ASSESSING THE TREATMENT EFFECT HETEROGENEITY WITH A LATENT VARIABLE

Yunjian Yin[1,2], Lan Liu[2,*], and Zhi Geng[1]

[1]*School of Mathematical Sciences, Peking University, Beijing 100871, China*
[2]*School of Statistics, University of Minnesota, Minneapolis, Minnesota 55455, USA*

Appendix A Proof of Formulas

Proof of Formulas (F1) and (F2)

$$
\begin{aligned}
& \mathrm{TBR}_c(x) = P(Y_1 - Y_0 > c|X = x) \\
= \; & P\big((\alpha_{1,0}-\alpha_{0,0}) + (\alpha_{1,1}-\alpha_{0,1})^T x + (\alpha_{1,2}-\alpha_{0,2})U + (\alpha_{1,3}-\alpha_{0,3})^T xU + (\epsilon_1 - \epsilon_0\big) > c) \\
= \; & P\Big(\frac{\epsilon_0-\epsilon_1}{\sqrt{\sigma_0^2+\sigma_1^2}} < \frac{(\alpha_{1,0}-\alpha_{0,0}) + (\alpha_{1,1}-\alpha_{0,1})^T x + (\alpha_{1,2}-\alpha_{0,2})U + (\alpha_{1,3}-\alpha_{0,3})^T xU - c}{\sqrt{\sigma_0^2+\sigma_1^2}}\Big) \\
= \; & \int \Phi\Big(\frac{(\alpha_{1,0}-\alpha_{0,0}) + (\alpha_{1,1}-\alpha_{0,1})^T x + (\alpha_{1,2}-\alpha_{0,2})u + (\alpha_{1,3}-\alpha_{0,3})^T xu - c}{\sqrt{\sigma_0^2+\sigma_1^2}}\Big) f_U(u)du \\
= \; & \int \Phi\Big(\big(w_1 + w_2 u\big)/w_3\Big) f_U(u)du \\
= \; & \int \int_{-\infty}^{(w_1+w_2u)/w_3} \frac{1}{\sqrt{2\pi}} \exp\big(-s^2/2\big) f_U(u) ds du \\
= \; & \int \int_{-\infty}^{0} \frac{1}{2\pi} \exp\Big[-\frac{1}{2w_3^2}\Big\{(w_2^2+w_3^2)\big(u + \frac{w_2(w_3 s + w_1)}{w_2^2+w_3^2}\big)^2 + \frac{w_3^2(w_3 s + w_1)^2}{w_2^2+w_3^2}\Big\}\Big] ds du \\
= \; & \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi}} \sqrt{\frac{w_3^2}{w_2^2+w_3^2}} \exp\Big\{-\frac{(w_3 s + w_1)^2}{2(w_2^2+w_3^2)}\Big\} ds \\
= \; & \Phi\Big(\frac{w_1}{\sqrt{w_2^2+w_3^2}}\Big),
\end{aligned}
$$

where $f_U(\cdot)$ is the density functions of U, $w_1 = (\alpha_{1,0}-\alpha_{0,0}) + (\alpha_{1,1}-\alpha_{0,1})^T x - c, w_2 = (\alpha_{1,2}-\alpha_{0,2}) + (\alpha_{1,3}-\alpha_{0,3})^T x, w_3 = \sqrt{\sigma_0^2+\sigma_1^2}$. Similarly, we can derive the form for $\mathrm{THR}_c(x)$.

Proof of Formulas (F3) and (F4)

Let $K(\alpha_t,x,u)=\alpha_{t,0}+\alpha_{t,1}^Tx+\alpha_{t,2}u+\alpha_{t,3}^Txu$, we have

$$\begin{aligned}
&\mathrm{TBR}(x)=\int\big\{1-\Phi\big(K(\alpha_0,x,u)\big)\big\}\Phi\big(K(\alpha_1,x,u)\big)f_U(u)du\\
=&\int\Big\{\int_{K(\alpha_0,x,u)}^{\infty}\frac{1}{\sqrt{2\pi}}\exp(-s_0^2/2)ds_0\Big\}\Big\{\int_{-\infty}^{K(\alpha_1,x,u)}\frac{1}{\sqrt{2\pi}}\exp(-s_1^2/2)ds_1\Big\}f_U(u)du\\
=&\int\int_{-\infty}^{0}\int_{0}^{\infty}\frac{1}{(2\pi)^{3/2}}\exp\Big\{-\frac{(s_0+K(\alpha_0,x,u))^2+(s_1+K(\alpha_1,x,u))^2+u^2}{2}\Big\}ds_0ds_1du.
\end{aligned}$$

Let $K_1(\alpha_t,x)=\alpha_{t,0}+\alpha_{t,1}^Tx$, $K_2(\alpha_t,x)=\alpha_{t,2}+\alpha_{t,3}^Tx$, thus $K(\alpha_t,x,u)=K_1(\alpha_t,x)+uK_2(\alpha_t,x)$.

Then

$$\begin{aligned}
&\{s_0+K(\alpha_0,x,u)\}^2+\{s_1+K(\alpha_1,x,u)\}^2+u^2\\
=&\big\{1+K_2(\alpha_0,x)^2+K_2(\alpha_1,x)^2\big\}u^2+2\big\{(s_0+K_1(\alpha_0,x))K_2(\alpha_0,x)+(s_1+K_1(\alpha_1,x))K_2(\alpha_1,x)\big\}u\\
&+\{s_0+K_1(\alpha_0,x)\}^2+\{s_1+K_1(\alpha_1,x)\}^2\\
=&\big\{1+K_2(\alpha_0,x)^2+K_2(\alpha_1,x)^2\big\}\big\{u+\frac{(s_0+K_1(\alpha_0,x))K_2(\alpha_0,x)+(s_1+K_1(\alpha_1,x))K_2(\alpha_1,x)}{1+K_2(\alpha_0,x)^2+K_2(\alpha_1,x)^2}\big\}^2\\
&+\frac{1}{1+K_2(\alpha_0,x)^2+K_2(\alpha_1,x)^2}\Big[\{s_0+K_1(\alpha_0,x)\}^2\{1+K_2(\alpha_1,x)^2\}\\
&\qquad+\{s_1+K_1(\alpha_1,x)\}^2\{1+K_2(\alpha_0,x)^2\}\\
&\qquad-2\{s_0+K_1(\alpha_0,x)\}K_2(\alpha_0,x)\{s_1+K_1(\alpha_1,x)\}K_2(\alpha_1,x)\Big].
\end{aligned}$$

