Estimating the Asymptotic Variance of the ATE Estimator under the Cluster-Level Treatment Assignment

Juri Trifonov

March 25, 2024

Proof of Theorem 3.1 under the perfect compliance and clusters

Let us denote:

$$Q \equiv \mathbb{E}[N_g \bar{Y}_g(1) - N_g \bar{Y}_g(0)],$$
$$H \equiv \mathbb{E}[N_g],$$
$$\hat{Q} \equiv \frac{1}{G}\sum_{g=1}^{G}\left[\frac{A_g(N_g\bar{Y}_g - \hat{\mu}(1,S_g,X_g,N_g))}{\hat{\pi}(S_g)} - \frac{(1-A_g)(N_g\bar{Y}_g - \hat{\mu}(0,S_g,X_g,N_g))}{1-\hat{\pi}(S_g)} + \hat{\mu}(1,S_g,X_g,N_g) - \hat{\mu}(0,S_g,X_g,N_g)\right],$$
$$\hat{H} \equiv \frac{1}{G}\sum_{g=1}^{G} N_g.$$

The ATE estimator, $\hat{\tau}$, takes the following form:

$$\begin{aligned}
\hat{\tau} &= \frac{1}{\sum_{g=1}^{G} N_g}\sum_{g=1}^{G}\hat{\Xi}_g,\\
\hat{\Xi}_g &= \frac{A_g\left(N_g\bar{Y}_g - \hat{\mu}(1,S_g,X_g,N_g)\right)}{\hat{\pi}(S_g)} - \frac{(1-A_g)\left(N_g\bar{Y}_g - \hat{\mu}(0,S_g,X_g,N_g)\right)}{1-\hat{\pi}(S_g)}\\
&\quad + \hat{\mu}(1,S_g,X_g,N_g) - \hat{\mu}(0,S_g,X_g,N_g),
\end{aligned}$$

where $\hat{\mu}(a,S_g,X_g,N_g) = X_g'\hat{\theta}_{a,s} + N_g'\hat{\Delta}_{a,s}$. Where $\hat{\theta}_{a,s}$ and $\hat{\Delta}_{a,s}$ are obtained from:

$$N_g\bar{Y}_g \sim \gamma_{a,s} + X_{g,s}'\theta_{a,s} + N_g'\Delta_{a,s}.$$

Then, we can write the estimator as follows:

$$\begin{aligned}
\sqrt{G}(\hat{\tau} - \tau) &= \sqrt{G}\left(\frac{\hat{Q}}{\hat{H}} - \frac{Q}{H}\right)\\
&= \frac{1}{\hat{H}}\left[\sqrt{G}(\hat{Q} - Q) - \tau\sqrt{G}(\hat{H} - H)\right].
\end{aligned}$$

Step 1. Obtain the linear expansion of $\sqrt{G}(\hat{Q}-Q)$ and $\sqrt{G}(\hat{H}-H)$

Consider $\sqrt{G}(\hat{Q}-Q)$:

$$\begin{aligned}
\sqrt{G}(\hat{Q}-Q) &= \sqrt{G}\Bigg\{\frac{1}{G}\sum_{g=1}^{G}\Bigg[\frac{A_g(N_g\bar{Y}_g-\hat{\mu}(1,S_g,X_g,N_g))}{\hat{\pi}(S_g)} - \frac{(1-A_g)(N_g\bar{Y}_g-\hat{\mu}(0,S_g,X_g,N_g))}{1-\hat{\pi}(S_g)} \\
&\quad + \hat{\mu}(1,S_g,X_g,N_g)-\hat{\mu}(0,S_g,X_g,N_g)\Bigg] - Q\Bigg\} \\
&= \frac{1}{\sqrt{G}}\sum_{g=1}^{G}\left[\hat{\mu}(1,S_g,X_g,N_g) - \frac{A_g\hat{\mu}(1,S_g,X_g,N_g)}{\hat{\pi}(S_g)}\right] \\
&\quad + \frac{1}{\sqrt{G}}\sum_{g=1}^{G}\left[\frac{(1-A_g)\hat{\mu}(0,S_g,X_g,N_g)}{1-\hat{\pi}(S_g)} - \hat{\mu}(0,S_g,X_g,N_g)\right] \\
&\quad + \frac{1}{\sqrt{G}}\sum_{g=1}^{G}\frac{A_gN_g\bar{Y}_g}{\hat{\pi}(S_g)} - \frac{1}{\sqrt{G}}\sum_{g=1}^{G}\frac{(1-A_g)N_g\bar{Y}_g}{1-\hat{\pi}(S_g)} - \sqrt{G}Q \\
&\equiv R_{n,1}+R_{n,2}+R_{n,3}
\end{aligned}$$

