0.41, theta3=0.58, theta4=0.41, theta5=0.58, beta0=-.3, beta1=1.2, alpha0=0.00,
alpha1=0.41, alpha2=0.41, nsubjects=100000, nsamples=10000);

*Scenario 3 Strong Positive Confounding;
%samples(log=M3,theta0=-6,theta1=0.69, theta2=0.69, theta3=0.51, theta4=0.69, theta5=0.51, beta0=-.3, beta1=1.2, alpha0=0.00,
alpha1=0.41, alpha2=0.41, nsubjects=100000, nsamples=10000);

*Scenario 4 Negative Confounding;
%samples(log=M4,theta0=-6,theta1=0.69, theta2=-0.41,theta3=0.81, theta4=-0.41,theta5=0.81, beta0=-.3, beta1=1.2, alpha0=0.00,
alpha1=0.41, alpha2=0.41, nsubjects=100000, nsamples=10000);

*Scenario 5 Greater confounding earlier in life;
%samples(log=M5,theta0=-6,theta1=0.69, theta2=0.41, theta3=0.58, theta4=0.69, theta5=0.51, beta0=-.3, beta1=1.2, alpha0=0.00,
alpha1=0.41, alpha2=0.41, nsubjects=100000, nsamples=10000);

*Scenario 6 Greater confounding later in life;
%samples(log=M6,theta0=-6,theta1=0.69, theta2=0.69, theta3=0.51, theta4=0.41, theta5=0.58, beta0=-.3, beta1=1.2, alpha0=0.00,
alpha1=0.41, alpha2=0.41, nsubjects=100000, nsamples=10000);



*************************************************************************************************************************************

Table 1 Scenario 7: No effect of depression on AUD

*************************************************************************************************************************************
options nonotes;
*options nosource;

%macro samples(log=,theta0=,theta1=,theta2=,theta3=,theta4=,theta5=,beta0=,beta1=,alpha0=,alpha1=,alpha2=,nsubjects=,nsamples=);
       *calculates the true value of psi based on input choices of theta and beta;
       %calculate_truth;
       %trueptrtrange;
       *empty data set that will hold the nsamples IPW estimates of psi;
       data psifinal;
       run;
       *empty data set that will hold the nsamples MLEs of theta;
       data thetafinal;
       run;
       %let seed = 20;
       %do sample = 1 %to &nsamples;
           %put sample = &sample;
            %let seed = %eval(&seed + 3);
            %put seed= &seed;
            %simulate;
            %estimate;
       %end;
/*     PSI FINAL ------------------------------------------------------- */
       title "psihat means";
       proc means data=psifinal n mean std STDERR clm;
       var A2 A2_stderr A2_pct_bias A2_CI_coverage A1 A1_stderr A1_pct_bias A1_CI_coverage A0 A0_stderr A0_pct_bias A0_CI_coverage;
```

```
run;
%let true = psi1;
%let pred = A2;
%let dataset = psifinal;
proc sql;
        create table psi_mse_A2 as
        select mean((&pred.-&true.)**2) as MSE_A2, sqrt(mean((&pred.-&true.)**2)) as RMSE_A2 from &dataset.;
        quit;
%let true = psi2;
%let pred = A1;
%let dataset = psifinal;
proc sql;
        create table psi_mse_A1 as
        select mean((&pred.-&true.)**2) as MSE_A1, sqrt(mean((&pred.-&true.)**2)) as RMSE_A1 from &dataset.;
        quit;
%let true = psi3;
%let pred = A0;
%let dataset = psifinal;
proc sql;
        create table psi_mse_A0 as
        select mean((&pred.-&true.)**2) as MSE_A0, sqrt(mean((&pred.-&true.)**2)) as RMSE_A0 from &dataset.;
        quit;
title "psi hat MSE";
proc print data=psi_mse_A0; run;
proc print data=psi_mse_A1; run;
proc print data=psi_mse_A2; run;
/*      THETA FINAL ------------------------------------------------------------ */
title 'thetahat means';
proc means data=thetafinal n mean std STDERR clm;
var A2 A2_stderr A2_pct_bias A2_CI_coverage A1 A1_stderr A1_pct_bias A1_CI_coverage A0 A0_stderr A0_pct_bias A0_CI_coverage;
run;
%let true = psi1;
%let pred = A2;
%let dataset = thetafinal;
proc sql;
        create table theta_mse_A2 as
        select mean((&pred.-&true.)**2) as MSE_A2, sqrt(mean((&pred.-&true.)**2)) as RMSE_A2 from &dataset.;
        quit;
%let true = psi2;
%let pred = A1;
%let dataset = thetafinal;
proc sql;
        create table theta_mse_A1 as
        select mean((&pred.-&true.)**2) as MSE_A1, sqrt(mean((&pred.-&true.)**2)) as RMSE_A1 from &dataset.;
        quit;
%let true = psi3;
%let pred = A0;
%let dataset = thetafinal;
proc sql;
        create table theta_mse_A0 as
        select mean((&pred.-&true.)**2) as MSE_A0, sqrt(mean((&pred.-&true.)**2)) as RMSE_A0 from &dataset.;
```

```
		quit;
	title "theta hat MSE";
	proc print data=theta_mse_A0; run;
	proc print data=theta_mse_A1; run;
	proc print data=theta_mse_A2; run;
	title;
DM "log; clear; ";
%mend samples;


%macro simulate;
data simsample;
	do id= 1 to %eval(&nsubjects);
	call streaminit(&seed);
		* A0 = Early-life depression ------------------------------------------------;
		logitA0 = &alpha0 ;
				pA0 = 1/(1+exp(-logitA0)) ;
				uA0=rand('uniform');
				if uA0<=pA0 then A0=1;
				if uA0>pA0>. then A0=0;
		* L1 = Early-life AUD -------------------------------------------------------;
		logitL1 = &beta0 + (&beta1 * A0) ;
				pL1= 1/(1+exp(-logitL1)) ;
				uL1=rand('uniform');
				if uL1<=pL1  then L1=1;
				if uL1>pL1>. then L1=0;
		* A1 = Mid-life depression --------------------------------------------------;
		logitA1 = &alpha0 + (&alpha1 * L1) + (&alpha2 * A0);
				pA1 = 1/(1+exp(-logitA1)) ;
				uA1=rand('uniform');
				if uA1<=pA1 then A1=1;
				if uA1>pA1>. then A1=0;
		* L2 = Mid-life AUD ---------------------------------------------------------;
		logitL2 = &beta0 + (&beta1 * A1) ;
				pL2= 1/(1+exp(-logitL2)) ;
				uL2=rand('uniform');
				if uL2<=pL2  then L2=1;
				if uL2>pL2>. then L2=0;
		* A2 = Late-life depression -------------------------------------------------;
		logitA2 = &alpha0 + (&alpha1 * L2) + (&alpha2 * A1);
				pA2 = 1/(1+exp(-logitA2)) ;
				uA2=rand('uniform');
				if uA2<=pA2 then A2=1;
				if uA2>pA2>. then A2=0;
		* Y -------------------------------------------------------------------------;
		logitY = &theta0 + (&theta1 * A2) + (&theta2 * L2) + (&theta3 * A1)  + (&theta4 * L1)  + (&theta5 * A0)    ;
				pY = 1/(1+exp(-logitY)) ;
				uY=rand('uniform');
				if uY<=pY  then Y=1;
				if uY>pY>. then Y=0;
			output;
```

```
        end; /*subjects*/
        keep id A0 L1 A1 L2 A2 Y ;
run;
%mend;

