Doubly Robust Difference-in-Differences Estimators: Supplemental Appendix

Pedro H. C. Sant'Anna* Jun B. Zhao†

Vanderbilt University Vanderbilt University

May 5, 2020

This supplemental appendix collects all the proofs of the main results of the paper.

*Department of Economics, Vanderbilt University. E-mail: pedro.h.santanna@vanderbilt.edu.

†Department of Economics, Vanderbilt University. E-mail: jun.zhao@vanderbilt.edu.

Proofs of Main Results

Proof of Theorem 1: We prove the results for panel and repeated cross-section data separately.

Case 1: Panel data are available and propensity score model is correctly specified

In this case, we have that $\pi(X) = p(X)$ $a.s.$. In order to show that $\tau^{dr,p} = \tau \equiv ATT$, first note that, by the law of iterated expectations,

$$\mathbb{E}\left[\frac{\pi(X)(1-D)}{1-\pi(X)}\right] = \mathbb{E}\left[\mathbb{E}\left[\frac{p(X)(1-D)}{1-p(X)}\middle| X\right]\right] = \mathbb{E}\left[p(X)\right] = \mathbb{E}\left[D\right],$$

which yields

$$w_0^p(D,X;\pi) = \frac{\pi(X)(1-D)}{1-\pi(X)} \Big/ \mathbb{E}\left[\frac{\pi(X)(1-D)}{1-\pi(X)}\right] = \frac{1}{\mathbb{E}\left[D\right]}\frac{p(X)(1-D)}{1-p(X)}.$$

Therefore, we have that

$$\begin{aligned}
\tau^{dr,p} &= \frac{1}{\mathbb{E}\left[D\right]}\mathbb{E}\left[\left(D - \frac{p(X)(1-D)}{(1-p(X))}\right)(\Delta Y - \mu_{0,\Delta}^p(X))\right] \\
&= \frac{1}{\mathbb{E}\left[D\right]}\mathbb{E}\left[\left(D - \frac{p(X)(1-D)}{(1-p(X))}\right)\Delta Y\right] - \frac{1}{\mathbb{E}\left[D\right]}\mathbb{E}\left[\left(D - \frac{p(X)(1-D)}{(1-p(X))}\right)\mu_{0,\Delta}^p(X)\right] \\
&= \tau - \mathbb{E}\left[(p(X) - p(X)) \cdot \mu_{0,\Delta}^p(X)\right] \\
&= \tau,
\end{aligned}$$

where the second to last equality follows from Lemma 3.1 and equation (10) in Abadie (2005) and the law of iterated expectations.

Case 2: Panel data are available and outcome regression models are correctly specified.

In this case, we have that $\mu_{0,\Delta}^p(X) = m_{0,\Delta}^p(X)$ $a.s.$, i.e. the outcome regression models are correctly specified. Let $B = \mathbb{E}\left[\frac{\pi(X)(1-D)}{(1-\pi(X))}\right]^{-1}$. Therefore,

$$\begin{aligned}
\tau^{dr,p} &= \mathbb{E}\left[\frac{D}{\mathbb{E}\left[D\right]}\left(\Delta Y - m_{0,\Delta}^p(X)\right)\right] - \mathbb{E}\left[w_0^p(D,X;\pi)\left(\Delta Y - m_{0,\Delta}^p(X)\right)\right] \\
&= \mathbb{E}\left[\Delta Y - m_{0,\Delta}^p(X)|D=1\right] - B\cdot\mathbb{E}\left[\frac{1-\pi(X)}{\pi(X)}\left(\Delta Y - m_{0,\Delta}^p(X)\right)\middle| D=0\right](1-p(X)) \\
&= \mathbb{E}\left[m_{1,\Delta}^p(X) - m_{0,\Delta}^p(X)|D=1\right] - B\cdot\mathbb{E}\left[\frac{1-\pi(X)}{\pi(X)}\left(m_{0,\Delta}^p(X) - m_{0,\Delta}^p(X)\right)\middle| D=0\right](1-p(X)) \\
&= \tau
\end{aligned}$$

where the third equality follows from the law of iterated expectations, and the last one from the conditional parallel trends assumption, i.e., Assumption 2.

Case 3: Repeated cross-section data are available and propensity score model is correctly specified.

In this case, we have that $\pi(X) = p(X)$ $a.s.$. First of all, recall that $\lambda \equiv \mathbb{P}(T=1) = \mathbb{E}[T]$, and notice that, by the law of iterated expectations and stationarity of (D,X),

$$\begin{aligned}
\mathbb{E}\left[\frac{\pi(X)(1-D)T}{1-\pi(X)}\right] &= \mathbb{E}\left[\left.\frac{p(X)(1-D)}{1-p(X)}\right| T=1\right]\lambda \\
&= \mathbb{E}[p(X)]\lambda \\
&= \mathbb{E}[D]\lambda.
\end{aligned}$$

Similarly, we have that

$$\mathbb{E}\left[\frac{\pi(X)(1-D)(1-T)}{1-\pi(X)}\right] = \mathbb{E}[D](1-\lambda),$$

together with $\mathbb{E}[DT] = \mathbb{E}[D]\lambda$ and $\mathbb{E}[D(1-T)] = \mathbb{E}[D](1-\lambda)$. Thus, when the propensity score is correctly specified, we can write the four weights in $\tau^{dr,rc}$ as

$$w_{1,1}^{rc}(D,T) = \frac{D}{\mathbb{E}[D]}\frac{T}{\lambda}, \qquad w_{1,0}^{rc}(D,T) = \frac{D}{\mathbb{E}[D]}\frac{(1-T)}{(1-\lambda)},$$

$$w_{0,1}^{rc}(D,T,X;p) = \frac{1}{\mathbb{E}[D]}\frac{p(X)(1-D)}{(1-p(X))}\frac{T}{\lambda}, \qquad w_{0,0}^{rc}(D,T,X;p) = \frac{1}{\mathbb{E}[D]}\frac{p(X)(1-D)}{1-p(X)}\frac{(1-T)}{(1-\lambda)}.$$

Next, after some straightforward manipulation, we have that

$$\begin{aligned}
\tau_1^{dr,rc} &= \frac{1}{\mathbb{E}[D]}\mathbb{E}\left[\frac{D-p(X)}{1-p(X)}\frac{T-\lambda}{1-\lambda}(Y-\mu_{0,Y}^{rc}(T,X))\right] \\
&= \tau - \frac{1}{\mathbb{E}[D]}\mathbb{E}\left[\frac{D-p(X)}{1-p(X)}\frac{T-\lambda}{1-\lambda}\mu_{0,Y}^{rc}(T,X)\right] \\
&= \tau - \frac{1}{\mathbb{E}[D]}\mathbb{E}\left[\mathbb{E}\left[\left.\frac{D-p(X)}{1-p(X)}\right| T,X\right]\frac{T-\lambda}{1-\lambda}\mu_{0,Y}^{rc}(T,X)\right] \\
&= \tau - \frac{1}{\mathbb{E}[D]}\mathbb{E}\left[\frac{p(X)-p(X)}{1-p(X)}\frac{T-\lambda}{1-\lambda}\mu_{0,Y}^{rc}(T,X)\right] \\
&= \tau
\end{aligned}$$

where the second equality follows from Lemma 3.2 and equation (12) in Abadie (2005), and the third and forth equality from the law of iterated expectations, and stationarity condition in Assumption 1(b). Indeed, under the stationarity condition, we have that $\tau_2^{dr,rc} = \tau_1^{dr,rc}$, since, for any generic integrable and measurable function g, $\mathbb{E}[g(X)|D=1] = \mathbb{E}[g(X)|D=1,T=t]$, $t=0,1$.

Case 4: Repeated cross-section data are available and outcome regression models are correctly specified.

We now consider the case where $\mu_{0,\Delta}^{rc}(X) = m_{0,\Delta}^{rc}(X)$ $a.s.$, where $m_{0,\Delta}^{rc}(x) = m_{0,1}^{rc}(x) - m_{0,0}^{rc}(x)$, where $m_{0,t}^{rc}(x) = \mathbb{E}[Y|D=0,T=t,X=x]$, $t=0,1$. By the total law of expectations, we can write $\tau_1^{dr,rc}$ as

$$\begin{aligned}
\tau_1^{dr,rc} = \big(&\mathbb{E}[w_{1,1}^{rc}(D,1)(Y-\mu_{0,1}^{rc}(X))|T=1]\lambda - \mathbb{E}[w_{1,0}^{rc}(D,0)(Y-\mu_{0,0}^{rc}(X))|T=0](1-\lambda)\big) \\
&- \mathbb{E}[w_0^{rc}(D,T,X;\pi)(Y-\mu_{0,Y}^{rc}(T,X))].
\end{aligned}$$

Next, note that, by the stationarity condition, for $t = 0, 1$,

$$\begin{aligned}\mathbb{E}\left[w_{1,t}^{rc}(D,1)\left(Y-\mu_{0,t}^{rc}(X)\right)\middle|T=t\right]\mathbb{P}(T=t) &= \mathbb{E}\left[\frac{D}{\mathbb{E}[D]\,\mathbb{P}(T=t)}\left(Y-\mu_{0,t}^{rc}(X)\right)\middle|T=t\right]\mathbb{P}(T=t)\\ &= \mathbb{E}\left[\left(Y-\mu_{0,t}^{rc}(X)\right)\middle|T=t,D=1\right].\\ &= \mathbb{E}\left[Y|T=t,D=1\right]-\mathbb{E}\left[\mu_{0,t}^{rc}(X)\middle|D=1\right].\end{aligned}$$

Thus, given that $\mu_{0,\Delta}^{rc}(X) = m_{0,\Delta}^{rc}(X)$ $a.s.$ and Assumption 2 holds,

$$\begin{aligned}&\mathbb{E}[w_{1,1}^{rc}(D,1)(Y-\mu_{0,1}^{rc}(X))|T=1]\lambda-\mathbb{E}[w_{1,0}^{rc}(D,0)(Y-\mu_{0,0}^{rc}(X))|T=0](1-\lambda)\\ =\ &\mathbb{E}\left[Y|T=1,D=1\right]-\mathbb{E}\left[Y|T=0,D=1\right]-\mathbb{E}\left[\mu_{0,\Delta}^{rc}(X)\middle|D=1\right]\\ =\ &\tau.\end{aligned}$$

Next, note that under the stationarity Assumption,

$$w_0^{rc}(D,T,X;\pi)=\mathbb{E}\left[\frac{\pi(X)(1-D)}{1-\pi(X)}\right]^{-1}\frac{\pi(X)(1-D)}{(1-\pi(X))}\left(\frac{T}{\lambda}-\frac{1-T}{1-\lambda}\right),$$

implying that to complete the proof of Case 4 for $\tau_1^{dr,rc}$, it suffices to show that

$$\mathbb{E}\left[\frac{\pi(X)(1-D)}{(1-\pi(X))}\left(\frac{T}{\lambda}-\frac{1-T}{1-\lambda}\right)\left(Y-\mu_{0,Y}^{rc}(T,X)\right)\right]=0, \tag{S.1}$$

when $\mu_{0,\Delta}^{rc}(X) = m_{0,\Delta}^{rc}(X)$ $a.s.$. By the law of iterated expectations

$$\begin{aligned}\mathbb{E}\left[\frac{\pi(X)(1-D)}{(1-\pi(X))}\left(\frac{T}{\lambda}-\frac{1-T}{1-\lambda}\right)Y\right] &= \mathbb{E}\left[\frac{\pi(X)(1-D)}{(1-\pi(X))}\left(\mathbb{E}[Y|T=1,D=0,X]-\mathbb{E}[Y|T=0,D=0,X]\right)\right]\\ &= \mathbb{E}\left[\frac{\pi(X)(1-D)}{(1-\pi(X))}m_{0,\Delta}^{rc}(X)\right],\end{aligned}$$

and, given that $\mu_{0,Y}^{rc}(T,X) \equiv T\cdot\mu_{0,1}^{rc}(X)+(1-T)\,\mu_{0,0}^{rc}(X)$,

$$\begin{aligned}\mathbb{E}\left[\frac{\pi(X)(1-D)}{(1-\pi(X))}\left(\frac{T}{\lambda}-\frac{1-T}{1-\lambda}\right)\mu_{0,Y}^{rc}(T,X)\right] &= \mathbb{E}\left[\frac{\pi(X)(1-D)}{(1-\pi(X))}\left(\frac{T\mu_{0,1}^{rc}(X)}{\lambda}-\frac{(1-T)\,\mu_{0,0}^{rc}(X)}{1-\lambda}\right)\right]\\ &= \mathbb{E}\left[\frac{\pi(X)(1-D)}{(1-\pi(X))}\mu_{0,\Delta}^{rc}(X)\right]\\ &= \mathbb{E}\left[\frac{\pi(X)(1-D)}{(1-\pi(X))}m_{0,\Delta}^{rc}(X)\right],\end{aligned}$$

where the second equality follows from stationarity and the third one from $\mu_{0,\Delta}^{rc}(X) = m_{0,\Delta}^{rc}(X)$ $a.s.$. It is then clear that (S.1) holds, and we have that $\tau_1^{dr,rc} = \tau$ when $\mu_{0,\Delta}^{rc}(X) = m_{0,\Delta}^{rc}(X)$ $a.s.$. Finally, by the stationarity condition, we have that $\tau_2^{dr,rc} = \tau_1^{dr,rc}$ as well.■

Proof of Proposition 1: Our proof follows closely the structure of semiparametric efficiency bound derivation of Newey (1990), which is also used by Hahn (1998) and Chen et al. (2008), among others. As before, we derive the results for panel and repeated cross-section data separately.

