Econometrica Supplementary Material

SUPPLEMENT TO "LONG-RUN EFFECTS OF DYNAMICALLY ASSIGNED TREATMENTS: A NEW METHODOLOGY AND AN EVALUATION OF TRAINING EFFECTS ON EARNINGS"

(*Econometrica*, Vol. 90, No. 3, May 2022, 1337–1354)

GERARD J. VAN DEN BERG
Department of Economics, University of Groningen, Department of Epidemiology, University Medical Center Groningen, IFAU, IZA, ZEW, CEPR, and CESifo

JOHAN VIKSTRÖM
IFAU and Department of Economics and UCLS, Uppsala University

APPENDIX B: ADDITIONAL PROOFS AND DERIVATIONS

B.1. *Estimation of ATET(t_s)*

WE SHOW THAT IF ASSUMPTIONS 1 AND 2 HOLD, the IPW estimator, $\widehat{\text{ATET}}$, is an unbiased estimator of $\text{ATET}(t_s) = E(Y(t_s) - Y(\infty) \mid T_s = t_s, T_u(t_s) \geq t_s)$.

For the first part of ATET(t_s), the estimator is

$$\frac{1}{\pi(t_s)N_{t_s}} \sum_{i \in T_{s,i}=t_s, T_{u,i}\geq t_s} Y_i,$$

for which we have

$$\begin{aligned} &E\left[\frac{1}{\pi(t_s)N_{t_s}} \sum_{i \in T_{s,i}=t_s, T_{u,i}\geq t_s} Y_i\right] \\ &= \frac{1}{\pi(t_s)} E\left[\frac{1}{N_{t_s}} \sum_{i \in T_{s,i}\geq t_s, T_{u,i}\geq t_s,} \mathbf{I}(T_{s,i} = t_s) Y_i\right] \\ &= \frac{1}{\pi(t_s)} E\left[\mathbf{I}(T_s = t_s) Y \middle| T_s \geq t_s, T_u \geq t_s\right] \\ &= \frac{1}{\pi(t_s)} \Pr(T_s = t_s | T_u \geq t_s, T_s \geq t_s) E[Y | T_s = t_s, T_u \geq t_s] \\ &= E[Y | T_s = t_s, T_u \geq t_s] \\ &= E\left[Y(t_s) | T_s = t_s, T_u(t_s) \geq t_s\right], \end{aligned} \tag{B.1}$$

where the last equality follows by Assumption 1 and the observational rule. Note that $\pi(t_s) = \Pr(T_s = t_s | T_s \geq t_s, T_u \geq t_s)$.

For the second part of ATET(t_s), the estimator without the normalization is

$$\frac{1}{\pi(t_s)N_{t_s}} \sum_{i \in T_{s,i} > T_{u,i} \geq t_s} w^{t_s}(T_{u,i}, X_i) Y_i, \tag{B.2}$$

Gerard J. van den Berg: gerard.van.den.berg@rug.nl
Johan Vikström: johan.vikstrom@ifau.uu.se

for which we have

$$
\begin{aligned}
&E\left[\frac{1}{\pi(t_s)N_{t_s}}\sum_{i\in T_{s,i}>T_{u,i}\geq t_s} w^{t_s}(T_{u,i},X_i)Y_i\right]\\
&\quad = E\left[\frac{1}{\pi(t_s)N_{t_s}}\sum_{i\in T_{s,i}\geq t_s,T_{u,i}\geq t_s} w^{t_s}(T_{u,i},X_i)\mathbf{I}(T_{s,i}>T_{u,i})Y_i\right]\\
&\quad = E\left[\frac{1}{\pi(t_s)N_{t_s}}\sum_{i\in T_{s,i}\geq t_s,T_{u,i}\geq t_s}\sum_{t_u=t_s}^{T_u^{\max}} w^{t_s}(t_u,X_i)\mathbf{I}(T_{s,i}>t_u,T_{u,i}=t_u)Y_i\right]\\
&\quad = \frac{1}{\pi(t_s)}E\left[\sum_{t_u=t_s}^{T_u^{\max}} w^{t_s}(t_u,X)\mathbf{I}(T_s>t_u,T_u=t_u)Y\middle|T_s\geq t_s,T_u\geq t_s\right]\\
&\quad = E_{X|T_s\geq t_s,T_u\geq t_s}\left[\frac{1}{\pi(t_s)}E\left[\sum_{t_u=t_s}^{T_u^{\max}} w^{t_s}(t_u,X)\mathbf{I}(T_s>t_u,T_u=t_u)Y\middle|T_s\geq t_s,T_u\geq t_s,X\right]\right].
\end{aligned}
$$

For sake of presentation, use the notation

$$h(t,X)=\Pr(T_u=t|T_s>t,T_u\geq t,X).$$

Next, using Assumptions 1 and 2, and using that $w^{t_s}(t_u,X)=\frac{p(t_s,X)}{\prod_{m=t_s}^{t_u}[1-p(m,X)]}$:

$$
\begin{aligned}
&E\left[w^{t_s}(t_u,X)\mathbf{I}(T_s>t_u,T_u=t_u)Y|T_s\geq t_s,T_u\geq t_s,X\right]\\
&\quad = w^{t_s}(t_u,X)\Pr(T_s>t_u,T_u=t_u|T_s\geq t_s,T_u\geq t_s,X)E[Y|T_s>t_u,T_u=t_u,X]\\
&\quad = \frac{p(t_s,X)}{\prod_{m=t_s}^{t_u}[1-p(m,X)]}h(t_u,X)\left[\prod_{m=t_s}^{t_u-1}[1-h(m,X)]\right]\\
&\qquad \times\left[\prod_{m=t_s}^{t_u}[1-p(m,X)]\right]E[Y|T_s>t_u,T_u=t_u,X]\\
&\quad = p(t_s,X)h(t_u,X)\left[\prod_{m=t_s}^{t_u-1}[1-h(m,X)]\right]E[Y|T_s>t_u,T_u=t_u,X]\\
&\quad = p(t_s,X)h(t_u,X)\left[\prod_{m=t_s}^{t_u-1}[1-h(m,X)]\right]E\left[Y(\infty)|T_s>t_u,T_u(\infty)=t_u,X\right]\\
&\quad = p(t_s,X)h(t_u,X)\left[\prod_{m=t_s}^{t_u-1}[1-h(m,X)]\right]\\
&\qquad \times E\left[Y(\infty)|T_s=t_s,T_u(\infty)=t_u,X\right]. \qquad\qquad \text{(B.3)}
\end{aligned}
$$

Note that the second equality follows from the definition of $w^{t_s}(t_u, X)$, the third equality by simplifying, the fourth equality by Assumption 1, and the fifth equality by applying Assumption 2 for $t_s, \dots, t_u$.

From (B.2) and (B.3),

$$E\left[\frac{1}{\pi(t_s)N_{t_s}} \sum_{i\in T_{s,i}>T_{u,i},T_{u,i}\geq t_s} w^{t_s}(T_{u,i}, X_i)Y_i\right]$$
$$= E_{X|T_s\geq t_s, T_u\geq t_s}\left[\frac{p(t_s, X)}{\pi(t_s)} \sum_{t_u=t_s}^{T_u^{\max}} h(t_u, X)\left[\prod_{m=t_s}^{t_u-1}[1-h(m, X)]\right] \times E\big[Y(\infty)|T_s = t_s, T_u(\infty) = t_u, X\big]\right]. \quad \text{(B.4)}$$

For sake of presentation, introduce the notation:

$$y\big(T_u(\infty) = t, X\big) = E\big[Y(\infty)|T_s = t_s, T_u(\infty) = t, X\big],$$
$$y\big(T_u(\infty) > t, X\big) = E\big[Y(\infty)|T_s = t_s, T_u(\infty) > t, X\big],$$
$$y\big(T_u(\infty) \geq t, X\big) = E\big[Y(\infty)|T_s = t_s, T_u(\infty) \geq t, X\big].$$

Using this notation, we have using that by construction $h(T_u^{\max}, X) = 1$,

$$h\big(T_u^{\max}, X\big)\left[\prod_{m=t_s}^{T_u^{\max}-1}[1-h(m, X)]\right]y\big(T_u(\infty) = T_u^{\max}, X\big)$$
$$= \left[\prod_{m=t_s}^{T_u^{\max}-1}[1-h(m)]\right]y\big(T_u(\infty) = T_u^{\max}, X\big). \quad \text{(B.5)}$$

Next, for time periods $T_u^{\max} - 1$ and $T_u^{\max} - 2$,

$$\left[\prod_{m=t_s}^{T_u^{\max}-1}[1-h(m, X)]\right]y\big(T_u(\infty) = T_u^{M}, X\big)$$
$$+ h\big(T_u^{\max} - 1, X\big)\left[\prod_{m=t_s}^{T_u^{\max}-2}[1-h(m, X)]\right]y\big(T_u(\infty) = T_u^{\max} - 1, X\big)$$
$$= \left[\prod_{m=t_s}^{T_u^{\max}-2}[1-h(m, X)]\right]y\big(T_u(\infty) \geq T_u^{\max} - 1, X\big), \quad \text{(B.6)}$$

and for arbitrary time periods t and $t - 1$,

$$\left[\prod_{m=t_s}^{t}[1-h(m, X)]\right]y\big(T_u(\infty) > t, X\big) + h(t, X)]$$

$$\times \left[\prod_{m=t_s}^{t-1} [1 - h(m, X)] \right] y(T_u(\infty) = t - 1, X)$$
$$= \left[\prod_{m=t_s}^{t-1} [1 - h(m, X)] \right] y(T_u(\infty) \geq t - 1, X). \tag{B.7}$$

Thus, using (B.5) for $T_u^{\max}$, (B.6) for $T_u^{\max} - 1$ and (B.7) for $t_s, \dots, T_u^{\max} - 2$, we have

$$\sum_{t_u=t_s}^{T_u^{\max}} h(t_u, X) \left[\prod_{m=t_s}^{t_u-1} [1 - h(m, X)] \right] E[Y(\infty) | T_s = t_s, T_u(\infty) = t_u, X]$$
$$= \sum_{t_u=t_s}^{T_u^{\max}} h(t_u, X) \left[\prod_{m=t_s}^{t_u-1} [1 - h(m, X)] \right] y(T_u(\infty) = t_u, X)$$
$$\overset{(B.5)}{=} \left[\prod_{m=t_s}^{T_u^{\max}-1} [1 - h(m, X)] \right] y(T_u(\infty) = T_u^{\max}, X)$$
$$+ \sum_{t_u=t_s}^{T_u^{\max}-1} h(t_u, X) \left[\prod_{m=t_s}^{t_u-1} [1 - h(m, X)] \right] y(T_u(\infty) = t_u, X)$$
$$\overset{(B.6)}{=} \left[\prod_{m=t_s}^{T_u^{\max}-2} [1 - h(m, X)] \right] y(T_u(\infty) \geq T_u^{\max} - 1, X)$$
$$+ \sum_{t_u=t_s}^{T_u^{\max}-2} h(t_u, X) \left[\prod_{m=t_s}^{t_u-1} [1 - h(m, X)] \right] y(T_u(\infty) = t_u, X)$$
$$\overset{(B.7)}{=} y(T_u(\infty) \geq t_s, X). \tag{B.8}$$

Thus, from (B.4) and (B.8),

$$E\left[\frac{1}{N_{t_s}} \sum_{i \in T_{s,i} > T_{u,i}, T_{u,i} \geq t_s} w^{t_s}(T_{u,i}, X_i) Y_i \right]$$
$$= E_{X|T_s \geq t_s, T_u \geq t_s} \left[\frac{p(t_s, X)}{\pi(t_s)} y(T_u(\infty) \geq t_s, X) \right]$$
$$= \frac{1}{\pi(t_s)} E_{X|T_s \geq t_s, T_u \geq t_s} \left[p(t_s, X) E[Y(\infty) | T_s = t_s, T_u(\infty) \geq t_s, X] \right]$$
$$= \frac{1}{\pi(t_s)} E_{X|T_s \geq t_s, T_u \geq t_s} \left[\Pr(T_s = t_s | T_s \geq t_s, T_u \geq t_s, X) \right.$$
$$\left. \times E[Y(\infty) | T_s = t_s, T_u(\infty) \geq t_s, X] \right]$$
$$= \frac{1}{\pi(t_s)} \Pr(T_s = t_s | T_s \geq t_s, T_u \geq t_s) E[Y(\infty) | T_s = t_s, T_u(\infty) \geq t_s]$$

$$= E\big[Y(\infty)|T_s = t_s, T_u(\infty) \geq t_s\big]. \tag{B.9}$$

This averaging over X is admitted by the common support assumption.