%macro estimate;
*numerator IPTW A1;
proc logistic NOPRINT descending data=simsample;
        model A1= A0;
        output out=modtn1 p=ptrt_n1;
        run;
*denominator IPTW A1;
proc logistic NOPRINT descending data=simsample;
        model A1=L1 A0 ;
        output out=modtd1 p=ptrt_d1;
        run;
*numerator IPTW A2;
proc logistic NOPRINT descending data=simsample;
        model A2= A1 ;
        output out=modtn2 p=ptrt_n2;
        run;
*denominator IPTW A2;
proc logistic NOPRINT descending data=simsample;
        model A2=L2 A1 ;
        output out=modtd2 p=ptrt_d2;
        run;
data new_sim;
        merge modtn1 modtd1 modtn2 modtd2;
        by id;
        /*calculate time 1 part of weight*/
        if A1=1 then do;
           pan1=ptrt_n1;
           pad1=ptrt_d1;
           end;
        else do;
           pan1=1-ptrt_n1;
           pad1=1-ptrt_d1;
           end;
        /*calculate time 2 part of weight*/
        if A2=1 then do;
           pan2=ptrt_n2;
           pad2=ptrt_d2;
           end;
         else do;
            pan2=1-ptrt_n2;
            pad2=1-ptrt_d2;
            end;
        /*calculate products and stabilized weight*/
        prod_n=pan1*pan2;
        prod_d=pad1*pad2;
        stabw=prod_n/prod_d;
```

```
        run;
/*      GET PSIHAT ------------------------------------------------------------------------------------------------*/
        ods exclude all;
        proc genmod data=new_sim descending;
                model Y = A2 A1 A0 / link=logit dist=binomial cl;
                weight stabw;
                ods output parameterestimates=psidata_out;
                run;
        ods exclude none;
        * transpose parameter estimates;
        proc transpose data=psidata_out out=psidata_wide;
                id Parameter;
                var Estimate StdErr LowerWaldCL UpperWaldCL;
                run;
        data estimate; set psidata_wide(drop=Intercept Scale); where _NAME_='Estimate'; run;
        data stderr;   set psidata_wide(drop=Intercept Scale); where _NAME_='StdErr'; run;
        data lcl;              set psidata_wide(drop=Intercept Scale); where _NAME_='LowerWaldCL'; run;
        data ucl;              set psidata_wide(drop=Intercept Scale); where _NAME_='UpperWaldCL'; run;
        data psidata;
                set estimate;
                if _N_=1 then set truepsi;
                if _N_=1 then set stderr(rename=(A0=A0_stderr A1=A1_stderr A2=A2_stderr));
                if _N_=1 then set lcl  (rename=(A0=A0_lcl A1=A1_lcl A2=A2_lcl));
                if _N_=1 then set ucl  (rename=(A0=A0_ucl A1=A1_ucl A2=A2_ucl));
                drop _NAME_ _LABEL_;
                run;
        data psidata;
                set psidata;
                sample=&sample;
                * CI coverage ;
                A0_CI_coverage=.; if A0_lcl<psi3 and A0_ucl>psi3 then A0_CI_coverage=1; else A0_CI_coverage=0;
                A1_CI_coverage=.; if A1_lcl<psi2 and A1_ucl>psi2 then A1_CI_coverage=1; else A1_CI_coverage=0;
                A2_CI_coverage=.; if A2_lcl<psi1 and A2_ucl>psi1 then A2_CI_coverage=1; else A2_CI_coverage=0;
                * pct bias ;
                format A1_pct_bias percentn. A2_pct_bias percentn. A0_pct_bias percentn.;
                A0_pct_bias=((A0-psi3)/psi3);
                A1_pct_bias=((A1-psi2)/psi2);
                A2_pct_bias=((A2-psi1)/psi1);
                run;
                data psifinal;
                set psifinal psidata;
                if sample ^= .;
        /*      GET THETAHAT ----------------------------------------------------------------------------------------*/
        ods exclude all;
        proc genmod data=new_sim descending;
                model Y = A2 L2 A1 L1 A0 / link=logit dist=binomial cl;
                ods output parameterestimates=simsample_out;
                run;
        ods exclude none;
        * transpose parameter estimates;
        proc transpose data=simsample_out out=simsample_wide;
```

```
            id Parameter;
            var Estimate StdErr LowerWaldCL UpperWaldCL;
            run;
        data estimate; set simsample_wide(drop=Intercept Scale); where _NAME_='Estimate'; run;
        data stderr;   set simsample_wide(drop=Intercept Scale); where _NAME_='StdErr'; run;
        data lcl;              set simsample_wide(drop=Intercept Scale); where _NAME_='LowerWaldCL'; run;
        data ucl;              set simsample_wide(drop=Intercept Scale); where _NAME_='UpperWaldCL'; run;
        data thetadata;
            set estimate;
            if _N_=1 then set truepsi;
            if _N_=1 then set stderr(rename=(A0=A0_stderr A1=A1_stderr A2=A2_stderr));
            if _N_=1 then set lcl  (rename=(A0=A0_lcl A1=A1_lcl A2=A2_lcl));
            if _N_=1 then set ucl  (rename=(A0=A0_ucl A1=A1_ucl A2=A2_ucl));
            drop _NAME_ _LABEL_;
            run;
        data thetadata;
            set thetadata;
            sample=&sample;
            *keep sample A2 A1 ;
            * CI coverage ;
            A0_CI_coverage=.; if A0_lcl<psi3 and A0_ucl>psi3 then A0_CI_coverage=1; else A0_CI_coverage=0;
            A1_CI_coverage=.; if A1_lcl<psi2 and A1_ucl>psi2 then A1_CI_coverage=1; else A1_CI_coverage=0;
            A2_CI_coverage=.; if A2_lcl<psi1 and A2_ucl>psi1 then A2_CI_coverage=1; else A2_CI_coverage=0;
            * pct bias ;
            format A1_pct_bias percentn. A2_pct_bias percentn. A0_pct_bias percentn.;
            A0_pct_bias=((A0-psi3)/psi3);
            A1_pct_bias=((A1-psi2)/psi2);
            A2_pct_bias=((A2-psi1)/psi1);
            run;
            data thetafinal;
            set thetafinal thetadata;
            if sample ^= .;
            run;
%mend;

%macro calculate_truth;
data truepsi; run;
data truepsi;
set truepsi;
 *      formula for psi0 --intercept of the MSM;
        psi0fracnum1=exp(&theta2 + &beta0) ;
        psi0fracden1=1+exp(&beta0);
        psi0frac1=psi0fracnum1/psi0fracden1;
        psi0frac2=1/psi0fracden1;
        psi0sum1=psi0frac1+psi0frac2;
        psi0fracnum3=exp(&theta4+&beta0);
        psi0frac3=psi0fracnum3/psi0fracden1;
        psi0sum2=psi0frac3+psi0frac2;
        exppsi0=exp(&theta0)*psi0sum1*psi0sum2;
        psi0=log(exppsi0);
 *      formula for psi1 -- coefficient of late life exposure ;
```

```
       psi1=&theta1;
 *     formula for psi2 -- coefficient of mid life exposure;
       psi2numfrac1=exp(&theta2 + &beta0 + &beta1) ;
       psi2denfrac1=1+exp(&beta0+&beta1);
       psi2frac1=psi2numfrac1/psi2denfrac1;
       psi2frac2=1/psi2denfrac1;
       psi2sum1=psi2frac1+psi2frac2;
       psi2frac=psi2sum1/psi0sum1;
       exppsi2=exp(&theta3)*psi2frac;
       psi2=log(exppsi2);
 *     formula for psi3 -- coefficient of early life exposure;
       psi3numfrac1=exp(&theta4 + &beta0 +&beta1);
       psi3frac1=psi3numfrac1/psi2denfrac1;
       psi3sum1=psi3frac1+psi2frac2;
       psi3frac=psi3sum1/psi0sum2;
       exppsi3=exp(&theta5)*psi3frac;
       psi3=log(exppsi3);
keep /*psi0*/ psi1 psi2 psi3;
run;
proc print data=truepsi;
run;
%mend calculate_truth;


%macro trueptrtrange;
data pat;
run;
data pat;
set pat;
       pat00=exp(&alpha0)/(1+exp(&alpha0));
       pat10=exp(&alpha0+&alpha1)/(1+exp(&alpha0+&alpha1));
       pat01=exp(&alpha0+&alpha2)/(1+exp(&alpha0+&alpha2));
       pat11=exp(&alpha0+&alpha1+&alpha2)/(1+exp(&alpha0+&alpha1+&alpha2));
run;
proc print data=pat;
run;
%mend;