So

$$\mathrm{TBR}(x)=\int_0^{\infty}\int_{-\infty}^{0}\frac{1}{(2\pi)S}\exp\Big(-\frac{F}{2}\Big)ds_0ds_1, \tag{A1.1}$$

where $S^2=1+K_2(\alpha_0,x)^2+K_2(\alpha_1,x)^2$,

$$\begin{aligned}
F\;=\;&\Big[\{s_0+K_1(\alpha_0,x)\}^2\{1+K_2(\alpha_1,x)^2\}+\{s_1+K_1(\alpha_1,x)\}^2\{1+K_2(\alpha_0,x)^2\}\\
&-2\{s_0+K_1(\alpha_0,x)\}K_2(\alpha_0,x)\{s_1+K_1(\alpha_1,x)\}K_2(\alpha_1,x)\Big]/S^2\\
=\;&\Big\{(s_0,s_1)-\mu\Big\}\Sigma^{-1}\Big\{(s_0,s_1)-\mu\Big\}^T,
\end{aligned}$$

$$\mu=(-K_1(\alpha_0,x),\ -K_1(\alpha_1,x)),$$

$$\Sigma = \begin{pmatrix} 1 + K_2(\alpha_0, x)^2 & K_2(\alpha_0, x)K_2(\alpha_1, x) \\ K_2(\alpha_0, x)K_2(\alpha_1, x) & 1 + K_2(\alpha_1, x)^2 \end{pmatrix}.$$

Thus, $\mathrm{TBR}(x) = \Phi_2\big((0,\infty), (-\infty, 0); \mu, \Sigma\big)$, where $\Phi_2(A_0, A_1; \mu, \Sigma)$ is the distribution function of bivariate normal vector with mean μ, covariance matrix Σ and integral region $A_0 \times A_1$. Similarly, we can derive the form for $\mathrm{THR}(x)$.

Relationship of $\mathrm{ATE}(x)$, $\mathrm{TBR}_c(x)$ and $\mathrm{THR}_c(x)$

Note for any random variable Z, we have

$$E(Z) = \int_0^\infty \{1 - F_Z(z)\}dz - \int_{-\infty}^0 F_Z(z)dz,$$

where F_Z is the cumulative distribution function of Z. Thus,

$$\begin{aligned}
\mathrm{ATE}(x) &= E(Y_1 - Y_0 | X = x) \\
&= \int_0^\infty \{1 - F_{Y_1 - Y_0|x}(c)\}dc - \int_{-\infty}^0 F_{Y_1-Y_0|x}(c)dc \\
&= \int_0^\infty \mathrm{TBR}_c(x)dc - \int_{-\infty}^0 \{1 - \mathrm{TBR}_c(x)\}dc \\
&= \int_0^\infty \mathrm{TBR}_c(x)dc - \int_{-\infty}^0 \mathrm{THR}_{-c}(x)dc \\
&= \int_0^\infty \{\mathrm{TBR}_c(x) - \mathrm{THR}_c(x)\}dc,
\end{aligned}$$

where the penultimate step holds since $Y_1 - Y_0$ is continuous.

Appendix B Proof of Theorem 1

Instead of proving Theorem 1 directly, we first provide sufficient and necessary identification conditions of $(g_t(X); h_t(X))$ in the general models (3) and (4).

Theorem B.1. *Under Assumption 2,*

(i) *When the outcome is continuous, if the following model (A2.2) holds for t=0,1,*

$$\begin{cases} Y_t = g_t(X) + Uh_t(X) + \epsilon_t, \\ \epsilon_t \perp (X,U), \epsilon_t \sim N(0,\sigma_t^2), U \sim N(0,1),\ \ h_t(0) > 0, \end{cases} \tag{A2.2}$$

then the following Condition A is the sufficient and necessary condition to identify $(g_0(X), h_0(X), \sigma_0^2, g_1(X), h_1(X), \sigma_1^2)$

Condition A. $h_t(X)$ *belongs to the family* $\mathcal{S}(X)$ *for* $t = 0,1$, *where*

$$\mathcal{S}(X) = \big\{h(X) : h(X) \text{ can be identified if } h(X)h'(X) \text{ is known.}\big\}$$

(ii) *When the outcome is continuous, if the following model (A2.3) holds for t=0,1,*

$$\begin{cases} Y_t^* = g_t(X) + Uh_t(X) + \epsilon_t, \\ Y_t = I(Y_t^* > 0), \\ \epsilon_t \perp (X,U), \epsilon_t \sim N(0,\sigma^2), U \sim N(0,1),\ \ h_t(0) > 0, \end{cases} \tag{A2.3}$$

then the following Condition B is the sufficient and necessary condition to identify $(g_0(X), h_0(X), g_1(X), h_1(X))$.

Condition B. $(g_t(X), h_t(X))$ *belongs to the family* $\big(\mathcal{S}_1(X), \mathcal{S}_2(X)\big)$ *for* $t = 0,1$,

where

$$\begin{aligned} &\big(\mathcal{S}_1(X), \mathcal{S}_2(X)\big) \\ = \ &\Big\{\big(g(X;\alpha_1), h(X;\alpha_2)\big)\big|(\alpha_1,\alpha_2) \in \mathcal{A}, \forall(\alpha_1,\alpha_2) \neq (\beta_1,\beta_2) \in \mathcal{A}, \frac{g(X;\alpha_1)}{\sqrt{1+h^2(X;\alpha_2)}} \neq \frac{g(X;\beta_1)}{\sqrt{1+h^2(X;\beta_2)}}\Big\}. \end{aligned}$$

Proof.

(i) Since $E[Y|X,T=t] = E[Y_t|X] = g_t(X)$, we can identify $g_t(X)$ and we have

$$\big(Y - g_t(X)\big)\big|\big(X, T=t\big) \sim N(0, h_t^2(X) + \sigma_t^2).$$

Thus $A_t(X) = h_t^2(X) + \sigma_t^2$ can also be identified, so is $A_t'(X) = h_t(X)h_t'(X)$.