Finally, (B.1) and (B.9) imply that $E[\widehat{\text{ATET}}(t_s)] = \text{ATET}(t_s)$.

B.2. *Average Treatment Effect* ATE(t_s)

Section B.2 provides identification and estimation results for the average treatment effect of t_s on Y among all those who, if they were assigned to t_s, would still be in the initial state at that time t_s:

$$\text{ATE}(t_s) = E\big(Y(t_s) - Y(\infty) \mid T_u(t_s) \geq t_s\big).$$

In this section, the sequential unconfoundedness Assumption 2 refers to the variety for the ATE(t_s), that is, with conditional independence of P_t from both $Y(t)$ and $Y(\infty)$.

B.2.1. *Identification*

Identification of ATE(t_s) follows using similar reasoning as for ATET(t_s). For the second component of ATE(t_s), our assumptions give

$$\begin{aligned} E\big(Y(\infty)|T_u(t_s) \geq t_s\big) &= E_{X|T_u \geq t_s}\big[E\big(Y(\infty)|T_u(\infty) \geq t_s, X\big)\big] \\ &= E_{X|T_u \geq t_s}\big[E\big(Y(\infty)|T_s > t_s, T_u(\infty) \geq t_s, X\big)\big], \end{aligned} \tag{B.10}$$

and from (A.6):

$$\begin{aligned} &E\big(Y(\infty)|T_s > t_s, T_u(\infty) \geq t_s, X\big) \\ &\quad = \sum_{k=t_s}^{T_u^{\max}} h(k, X)\left[\prod_{m=t_s}^{k-1}\big[1 - h(m, X)\big]\right] E(Y|T_s > k, T_u = k, X). \end{aligned} \tag{B.11}$$

Thus, from (B.10)–(B.11) we have

$$\begin{aligned} &E\big(Y(\infty)|T_u(t_s) \geq t_s\big) \\ &\quad = E_{X|T_u \geq t_s}\left[\sum_{k=t_s}^{T_u^{\max}} h(k, X)\left[\prod_{m=t_s}^{k-1}\big[1 - h(m, X)\big]\right] E(Y|T_s > k, T_u = k, X)\right]. \end{aligned} \tag{B.12}$$

For the first component of ATE(t_s),

$$\begin{aligned} E\big(Y(t_s)|T_u(t_s) \geq t_s\big) &= E\big(Y(t_s)|T_u(\infty) \geq t_s\big) \\ &= E_{X|T_u \geq t_s}\big[E\big(Y(t_s)|T_u(\infty) \geq t_s, X\big)\big] \\ &= E_{X|T_u \geq t_s}\big[E\big(Y(t_s)|T_s = t_s, T_u(\infty) \geq t_s, X\big)\big] \\ &= E_{X|T_u \geq t_s}\big[E(Y|T_s = t_s, T_u \geq t_s, X)\big], \end{aligned} \tag{B.13}$$

where we apply Assumption 1 multiple times and where the second equality follows from the law of iterated expectations, the third from Assumption 2, and the fourth from the observational rule.

From (B.12) and (B.13), we thus obtain the following.

THEOREM 1—ATE version: *If Assumptions* 1 *and* 2 *hold*, *then*

$$\mathrm{ATE}(t_s) = E_{X|T_u \geq t_s}\big[E(Y|T_s = t_s, T_u \geq t_s, X)\big] - E_{X|T_u \geq t_s}\left[\sum_{k=t_s}^{T_u^{\max}} h(k, X)\left[\prod_{m=t_s}^{k-1}\big[1 - h(m, X)\big]\right]E(Y|T_s > k, T_u = k, X)\right],$$

where

$$h(t, X) = \Pr(T_u = t|T_s > t, T_u \geq t, X).$$

B.3. *Estimation*

If Assumptions 1 and 2 hold, then

$$\widehat{\mathrm{ATE}}(t_s) = \frac{1}{\sum_{i \in T_{s,i}=t_s, T_{u,i} \geq t_s} w_{\mathrm{ATE1}}^{t_s}(X_i)} \sum_{i \in T_{s,i}=t_s, T_{u,i} \geq t_s} w_{\mathrm{ATE1}}^{t_s}(X_i) Y_i - \frac{1}{\sum_{i \in T_{s,i}>T_{u,i} \geq t_s} w_{\mathrm{ATE0}}^{t_s}(T_{u,i}, X_i)} \sum_{i \in T_{s,i}>T_{u,i} \geq t_s} w_{\mathrm{ATE0}}^{t_s}(T_{u,i}, X_i) Y_i, \tag{B.14}$$

where

$$w_{\mathrm{ATE1}}^{t}(X) = \frac{1}{p(t, X)},$$

$$w_{\mathrm{ATE0}}^{t}(t_u, X) = \frac{1}{1 - p(t, X)} \frac{1}{\prod_{m=t+1}^{t_u}\big[1 - p(m, X)\big]},$$

is an unbiased estimator of $\mathrm{ATE}(t_s) = E(Y(t_s) - Y(\infty) \mid T_s \geq t_s, T_u(\infty) \geq t_s)$.

PROOF: For the first part of $\mathrm{ATE}(t_s)$, the estimator without the normalization is

$$\frac{1}{N_{t_s}} \sum_{i \in T_{s,i}=t_s, T_{u,i} \geq t_s} w_{\mathrm{ATE1}}^{t_s}(X_i) Y_i,$$

under Assumptions 1 and 2 we have

$$\begin{aligned}
&E\left[\frac{1}{N_{t_s}} \sum_{i \in T_{s,i}=t_s, T_{u,i} \geq t_s} w_{\mathrm{ATE1}}^{t_s}(X_i) Y_i\right] \\
&= E\left[\frac{1}{N_{t_s}} \sum_{i \in T_{s,i} \geq t_s, T_{u,i} \geq t_s,} w_{\mathrm{ATE1}}^{t_s}(X_i)\mathbf{I}(T_{s,i} = t_s) Y_i\right] \\
&= E\big[w_{\mathrm{ATE1}}^{t_s}(X)\mathbf{I}(T_s = t_s) Y|T_s \geq t_s, T_u \geq t_s\big] \\
&= E_{X|T_s \geq t_s, T_u \geq t_s}\big[E\big[w_{\mathrm{ATE1}}^{t_s}(X)\mathbf{I}(T_s = t_s) Y|T_s \geq t_s, T_u \geq t_s, X\big]\big]
\end{aligned}$$

$$
\begin{aligned}
&= E_{X|T_s\geq t_s, T_u\geq t_s}\left[\frac{1}{p(t_s, X)} p(t_s, X) E[Y|X, T_s = t_s, T_u \geq t_s, X]\right] \\
&= E_{X|T_s\geq t_s, T_u\geq t_s}\big[E\big[Y(t_s)|T_s = t_s, T_u(\infty) \geq t_s, X\big]\big] \\
&= E_{X|T_s\geq t_s, T_u\geq t_s}\big[E\big[Y(t_s)|T_s \geq t_s, T_u(\infty) \geq t_s, X\big]\big] \\
&= E\big[Y(t_s)|T_s \geq t_s, T_u(\infty) \geq t_s\big],
\end{aligned}
\tag{B.15}
$$

where the first three equalities follow by rewriting, the fourth by substituting for $w^{t_s}_{\text{ATE1}}(X) = \frac{1}{p(t_s,X)}$, the fifth by Assumption 1 and the observational rule, the sixth equality by Assumption 2 for period t_s, and the seventh by averaging over X.

For the second part of $\text{ATE}(t_s)$, the estimator without the normalization is

$$
\frac{1}{N_{t_s}} \sum_{i\in T_{s,i}>T_{u,i}\geq t_s} w^{t_s}_{\text{ATE0}}(T_{u,i}, X_i) Y_i,
$$

using similar reasoning as in (B.2) we have

$$
\begin{aligned}
&E\left[\frac{1}{N_{t_s}} \sum_{i\in T_{s,i}>T_{u,i}\geq t_s} w^{t_s}_{\text{ATE0}}(T_{u,i}, X_i) Y_i\right] \\
&\quad= E_{X|T_s\geq t_s, T_u\geq t_s} \\
&\qquad\times \left[E\left[\sum_{t_u=t_s}^{T_u^{\max}} w^{t_s}_{\text{ATE0}}(t_u, X)\mathbf{I}(T_s > t_u, T_u = t_u)Y \,\middle|\, T_s \geq t_s, T_u \geq t_s, X\right]\right].
\end{aligned}
\tag{B.16}
$$

Under Assumptions 1 and 2, and using the fact that $w^{t_s}_{\text{ATE0}}(t_u, X) = \frac{1}{\prod_{m=t_s}^{t_u}[1-p(m,X)]}$:

$$
\begin{aligned}
&E\big[w^{t_s}_{\text{ATE0}}(t_u, X)\mathbf{I}(T_s > t_u, T_u = t_u)Y|T_s \geq t_s, T_u \geq t_s, X\big] \\
&\quad= w^{t_s}_{\text{ATE0}}(t_u, X) \Pr(T_s > t_u, T_u = t_u|T_u \geq t_s, T_s \geq t_s, X) E[Y|T_s > t_u, T_u = t_u, X] \\
&\quad= \frac{1}{\prod_{m=t_s}^{t_u}[1 - p(m|X)]} h(t_u, X)\left[\prod_{m=t_s}^{t_u-1}[1 - h(m, X)]\right] \\
&\qquad\times \left[\prod_{m=t_s}^{t_u}[1 - p(m, X)]\right] E[Y|T_s > t_u, T_u = t_u, X] \\
&\quad= h(t_u, X)\left[\prod_{m=t_s}^{t_u-1}[1 - h(m, X)]\right] E[Y|T_s > t_u, T_u = t_u, X] \\
&\quad= h(t_u, X)\left[\prod_{m=t_s}^{t_u-1}[1 - h(m, X)]\right] E\big[Y(\infty)|T_s > t_u, T_u(\infty) = t_u, X\big] \\
&\quad= h(t_u, X)\left[\prod_{m=t_s}^{t_u-1}[1 - h(m, X)]\right] E\big[Y(\infty)|T_s \geq t_s, T_u(\infty) = t_u, X\big].
\end{aligned}
\tag{B.17}
$$

Note that the second equality follows from the definition of $w^{t_s}_{\text{ATE0}}(t_u, X)$, the third equality by simplifying, the fourth equality by Assumption 1, and the fifth equality by applying Assumption 2 for $t_s, \dots, t_u$.

Thus, from (B.16) and (B.17),

$$E\left[\frac{1}{N_{t_s}} \sum_{i \in T_{s,i} > T_{u,i}, T_{u,i} \geq t_s} w^{t_s}_{\text{ATE0}}(T_{u,i}, X_i) Y_i\right] = E_{X|T_s \geq t_s, T_u \geq t_s}\left[\sum_{t_u = t_s}^{T_u^{\max}} h(t_u, X)\left[\prod_{m=t_s}^{t_u - 1}[1 - h(m, X)]\right] \times E\big[Y(\infty)|T_s \geq t_s, T_u(\infty) = t_u, X\big]\right]. \quad \text{(B.18)}$$

Next, using similar reasoning as for (B.8) we have

$$\sum_{t_u = t_s}^{T_u^{\max}} h(t_u, X)\left[\prod_{m=t_s}^{t_u - 1}[1 - h(m, X)]\right] E\big[Y(\infty)|T_s \geq t_s, T_u(\infty) = t_u, X\big] = E\big[Y(\infty)|T_s \geq t_s, T_u(\infty) \geq t_s, X\big], \quad \text{(B.19)}$$

so that from (B.18) and (B.19),

$$E\left[\frac{1}{N_{t_s}} \sum_{i \in T_{s,i} > T_{u,i}, T_{u,i} \geq t_s} w^{t_s}_{\text{ATE0}}(T_{u,i}, X_i) Y_i\right] = E_{X|T_s \geq t_s, T_u \geq t_s}\big[E\big[Y(\infty)|T_s \geq t_s, T_u(\infty) \geq t_s, X\big]\big] = E\big[Y(\infty)|T_s \geq t_s, T_u(\infty) \geq t_s\big]. \quad \text{(B.20)}$$

Finally, (B.15) and (B.20) imply that $E[\widehat{\text{ATE}}(t_s)] = \text{ATE}(t_s)$. *Q.E.D.*