*Scenario 7 No effect of mid- and late-life depression on AUD;
%samples(log=M7,theta0=-6,theta1=0.69, theta2=0.41, theta3=0.69, theta4=0.41, theta5=0.69, beta0=-.3, beta1=0.00, alpha0=0.00,
alpha1=0.41, alpha2=0.41, nsubjects=100000, nsamples=10000);

*****************************************************************************************************************************
Table 1 Scenario 8: Moderate positive unmeasured confounder of AUD and stroke
*****************************************************************************************************************************
options nonotes;

%macro samples(theta0=,theta1=,theta2=,theta3=,theta4=,theta5=,theta6=,beta0=,beta1=,beta2=,beta3=,alpha0=,alpha1=,alpha2=,omega0=,
nsubjects=,nsamples=);
       *calculates the true value of psi based on input choices of theta and beta;
       %calculate_truth;
```

```
%trueptrtrange;
*empty data set that will hold the nsamples IPW estimates of psi;
data psifinal;
run;
*empty data set that will hold the nsamples MLEs of theta;
data thetafinal;
run;
%let seed = 20;
%do sample = 1 %to &nsamples;
    %put sample = &sample;
     %let seed = %eval(&seed + 3);
     %put seed= &seed;
     %simulate;
     %estimate;
%end;
/*      PSI FINAL ---------------------------------------------------------- */
title "psihat means";
proc means data=psifinal n mean std STDERR clm;
var A2 A2_stderr A2_pct_bias A2_CI_coverage A1 A1_stderr A1_pct_bias A1_CI_coverage A0 A0_stderr A0_pct_bias A0_CI_coverage;
run;
%let true = psi1;
%let pred = A2;
%let dataset = psifinal;
proc sql;
       create table psi_mse_A2 as
       select mean((&pred.-&true.)**2) as MSE_A2, sqrt(mean((&pred.-&true.)**2)) as RMSE_A2 from &dataset.;
       quit;
%let true = psi2;
%let pred = A1;
%let dataset = psifinal;
proc sql;
       create table psi_mse_A1 as
       select mean((&pred.-&true.)**2) as MSE_A1, sqrt(mean((&pred.-&true.)**2)) as RMSE_A1 from &dataset.;
       quit;
%let true = psi3;
%let pred = A0;
%let dataset = psifinal;
proc sql;
       create table psi_mse_A0 as
       select mean((&pred.-&true.)**2) as MSE_A0, sqrt(mean((&pred.-&true.)**2)) as RMSE_A0 from &dataset.;
       quit;
title "psi hat MSE";
proc print data=psi_mse_A0; run;
proc print data=psi_mse_A1; run;
proc print data=psi_mse_A2; run;
/*      THETA FINAL -------------------------------------------------------- */
title 'thetahat means';
proc means data=thetafinal n mean std STDERR clm;
var A2 A2_stderr A2_pct_bias A2_CI_coverage A1 A1_stderr A1_pct_bias A1_CI_coverage A0 A0_stderr A0_pct_bias A0_CI_coverage;
run;
%let true = psi1;
```

```
	%let pred = A2;
	%let dataset = thetafinal;
	proc sql;
		create table theta_mse_A2 as
		select mean((&pred.-&true.)**2) as MSE_A2, sqrt(mean((&pred.-&true.)**2)) as RMSE_A2 from &dataset.;
		quit;
	%let true = psi2;
	%let pred = A1;
	%let dataset = thetafinal;
	proc sql;
		create table theta_mse_A1 as
		select mean((&pred.-&true.)**2) as MSE_A1, sqrt(mean((&pred.-&true.)**2)) as RMSE_A1 from &dataset.;
		quit;
	%let true = psi3;
	%let pred = A0;
	%let dataset = thetafinal;
	proc sql;
		create table theta_mse_A0 as
		select mean((&pred.-&true.)**2) as MSE_A0, sqrt(mean((&pred.-&true.)**2)) as RMSE_A0 from &dataset.;
		quit;
	title "theta hat MSE";
	proc print data=theta_mse_A0; run;
	proc print data=theta_mse_A1; run;
	proc print data=theta_mse_A2; run;
	title;
DM "log; clear; ";
%mend samples;


%macro simulate;
data simsample;
	do id= 1 to %eval(&nsubjects);
	call streaminit(&seed);
		* A0 = Early-life depression ------------------------------------------------;
		logitA0 = &alpha0 ;
		 pA0 = 1/(1+exp(-logitA0)) ;
		 uA0=rand('uniform');
		 if uA0<=pA0 then A0=1;
		 if uA0>pA0>. then A0=0;
		* W = common cause of early-life AUD and stroke -----------------------------;
		logitW0 = &omega0 ;
		 pW0 = 1/(1+exp(-logitW0)) ;
		 uW0=rand('uniform');
		 if uW0<=pW0 then W0=1;
		 if uW0>pW0>. then W0=0;
		* L1 = Early-life AUD -------------------------------------------------------;
		logitL1 = &beta0 + (&beta1 * A0) + (&beta2 * W0) + (&beta3 * W0 * A0);
				pL1= 1/(1+exp(-logitL1)) ;
				uL1=rand('uniform');
				if uL1<=pL1  then L1=1;
				if uL1>pL1>. then L1=0;
```

```
               * A1 = Mid-life depression ------------------------------------------------;
               logitA1 = &alpha0 + (&alpha1 * L1) + (&alpha2 * A0);
                pA1 = 1/(1+exp(-logitA1)) ;
                uA1=rand('uniform');
                if uA1<=pA1 then A1=1;
                if uA1>pA1>. then A1=0;
               * L2 = Mid-life AUD -------------------------------------------------------;
               logitL2 = &beta0 + (&beta1 * A1) ;
                             pL2= 1/(1+exp(-logitL2)) ;
                             uL2=rand('uniform');
                             if uL2<=pL2  then L2=1;
                             if uL2>pL2>. then L2=0;
               * A2 = Late-life depression -----------------------------------------------;
        logitA2 = &alpha0 + (&alpha1 * L2) + (&alpha2 * A1);
                pA2 = 1/(1+exp(-logitA2)) ;
                uA2=rand('uniform');
                if uA2<=pA2 then A2=1;
                if uA2>pA2>. then A2=0;
               * Y -----------------------------------------------------------------------;
               logitY = &theta0 + (&theta1 * A2) + (&theta2 * L2) + (&theta3 * A1)  + (&theta4 * L1)  + (&theta5 * A0)  + (&theta6 *
W0) ;
                pY = 1/(1+exp(-logitY)) ;
                uY=rand('uniform');
                if uY<=pY  then Y=1;
                if uY>pY>. then Y=0;
                      output;
       end; /*subjects*/
       keep id A0 L1 A1 L2 A2 Y ;
run;
%mend;

%macro estimate;
*numerator IPTW A1;
proc logistic NOPRINT descending data=simsample;
        model A1= A0;
        output out=modtn1 p=ptrt_n1;
        run;
*denominator IPTW A1;
proc logistic NOPRINT descending data=simsample;
        model A1=L1|A0 ;
        output out=modtd1 p=ptrt_d1;
        run;
*numerator IPTW A2;
proc logistic NOPRINT descending data=simsample;
        model A2= A1 ;
        output out=modtn2 p=ptrt_n2;
        run;
*denominator IPTW A2;
proc logistic NOPRINT descending data=simsample;
        model A2=L2 A1 ;
        output out=modtd2 p=ptrt_d2;
```

