Web-based Supplementary Materials for

"Sharpening bounds on principal effects with covariates,"

by Dustin M. Long and Michael G. Hudgens*

* email: mhudgens@bios.unc.edu

Web Appendix A: Proposition Proofs

Proof of Proposition 1

Note

$$\begin{aligned}\theta^u_{100X} &= \sum_x \theta^u_{100x}\phi_x = \sum_x \min\left\{\frac{\pi_{1x}}{\gamma_x}, 1\right\}\phi_x \leq \min\left\{\sum_x \frac{\pi_{1x}}{\gamma_x}\phi_x, \sum_x \phi_x\right\} \\ &= \min\left\{\frac{\pi_1}{\gamma}, 1\right\} = \theta^u_{100},\end{aligned}$$

where the inequality holds since $\min\{a_1, b_1\} + \min\{a_2, b_2\} \leq \min\{a_1 + a_2, b_1 + b_2\}$ and the third equality holds because

$$\begin{aligned}&\sum_x \frac{\pi_{1x}}{\gamma_x}\phi_x = \sum_x \frac{\Pr[Y_i(1) = 1|S_i(1) = 0, X_i = x]}{\Pr[S_i(0) = 0|S_i(1) = 0, X_i = x]}\Pr[X_i = x|S_i(0) = S_i(1) = 0] \\ &= \sum_x \frac{\Pr[Y_i(1) = 1, S_i(1) = 0, X_i = x]}{\Pr[S_i(0) = S_i(1) = 0, X_i = x]}\frac{\Pr[X_i = x, S_i(0) = S_i(1) = 0]}{\Pr[S_i(0) = S_i(1) = 0]} \\ &= \sum_x \frac{\Pr[Y_i(1) = 1, S_i(1) = 0, X_i = x]}{\Pr[S_i(0) = S_i(1) = 0]} = \sum_x \frac{\Pr[Y_i(1) = 1, X_i = x|S_i(1) = 0]}{\Pr[S_i(0) = 0|S_i(1) = 0]} \\ &= \frac{\pi_1}{\gamma}. \qquad \text{(A-1)}\end{aligned}$$

Similarly for the lower bound,

$$\begin{aligned}\theta^l_{100X} &= \sum_x \theta^l_{100x}\phi_x = \sum_x \max\left\{\frac{\pi_{1x} - (1-\gamma_x)}{\gamma_x}, 0\right\}\phi_x \\ &\geq \max\left\{\sum_x \frac{\pi_{1x} - (1-\gamma_x)}{\gamma_x}\phi_x, 0\right\} = \max\left\{\frac{\pi_1 - (1-\gamma)}{\gamma}, 0\right\} = \theta^l_{100},\end{aligned}$$

where the inequality holds because $\max\{a_1,0\}+\max\{a_2,0\} \geq \max\{a_1+a_2,0\}$ and the third equality holds because of (A-1) and

$$\begin{aligned}
\sum_x \frac{1-\gamma_x}{\gamma_x}\phi_x &= \sum_x \frac{\Pr[S_i(0)=1|S_i(1)=0,X_i=x]}{\Pr[S_i(0)=0|S_i(1)=0,X_i=x]}\Pr[X_i=x|S_i(0)=S_i(1)=0] \\
&= \sum_x \frac{\Pr[S_i(0)=1,S_i(1)=0,X_i=x]}{\Pr[S_i(0)=0,S_i(1)=0,X_i=x]}\frac{\Pr[X_i=x,S_i(0)=S_i(1)=0]}{\Pr[S_i(0)=S_i(1)=0]} \\
&= \sum_x \frac{\Pr[S_i(0)=1,S_i(1)=0,X_i=x]}{\Pr[S_i(0)=S_i(1)=0]} = \sum_x \frac{\Pr[S_i(0)=1,X_i=x|S_i(1)=0]}{\Pr[S_i(0)=0|S_i(1)=0]} \\
&= \frac{1-\gamma}{\gamma}. \quad \square
\end{aligned}$$

Proof of Propositions 2 and 3

The proof of Proposition 2 is given below. The proof of Proposition 3 is similar and is omitted for brevity.

First, suppose equation (8) from the main text holds. Without loss of generality, assume $\pi_{10} < \gamma_0$ and $\pi_{11} > \gamma_1$ which implies that $\theta^u_{1000} = \pi_{10}/\gamma_0$ and $\theta^u_{1001} = 1$. If $\theta^u_{100} = \pi_1/\gamma$ then,

$$\theta^u_{100X} = \sum_x \theta^u_{100x}\phi_x = \frac{\pi_{10}}{\gamma_0}\phi_0 + \phi_1 < \frac{\pi_{10}}{\gamma_0}\phi_0 + \frac{\pi_{11}}{\gamma_1}\phi_1 = \frac{\pi_1}{\gamma} = \theta^u_{100},$$

where the inequality holds because $\pi_{11}/\gamma_1 > 1$. Likewise, if $\theta^u_{100} = 1$ then,

$$\theta^u_{100X} = \sum_x \theta^u_{100x}\phi_x = \frac{\pi_{10}}{\gamma_0}\phi_0 + \phi_1 < \phi_0 + \phi_1 = 1 = \theta^u_{100},$$

where the inequality holds since $\pi_{10}/\gamma_0 < 1$. Thus, if (8) is satisfied by X then $\theta^u_{100X} < \theta^u_{100}$.