Applying Lemma P.1, we have that:

$$\begin{aligned}
R_{n,1} &= \frac{1}{\sqrt{G}}\sum_{g=1}^{G}\left(1-\frac{1}{\pi(S_g)}\right)A_g\tilde{\mu}(1,S_g,X_g,N_g) + \frac{1}{\sqrt{G}}\sum_{g=1}^{G}(1-A_g)\tilde{\mu}(1,S_g,X_g,N_g) + o_p(1) \\
R_{n,2} &= \frac{1}{\sqrt{G}}\sum_{g=1}^{G}\left(\frac{1}{1-\pi(S_g)}-1\right)(1-A_g)\tilde{\mu}(0,S_g,X_g,N_g) + \frac{1}{\sqrt{G}}\sum_{g=1}^{G}A_g\tilde{\mu}(0,S_g,X_g,N_g) + o_p(1) \\
R_{n,3} &= \frac{1}{\sqrt{G}}\sum_{g=1}^{G}\frac{A_g}{\pi(S_g)}\tilde{W}_g - \frac{1}{\sqrt{G}}\sum_{g=1}^{G}\frac{1-A_g}{1-\pi(S_g)}\tilde{Z}_g + \frac{1}{\sqrt{G}}\sum_{g=1}^{G}(\mathbb{E}[W_g-Z_g|S_g]-\mathbb{E}[W_g-Z_g]),
\end{aligned}$$

where $W_g \equiv N_g\bar{Y}_g(1)$, $Z_g \equiv N_g\bar{Y}_g(0)$, and $\tilde{W}_g = W_g - \mathbb{E}[W_g|S_g]$, $\tilde{Z}_g = Z_g - \mathbb{E}[Z_g|S_g]$.

Thus, it implies that:

$$\begin{aligned}
\sqrt{G}(\hat{Q}-Q) &= \frac{1}{\sqrt{G}}\sum_{g=1}^{G}\left[\left(1-\frac{1}{\pi(S_g)}\right)\tilde{\mu}(1,S_g,X_g,N_g) - \tilde{\mu}(0,S_g,X_g,N_g) + \frac{\tilde{W}_g}{\pi(S_g)}\right]A_g \\
&\quad + \frac{1}{\sqrt{G}}\sum_{g=1}^{G}\left[\left(\frac{1}{1-\pi(S_g)}-1\right)\tilde{\mu}(0,S_g,X_g,N_g) + \tilde{\mu}(1,S_g,X_g,N_g) - \frac{\tilde{Z}_g}{1-\pi(S_g)}\right](1-A_g) \\
&\quad + \left\{\frac{1}{\sqrt{G}}\sum_{g=1}^{G}(\mathbb{E}[W_g-Z_g|S_g]-\mathbb{E}[W_g-Z_g])\right\} + o_p
\end{aligned}$$

Now let us consider $\sqrt{G}(\hat{H}-H)$. Note that we can rewrite it as

$$\begin{aligned}
\sqrt{G}(\hat{H}-H) &= \sqrt{G}\left(\frac{1}{G}\sum_{g=1}^{G} N_g - \mathbb{E}[N_g]\right) \\
&= \sqrt{G}\Bigg\{\frac{1}{G}\sum_{g=1}^{G}(N_g A_g - \mathbb{E}[N_g|S_g]A_g) \\
&+ \frac{1}{G}\sum_{g=1}^{G}(N_g(1-A_g) - \mathbb{E}[N_g|S_g](1-A_g)) \\
&+ \frac{1}{G}\sum_{g=1}^{G}(\mathbb{E}[N_g|S_g] - \mathbb{E}[N_g])\Bigg\}
\end{aligned}$$