Panel data are available:

The density of $(Y_1(1), Y_1(0), Y_0(0), D, X')'$ (with respect to a sigma-finite measure) on $\mathscr{Z} \in \mathbb{R}^3 \times \{0,1\} \times \mathbb{R}^k$, is given by

$$\bar{f}(y_1(1), y_1(0), y_0(0), d, x) =$$
$$\bar{f}(y_1(1), y_1(0), y_0(0)|D=1,x)^d p(x)^d \bar{f}(y_1(1), y_1(0), y_0(0)|D=0,x)^{1-d} (1-p(x))^{1-d} f(x),$$

where $\bar{f}(y_1(1), y_1(0), y_0(0)|D=d,x)$ and $f(x)$ denote the conditional density of $Y_1(1), Y_1(0), Y_0(0)$ given $D=d, d=0,1$ and $X=x$, as well as the marginal density of X respectively. Given that $Y_1 = DY_1(1) + (1-D)Y_1(0)$, and $Y_0 = Y_0(0)$, the density of observed data $(Y_1, Y_0, D, X')'$ is thus given by

$$f(y_1, y_0, d, x) = f_1(y_1, y_0|D=1,x)^d p(x)^d \; f_0(y_1, y_0|D=0,x)^{1-d} (1-p(x))^{1-d} \; f(x),$$

where,

$$f_1(\cdot,\cdot|D=d,x) = \int \bar{f}(\cdot, y_0(0), \cdot|D=1,x) dy_0(0), \text{ and } f_0(\cdot,\cdot|D=0,x) = \int \bar{f}(y_1(1), \cdot, \cdot|D=0,x) dy_1(1).$$

Note that $Y_1(0)$ is not observed whenever $D=1$. However, under Assumption 2, we have that

$$\mathbb{E}[Y_1(0)|D=1,X] = \mathbb{E}[Y_0|D=1,X] + \mathbb{E}[Y_1 - Y_0|D=0,X] \; a.s.,$$

implying that we can recover $\mathbb{E}[Y_1(0)|D=1,X]$ from the density of $(Y_1, Y_0, D, X')'$.

The first step of our proof is to characterize the tangent space. To this end, consider a regular parametric submodel

$$f_{1,\theta}(y_1, y_0|D=1,x)^d p_\theta(x)^d \times f_{0,\theta}(y_1, y_0|D=0,x)^{1-d} (1-p_\theta(x))^{1-d} \times f_\theta(x),$$

which equals $f(y_1, y_0, d, x)$ when $\theta = \theta_0$. The corresponding score is given by

$$s_\theta(y_1, y_0, d, x) \equiv d \cdot s_{1,\theta}(y_1, y_0|D=1,x) + (1-d) \cdot s_{0,\theta}(y_1, y_0|D=0,x) + \frac{d - p_\theta(x)}{p_\theta(x)(1-p_\theta(x))} \dot{p}_\theta(x) + t_\theta(x),$$

where, for $d=0,1$,

$$s_{d,\theta}(y_1, y_0|D=d,x) = \frac{d}{d\theta} log\; f_{d,\theta}(y_1, y_0|D=d,x), \quad \dot{p}_\theta(x) = \frac{d}{d\theta} p_\theta(x), \quad \text{and} \quad t_\theta(x) = \frac{d}{d\theta} log f_\theta(x).$$

From the above calculations, the tangent space of this model is given by

$$\mathscr{T} = \{d \cdot s_1(y_1, y_0|D=1,x) + (1-d) \cdot s_0(y_1, y_0|D=0,x) + a(x) \cdot (d - p(x)) + t(x)\},$$

where $\int\int s_d(y_1, y_0|D=d,x) f_d(y_1, y_0|D=d,x) dy_1 dy_0 = 0 \; \forall x, \; d=0,1, \; \int t(x) f(x) dx = 0$, and $a(x)$ is any square-integrable measurable function of x.

Under Assumptions 1-2, when panel data are available the ATT is given by

$$\tau = \mathbb{E}\left[\mathbb{E}\left[Y_1 - Y_0 | D=1, X\right] - \mathbb{E}\left[Y_1 - Y_0 | D=0, X\right] | D=1\right].$$

The next step of the proof is to show that τ is pathwise differentiable. For the parametric submodel under consideration, we have that

$$\tau(\theta) = \frac{\int\int\int (y_1 - y_0) p_\theta(x) f_{1,\theta}(y_1, y_0|D=1,x) f_\theta(x) dy_1 dy_0 dx}{\int p_\theta(x) f_\theta(x) dx} - \frac{\int\int\int (y_1 - y_0) p_\theta(x) f_{0,\theta}(y_1, y_0|D=0,x) f_\theta(x) dy_1 dy_0 dx}{\int p_\theta(x) f_\theta(x) dx}.$$

Thus,

$$\begin{aligned}\frac{\partial \tau(\theta_0)}{\partial \theta} =& \frac{\int\int\int (y_1 - y_0) p(x) s_1(y_1, y_0|D=1,x) f_1(y_1, y_0|D=1,x) f(x) dy_1 dy_0 dx}{p} \\ &- \frac{\int\int\int (y_1 - y_0) p(x) s_0(y_1, y_0|D=0,x) f_0(y_1, y_0|D=0,x) f(x) dy_1 dy_0 dx}{p} \\ &+ \frac{\int (m_{1,\Delta}(x) - m_{0,\Delta}(x) - \tau) \dot{p}(x) f(x) dx}{p} + \frac{\int (m_{1,\Delta}(x) - m_{0,\Delta}(x) - \tau) p(x) t(x) f(x) dx}{p},\end{aligned}$$

where $p \equiv \mathbb{E}[D]$, and for $d = 0, 1$, $m_{d,\Delta}(x) = \mathbb{E}[Y_1 - Y_0 | D = d, X = x]$. Let

$$\begin{aligned}F_\tau(Y_1, Y_0, D, X) =& \frac{D}{p}(Y_1 - Y_0 - m_{1,\Delta}(X)) + \frac{1}{p}\frac{(1-D)p(X)}{(1-p(X))}(Y_1 - Y_0 - m_{0,\Delta}(X)) \\ &+ \frac{m_{1,\Delta}(x) - m_{0,\Delta}(x) - \tau}{p}(D - p(X)) + \frac{m_{1,\Delta}(x) - m_{0,\Delta}(x) - \tau}{p} \cdot p(X).\end{aligned}$$

For the parametric submodel whose score is $s_\theta(y_1, y_0, d, x)$, we have

$$\frac{\partial \tau(\theta_0)}{\partial \theta} = \mathbb{E}[F_\tau(Y_1, Y_0, D, X) \cdot s_{\theta_0}(Y_1, Y_0, D, X)],$$

from which we conclude that τ is differentiable parameter.

Given that $F_\tau \in \mathscr{T}$, all the conditions of Theorem 3.1 in Newey (1990) hold, so that F_τ is the efficient influence function for the ATT. Note that by simple manipulation and an application of the law of iterated expectations, one can write

$$\begin{aligned}F_\tau(Y_1, Y_0, D, X) &= \frac{D}{p}(m_{1,\Delta}(x) - m_{0,\Delta}(x) - \tau) + \frac{D}{p}(\Delta Y - m_{1,\Delta}(X)) + \frac{1}{p}\frac{(1-D)p(X)}{(1-p(X))}(\Delta Y - m_0(X)) \\ &= \eta^{e,p}(D, X).\end{aligned}$$

The semiparametric efficiency bound now follows from taking the expected value of F_τ squared.

Repeated cross-section data are available:

The density of the observed (pooled) data $(Y, D, X', T)'$ is given by

$$\begin{aligned}f(y,d,x,t) =& f(y,d,x|T=1)^t f(y,d,x|T=0)^{1-t} \lambda^t (1-\lambda)^{1-t} \\ =& f(y|D=1,x,T=1)^{dt} f(y|D=1,x,T=0)^{d(1-t)} \\ &\times f(y|D=0,x,T=1)^{(1-d)t} f(y|D=0,x,T=0)^{(1-d)(1-t)} \\ &\times p(x)^{dt}(1-p(x))^{(1-d)t} p(x)^{d(1-t)}(1-p(x))^{(1-d)(1-t)} \\ &\times f(x|T=1)^t f(x|T=0)^{(1-t)}\end{aligned}$$

$$\times \lambda^t(1-\lambda)^{1-t}.$$

Given that (D,X) is strictly stationary by Assumption 1(b), $f(y,d,x,t)$ reduces to

$$\begin{aligned}
f(y,d,x,t) =& f(y|D=1,x,T=1)^{dt} f(y|D=1,x,T=0)^{d(1-t)} \\
&\times f(y|D=0,x,T=1)^{(1-d)t} f(y|D=0,x,T=0)^{(1-d)(1-t)} \\
&\times p(x)^d(1-p(x))^{(1-d)} \times f(x) \times \lambda^t(1-\lambda)^{1-t}.
\end{aligned}$$

As before, the first step of our proof is to characterize the tangent space. Consider a regular parametric submodel

$$\begin{aligned}
&f_\theta(y|D=1,x,T=1)^{dt} f_\theta(y|D=1,x,T=0)^{d(1-t)} \\
&\times f_\theta(y|D=0,x,T=1)^{(1-d)t} f_\theta(y|D=0,x,T=0)^{(1-d)(1-t)} \\
&\times p_\theta(x)^d(1-p_\theta(x))^{(1-d)} \times f_\theta(x) \times \lambda_\theta^t(1-\lambda_\theta)^{1-t},
\end{aligned}$$

which equals $f(y,d,x,t)$ when $\theta=\theta_0$. The corresponding score is given by

$$\begin{aligned}
s_\theta(y,d,x,t) \equiv& dt \cdot s_\theta(y|D=1,x,T=1) + d(1-t)\cdot s_\theta(y|D=1,x,T=0) \\
&+ (1-d)t \cdot s_\theta(y|D=0,x,T=1) + (1-d)(1-t)\cdot s_\theta(y|D=0,x,T=0) \\
&+ \frac{d-p_\theta(x)}{p_\theta(x)(1-p_\theta(x))}\dot{p}_\theta(x) + t_\theta(x) + \frac{t-\lambda_\theta}{\lambda_\theta(1-\lambda_\theta)}\dot{\lambda}_\theta,
\end{aligned}$$

where, for $d,t \in \{0,1\}^2$

$$s_\theta(y|D=d,x,T=t) = \frac{d}{d\theta} f_\theta(y|D=d,x,T=t), \quad \dot{p}_\theta(x) = \frac{d}{d\theta}p_\theta(x), \quad t_\theta(x) = \frac{d}{d\theta} log f_\theta(x), \quad \text{and} \quad \dot{\lambda}_\theta = \frac{d}{d\theta}\lambda_\theta.$$

The tangent space of this model is then given by

$$\begin{aligned}
\mathscr{T}^{rc} =& \{dt \cdot s(y|D=1,x,T=1) + d(1-t)\cdot s(y|D=1,x,T=0) \\
&+ (1-d)ts(y|D=0,x,T=1) + (1-d)(1-t)\cdot s(y|D=0,x,T=0) \\
&+ a(x)\cdot(d-p(x)) + t(x) + a(t-\lambda)\},
\end{aligned}$$

where $\int s(y|D=d,x,T=t)f(y|D=d,x,T=t)dy = 0 \ \forall x,\ d=0,1,\ t=0,1,\ \int t(x)f(x)dx = 0,\ a(x)$ is any square-integrable measurable function of x, and a is a finite constant.

Under Assumptions 1-2, the ATT is given by

$$\tau = \mathbb{E}\left[(\mathbb{E}[Y|D=1,T=1,X] - \mathbb{E}[Y|D=1,T=0,X]) - (\mathbb{E}[Y|D=0,T=1,X] - \mathbb{E}[Y|D=0,T=0,X])|D=1\right].$$

Our next step is to show that τ is pathwise differentiable. Towards this end, for the parametric submodel under consideration, we find that

$$\tau(\theta) = \left(\frac{\int\int y p_\theta(x) f_\theta(y|D=1,x,T=1) f_\theta(x)dydx}{\int p_\theta(x)f_\theta(x)dx} - \frac{\int\int y p_\theta(x) f_\theta(y|D=1,x,T=0) f_\theta(x)dydx}{\int p_\theta(x)f_\theta(x)dx}\right)$$

$$-\left(\frac{\int\int y p_\theta(x) f_\theta(y|D=0,x,T=1) f_\theta(x)dydx}{\int p_\theta(x) f_\theta(x)dx} - \frac{\int\int y p_\theta(x) f_\theta(y|D=0,x,T=0) f_\theta(x)dydx}{\int p_\theta(x) f_\theta(x)dx}\right).$$

Therefore,

$$\begin{aligned}\frac{\partial \tau(\theta_0)}{\partial\theta} =& \sum_{t=0}^{1}(-1)^{t+1}\frac{\int\int y p(x) s(y|D=1,x,T=t) f(y|D=1,x,T=t) f(x)dydx}{p}\\ &-\sum_{j=0}^{1}(-1)^{j+1}\frac{\int\int y p(x) s(y|D=0,x,T=j) f(y|D=0,x,T=j) f(x)dydx}{p}\\ &+\frac{\int(m^{rc}_{1,\Delta}(x)-m^{rc}_{0,\Delta}(x)-\tau)\dot{p}(x)f(x)dx}{p}\\ &+\frac{\int(m^{rc}_{1,\Delta}(x)-m^{rc}_{0,\Delta}(x)-\tau)p(x)t(x)f(x)dx}{p}.\end{aligned}$$

where, for $d,t=0,1$, $m^{rc}_{d,\Delta}(x)\equiv m^{rc}_{d,1}(x)-m^{rc}_{d,0}(x)$, $m^{rc}_{d,t}(x)\equiv\mathbb{E}[Y|D=d,T=t,X]$, and $p\equiv\mathbb{E}[D]$. Let

$$\begin{aligned}F^{rc}_\tau(Y,D,T,X) =& \frac{1}{p\lambda}DT(Y-m^{rc}_{1,1}(X)) - \frac{1}{p(1-\lambda)}D(1-T)(Y-m^{rc}_{1,0}(X))\\ &-\frac{1}{p\lambda}\frac{(1-D)p(X)T}{1-p(X)}(Y-m^{rc}_{0,1}(X)) + \frac{1}{p(1-\lambda)}\frac{(1-D)p(X)(1-T)}{(1-p(X)}(Y-m^{rc}_{0,0}(X))\\ &+\frac{m^{rc}_{1,\Delta}(x)-m^{rc}_{0,\Delta}(x)-\tau}{p}(D-p(X)) + \frac{m^{rc}_{1,\Delta}(x)-m^{rc}_{0,\Delta}(x)-\tau}{p}p(X).\end{aligned}$$

For the parametric submodel whose score is given by $s_\theta(y,d,x,t)$, we have

$$\frac{\partial \tau(\theta_0)}{\partial\theta} = \mathbb{E}[F^{rc}_\tau(Y,D,T,X)\cdot s_{\theta_0}(Y,D,X,T)],$$

from which we conclude that τ is differentiable parameter.