Next we show that Condition A is sufficient and necessary to identify $h_t(x), t = 0,1$. It is easy to see that if $h_t(X)$ belongs to $\mathcal{S}(X)$, then $h_t(X)$ is also identified. On the other hand, if $h_t(X)$ does not belong to $\mathcal{S}(X)$, then $h_t(X)$ can not be decided uniquely from $h_t(X)h_t'(X)$. Besides, knowing $h_t(X)h_t'(X)$ is equivalent to knowing $h_t^2(X)$ up to a constant, i.e., $h_t^2(X_1) - h_t^2(X_2)$ for all X_1, X_2. Note that $\Big(Y - g_t(X)\Big)\Big|(X, T = t) \sim N(0, h_t^2(X) + \sigma_t^2)$, the distribution of $Y - g_t(X)$ condition on $(X, T = t)$ is determined by the variance, so all the information we have about $h_t(X)$ is $h_t^2(X) + \sigma_t^2$, which is the same as knowing $h_t^2(X_1) - h_t^2(X_2)$ for all X_1, X_2. Thus, we can not identify $h_t(X)$. So the sufficient and necessary condition is that $h_t(X) \in \mathcal{S}(X)$ for $t = 0, 1$.

(ii) Since $P(Y = 1|X, U, T = t) = \Phi\big(g_t(X) + Uh_t(X)\big)$, we have

$$P(Y = 1|X, T = t) = \Phi\Big(\frac{g_t(X)}{\sqrt{1 + h_t^2(X)}}\Big),$$

It is easy to see that $(g_0(X), h_0(X), g_1(X), h_1(X))$ can be identified if and only if the Condition B holds. □

The identification of heterogeneous treatment effects given in Theorem 1 follows from the following corollaries.

Corollary 1. *When $h(X) = h(X; \eta) = \eta_0 + \eta_1^T X$, where $\eta = (\eta_0, \eta_1^T)^T, \eta_1 = \big(\eta_{1,1}, \cdots, \eta_{1,p}\big)^T$ and $\eta_0 > 0$, we have $h(X) \in \mathcal{S}$.*

Proof. Since $h(X)h'(X) = \big(\eta_0 + \eta_1^T X\big)\eta_1 = \eta_0\eta_1 + \eta_1\eta_1^T X$, we can identify $(\eta_0\eta_1, \eta_1\eta_1^T)$ if $h(X)h'(X)$ is known. Besides, $h(0) = \eta_0 > 0$, so the sign of every component of η_1 can be determined since we know $\eta_0\eta_1$. Then η_1 can be identified since we know the diagonal elements of $\eta_1\eta_1^T$. Then η_0 can also be identified from $\eta_0\eta_1$. Thus (η_0, η_1) is identifiable, so is $h(X)$. This completes the proof of the part (i) in Theorem 1. □

We impose the following regularity condition on $\mathcal{X}$ which is the domain of X.

Condition C. *There exists linear independent* $(\tau_1, \cdots, \tau_p) \subset \mathcal{X}$*, where* $\mathcal{X}$ *is the domain of* X*, s.t.* $P(Y = 1|X = \tau_i) = 0, i = 1, \ldots, p$.

Corollary 2. *When* $g(X) = g(X; \alpha) = \alpha_0 + \alpha_1^T X,\ h(X) = h(X; \alpha) = \alpha_2 + \alpha_3^T X$ *with* $(\alpha_0, \alpha_1) \neq 0, \alpha_2 > 0, \alpha_3 \neq 0,$ *where* $\alpha = (\alpha_0, \alpha_1^T, \alpha_2, \alpha_3^T)^T, \alpha_1 = (\alpha_{1,1}, \cdots, \alpha_{1,p})^T,\ \alpha_3 = (\alpha_{3,1}, \cdots, \alpha_{3,p})^T$*, if the Condition C holds, we have* $\{g(X), h(X)\} \in \{\mathcal{S}_1(X), \mathcal{S}_2(X)\}$.

Proof. It is enough to show that if $\alpha = (\alpha_0, \alpha_1^T, \alpha_2, \alpha_3^T)^T, \beta = (\beta_0, \beta_1^T, \beta_2, \beta_3^T)^T$ satisfy:

$$\frac{\alpha_0 + \alpha_1^T X}{\sqrt{1 + (\alpha_2 + \alpha_3^T X)^2}} = \frac{\beta_0 + \beta_1^T X}{\sqrt{1 + (\beta_2 + \beta_3^T X)^2}},\ \forall X \in \mathcal{X}, \tag{A2.4}$$

then $\alpha = \beta$. To keep the same signs on both sides, the following two subsets of a hyperplane (H_0, H_1) must be the same,

$$H_0 = \{X \subset \mathcal{X} | \alpha_0 + \alpha_1^T X = 0\},\ H_1 = \{X \subset \mathcal{X} | \beta_0 + \beta_1^T X = 0\},$$

since there exists linear independent $(\tau_1, \cdots, \tau_p) \subset \mathcal{X}$ such that $P(Y = 1|X = \tau_i) = 0.5, i = 1, \ldots, p$, thus, the following two hyperplane $(\widetilde{H}_0, \widetilde{H}_1)$ must be the same,

$$\widetilde{H}_0 = \{X \subset \mathbb{R}^p | \alpha_0 + \alpha_1^T X = 0\},\ \widetilde{H}_1 = \{X \subset \mathbb{R}^p | \beta_0 + \beta_1^T X = 0\},$$

which means $(\alpha_0, \alpha_1^T) = k(\beta_0, \beta_1^T)$, and $k \geq 0$ since the signs on the two sides of equations (A2.4) must be the same. And $(\alpha_0, \alpha_1) \neq 0$ exclude the case $k = 0$. Thus from equation (A2.4) we have

$$k^2 = \frac{1 + (\alpha_2 + \alpha_3^T X)^2}{1 + (\beta_2 + \beta_3^T X)^2}.$$

By arranging the equation above we have

$$X^T(\alpha_3\alpha_3^T - k^2\beta_3\beta_3^T)X + 2(\alpha_2\alpha_3^T - k^2\beta_2\beta_3^T)X + 1 + \alpha_2^2 - k - k\beta_2^2 = 0.$$

So

$$\alpha_3\alpha_3^T - k^2\beta_3\beta_3^T = 0, \tag{A2.5a}$$

$$\alpha_2\alpha_3^T - k^2\beta_2\beta_3^T = 0, \tag{A2.5b}$$

$$1 + \alpha_2^2 - k^2 - k^2\beta_2^2 = 0. \tag{A2.5c}$$

With a little abuse of notation, we use 0 to denote not only the number 0 but also the matrix and vector of 0 in (A2.5a) and (A2.5b) respectively. Take the (i, i) element of (A2.5a) and the i-th component of (A2.5b), with a little arrangement we have

$$\alpha_{3i}^2 = k^2\beta_{3i}^2, \tag{A2.5d}$$

$$\alpha_2\alpha_{3i} = k^2\beta_2\beta_{3i}, \tag{A2.5e}$$

$$\alpha_2^2 = k^2 + k^2\beta_2^2 - 1. \tag{A2.5f}$$

Note $(A2.5d)\cdot(A2.5f) - (A2.5e)^2 = k^2\beta_{3i}^2(k^2-1) = 0$, since $k > 0$ we have $k = 1$. And since $\alpha_2, \beta_2 \geq 0$, from (A2.5c) we have $\alpha_2 = \beta_2$, then from (A2.5b) we have $\alpha_3 = \beta_3$. Thus, $\alpha = \beta$. This completes the proof of part (ii) in Theorem 1. $\square$

Appendix C Non-identification without interaction term between X and U

Theorem C.1. *Under the same assumptions as in Theorem B.1,*

(i) If there is no interaction term between X and U in model (A2.2), i.e., $h_t(X) = h_t$ is a constant, the $(\mathrm{TBR}_c(x), \mathrm{THR}_c(x))$ can not be identified for any $c \neq \pm E[Y_1 - Y_0]$.