```
        run;
data new_sim;
        merge modtn1 modtd1 modtn2 modtd2;
        by id;
        /*calculate time 1 part of weight*/
        if A1=1 then do;
           pan1=ptrt_n1;
           pad1=ptrt_d1;
           end;
        else do;
           pan1=1-ptrt_n1;
           pad1=1-ptrt_d1;
           end;
        /*calculate time 2 part of weight*/
        if A2=1 then do;
           pan2=ptrt_n2;
           pad2=ptrt_d2;
           end;
         else do;
            pan2=1-ptrt_n2;
            pad2=1-ptrt_d2;
            end;
        /*calculate products and stabilized weight*/
        prod_n=pan1*pan2;
        prod_d=pad1*pad2;
        stabw=prod_n/prod_d;
        run;
        proc univariate data=new_sim noprint;
        var stabw;
        output out=stabw_univ skewness=stabw_skew kurtosis=stabw_kurtosis range=stabw_range min=stabw_min max=stabw_max pctlpre=stabw_p
pctlpts=0 to 5 by 0.25, 95 to 100 by 0.25;
        run;
/*      GET PSIHAT ------------------------------------------------------------------------------------------------------*/
        ods exclude all;
        proc genmod data=new_sim descending;
                model Y = A2 A1 A0 / link=logit dist=binomial cl;
                weight stabw;
                output out=psidata_pred predicted=y_pred;
                ods output parameterestimates=psidata_out;
                run;
        ods exclude none;
        proc print data=psidata_out; run;
        * transpose parameter estimates;
        proc transpose data=psidata_out out=psidata_wide;
                id Parameter;
                var Estimate StdErr LowerWaldCL UpperWaldCL;
                run;
        data estimate; set psidata_wide(drop=Intercept Scale); where _NAME_='Estimate'; run;
        data stderr;   set psidata_wide(drop=Intercept Scale); where _NAME_='StdErr'; run;
        data lcl;              set psidata_wide(drop=Intercept Scale); where _NAME_='LowerWaldCL'; run;
        data ucl;              set psidata_wide(drop=Intercept Scale); where _NAME_='UpperWaldCL'; run;
```

```
data psidata;
        set estimate;
        if _N_=1 then set truepsi;
        if _N_=1 then set stderr(rename=(A0=A0_stderr A1=A1_stderr A2=A2_stderr));
        if _N_=1 then set lcl  (rename=(A0=A0_lcl A1=A1_lcl A2=A2_lcl));
        if _N_=1 then set ucl  (rename=(A0=A0_ucl A1=A1_ucl A2=A2_ucl));
        if _N_=1 then set stabw_univ;
        drop _NAME_ _LABEL_;
        run;
data psidata;
        set psidata;
        sample=&sample;
        * CI coverage ;
        A0_CI_coverage=.; if A0_lcl<psi3 and A0_ucl>psi3 then A0_CI_coverage=1; else A0_CI_coverage=0;
        A1_CI_coverage=.; if A1_lcl<psi2 and A1_ucl>psi2 then A1_CI_coverage=1; else A1_CI_coverage=0;
        A2_CI_coverage=.; if A2_lcl<psi1 and A2_ucl>psi1 then A2_CI_coverage=1; else A2_CI_coverage=0;
        * pct bias ;
        format A1_pct_bias percentn. A2_pct_bias percentn. A0_pct_bias percentn.;
        A0_pct_bias=((A0-psi3)/psi3);
        A1_pct_bias=((A1-psi2)/psi2);
        A2_pct_bias=((A2-psi1)/psi1);
        run;
        data psifinal;
        set psifinal psidata;
        if sample ^= .;
/*      GET THETAHAT ------------------------------------------------------------------------------------------------*/
ods exclude all;
proc genmod data=new_sim descending;
        model Y = A2 L2 A1 L1 A0 / link=logit dist=binomial cl;
        output out=simsample_pred predicted=y_pred;
        ods output parameterestimates=simsample_out;
        run;
ods exclude none;
* transpose parameter estimates;
proc transpose data=simsample_out out=simsample_wide;
        id Parameter;
        var Estimate StdErr LowerWaldCL UpperWaldCL;
        run;
data estimate; set simsample_wide(drop=Intercept Scale); where _NAME_='Estimate'; run;
data stderr;   set simsample_wide(drop=Intercept Scale); where _NAME_='StdErr'; run;
data lcl;              set simsample_wide(drop=Intercept Scale); where _NAME_='LowerWaldCL'; run;
data ucl;              set simsample_wide(drop=Intercept Scale); where _NAME_='UpperWaldCL'; run;

data thetadata;
        set estimate;
        if _N_=1 then set truepsi;
        if _N_=1 then set stderr(rename=(A0=A0_stderr A1=A1_stderr A2=A2_stderr));
        if _N_=1 then set lcl  (rename=(A0=A0_lcl A1=A1_lcl A2=A2_lcl));
        if _N_=1 then set ucl  (rename=(A0=A0_ucl A1=A1_ucl A2=A2_ucl));
        if _N_=1 then set stabw_univ;
        drop _NAME_ _LABEL_;
```

```
		run;
	data thetadata;
		set thetadata;
		sample=&sample;
		* CI coverage ;
		A0_CI_coverage=.; if A0_lcl<psi3 and A0_ucl>psi3 then A0_CI_coverage=1; else A0_CI_coverage=0;
		A1_CI_coverage=.; if A1_lcl<psi2 and A1_ucl>psi2 then A1_CI_coverage=1; else A1_CI_coverage=0;
		A2_CI_coverage=.; if A2_lcl<psi1 and A2_ucl>psi1 then A2_CI_coverage=1; else A2_CI_coverage=0;
		* pct bias ;
		format A1_pct_bias percentn. A2_pct_bias percentn. A0_pct_bias percentn.;
		A0_pct_bias=((A0-psi3)/psi3);
		A1_pct_bias=((A1-psi2)/psi2);
		A2_pct_bias=((A2-psi1)/psi1);
		run;
		data thetafinal;
		set thetafinal thetadata;
		if sample ^= .;
		run;
%mend;

%macro calculate_truth;
data truepsi; run;
data truepsi;
set truepsi;
 *	formula for psi0 --intercept of the MSM;
	anum=1+exp(&theta2+&beta0)+exp(&theta4+&beta0)+exp(&theta2+&theta4+(2*&beta0));
	aden=(1+exp(&beta0))*(1+exp(&beta0))*(1+exp(&omega0));
	a= anum/aden;
	bnum=exp(&theta6+&omega0)+exp(&theta4+&theta6+&beta0+&beta2+&omega0)+exp(&theta2+&theta6+&beta0+&omega0)+exp(&theta2+&theta4+&t
heta6+(2*&beta0)+&beta2+&omega0);
	bden=(1+exp(&beta0))*(1+exp(&beta0+&beta2))*(1+exp(&omega0));
	b= bnum/bden;
	psi0=log(exp(&theta0)*(a/b));
 *	formula for psi1 -- coefficient of late life exposure (this does not change from the no W scenario);
	psi1=&theta1;
 *	formula for psi2 -- coefficient of mid life exposure;
	cnum=1+exp(&theta2+&beta0+&beta1)+exp(&theta4+&beta0)+exp(&theta2+&theta4+(2*&beta0)+&beta1);
	cden= (1+exp(&beta0+&beta1))*(1+exp(&beta0))*(1+exp(&omega0));
	c= cnum/cden;
	dnum=exp(&theta6+&omega0)+exp(&theta4+&theta6+&beta0+&beta2+&omega0)+exp(&theta2+&theta6+&beta0+&beta1+&omega0)+exp(&theta2+&th
eta4+&theta6+(2*&beta0)+&beta1+&beta2+&omega0);
	dden= (1+exp(&beta0+&beta1))*(1+exp(&beta0+&beta2))*(1+exp(&omega0));
	d= dnum/dden;
	psi2=log(exp(&theta3)*((c+d)/(a+b)));
 *	formula for psi3 -- coefficient of early life exposure;
	enum=1+exp(&theta2+&beta0)+exp(&theta4+&beta0+&beta1)+exp(&theta2+&theta4+(2*&beta0)+&beta1);
	eden=(1+exp(&beta0))*(1+exp(&beta0+&beta1))*(1+exp(&omega0));
	e=enum/eden;
	fnum=exp(&theta6+&omega0)+exp(&theta4+&theta6+&beta0+&beta1+&beta2+&beta3+&omega0)+exp(&theta2+&theta6+&beta0+&omega0)+exp(&the
ta2+&theta4+&theta6+(2*&beta0)+&beta1+&beta2+&beta3+&omega0);
	fden= (1+exp(&beta0))*(1+exp(&beta0+&beta1+&beta2+&beta3))*(1+exp(&omega0));
```