Now recall that $\sqrt{G}(\hat{\tau}-\tau) = \frac{1}{\hat{H}}\left[\sqrt{G}(\hat{Q}-Q) - \tau\sqrt{G}(\hat{H}-H)\right]$ and define $\mathcal{D}_g \equiv \{W_g, Z_g, A_g, X_g, N_g\}$

$$\begin{aligned}
\Xi_1(\mathcal{D}_g, S_g) &= \left[\left(1 - \frac{1}{\pi(S_g)}\right)\tilde{\mu}(1, S_g, X_g, N_g) - \tilde{\mu}(0, S_g, X_g, N_g) + \frac{\tilde{W}_g}{\pi(S_g)}\right] - \tau(N_g - \mathbb{E}[N_g|S_g]) \\
\Xi_0(\mathcal{D}_g, S_g) &= \left[\left(\frac{1}{1-\pi(S_g)} - 1\right)\tilde{\mu}(0, S_g, X_g, N_g) + \tilde{\mu}(1, S_g, X_g, N_g) - \frac{\tilde{Z}_g}{1-\pi(S_g)}\right] - \tau(N_g - \mathbb{E}[N_g|S_g]) \\
\Xi_2(\mathcal{D}_g, S_g) &= (\mathbb{E}[W_g - Z_g|S_g] - \mathbb{E}[W_g - Z_g]) - \tau(\mathbb{E}[N_g|S_g] - \mathbb{E}[N_g]).
\end{aligned}$$

Then, we can define the variance estimand as follows:

$$\sqrt{G}(\hat{\tau}-\tau) = \frac{1}{\sum_{g=1}^{G} N_g}\left[\frac{1}{\sqrt{G}}\sum_{g=1}^{G}\Xi_1(\mathcal{D}_g, S_g)A_g + \frac{1}{\sqrt{G}}\sum_{g=1}^{G}\Xi_0(\mathcal{D}_g, S_g)(1-A_g) + \frac{1}{\sqrt{G}}\sum_{g=1}^{G}\Xi_2(\mathcal{D}_g, S_g)\right]$$

■

Step 2. Obtain the asymptotic distribution of $\sqrt{G}(\hat{\tau}-\tau)$

Applying Lemma P.2, we get that three terms are asymptotically normally distributed and independent from each other:

$$\begin{aligned}
&\frac{1}{\sqrt{G}}\sum_{g=1}^{G}\Xi_1(\mathcal{D}_g, S_g)A_g \xrightarrow{\text{d}} N(0, \sigma_1^2) \\
&\frac{1}{\sqrt{G}}\sum_{g=1}^{G}\Xi_0(\mathcal{D}_g, S_g)(1-A_g) \xrightarrow{\text{d}} N(0, \sigma_0^2) \\
&\frac{1}{\sqrt{G}}\sum_{g=1}^{G}\Xi_2(\mathcal{D}_g, S_g) \xrightarrow{\text{d}} N(0, \sigma_2^2),
\end{aligned}$$

where $\sigma_1^2 = \mathbb{E}[\pi(S_g)\Xi_1^2(\mathcal{D}_g, S_g)]$; $\sigma_0^2 = \mathbb{E}[(1-\pi(S_g))\Xi_0^2(\mathcal{D}_g, S_g)]$; $\sigma_2^2 = \mathbb{E}[\Xi_2^2(\mathcal{D}_g, S_g)]$. Finally, we can state the asymptotic normality

$$\sqrt{G}(\hat{\tau}-\tau) \xrightarrow{\text{d}} N\left(0, \frac{\sigma_1^2 + \sigma_0^2 + \sigma_2^2}{\mathbb{E}[N_g]^2}\right), \sigma^2 = \frac{\sigma_1^2 + \sigma_0^2 + \sigma_2^2}{\mathbb{E}[N_g]^2}.$$

■

Step 3. Obtain a consistent estimator

Note that we can write:

$$\begin{aligned}\Xi_2(\mathcal{D}_g, S_g) &= (\mathbb{E}[W_g - Z_g|S_g] - \mathbb{E}[W_g - Z_g]) - \tau(\mathbb{E}[N_g|S_g] - \mathbb{E}[N_g]) \\ &= (\mathbb{E}[W_g - Z_g|S_g] - \tau\mathbb{E}[N_g|S_g]) - (\mathbb{E}[W_g - Z_g] - \tau\mathbb{E}[N_g])\end{aligned}$$