Given that $F^{rc}_\tau \in \mathscr{T}^{rc}$, all the conditions of Theorem 3.1 in Newey (1990) hold, so that F^{rc}_τ is the efficient influence function for the ATT when repeated cross-section are available, under our maintained Assumptions. To show that $F^{rc}_\tau(Y,D,T,X)=\eta^{e,rc}(Y,D,T,X)$ $a.s.$, note that by the stationarity condition,

$$p\lambda = \mathbb{E}[DT] = \mathbb{E}\left[\frac{p(X)(1-D)T}{1-p(X)}\right], \text{ and } p(1-\lambda) = \mathbb{E}[D(1-T)] = \mathbb{E}\left[\frac{p(X)(1-D)(1-T)}{1-p(X)}\right],$$

and the result now follows from straightforward manipulation. The semiparametric efficiency bound is obtained from taking the expected value of F^{rc}_τ squared, concluding the proof.■

Proof of Corollary 1: In what follows, we write $m_{d,t}(\cdot)=m^p_{d,t}(\cdot)=m^{rc}_{d,t}(\cdot)$, and denote

$$\begin{aligned}\mathbb{E}\left[\eta^{e,rc}(Y,D,T,X)^2\right] &= \frac{V^{rc}_1+V^{rc}_2+V^{rc}_3}{\mathbb{E}[D]^2},\\ \mathbb{E}\left[\eta^{e,p}(Y_1,Y_0,D,X)^2\right] &= \frac{V^p_1+V^p_2+V^p_3}{\mathbb{E}[D]^2},\end{aligned}$$

where

$$V^{rc}_1 = V^p_1 = \mathbb{E}\left[D(m_{1,\Delta}(X)-m_{0,\Delta}(X)-\tau)^2\right],$$

$$V_2^{rc} = \mathbb{E}\left[D\frac{T}{\lambda^2}(Y - m_{1,1}(X))^2 + D\frac{(1-T)}{(1-\lambda)^2}(Y - m_{1,0}(X))^2\right],$$

$$V_2^{p} = \mathbb{E}\left[D(\Delta Y - m_{1,\Delta}(X))^2\right],$$

$$V_3^{rc} = \mathbb{E}\left[\frac{(1-D)\,p(X)^2}{(1-p(X))^2}\frac{T}{\lambda^2}(Y - m_{0,1}(X))^2 + \frac{(1-D)\,p(X)^2}{(1-p(X))^2}\frac{(1-T)}{(1-\lambda)^2}(Y - m_{0,0}(X))^2\right],$$

$$V_3^{p} = \mathbb{E}\left[\frac{(1-D)\,p(X)^2}{(1-p(X))^2}(\Delta Y - m_{0,\Delta}(X))^2\right].$$

Now, V_2^p and V_2^{rc} can be written as

$$V_2^{p} = \mathbb{E}\left[D(Y_1 - m_{1,1}(X))^2\right] + \mathbb{E}\left[D(Y_0 - m_{1,0}(X))^2\right] - 2\mathbb{E}\left[D(Y_0 - m_{1,0}(X))(Y_1 - m_{1,1}(X))\right],$$

$$V_2^{rc} = \frac{1}{\lambda}\mathbb{E}\left[D(Y_1 - m_{1,1}(X))^2\right] + \frac{1}{1-\lambda}\mathbb{E}\left[D(Y_0 - m_{1,0}(X))^2\right]$$

Thus, by simple manipulation, we have that

$$\begin{aligned} V_2^{rc} - V_2^{p} &= \mathbb{E}\left[D(Y_1 - m_{1,1}(X))^2\right]\frac{1-\lambda}{\lambda} \\ &\quad + \mathbb{E}\left[D(Y_0 - m_{1,0}(X))^2\right]\frac{\lambda}{1-\lambda} \\ &\quad + 2\mathbb{E}\left[D(Y_0 - m_{1,0}(X))(Y_1 - m_{1,1}(X))\right] \\ &= \mathbb{E}\left[D\left(\sqrt{\frac{1-\lambda}{\lambda}}(Y_1 - m_{1,1}(X)) + \sqrt{\frac{\lambda}{1-\lambda}}(Y_0 - m_{1,0}(X))\right)^2\right]. \end{aligned}$$

Analogously, we have that

$$\begin{aligned} V_3^{p} &= \mathbb{E}\left[\frac{(1-D)\,p(X)^2}{(1-p(X))^2}(Y_1 - m_{0,1}(X))^2\right] + \mathbb{E}\left[\frac{(1-D)\,p(X)^2}{(1-p(X))^2}(Y_0 - m_{0,0}(X))^2\right] \\ &\quad - 2\mathbb{E}\left[\frac{(1-D)\,p(X)^2}{(1-p(X))^2}(Y_0 - m_{0,1}(X))(Y_1 - m_{0,0}(X))\right], \\ V_3^{rc} &= \frac{1}{\lambda}\mathbb{E}\left[\frac{(1-D)\,p(X)^2}{(1-p(X))^2}(Y - m_{0,1}(X))^2\right] + \frac{1}{1-\lambda}\mathbb{E}\left[\frac{(1-D)\,p(X)^2}{(1-p(X))^2}(Y_0 - m_{0,0}(X))^2\right], \end{aligned}$$

and

$$V_3^{rc} - V_3^{p} = \mathbb{E}\left[\frac{(1-D)\,p(X)^2}{(1-p(X))^2}\left(\sqrt{\frac{1-\lambda}{\lambda}}(Y_1 - m_{0,1}(X)) + \sqrt{\frac{\lambda}{1-\lambda}}(Y_0 - m_{0,0}(X))\right)^2\right].$$

Hence,

$$\begin{aligned} &\mathbb{E}\left[\eta^{e,rc}(Y,D,T,X)^2\right] - \mathbb{E}\left[\eta^{e,p}(Y_1,Y_0,D,X)^2\right] = \\ &\qquad \frac{1}{\mathbb{E}[D]^2}\mathbb{E}\left[D\left(\sqrt{\frac{1-\lambda}{\lambda}}(Y_1 - m_{1,1}(X)) + \sqrt{\frac{\lambda}{1-\lambda}}(Y_0 - m_{1,0}(X))\right)^2\right] \\ &\quad + \frac{1}{\mathbb{E}[D]^2}\mathbb{E}\left[\frac{(1-D)\,p(X)^2}{(1-p(X))^2}\left(\sqrt{\frac{1-\lambda}{\lambda}}(Y_1 - m_{0,1}(X)) + \sqrt{\frac{\lambda}{1-\lambda}}(Y_0 - m_{0,0}(X))\right)^2\right] \end{aligned}$$

$$\geq 0,$$

and the proof is complete.■

Before proving Theorems 2 and 3, we prove Theorems A.1 and A.2 stated in Appendix A, as they cover generic first-step estimators.

Proof of Theorem A.1: Under Assumptions 1-A.2, we prove the large sample properties of our generic DR estimator, $\widehat{\tau}^{dr,p}$, when panel data are available.

First of all, recall that the estimator takes the following form:

$$\widehat{\tau}^{dr,p} = \mathbb{E}_n\left[\left(\widehat{w}_1^p(D) - \widehat{w}_0^p(D,X;\widehat{\gamma})\right)\left(\Delta Y - \mu_{0,\Delta}^p(X;\widehat{\beta}_{0,0}^p,\widehat{\beta}_{0,1}^p)\right)\right],$$

where

$$\widehat{w}_1^p(D) = \frac{D}{\mathbb{E}_n[D]},$$
$$\widehat{w}_0^p(D,X;\gamma) = \frac{\pi(X;\gamma)(1-D)}{1-\pi(X;\gamma)} \Bigg/ \mathbb{E}_n\left[\frac{\pi(X;\gamma)(1-D)}{1-\pi(X;\gamma)}\right];$$

and where $\widehat{\gamma}$, $\widehat{\beta}_{0,0}^p$, and $\widehat{\beta}_{0,1}^p$ are estimators for pseudo-true γ^*, $\beta_{0,0}^{*,p}$, and $\beta_{0,1}^{*,p}$, and for generic β_0^p and β_1^p, $\mu_{0,\Delta}^p(\cdot;\beta_0^p,\beta_1^p) = \mu_{0,1}^p(\cdot;\beta_1^p) - \mu_{0,0}^p(\cdot;\beta_0^p)$.

By weak law of large numbers and continuous mapping theorem, as $n \to \infty$,

$$\mathbb{E}_n[D] \xrightarrow{p} \mathbb{E}[D],$$
$$\mathbb{E}_n\left[\frac{\pi(X;\hat{\gamma})(1-D)}{1-\pi(X;\hat{\gamma})}\right] \xrightarrow{p} \mathbb{E}\left[\frac{\pi(X;\gamma^*)(1-D)}{1-\pi(X;\gamma^*)}\right],$$

and

$$\hat{\tau}^{dr,p} \xrightarrow{p} \mathbb{E}_n\left[\left(w_1^p(D) - w_0^p(D,X;\gamma^*)\right)\left(\Delta Y - \mu_{0,\Delta}^p(X;\beta_{0,0}^{*,p},\beta_{0,1}^{*,p})\right)\right],$$

where

$$w_1^p(D) = \frac{D}{\mathbb{E}[D]},$$
$$w_0^p(D,X;\gamma^*) = \frac{\pi(X;\gamma^*)(1-D)}{1-\pi(X;\gamma^*)} \Bigg/ \mathbb{E}\left[\frac{\pi(X;\gamma^*)(1-D)}{1-\pi(X;\gamma^*)}\right].$$

Therefore, if either $\pi(X;\gamma^*) = p(X)$ a.s. or $\mu_{0,\Delta}^p(X;\beta_{0,0}^{*,p},\beta_{0,1}^{*,p}) = m_{0,1}^p(X) - m_{0,0}^p(X)$ *a.s.*, from Theorem 1, we have that

$$\mathbb{E}\left[\left(w_1^p(D) - w_0^p(D,X;\gamma^*)\right)\left(\Delta Y - \mu_{0,\Delta}^p(X;\beta_{0,0}^{*,p},\beta_{0,1}^{*,p})\right)\right] \equiv \tau^{dr,p} = ATT,$$

which completes the consistency proof.

Next, we derive the asymptotically linear representation of $\widehat{\tau}^{dr,p}$, which leads to the asymptotic distribution result. Towards this end, first notice that

$$\begin{aligned}
&\widehat{\tau}^{dr,p}-\tau^{dr,p}\\
&=\left(\mathbb{E}_n\left[\widehat{w}_1^p(D)\Delta Y\right]-\mathbb{E}\left[w_1^p(D)\Delta Y\right]\right)\\
&\quad-\left(\mathbb{E}_n\left[\widehat{w}_0^p(D,X;\widehat{\gamma})\Delta Y\right]-\mathbb{E}\left[w_0^p(D,X;\gamma^*)\Delta Y\right]\right)\\
&\quad-\left(\mathbb{E}_n\left[\widehat{w}_1^p(D)\mu_{0,\Delta}^p\left(X;\widehat{\beta}_{0,0}^p,\widehat{\beta}_{0,1}^p\right)\right]-\mathbb{E}\left[w_1^p(D)\mu_{0,\Delta}^p\left(X;\beta_{0,0}^{*,p},\beta_{0,1}^{*,p}\right)\right]\right)\\
&\quad+\left(\mathbb{E}_n[\widehat{w}_0^p(D,X;\widehat{\gamma})\mu_{0,\Delta}^p\left(X;\widehat{\beta}_{0,0}^p,\widehat{\beta}_{0,1}^p\right)]-\mathbb{E}[w_0^p(D,X;\gamma^*)\mu_{0,\Delta}^p\left(X;\beta_{0,0}^{*,p},\beta_{0,1}^{*,p}\right)]\right)\\
&\equiv\left(\widehat{ATT}_1-ATT_1\right)-\left(\widehat{ATT}_2-ATT_2\right)-\left(\widehat{ATT}_3-ATT_3\right)+\left(\widehat{ATT}_4-ATT_4\right).
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\sqrt{n}(\hat{\tau}^{dr,p}-\tau^{dr,p})\\
=&\sqrt{n}\left(\widehat{ATT}_1-ATT_1\right)-\sqrt{n}\left(\widehat{ATT}_2-ATT_2\right)-\sqrt{n}\left(\widehat{ATT}_3-ATT_3\right)+\sqrt{n}\left(\widehat{ATT}_4-ATT_4\right),
\end{aligned}\tag{S.2}$$

and we next obtain the asymptotically linear representation for each component in the above decomposition.

We first analyze $\sqrt{n}\left(\widehat{ATT}_1-ATT_1\right)$. Note that by Lindeberg–Lévy central limit theorem, it follows that

$$\begin{aligned}
\mathbb{E}_n[D]-\mathbb{E}[D]&=O_p(n^{-1/2}),\\
\mathbb{E}_n\left[\frac{D}{E[D]^2}\Delta Y\right]-\mathbb{E}\left[\frac{D}{E[D]^2}\Delta Y\right]&=O_p(n^{-1/2})=o_p(1),
\end{aligned}\tag{S.3}$$

which yields

$$(\mathbb{E}_n[D]-\mathbb{E}[D])^2=O_p(n^{-1})=o_p(n^{-1/2}).$$

Next, by a second-order Taylor expansion of $\widehat{ATT}_1$ around $\mathbb{E}[D]$, we have that

$$\begin{aligned}
\widehat{ATT}_1&=\mathbb{E}_n\left[\frac{D}{\mathbb{E}_n[D]}\Delta Y\right]\\
&=\mathbb{E}_n\left[\frac{D}{\mathbb{E}[D]}\Delta Y\right]-(\mathbb{E}_n[D]-\mathbb{E}[D])\mathbb{E}_n\left[\frac{D}{\mathbb{E}[D]^2}\Delta Y\right]+O_p((\mathbb{E}_n[D]-\mathbb{E}[D])^2)\\
&=\mathbb{E}_n\left[\frac{D}{\mathbb{E}[D]}\Delta Y\right]-(\mathbb{E}_n[D]-\mathbb{E}[D])\mathbb{E}_n\left[\frac{D}{\mathbb{E}[D]^2}\Delta Y\right]+o_p(n^{-1/2}),
\end{aligned}$$

and therefore

$$\begin{aligned}
&\sqrt{n}\left(\widehat{ATT}_1-ATT_1\right)\\
&=\frac{1}{\sqrt{n}}\sum_{i=1}^n\left(\frac{D_i}{\mathbb{E}[D]}\Delta Y_i-(D_i-\mathbb{E}[D])\mathbb{E}_n\left[\frac{D}{\mathbb{E}[D]^2}\Delta Y\right]-\mathbb{E}\left[\frac{D}{\mathbb{E}[D]}\Delta Y\right]\right)+o_p(1)\\
&=\frac{1}{\sqrt{n}}\sum_{i=1}^n\left(\frac{D_i}{\mathbb{E}[D]}\Delta Y_i-(D_i-\mathbb{E}[D])\mathbb{E}\left[\frac{D}{\mathbb{E}[D]^2}\Delta Y\right]-\mathbb{E}\left[\frac{D}{\mathbb{E}[D]}\Delta Y\right]\right)+o_p(1)
\end{aligned}$$

$$= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\frac{D_i}{\mathbb{E}[D]} \Delta Y_i - \frac{D_i}{\mathbb{E}[D]} \mathbb{E}\left[\frac{D}{\mathbb{E}[D]} \Delta Y \right] \right) + o_p(1)$$

$$= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(w_1^p(D_i) \Delta Y_i - w_1^p(D_i) \mathbb{E}[w_1^p(D) \Delta Y] \right) + o_p(1)$$

$$= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} w_1^p(D_i) \left(\Delta Y_i - \mathbb{E}[w_1^p(D) \Delta Y] \right) + o_p(1), \tag{S.4}$$

where the second equality follows from (S.3).