(ii) If there is no interaction term between X and U in model (A2.3), i.e., $h_t(X) = h_t$ is a constant, the $(\mathrm{TBR}(x), \mathrm{THR}(x))$ can not be identified for any $\big(g_0(x), g_1(x)\big) \neq (0, 0)$.

Proof.

(i) We have

$$Y\Big|\Big(X, T=t\Big) \sim N\big(g_t(X),\ h_t^2+\sigma_t^2\big).$$

Since $P(Y,X,T) = P(Y|X,T)P(X,T)$ and $P(X,T)$ is not related to the parameters in the model, we can only identify $g_t(X)$ and $h_t^2+\sigma_t^2$ for $t=0,1$. Since h_t^2 is a constant, we can no longer separate h_t and σ_t^2 from $\big(h_t^2+\sigma_t^2\big)$ without further assumptions, i.e., (h_t, σ_t^2) can not be identified. Additionally, we have

$$(Y_0, Y_1)\Big|X = x \sim N(\mu(x), \Sigma(x)),$$

where

$$\mu(x) = (g_0(x), g_1(x)),\ \Sigma(x) = \begin{pmatrix} h_0^2+\sigma_0^2 & h_0h_1 \\ h_0h_1 & h_1^2+\sigma_1^2 \end{pmatrix}.$$

Thus,

$$(Y_1-Y_0)|X = x \sim N\Big(g_1(x)-g_0(x),\ \ (h_0^2+\sigma_0^2)+(h_1^2+\sigma_1^2)-2h_0h_1\Big).$$

Since $h_t^2+\sigma_t^2$ can be identified while (h_t^2, σ_t^2) can not, the joint distribution of (Y_0, Y_1) given $X=x$ can not be identified, so is the distribution of Y_1-Y_0 given $X=x$.

Since $\mathrm{TBR}_c(x) = P(Y_1-Y_0 > c|X=x)$ and Y_1-Y_0 given $X=x$ is normally distributed with mean identified and variance unidentified, so $\mathrm{TBR}_c(x)$ is unidentified if $c \neq E[Y_1-Y_0]$. Similarly, $\mathrm{THR}_c(x)$ is unidentified if $c \neq -E[Y_1-Y_0]$.

(ii) Since

$$P(Y=1|X, T=t) = \Phi\Big(\frac{g_t(X)}{\sqrt{1+h_t^2}}\Big),$$

we can only identify $g_t(X)/\sqrt{1+h_t^2}$ in the model with the numerator and denominator unseparate, which means $(g_t(X), h_t^2)$ can not be identified. Additionally, we have

$$\begin{aligned}\text{TBR}(x) &= P(Y_0 = 0, Y_1 = 1 | X = x) \\ &= \int_{-\infty}^{0}\int_{0}^{\infty} \frac{1}{2\pi|\Sigma_b|^{1/2}} \exp\Big\{ -\frac{1}{2}\big((s_0, s_1) - \mu_b\big)\Sigma_b^{-1}\big((s_0, s_1) - \mu_b\big)\Big\} ds_0 ds_1,\end{aligned}$$

where

$$\mu_b = (-g_0(x), -g_1(x)), \quad \Sigma_b = \begin{pmatrix} 1 + h_0^2 & h_0 h_1 \\ h_0 h_1 & 1 + h_1^2 \end{pmatrix}.$$

Let $(t_0 = s_0/\sqrt{1 + h_0^2},\ t_1 = s_1/\sqrt{1 + h_1^2})$, we have

$$\begin{aligned}\text{TBR}(x) &= P(Y_0 = 0, Y_1 = 1 | X = x) \\ &= \int_{-\infty}^{0}\int_{0}^{\infty} \frac{1}{2\pi|\widetilde{\Sigma}_b|^{1/2}} \exp\Big\{ -\frac{1}{2}\big((t_0, t_1) - \widetilde{\mu}_b\big)\widetilde{\Sigma}_b^{-1}\big((t_0, t_1) - \widetilde{\mu}_b\big)\Big\} dt_0 dt_1,\end{aligned}$$

where

$$\widetilde{\mu}_b = \big(-g_0(x)/\sqrt{1 + h_0^2}, -g_1(x)/\sqrt{1 + h_1^2}\big), \quad \widetilde{\Sigma}_b = \begin{pmatrix} 1 & \frac{h_0 h_1}{\sqrt{1+h_0^2}\sqrt{1+h_1^2}} \\ \frac{h_0 h_1}{\sqrt{1+h_0^2}\sqrt{1+h_1^2}} & 1 \end{pmatrix}.$$

So $\widetilde{\mu}_b$ is identified while $\widetilde{\Sigma}_b$ not. Thus, we can easily conclude that TBR(x) can not be identified when $\big(g_0(x), g_1(x)\big) \neq (0, 0)$, so is THR$(x)$ and the joint distribution of (Y_0, Y_1) given $X = x$. □

Appendix D Proof of Theorem 2

Proof. The estimator $\widehat{\theta} = (\widehat{\alpha}_{0,0}, \widehat{\alpha}_{0,1}^T, \widehat{\alpha}_{0,2}, \widehat{\alpha}_{0,3}^T, \widehat{\sigma}_0^2, \widehat{\alpha}_{1,0}, \widehat{\alpha}_{1,1}^T, \widehat{\alpha}_{1,2}, \widehat{\alpha}_{1,3}^T, \widehat{\sigma}_1^2)^T$ maximize the following likelihood

$$\begin{aligned}&\ell = \log L(Y|X) \\ = &\sum_{i=1}^{n}\sum_{t=0,1} \frac{1}{2}\Big[I(T_i = t)\Big\{ -\log(2\pi) - \log\big((\alpha_{t,2} + \alpha_{t,3}^T X_i)^2 + \sigma_t^2\big) - \frac{(Y_i - \alpha_{t,0} - \alpha_{t,1}^T X_i)^2}{(\alpha_{t,2} + \alpha_{t,3}^T X_i)^2 + \sigma_t^2}\Big\}\Big].\end{aligned}$$