```
        f= fnum/fden;
        psi3= log(exp(&theta5)*((e+f)/(a+b)));
keep /*psi0*/ psi1 psi2 psi3;
run;
proc print data=truepsi;
run;
%mend calculate_truth;


%macro trueptrtrange;
data pat;
run;
data pat;
set pat;
        pat00=exp(&alpha0)/(1+exp(&alpha0));
        pat10=exp(&alpha0+&alpha1)/(1+exp(&alpha0+&alpha1));
        pat01=exp(&alpha0+&alpha2)/(1+exp(&alpha0+&alpha2));
        pat11=exp(&alpha0+&alpha1+&alpha2)/(1+exp(&alpha0+&alpha1+&alpha2));
run;
proc print data=pat;
run;
%mend;

*Scenario 8 - collider;
%samples(theta0=-6,theta1=0.69, theta2=0.41, theta3=0.58, theta4=0.41, theta5=0.58, theta6=-1.5, beta0=-.3, beta1=1.2, beta2=0.5,
beta3= 0.0, alpha0=0.00, alpha1=0.41, alpha2=0.41, omega0=0.25, nsubjects=100000, nsamples=10000);
```

****
SAS code for Table 2 and Supplemental Table 2. Difference in "true" versus estimated effects on the log odds scale under different theoretical life-course model when information regarding early-life exposure is missing

MACRO ORDER
- %samples
- %simulate
- %estimate

We generated data to assess possible differences in estimates obtained from conventional models versus MSMs estimated with IPW when information on early-life depression was not available. We assumed that depression approximately tripled odds of AUD, AUD increased odds of depression at the subsequent time point by 50%, and that AUD also doubled stroke risk. For each scenario, we approximated log linear models with conventional logistic models under the rare disease assumption to estimate the lifetime effect of depression on stroke, contrasting "depressed throughout life" with "not depressed at any point". We then re-estimated the lifetime effect using information only from mid- and late-life to mimic a real-world setting where follow-up started at mid-life. The percent bias was calculated comparing the estimated effects to the "true" effects on the log odds scale.

Table 2 key

EARLY-LIFE	MID-LIFE		LATE-LIFE	
true effect 3 theta5	true effect 2 theta3	estimated 2 etahat A1	true effect 1 theta1	estimated 1 etahat A2

**********;

```
options nonotes;

%macro samples(theta0=,theta1=,theta3=,theta5=,alpha0=,alpha1=,alpha2=,nsubjects=,nsamples=);
*       empty data set that will hold the nsamples MLEs of theta;
        data thetafinal;
        run;
        data etafinal;
        run;
        %let seed = 20;
        %do sample = 1 %to &nsamples;
            %put sample = &sample;
             %let seed = %eval(&seed + 3);
             %put seed= &seed;
             %simulate;
             %estimate;
        %end;
/*      THETA FINAL ------------------------------------------------------- */
        title 'thetahat means';
        proc means data=thetafinal stderr;
        var A1_A2;
        output out=theta_lifetime_stderr stderr=A1_A2_stderr;
        run;
        data thetafinal;
        set thetafinal;
```

```
	if _n_=1 then set theta_lifetime_stderr (drop=_type_ _freq_);
	A1_A2_lower = A1_A2 - (A1_A2_stderr*1.96);
	A1_A2_upper = A1_A2 + (A1_A2_stderr*1.96);
	if true_lifetime>A1_A2_lower and true_lifetime<A1_A2_upper then A1_A2_CI_coverage=1; else A1_A2_CI_coverage=0;
	run;
	proc means data=thetafinal n mean std STDERR clm;
	var A2 A2_stderr A2_pct_bias A2_CI_coverage A1 A1_stderr A1_pct_bias A1_CI_coverage A1_A2 A1_A2_pct_bias A1_A2_lower
A1_A2_upper true_lifetime A1_A2_CI_coverage ;
	run;
	%let true = &theta1;
	%let pred = A2;
	%let dataset = thetafinal;
	proc sql;
		create table theta_mse_A2 as
		select mean((&pred.-&true.)**2) as MSE_A2, sqrt(mean((&pred.-&true.)**2)) as RMSE_A2 from &dataset.;
		quit;
	%let true = &theta3;
	%let pred = A1;
	%let dataset = thetafinal;
	proc sql;
		create table theta_mse_A1 as
		select mean((&pred.-&true.)**2) as MSE_A1, sqrt(mean((&pred.-&true.)**2)) as RMSE_A1 from &dataset.;
		quit;
	%let true = true_lifetime;
	%let pred = A1_A2;
	%let dataset = thetafinal;
	proc sql;
		create table theta_mse_A1_A2 as
		select mean((&pred.-&true.)**2) as MSE_A1_A2, sqrt(mean((&pred.-&true.)**2)) as RMSE_A1_A2 from &dataset.;
		quit;
	title "theta hat MSE";
	proc print data=theta_mse_A2; run;
	proc print data=theta_mse_A1; run;
	proc print data=theta_mse_A1_A2; run;
/*	ETA FINAL ------------------------------------------------------------ */
	title 'etahat means';
	proc means data=etafinal stderr;
	var A1_A2;
	output out=eta_lifetime_stderr stderr=A1_A2_stderr;
	run;
	data etafinal;
	set etafinal;
	if _n_=1 then set eta_lifetime_stderr (drop=_type_ _freq_);
	A1_A2_lower = A1_A2 - (A1_A2_stderr*1.96);
	A1_A2_upper = A1_A2 + (A1_A2_stderr*1.96);
	if true_lifetime>A1_A2_lower and true_lifetime<A1_A2_upper then A1_A2_CI_coverage=1; else A1_A2_CI_coverage=0;
	run;
	proc means data=etafinal n mean std STDERR clm;
	var A2 A2_stderr A2_pct_bias A2_CI_coverage A1 A1_stderr A1_pct_bias A1_CI_coverage A1_A2 A1_A2_pct_bias A1_A2_lower
A1_A2_upper true_lifetime A1_A2_CI_coverage ;
	run;
```

```
        %let true = &theta1;
        %let pred = A2;
        %let dataset = etafinal;
        proc sql;
                create table eta_mse_A2 as
                select mean((&pred.-&true.)**2) as MSE_A2, sqrt(mean((&pred.-&true.)**2)) as RMSE_A2 from &dataset.;
                quit;
        %let true = &theta3;
        %let pred = A1;
        %let dataset = etafinal;
        proc sql;
                create table eta_mse_A1 as
                select mean((&pred.-&true.)**2) as MSE_A1, sqrt(mean((&pred.-&true.)**2)) as RMSE_A1 from &dataset.;
                quit;
        %let true = true_lifetime;
        %let pred = A1_A2;
        %let dataset = etafinal;
        proc sql;
                create table eta_mse_A1_A2 as
                select mean((&pred.-&true.)**2) as MSE_A1_A2, sqrt(mean((&pred.-&true.)**2)) as RMSE_A1_A2 from &dataset.;
                quit;
        title "eta hat MSE";
        proc print data=eta_mse_A2; run;
        proc print data=eta_mse_A1; run;
        proc print data=eta_mse_A1_A2; run;
%mend samples;