We next analyze $\sqrt{n}\left(\widehat{ATT}_2 - ATT_2\right)$. Following the same arguments as above, we have that

$$\sqrt{n}\left(\widehat{ATT}_2 - ATT_2\right)$$

$$= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\widehat{w}_0^p(D_i, X_i; \widehat{\gamma}) \Delta Y_i - \mathbb{E}[w_0^p(D, X; \gamma^*) \Delta Y] \right)$$

$$= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\widetilde{w}_0^p(D_i, X_i; \widehat{\gamma}) \Delta Y_i - \mathbb{E}[w_0^p(D, X; \gamma^*) \Delta Y] \right)$$

$$- \sqrt{n} \left(\mathbb{E}_n \left[\frac{\pi(X; \widehat{\gamma})(1-D)}{1-\pi(X; \widehat{\gamma})} \right] - \mathbb{E}\left[\frac{\pi(X; \gamma^*)(1-D)}{1-\pi(X; \gamma^*)} \right] \right) \cdot \frac{\mathbb{E}\left[\frac{\pi(X; \gamma^*)(1-D)}{1-\pi(X; \gamma^*)} \Delta Y \right]}{\mathbb{E}\left[\frac{\pi(X; \gamma^*)(1-D)}{1-\pi(X; \gamma^*)} \right]^2} + o_p(1)$$

$$= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\widetilde{w}_0^p(D_i, X_i; \widehat{\gamma}) \Delta Y_i - \mathbb{E}[w_0^p(D, X; \gamma^*) \Delta Y] \right)$$

$$- \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\left(\widetilde{w}_0^p(D_i, X_i; \widehat{\gamma}) - 1 \right) \mathbb{E}[w_0^p(D, X; \gamma^*) \Delta Y] \right) + o_p(1)$$

$$= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \widetilde{w}_0^p(D_i, X_i; \widehat{\gamma}) \left(\Delta Y_i - \mathbb{E}[w_0^p(D, X; \gamma^*) \Delta Y] \right) + o_p(1), \tag{S.5}$$

where

$$\widetilde{w}_0^p(D, X; \widehat{\gamma}) = \frac{\pi(X; \widehat{\gamma})(1-D)}{1-\pi(X; \widehat{\gamma})} \Bigg/ \mathbb{E}\left[\frac{\pi(X; \gamma^*)(1-D)}{1-\pi(X; \gamma^*)} \right].$$

Then, by doing a second-order Taylor expansion of the above expression around pseudo-true γ^*, we have that

$$\sqrt{n}\left(\widehat{ATT}_2 - ATT_2\right)$$

$$= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} w_0^p(D_i, X_i; \gamma^*) \left(\Delta Y_i - \mathbb{E}[w_0^p(D, X; \gamma^*) \Delta Y] \right)$$

$$+ (\widehat{\gamma} - \gamma^*)' \cdot \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \dot{w}_0^p(D_i, X_i; \gamma^*) \left(\Delta Y_i - \mathbb{E}[w_0^p(D, X; \gamma^*) \Delta Y] \right) + o_p(1)$$

$$= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} w_0^p(D_i, X_i; \gamma^*) \left(\Delta Y_i - \mathbb{E}[w_0^p(D, X; \gamma^*) \Delta Y] \right)$$

$$+ \sqrt{n} (\widehat{\gamma} - \gamma^*)' \cdot \frac{1}{n} \sum_{i=1}^{n} \dot{w}_0^p(D_i, X_i; \gamma^*) \left(\Delta Y_i - \mathbb{E}[w_0^p(D, X; \gamma^*) \Delta Y] \right) + o_p(1)$$

$$= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} w_0^p(D_i, X_i; \gamma^*) \left(\Delta Y_i - \mathbb{E}[w_0^p(D, X; \gamma^*) \Delta Y] \right)$$

$$+\frac{1}{\sqrt{n}}\sum_{i=1}^{n} l_{ps}(D_i,X_i;\gamma^*)'\cdot\mathbb{E}\left[\dot{w}_0^p(D,X;\gamma^*)\left(\Delta Y-\mathbb{E}[w_0^p(D,X;\gamma^*)\Delta Y]\right)\right]+o_p(1)$$
$$=\frac{1}{\sqrt{n}}\sum_{i=1}^{n} w_0^p(D_i,X_i;\gamma^*)\left(\Delta Y_i-\mathbb{E}[w_0^p(D,X;\gamma^*)\Delta Y]\right)$$
$$+\frac{1}{\sqrt{n}}\sum_{i=1}^{n} l_{ps}(D_i,X_i;\gamma^*)'\cdot\mathbb{E}\left[\alpha_{ps}^p(D,X;\gamma^*)\left(\Delta Y-\mathbb{E}[w_0^p(D,X;\gamma^*)\Delta Y]\right)\dot{\pi}(X;\gamma^*)\right]+o_p(1)$$
$$=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(w_0^p(D_i,X_i;\gamma^*)\left(\Delta Y_i-\mathbb{E}[w_0^p(D,X;\gamma^*)\Delta Y]\right)\right.$$
$$\left.+l_{ps}(D_i,X_i;\gamma^*)'\cdot\mathbb{E}\left[\alpha_{ps}^p(D,X;\gamma^*)\left(\Delta Y-\mathbb{E}[w_0^p(D,X;\gamma^*)\Delta Y]\right)\dot{\pi}(X;\gamma^*)\right]\right)+o_p(1), \quad \text{(S.6)}$$

where the third equality follows from Assumption A.1, and the fourth equality holds because the derivative of $w_0^p(D,X;\gamma)$ with respect to γ, denoted by $\dot{w}_0^p(D,X;\gamma)$, can be written as

$$\dot{w}_0^p(D,X;\gamma)=\alpha_{ps}^r(D,X;\gamma)\dot{\pi}(X;\gamma),$$

where

$$\alpha_{ps}^p(D,X;\gamma)=\frac{(1-D)}{(1-\pi(X;\gamma))^2}\Big/\mathbb{E}\left[\frac{\pi(X;\gamma)(1-D)}{1-\pi(X;\gamma)}\right].$$

For $\sqrt{n}\left(\widehat{ATT}_3-ATT_3\right)$, we have that, by using similar arguments as in (S.4), it follows that

$$\sqrt{n}\left(\widehat{ATT}_3-ATT_3\right)=\frac{1}{n}\sum_{i=1}^{n} w_1^p(D_i)\left(\mu_{0,\Delta}^p(X_i;\widehat{\beta}_{0,0}^p,\widehat{\beta}_{0,1}^p)-\mathbb{E}\left[w_1^p(D)\mu_{0,\Delta}^p(X;\beta_{0,0}^{*,p},\beta_{0,1}^{*,p})\right]\right)+o_p(1).$$

Denote $\widehat{\beta}\equiv(\widehat{\beta}_{0,1}^p,\widehat{\beta}_{0,0}^p)$. Then, from a second-order Taylor expansion of the above expression around $\beta^*\equiv(\beta_{0,1}^{*,p},\beta_{0,0}^{*,p})$, we have that

$$\sqrt{n}\left(\widehat{ATT}_3-ATT_3\right)=\frac{1}{\sqrt{n}}\sum_{i=1}^{n} w_1^p(D_i)\left(\mu_{0,\Delta}^p(X_i;\beta^*)-\mathbb{E}[w_1^p(D)\mu_{0,\Delta}^p(X;\beta^*)]\right)$$
$$+\sqrt{n}\left(\widehat{\beta}-\beta^*\right)'\cdot\mathbb{E}_n[w_1^p(D)\dot{\mu}_{0,\Delta}^p(X;\beta^*)]+o_p(1)$$
$$=\frac{1}{\sqrt{n}}\sum_{i=1}^{n} w_1^p(D_i)\left(\mu_{0,\Delta}^p(X_i;\beta^*)-\mathbb{E}[w_1^p(D)\mu_{0,\Delta}^p(X;\beta^*)]\right)$$
$$+\frac{1}{\sqrt{n}}\sum_{i=1}^{n} l_{reg}(W_i;\beta^*)'\cdot\mathbb{E}[w_1^p(D)\dot{\mu}_{0,\Delta}^p(X;\beta^*)]+o_p(1)$$
$$=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(w_1^p(D_i)\left(\mu_{0,\Delta}^p(X_i;\beta^*)-\mathbb{E}[w_1^p(D)\mu_{0,\Delta}^p(X;\beta^*)]\right)\right.$$
$$\left.+l_{reg}(W_i;\beta^*)'\cdot\mathbb{E}[w_1^p(D)\dot{\mu}_{0,\Delta}^p(X;\beta^*)]\right)\}+o_p(1), \quad \text{(S.7)}$$

where the second equality holds because of Assumption A.1.

Finally, for $\sqrt{n}\left(\widehat{ATT}_4-ATT_4\right)$, using the same arguments as in (S.5)-(S.7), we have that

$$\sqrt{n}\left(\widehat{ATT}_4-ATT_4\right)$$

$$
\begin{aligned}
=&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\widetilde{w}_0^p(D_i,X_i;\widehat{\gamma})\left(\mu_{0,\Delta}^p(X_i;\widehat{\beta})-\mathbb{E}[w_0^p(D,X;\gamma^*)\mu_{0,\Delta}^p(X;\beta^*)]\right)+o_p(1)\\
=&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}w_0^p(D_i,X_i;\gamma^*)\left(\mu_{0,\Delta}^p(X_i;\beta^*)-\mathbb{E}[w_0^p(D,X;\gamma^*)\mu_{0,\Delta}^p(X;\beta^*)]\right)\\
&+\sqrt{n}(\widehat{\gamma}-\gamma^*)'\cdot\mathbb{E}\left[\alpha_{ps}^p(D,X;\gamma^*)\left(\mu_{\Delta}^p(X;\beta^*)-\mathbb{E}[w_0^p(D,X;\gamma^*)\mu_{0,\Delta}^p(X;\beta^*)]\right)\dot{\pi}(X;\gamma^*)\right]\\
&+\sqrt{n}\left(\widehat{\beta}-\beta^*\right)'\cdot\mathbb{E}[w_0^p(D,X;\gamma^*)\dot{\mu}_{0,\Delta}^p(X;\beta^*)]+o_p(1)\\
=&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(w_0^p(D_i,X_i;\gamma^*)\left(\mu_{0,\Delta}^p(X_i;\beta^*)-\mathbb{E}[w_0^p(D,X;\gamma^*)\mu_{0,\Delta}^p(X;\beta^*)]\right)\right.\\
&+l_{ps}(D_i,X_i;\gamma^*)'\cdot\mathbb{E}\left[\alpha_{ps}^p(D,X;\gamma^*)\left(\mu_{0,\Delta}^p(X;\beta^*)-\mathbb{E}[w_0^p(D,X;\gamma^*)\mu_{0,\Delta}^p(X;\beta^*)]\right)\dot{\pi}(X;\gamma^*)\right]\\
&\left.+l_{reg}(W_i;\beta^*)'\cdot\mathbb{E}[w_0^p(D,X;\gamma^*)\dot{\mu}_{0,\Delta}^p(X;\beta^*)]\right)+o_p(1).
\end{aligned}\tag{S.8}
$$

Taking (S.4)-(S.8) together, we conclude that

$$
\begin{aligned}
&\sqrt{n}(\widehat{\tau}^{dr,p}-\tau^{dr,p})\\
=&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(w_1^p(D_i)\left(\Delta Y_i-\mathbb{E}[w_1^p(D)\Delta Y]\right)\right.\\
&-w_0^p(D_i,X_i;\gamma^*)\left(\Delta Y_i-\mathbb{E}[w_0^p(D,X;\gamma^*)\Delta Y]\right)\\
&-l_{ps}(D_i,X_i;\gamma^*)'\cdot\mathbb{E}\left[\alpha_{ps}^p(D,X;\gamma^*)\left(\Delta Y-\mathbb{E}[w_0^p(D,X;\gamma^*)\Delta Y]\right)\dot{\pi}(X;\gamma^*)\right]\\
&-w_1^p(D_i)\left(\mu_{0,\Delta}^p(X_i;\beta^*)-\mathbb{E}[w_1^p(D)\mu_{0,\Delta}^p(X;\beta^*)]\right)\\
&-l_{reg}(W_i;\beta^*)'\cdot\mathbb{E}[w_1^p(D)\dot{\mu}_{0,\Delta}^p(X;\beta^*)]\\
&+w_0^p(D_i,X_i;\gamma^*)\left(\mu_{0,\Delta}^p(X_i;\beta^*)-\mathbb{E}[w_0^p(D,X;\gamma^*)\mu_{0,\Delta}^p(X;\beta^*)]\right)\\
&+l_{ps}(D_i,X_i;\gamma^*)'\cdot\mathbb{E}\left[\alpha_{ps}^p(D,X;\gamma^*)\left(\mu_{0,\Delta}^p(X;\beta^*)-\mathbb{E}[w_0^p(D,X;\gamma^*)\mu_{0,\Delta}^p(X;\beta^*)]\right)\dot{\pi}(X;\gamma^*)\right]\\
&\left.+l_{reg}(W_i;\beta^*)'\cdot\mathbb{E}[w_0^p(D,X;\gamma^*)\dot{\mu}_{0,\Delta}^p(X;\beta^*)]\right)+o_p(1).
\end{aligned}
$$

