1 Setup

We observe a random sample consisting of N_0 untreated observations and N_1 treated observations, such that $N_0 + N_1 = N$. Each observation i is characterized by potential outcomes $Y_i(1)$ and $Y_i(0)$ (treated and untreated outcome, respectively). Thus, $Y_i(0)$ is observed for control observations, and $Y_i(1)$ for treated observations.

For all observations, we observe both a scalar covariate X_i and a binary treatment indicator W_i ($W_i = 1$ if i received treatment, $W_i = 0$ otherwise). The vector of all observed outcomes, covariates, or treatment indicators is represented as $\mathbf{Y}, \mathbf{X}$, or $\mathbf{W}$ respectively.

We are concerned with estimation of the average treatment effect on the treated (henceforth ATET):

$$\tau = \mathbb{E}\left[Y_i(1) - Y_i(0) | W_i = 1\right] \tag{1}$$

Since we are using a matching estimator approach, I may sometimes refer to "the matching matrix". The matching matrix is a $N_0 \times N_1$ matrix whose ij'th entry is 1 if the following holds (and 0 otherwise):

$$|X_j - X_i| \leq |X_j - X_k| \ \forall \ k \neq i \text{s.t } W_k = 0$$

Intuitively, the ij'th entry in this matrix is 1 if the j'th treated unit is matched with the i'th control unit, and 0 otherwise. The matching matrix is a function only of $\mathbf{X}$ and $\mathbf{W}$, not $\mathbf{Y}$.

Identifying assumptions are:

1. Unconfoundedness - *For almost all x, $(Y_i(1), Y_i(0))$ is independent of W_i conditional on $X_i = x$.*
2. Overlap - *For some $c > 0$, and almost all x, $c \leq \mathbb{P}[W_i = 1 | X_i = x] \leq 1 - c$.*

For simplicity, we ignore the possibility of multiple control units matching to a single treatment unit, as this occurs with probability 0 in both the original sample and in the bootstrapped sample. Ignoring ties, we construct the estimated counterfactual outcome of interest:

$$\widehat{Y_i(0)} = \begin{cases} Y_k & \text{if } W_i = 1 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

Where Y_k is the unique (since we ignore ties) control unit that matches to unit i. Thus, the matching estimator of the ATET is:

$$\hat{\tau} = \frac{1}{N_1} \sum_{W_i = 1} \left(Y_i - \widehat{Y_i(0)} \right) \tag{3}$$

Let:

$$K_i = \begin{cases} 0 & \text{if } W_i = 1 \\ \sum_{W_j = 1} 1\left\{ i \text{ matches to } j \right\} & \text{if } W_i = 0 \end{cases} \tag{4}$$

When we ignore ties, K_i has a simple interpretation as the number of times unit i is used to construct $\widehat{Y_j(0)}$. We can now represent $\hat{\tau}$ as follows:

$$\hat{\tau} = \frac{1}{N_1} \sum_{i=1}^{N} \left(W_i - (1 - W_i)K_i\right) Y_i \tag{5}$$

It is easy to see that this is correct if one notes that this can be split up into $\frac{1}{N_1} \sum_{W_i=1} Y_i - \frac{1}{N_1} \sum_{W_i=0} K_i Y_i$. The first part gives the average treated outcome, and the latter part the average estimated counterfactual outcome.

2 Bootstrap Variance

The bootstrap algorithm is as follows:

1. Estimate $\hat{\tau}$ on the original sample.
2. Get residuals $\varepsilon_i = Y_i - \widehat{Y_i(0)} - \hat{\tau}$ for each treated observation i.
3. Draw, for each i, μ_i from a Bernoulli distribution with success probability $\frac{1}{2}$.
4. If $\mu_i \neq 1$, set it to -1. This simply changes failures from being reported as 0 to being reported as -1.
5. Create perturbed outcomes $Y_i^* = Y_i(0) + \hat{\tau} + \mu_i \varepsilon_i$.
6. Now estimate $\hat{\tau}_b$ on the bootstrapped sample $\{\mathbf{Y}^*, \mathbf{W}, \mathbf{X}\}$.
7. Repeat steps 1-6 for b bootstraps, and estimate the variance of $\hat{\tau}$ as the sample variance of $\hat{\tau}_b$ over the b bootstraps.