According to the M-estimator property, we have

$$\sqrt{n}(\widehat{\theta}-\theta) \xrightarrow{d} N\Big(0, \Big[P_0\big\{\frac{\partial^2\psi}{\partial\theta\partial\theta^T}\big\}\Big]^{-1} P_0\big\{\frac{\partial\psi}{\partial\theta}\frac{\partial\psi}{\partial\theta^T}\big\}\Big[P_0\big\{\frac{\partial^2\psi}{\partial\theta\partial\theta^T}\big\}\Big]^{-1}\Big),$$

where P_0 is the true mean and

$$\begin{aligned}&\psi(T,X,Y;\theta)\\ =\;& \sum_{t=0,1}\frac{1}{2}\Big[I(T=t)\Big\{-\log(2\pi)-\log\big((\alpha_{t,2}+\alpha_{t,3}^TX)^2+\sigma_t^2\big)-\frac{(Y-\alpha_{t,0}-\alpha_{t,1}^TX)^2}{(\alpha_{t,2}+\alpha_{t,3}^TX)^2+\sigma_t^2}\Big\}\Big].\end{aligned}$$

Let

$$m_B(X;\theta)=\Phi\Big(\frac{(\alpha_{1,0}-\alpha_{0,0})+(\alpha_{1,1}-\alpha_{0,1})^TX-c}{\sqrt{((\alpha_{1,2}-\alpha_{0,2})+(\alpha_{1,3}-\alpha_{0,3})^TX)^2+(\sigma_0^2+\sigma_1^2)}}\Big),$$

and

$$m_H(X;\theta)=\Phi\Big(\frac{(\alpha_{0,0}-\alpha_{1,0})+(\alpha_{0,1}-\alpha_{1,1})^TX-c}{\sqrt{((\alpha_{0,2}-\alpha_{1,2})+(\alpha_{0,3}-\alpha_{1,3})^TX)^2+(\sigma_0^2+\sigma_1^2)}}\Big),$$

thus, $\widehat{\mathrm{TBR}}_c(x)-\mathrm{TBR}_c(x)=m_B(x;\widehat{\theta})-m_B(x;\theta)$ and $\widehat{\mathrm{THR}}_c(x)-\mathrm{THR}_c(x)=m_H(x;\widehat{\theta})-m_H(x;\theta)$.

By the Delta-Method, we have

$$\sqrt{n}\big(\widehat{\mathrm{TBR}}(x)-\mathrm{TBR}(x)\big)\xrightarrow{d} N(0,\sigma_{cB}^2(x;\theta)),$$

$$\sqrt{n}\big(\widehat{\mathrm{THR}}(x)-\mathrm{THR}(x)\big)\xrightarrow{d} N(0,\sigma_{cH}^2(x;\theta)),$$

where

$$\sigma_{cB}^2(x;\theta)=\frac{\partial}{\partial\theta^T}m_B(x;\theta)\Big[P_0\big\{\frac{\partial^2\psi}{\partial\theta\partial\theta^T}\big\}\Big]^{-1}P_0\big\{\frac{\partial\psi}{\partial\theta}\frac{\partial\psi}{\partial\theta^T}\big\}\Big[P_0\big\{\frac{\partial^2\psi}{\partial\theta\partial\theta^T}\big\}\Big]^{-1}\frac{\partial}{\partial\theta}m_B(x;\theta),$$

and

$$\sigma_{cH}^2(x;\theta)=\frac{\partial}{\partial\theta^T}m_H(x;\theta)\Big[P_0\big\{\frac{\partial^2\psi}{\partial\theta\partial\theta^T}\big\}\Big]^{-1}P_0\big\{\frac{\partial\psi}{\partial\theta}\frac{\partial\psi}{\partial\theta^T}\big\}\Big[P_0\big\{\frac{\partial^2\psi}{\partial\theta\partial\theta^T}\big\}\Big]^{-1}\frac{\partial}{\partial\theta}m_H(x;\theta).$$

Appendix E Proof of Theorem 3

Proof. The estimator $\widehat{\theta} = (\widehat{\alpha}_{0,0}, \widehat{\alpha}_{0,1}^T, \widehat{\alpha}_{0,2}, \widehat{\alpha}_{0,3}^T, \widehat{\alpha}_{1,0}, \widehat{\alpha}_{1,1}^T, \widehat{\alpha}_{1,2}, \widehat{\alpha}_{1,3}^T)^T$ maximize the following likelihood

$$\ell = \log L(Y|X) = \sum_{i=1}^{n} \sum_{t=0,1} \Big[I(T_i = t) \Big\{ Y_i \log \big(G(X_i; \theta_t) \big) + (1 - Y_i) \log \big(1 - G(X_i; \theta_t) \big) \Big\} \Big],$$

where

$$G(X; \theta_t) = \Phi\Big(\frac{\alpha_{t,0} + \alpha_{t,1}^T X}{\sqrt{1 + (\alpha_{t,2} + \alpha_{t,3}^T X)^2}} \Big).$$

According to the M-estimator property, we have

$$\widehat{\theta} - \theta = -\Big[P_0 \big\{ \frac{\partial^2}{\partial\theta\partial\theta^T} \psi(T, X, Y; \theta) \big\} \Big]^{-1} \frac{1}{n} \sum_{i=1}^{n} \frac{\partial}{\partial\theta} \psi(T_i, X_i, Y_i; \theta) + o_p(1/\sqrt{n}),$$

where

$$\psi(T, X, Y; \theta) = \sum_{t=0,1} \Big[I(T = t) \Big\{ Y \log \big(G(X; \theta_t) \big) + (1 - Y) \log \big(1 - G(X; \theta_t) \big) \Big\} \Big].$$

Let $m_B(X; \theta) = \Phi_b\big(\mu(x; \theta), \Sigma(x; \theta)\big)$, and $m_H(X; \theta) = \Phi_h\big(\mu(x; \theta), \Sigma(x; \theta)\big)$, we have $\widehat{\mathrm{TBR}}(x) - \mathrm{TBR}(x) = m_B(x; \widehat{\theta}) - m_B(x; \theta)$ and $\widehat{\mathrm{THR}}(x) - \mathrm{THR}(x) = m_H(x; \widehat{\theta}) - m_H(x; \theta)$.