After rearrangement and dropping the dependence of the functionals on W within $\mathbb{E}[\cdot]$, we have that

$$
\begin{aligned}
&\sqrt{n}(\widehat{\tau}^{dr,p}-\tau^{dr,p})\\
=&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(w_1^p(D_i)\left(\Delta Y_i-\mu_{0,\Delta}^p(X_i;\beta^*)-\mathbb{E}[w_1^p\cdot\left(\Delta Y-\mu_{0,\Delta}^p(\beta^*)\right)]\right)\right.\\
&-w_0^p(D_i,X_i;\gamma^*)\left(\Delta Y_i-\mu_{0,\Delta}^p(X_i;\beta^*)-\mathbb{E}[w_0^p(\gamma^*)\cdot\left(\Delta Y-\mu_{0,\Delta}^p(\beta^*)\right)]\right)\\
&-l_{reg}(W_i;\beta^*)'\cdot\mathbb{E}[\left(w_1^p-w_0^p(\gamma^*)\right)\cdot\dot{\mu}_{0,\Delta}^p(\beta^*)]\\
&\left.-l_{ps}(D_i,X_i;\gamma^*)'\cdot\mathbb{E}\left[\alpha_{ps}^p(\gamma^*)\cdot\left(\Delta Y-\mu_{0,\Delta}^p(\beta^*)-\mathbb{E}[w_0^p(\gamma^*)\cdot\left(\Delta Y-\mu_{0,\Delta}^p(\beta^*)\right)]\right)\dot{\pi}(\gamma^*)\right]\right)+o_p(1)\\
=&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta^p(W_i;\gamma^*,\beta^*)+o_p(1),
\end{aligned}\tag{S.9}
$$

since

$$
\begin{aligned}
\eta^{p}(W_i;\gamma^*,\beta^*) \equiv & \left(w_1^p(D_i)\left(\Delta Y_i - \mu_{0,\Delta}^p(X_i;\beta^*) - \mathbb{E}[w_1^p\cdot\left(\Delta Y - \mu_{0,\Delta}^p(\beta^*)\right)]\right)\right.\\
& - w_0^p(D_i,X_i;\gamma^*)\left(\Delta Y_i - \mu_{0,\Delta}^p(X_i;\beta^*) - \mathbb{E}[w_0^p(\gamma^*)\cdot\left(\Delta Y - \mu_{0,\Delta}^p(\beta^*)\right)]\right)\\
& - l_{reg}(W_i;\beta^*)'\cdot\mathbb{E}[\left(w_1^p - w_0^p(\gamma^*)\right)\cdot\dot{\mu}_{0,\Delta}^p(\beta^*)]\\
& - l_{ps}(D_i,X_i;\gamma^*)'\cdot\mathbb{E}\left[\alpha_{ps}^p(\gamma^*)\cdot\left(\Delta Y - \mu_{0,\Delta}^p(\beta^*) - \mathbb{E}[w_0^p(\gamma^*)\cdot\left(\Delta Y - \mu_{0,\Delta}^p(\beta^*)\right)]\right)\dot{\pi}(\gamma^*)\right]\\
& = \eta_1^p(W_i;\beta^*) - \eta_0^p(W_i;\gamma^*,\beta^*) - \eta_{est}^p(W_i;\gamma^*,\beta^*),
\end{aligned}
$$

where

$$
\begin{aligned}
\eta_1^p(W_i;\beta^*) &\equiv w_1^p(D_i)\left(\Delta Y_i - \mu_{0,\Delta}^p(X_i;\beta^*) - \mathbb{E}[w_1^p\left(\Delta Y - \mu_{0,\Delta}^p(\beta^*)\right)]\right),\\
\eta_0^p(W_i;\gamma^*,\beta^*) &\equiv w_0^p(D_i,X_i;\gamma^*)\left(\Delta Y_i - \mu_{0,\Delta}^p(X_i;\beta^*) - \mathbb{E}\left[w_0^p(\gamma^*)\left(\Delta Y - \mu_{0,\Delta}^p(\beta^*)\right)\right]\right),
\end{aligned}
$$

as well as

$$
\begin{aligned}
\eta_{est}^p(W_i;\gamma^*,\beta^*) \equiv & l_{reg}(Y_{1i},Y_{0i},X_i;\beta^*)'\cdot\mathbb{E}[\left(w_1^p - w_0^p(\gamma^*)\right)\cdot\dot{\mu}_{0,\Delta}^p(\beta^*)]\\
& + l_{ps}(D_i,X_i;\gamma^*)'\cdot\mathbb{E}\left[\alpha_{ps}^p(\gamma^*)\cdot\left(\Delta Y - \mu_{0,\Delta}^p(\beta^*) - \mathbb{E}\left[w_0^p(\gamma^*)\cdot\left(\Delta Y - \mu_{0,\Delta}^p(\beta^*)\right)\right]\right)\dot{\pi}(\gamma^*)\right].
\end{aligned}
$$

Thus, from (S.9) we conclude the proof of the asymptotic linear representation of $\sqrt{n}(\widehat{\tau}^{dr,p} - \tau^{dr,p})$. The asymptotic normality now follows directly from an application of the Lindeberg–Lévy central limit theorem, and this concludes the proof of part (a) of the Theorem.

For part (b), note that when both the propensity score working model and the working model for the outcome evolution for the comparison group are correctly specified, it follows that

$$
\begin{aligned}
\mathbb{E}[\left(w_1^p(D) - w_0^p(D,X;\gamma^*)\right)\dot{\mu}_{0,\Delta}^p(X;\beta^*)] &= \mathbb{E}\left[\frac{1}{\mathbb{E}[D]}\mathbb{E}\left[D - \frac{(1-D)\,p(X)}{1-p(X)}\middle|X\right]\dot{\mu}_{0,\Delta}^p(X;\beta^*)\right]\\
&= 0,
\end{aligned}
$$

where both equalities follow from the law of iterated expectations. Analogously, by the law of iterated expectations,

$$
\begin{aligned}
& \mathbb{E}\left[\alpha_{ps}^p(D,X;\gamma^*)\left(\Delta Y - \mu_{0,\Delta}^p(X;\beta^*)\right)\dot{\pi}(D,X;\gamma^*)\right]\\
= & \frac{\mathbb{P}(D=0)}{\mathbb{E}[D]}\cdot\mathbb{E}\left[\frac{1}{(1-\pi(X;\gamma^*))^2}\left(\Delta Y - \mu_{0,\Delta}^p(X;\beta^*)\right)\dot{\pi}(D,X;\gamma^*)\middle|D=0\right],\\
= & \frac{\mathbb{P}(D=0)}{\mathbb{E}[D]}\cdot\mathbb{E}\left[\frac{1}{(1-\pi(X;\gamma^*))^2}\left(\mathbb{E}[\Delta Y|D=0,X] - \mu_{0,\Delta}^p(X;\beta^*)\right)\dot{\pi}(0,X;\gamma^*)\middle|D=0\right]\\
= & 0
\end{aligned}
$$

and using the same arguments,

$$
\mathbb{E}\left[w_0^p(D,X;\gamma^*)\left(\Delta Y - \mu_{0,\Delta}^p(\beta^*)\right)\right] = 0.
$$

These observations imply that $\eta^p_{est}(W_i;\gamma^*,\beta^*) = 0$ $a.s.$, and therefore,

$$\begin{aligned}
\eta^p(W_i;\gamma^*,\beta^*) &= \left(w_1^p(D_i) - w_0^p(D_i,X_i;p)\right)\left(\Delta Y_i - m^p_{0,\Delta}(X_i)\right) \\
&\quad - w_1^p(D_i)\mathbb{E}[w_1^p \cdot \left(\Delta Y - m^p_{0,\Delta}(\beta^*)\right) \\
&= \left(w_1^p(D_i) - w_0^p(D_i,X_i;p)\right)\left(\Delta Y_i - m^p_{0,\Delta}(X_i)\right) - w_1^p(D_i)\tau \\
&= \eta^{e,p}(Y_1,Y_0,D,X),
\end{aligned}$$

where the first equality follows from simple algebra, the second one from the law of iterated expectation and Assumption 2, and the third one from simple manipulation. Given that the $\eta^p(W_i;\gamma^*,\beta^*) = \eta^{e,p}(Y_1,Y_0,D,X)$ $a.s.$, their asymptotic variance must coincide, which concludes the proof.■

Proof of Theorem A.2: Under Assumptions 1-A.2 and the additional assumption that $n_1/n \xrightarrow{p} \lambda \in (0,1)$, we first establish the large sample properties of the estimator $\widehat{\tau}_1^{dr,rc}$ when repeated cross-section data are available.

First of all, recall that the estimator takes the following form:

$$\widehat{\tau}_1^{dr,rc} = \mathbb{E}_n\left[\left(\widehat{w}_1^{rc}(D,T) - \widehat{w}_0^{rc}(D,T,X;\widehat{\gamma})\right)\left(Y - \mu^{rc}_{0,Y}(T,X;\widehat{\beta}^{rc}_{0,0},\widehat{\beta}^{rc}_{0,1})\right)\right],$$

where

$$\widehat{w}_1^{rc}(D,T) = \widehat{w}_{1,1}^{rc}(D,T) - \widehat{w}_{1,0}^{rc}(D,T),$$
$$\widehat{w}_0^{rc}(D,T,X;\gamma) = \widehat{w}_{0,1}^{rc}(D,T,X;\gamma) - \widehat{w}_{0,0}^{rc}(D,T,X;\gamma);$$

with

$$\begin{aligned}
\widehat{w}_{1,1}^{rc}(D,T) &= \frac{DT}{\mathbb{E}_n[DT]}, \\
\widehat{w}_{1,0}^{rc}(D,T) &= \frac{D(1-T)}{\mathbb{E}_n[D(1-T)]}, \\
\widehat{w}_{0,1}^{rc}(D,T,X;\gamma) &= \frac{\pi(X;\gamma)(1-D)T}{1-\pi(X;\gamma)} \bigg/ \mathbb{E}_n\left[\frac{\pi(X;\gamma)(1-D)T}{1-\pi(X;\gamma)}\right], \\
\widehat{w}_{0,1}^{rc}(D,T,X;\gamma) &= \frac{\pi(X;\gamma)(1-D)(1-T)}{1-\pi(X;\gamma)} \bigg/ \mathbb{E}_n\left[\frac{\pi(X;\gamma)(1-D)(1-T)}{1-\pi(X;\gamma)}\right];
\end{aligned}$$

and $\widehat{\gamma}$, $\widehat{\beta}^{rc}_{0,0}$ and $\widehat{\beta}^{rc}_{0,1}$ are estimators for the pseudo-true parameters γ^*, $\beta^{*,rc}_{0,0}$ and $\beta^{*,rc}_{0,1}$; and, for generic β^{rc}_0 and β^{rc}_1, $\mu^{rc}_{0,Y}(T,\cdot;\beta^{rc}_0,\beta^{rc}_1) = T\cdot\mu^{rc}_{0,1}(\cdot;\beta^{rc}_1) + (1-T)\cdot\mu^{rc}_{0,0}(\cdot;\beta^{rc}_0)$.

By the weak law of large numbers and continuous mapping theorem, we have that, as $n\to\infty$,

$$\widehat{\tau}_1^{dr,rc} \xrightarrow{p} \mathbb{E}\left[\left(w_1^{rc}(D,T) - w_0^{rc}(D,T,X;\gamma^*)\right)\left(Y - \mu^{rc}_{0,Y}(T,X;\beta^{*,rc}_{0,0},\beta^{*,rc}_{0,1})\right)\right], \tag{S.10}$$

where

$$w_1^{rc}(D,T) = w_{1,1}^{rc}(D,T) - w_{1,0}^{rc}(D,T),$$
$$w_0^{rc}(D,T,X;\gamma) = w_{0,1}^{rc}(D,T,X;\gamma) - w_{0,0}^{rc}(D,T,X;\gamma),$$

and

$$\begin{aligned}
w_{1,1}^{rc}(D,T) &= \frac{DT}{\mathbb{E}[DT]},\\
w_{1,0}^{rc}(D,T) &= \frac{D(1-T)}{\mathbb{E}[D(1-T)]},\\
w_{0,1}^{rc}(D,T,X;\gamma) &= \frac{\pi(X;\gamma)(1-D)T}{1-\pi(X;\gamma)} \Big/ \mathbb{E}[\frac{\pi(X;\gamma)(1-D)T}{1-\pi(X;\gamma)}],\\
w_{0,0}^{rc}(D,T,X;\gamma) &= \frac{\pi(X;\gamma)(1-D)(1-T)}{1-\pi(X;\gamma)} \Big/ \mathbb{E}[\frac{\pi(X;\gamma)(1-D)(1-T)}{1-\pi(X;\gamma)}].
\end{aligned}$$

Thus, if either $\pi(X;\gamma^*)=p(X)$ a.s. or $\mu_{0,\Delta}^{rc}(T,X;\beta_{0,0}^{*,rc},\beta_{0,1}^{*rc})=m_{0,1}^{*,rc}(X)-m_{0,0}^{*,rc}(X)$ a.s., it follows from Theorem 1 that

$$\mathbb{E}\left[\left(w_1^{rc}(D,T)-w_0^{rc}(D,T,X;\gamma^*)\right)\left(Y-\mu_{0,Y}^{rc}(T,X;\beta_{0,0}^{*,rc},\beta_{0,1}^{*,rc})\right)\right]\equiv\tau_1^{dr,rc}=ATT, \tag{S.11}$$

which completes the convergence in probability result.