Recalling that $\tau = \frac{\mathbb{E}[W_g - Z_g]}{\mathbb{E}[N_g]}$, it is clear that both terms are mean-zero. Since the second term is a constant, the variance of $\Xi_2(\mathcal{D}_g, S_g)$ becomes:

$$\begin{aligned}\sigma_2^2 &= Var\left[(\mathbb{E}[W_g - Z_g|S_g] - \tau\mathbb{E}[N_g|S_g]) - (\mathbb{E}[W_g - Z_g] - \tau\mathbb{E}[N_g])\right] \\ &= \mathbb{E}[(\mathbb{E}[W_g - Z_g|S_g] - \tau\mathbb{E}[N_g|S_g])^2].\end{aligned}$$

Then, the consistent estimator for σ_2^2 can be defined as:

$$\hat{\sigma}_2^2 = \left(\frac{1}{G_1(s)} \sum_{j \in I_1(s)} (N_j \bar{Y}_j - \hat{\tau} N_j)\right) - \left(\frac{1}{G_0(s)} \sum_{j \in I_0(s)} (N_j \bar{Y}_j - \hat{\tau} N_j)\right).$$

Let us define $I_a(s) \equiv \{j \in [g] : A_j = a, S_j = s\}$, $I(s) \equiv \{j \in [g] : S_j = s\}$, $G(s) \equiv \sum_{j \in [g]} I\{S_j = s\}$, $G_1(s) \equiv \sum_{j \in [g]} A_j I\{S_j = s\}$, $G_0(s) \equiv G(s) - G_1(s)$. Thus, following the results from (Jiang et all, 2023) and combining the terms, we can define the variance estimator, $\hat{\sigma}^2$, as follows:

$$\hat{\sigma}^2 = \frac{\frac{1}{G}\sum_{g=1}^{G}\left[A_g \hat{\Xi}_1^2(\mathcal{D}_g, S_g) + (1 - A_g)\hat{\Xi}_0^2(\mathcal{D}_g, S_g) + \hat{\Xi}_2^2(\mathcal{D}_g, S_g)\right]}{(\frac{1}{G}\sum_{g=1}^{G} N_g)^2},$$

where

$$\begin{aligned}\hat{\Xi}_1(\mathcal{D}_g, s) &= \tilde{\Xi}_1(s) - \frac{1}{G_1(s)} \sum_{j \in I_1(s)} \tilde{\Xi}_{1,j}(s) - \hat{\tau}\left(N_g - \frac{1}{G(s)} \sum_{j \in I(s)} N_j\right), \\ \hat{\Xi}_0(\mathcal{D}_g, s) &= \tilde{\Xi}_0(s) - \frac{1}{G_0(s)} \sum_{j \in I_0(s)} \tilde{\Xi}_{0,j}(s) - \hat{\tau}\left(N_g - \frac{1}{G(s)} \sum_{j \in I(s)} N_j\right), \\ \hat{\Xi}_2(s) &= \left(\frac{1}{G_1(s)} \sum_{j \in I_1(s)} N_j \bar{Y}_j\right) - \left(\frac{1}{G_0(s)} \sum_{j \in I_0(s)} N_j \bar{Y}_j\right) - \hat{\tau} \times \left(\frac{1}{G(s)} \sum_{j \in I(s)} N_j\right), \\ \tilde{\Xi}_1(\mathcal{D}_g, s) &= \left(1 - \frac{1}{\hat{\pi}(s)}\right)\hat{\mu}(1, s, X_g, N_g) - \hat{\mu}(0, s, X_g, N_g) + \frac{N_g \bar{Y}_g}{\hat{\pi}(s)}, \\ \tilde{\Xi}_0(\mathcal{D}_g, s) &= \left(\frac{1}{1 - \hat{\pi}(s)} - 1\right)\hat{\mu}(0, s, X_g, N_g) + \hat{\mu}(1, s, X_g, N_g) - \frac{N_g \bar{Y}_g}{1 - \hat{\pi}(s)}.\end{aligned}$$

■