%macro simulate;
data simsample;
 *      generating simulation dataset ;
        do id= 1 to %eval(&nsubjects);
        call streaminit(&seed);
                * A0 = Early-life depression ------------------------------------------------;
                logitA0 = &alpha0 ;
                                pA0 = 1/(1+exp(-logitA0)) ;
                                uA0=rand('uniform');
                                if uA0<=pA0 then A0=1;
                                if uA0>pA0>. then A0=0;
                * A1 = Mid-life depression --------------------------------------------------;
                logitA1 = &alpha0 + (&alpha1 * A0);
                                pA1 = 1/(1+exp(-logitA1)) ;
                                uA1=rand('uniform');
                                if uA1<=pA1 then A1=1;
                                if uA1>pA1>. then A1=0;
                * A2 = Late-life depression -------------------------------------------------;
                logitA2 = &alpha0 + (&alpha1 * A0)+ (&alpha2 * A1);
                                pA2 = 1/(1+exp(-logitA2)) ;
                                uA2=rand('uniform');
                                if uA2<=pA2 then A2=1;
                                if uA2>pA2>. then A2=0;
```

```
                * Y ------------------------------------------------------------------------;
                logitY = &theta0 + (&theta1 * A2) + (&theta3 * A1) + (&theta5 * A0);
                                pY = 1/(1+exp(-logitY)) ;
                                uY=rand('uniform');
                                if uY<=pY  then Y=1;
                                if uY>pY>. then Y=0;
                output;
        end; /*subjects*/
        keep id A0 A1 A2 Y ;
run;
%mend;

%macro estimate;
/*      GET THETAHAT ---------------------------------------------------------------------------------------------*/
        ods exclude all;
        proc genmod data=simsample descending;
                model Y = A2 A1 / link=logit dist=binomial cl;
                output out=simsample_pred predicted=y_pred;
                ods output parameterestimates=simsample_out;
                run;
        ods exclude none;
*       transpose parameter estimates;
        proc transpose data=simsample_out out=simsample_wide;
                id Parameter;
                var Estimate StdErr LowerWaldCL UpperWaldCL;
                run;
        data estimate; set simsample_wide(drop=Intercept Scale); where _NAME_='Estimate'; run;
        data stderr;   set simsample_wide(drop=Intercept Scale); where _NAME_='StdErr'; run;
        data lcl;              set simsample_wide(drop=Intercept Scale); where _NAME_='LowerWaldCL'; run;
        data ucl;              set simsample_wide(drop=Intercept Scale); where _NAME_='UpperWaldCL'; run;
        data thetadata;
                set estimate;
                if _N_=1 then set stderr(rename=(A1=A1_stderr A2=A2_stderr));
                if _N_=1 then set lcl(rename=(A1=A1_lcl A2=A2_lcl));
                if _N_=1 then set ucl(rename=(A1=A1_ucl A2=A2_ucl));
                drop _NAME_ _LABEL_;
                run;
        data thetadata;
                set thetadata;
                sample=&sample;
                *keep sample A2 A1 ;
                A1_A2=A1+A2;
                true_lifetime=&theta1 + &theta3 + &theta5;
                * CI coverage ;
                A1_CI_coverage=.; if A1_lcl<&theta3. and A1_ucl>&theta3. then A1_CI_coverage=1; else A1_CI_coverage=0;
                A2_CI_coverage=.; if A2_lcl<&theta1. and A2_ucl>&theta1. then A2_CI_coverage=1; else A2_CI_coverage=0;
                * pct bias ;
                format A1_pct_bias percentn. A2_pct_bias percentn. A1_A2_pct_bias percentn.;
                A1_pct_bias=(A1/&theta3)-1;
                A2_pct_bias=(A2/&theta1)-1;
                A1_A2_pct_bias=((A1+A2)/(&theta3 + &theta1 + &theta5))-1;
```

```
               run;
               data thetafinal;
               set thetafinal thetadata;
               if sample ^= .;
               run;
/*      GET ETAHAT ------------------------------------------------------------------------------------------------*/
ods exclude all;
       proc genmod data=simsample descending;
              model Y = A2 A1 A0 / link=logit dist=binomial cl;
              output out=simsample_pred predicted=y_pred;
              ods output parameterestimates=simsample_out;
              run;
       ods exclude none;
 *     transpose parameter estimates;
       proc transpose data=simsample_out out=simsample_wide;
              id Parameter;
              var Estimate StdErr LowerWaldCL UpperWaldCL;
              run;
       data estimate; set simsample_wide(drop=Intercept Scale); where _NAME_='Estimate'; run;
       data stderr;   set simsample_wide(drop=Intercept Scale); where _NAME_='StdErr'; run;
       data lcl;              set simsample_wide(drop=Intercept Scale); where _NAME_='LowerWaldCL'; run;
       data ucl;              set simsample_wide(drop=Intercept Scale); where _NAME_='UpperWaldCL'; run;
       data etadata;
              set estimate;
              if _N_=1 then set mse_A1;
              if _N_=1 then set mse_A2;
              if _N_=1 then set stderr(rename=(A1=A1_stderr A2=A2_stderr));
              if _N_=1 then set lcl(rename=(A1=A1_lcl A2=A2_lcl));
              if _N_=1 then set ucl(rename=(A1=A1_ucl A2=A2_ucl));
              drop _NAME_ _LABEL_;
              run;
       data etadata;
              set etadata;
              sample=&sample;
              A1_A2=A1+A2;
              true_lifetime=&theta1 + &theta3 + &theta5;
              * CI coverage ;
              A1_CI_coverage=.; if A1_lcl<&theta3. and A1_ucl>&theta3. then A1_CI_coverage=1; else A1_CI_coverage=0;
              A2_CI_coverage=.; if A2_lcl<&theta1. and A2_ucl>&theta1. then A2_CI_coverage=1; else A2_CI_coverage=0;
              * pct bias ;
              format A1_pct_bias percentn. A2_pct_bias percentn. A1_A2_pct_bias percentn.;
              A1_pct_bias=(A1/&theta3)-1;
              A2_pct_bias=(A2/&theta1)-1;
              A1_A2_pct_bias=((A1+A2)/(&theta3 + &theta1 + &theta5))-1;
              run;
              data etafinal;
              set etafinal etadata;
              if sample ^= .;
              run;
%mend;
```

```
*      Accumulation model with equal effects of depression on stroke across the life-course;
       title 'Accumulation model with equal effects of depression on stroke across the life-course';
       %samples(theta0=-6,theta1=0.69,theta3=0.69,theta5=0.69,alpha0=0.00,alpha1=0.41,alpha2=0.41,nsubjects=100000,nsamples=10000);

*      Early life critical period;
       title 'Early life critical period';
       %samples(theta0=-6,theta1=0.00,theta3=0.00,theta5=1.10,alpha0=0.00,alpha1=0.41,alpha2=0.41,nsubjects=100000,nsamples=10000);

*      Pathway model;
       title 'Pathway model';
       %samples(theta0=-6,theta1=1.10,theta3=0.00,theta5=0.00,alpha0=0,alpha1=0.41,alpha2=0.41,nsubjects=100000,nsamples=10000);

*      Accumulation model with increasing effect across life-course;
       title 'Accumulation model with increasing effect across life-course';
       %samples(theta0=-6,theta1=0.92,theta3=0.69,theta5=0.41,alpha0=0.00,alpha1=0.41,alpha2=0.41,nsubjects=100000,nsamples=10000);

*      Accumulation model with decreasing effect across life-course;
       title 'Accumulation model with decreasing effect across life-course';
       %samples(theta0=-6,theta1=0.41,theta3=0.69,theta5=0.92,alpha0=0.00,alpha1=0.41,alpha2=0.41,nsubjects=100000,nsamples=10000);
```

```
*************************************************************************************************************************************************;
```
SAS code for Table 3 and Supplemental Table 3. Estimated effects of depression on stroke under an accumulation model, expressed on the log odds scale, comparing conventional regression and inverse probability weighting to estimate MSM parameters, in the presence of a late-study-start and time-varying confounders.
```
*************************************************************************************************************************************************;

options nonotes;