Next, we establish the asymptotically linear representation of $\widehat{\tau}_1^{dr,rc}$. Following the same procedure of the proof of Theorem A.1, we first obtain the decomposition

$$\begin{aligned}
&\widehat{\tau}_1^{dr,rc}-\tau_1^{dr,rc}\\
&=\left(\mathbb{E}_n\left[\widehat{w}_{1,1}^{rc}(D,T)Y\right]-\mathbb{E}[w_{1,1}^{rc}(D,T)Y]\right)\\
&\quad-\left(\mathbb{E}_n\left[\widehat{w}_{1,0}^{rc}(D,T)Y\right]-\mathbb{E}[w_{1,0}^{rc}(D,T)Y]\right)\\
&\quad-\left(\mathbb{E}_n\left[\widehat{w}_{0,1}^{rc}(D,T,X;\widehat{\gamma})Y\right]-\mathbb{E}\left[w_{0,1}^{rc}(D,T,X;\gamma^*)Y\right]\right)\\
&\quad+\left(\mathbb{E}_n\left[\widehat{w}_{0,0}^{rc}(D,T,X;\widehat{\gamma})Y\right]-\mathbb{E}[w_{0,0}^{rc}(D,T,X;\gamma^*)Y]\right)\\
&\quad-\left(\mathbb{E}_n\left[\widehat{w}_{1,1}^{rc}(D,T)\mu_{0,Y}^{rc}\left(T,X;\widehat{\beta}_0^{rc},\widehat{\beta}_1^{rc}\right)\right]-\mathbb{E}[w_{1,1}^{rc}(D,T)\mu_{0,Y}^{rc}(T,X;\beta_{0,0}^{*,rc},\beta_{0,1}^{*,rc})]\right)\\
&\quad+\left(\mathbb{E}_n\left[\widehat{w}_{1,0}^{rc}(D,T)\mu_{0,Y}^{rc}\left(T,X;\widehat{\beta}_0^{rc},\widehat{\beta}_1^{rc}\right)\right]-\mathbb{E}[w_{1,0}^{rc}(D,T)\mu_{0,Y}^{rc}(T,X;\beta_{0,0}^{*,rc},\beta_{0,1}^{*,rc})]\right)\\
&\quad+\left(\mathbb{E}_n\left[\widehat{w}_{0,1}^{rc}(D,T,X;\widehat{\gamma})\mu_{0,Y}^{rc}\left(T,X;\widehat{\beta}_0^{rc},\widehat{\beta}_1^{rc}\right)\right]-\mathbb{E}[w_{0,1}^{rc}(D,T,X;\gamma^*)\mu_{0,Y}^{rc}(T,X;\beta_{0,0}^{*,rc},\beta_{0,1}^{*,rc})]\right)\\
&\quad-\left(\mathbb{E}_n\left[\widehat{w}_{0,0}^{rc}(D,T,X;\widehat{\gamma})\mu_{0,Y}^{rc}\left(T,X;\widehat{\beta}_0^{rc},\widehat{\beta}_1^{rc}\right)\right]-\mathbb{E}[w_{0,0}^{rc}(D,T,X;\gamma^*)\mu_{0,Y}^{rc}(T,X;\beta_{0,0}^{*,rc},\beta_{0,1}^{*,rc})]\right)\\
&\equiv\left(\widehat{ATT}_1-ATT_1\right)-\left(\widehat{ATT}_2-ATT_2\right)-\left(\widehat{ATT}_3-ATT_3\right)+\left(\widehat{ATT}_4-ATT_4\right)\\
&\quad-\left(\widehat{ATT}_5-ATT_5\right)+\left(\widehat{ATT}_6-ATT_6\right)+\left(\widehat{ATT}_7-ATT_7\right)-\left(\widehat{ATT}_8-ATT_8\right).
\end{aligned}$$

Hence, in obvious notation,

$$\begin{aligned}
&\sqrt{n}(\widehat{\tau}_1^{dr,rc}-\tau_1^{dr,rc})\\
=&\sqrt{n}\left(\widehat{ATT}_1-ATT_1\right)-\sqrt{n}\left(\widehat{ATT}_2-ATT_2\right)-\sqrt{n}\left(\widehat{ATT}_3-ATT_3\right)+\sqrt{n}\left(\widehat{ATT}_4-ATT_4\right)\\
&-\sqrt{n}\left(\widehat{ATT}_5-ATT_5\right)+\sqrt{n}\left(\widehat{ATT}_6-ATT_6\right)+\sqrt{n}\left(\widehat{ATT}_7-ATT_7\right)-\sqrt{n}\left(\widehat{ATT}_8-ATT_8\right),
\end{aligned}$$

and we next obtain the asymptotically linear representation for each component in the above decomposition.

For $\sqrt{n}\left(\widehat{ATT}_1 - ATT_1\right)$, following the analogous steps to derive (S.4) in the panel data case, we have that

$$\sqrt{n}\left(\widehat{ATT}_1 - ATT_1\right) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} w_{1,1}^{rc}(D_i,T_i)\left(Y_i - \mathbb{E}[w_{1,1}^{rc}(D,T)Y]\right) + o_p(1). \tag{S.12}$$

Analogously, for $\sqrt{n}\left(\widehat{ATT}_2 - ATT_2\right)$ we have that

$$\sqrt{n}\left(\widehat{ATT}_2 - ATT_2\right) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} w_{1,0}^{rc}(D_i,T_i)\left(Y_i - \mathbb{E}[w_{1,0}^{rc}(D,T)Y]\right) + o_p(1). \tag{S.13}$$

The case of $\sqrt{n}\left(\widehat{ATT}_3 - ATT_3\right)$ is similar to (S.5)-(S.6) in the panel data case. That is, by following similar steps as above and a second-order Taylor expansion argument, we have that

$$\begin{aligned}
&\sqrt{n}\left(\widehat{ATT}_3 - ATT_3\right)\\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n} \widetilde{w}_{0,1}^{rc}(D_i,T_i,X_i;\widehat{\gamma})\left(Y_i - \mathbb{E}[w_{0,1}^{rc}(D,T,X;\gamma^*)Y]\right) + o_p(1)\\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(w_{0,1}^{rc}(D_i,T_i,X_i;\gamma^*)\left(Y_i - \mathbb{E}[w_{0,1}^{rc}(D,T,X;\gamma^*)Y]\right)\right.\\
&\left. + l_{ps}(D_i,X_i;\gamma^*)' \cdot \mathbb{E}\left[\alpha_{ps,1}^{rc}(D,T,X;\gamma^*)\left(Y - \mathbb{E}\left[w_{0,1}^{rc}(D,T,X;\gamma^*)Y\right]\right)\dot{\pi}(X;\gamma^*)\right]\right) + o_p(1)
\end{aligned} \tag{S.14}$$

where

$$\begin{aligned}
\widetilde{w}_{0,1}^{rc}(D,T,X;\widehat{\gamma}) &= \frac{\pi(X;\widehat{\gamma})(1-D)T}{1-\pi(X;\widehat{\gamma})} \Bigg/ \mathbb{E}\left[\frac{\pi(X;\gamma^*)(1-D)T}{1-\pi(X;\gamma^*)}\right],\\
\alpha_{ps,1}^{rc}(D,T,X;\gamma) &= \frac{(1-D)T}{(1-\pi(X;\gamma))^2} \Bigg/ \mathbb{E}\left[\frac{\pi(X;\gamma)(1-D)T}{1-\pi(X;\gamma)}\right].
\end{aligned}$$

Using analogous arguments as in (S.14), we get

$$\begin{aligned}
&\sqrt{n}\left(\widehat{ATT}_4 - ATT_4\right)\\
=&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(w_{0,0}^{rc}(D_i,T_i,X_i;\gamma^*)\left(Y_i - \mathbb{E}[w_{0,0}^{rc}(D,T,X;\gamma^*)Y]\right)\right.\\
&\left.+l_{ps}(D_i,X_i;\gamma^*)' \cdot \mathbb{E}\left[\alpha_{ps,0}^{rc}(D,T,X;\gamma^*)\left(Y - \mathbb{E}[w_{0,0}^{rc}(D,T,X;\gamma^*)Y]\right)\dot{\pi}(X;\gamma^*)\right]\right) + o_p(1),
\end{aligned} \tag{S.15}$$

where

$$\alpha_{ps,0}^{rc}(D,T,X;\gamma) = \frac{(1-D)(1-T)}{(1-\pi(X;\gamma))^2} \Bigg/ \mathbb{E}\left[\frac{\pi(X;\gamma)(1-D)(1-T)}{1-\pi(X;\gamma)}\right].$$

We next derive the asymptotic linear representation of the terms associated with the "regression components". Starting with $\sqrt{n}\left(\widehat{ATT}_5 - ATT_5\right)$, by following similar steps as in (S.7), it is easy to show that

$$\begin{aligned}
&\sqrt{n}\left(\widehat{ATT}_5 - ATT_5\right)\\
=&\frac{1}{\sqrt{n}}\sum_{i=1}^{n} w_{1,1}^{rc}(D_i,T_i)\left(\mu_{0,Y}^{rc}(T_i,X_i;\widehat{\beta}_0^{rc},\widehat{\beta}_1^{rc}) - \mathbb{E}[w_{1,1}^{rc}(D,T)\mu_{0,Y}^{rc}(T,X;\beta_{0,0}^{*,rc},\beta_{0,1}^{*,rc})]\right) + o_p(1).
\end{aligned}$$

Let $\widehat{\beta} = (\widehat{\beta}_{0,1}^{rc}, \widehat{\beta}_{0,0}^{rc})$, and $\beta^* = (\beta_{0,0}^{*,rc}, \beta_{0,1}^{*,rc})$. Next, from a second-order Taylor expansion argument, we obtain that

$$\sqrt{n}\left(\widehat{ATT}_5 - ATT_5\right)$$
$$= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(w_{1,1}^{rc}(D_i,T_i)\left(\mu_{0,Y}^{rc}(T_i,X_i;\beta^*) - \mathbb{E}[w_{1,1}^{rc}(D,T)\mu_{0,Y}^{rc}(T,X;\beta^*)]\right)\right.$$
$$\left. + l_{reg}(W_i;\beta^*)'\cdot\mathbb{E}\left[w_{1,1}^{rc}(D,T)\dot{\mu}_{0,Y}^{rc}(T,X;\beta^*)\right]\right) + o_p(1). \quad \text{(S.16)}$$

Using the same arguments as in (S.16), we have

$$\sqrt{n}\left(\widehat{ATT}_6 - ATT_6\right)$$
$$= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(w_{1,0}^{rc}(D_i,T_i)\left(\mu_{0,Y}^{rc}(T_i,X_i;\beta^*) - \mathbb{E}[w_{1,0}^{rc}(D,T)\mu_{0,Y}^{rc}(T,X;\beta^*)]\right)\right.$$
$$\left. + l_{reg}(W_i;\beta^*)'\cdot\mathbb{E}\left[w_{1,0}^{rc}(D,T)\dot{\mu}_{0,Y}^{rc}(T,X;\beta^*)\right]\right) + o_p(1). \quad \text{(S.17)}$$

The asymptotically linear representation of $\sqrt{n}\left(\widehat{ATT}_7 - ATT_7\right)$ can be derived following similar steps as in (S.8). More precisely, one can easily show that

$$\sqrt{n}\left(\widehat{ATT}_7 - ATT_7\right)$$
$$= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\widetilde{w}_{0,1}^{rc}(D_i,T_i,X_i;\widehat{\gamma})\left(\mu_{0,Y}^{rc}(T_i,X_i;\widehat{\beta}) - \mathbb{E}[w_{0,1}^{rc}(D,T,X;\gamma^*)\mu_{0,Y}^{rc}(T,X;\beta^*)]\right) + o_p(1).$$

Then, by doing a second-order Taylor expansion of the above expression around pseudo-true γ^* and β^*, and plugging the asymptotically linear representations of $\sqrt{n}(\widehat{\gamma} - \gamma^*)$ and $\sqrt{n}\left(\widehat{\beta} - \beta^*\right)$, we have

$$\sqrt{n}\left(\widehat{ATT}_7 - ATT_7\right)$$
$$= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(w_{0,1}^{rc}(D_i,T_i,X_i;\gamma^*)\left(\mu_{0,Y}^{rc}(T_i,X_i;\beta^*) - \mathbb{E}[w_{0,1}^{rc}(\gamma^*)\mu_{0,Y}^{rc}(\beta^*)]\right)\right.$$
$$+ l_{ps}(D_i,X_i;\gamma^*)'\cdot\mathbb{E}\left[\alpha_{ps,1}^{rc}(\gamma^*)\cdot\left(\mu_{0,Y}^{rc}(\beta^*) - \mathbb{E}[w_{0,1}^{rc}(\gamma^*)\cdot\mu_{0,Y}^{rc}(\beta^*)]\right)\dot{\pi}(\gamma^*)\right]$$
$$\left. + l_{reg}(W_i;\beta^*)'\cdot\mathbb{E}\left[w_{0,1}^{rc}(\gamma^*)\cdot\dot{\mu}_{0,Y}^{rc}(\beta^*)\right]\right) + o_p(1), \quad \text{(S.18)}$$

where the dependence of the functionals on W within $\mathbb{E}[\cdot]$ is dropped to ease the notation. Analogously,

$$\sqrt{n}\left(\widehat{ATT}_8 - ATT_8\right)$$
$$= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(w_{0,0}^{rc}(D_i,T_i,X_i;\gamma^*)\left(\mu_{0,Y}^{rc}(T_i,X_i;\beta^*) - \mathbb{E}[w_{0,0}^{rc}(\gamma^*)\cdot\mu_{0,Y}^{rc}(\beta^*)]\right)\right.$$
$$+ l_{ps}(D_i,X_i;\gamma^*)'\cdot\mathbb{E}\left[\alpha_{ps,0}^{rc}(\gamma^*)\cdot\left(\mu_{0,Y}^{rc}(\beta^*) - \mathbb{E}[w_{0,0}^{rc}(\gamma^*)\cdot\mu_{0,Y}^{rc}(\beta^*)]\right)\dot{\pi}(\gamma^*)\right]$$
$$\left. + l_{reg}(W_i;\beta^*)'\cdot\mathbb{E}\left[w_{0,0}^{rc}(\gamma^*)\cdot\dot{\mu}_{0,Y}^{rc}(\beta^*)\right]\right) + o_p(1). \quad \text{(S.19)}$$