B.4. *Time-Varying Covariates*

B.4.1. *Identification*

Consider identification of $\text{ATET}(t_s) = E(Y(t_s) - Y(\infty) \mid T_s = t_s, T_u(t_s) \geq t_s)$. For the first component, as before

$$E\big(Y(t_s)|T_s = t_s, T_u(t_s) \geq t_s\big) = E(Y|T_s = t_s, T_u \geq t_s). \quad \text{(B.21)}$$

For the second component, by no-anticipation and averaging over $X^-_{t_s}$,

$$E\big(Y(\infty)|T_s = t_s, T_u(t_s) \geq t_s\big) = E\big(Y(\infty)|T_s = t_s, T_u(\infty) \geq t_s\big) = E_{X^-_{t_s}|T_s = t_s, T_u \geq t_s}\big[E\big(Y(\infty)|T_s = t_s, T_u(\infty) \geq t_s, X^-_{t_s}\big)\big]. \quad \text{(B.22)}$$

Then, by Assumption 4 for period t_s,

$$E(Y(\infty)|T_s = t_s, T_u(\infty) \geq t_s, X_{t_s}^-) = E(Y(\infty)|T_s > t_s, T_u(\infty) \geq t_s, X_{t_s}^-).$$

Next, using the notation $h(X_t^-) = \Pr(T_u = t|T_s > t, T_u \geq t, X_t^-)$ we have

$$\begin{aligned} &E(Y(\infty)|T_s > t_s, T_u(\infty) \geq t_s, X_{t_s}^-) \\ &\quad = h(X_{t_s}^-)E(Y(\infty)|T_s > t_s, T_u(\infty) = t_s, X_{t_s}^-) \\ &\qquad + [1 - h(X_{t_s}^-)]E(Y(\infty)|T_s > t_s, T_u(\infty) > t_s, X_{t_s}^-) \\ &\quad = h(X_{t_s}^-)E(Y|T_s > t_s, T_u = t_s, X_{t_s}^-) \\ &\qquad + [1 - h(X_{t_s}^-)]E(Y(\infty)|T_s > t_s, T_u(\infty) > t_s, X_{t_s}^-), \end{aligned} \tag{B.23}$$

where the second equality follows from no-anticipation and the observational rule. Also, under no-anticipation, $\Pr(T_u = t|T_s > t, T_u(\infty) \geq t, X_t^-) = \Pr(T_u = t|T_s > t, T_u \geq t, X_t^-) = h(X_t^-)$, and the treatment probability $h(X_{t_s}^-)$ is observed. Next,

$$\begin{aligned} &E(Y(\infty)|T_s > t_s, T_u(\infty) > t_s, X_{t_s}^-) \\ &\quad = E_{X_{t_s+1}^-|T_s>t_s, T_u>t_s, X_{t_s}^-}[E(Y(\infty)|T_s > t_s, T_u(\infty) > t_s, X_{t_s+1}^-)] \\ &\quad = E_{X_{t_s+1}^-|T_s>t_s, T_u>t_s, X_{t_s}^-}[E(Y(\infty)|T_s > t_s + 1, T_u(\infty) > t_s, X_{t_s+1}^-)] \\ &\quad = E_{X_{t_s+1}^-|T_s>t_s, T_u>t_s, X_{t_s}^-}[E(Y(\infty)|T_s > t_s + 1, T_u(\infty) \geq t_s + 1, X_{t_s+1}^-)], \end{aligned} \tag{B.24}$$

where the first equality follows from the law of iterated expectations, the second equality from Assumption 4 for period $t_s + 1$, and the third equality by rewriting. Here, the covariates $X_{t_s+1}^-$ may include $X_{t_s}^-$. From (B.23), by replacing t_s with $t_s + 1$,

$$\begin{aligned} &E(Y(\infty)|T_s > t_s + 1, T_u(\infty) \geq t_s + 1, X_{t_s+1}^-) \\ &\quad = h(X_{t_{s+1}}^-)E(Y(\infty)|T_s > t_s + 1, T_u(\infty) = t_s + 1, X_{t_s+1}^-) \\ &\qquad + [1 - h(X_{t_{s+1}}^-)]E(Y(\infty)|T_s > t_s + 1, T_u(\infty) > t_s + 1, X_{t_s+1}^-). \end{aligned} \tag{B.25}$$

Next, from (B.23) and (B.25),

$$\begin{aligned} &E(Y(\infty)|T_s > t_s, T_u(\infty) \geq t_s, X_{t_s}^-) \\ &\quad = h(X_{t_s}^-)E(Y|T_s > t_s, T_u = t_s, X_{t_s}^-) \\ &\qquad + [1 - h(X_{t_s}^-)]E_{X_{t_s+1}^-|T_s>t_s, T_u>t_s, X_{t_s}^-}[h(X_{t_s+1}^-) \\ &\qquad \times E(Y|T_s > t_s + 1, T_u = t_s + 1, X_{t_s+1}^-) \\ &\qquad + [1 - h(X_{t_s+1}^-)]E(Y(\infty)|T_s > t_s + 1, T_u(\infty) > t_s + 1, X_{t_s+1}^-)]. \end{aligned}$$

Then, using (B.24) for $t_s + 1$ gives

$$\begin{aligned} &E(Y(\infty)|T_s > t_s, T_u \geq t_s, X_{t_s}^-) \\ &\quad = h(X_{t_s}^-)E(Y|T_s > t_s, T_u = t_s, X_{t_s}^-) \end{aligned}$$

$$
\begin{aligned}
&+ [1-h(X_{t_s}^-)]E_{X_{t_s+1}^-|T_s>t_s,T_u>t_s,X_{t_s}^-}\big[h(X_{t_s+1}^-)E(Y|T_s>t_s+1,T_u=t_s+1,X_{t_s+1}^-)\\
&+ [1-h(X_{t_s+1}^-)]\\
&\times E_{X_{t_s+2}^-|T_s>t_s+1,T_u>t_s+1,X_{t_s+1}^-}\big[E(Y(\infty)|T_s>t_s+2,T_u(\infty)\geq t_s+2,X_{t_s+2}^-)\big]\big],
\end{aligned}
$$

and (B.23) for t_s+2 gives

$$
\begin{aligned}
&E(Y(\infty)|T_s>t_s,T_u\geq t_s,X_{t_s}^-)\\
&\quad= h(X_{t_s}^-)E(Y|T_s>t_s,T_u=t_s,X_{t_s}^-)\\
&\qquad+ [1-h(X_{t_s}^-)]E_{X_{t_s+1}^-|T_s>t_s,T_u>t_s,X_{t_s}^-}\big[h(X_{t_s+1}^-)E(Y|T_s>t_s+1,T_u=t_s+1,X_{t_s+1}^-)\\
&\qquad+ [1-h(X_{t_s+1}^-)]E_{X_{t_s+2}^-|T_s>t_s+1,T_u>t_s+1,X_{t_s+1}^-}\big[h(X_{t_s+2}^-)\\
&\qquad\times E(Y|T_s>t_s+2,T_u=t_s+2,X_{t_s+2}^-)\\
&\qquad+ [1-h(X_{t_s+2}^-)]E(Y(\infty)|T_s>t_s+2,T_u(\infty)>t_s+2,X_{t_s+2}^-)\big]\big],
\end{aligned}
$$

and (B.24) for t_s+2 gives

$$
\begin{aligned}
&E(Y(\infty)|T_s>t_s,T_u\geq t_s,X_{t_s}^-)\\
&\quad= h(X_{t_s}^-)E(Y|T_s>t_s,T_u=t_s,X_{t_s}^-)\\
&\qquad+ [1-h(X_{t_s}^-)]E_{X_{t_s+1}^-|T_s>t_s,T_u>t_s,X_{t_s}^-}\big[h(X_{t_s+1}^-)E(Y|T_s>t_s+1,T_u=t_s+1,X_{t_s+1}^-)\\
&\qquad+ [1-h(X_{t_s+1}^-)]E_{X_{t_s+2}^-|T_s>t_s+1,T_u>t_s+1,X_{t_s+1}^-}\big[h(X_{t_s+2}^-)\\
&\qquad\times E(Y|T_s>t_s+2,T_u=t_s+2,X_{t_s+2}^-)\\
&\qquad+ [1-h(X_{t_s+2}^-)]\\
&\qquad\times E_{X_{t_s+3}^-|T_s>t_s+2,T_u>t_s+2,X_{t_s+2}^-}\big[E(Y(\infty)|T_s>t_s+3,T_u(\infty)\geq t_s+3,X_{t_s+1}^-)\big]\big]\big].
\end{aligned}
$$

and iteratively using (B.23) and (B.24) for $t_s+3,\ldots,T_u^{\max}$ we have

$$
\begin{aligned}
&E(Y(\infty)|T_s>t_s,T_u\geq t_s,X_{t_s}^-)\\
&\quad= h(X_{t_s}^-)E(Y|T_s>t_s,T_u=t_s,X_{t_s}^-)\\
&\qquad+ [1-h(X_{t_s}^-)]E_{X_{t_s+1}^-|T_s>t_s,T_u>t_s,X_{t_s}^-}\big[h(X_{t_s+1}^-)E(Y|T_s>t_s+1,T_u=t_s+1,X_{t_s+1}^-)\\
&\qquad+ [1-h(X_{t_s+1}^-)]E_{X_{t_s+2}^-|T_s>t_s+1,T_u>t_s+1,X_{t_s+1}^-}\big[h(X_{t_s+2}^-)\\
&\qquad\times E(Y|T_s>t_s+2,T_u=t_s+2,X_{t_s+2}^-)+\ldots\\
&\qquad+ [1-h(X_{T_u^{\max}-1}^-)]E_{X_{T_u^{\max}}^-|T_s>T_u^{\max}-1,T_u>T_u^{\max}-1,X_{T_u^{\max}-1}^-}\\
&\qquad\times \big[p(X_{T_u^{\max}}^-)E(Y|T_s>T_u^{\max},T_u=T_u^{\max},X_{T_u^{\max}}^-)\big]\ldots\big]\big].
\end{aligned}
\tag{B.26}
$$

Finally, combining (B.22), (B.26) and (B.21) gives the result in Theorem 2.

B.4.2. *Estimation*

If no-anticipation and Assumption 4 hold, an unbiased estimator of $\text{ATET}(t_s)$ is

$$\begin{aligned}\widehat{\text{ATET}}(t_s) &= \frac{1}{\pi(t_s)N_{t_s}} \sum_{i \in T_{s,i}=t_s, T_{u,i} \geq t_s} Y_i \\ &\quad - \frac{1}{\sum_{i \in T_{s,i} > T_{u,i} \geq t_s} w^{t_s}\left(T_{u,i}, X_i^-\right)} \sum_{i \in T_{s,i} > T_{u,i} \geq t_s} w^{t_s}\left(T_{u,i}, X_i^-\right) Y_i \end{aligned} \tag{B.27}$$

where

$$w^{t_s}\left(t_u, X^-\right) = \frac{p\left(t_s, X_{t_s}^-\right)}{\prod_{m=t_s}^{t_u}\left[1 - p\left(m, X_m^-\right)\right]}.$$

PROOF: For the first part of $\text{ATET}(t_s)$, we have from (B.1),

$$E\left[\frac{1}{\pi(t_s)N_{t_s}} \sum_{i \in T_{s,i}=t_s, T_{u,i} \geq t_s} Y_i\right] = E\left[Y(t_s) | T_s = t_s, T_u(t_s) \geq t_s\right]. \tag{B.28}$$

For the second part of $\text{ATET}(t_s)$, the estimator without the normalization is

$$\frac{1}{\pi(t_s)N_{t_s}} \sum_{i \in T_{s,i} > T_{u,i} \geq t_s} w^{t_s}\left(T_{u,i}, X_i^-\right) Y_i,$$

using similar reasoning as for (B.2) we have

$$\begin{aligned} &E\left[\frac{1}{\pi(t_s)N_{t_s}} \sum_{i \in T_{s,i} > T_{u,i} \geq t_s} w^{t_s}\left(T_{u,i}, X_i^-\right) Y_i\right] \\ &\quad = E_{X_{t_s}^- | T_s \geq t_s, T_u \geq t_s} \\ &\qquad \times \left[\frac{1}{\pi(t_s)} E\left[\sum_{t_u=t_s}^{T_u^{\max}} w^{t_s}\left(t_u, X^-\right) \mathbf{I}(T_s > t_u, T_u = t_u) Y | T_s \geq t_s, T_u \geq t_s, X_{t_s}^-\right]\right]. \end{aligned} \tag{B.29}$$

We use the notation

$$h\left(t, X_t^-\right) = \Pr\left(T_u = t | T_s > t, T_u \geq t, X_t^-\right).$$