%macro samples(theta0=,theta1=,theta2=,theta3=,theta4=,theta5=,beta0=,beta1=,alpha0=,alpha1=,alpha2=,nsubjects=,nsamples=);
        *calculates the true value of psi based on input choices of theta and beta;
        %calculate_truth;
        %trueptrtrange;
        *empty data set that will hold the nsamples IPW estimates of psi;
        data psifinal;
        run;
        *empty data set that will hold the nsamples MLEs of theta;
        data thetafinal;
        run;
        *running simulations;
        %let seed = 20;
        %do sample = 1 %to &nsamples;
            %put sample = &sample;
             %let seed = %eval(&seed + 3);
             %put seed= &seed;
             %simulate;
             %estimate;
        %end;
/*      PSI FINAL ------------------------------------------------------------ */
        proc means data=psifinal stderr noprint;
        var A1_A2;
        output out=psi_lifetime_stderr stderr=A1_A2_stderr;
        run;
        data psifinal;
        set psifinal;
        if _n_=1 then set psi_lifetime_stderr (drop=_type_ _freq_);
        A1_A2_lower = A1_A2 - (A1_A2_stderr*1.96);
        A1_A2_upper = A1_A2 + (A1_A2_stderr*1.96);
        if true_lifetime>A1_A2_lower and true_lifetime<A1_A2_upper then A1_A2_CI_coverage=1; else A1_A2_CI_coverage=0;
        run;
        title "psihat means";
        proc means data=psifinal n mean std STDERR clm;
        var A2 A2_stderr A2_pct_bias A2_CI_coverage A1 A1_stderr A1_pct_bias A1_CI_coverage A1_A2 A1_A2_stderr A1_A2_pct_bias
A1_A2_lower A1_A2_upper true_lifetime A1_A2_CI_coverage;
        run;
        %let true = psi1;
        %let pred = A2;
        %let dataset = psifinal;
        proc sql;
                create table psi_mse_A2 as
                select mean((&pred.-&true.)**2) as MSE_A2, sqrt(mean((&pred.-&true.)**2)) as RMSE_A2 from &dataset.;
                quit;
```

```
	%let true = psi2;
	%let pred = A1;
	%let dataset = psifinal;
	proc sql;
		create table psi_mse_A1 as
		select mean((&pred.-&true.)**2) as MSE_A1, sqrt(mean((&pred.-&true.)**2)) as RMSE_A1 from &dataset.;
		quit;
	%let true = true_lifetime;
	%let pred = A1_A2;
	%let dataset = psifinal;
	proc sql;
		create table psi_mse_A1_A2 as
		select mean((&pred.-&true.)**2) as MSE_A1_A2, sqrt(mean((&pred.-&true.)**2)) as RMSE_A1_A2 from &dataset.;
		quit;
	title "psi hat MSE";
	proc print data=psi_mse_A1; run;
	proc print data=psi_mse_A2; run;
	proc print data=psi_mse_A1_A2; run;
/*	THETA FINAL ------------------------------------------------------- */
	proc means data=thetafinal stderr noprint;
	var A1_A2;
	output out=theta_lifetime_stderr stderr=A1_A2_stderr;
	run;
	data thetafinal;
	set thetafinal;
	if _n_=1 then set theta_lifetime_stderr (drop=_type_ _freq_);
	A1_A2_lower = A1_A2 - (A1_A2_stderr*1.96);
	A1_A2_upper = A1_A2 + (A1_A2_stderr*1.96);
	if true_lifetime>A1_A2_lower and true_lifetime<A1_A2_upper then A1_A2_CI_coverage=1; else A1_A2_CI_coverage=0;
	run;
	title 'thetahat means';
	proc means data=thetafinal n mean std STDERR clm;
	var A2 A2_stderr A2_pct_bias A2_CI_coverage A1 A1_stderr A1_pct_bias A1_CI_coverage A1_A2 A1_A2_stderr A1_A2_pct_bias
A1_A2_lower A1_A2_upper true_lifetime A1_A2_CI_coverage;
	run;
	%let true = psi1;
	%let pred = A2;
	%let dataset = thetafinal;
	proc sql;
		create table theta_mse_A2 as
		select mean((&pred.-&true.)**2) as MSE_A2, sqrt(mean((&pred.-&true.)**2)) as RMSE_A2 from &dataset.;
		quit;
	%let true = psi2;
	%let pred = A1;
	%let dataset = thetafinal;
	proc sql;
		create table theta_mse_A1 as
		select mean((&pred.-&true.)**2) as MSE_A1, sqrt(mean((&pred.-&true.)**2)) as RMSE_A1 from &dataset.;
		quit;
	%let true = true_lifetime;
	%let pred = A1_A2;
```

```
	%let dataset = thetafinal;
	proc sql;
		create table theta_mse_A1_A2 as
		select mean((&pred.-&true.)**2) as MSE_A1_A2, sqrt(mean((&pred.-&true.)**2)) as RMSE_A1_A2 from &dataset.;
		quit;
	title "theta hat MSE";
	proc print data=theta_mse_A1; run;
	proc print data=theta_mse_A2; run;
	proc print data=theta_mse_A1_A2; run;
	title;
DM "log; clear; ";
%mend samples;


%macro simulate;
data simsample;
	do id= 1 to %eval(&nsubjects);
	call streaminit(&seed);
		* A0 = Early-life depression ------------------------------------------------;
		logitA0 = &alpha0 ;
		 pA0 = 1/(1+exp(-logitA0)) ;
		 uA0=rand('uniform');
		 if uA0<=pA0 then A0=1;
		 if uA0>pA0>. then A0=0;
		* L1 = Early-life AUD -------------------------------------------------------;
		logitL1 = &beta0 + (&beta1 * A0) ;
		 pL1= 1/(1+exp(-logitL1)) ;
		 uL1=rand('uniform');
		 if uL1<=pL1  then L1=1;
		 if uL1>pL1>. then L1=0;
		* A1 = Mid-life depression --------------------------------------------------;
		logitA1 = &alpha0 + (&alpha1 * L1) + (&alpha2 * A0);
		 pA1 = 1/(1+exp(-logitA1)) ;
		 uA1=rand('uniform');
		 if uA1<=pA1 then A1=1;
		 if uA1>pA1>. then A1=0;
		* L2 = Mid-life AUD ---------------------------------------------------------;
		logitL2 = &beta0 + (&beta1 * A1) ;
		 pL2= 1/(1+exp(-logitL2)) ;
		 uL2=rand('uniform');
		 if uL2<=pL2  then L2=1;
		 if uL2>pL2>. then L2=0;
		* A2 = Late-life depression -------------------------------------------------;
		logitA2 = &alpha0 + (&alpha1 * L2) + (&alpha2 * A1);
		 pA2 = 1/(1+exp(-logitA2)) ;
		 uA2=rand('uniform');
		 if uA2<=pA2 then A2=1;
		 if uA2>pA2>. then A2=0;
		* Y -------------------------------------------------------------------------;
		logitY = &theta0 + (&theta1 * A2) + (&theta2 * L2) + (&theta3 * A1)  + (&theta4 * L1)  + (&theta5 * A0)    ;
				pY = 1/(1+exp(-logitY)) ;
```

```
                              uY=rand('uniform');
                              if uY<=pY  then Y=1;
                              if uY>pY>. then Y=0;
       output;
       end; /*subjects*/
keep id A0 L1 A1 L2 A2 Y ;
run;
%mend;

%macro estimate;
*numerator IPTW A1;
proc logistic NOPRINT descending data=simsample;
model A1= /*A0*/;
output out=modtn1 p=ptrt_n1;
run;
*denominator IPTW A1;
proc logistic NOPRINT descending data=simsample;
model A1=L1 /*A0*/ ;
output out=modtd1 p=ptrt_d1;
run;
*numerator IPTW A2;
proc logistic NOPRINT descending data=simsample;
model A2= A1 ;
output out=modtn2 p=ptrt_n2;
run;
*denominator IPTW A2;
proc logistic NOPRINT descending data=simsample;
model A2=L2 A1 ;
output out=modtd2 p=ptrt_d2;
run;
data new_sim;
merge modtn1 modtd1 modtn2 modtd2;
by id;
/*calculate time 1 part of weight*/
if A1=1 then do;
pan1=ptrt_n1;
pad1=ptrt_d1;
end;
else do;
pan1=1-ptrt_n1;
pad1=1-ptrt_d1;
end;
/*calculate time 2 part of weight*/
if A2=1 then do;
pan2=ptrt_n2;
pad2=ptrt_d2;
end;
else do;
pan2=1-ptrt_n2;
pad2=1-ptrt_d2;
end;
```