Finally, by combining and rearranging (S.12)-(S.19), it follows that

$$\sqrt{n}(\widehat{\tau}_1^{dr,rc} - \tau_1^{dr,rc})$$

$$
\begin{aligned}
=&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Big(w_{1,1}^{rc}(D_i,T_i)\left(Y_i-\mu_{0,1}^{rc}(X_i;\beta^*)-\mathbb{E}[w_{1,1}^{rc}\cdot\left(Y-\mu_{0,1}^{rc}(\beta^*)\right)]\right)\\
&-w_{1,0}^{rc}(D_i,T_i)\left(Y_i-\mu_{0,0}^{rc}(X_i;\beta^*)-\mathbb{E}[w_{1,0}^{rc}\cdot\left(Y-\mu_{0,0}^{rc}(\beta^*)\right)]\right)\\
&-w_{0,1}^{rc}(D_i,T_i,X_i;\gamma^*)\left(Y_i-\mu_{0,1}^{rc}(X_i;\beta^*)-\mathbb{E}[w_{0,1}^{rc}(\gamma^*)\cdot\left(Y-\mu_{0,1}^{rc}(\beta^*)\right)]\right)\\
&+w_{0,0}^{rc}(D_i,T_i,X_i;\gamma^*)\left(Y_i-\mu_{0,0}^{rc}(T_i,X_i;\beta^*)-\mathbb{E}[w_{0,0}^{rc}(\gamma^*)\cdot\left(Y-\mu_{0,0}^{rc}(\beta^*)\right)]\right)\\
&-l_{reg}(W_i;\beta^*)'\cdot\mathbb{E}[(w_{1,1}^{rc}-w_{1,0}^{rc})-(w_{0,1}^{rc}(\gamma^*)-w_{0,0}^{rc}(\gamma^*))\cdot\dot{\mu}_{0,Y}^{rc}(\beta^*)]\\
&-l_{ps}(D_i,X_i;\gamma^*)'\cdot\mathbb{E}\left[\alpha_{ps,1}^{rc}(\gamma^*)\cdot\left(Y-\mu_{0,Y}^{rc}(\beta^*)-\mathbb{E}[w_{0,1}^{rc}(\gamma^*)\cdot\left(Y-\mu_{0,Y}^{rc}(\beta^*)\right)]\right)\dot{\pi}(\gamma^*)\right]\\
&+l_{ps}(D_i,X_i;\gamma^*)'\cdot\mathbb{E}\left[\alpha_{ps,0}^{rc}(\gamma^*)\cdot\left(Y-\mu_{0,Y}^{rc}(\beta^*)-\mathbb{E}[w_{0,0}^{rc}(\gamma^*)\cdot\left(Y-\mu_{0,Y}^{rc}(\beta^*)\right)]\right)\dot{\pi}(\gamma^*)\right]\Big)\\
&+o_p(1)\\
=&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta_1^{rc}(W_i;\gamma^*,\beta^*)+o_p(1),
\end{aligned}
\tag{S.20}
$$

where

$$
\eta_1^{rc}(W_i;\gamma^*,\beta^*)=\eta_1^{rc,1}(W_i;\beta^*)-\eta_0^{rc,1}(W_i;\gamma^*,\beta^*)-\eta_{est}^{rc,1}(W_i;\gamma^*,\beta^*),
$$

with

$$
\begin{aligned}
\eta_1^{rc,1}(W_i;\beta^*)&=\eta_{1,1}^{rc,1}(W_i;\beta^*)-\eta_{1,0}^{rc,1}(W_i;\beta^*),\\
\eta_0^{rc,1}(W_i;\gamma^*,\beta^*)&=\eta_{0,1}^{rc,1}(W_i;\gamma^*,\beta^*)-\eta_{0,0}^{rc,1}(W_i;\gamma^*,\beta^*),\\
\eta_{est}^{rc,1}(W_i;\gamma^*,\beta^*)&=\eta_{est,reg}^{rc,1}(W_i;\gamma^*,\beta^*)+\eta_{est,ps}^{rc,1}(W_i;\gamma^*,\beta^*),
\end{aligned}
$$

and, for $t=0,1$,

$$
\begin{aligned}
\eta_{1,t}^{rc,1}(W_i;\gamma^*,\beta^*)&=w_{1,t}^{rc}(D_i,T_i)\left(Y_i-\mu_{0,t}^{rc}(X_i;\beta^*)-\mathbb{E}[w_{1,t}^{rc}\cdot\left(Y-\mu_{0,t}^{rc}(\beta^*)\right)]\right),\\
\eta_{0,t}^{rc,1}(W_i;\gamma^*,\beta^*)&=w_{0,t}^{rc}(D_i,T_i,X_i;\gamma^*)\left(Y_i-\mu_{0,t}^{rc}(X_i;\beta^*)-\mathbb{E}[w_{0,t}^{rc}(\gamma^*)\cdot\left(Y-\mu_{0,t}^{rc}(\beta^*)\right)]\right),
\end{aligned}
$$

and

$$
\begin{aligned}
&\eta_{est,reg}^{rc,1}(W_i;\gamma^*,\beta^*)=l_{reg}(W_i;\beta^*)'\cdot\mathbb{E}[(w_{1,1}^{rc}-w_{1,0}^{rc})-(w_{0,1}^{rc}(\gamma^*)-w_{0,0}^{rc}(\gamma^*))\cdot\dot{\mu}_{0,Y}^{rc}(\beta^*)],\\
&\eta_{est,ps}^{rc,1}(W_i;\gamma^*,\beta^*)\\
=&l_{ps}(D_i,X_i;\gamma^*)'\cdot\mathbb{E}\left[\alpha_{ps,1}^{rc}(\gamma^*)\cdot\left(Y-\mu_{0,1}^{rc}(\beta^*)-\mathbb{E}[w_{0,1}^{rc}(\gamma^*)\cdot\left(Y-\mu_{0,1}^{rc}(\beta^*)\right)]\right)\dot{\pi}(\gamma^*)\right]\\
&-l_{ps}(D_i,X_i;\gamma^*)'\cdot\mathbb{E}\left[\alpha_{ps,0}^{rc}(\gamma^*)\cdot\left(Y-\mu_{0,0}^{rc}(\beta^*)-\mathbb{E}[w_{0,0}^{rc}(\gamma^*)\cdot\left(Y-\mu_{0,0}^{rc}(\beta^*)\right)]\right)\dot{\pi}(\gamma^*)\right].
\end{aligned}
$$

The asymptotic normality result for $\sqrt{n}(\widehat{\tau}_1^{dr,rc}-\tau_1^{dr,rc})$ now follows immediately from (S.20) and the Lindeberg–Lévy central limit theorem.

Next, we study the asymptotic properties of

$$\widehat{\tau}_2^{dr,rc} = \widehat{\tau}_1^{dr,rc} + \left(\mathbb{E}_n \left[\left(\frac{D}{\mathbb{E}_n[D]} - \frac{DT}{\mathbb{E}_n[DT]} \right) \left(\mu_{1,1}^{rc}\left(X;\widehat{\beta}_{1,1}^{rc}\right) - \mu_{0,1}^{rc}\left(X;\widehat{\beta}_{0,1}^{rc}\right) \right) \right] \right) - \left(\mathbb{E}_n \left[\left(\frac{D}{\mathbb{E}_n[D]} - \frac{D(1-T)}{\mathbb{E}_n[D(1-T)]} \right) \left(\mu_{1,0}^{rc}\left(X;\widehat{\beta}_{1,0}^{rc}\right) - \mu_{0,0}^{rc}\left(X;\widehat{\beta}_{0,0}^{rc}\right) \right) \right] \right),$$

where $\widehat{\beta}_{1,1}^{rc}$ and $\widehat{\beta}_{1,0}^{rc}$ are estimators for the pseudo-true parameters $\beta_{1,1}^{*,rc}$ and $\beta_{1,0}^{rc}$, respectively, and all other finite dimensional parameters are defined as before. By weak law of large numbers, continuous mapping theorem, and (S.10), as $n \to \infty$,

$$\widehat{\tau}_2^{dr,rc} \xrightarrow{p} \tau_1^{dr,rc} + \left(\mathbb{E} \left[\left(\frac{D}{\mathbb{E}[D]} - \frac{DT}{\mathbb{E}[DT]} \right) \left(\mu_{1,1}^{rc}\left(X;\beta_{1,1}^{*,rc}\right) - \mu_{0,1}^{rc}\left(X;\beta_{0,1}^{*,rc}\right) \right) \right] \right) - \left(\mathbb{E} \left[\left(\frac{D}{\mathbb{E}[D]} - \frac{D(1-T)}{\mathbb{E}[D(1-T)]} \right) \left(\mu_{1,0}^{rc}\left(X;\beta_{1,0}^{*,rc}\right) - \mu_{0,0}^{rc}\left(X;\beta_{0,0}^{*,rc}\right) \right) \right] \right).$$

In turn, by the total law of expectation and the stationarity condition in Assumption 1(b), it follows that

$$\mathbb{E} \left[\left(\frac{D}{\mathbb{E}[D]} - \frac{DT}{\mathbb{E}[DT]} \right) \left(\mu_{1,1}^{rc}\left(X;\beta_{1,1}^{*,rc}\right) - \mu_{0,1}^{rc}\left(X;\beta_{0,1}^{*,rc}\right) \right) \right] = 0, \tag{S.21}$$

$$\mathbb{E} \left[\left(\frac{D}{\mathbb{E}[D]} - \frac{D(1-T)}{\mathbb{E}[D(1-T)]} \right) \left(\mu_{1,0}^{rc}\left(X;\beta_{1,0}^{*,rc}\right) - \mu_{0,0}^{rc}\left(X;\beta_{0,0}^{*,rc}\right) \right) \right] = 0, \tag{S.22}$$

implying that provided that either $\pi(X;\gamma^*) = p(X)$ a.s. or $\mu_{0,\Delta}^{rc}(T,X;\beta_{0,0}^{*,rc},\beta_{0,1}^{*,rc}) = m_{0,1}^{*,rc}(X) - m_{0,0}^{*,rc}(X)$ a.s..

$$\widehat{\tau}_2^{dr,rc} \xrightarrow{p} \tau \equiv ATT.$$

Next, we establish the asymptotically linear representation of $\widehat{\tau}_2^{dr,rc}$. Given that we have already establish the asymptotically linear representation of $\widehat{\tau}_1^{dr,rc}$, and (S.21)-(S.22) hold, it remains to study these two additional terms

$$\begin{aligned} \sqrt{n}\widehat{ATT}_9 &\equiv \sqrt{n}\mathbb{E}_n \left[\left(\frac{D}{\mathbb{E}_n[D]} - \frac{DT}{\mathbb{E}_n[DT]} \right) \left(\mu_{1,1}^{rc}\left(X;\widehat{\beta}_{1,1}^{rc}\right) - \mu_{0,1}^{rc}\left(X;\widehat{\beta}_{0,1}^{rc}\right) \right) \right], \\ \sqrt{n}\widehat{ATT}_{10} &\equiv \sqrt{n}\mathbb{E}_n \left[\left(\frac{D}{\mathbb{E}_n[D]} - \frac{D(1-T)}{\mathbb{E}_n[D(1-T)]} \right) \left(\mu_{1,0}^{rc}\left(X;\widehat{\beta}_{1,0}^{rc}\right) - \mu_{0,0}^{rc}\left(X;\widehat{\beta}_{0,0}^{rc}\right) \right) \right]. \end{aligned}$$

We first analyze $\sqrt{n}\widehat{ATT}_9$. By doing a second-order Taylor expansion of $\sqrt{n}\widehat{ATT}$ around the pseudo-true $\left(\beta_{1,1}^{*,rc}, \beta_{0,1}^{*rc}\right)$, we have that

$$\begin{aligned} \sqrt{n}\widehat{ATT}_9 &= \sqrt{n}\mathbb{E}_n \left[\left(\frac{D}{\mathbb{E}_n[D]} - \frac{DT}{\mathbb{E}_n[DT]} \right) \left(\mu_{1,1}^{rc}\left(X;\beta_{1,1}^{*,rc}\right) - \mu_{0,1}^{rc}\left(X;\beta_{0,1}^{*,rc}\right) \right) \right] \\ &\quad + \mathbb{E} \left[\left(\frac{D}{\mathbb{E}[D]} - \frac{DT}{\mathbb{E}[DT]} \right) \dot{\mu}_{1,1}^{rc}\left(X;\beta_{1,1}^{*,rc}\right) \right] \sqrt{n}\left(\widehat{\beta}_{1,1}^{rc} - \beta_{1,1}^{*,rc}\right) \\ &\quad + \mathbb{E} \left[\left(\frac{D}{\mathbb{E}[D]} - \frac{DT}{\mathbb{E}[DT]} \right) \dot{\mu}_{0,1}^{rc}\left(X;\beta_{0,1}^{*,rc}\right) \right] \sqrt{n}\left(\widehat{\beta}_{0,1}^{rc} - \beta_{0,1}^{*,rc}\right) + o_p(1) \\ &= \sqrt{n}\mathbb{E}_n \left[\left(\frac{D}{\mathbb{E}_n[D]} - \frac{DT}{\mathbb{E}_n[DT]} \right) \left(\mu_{1,1}^{rc}\left(X;\beta_{1,1}^{*,rc}\right) - \mu_{0,1}^{rc}\left(X;\beta_{0,1}^{*,rc}\right) \right) \right] + o_p(1)\,O_p(1) \\ &= \sqrt{n}\mathbb{E}_n \left[\left(\frac{D}{\mathbb{E}_n[D]} - \frac{DT}{\mathbb{E}_n[DT]} \right) \left(\mu_{1,1}^{rc}\left(X;\beta_{1,1}^{*,rc}\right) - \mu_{0,1}^{rc}\left(X;\beta_{0,1}^{*,rc}\right) \right) \right] + o_p(1), \end{aligned} \tag{S.23}$$

where the second equality follows from the stationarity condition in $1(b)$, and Assumptions A.1-A.2. Next, by a second-order Taylor expansion around $\mathbb{E}[D]$,

$$\begin{aligned}
&\sqrt{n}\mathbb{E}_n\left[\frac{D}{\mathbb{E}_n[D]}\left(\mu_{1,1}^{rc}\left(X;\beta_{1,1}^{*,rc}\right)-\mu_{0,1}^{rc}\left(X;\beta_{0,1}^{*,rc}\right)\right)\right]\\
=\;&\sqrt{n}\mathbb{E}_n\left[\frac{D}{\mathbb{E}[D]}\left(\mu_{1,1}^{rc}\left(X;\beta_{1,1}^{*,rc}\right)-\mu_{0,1}^{rc}\left(X;\beta_{0,1}^{*,rc}\right)\right)\right]\\
&-\mathbb{E}\left[\frac{D}{\mathbb{E}[D]}\left(\mu_{1,1}^{rc}\left(X;\beta_{1,1}^{*,rc}\right)-\mu_{0,1}^{rc}\left(X;\beta_{0,1}^{*,rc}\right)\right)\right]\frac{\sqrt{n}\mathbb{E}_n[D-\mathbb{E}[D]]}{\mathbb{E}[D]}+o_p(1).
\end{aligned}\tag{S.24}$$

Analogously, by a second-order Taylor expansion around $\mathbb{E}[DT]$,

$$\begin{aligned}
&\sqrt{n}\mathbb{E}_n\left[\frac{DT}{\mathbb{E}_n[DT]}\left(\mu_{1,1}^{rc}\left(X;\beta_{1,1}^{*,rc}\right)-\mu_{0,1}^{rc}\left(X;\beta_{0,1}^{*,rc}\right)\right)\right]\\
=\;&\sqrt{n}\mathbb{E}_n\left[w_{1,1}^{rc}(D,T)\left(\mu_{1,1}^{rc}\left(X;\beta_{1,1}^{*,rc}\right)-\mu_{0,1}^{rc}\left(X;\beta_{0,1}^{*,rc}\right)\right)\right]\\
&-\mathbb{E}\left[w_{1,1}^{rc}(D,T)\left(\mu_{1,1}^{rc}\left(X;\beta_{1,1}^{*,rc}\right)-\mu_{0,1}^{rc}\left(X;\beta_{0,1}^{*,rc}\right)\right)\right]\frac{\sqrt{n}\mathbb{E}_n[DT-\mathbb{E}[DT]]}{\mathbb{E}[DT]}+o_p(1).
\end{aligned}\tag{S.25}$$