If no-anticipation and Assumption 4 hold, and since $w^{t_s}(t_u, X^-) = \frac{p(t_s, X_{t_s}^-)}{\prod_{m=t_s}^{t_u}[1-p(m, X_m^-)]}$:

$$\begin{aligned} &E\left[w^{t_s}\left(t_s+1, X^-\right) \mathbf{I}(T_s > t_s+1, T_u = t_s+1) Y | T_s \geq t_s, T_u \geq t_s, X_{t_s}^-\right] \\ &\quad = w^{t_s}\left(t_s+1, X^-\right) \Pr\left(T_s > t_s, T_u > t_s | T_u \geq t_s, T_s \geq t_s, X_{t_s}^-\right) \\ &\qquad \times \Pr\left(T_s > t_s+1, T_u = t_s+1 | T_s > t_s, T_u > t_s, X_{t_s}^-\right) \end{aligned}$$

$$
\begin{aligned}
&\times E\big[Y|T_s > t_s+1, T_u = t_s+1, X_{t_s}^-\big] \\
&= \frac{p(t_s, X_{t_s}^-)}{\prod_{m=t_s}^{t_s+1}\big[1-p(m, X_m^-)\big]}\big[1-h(t_s, X_{t_s}^-)\big]\big[1-p(t_s, X_{t_s}^-)\big] \\
&\quad \times \Pr\big(T_s > t_s+1, T_u = t_s+1|T_s > t_s, T_u > t_s, X_{t_s}^-\big) \\
&\quad \times E\big[Y|T_s > t_s+1, T_u = t_s+1, X_{t_s}^-\big] \\
&= \frac{p(t_s, X_{t_s}^-)\big[1-h(t_s, X_{t_s}^-)\big]}{\big[1-p(t_s+1, X_{t_s+1}^-)\big]}\Pr\big(T_s > t_s+1, T_u = t_s+1|T_s > t_s, T_u > t_s, X_{t_s}^-\big) \\
&\quad \times E\big[Y|T_s > t_s+1, T_u = t_s+1, X_{t_s}^-\big]. \qquad \text{(B.30)}
\end{aligned}
$$

Next,

$$
\begin{aligned}
&\frac{\Pr\big(T_s > t_s+1, T_u = t_s+1|T_s > t_s, T_u > t_s, X_{t_s}^-\big)}{1-p(t_s+1, X_{t_s+1}^-)} \\
&= E_{X_{t_s+1}^-|T_s>t_s, T_u>t_s, X_{t_s}^-}\left[\frac{\Pr\big(T_s > t_s+1, T_u = t_s+1|T_s > t_s, T_u > t_s, X_{t_s+1}^-\big)}{1-p(t_s+1, X_{t_s+1}^-)}\right] \\
&= E_{X_{t_s+1}^-|T_s>t_s, T_u>t_s, X_{t_s}^-}\left[\frac{1-p(t_s+1, X_{t_s+1}^-)\Pr\big(T_u = t_s+1|T_s > t_s+1, T_u > t_s, X_{t_s+1}^-\big)}{1-p(t_s+1, X_{t_s+1}^-)}\right] \\
&= E_{X_{t_s+1}^-|T_s>t_s, T_u>t_s, X_{t_s}^-}\big[\Pr\big(T_u = t_s+1|T_s > t_s+1, T_u > t_s, X_{t_s+1}^-\big)\big] \\
&= E_{X_{t_s+1}^-|T_s>t_s, T_u>t_s, X_{t_s}^-}\big[\Pr\big(T_u = t_s+1|T_s > t_s, T_u > t_s, X_{t_s+1}^-\big)\big] \\
&= \Pr\big(T_u = t_s+1|T_s > t_s, T_u > t_s, X_{t_s}^-\big) = h\big(t_s+1, X_{t_s}^-\big). \qquad \text{(B.31)}
\end{aligned}
$$

Note that the fourth equality follows from Assumption 4. Then, by (B.30) and (B.31), and using no-anticipation and Assumption 4,

$$
\begin{aligned}
&E\big[w^{t_s}(t_s+1, X^-)\mathbf{I}(T_s > t_s+1, T_u = t_s+1)Y|T_s \geq t_s, T_u \geq t_s, X_{t_s}^-\big] \\
&= p(t_s, X_{t_s}^-)h(t_s+1, X_{t_s}^-)\big[1-h(t_s, X_{t_s}^-)\big]E\big[Y|T_s > t_s+1, T_u = t_s+1, X_{t_s}^-\big] \\
&= p(t_s, X_{t_s}^-)h(t_s+1, X_{t_s}^-)\big[1-h(t_s, X_{t_s}^-)\big] \\
&\quad \times E\big[Y(\infty)|T_s > t_s+1, T_u(\infty) = t_s+1, X_{t_s}^-\big] \\
&= p(t_s, X_{t_s}^-)h(t_s+1, X_{t_s}^-)\big[1-h(t_s, X_{t_s}^-)\big] \\
&\quad \times E\big[Y(\infty)|T_s = t_s, T_u(\infty) = t_s+1, X_{t_s}^-\big], \qquad \text{(B.32)}
\end{aligned}
$$

where the second equality follows from no-anticipation, and the third by Assumption 4.

By similar reasoning as for (B.30)–(B.32), we have

$$
E\big[w^{t_s}(t_u, X^-)\mathbf{I}(T_s > t_u, T_u = t_u)Y|T_s \geq t_s, T_u \geq t_s, X_{t_s}^-\big]
$$

$$= p\big(t_s, X_{t_s}^-\big)h\big(t_u, X_{t_s}^-\big)\left[\prod_{m=t_s}^{t_u-1}\big[1-h\big(m, X_{t_s}^-\big)\big]\right]$$
$$\times E\big[Y(\infty)|T_s = t_s, T_u(\infty) = t_u, X_{t_s}^-\big]. \tag{B.33}$$

Thus, from (B.29)–(B.33),

$$E\left[\frac{1}{\pi(t_s)N_{t_s}} \sum_{i\in T_{s,i}>T_{u,i}\geq t_s} w^{t_s}\big(T_{u,i}, X_i^-\big)Y_i\right]$$
$$= E_{X_{t_s}^-|T_s\geq t_s, T_u\geq t_s}\left[\frac{p\big(t_s, X_{t_s}^-\big)}{\pi(t_s)} \sum_{t_u=t_s}^{T_u^{\max}} h\big(t_u, X_{t_s}^-\big)\left[\prod_{m=t_s}^{t_u-1}\big[1-h\big(m, X_{t_s}^-\big)\big]\right]\right.$$
$$\left.\times E\big[Y(\infty)|T_s = t_s, T_u(\infty) = t_u, X_{t_s}^-\big]\right]. \tag{B.34}$$

Next, by similar reasoning as for (B.5)–(B.8) we have

$$\sum_{t_u=t_s}^{T_u^{\max}} h\big(t_u, X_{t_s}^-\big)\left[\prod_{m=t_s}^{t_u-1}\big[1-h\big(m, X_{t_s}^-\big)\big]\right]E\big[Y(\infty)|T_s = t_s, T_u(\infty) = t_u, X_{t_s}^-\big]$$
$$= E\big[Y(\infty)|T_s = t_s, T_u(\infty) \geq t_s, X_{t_s}^-\big], \tag{B.35}$$

so that from (B.34) and (B.35),

$$E\left[\frac{1}{\pi(t_s)N_{t_s}} \sum_{i\in T_{s,i}>T_{u,i}\geq t_s} w^{t_s}\big(T_{u,i}, X_i^-\big)Y_i\right]$$
$$= E_{X_{t_s}^-|T_s\geq t_s, T_u\geq t_s}\left[\frac{p\big(t_s, X_{t_s}^-\big)}{\pi(t_s)}E\big[Y(\infty)|T_s = t_s, T_u(\infty) \geq t_s, X_{t_s}^-\big]\right]$$
$$= \frac{1}{\pi(t_s)}E_{X_{t_s}^-|T_s\geq t_s, T_u\geq t_s}\big[\Pr\big(T_s = t_s|T_s \geq t_s, T_u \geq t_s, X_{t_s}^-\big)$$
$$\times E\big[Y(\infty)|T_s = t_s, T_u(\infty) \geq t_s, X_{t_s}^-\big]$$
$$= \frac{1}{\pi(t_s)}\Pr(T_s = t_s|T_s \geq t_s, T_u \geq t_s)E\big[Y(\infty)|T_s = t_s, T_u(\infty) \geq t_s\big]$$
$$= E\big[Y(\infty)|T_s = t_s, T_u(\infty) \geq t_s\big]. \tag{B.36}$$

Finally, (B.28) and (B.36) imply that $E[\widehat{\text{ATET}}(t_s)] = \text{ATET}(t_s)$. *Q.E.D.*

B.5. *Right-Censored Durations*

B.5.1. *Identification*

Consider identification of $\text{ATET}(t_s) = E(Y(t_s) - Y(\infty) \mid T_s = t_s, T_c > t_s, T_u(t_s) \geq t_s)$ under Assumptions 1, 2, and 5. First, consider $E(Y(\infty)|T_s = t_s, T_c > t_s, T_u(t_s) \geq t_s)$. Initially, by Assumption 1 and the law of iterated expectations:

$$\begin{aligned} &E\big(Y(\infty) \mid T_s = t_s, T_c > t_s, T_u(t_s) \geq t_s\big) \\ &\quad = E\big(Y(\infty) \mid T_s = t_s, T_c > t_s, T_u(\infty) \geq t_s\big) \\ &\quad = E_{X|T_s=t_s,T_c>t_s,T_u\geq t_s}\big[E\big(Y(\infty)|_s = t_s, T_c > t_s, T_u(\infty) \geq t_s, X\big)\big], \end{aligned} \tag{B.37}$$

where the averaging over X is possible given common support. Next, if Assumption 2 holds for period t_s we have

$$E\big(Y(\infty) \mid T_s = t_s, T_c > t_s, T_u(\infty) \geq t_s, X\big) = E\big(Y(\infty) \mid T_s > t_s, T_c > t_s, T_u(\infty) \geq t_s, X\big).$$

Then, by the law of iterated expectations,

$$\begin{aligned} &E\big(Y(\infty) \mid T_s > t_s, T_c > t_s, T_u(\infty) \geq t_s, X\big) \\ &\quad = \Pr\big(T_u = t_s | T_s > t_s, T_c > t_s, T_u(\infty) \geq t_s, X\big) \\ &\qquad \times E\big(Y(\infty) | T_s > t_s, T_c > t_s, T_u(\infty) = t_s, X\big) \\ &\qquad + \Pr\big(T_u > t_s | T_s > t_s, T_c > t_s, T_u(\infty) \geq t_s, X\big) \\ &\qquad \times E\big(Y(\infty) | T_s > t_s, T_c > t_s, T_u(\infty) > t_s, X\big), \end{aligned} \tag{B.38}$$

decomposing the counterfactual outcome under never treatment into average outcomes for individuals with $T_u = t_s$ and $T_u > t_s$. For the group with $T_u = t_s$ in (B.38), we have by Assumption 1,

$$E\big(Y(\infty) | T_s > t_s, T_c > t_s, T_u = t_s, X\big) = E(Y | T_s > t_s, T_c > t_s, T_u = t_s, X), \tag{B.39}$$

and the probabilities $\Pr(T_u = t_s | T_s > t_s, T_c > t_s, T_u(\infty) \geq t_s, X)$ and $\Pr(T_u > t_s | T_s > t_s, T_c > t_s, T_u(\infty) \geq t_s, X)$ are also observed.