```
/*calculate products and stabilized weight*/
prod_n=pan1*pan2;
prod_d=pad1*pad2;
stabw=prod_n/prod_d;
run;
/*      GET PSIHAT ------------------------------------------------------------------------------------------------*/
        ods exclude all;
        proc genmod data=new_sim descending;
                model Y = A2 A1 /* A0 */ / link=logit dist=binomial cl;
                weight stabw;
                ods output parameterestimates=psidata_out;
                run;
        ods exclude none;
        * transpose parameter estimates;
        proc transpose data=psidata_out out=psidata_wide;
                id Parameter;
                var Estimate StdErr LowerWaldCL UpperWaldCL;
                run;
        data estimate; set psidata_wide(drop=Intercept Scale); where _NAME_='Estimate'; run;
        data stderr;   set psidata_wide(drop=Intercept Scale); where _NAME_='StdErr'; run;
        data lcl;              set psidata_wide(drop=Intercept Scale); where _NAME_='LowerWaldCL'; run;
        data ucl;              set psidata_wide(drop=Intercept Scale); where _NAME_='UpperWaldCL'; run;
        data psidata;
                set estimate;
                if _N_=1 then set truepsi;
                if _N_=1 then set stderr(rename=(A1=A1_stderr A2=A2_stderr));
                if _N_=1 then set lcl  (rename=(A1=A1_lcl A2=A2_lcl));
                if _N_=1 then set ucl  (rename=(A1=A1_ucl A2=A2_ucl));
                drop _NAME_ _LABEL_;
                run;
        data psidata;
                set psidata;
                sample=&sample;
                true_lifetime=psi1+psi2+psi3;
                A1_A2 = A1+A2;
                * CI coverage ;
                A1_CI_coverage=.; if A1_lcl<psi2 and A1_ucl>psi2 then A1_CI_coverage=1; else A1_CI_coverage=0;
                A2_CI_coverage=.; if A2_lcl<psi1 and A2_ucl>psi1 then A2_CI_coverage=1; else A2_CI_coverage=0;
                * pct bias ;
                format A1_pct_bias percentn. A2_pct_bias percentn. ;
                A1_pct_bias=((A1-psi2)/psi2);
                A2_pct_bias=((A2-psi1)/psi1);
                A1_A2_pct_bias=((A1_A2-true_lifetime)/true_lifetime);
                run;
                data psifinal;
                set psifinal psidata;
                if sample ^= .;
/*      GET THETAHAT ----------------------------------------------------------------------------------------------*/
ods exclude all;
        proc genmod data=new_sim descending;
                 model Y = A2 L2 A1 L1 /*A0*/ / link=logit dist=binomial cl;
```

```
		ods output parameterestimates=simsample_out;
		run;
	ods exclude none;
	* transpose parameter estimates;
	proc transpose data=simsample_out out=simsample_wide;
		id Parameter;
		var Estimate StdErr LowerWaldCL UpperWaldCL;
		run;
	data estimate; set simsample_wide(drop=Intercept Scale); where _NAME_='Estimate'; run;
	data stderr;   set simsample_wide(drop=Intercept Scale); where _NAME_='StdErr'; run;
	data lcl;              set simsample_wide(drop=Intercept Scale); where _NAME_='LowerWaldCL'; run;
	data ucl;              set simsample_wide(drop=Intercept Scale); where _NAME_='UpperWaldCL'; run;
	data thetadata;
		set estimate;
		if _N_=1 then set truepsi;
		if _N_=1 then set stderr(rename=(A1=A1_stderr A2=A2_stderr));
		if _N_=1 then set lcl  (rename=(A1=A1_lcl A2=A2_lcl));
		if _N_=1 then set ucl  (rename=(A1=A1_ucl A2=A2_ucl));
		drop _NAME_ _LABEL_;
		run;
	data thetadata;
		set thetadata;
		sample=&sample;
		*keep sample A2 A1 ;
		true_lifetime=psi1+psi2+psi3;
		A1_A2 = A1+A2;
		* CI coverage ;
		A1_CI_coverage=.; if A1_lcl<psi2 and A1_ucl>psi2 then A1_CI_coverage=1; else A1_CI_coverage=0;
		A2_CI_coverage=.; if A2_lcl<psi1 and A2_ucl>psi1 then A2_CI_coverage=1; else A2_CI_coverage=0;
		* pct bias ;
		format A1_pct_bias percentn. A2_pct_bias percentn. ;
		A1_pct_bias=((A1-psi2)/psi2);
		A2_pct_bias=((A2-psi1)/psi1);
		A1_A2_pct_bias=((A1_A2-true_lifetime)/true_lifetime);
		run;
		data thetafinal;
		set thetafinal thetadata;
		if sample ^= .;
		run;
%mend;

%macro calculate_truth;
data truepsi;
run;
data truepsi;
set truepsi;
 *      formula for psi0 --intercept of the MSM;
	psi0fracnum1=exp(&theta2+&beta0);
	psi0fracden1=1+exp(&beta0);
	psi0frac1=psi0fracnum1/psi0fracden1;
	psi0frac2=1/psi0fracden1;
```

```
       psi0sum1=psi0frac1+psi0frac2;
       psi0fracnum3=exp(&theta4+&beta0);
       psi0frac3=psi0fracnum3/psi0fracden1;
       psi0sum2=psi0frac3+psi0frac2;
       exppsi0=exp(&theta0)*psi0sum1*psi0sum2;
       psi0=log(exppsi0);
 *     formula for psi1 -- coefficient of late life exposure;
       psi1=&theta1;
 *     formula for psi2 -- coefficient of mid life exposure;
       psi2numfrac1=exp(&theta2+&beta0+&beta1);
       psi2denfrac1=1+exp(&beta0+&beta1);
       psi2frac1=psi2numfrac1/psi2denfrac1;
       psi2frac2=1/psi2denfrac1;
       psi2sum1=psi2frac1+psi2frac2;
       psi2frac=psi2sum1/psi0sum1;
       exppsi2=exp(&theta3)*psi2frac;
       psi2=log(exppsi2);
 *     formula for psi3 -- coefficient of early life exposure;
       psi3numfrac1=exp(&theta4+&beta0+&beta1);
       psi3frac1=psi3numfrac1/psi2denfrac1;
       psi3sum1=psi3frac1+psi2frac2;
       psi3frac=psi3sum1/psi0sum2;
       exppsi3=exp(&theta5)*psi3frac;
       psi3=log(exppsi3);
keep /*psi0*/ psi1 psi2 psi3;
run;
proc print data=truepsi;
run;
%mend calculate_truth;

%macro trueptrtrange;
data pat;
run;
data pat;
set pat;
pat00=exp(&alpha0)/(1+exp(&alpha0));
pat10=exp(&alpha0+&alpha1)/(1+exp(&alpha0+&alpha1));
pat01=exp(&alpha0+&alpha2)/(1+exp(&alpha0+&alpha2));
pat11=exp(&alpha0+&alpha1+&alpha2)/(1+exp(&alpha0+&alpha1+&alpha2));
run;
proc print data=pat;
run;
%mend;

*Late-study-start scenario with time-varying confounding under the accumulation model;
%samples(theta0=-6,theta1=0.69, theta2=0.69, theta3=0.51, theta4=0.69, theta5=0.51, beta0=-.3, beta1=1.2, alpha0=0.00, alpha1=0.41,
alpha2=0.41, nsubjects=100000, nsamples=10000);
```