Now suppose that (8) is not satisfied. We consider three possible cases. The first case is $\pi_1 < \gamma$, which implies that $\theta^u_{100} = \pi_1/\gamma$. Recall $\lambda_x = \Pr[X_i = x|S_i(1) = 0]$ for $x = 0, 1$. Suppose $\pi_{1x} \geq \gamma_x$ for $x = 0, 1$. Under this supposition, $\lambda_0\pi_{10} \geq \lambda_0\gamma_0$ and $\lambda_1\pi_{11} \geq \lambda_1\gamma_1$, implying that

$$\pi_1 = \lambda_0\pi_{10} + \lambda_1\pi_{11} \geq \lambda_0\gamma_0 + \lambda_1\gamma_1 = \gamma,$$

which is a contradiction. Thus, because we are assuming (8) is not satisfied, it must be that $\pi_{1x} \leq \gamma_x$ for $x = 0, 1$. This implies $\theta^u_{100x} = \pi_{1x}/\gamma_x$ for $x = 0, 1$ and thus

$$\theta^u_{100X} \;=\; \sum_x \theta^u_{100x}\phi_x \;=\; \frac{\pi_{10}}{\gamma_0}\phi_0 + \frac{\pi_{11}}{\gamma_1}\phi_1 \;=\; \frac{\pi_1}{\gamma} \;=\; \theta^u_{100}.$$

For the second case $\pi_1 > \gamma$ an analogous argument leads to the conclusion $\theta^u_{100X} = 1 = \theta^u_{100}$. Finally consider the third case $\pi_1 = \gamma$. For this case it is helpful to recall that throughout we have assumed $\Pr[S_i(0) = 0, X_i = x] > 0$ for all x, which implies $\lambda_x > 0$ for $x = 0, 1$. Now, suppose $\pi_{1x} \geq \gamma_x$ for $x = 0, 1$ where at least one of the inequalities is strict. Then

$$\pi_1 = \lambda_0\pi_{10} + \lambda_1\pi_{11} \;>\; \lambda_0\gamma_0 + \lambda_1\gamma_1 = \gamma,$$

which is a contradiction. Therefore $\pi_{1x} = \gamma_x$ for $x = 0, 1$, implying $\theta^u_{100X} = 1 = \theta^u_{100}$. Similarly, if we suppose $\pi_{1x} \leq \gamma_x$ for $x = 0, 1$ where at least one of the inequalities is strict, then a contradiction is reached, implying $\pi_{1x} = \gamma_x$ for $x = 0, 1$ and thus $\theta^u_{100X} = 1 = \theta^u_{100}$. Thus for all three cases it has been shown that if (8) is not satisfied then $\theta^u_{100X} = \theta^u_{100}$, completing the proof. □

Web Appendix B: Principal Stratification on S and Y

In the main text, principal strata are formed by considering the cross-classification of $S_i(0)$ and $S_i(1)$, i.e., the potential infection status of an infant at τ_0 for each possible treatment assignment. Individuals can be stratified even further by also cross-classifying by $Y_i(0)$ and $Y_i(1)$, i.e., the potential infection status of an infant at τ for each possible treatment assignment. Because no cure yet exists for HIV infection and individuals who become infected do not subsequently ever clear the virus, we can assume that $S_i(z) = 1$ implies $Y_i(z) = 1$ for $z = 0, 1$. This limits the total number of possible principal strata based on S_i and Y_i to seven, as enumerated in Web Table 1. For instance, the fourth row on Web Table 1 defines the NI/AI principal stratum as those infants who would never be infected by τ_0 but would always be infected by τ regardless of treatment assignment.

This cross-classification based on S and Y provides additional insight into some of the results given in the main text. For example, in Section 3 it was noted that CE is identifiable if and only if $\gamma = 1$, $\pi_1 = 1$, or $\pi_1 = 0$. The condition $\pi_1 = 1$ implies that individuals who are not in AI/AI must be members of PR/AI, NI/AI, or NI/HA. Conversely, the condition $\pi_1 = 0$ implies that individuals who are not in AI/AI cannot be members of PR/AI, NI/AI, or NI/HA (i.e., these strata are empty). In general, π_1 is the proportion of the population not in the AI/AI stratum who are members of PR/AI, NI/AI, or NI/HA. Similarly, γ is the proportion of the population not in the AI/AI stratum who are members of NI/AI, NI/PR, NI/HA, or NI/NI. Thus π_1 will be greater (less) than γ if and only if the proportion of the population in the PR/AI stratum is greater (less) than the proportion in the NI/PR and NI/NI strata combined.

Recall from Proposition 2 of the main text, the adjusted upper bound will be less than the unadjusted upper bound if and only if (8) holds. Based on the interpretations given above, (8) is equivalent to the proportion of the population in the PR/AI stratum being greater than the proportion in the NI/PR and NI/NI strata combined for one value of X and the opposite holding for another value of X. Similarly, Proposition 3 from the main text indicates the adjusted lower bound will be greater than the unadjusted lower bound if and only if (9) holds. Note $1 - \gamma$ is the proportion of the population not in the AI/AI stratum who are members of PR/PR or PR/AI. Thus condition (9) is equivalent to the proportion of the population in the PR/PR stratum being greater than the proportion in the NI/AI and NI/HA strata combined for one value of X and the opposite holding for another value of X.

Web Tables

Web Table 1

Principal Stratification on S and Y

$S_i(0)$	$S_i(1)$	$Y_i(0)$	$Y_i(1)$	Interpretation
1	1	1	1	Always infected/always infected (AI/AI)
1	0	1	0	Protected/protected (PR/PR)
1	0	1	1	Protected/always infected (PR/AI)
0	0	1	1	Never infected/always infected (NI/AI)
0	0	1	0	Never infected/protected (NI/PR)
0	0	0	1	Never infected/harmed (NI/HA)
0	0	0	0	Never infected/never infected (NI/NI)