For the group, with $T_u > t_s$, in (B.38), we have

$$\begin{aligned} &E\big(Y(\infty) | T_s > t_s, T_c > t_s, T_u > t_s, X\big) \\ &\quad = E\big(Y(\infty) | T_s > t_s, T_c > t_s + 1, T_u > t_s, X\big) \\ &\quad = E\big(Y(\infty) | T_s > t_s + 1, T_c > t_s + 1, T_u > t_s, X\big) \\ &\quad = E\big(Y(\infty) | T_s > t_s + 1, T_c > t_s + 1, T_u \geq t_s + 1, X\big), \end{aligned}$$

where the first equality follows from Assumption 5 for period $t_s + 1$, the second from Assumption 2 for period $t_s + 1$, and the third equality by rewriting. Next, for sake of presentation, let us introduce some auxiliary notation:

$$h_c(t, X) = \Pr(T_u = t \mid T_s > t, T_c > t, T_u \geq t, X).$$

Using this notation and using (B.38) by replacing t_s with $t_s + 1$, we have

$$\begin{aligned} &E\big(Y(\infty) \mid T_s > t_s + 1, T_c > t_s + 1, T_u \geq t_s + 1, X\big) \\ &\quad = h_c(t_s + 1, X) E\big(Y(\infty) | T_s > t_s + 1, T_c > t_s + 1, T_u = t_s + 1, X\big) \\ &\qquad + \big[1 - h_c(t_s + 1, X)\big] \\ &\qquad \times E\big(Y(\infty) | T_s > t_s + 1, T_c > t_s + 1, T_u > t_s + 1, X\big), \end{aligned} \tag{B.40}$$

so that (B.38)–(B.40) give

$$\begin{aligned}
&E\big(Y(\infty)|T_s > t_s, T_c > t_s, T_u \geq t_s, X\big) \\
&\quad = h_c(t_s, X)E(Y|T_s > t_s, T_c > t_s, T_u = t_s, X) \\
&\qquad + \big[1 - h_c(t_s, X)\big]h_c(t_s + 1, X) \\
&\qquad \times E(Y|T_s > t_s + 1, T_c > t_s + 1, T_u = t_s + 1, X) \\
&\qquad + \big[1 - h_c(t_s, X)\big]\big[1 - h_c(t_s + 1, X)\big] \\
&\qquad \times E\big(Y(\infty)|T_s > t_s + 1, T_c > t_s + 1, T_u > t_s + 1, X\big).
\end{aligned}$$

Then, iteratively using (B.38) and (B.39) for $t_s + 2, \dots, T_u^{\max}$ we have

$$\begin{aligned}
&E\big(Y(\infty)|T_s > t_s, T_c > t_s, T_u \geq t_s, X\big) \\
&\quad = \sum_{k=t_s}^{T_u^{\max}} h_c(k, X)\left[\prod_{m=t_s}^{k-1}\big[1 - h_c(m, X)\big]\right] \\
&\qquad \times E(Y|T_s > k, T_c > k, T_u = k, X).
\end{aligned} \tag{B.41}$$

Then, from (B.37)–(B.41),

$$\begin{aligned}
&E\big(Y(\infty) \mid T_s = t_s, T_c > t_s, T_u(t_s) \geq t_s, X\big) \\
&\quad = E_{X|T_s=t_s, T_c>t_s, T_u\geq t_s}\left[\sum_{k=t_s}^{T_u^{\max}} h_c(k, X)\left[\prod_{m=t_s}^{k-1}\big[1 - h_c(m, X)\big]\right]\right. \\
&\qquad \left. \times E(Y|T_s > k, T_c > k, T_u = k, X)\right].
\end{aligned} \tag{B.42}$$

Second, for $E(Y(t_s)|T_s = t_s, T_c > t_s, T_u(t_s) \geq t_s)$, Assumption 1 and the law of iterated expectations give

$$\begin{aligned}
&E\big(Y(t_s) \mid T_s = t_s, T_c > t_s, T_u(t_s) \geq t_s\big) \\
&\quad = E\big(Y(t_s) \mid T_s = t_s, T_c > t_s, T_u \geq t_s\big) \\
&\quad = E_{X|T_s=t_s, T_c>t_s, T_u\geq t_s}\big[E\big(Y(t_s) \mid T_s = t_s, T_c > t_s, T_u \geq t_s, X\big)\big].
\end{aligned} \tag{B.43}$$

Then, by the law of iterated expectations,

$$\begin{aligned}
&E\big(Y(t_s) \mid T_s = t_s, T_c > t_s, T_u \geq t_s, X\big) \\
&\quad = \Pr(T_u = t_s|T_s = t_s, T_c > t_s, T_u \geq t_s, X)E\big(Y(t_s)|T_s = t_s, T_c > t_s, T_u = t_s, X\big) \\
&\qquad + \Pr(T_u > t_s|T_s = t_s, T_c > t_s, T_u \geq t_s, X) \\
&\qquad \times E\big(Y(t_s)|T_s = t_s, T_c > t_s, T_u > t_s, X\big),
\end{aligned} \tag{B.44}$$

as above decomposing the outcome of interest into average outcomes for individuals with $T_u = t_s$ and $T_u > t_s$.

Next,

$$E\big(Y(t_s)|T_s=t_s, T_c>t_s, T_u=t_s, X\big)=E(Y|T_s=t_s, T_c>t_s, T_u=t_s, X). \tag{B.45}$$

For sake of presentation, let us introduce some additional auxiliary notation:

$$h_{c1}(t, X, t_s)=\Pr(T_u=t \mid T_s=t_s, T_c>t, T_u\geq t, X).$$

Then, using this notation iteratively, using (B.44) and (B.45), and with Assumption 5 holding for for $t_s+2, \dots, T_u^{\max}$, we obtain

$$\begin{aligned} &E\big(Y(t_s)|T_s=t_s, T_c>t_s, T_u\geq t_s, X\big) \\ &\quad=\sum_{k=t_s}^{T_u^{\max}} h_{c1}(t, X, t_s)\left[\prod_{m=t_s}^{k-1}[1-h_{c1}(t, X, t_s)]\right] \\ &\qquad\times E(Y|T_s=t_s, T_c>k, T_u=k, X), \end{aligned} \tag{B.46}$$

so that by (B.43)–(B.46),

$$\begin{aligned} &E\big(Y(t_s)\mid T_s=t_s, T_c>t_s, T_u(t_s)\geq t_s\big) \\ &\quad=E_{X|T_s=t_s,T_c>t_s,T_u\geq t_s}\left[\sum_{k=t_s}^{T_u^{\max}} h_{c1}(t, X, t_s)\left[\prod_{m=t_s}^{k-1}[1-h_{c1}(t, X, t_s)]\right]\right. \\ &\qquad\left.\times E(Y|T_s=t_s, T_c>k, T_u=k, X)\right]. \end{aligned} \tag{B.47}$$

Finally, (B.42) and (B.47) give the result in Theorem 3.

B.5.2. *Estimation*

We now show that if Assumptions 1, 2, and 5 hold, then an unbiased estimator of $\text{ATET}(t_s)$ is

$$\begin{aligned} \widehat{\text{ATET}}(t_s)=&\frac{1}{\sum_{i\in T_{s,i}=t_s,T_{c,i}>T_{u,i},T_{u,i}\geq t_s} w_{c_1}^{t_s}(T_{u,i}, X_i)} \\ &\times \sum_{i\in T_{s,i}=t_s,T_{c,i}>T_{u,i},T_{u,i}\geq t_s} w_{c_1}^{t_s}(T_{u,i}, X_i)Y_i \\ &-\frac{1}{\sum_{i\in T_{s,i}>T_{u,i},T_{c,i}>T_{u,i},T_{u,i}\geq t_s} w_{c_0}^{t_s}(T_{u,i}, X_i)} \\ &\times \sum_{i\in T_{s,i}>T_{u,i},T_{c,i}>T_{u,i},T_{u,i}\geq t_s} w_{c_0}^{t_s}(T_{u,i}, X_i)Y_i, \end{aligned}$$

$$w_{c_1}^{t_s}(t_u, X) = \frac{1}{\prod_{m=t_{s+1}}^{t_u} [1 - e_{c1}(m, t_s, X)]}, \tag{B.48}$$

$$w_{c_0}^{t_s}(t_u, X) = \frac{p_c(t_s, X)}{[1 - p_c(t_s, X)] \prod_{m=t_s+1}^{t_u} [1 - p_c(m, X)][1 - e_{c0}(m, X)]},$$

$$p_c(t, X) = \Pr(T_s = t | T_s \geq t, T_c > t, T_u \geq t, X),$$

$$e_{c1}(t, t_s, X) = \Pr(T_c = t | T_s = t_s, T_c \geq t, T_u \geq t, X),$$

$$e_{c0}(t, X) = \Pr(T_c = t | T_s \geq t, T_c \geq t, T_u \geq t, X).$$

PROOF: First, for the first component of ATET(t_s), the estimator without the normalization is

$$\frac{1}{\rho_{t_s}^c N_{t_s}^c} \sum_{i \in T_{s,i} = t_s, T_{c,i} > T_{u,i}, T_{u,i} \geq t_s} w_{c_1}^{t_s}(T_{u,i}, X_i) Y_i,$$

where $N_{t_s}^c$ is the number of nontreated survivors at the beginning of t_s with durations censored after t_s and $\rho_t^c = \Pr(T_s = t | T_u \geq t, T_c > t, T_s \geq t)$.

Using similar reasoning as above, we have

$$\begin{aligned}
&E\left[\frac{1}{\rho_{t_s}^c N_{t_s}^c} \sum_{i \in T_{s,i} = t_s, T_{c,i} > T_{u,i}, T_{u,i} \geq t_s} w_{c_1}^{t_s}(T_{u,i}, X_i) Y_i\right] \\
&= E\left[\frac{1}{\rho_{t_s}^c N_{t_s}^c} \sum_{i \in T_{s,i} = t_s, T_{c,i} > t_s, T_{u,i} \geq t_s} w_{c_1}^{t_s}(T_{u,i}, X_i) \mathbf{I}(T_{c,i} > T_{u,i}) Y_i\right] \\
&= E\left[w_{c_1}^{t_s}(T_u, X) \mathbf{I}(T_c > T_u) Y | T_s = t_s, T_c > t_s, T_u \geq t_s\right] \\
&= E\left[\sum_{t_u = t_s}^{T_u^{\max}} w_{c_1}^{t_s}(t_u, X) \mathbf{I}(T_c > t_u, T_u = t_u) Y | T_s = t_s, T_c > t_s, T_u \geq t_s\right] \\
&= E_{X | T_s = t_s, T_c > t_s, T_u \geq t_s} \\
&\quad \times \left[E\left[\sum_{t_u = t_s}^{T_u^{\max}} w_{c_1}^{t_s}(t_u, X) \mathbf{I}(T_c > t_u, T_u = t_u) Y | T_s = t_s, T_c > t_s, T_u \geq t_s, X\right]\right].
\end{aligned} \tag{B.49}$$

Introduce the notation

$$h_c(t, X) = \Pr(T_u = t | T_u \geq t, T_c > t, T_s > t, X).$$

Then, if Assumptions 1, 2, and 5 hold, and noting that $w_{c_1}^{t_s}(t_u, X) = \frac{1}{\prod_{m=t_{s+1}}^{t_u} [1 - e_{c1}(m, t_s, X)]}$:

$$\begin{aligned}
&E\left[w_{c_1}^{t_s}(t_u, X) \mathbf{I}(T_c > t_u, T_u = t_u) Y | T_s = t_s, T_c > t_s, T_u \geq t_s, X\right] \\
&= w_{c_1}^{t_s}(t_u, X) \Pr(T_c > t_u, T_u = t_u | T_u = t_s, T_c > t_s, T_s \geq t_s, X)
\end{aligned}$$

$$
\begin{aligned}
&\times E[Y|T_s = t_s, T_c > t_u, T_u = t_u, X] \\
&= \frac{1}{\prod_{m=t_s+1}^{t_u} [1 - e_{c1}(m, t_s, X)]} h_c(t_u, X) \\
&\quad \times \left[\prod_{m=t_s}^{t_u-1} [1 - h_c(m, X)]\right]\left[\prod_{m=t_s+1}^{t_u} [1 - e_{c1}(m, t_s, X)]\right] \\
&\quad \times E[Y|T_s = t_s, T_c > t_u, T_u = t_u, X] \\
&= h_c(t_u, X)\left[\prod_{m=t_s}^{t_u-1} [1 - h_c(m, X)]\right] \\
&\quad \times E[Y|T_s = t_s, T_c > t_u, T_u = t_u, X] \\
&= h_c(t_u, X)\left[\prod_{m=t_s}^{t_u-1} [1 - h_c(m, X)]\right] \\
&\quad \times E\big[Y(t_s)|T_s = t_s, T_c > t_u, T_u(t_s) = t_u, X\big] \\
&= h_c(t_u, X)\left[\prod_{m=t_s}^{t_u-1} [1 - h_c(m, X)]\right] \\
&\quad \times E\big[Y(t_s)|T_s = t_s, T_c > t_s, T_u(t_s) = t_u, X\big],
\end{aligned}
\tag{B.50}
$$

where the second equality follows from the definition of $w_{c_0}^{t_s}(t_u, X)$, the third equality by simplifying, the fourth equality by Assumption 1, and the fifth equality by applying Assumption 5 for $t_s, \dots, t_u$.