By the Delta-Method, we have

$$\sqrt{n}\big(\widehat{\mathrm{TBR}}(x) - \mathrm{TBR}(x)\big) \xrightarrow{d} N(0, \sigma_{bB}^2(x; \theta)),$$

$$\sqrt{n}\big(\widehat{\mathrm{THR}}(x) - \mathrm{THR}(x)\big) \xrightarrow{d} N(0, \sigma_{bH}^2(x; \theta)),$$

where

$$\sigma_{bB}^2(x; \theta) = \frac{\partial}{\partial\theta^T} m_B(x; \theta) \Big[P_0 \big\{ \frac{\partial^2 \psi}{\partial\theta\partial\theta^T} \big\} \Big]^{-1} P_0 \big\{ \frac{\partial\psi}{\partial\theta} \frac{\partial\psi}{\partial\theta^T} \big\} \Big[P_0 \big\{ \frac{\partial^2 \psi}{\partial\theta\partial\theta^T} \big\} \Big]^{-1} \frac{\partial}{\partial\theta} m_B(x; \theta),$$

and

$$\sigma_{bH}^2(x; \theta) = \frac{\partial}{\partial\theta^T} m_H(x; \theta) \Big[P_0 \big\{ \frac{\partial^2 \psi}{\partial\theta\partial\theta^T} \big\} \Big]^{-1} P_0 \big\{ \frac{\partial\psi}{\partial\theta} \frac{\partial\psi}{\partial\theta^T} \big\} \Big[P_0 \big\{ \frac{\partial^2 \psi}{\partial\theta\partial\theta^T} \big\} \Big]^{-1} \frac{\partial}{\partial\theta} m_H(x; \theta).$$

Appendix F Estimates of Parameters and Their Asymptotic Properties in Models (3) and (4)

The corresponding formulas of (F1), (F2), (F3) and (F4) for the general models are:

$$\begin{cases} \mathrm{TBR}_c(x) &= \Phi\Big(\dfrac{\big(g_1(x)-g_0(x)\big)-c}{\sqrt{\big(h_1(x)-h_0(x)\big)^2+\sigma_0^2+\sigma_1^2}}\Big), \\ \mathrm{THR}_c(x) &= \Phi\Big(\dfrac{\big(g_0(x)-g_1(x)\big)-c}{\sqrt{\big(h_0(x)-h_1(x)\big)^2+\sigma_0^2+\sigma_1^2}}\Big), \end{cases}$$

and

$$\begin{cases} \mathrm{TBR}(x) &= \Phi_b\big(\widetilde{\mu}(x),\widetilde{\Sigma}(x)\big), \\ \mathrm{THR}(x) &= \Phi_h\big(\widetilde{\mu}(x),\widetilde{\Sigma}(x)\big), \end{cases}$$

where

$$\widetilde{\mu}(x) = -\big(g_0(x),\ g_1(x)\big),$$

$$\widetilde{\Sigma}(x) = \begin{pmatrix} 1+h_0^2(x) & h_0(x)h_1(x) \\ h_0(x)h_1(x) & 1+h_1^2(x) \end{pmatrix}.$$

In estimation, we first model $g_t(X)$ and $h_t(X)$ as $g_t(X;\alpha_{t,1})$ and $h_t(X;\alpha_{t,2})$. Also let $\psi(T,X,Y;\theta)$ denote the log-density function, where $\theta=(\alpha_{0,1},\alpha_{0,2},\sigma_0^2,\alpha_{1,1},\alpha_{1,2},\sigma_1^2)^T$ in the continuous case and $\theta=(\alpha_{0,1},\alpha_{0,2},\alpha_{1,1},\alpha_{1,2})^T$ in the binary case. The estimation for θ can be obtained by maximizing $P_n[\psi(T,X,Y;\theta)]$, denote as $\widehat{\theta}$. Then $\mathrm{TBR}(x)$, $\mathrm{THR}(x)$, $\mathrm{TBR}_c(x)$ and $\mathrm{THR}_c(x)$ can be estimated by:

$$\begin{cases} \widehat{\mathrm{TBR}_c}(x) &= \Phi\Big(\dfrac{\big(g_1(X;\widehat{\alpha}_{1,1})-g_0(X;\widehat{\alpha}_{0,1})\big)-c}{\sqrt{\big(h_1(X;\widehat{\alpha}_{1,2})-h_0(X;\widehat{\alpha}_{0,2})\big)^2+\widehat{\sigma}_0^2+\widehat{\sigma}_1^2}}\Big), \\ \widehat{\mathrm{THR}_c}(x) &= \Phi\Big(\dfrac{\big(g_1(X;\widehat{\alpha}_{0,1})-g_0(X;\widehat{\alpha}_{1,1})\big)-c}{\sqrt{\big(h_1(X;\widehat{\alpha}_{0,2})-h_0(X;\widehat{\alpha}_{1,2})\big)^2+\widehat{\sigma}_0^2+\widehat{\sigma}_1^2}}\Big), \\ \widehat{\mathrm{TBR}}(x) &= \Phi_b\big(\widetilde{\mu}(X;\widehat{\theta}),\widetilde{\Sigma}(X;\widehat{\theta})\big), \\ \widehat{\mathrm{THR}}(x) &= \Phi_h\big(\widetilde{\mu}(X;\widehat{\theta}),\widetilde{\Sigma}(X;\widehat{\theta})\big), \end{cases}$$

where

$$\widetilde{\mu}(X;\widehat{\theta}) = \big(-g_0(X;\widehat{\alpha}_{0,1}),\ -g_1(X;\widehat{\alpha}_{1,1})\big),$$

$$\widetilde{\Sigma}(X;\widehat{\theta}) = \begin{pmatrix} 1+h_0^2(X;\widehat{\alpha}_{0,2}) & h_0(X;\widehat{\alpha}_{0,2})h_1(X;\widehat{\alpha}_{1,2}) \\ h_0(X;\widehat{\alpha}_{0,2})h_1(X;\widehat{\alpha}_{1,2}) & 1+h_1^2(X;\widehat{\alpha}_{1,2}) \end{pmatrix}.$$

We estimate the variances of $\widehat{\mathrm{TBR}}_c(x)$, $\widehat{\mathrm{THR}}_c(x)$, $\widehat{\mathrm{TBR}}(x)$ and $\widehat{\mathrm{THR}}(x)$ by the plug-in estimator respectively.