Note first that this algorithm does nothing to $\mathbf{X}$ and $\mathbf{W}$. Thus, it does nothing to the matching matrix. For each treated unit i, the estimated counterfactual outcome $\widehat{Y_i(0)}$ is unchanged by this procedure.

From (5) it follows that:

$$\begin{aligned}
\hat{\tau}_b &= \frac{1}{N_1} \sum_{W_i=1} Y_i^* - \frac{1}{N_1} \sum_{W_i=0} K_i Y_i && (6) \\
&= \frac{1}{N_1} \sum_{W_i=1} \left(\widehat{Y_i(0)} + \mu_i \varepsilon_i + \hat{\tau}\right) - \frac{1}{N_1} \sum_{W_i=0} K_i Y_i \\
&= \hat{\tau} + \frac{1}{N_1} \sum_{W_i=1} \widehat{Y_i(0)} + \frac{1}{N_1} \sum_{W_i=1} \mu_i \varepsilon_i - \frac{1}{N_1} \sum_{W_i=0} K_i Y_i && (7)
\end{aligned}$$

Note that the residuals ε_i are:

$$\begin{aligned}
\varepsilon_i &= Y_i - \widehat{Y_i(0)} - \hat{\tau} \\
&= Y_i - \widehat{Y_i(0)} - \frac{1}{N_1} \sum_{W_i=1} Y_i - \frac{1}{N_1} \sum_{W_i=0} K_i Y_i
\end{aligned}$$

Thus $\sum_{W_i=1} \varepsilon_i = 0$. This does not let us claim that $\sum_{W_i=1} \mu_i \varepsilon_i$ is 0, however.

Note also that by the logic of (5), it follows that:

$$\frac{1}{N_1}\left(\sum_{W_i=1} \widehat{Y_i(0)} - \sum_{W_i=0} K_i Y_i\right) = 0$$

Because both sums are adding up the same things. Thus, (7) becomes:

$$\hat{\tau}_b = \hat{\tau} + \frac{1}{N_1}\sum_{W_i=1} \mu_i \varepsilon_i \tag{8}$$

It should at this point be obvious that $\mathbb{E}^*[\hat{\tau}_b] = \hat{\tau}$, where $\mathbb{E}^*$ is the expectation over the bootstrap (or the expectation conditioned on $\mathbf{Y}$, $\mathbf{W}$, and $\mathbf{X}$).

Thus, using similar notation:

$$\begin{aligned}
\mathbb{V}^*(\hat{\tau}_b) &= \mathbb{E}^*\left[(\hat{\tau}_b - \hat{\tau})^2\right] \\
&= \mathbb{E}^*\left[\frac{1}{N_1^2}\left(\sum_{W_i=1} \mu_i \varepsilon_i\right)^2\right] \qquad \text{by substituting (8) and distributing}
\end{aligned}$$

Without loss of generality, assume that $W_i = 1$ for $i \in \{1, 2, ..., N_1\}$ - that is, the first N_1 units are the treated units. Then:

$$\begin{aligned}
\mathbb{V}^*(\hat{\tau}_b) &= \mathbb{E}^*\left[\frac{1}{N_1^2}\left(\mu_1\varepsilon_1 + \mu_2\varepsilon_2 + ... + \mu_{N_1}\varepsilon_{N_1}\right)^2\right] \qquad (9) \\
&= \mathbb{E}^*\Bigg[\frac{1}{N_1^2}\left(\mu_1^2\varepsilon_1^2 + \mu_1\varepsilon_1\mu_2\varepsilon_2 + ... + \mu_1\varepsilon_1\mu_{N_1}\varepsilon_{N_1}\right) \\
&\quad + \frac{1}{N_1^2}\left(\mu_2\varepsilon_2\mu_1\varepsilon_2 + \mu_2^2\varepsilon_2^2 + ... + \mu_2\varepsilon_2\mu_{N_1}\varepsilon_{N_1}\right) \\
&\quad + ... \\
&\quad + \frac{1}{N_1^2}\left(\mu_{N_1}\varepsilon_{N_1}\mu_1\varepsilon_1 + \mu_{N_1}\varepsilon_{N_1}\mu_2\varepsilon_2 + ... + \mu_{N_1}^2\varepsilon_{N_1}^2\right)\Bigg]
\end{aligned}$$

To get further into this, consider first $\mu_i^2\varepsilon_i^2$. Note that μ_i^2 is always 1, and thus:

$$\begin{aligned}
\mu_i^2\varepsilon_i^2 &= \varepsilon_i^2 \\
&= \left(Y_i - \widehat{Y_i(0)} - \hat{\tau}\right)^2 \\
&= Y_i^2 + \widehat{Y_i(0)}^2 + \hat{\tau}^2 + 2\hat{\tau}\widehat{Y_i(0)} - 2Y_i\widehat{Y_i(0)} - 2Y_i\hat{\tau} \qquad (10)
\end{aligned}$$

The other case that matters is $\mu_i \mu_j \varepsilon_i \varepsilon_j$. Note first that $\mu_i \mu_j$ is, just like μ_i, -1 or +1 with equal probability. For simplicitly, let $\mu_{ij} = \mu_i \mu_j$. Thus:

$$\mu_i \mu_j \varepsilon_i \varepsilon_j = \mu_{ij} \left(Y_i - \widehat{Y_i(0)} - \hat{\tau} \right) \left(Y_j - \widehat{Y_j(0)} - \hat{\tau} \right) \tag{11}$$

Beyond this point I am no longer entirely confident of what I'm doing.

Under the bootstrap, ε_i is a known quantity, not a random variable. The only randomness present in (10) or (11) is that brought in by μ_i and μ_j. Conveniently, that randomness disappears by definition in (10), but not in (11). Thus, from (9), we can go to:

$$\begin{aligned}
\mathbb{V}^*(\hat{\tau}_b) = \frac{1}{N_1^2} \big(& \mathbb{E}^* \left[\mu_1^2 \varepsilon_1^2\right] + \mathbb{E}^* \left[\mu_1 \mu_2 \varepsilon_1 \varepsilon_2\right] + ... + \mathbb{E}^* \left[\mu_1 \mu_{N_1} \varepsilon_1 \varepsilon_{N_1}\right] \\
& + \mathbb{E}^* \left[\mu_2 \mu_1 \varepsilon_2 \varepsilon_1\right] + \mathbb{E}^* \left[\mu_2^2 \varepsilon_2^2\right] + ... + \mathbb{E}^* \left[\mu_2 \mu_{N_1} \varepsilon_2 \varepsilon_{N_1}\right] \\
& + ... \\
& + \mathbb{E}^* \left[\mu_{N_1} \mu_1 \varepsilon_{N_1} \varepsilon_1\right] + \mathbb{E}^* \left[\mu_{N_1} \mu_2 \varepsilon_{N_1} \varepsilon_2\right] + ... + \mathbb{E}^* \left[\mu_{N_1}^2 \varepsilon_{N_1}^2\right] \big)
\end{aligned} \tag{12}$$

And by taking $\varepsilon_i \varepsilon_j$ out of the expectation (which I *think* I can do), we get:

$$\mathbb{V}^*(\hat{\tau}_b) = \frac{1}{N_1^2} \left(\sum_{i=1}^{N_1} \varepsilon_i^2 \right) \tag{13}$$

This is because $\mathbb{E}^* \left[\mu_i \mu_j\right] = 0$ by design, while $\mathbb{E}^* \left[\mu_i^2\right] = 1$.

This is quite promising.