From (B.49) and (B.50),

$$
\begin{aligned}
&E\left[\frac{1}{\rho_{t_s}^c N_{t_s}^c} \sum_{i \in T_{s,i}=t_s, T_{c,i}>T_{u,i}, T_{u,i}\geq t_s} w_{c_1}^{t_s}(T_{u,i}, X_i) Y_i\right] \\
&\quad = E_{X|T_s=t_s, T_c>t_s, T_u\geq t_s}\left[\sum_{t_u=t_s}^{T_u^{\max}} h_c(t_u, X)\left[\prod_{m=t_s}^{t_u-1} [1 - h_c(m, X)]\right]\right. \\
&\quad\quad \left. \times E\big[Y(t_s)|T_s = t_s, T_c > t_s, T_u(t_s) = t_u, X\big]\right].
\end{aligned}
\tag{B.51}
$$

Next, by similar reasoning as for (B.8) we have

$$
\begin{aligned}
&\sum_{t_u=t_s}^{T_u^{\max}} h_c(t_u, X)\left[\prod_{m=t_s}^{t_u-1} [1 - h_c(m, X)]\right] \\
&\quad \times E\big[Y(t_s)|T_s = t_s, T_c > t_s, T_u(t_s) = t_u, X\big] \\
&\quad = E\big[Y(t_s)|T_s = t_s, T_c > t_s, T_u(t_s) \geq t_s, X\big],
\end{aligned}
\tag{B.52}
$$

so that from (B.51) and (B.52),

$$E\left[\frac{1}{\rho^c_{t_s} N^c_{t_s}} \sum_{i\in T_{s,i}=t_s, T_{c,i}>T_{u,i}, T_{u,i}\geq t_s} w^{t_s}_{c_1}(T_{u,i}, X_i) Y_i\right]$$
$$= E_{X|T_s=t_s, T_c>t_s, T_u\geq t_s}\left[E\left[Y(t_s)|T_s=t_s, T_c>t_s, T_u(t_s)\geq t_s, X\right]\right]$$
$$= E\left[Y(t_s)|T_s=t_s, T_c>t_s, T_u(t_s)\geq t_s\right]. \tag{B.53}$$

Second, for the second component of ATET(t_s) the estimator without the normalization is

$$\frac{1}{\rho^c_{t_s} N^c_{t_s}} \sum_{i\in T_{s,i}>T_{u,i}, T_{c,i}>T_{u,i}, T_{u,i}\geq t_s} w^{t_s}_{c_0}(T_{u,i}, X_i) Y_i.$$

Using similar reasoning as for (B.2), we have

$$E\left[\frac{1}{\rho^c_{t_s} N_{t_s}} \sum_{i\in T_{s,i}>T_{u,i}, T_{c,i}>T_{u,i}, T_{u,i}\geq t_s} w^{t_s}_{c_0}(T_{u,i}, X_i) Y_i\right]$$
$$= E_{X|T_s\geq t_s, T_c>t_s, T_u\geq t_s}\left[\frac{1}{\rho^c_{t_s}} E\left[\sum_{t_u=t_s}^{T_u^{\max}} w^{t_s}_{c_0}(t_u, X)\right.\right.$$
$$\left.\left.\times \mathbf{I}(T_s>t_u, T_c>t_u, T_u=t_u) Y \Big| T_s\geq t_s, T_c>t_s, T_u\geq t_s, X\right]\right]. \tag{B.54}$$

Next, if Assumptions 1, 2, and 5 hold, and since

$$w^{t_s}_{c_0}(t_u, X) = \frac{p_c(t_s, X)}{[1-p_c(t_s, X)] \prod_{m=t_s+1}^{t_u} [1-p_c(m, X)][1-e_{c0}(m, X)]}$$

we have

$$E\left[w^{t_s}_{c_0}(t_u, X)\mathbf{I}(T_s>t_u, T_c>t_u, T_u=t_u)Y|T_s\geq t_s, T_c>t_s, T_u\geq t_s, X\right]$$
$$= w^{t_s}_{c_0}(t_u, X)\Pr(T_s>t_u, T_c>t_u, T_u=t_u|T_u\geq t_s, T_c>t_s, T_s\geq t_s, X)$$
$$\times E[Y|T_s>t_u, T_c>t_u, T_u=t_u, X]$$
$$= \frac{p_c(t_s, X)}{[1-p_c(t_s, X)] \prod_{m=t_s+1}^{t_u} [1-p_c(m, X)][1-e_{c0}(m, X)]}$$
$$\times h_c(t_u, X)\left[\prod_{m=t_s}^{t_u-1}[1-h_c(m, X)]\right]$$
$$\times [1-p_c(t_s, X)]\left[\prod_{m=t_s+1}^{t_u} [1-p(m, X)][1-e_{c0}(m, X)]\right]$$

$$
\begin{aligned}
&\times E[Y|T_s > t_u, T_c > t_u, T_u = t_u, X] \\
&= p_c(t_s, X) h_c(t_u, X) \left[\prod_{m=t_s}^{t_u-1} [1 - h_c(m, X)] \right] \\
&\quad \times E[Y|T_s > t_u, T_c > t_u, T_u = t_u, X] \\
&= p_c(t_s, X) h_c(t_u, X) \left[\prod_{m=t_s}^{t_u-1} [1 - h_c(m, X)] \right] \\
&\quad \times E\big[Y(\infty)|T_s > t_u, T_c > t_u, T_u(\infty) = t_u, X\big] \\
&= p_c(t_s, X) h_c(t_u, X) \left[\prod_{m=t_s}^{t_u-1} [1 - h_c(m, X)] \right] \\
&\quad \times E\big[Y(\infty)|T_s = t_s, T_c > t_u, T_u(\infty) = t_u, X\big] \\
&= p_c(t_s, X) h_c(t_u, X) \left[\prod_{m=t_s}^{t_u-1} [1 - h_c(m, X)] \right] \\
&\quad \times E\big[Y(\infty)|T_s = t_s, T_c > t_s, T_u(\infty) = t_u, X\big],
\end{aligned}
\tag{B.55}
$$

where the second equality follows from the definition of $w_{c0}^{t_s}(t_u, X)$, the third equality by simplifying, the fourth equality by Assumption 1, the fifth equality by applying Assumption 2 for $t_s, \ldots, t_u$, and the sixth equality by applying Assumption 5 for $t_s, \ldots, t_u$.

From (B.54) and (B.55),

$$
\begin{aligned}
&E\left[\frac{1}{\rho_{t_s}^c N_{t_s}} \sum_{i \in T_{s,i} > T_{u,i}, T_{c,i} > T_{u,i}, T_{u,i} \geq t} w_{c0}^{t_s}(T_{u,i}, X_i) Y_i\right] \\
&\quad = E_{X|T_s \geq t_s, T_c > t_s, T_u \geq t_s}\left[\frac{p_c(t_s, X)}{\rho_{t_s}^c} \sum_{t_u=t_s}^{T_u^{\max}} h(t_u, X) \left[\prod_{m=t_s}^{t_u-1} [1 - h(m, X)]\right]\right. \\
&\quad\quad \left. \times E\big[Y(\infty)|T_s = t_s, T_c > t_s, T_u(\infty) = t_u, X\big]\right].
\end{aligned}
\tag{B.56}
$$

Next, by similar reasoning as for (B.8) we have

$$
\begin{aligned}
&\sum_{t_u=t_s}^{T_u^{\max}} h(t_u, X) \left[\prod_{m=t_s}^{t_u-1} [1 - h(m, X)]\right] E\big[Y(\infty)|T_s = t_s, T_c > t_s, T_u(\infty) = t_u, X\big] \\
&\quad = E\big[Y(\infty)|T_s = t_s, T_c > t_s, T_u(\infty) \geq t_s, X\big].
\end{aligned}
\tag{B.57}
$$

Thus, from (B.56) and (B.57),

$$
\begin{aligned}
&E\left[\frac{1}{\rho_{t_s}^c N_{t_s}} \sum_{i \in T_{s,i} > T_{u,i}, T_{c,i} > T_{u,i}, T_{u,i} \geq t} w_{c0}^{t_s}(T_{u,i}, X_i) Y_i\right] \\
&\quad = E_{X|T_s \geq t_s, T_c > t_s, T_u \geq t_s}\left[\frac{p_c(t_s, X)}{\rho_{t_s}^c} E\big[Y(\infty)|T_s = t_s, T_c > t_s, T_u(\infty) \geq t_s, X\big]\right]
\end{aligned}
$$

$$= \frac{\Pr(T_s = t_s | T_u \geq t_s, T_c > t_s, T_s \geq t_s)}{\rho^c_{t_s}} \times E\big[Y(\infty) | T_s = t_s, T_c > t_s, T_u(\infty) \geq t_s\big]$$
$$= E\big[Y(\infty) | T_s = t_s, T_c > t_s, T_u(\infty) \geq t_s\big], \quad \text{(B.58)}$$

since $\rho^c_{t_s} = \Pr(T_s = t_s | T_u \geq t_s, T_c > t_s, T_s \geq t_s)$.

Finally, (B.53) and (B.58) imply that $E[\widehat{\text{ATET}}(t_s)] = \text{ATET}(t_s)$. *Q.E.D.*

B.6. *ATET(t_s) With Short-Run Outcomes*

B.6.1. *Identification*

The first component of $\text{ATET}(t_s, \tau)$ is identified from the observed outcomes, Y_t, of those treated at time t_s:

$$E\big(Y_{t_s+\tau}(t_s) | T_s = t_s, T_u(t_s) \geq t_s\big) = E(Y_{t_s+\tau} | T_s = t_s, T_u \geq t_s). \quad \text{(B.59)}$$

For the second component of $\text{ATET}(t_s, \tau)$, we condition on X. We assume sequential unconfoundedness.

ASSUMPTION B.1—Sequential unconfoundedness, short-run outcomes: *For all t*,

$$P_t \perp Y(\infty) \mid X, T_s \geq t, T_u \geq t.$$

Based on this and on appropriate no-anticipation assumptions,

$$\begin{aligned} &E\big(Y_{t_s+\tau}(\infty) | T_s = t_s, T_u(t_s) \geq t_s, X\big) \\ &\quad = E\big(Y_{t_s+\tau}(\infty) | T_s = t_s, T_u(\infty) \geq t_s, X\big) \\ &\quad = E\big(Y_{t_s+\tau}(\infty) | T_s > t_s, T_u(\infty) \geq t_s, X\big), \end{aligned} \quad \text{(B.60)}$$

and by the law of iterated expectations and the observational rule,

$$\begin{aligned} &E\big(Y_{t_s+\tau}(\infty) | T_s > t_s, T_u(\infty) \geq t_s, X\big) \\ &\quad = h(t_s, X) E\big(Y_{t_s+\tau}(\infty) | T_s > t_s, T_u(\infty) = t_s, X\big) \\ &\qquad + \big[1 - h(t_s, X)\big] E\big(Y_{t_s+\tau}(\infty) | T_s > t_s, T_u(\infty) > t_s, X\big), \end{aligned} \quad \text{(B.61)}$$

where $E(Y_{t_s+\tau}(\infty) | T_s > t_s, T_u(\infty) = t_s, X) = E(Y_{t_s+\tau} | T_s > t_s, T_u = t_s, X)$, and the probability $h(t_s, X)$ also is observed. By Assumption B.1 for period $t_s + 1$ and (B.61) by replacing t_s with $t_s + 1$, we have

$$\begin{aligned} &E\big(Y_{t_s+\tau}(\infty) | T_s > t_s, T_u \geq t_s, X\big) \\ &\quad = E\big(Y_{t_s+\tau}(\infty) | T_s > t_s + 1, T_u(\infty) \geq t_s + 1, X\big) \\ &\quad = h(t_s + 1, X) E\big(Y_{t_s+\tau}(\infty) | T_s > t_s + 1, T_u(\infty) = t_s + 1, X\big) \\ &\qquad + \big[1 - h(t_s + 1, X)\big] E\big(Y_{t_s+\tau}(\infty) | T_s > t_s + 1, T_u(\infty) > t_s + 1, X\big), \end{aligned} \quad \text{(B.62)}$$

where the first equality follows from Assumption B.1 and the second from (B.61). Iteratively, for $t_s+2,\dots$ gives

$$\begin{aligned}
&E\big(Y_{t_s+\tau}(\infty)|T_s>t_s, T_u(\infty)\geq t_s, X\big)\\
&\quad=\sum_{k=t_s}^{t_s+\tau} h(k,X)\left[\prod_{m=t_s}^{k-1}[1-h(m,X)]\right]E(Y|T_s>k, T_u=k, X)\\
&\qquad+\left[\prod_{m=t_s}^{t_s+\tau}[1-h(m,X)]\right]E(Y_{t_s+\tau}|T_s>t_s+\tau, T_u>t_s+\tau, X). \qquad \text{(B.63)}
\end{aligned}$$

Combining (B.60) and (B.63) and averaging over X gives the second component of equation (12).