Appendix G Identification When U Depends on X

Theorem G.1. *Under the Assumption 2:*

(i) When the outcome is continuous, if the following model (A7.6) holds for t=0,1

$$\begin{cases} Y_t = g_t(X) + h_t(X)U + \epsilon_t,\ \epsilon_t \sim N(\mu_t, \sigma_t^2),\ \epsilon_t \perp (X, U, \epsilon_u), \\ U = W(X) + \epsilon_u,\ \epsilon_u \sim N(\mu_u, \sigma^2),\ \epsilon_u \perp X \end{cases} \tag{A7.6}$$

then the Condition A in the Appendix B is sufficient to identify the joint distribution of (Y_0, Y_1) *given X.*

(ii) When the outcome is binary, if the following model (A7.7) holds for t=0,1

$$\begin{cases} Y_t^* = g_t(X) + h_t(X)U + \epsilon_t, \epsilon_t \sim N(\mu_t, \sigma_t^2), \epsilon_t \perp (X, U, \epsilon_u), \\ Y_t = I(Y_t^* > 0), \\ U = W(X) + \epsilon_u, \epsilon_u \sim N(\mu_u, \sigma^2),\ \epsilon_u \perp X, \end{cases} \tag{A7.7}$$

then the following Condition D is sufficient to identify the joint distribution of (Y_0, Y_1) *given X.*

Condition D. $\big(g_t(X) + W(X)h_t(X), h_t(X)\big)$ *belongs to the family* $\big(\mathcal{S}_1(X), \mathcal{S}_2(X)\big)$ *for* $t = 0, 1$*, where*

$$\big(\mathcal{S}_1(X), \mathcal{S}_2(X)\big) = \Big\{\big(S_1(X;\beta_1), S_2(X;\beta_2)\big)\big|(\beta_1, \beta_2) \in \mathcal{A},$$
$$\forall(\beta_1^{(1)}, \beta_2^{(1)}) \neq (\beta_1^{(2)}, \beta_2^{(2)}) \in \mathcal{A}, \frac{S_1(X;\beta_1^{(1)})}{\sqrt{1+S_2^2(X;\beta_2^{(1)})}} \neq \frac{S_1(X;\beta_1^{(2)})}{\sqrt{1+S_2^2(X;\beta_2^{(2)})}}\Big\}.$$

Proof.

(i) Without loss of generality, we assume $\sigma^2 = 1$ since otherwise it can be absorbed into $h_t(X)$, $\mu_u = 0$ since otherwise it can be absorbed into $W(X)$ and assume $\mu_t = 0$ since otherwise it can be absorbed into $g_t(X)$. Also, we assume $h_t(0) > 0$ since otherwise we use $U^* = -U$ to replace U. By a little arrangement, we have

$$Y_t = \big(g_t(X) + h_t(X)W(X)\big) + h_t(X)\epsilon_u + \epsilon_t.$$

Thus,

$$Y\Big|\Big(X, T = t\Big) \sim N\big(g_t(X) + h_t(X)W(X),\ h_t^2(X) + \sigma_t^2\big).$$

Then $\big(g_t(X) + h_t(X)W(X)\big)$ and $\big(h_t^2(X) + \sigma_t^2\big)$ can both be identified, so is $h_t(X)h_t'(X)$. Since $h_t(X)$ belongs to $\mathcal{S}(X)$, we can also identify $h_t(X)$ and σ_t^2.

Note that

$$P(Y_0, Y_1|X = x) = P\Big(\big(g_0(x)+h_0(x)W(x)\big)+h_0(x)\epsilon_u+\epsilon_0,\ \big(g_1(X)+h_1(X)W(X)\big)+h_1(X)\epsilon_u+\epsilon_1\Big).$$

Thus, we can identify the joint distribution of (Y_0, Y_1) given X.

(ii) Without loss of generality, we can assume that ϵ_u follows a standard normal distribution. Also, we assume $\mu_t = 0$ since otherwise it can be absorbed into $g_t(X)$, $\sigma_t^2 = 1$ since otherwise we can use $\widetilde{Y}_t^* = Y_t^*/\sigma_t$ to replace Y_t^* and $h_t(0) > 0$ since otherwise we can use $U^* = -U$ to replace U. By a little arrangement, we have

$$
\begin{aligned}
& P(Y=1|X,T=t) = P(Y_t=1|X) \\
= \;& P\big(g_t(X)+h_t(X)W(X)+h_t(X)\epsilon_u+\epsilon_t>0|X\big) \\
= \;& \int\int \frac{1}{2\pi}\exp\big(-\frac{s_u^2+s_t^2}{2}\big)I\big(g_t(X)+h_t(X)W(X)+h_t(X)s_u+s_t>0\big)ds_t ds_u \\
= \;& \int\int_{-\infty}^{g_t(X)+h_t(X)W(X)+h_t(X)s_u} \frac{1}{2\pi}\exp\big(-\frac{s_u^2+s_t^2}{2}\big)ds_t ds_u \\
= \;& \int\int_{-\infty}^{0}\frac{1}{2\pi}\exp(-\frac{s_u^2+\big(g_t(X)+h_t(X)W(X)+h_t(X)s_u+s_t\big)^2}{2})ds_t ds_u \\
= \;& \int_{-\infty}^{0}\int\frac{1}{2\pi}\exp\Big(-\frac{1}{2}\big((1+h_t^2(X))(s_u+\frac{h_t(X)(g_t(X)+h_t(X)W(X)+s_t)}{1+h_t^2(X)})^2 \\
& +\frac{\big(g_t(X)+h_t(X)W(X)+s_t\big)^2}{1+h_t^2(X)}\Big)ds_u ds_t \\
= \;& \int_{-\infty}^{0}\frac{1}{\sqrt{2\pi}\sqrt{1+h_t^2(X)}}\exp\Big(-\frac{1}{2}\frac{\big(s_t+g_t(X)+h_t(X)W(X)\big)^2}{1+h_t^2(X)}\Big)ds_t \\
= \;& \Phi\Big(\frac{g_t(X)+h_t(X)W(X)}{\sqrt{1+h_t^2(X)}}\Big).
\end{aligned}
$$