B.6.2. *Estimation*

Under the above assumptions, an unbiased estimator of ATET(t_s) is

$$\begin{aligned}
&\widehat{\text{ATET}}(t_s)\\
&\quad=\frac{1}{\pi(t_s)N_{t_s}}\sum_{i\in T_{s,i}=t_s, T_{u,i}\geq t_s} Y_{t_s+\tau,i}\\
&\qquad-\left(\sum_{i\in T_{s,i}>T_{u,i}, t_s+\tau\geq T_{u,i}\geq t_s} w^{t_s}(T_{u,i},X_i)Y_{t_s+\tau,i}\right.\\
&\qquad+\left.\sum_{i\in T_{s,i}>t_s+\tau, T_{u,i}>t_s+\tau} w_\tau^{t_s}(T_{u,i},X_i)Y_{t_s+\tau,i}\right)\\
&\qquad\bigg/\left(\sum_{i\in T_{s,i}>T_{u,i}, t_s+\tau\geq T_{u,i}\geq t_s} w^{t_s}(T_{u,i},X_i)+\sum_{i\in T_{s,i}>t_s+\tau, T_{u,i}>t_s+\tau} w_\tau^{t_s}(T_{u,i},X_i)\right), \qquad \text{(B.64)}
\end{aligned}$$

where $w^{t_s}(t_u, X)$ is given by (11) and

$$w_\tau^{t_s}(X)=\frac{p(t_s,X)}{\prod_{m=t_s}^{t_s+\tau}1-p(m,X)}.$$

PROOF: Consider estimation of ATET(t_s,τ) and the estimator in (B.64),

$$\text{ATET}(t_s,\tau)=E\big(Y_{t_s+\tau}(t_s)-Y_{t_s+\tau}(\infty)|T_s=t_s, T_u(t_s)\geq t_s\big).$$

For the first part of ATET(t_s,τ), the estimator is

$$\frac{1}{\pi(t_s)N_{t_s}}\sum_{i\in T_{s,i}=t_s, T_{u,i}\geq t_s} Y_{t_s+\tau,i}.$$

By similar reasoning as for (B.1), we have

$$E\left[\frac{1}{\pi(t_s)N_{t_s}}\sum_{i\in T_{s,i}=t_s T_{u,i}\geq t_s} Y_{t_s+\tau,i}\right] = E\left[Y_{t_s+\tau}(t_s)|T_s=t_s, T_u(t_s)\geq t_s\right]. \tag{B.65}$$

For the second part of ATET(t_s, τ), the estimator without the normalization is

$$\frac{1}{\pi(t_s)N_{t_s}}\left[\sum_{i\in T_{s,i}>T_{u,i}, t_s+\tau\geq T_{u,i}\geq t_s} w^{t_s}(T_{u,i}, X_i)Y_{t_s+\tau,i} + \sum_{i\in T_{s,i}>t_s+\tau, T_{u,i}>t_s+\tau} w_\tau^{t_s}(T_{u,i}, X_i)Y_{t_s+\tau,i}\right].$$

Initially, using similar reasoning as for (B.2),

$$\begin{aligned}
&E\left[\frac{1}{\pi(t_s)N_{t_s}}\sum_{i\in T_{s,i}>T_{u,i}, t_s+\tau\geq T_{u,i}\geq t_s} w^{t_s}(T_{u,i}, X_i)Y_{t_s+\tau,i}\right] \\
&\quad = E_{X|T_s\geq t_s, T_u\geq t_s}\left[\frac{1}{\pi(t_s)}\right. \\
&\qquad \left.\times E\left[\sum_{t_u=t_s}^{t_s+\tau} w^{t_s}(t_u, X)\mathbf{I}(T_s>t_u, T_u=t_u)Y_{t_s+\tau}\middle|T_s\geq t_s, T_u\geq t_s, X\right]\right],
\end{aligned} \tag{B.66}$$

and using similar reasoning as for (B.3), we have

$$\begin{aligned}
&E\left[w^{t_s}(t_u, X)\mathbf{I}(T_s>t_u, T_u=t_u)Y_{t_s+\tau}|T_s\geq t_s, T_u\geq t_s, X\right] \\
&\quad = p(t_s, X)h(t_u, X)\left[\prod_{m=t_s}^{t_u-1}[1-h(m, X)]\right] \\
&\qquad \times E\left[Y_{t_s+\tau}(\infty)|T_s=t_s, T_u(\infty)=t_u, X\right].
\end{aligned} \tag{B.67}$$

Next,

$$\begin{aligned}
&E\left[\frac{1}{\pi(t_s)N_{t_s}}\sum_{i\in T_{s,i}>t_s+\tau, T_{u,i}>t_s+\tau} w_\tau^{t_s}(T_{u,i}, X_i)Y_{t_s+\tau,i}\right] \\
&\quad = E\left[\frac{1}{\pi(t_s)N_{t_s}}\sum_{i\in T_{s,i}\geq t_s, T_{u,i}\geq t_s} w_\tau^{t_s}(T_{u,i}, X_i)\mathbf{I}(T_{s,i}>t_s+\tau, T_{u,i}>t_s+\tau)Y_{t_s+\tau,i}\right] \\
&\quad = E\left[\frac{1}{\pi(t_s)N_{t_s}}\sum_{i\in T_{s,i}\geq t_s, T_{u,i}\geq t_s}\sum_{t_u>t_s+\tau}^{T_u^{\max}} w_\tau^{t_s}(t_u, X_i)\mathbf{I}(T_{s,i}>t_s+\tau, T_{u,i}=t_u)Y_{t_s+\tau,i}\right] \\
&\quad = \frac{1}{\pi(t_s)}E\left[\sum_{t_u>t_s+\tau}^{T_u^{\max}} w_\tau^{t_s}(t_u, X)\mathbf{I}(T_s>t_s+\tau, T_u=t_u)Y_{t_s+\tau}|T_s\geq t_s, T_u\geq t_s\right]
\end{aligned}$$

$$= E_{X|T_s \geq t_s, T_u \geq t_s}\left[\frac{1}{\pi(t_s)} E\left[\sum_{t_u > t_s+\tau}^{T_u^{\max}} w_\tau^{t_s}(t_u, X) \right.\right.$$

$$\left.\left. \times \mathbf{I}(T_s > t_s + \tau, T_u = t_u) Y_{t_s+\tau} | T_s \geq t_s, T_u \geq t_s, X\right]\right], \tag{B.68}$$

Then Assumption B.1 and no-anticipation and the fact that $w_\tau^{t_s}(X) = \frac{p(t_s, X)}{\prod_{m=t_s}^{t_s+\tau}[1-p(m,X)]}$ jointly imply that

$$\begin{aligned}
&E\big[w_\tau^{t_s}(X)\mathbf{I}(T_s > t_s + \tau, T_u = t_u) Y_{t_s+\tau} | T_s \geq t_s, T_u \geq t_s, X\big] \\
&\quad = w_\tau^{t_s}(X) \Pr(T_s > t_s + \tau, T_u = t_u | T_u \geq t_s, T_s \geq t_s, X) \\
&\qquad \times E[Y_{t_s+\tau} | T_s > t_s + \tau, T_u = t_u, X] \\
&\quad = \frac{p(t_s, X)}{\prod_{m=t_s}^{t_s+\tau}[1 - p(m, X)]} h(t_u, X)\left[\prod_{m=t_s}^{t_u-1}[1 - h(m, X)]\right] \\
&\qquad \times \left[\prod_{m=t_s}^{t_s+\tau}[1 - p(m, X)]\right] E[Y_{t_s+\tau} | T_s > t_s + \tau, T_u = t_u, X] \\
&\quad = p(t_s, X) h(t_u, X)\left[\prod_{m=t_s}^{t_u-1}[1 - h(m, X)]\right] \\
&\qquad \times E[Y_{t_s+\tau} | T_s > t_s + \tau, T_u = t_u, X] \\
&\quad = p(t_s, X) h(t_u, X)\left[\prod_{m=t_s}^{t_u-1}[1 - h(m, X)]\right] \\
&\qquad \times E\big[Y_{t_s+\tau}(\infty) | T_s > t_s + \tau, T_u(\infty) = t_u, X\big] \\
&\quad = p(t_s, X) h(t_u, X)\left[\prod_{m=t_s}^{t_u-1}[1 - h(m, X)]\right] \\
&\qquad \times E\big[Y_{t_s+\tau}(\infty) | T_s = t_s, T_u(\infty) = t_u, X\big].
\end{aligned} \tag{B.69}$$

From (B.66)–(B.69), we have

$$E\left[\frac{1}{\pi(t_s) N_{t_s}}\left[\sum_{\substack{i \in T_{s,i} > T_{u,i}, \\ t_s+\tau \geq T_{u,i} \geq t_s}} w^{t_s}(T_{u,i}, X_i) Y_{t_s+\tau,i} + \sum_{\substack{i \in T_{s,i} > t_s+\tau, \\ T_{u,i} > t_s+\tau}} w_\tau^{t_s}(T_{u,i}, X_i) Y_{t_s+\tau,i}\right]\right]$$

$$= E_{X|T_s \geq t_s, T_u \geq t_s}\left[\frac{p(t_s, X)}{\pi(t_s)} \sum_{t_u=t_s}^{T_u^{\max}} h(t_u, X)\left[\prod_{m=t_s}^{t_u-1}[1 - h(m, X)]\right]\right.$$

$$\left. \times E\big[Y_{t_s+\tau}(\infty) | T_s = t_s, T_u(\infty) = t_u, X\big]\right]. \tag{B.70}$$

Then, using similar reasoning as for (B.8),

$$\sum_{t_u=t_s}^{T_u^{\max}} h(t_u, X)\left[\prod_{m=t_s}^{t_u-1}[1-h(m,X)]\right]$$
$$\times E\big[Y_{t_s+\tau}(\infty)|T_s=t_s, T_u(\infty)=t_u, X\big]$$
$$= E\big[Y_{t_s+\tau}(\infty)|T_s=t_s, T_u(\infty)\geq t, X\big], \tag{B.71}$$

and thus from (B.70) and (B.71),

$$E\left[\frac{1}{\pi(t_s)N_{t_s}}\left[\sum_{\substack{i\in T_{s,i}>T_{u,i},\\ t_s+\tau\geq T_{u,i}\geq t_s}} w^{t_s}(T_{u,i}, X_i)Y_{t_s+\tau,i} + \sum_{\substack{i\in T_{s,i}>t_s+\tau,\\ T_{u,i}>t_s+\tau}} w_\tau^{t_s}(T_{u,i}, X_i)Y_{t_s+\tau,i}\right]\right]$$
$$= \frac{1}{\pi(t_s)}E_{X|T_s\geq t_s, T_u\geq t_s}\big[p(t_s,X)E\big[Y_{t_s+\tau}(\infty)|T_s=t_s, T_u(\infty)\geq t_s, X\big]\big]$$
$$= \frac{1}{\pi(t_s)}E_{X|T_s\geq t_s, T_u\geq t_s}[\Pr(T_s=t_s|T_s\geq t_s, T_u\geq t_s, X)$$
$$\times E\big[Y_{t_s+\tau}(\infty)|T_s=t_s, T_u(\infty)\geq t_s, X\big]$$
$$= \frac{1}{\pi(t_s)}\Pr(T_s=t_s|T_s\geq t_s, T_u\geq t_s)$$
$$\times E\big[Y_{t_s+\tau}(\infty)|T_s=t_s, T_u(\infty)\geq t_s\big]$$
$$= E\big[Y_{t_s+\tau}(\infty)|T_s=t_s, T_u(\infty)\geq t_s\big]. \tag{B.72}$$

Finally, (B.65) and (B.72) imply that $E[\widehat{\text{ATET}}(t_s)] = \text{ATET}(t_s)$. *Q.E.D.*

APPENDIX C: MONTE CARLO SIMULATION

C.1. *Simulation Design*

This simulation study examines properties of the estimator introduced in the paper. We use the following notation for the conditional exit probability out of the initial state: $\theta_{T_u}(t) = \Pr(T_u = t_u|T_u \geq t_u)$, and the conditional treatment probability: $\theta_{T_s}(t) = \Pr(T_s = t_s|T_u \geq t_s, T_s \geq t_s)$. We consider the following discrete time DGP:

$$\begin{aligned}
&\theta_{T_u}(t) = f(-2.5 + X + v_u),\\
&\theta_{T_s}(t) = f(\alpha_s + \beta_s X + v_s), \quad t \leq 12,\\
&\theta_{T_s}(t) = 0, \quad t > 12,\\
&\quad Y = 100 + \beta_Y X + \delta I(T_u \geq T_s) + \beta_u v_u + v_y\\
&\qquad \text{with } X, v_u, v_s \sim \text{unif}(-1, 1), v_y \sim N(0, 5),
\end{aligned} \tag{C.73}$$

with X, v_u, v_s, v_y all independently distributed of each other, and $f(h) = [1 + \exp(-h)]^{-1}$.

This model has several properties worth noticing. First, the treatment can start at any point during the first 12 time periods, corresponding to a treatment in place during the

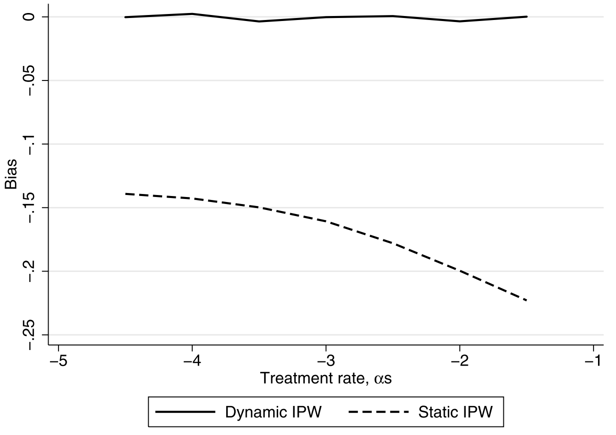

FIGURE C.1.—Simulated bias for the dynamic IPW estimator and a static IPW estimator. Model A: baseline treatment rate. Note: α_s is the conditional treatment probability parameter. The data generating processes are described in the text. Bias for aggregated effect of treatment over the first 12 months. Dynamic IPW is the estimator introduced in this paper. Static IPW is a standard static IPW estimator with normalized weights. Results are based on 2000 replications.

first year (if the time period is a month). Second, both durations, T_u and T_s, and the outcome, Y, depend on observed and unobserved characteristics. However, since the unobserved effect in the treatment equation is uncorrelated with the other unobserved effects, the unconfoundedness assumption holds. Third, the unobserved effect in the duration time equation also appears in the long-run outcome equation. This is consistent with the idea that some unobserved characteristics affect both time in the initial state and the long-run outcome. In the training for unemployed example, this may be unobserved motivation and/or unobserved ability.

In the baseline setting, the correlation between the unobserved characteristics in the exit and long-outcome equations β_u is 1, the baseline treatment probability parameter α_s is -3.0, the impact of the covariate on treatment β_s is 1, the treatment effect on the long-run outcome δ is 0, and impact of the covariate on the long-run outcomes β_Y is set to 1. These parameters are then varied in four different ways. Model A varies the baseline treatment parameter (α_s between -4.5 and -1.5). Model B varies the impact of the covariate on treatment (β_s between 0 and 2). With $\alpha_s = -4$, the conditional treatment probability in each period is 0.021 while with $\alpha_s = -2$ this is 0.13. If β_s equals 0.5, the conditional treatment probability varies between 0.029 and 0.076; and if $\beta_s = 1.5$, this probability varies between 0.011 and 0.18. Model C varies the correlation between the unobserved characteristics in the exit and long-outcome equations (β_u between 0 and 2). Finally, Model D allows the treatment effect on the long-run outcome, δ, to vary between 1 and 10.

We focus on the aggregated effect ATET. All propensity scores are estimated with a correct logistic model specification. We initially study the bias of each estimator. The sample size is set to 10,000 and the number of replications is 2000. Common support is imposed through the above mentioned variant of the three-step approach from Huber, Lechner, and Wunsch (2013), with the upper limit on the weight given to a certain observation set to 1%. After this, we study the size and variance of the dynamic estimator, using bootstrapped standard errors (500 replications). In that case, we allow the treatment to start at any point during the first 4 periods.

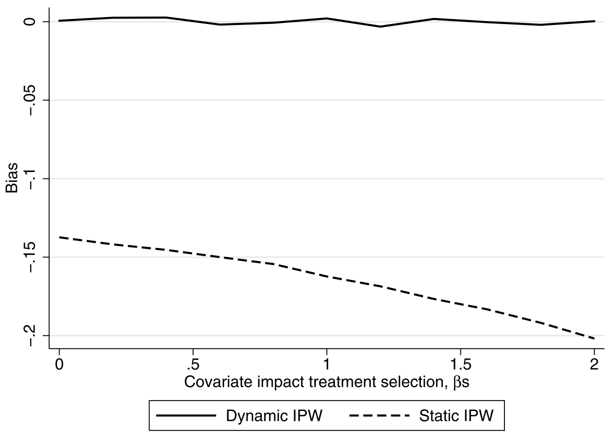

FIGURE C.2.—Simulated bias for the dynamic IPW estimator and a static IPW estimator. Model B: impact of the covariate on the conditional treatment probability. Note: β_s is the impact of the covariate on treatment. The data generating processes are described in the text. Bias for aggregated effect of treatment over the first 12 months. Dynamic IPW is the estimator introduced in this paper. Static IPW is a standard static IPW estimator with normalized weights. Results are based on 2000 replications.

C.2. *Simulation Results*

We compare the dynamic IPW estimator and a static IPW estimator. Figure C.1 reports how the bias of the two estimators are related to the baseline treatment rate. As expected, the static IPW estimator is biased, and the bias is increasing in the treatment parameter (higher α_s). This is because a higher conditional treatment probability implies more extensive dynamic treatment assignment. The bias of the dynamic IPW estimator, on the

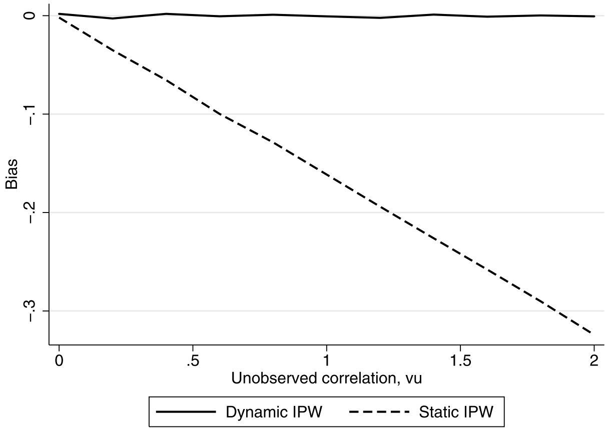

FIGURE C.3.—Simulated bias for the dynamic IPW estimator and a static IPW estimator. Model C: correlation between the unobserved characteristics in the exit and long-outcome equations. Note: β_u determines the correlation between the unobserved characteristics in the exit and long-outcome equations. The data generating processes are described in the text. Bias for aggregated effect of treatment over the first 12 months. Dynamic IPW is the estimator introduced in this paper. Static IPW is a standard static IPW estimator with normalized weights. Results are based on 2000 replications.

TABLE C.I

SIMULATED BIAS, SIZE, AND VARIANCE OF THE DYNAMIC IPW ESTIMATOR.

	1000 observations			4000 observations		
	bias [1]	se [2]	size [3]	bias [4]	se [5]	size [6]
Baseline model	0.004	0.198	0.056	−0.005	0.100	0.058
Unobserved correlation, $\beta_u = 2$	0.000	0.213	0.049	0.001	0.105	0.047
Treatment rate, $\alpha_s = -2$	−0.005	0.166	0.055	−0.001	0.084	0.057
Treatment selection, $\beta_s = 2$	0.004	0.209	0.058	0.000	0.106	0.049
Treatment effect, $\delta_s = 5$	0.002	0.190	0.040	−0.001	0.096	0.038

Note: IPW estimates with bootstrapped standard errors (500 replications). The data generating processes are described in the text. Size is for 5% level tests. The results are based on 2000 replications.

other hand, is virtually zero for all treatment probabilities and roughly 100 times smaller than for the static IPW estimator.

Figure C.2 also shows that the bias of the static IPW estimator increases with the variance of the treatment probability across units (larger β_s), while the dynamic approach is unbiased for all values of β_s. From Figure C.3, it can also be seen that the bias of the static approach is increasing in the correlation between the unobserved characteristics in the exit and long-outcome equations, β_u. Again, the bias of our dynamic approach is virtually zero.

Finally, Table C.I presents the bias, variance, and size of our dynamic IPW estimator. The simulation results are for sample sizes of 1000 and 4000. We vary the parameters of the DGP in a similar way for Models A–D, but we only report simulation results for the baseline case and one additional case for each model. First, as expected, based on the results in Figures C.1–C.4, the bias is small in all cases. Size is for a test with nominal size of 5%, so that the IPW estimator roughly has correct size (columns 3 and 6). The tables also show that standard error decreases by roughly 50% when the sample size is

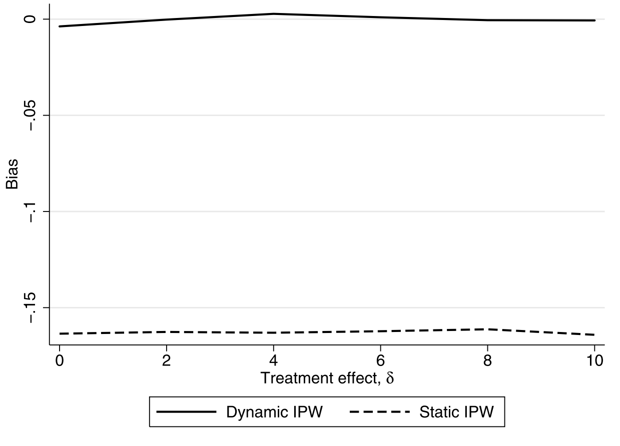

FIGURE C.4.—Simulated bias for the dynamic IPW estimator and a static IPW estimator. Model D: treatment effect on the long-run outcome. Note: δ is the treatment effect. The data generating processes are described in the text. Bias for aggregated effect of treatment over the first 12 months. Dynamic IPW is the estimator introduced in this paper. Static IPW is a standard static IPW estimator with normalized weights. Results are based on 2000 replications.

increased by a factor of four from 1000 to 4000, suggesting that the estimator is $\sqrt{N}$-convergent.

APPENDIX D: SAMPLE STATISTICS

TABLE D.I

SELECTED SAMPLE STATISTICS.

	Non-treated (averages)	Treated (averages)
# observations	735,547	57,033
Unemployment duration in months	8.7	25.6
Outcomes		
Earnings in unemployment entry year	69,354	55,817
Earnings +1 years	79,885	45,887
Earnings +2 years	101,123	78,256
Earnings +3 years	119,707	102,533
Earnings +4 years	133,284	117,453
Earnings +5 years	142,351	127,252
Earnings +6 years	147,865	132,906
Earnings +7 years	152,406	136,648
Earnings +8 years	158,094	141,483
Earnings +9 years	165,504	147,997
Background characteristics		
Male (%)	47.4	47.5
Age at the start of the spell		
25–34	57.8	50.3
35–44	26.4	30.8
45–54	15.8	18.9
Married (%)	36.5	41.2
Education (%)		
Less than high school	21.8	20.4
High school education	51.8	57.8
University education	26.4	21.8
Child in ages 0–3 (%)	25.6	30.5
Child in ages 4–6 (%)	17.7	19.4
Child in ages 7–15 (%)	24.7	27.8
UI eligible (%)	78.0	80.7
Only search in local area (%)	16.6	19.3
Preprogram earnings and unemployment		
Days unemployed year −1	58.0	64.7
Days unemployed year −2	81.1	96.7
Earnings year −1	72,095	72,121
Earnings year −2	74,202	71,798
Earnings year −3	77,121	74,655
Year of inflow (%; residual category is 1995)		
1996	24.9	21.7
1997	24.6	20.3
1998	26.0	24.9

(*Continues*)

TABLE D.I

Continued.

	Non-treated (averages)	Treated (averages)
Area of residence (%; residual category is "other")		
Stockholm MSA	21.1	17.2
Gothenburg MSA	16.4	13.9
Skane MSA	13.3	13.3
North	14.2	15.1
South	11.8	13.4

Note: Covariates recorded at the start of the unemployment spell. Earnings are in SEK.

REFERENCES

HUBER, MARTIN, MICHAEL LECHNER, AND CONNY WUNSCH (2013): "The Performance of Estimators Based on the Propensity Score," *Journal of Econometrics*, 175, 1–21. [26]

Co-editor Guido Imbens handled this manuscript.

Manuscript received 26 July, 2019; final version accepted 11 August, 2021; available online 21 October, 2021.