Thus, if the Condition D is satisfied, we can identify $\big(g_t(X)+W(X)h_t(X),h_t(X)\big)$. Let $K_t(x,\epsilon_u)=g_t(x)+h_t(x)W(x)+h_t(x)\epsilon_u$, we have

$$
\begin{aligned}
& \mathrm{TBR}(x) = \big\{1-\Phi\big(K_0(x,s)\big)\big\}\Phi\big(K_1(x,s)\big)f_{\epsilon_u}(s)ds \\
= \;& \int\Big\{\int_{K_0(x,s)}^{\infty}\frac{1}{\sqrt{2\pi}}\exp(-s_0^2/2)ds_0\Big\}\Big\{\int_{-\infty}^{K_1(x,s)}\frac{1}{\sqrt{2\pi}}\exp(-s_1^2/2)ds_1\Big\}f_{\epsilon_u}(s)ds \\
= \;& \int\int_{-\infty}^{0}\int_{0}^{\infty}\frac{1}{(2\pi)^{3/2}}\exp\Big\{-\frac{(s_0+K_0(x,s))^2+(s_1+K_1(x,s))^2+s^2}{2}\Big\}ds_0 ds_1 ds.
\end{aligned}
$$

Let $K_{t,1}(x)=g_t(x)+h_t(x)W(x)$, thus $K_t(x,s)=K_{t,1}(x)+sh_t(x)$. Then the term in

$\exp\big(-\frac{1}{2}(\cdot)\big)$ can be arranged as

$$\begin{aligned}
& \{s_0+K_0(x,s)\}^2+\{s_1+K_1(x,s)\}^2+s^2 \\
= \quad & \big\{1+h_0^2(x)+h_1^2(x)\big\}s^2+2\big\{(s_0+K_{0,1}(x))h_0(x)+(s_1+K_{1,1}(x))h_1(x)\big\}s \\
& +\{s_0+K_{0,1}(x)\}^2+\{s_1+K_{1,1}(x)\}^2 \\
= \quad & \big\{1+h_0^2(x)+h_1^2(x)\big\}\big\{s+\frac{(s_0+K_{0,1}(x))h_0(x)+(s_1+K_{1,1}(x))h_1(x)}{1+h_0^2(x)+h_1^2(x)}\big\}^2 \\
& +\frac{1}{1+h_0^2(x)+h_1^2(x)}\Big[\{s_0+K_{0,1}(x)\}^2\{1+h_1^2(x)\}+\{s_1+K_{1,1}(x)\}^2\{1+h_0^2(x)\} \\
& \qquad -2\{s_0+K_{0,1}(x)\}h_0(x)\{s_1+K_{1,1}(x)\}h_1(x)\Big].
\end{aligned}$$

So

$$\mathrm{TBR}(x)=\int_0^\infty\int_{-\infty}^0\frac{1}{(2\pi)S}\exp\Big(-\frac{F}{2}\Big)ds_0ds_1,$$

where

$$S^2=1+h_0^2(x)+h_1^2(x),$$

$$\begin{aligned}
F \quad = \quad & \Big[\{s_0+K_{0,1}(x)\}^2\{1+h_1^2(x)\}+\{s_1+K_{1,1}(x)\}^2\{1+h_0^2(x)\} \\
& -2\{s_0+K_{0,1}(x)\}h_0(x)\{s_1+K_{1,1}(x)\}h_1(x)\Big]/S^2 \\
= \quad & \Big\{(s_0,s_1)-\mu\Big\}\Sigma^{-1}\Big\{(s_0,s_1)-\mu\Big\}^T,
\end{aligned}$$

$$\mu=(-K_{0,1}(x),\ -K_{1,1}(x)),$$

$$\Sigma=\begin{pmatrix}1+h_0^2(x) & h_0(x)h_1(x)\\ h_0(x)h_1(x) & 1+h_1^2(x)\end{pmatrix}.$$

Thus, $\mathrm{TBR}(x)=\Phi_2\big((0,\infty),(-\infty,0);\mu,\Sigma\big)$, where $\Phi_2(A_0,A_1;\mu,\Sigma)$ is the distribution function of bivariate normal vector with mean μ, covariance matrix Σ and integral region $A_0\times A_1$. Similarly, we can derive the form for $\mathrm{THR}(x)$. Thus, we can identify the $\mathrm{TBR}(x)$ and $\mathrm{THR}(x)$, so the joint distribution of (Y_0,Y_1) given X are identifiable. □

Appendix H Additional Tables

Table 1: The true value, bias, average estimated standard error (ASE), empirical standard error (ESE) and 95% confidence interval (CI) coverage in continuous case. Every table cell contains two elements, which corresponds to the population TBR_c (first row in each cell) and THR_c (second row in each cell) ($c = 0.5$) respectively.

Distribution of U	true value	bias	ASE	ESE	95% CI coverage
Normal	0.501	−0.001	0.017	0.017	0.945
	0.397	−0.001	0.016	0.016	0.949
t(3)	0.500	−0.001	0.017	0.017	0.951
	0.396	0.002	0.016	0.016	0.948
t(10)	0.499	< 0.001	0.017	0.017	0.953
	0.395	0.002	0.016	0.016	0.939
$\chi^2(3)$	0.501	−0.001	0.017	0.016	0.955
	0.397	< 0.001	0.016	0.016	0.954
$\chi^2(10)$	0.502	−0.002	0.017	0.017	0.952
	0.397	< 0.001	0.016	0.016	0.956
P(3)	0.502	−0.002	0.017	0.017	0.951
	0.398	−0.002	0.016	0.016	0.943
P(10)	0.503	−3e-03	0.017	0.017	0.933
	0.397	−7e-04	0.016	0.016	0.942
B(0.5)	0.501	−8e-04	0.017	0.017	0.949
	0.395	6e-04	0.016	0.016	0.953

Table 2: Estimates, estimated standard deviation (SD) and p-value of parameters of the Mind Study

	$t=0$			$t=1$		
	Estimate	SD	p-value	Estimate	SD	p-value
Gender	-0.656	0.275	0.017	-0.248	0.321	0.439
CVD	0.581	0.353	0.100	0.100	0.395	0.801
Age	1.075	0.202	< 0.001	0.500	0.231	0.030
DSST	-0.483	0.190	0.011	-0.652	0.231	0.005
Race	0.619	0.309	0.045	0.355	0.383	0.354
U	1.768	0.791	0.025	0.148	0.374	0.693
UGender	-1.916	0.480	< 0.001	-1.742	0.414	< 0.001
UCVD	-0.321	0.398	0.420	-1.669	0.506	0.001
UAge	1.280	0.513	0.013	2.090	0.435	< 0.001
UDSST	-1.166	0.313	< 0.001	-1.157	0.339	0.001
URace	1.729	0.390	< 0.001	2.239	0.479	< 0.001
σ_t^2	1.080	0.277	< 0.001	1.992	0.491	< 0.001