Hence, from (S.21)-(S.25), the stationarity condition, and some simple algebra, we conclude that

$$\begin{aligned}
\sqrt{n}\widehat{ATT}_9=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Bigg(&\frac{D_i}{\mathbb{E}[D]}\left(\mu_{1,1}^{rc}\left(X_i;\beta_{1,1}^{*,rc}\right)-\mathbb{E}\left[\frac{D}{\mathbb{E}[D]}\mu_{1,1}^{rc}\left(X;\beta_{1,1}^{*,rc}\right)\right]\right)\\
&-\frac{D_i}{\mathbb{E}[D]}\left(\mu_{0,1}^{rc}\left(X_i;\beta_{0,1}^{*,rc}\right)-\mathbb{E}\left[\frac{D}{\mathbb{E}[D]}\mu_{0,1}^{rc}\left(X;\beta_{0,1}^{*,rc}\right)\right]\right)\\
&-w_{1,1}^{rc}(D_i,T_i)\left(\mu_{1,1}^{rc}\left(X;\beta_{1,1}^{*,rc}\right)-\mathbb{E}\left[w_{1,1}^{rc}(D,T)\mu_{1,1}^{rc}\left(X;\beta_{1,1}^{*,rc}\right)\right]\right)\\
&+w_{1,1}^{rc}(D_i,T_i)\left(\mu_{0,1}^{rc}\left(X_i;\beta_{0,1}^{*,rc}\right)-\mathbb{E}\left[w_{1,1}^{rc}(D,T)\mu_{0,1}^{rc}\left(X;\beta_{0,1}^{*,rc}\right)\right]\right)\Bigg)+o_p(1).
\end{aligned}\tag{S.26}$$

Analogously, we have that

$$\begin{aligned}
\sqrt{n}\widehat{ATT}_{10}=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Bigg(&\frac{D_i}{\mathbb{E}[D]}\left(\mu_{1,0}^{rc}\left(X_i;\beta_{1,0}^{*,rc}\right)-\mathbb{E}\left[\frac{D}{\mathbb{E}[D]}\mu_{1,0}^{rc}\left(X;\beta_{1,0}^{*,rc}\right)\right]\right)\\
&-\frac{D_i}{\mathbb{E}[D]}\left(\mu_{0,0}^{rc}\left(X_i;\beta_{0,0}^{*,rc}\right)-\mathbb{E}\left[\frac{D}{\mathbb{E}[D]}\mu_{0,0}^{rc}\left(X;\beta_{0,0}^{*,rc}\right)\right]\right)\\
&-w_{1,0}^{rc}(D_i,T_i)\left(\mu_{1,0}^{rc}\left(X;\beta_{1,0}^{*,rc}\right)-\mathbb{E}\left[w_{1,0}^{rc}(D,T)\mu_{1,0}^{rc}\left(X;\beta_{1,0}^{*,rc}\right)\right]\right)\\
&+w_{1,0}^{rc}(D_i,T_i)\left(\mu_{0,0}^{rc}\left(X_i;\beta_{0,0}^{*,rc}\right)-\mathbb{E}\left[w_{1,0}^{rc}(D,T)\mu_{0,0}^{rc}\left(X;\beta_{0,0}^{*,rc}\right)\right]\right)\Bigg)+o_p(1).
\end{aligned}\tag{S.27}$$

Hence, combining (S.20) with (S.26) and (S.27), it follows that

$$\sqrt{n}(\widehat{\tau}_2^{dr,rc}-\tau_2^{dr,rc})=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta_2^{rc}(W_i;\gamma^*,\beta^*)+o_p(1),\tag{S.28}$$

with

$$\eta_2^{rc}(W_i;\gamma^*,\beta^*)\equiv\eta_1^{rc,2}(W_i;\beta^*)-\eta_0^{rc,2}(W_i;\gamma^*,\beta^*)-\eta_{est}^{rc,2}(W_i;\gamma^*,\beta^*),$$

where $\beta^*=\left(\beta_{1,1}^{*,rc},\beta_{1,0}^{*,rc},\beta_{0,1}^{*,rc},\beta_{0,0}^{*,rc}\right)$,

$$\eta_1^{rc,2}(W_i;\beta^*)=\eta_{1,1}^{rc,2}(W_i;\beta^*)-\eta_{1,0}^{rc,2}(W_i;\beta^*),$$

$$\eta_0^{rc,2}(W_i;\gamma^*,\beta^*) = \eta_0^{rc,1}(W_i;\gamma^*,\beta^*)$$
$$\eta_{est}^{rc,2}(W_i;\gamma^*,\beta^*) = \eta_{est}^{rc,1}(W_i;\gamma^*,\beta^*),$$

and

$$\begin{aligned}
\eta_{1,1}^{rc,2}(W_i;\beta^*) &= \frac{D_i}{\mathbb{E}[D]}\left(\mu_{1,\Delta}^{rc}\left(X_i;\beta_{1,1}^{*,rc},\beta_{1,0}^{*,rc}\right)-\mathbb{E}\left[\frac{D}{\mathbb{E}[D]}\mu_{1,\Delta}^{rc}\left(X;\beta_{1,1}^{*,rc},\beta_{1,0}^{*,rc}\right)\right]\right)\\
&\quad +w_{1,1}^{rc}(D_i,T_i)\left(\left(Y_i-\mu_{1,1}^{rc}\left(X_i;\beta_{1,1}^{*,rc}\right)\right)-\mathbb{E}[w_{1,1}^{rc}\cdot\left(Y-\mu_{1,1}^{rc}\left(X;\beta_{1,1}^{*,rc}\right)\right)]\right),\\
\eta_{1,0}^{rc,2}(W_i;\beta^*) &= \frac{D_i}{\mathbb{E}[D]}\left(\mu_{0,\Delta}^{rc}\left(X_i;\beta_{0,1}^{*,rc},\beta_{0,0}^{*,rc}\right)-\mathbb{E}\left[\frac{D}{\mathbb{E}[D]}\mu_{0,\Delta}^{rc}\left(X;\beta_{0,1}^{*,rc},\beta_{0,0}^{*,rc}\right)\right]\right)\\
&\quad +w_{1,0}^{rc}(D_i,T_i)\left(Y_i-\mu_{1,0}^{rc}\left(X_i;\beta_{1,0}^{*,rc}\right)-\mathbb{E}[w_{1,0}^{rc}\cdot\left(Y-\mu_{1,0}^{rc}\left(X;\beta_{1,0}^{*,rc}\right)\right)]\right).
\end{aligned}$$

The asymptotic normality result for $\sqrt{n}(\widehat{\tau}_2^{dr,rc}-\tau_2^{dr,rc})$ now follows immediately from (S.28) and the Lindeberg–Lévy central limit theorem.

Now, we prove part (b) of Theorem A.2. When (A.3) and (A.7) are true, we have that, for $d,t=0,1$,

$$\begin{aligned}
\mathbb{E}\left[\frac{D}{\mathbb{E}[D]}\left(\mu_{1,\Delta}^{rc}\left(X;\beta_{1,1}^{*,rc},\beta_{1,0}^{*,rc}\right)-\mu_{0,\Delta}^{rc}\left(X;\beta_{0,1}^{*,rc},\beta_{0,0}^{*,rc}\right)\right)\right] &= \tau,\\
\mathbb{E}[w_{d,t}^{rc}\cdot\left(Y-\mu_{d,t}^{rc}\left(X;\beta_{d,t}^{*,rc}\right)\right)] &= 0,\\
\mathbb{E}\left[\frac{(1-D)\,T}{1-\pi(X;\gamma^*)}\left(Y-\mu_{0,1}^{rc}(X;\beta_{0,1}^{*,rc})\right)\dot{\pi}(X;\gamma^*)\right] &= 0,\\
\mathbb{E}\left[\frac{(1-D)\,(1-T)}{1-\pi(X;\gamma^*)}\left(Y-\mu_{0,0}^{rc}(X;\beta_{0,0}^{*,rc})\right)\dot{\pi}(X;\gamma^*)\right] &= 0,\\
\mathbb{E}[(w_{1,t}^{rc}(D,T)-w_{0,t}^{rc}(D,T,X;\gamma^*))\dot{\mu}_{0,t}^{rc}(X;\beta^*)] &= 0,
\end{aligned}$$

implying that

$$\begin{aligned}
\eta_1^{rc,2}(W_i;\beta^*) &= \frac{D_i}{\mathbb{E}[D]}\left(m_{1,\Delta}^{rc}(X_i)-m_{0,\Delta}^{rc}(X_i)-\tau\right)\\
&\quad +w_{1,1}^{rc}(D_i,T_i)\left(\left(Y_i-m_{1,1}^{rc}(X_i)\right)\right)-w_{1,0}^{rc}(D_i,T_i)\left(Y_i-m_{1,0}^{rc}(X_i)\right),\\
\eta_0^{rc,2}(W_i;\gamma^*,\beta^*) &= w_{0,1}^{rc}(D_i,T_i,X_i;p)\left(Y_i-m_{0,1}^{rc}(X_i)\right)-w_{0,0}^{rc}(D_i,T_i,X_i;p)\left(Y_i-m_{0,0}^{rc}(X_i)\right),\\
\eta_{est}^{rc,2}(W_i;\gamma^*,\beta^*) &= 0.
\end{aligned}$$

Therefore, $\eta_2^{rc}(W_i;\gamma^*,\beta^*)=\eta^{e,rc}(Y,D,T,X)$ $a.s.$, and we conclude that the asymptotic variance of the DR DID estimator $\widehat{\tau}_2^{dr,rc}$ attains the semiparametric efficiency bound. On the other hand, when (A.3) and (A.7) are true, we have that

$$\begin{aligned}
\eta_1^{rc,1}(W_i;\beta^*) &= w_{1,1}^{rc}(D_i,T_i)\left[\left(Y_i-m_{0,1}^{rc}(X_i)\right)-\mathbb{E}\left[w_{1,1}^{rc}(D,T)\left(Y-m_{0,1}^{rc}(X)\right)\right]\right]\\
&\quad -w_{1,0}^{rc}(D_i,T_i)\left[\left(Y_i-m_{0,0}^{rc}(X_i)\right)-\mathbb{E}\left[w_{1,0}^{rc}(D,T)\left(Y-m_{0,0}^{rc}(X)\right)\right]\right]\\
&\neq \eta_1^{rc,2}(W_i;\beta^*),
\end{aligned}$$

and given that $\eta_0^{rc,1} = \eta_0^{rc,2}$, and $\eta_{est}^{rc,1} = \eta_{est}^{rc,2} = 0$, it follows that $\eta_1^{rc}(W_i;\gamma^*,\beta^*) \neq \eta_2^{rc}(W_i;\gamma^*,\beta^*)$ and therefore the asymptotic variance of $\widehat{\tau}_1^{dr,rc}$ does not achieve the semiparametric efficiency bound.■

Proof of Theorem 2: The doubly-robust consistency and the locally semiparametric efficiency of $\widehat{\tau}_{imp}^{dr,p}$ follows directly from Theorem A.1. Next, we establish that $\widehat{\tau}_{imp}^{dr,p}$ admits an asymptotic linear representation that is insensitive to first-step estimators, which, in turn, implies that $\widehat{\tau}_{imp}^{dr,p}$ is also doubly-robust for inference.

From Theorem A.1, it is clear that if

$$\begin{aligned}
\mathbb{E}\left[\left(w_1^p(D) - w_0^p(D,X;\gamma^*)\right)\cdot \dot{\mu}_{0,\Delta}^p\left(X;\beta_{0,1}^{*,p},\beta_{0,0}^{*,p}\right)\right] &= 0,\\
\mathbb{E}\left[\frac{(1-D)}{(1-\pi(X;\gamma^*))^2}\left(\Delta Y - \mu_{0,\Delta}^p\left(X;\beta_{0,1}^{*,p},\beta_{0,0}^{*,p}\right)\right)\cdot\dot{\pi}(X;\gamma^*)\right] &= 0,\\
\mathbb{E}\left[w_0^p(D,X;\gamma^*)\cdot\left(\Delta Y - \mu_{0,\Delta}^p\left(X;\beta_{0,1}^{*,p},\beta_{0,0}^{*,p}\right)\right)\right] &= 0,
\end{aligned}$$

then η_{est}^p is asymptotically negligible. Since the first component of X is a constant and we adopt the working models (3.5), it follows these three vectors of moment conditions reduces to

$$\mathbb{E}\left[\left(\frac{D}{\mathbb{E}[D]} - \frac{\exp(X'\gamma^*)(1-D)}{\mathbb{E}[\exp(X'\gamma^*)(1-D)]}\right)X\right] = 0, \tag{S.29}$$

$$\mathbb{E}\left[\exp(X'\gamma^*)\left(\Delta Y - X'\beta_{0,\Delta}^*\right)X\middle|D=0\right] = 0. \tag{S.30}$$

Now, notice that the first order condition associated with $\widehat{\gamma}^{ipt}$ is given by

$$\mathbb{E}_n\left[\left(D - \exp\left(X'\widehat{\gamma}^{ipt}\right)(1-D)\right)X\right] = 0.$$

From Theorem 3.1 in Graham et al. (2012), it follows that $\widehat{\gamma}^{ipt} - \gamma^{*,ipt} = O_p\left(n^{-1/2}\right)$, and therefore, by the weak law of large numbers and continuous mapping theorem, as $n\to\infty$,

$$\mathbb{E}_n\left[\left(D - \exp\left(X'\widehat{\gamma}^{ipt}\right)(1-D)\right)X\right] = \mathbb{E}\left[\left(D - \exp\left(X'\gamma^{*,ipt}\right)(1-D)\right)X\right] + o_p(1).$$

Given that the first component of X is a constant, we have that

$$\mathbb{E}_n\left[\left(D - \exp\left(X'\widehat{\gamma}^{ipt}\right)(1-D)\right)\right] = \mathbb{E}\left[\left(D - \exp\left(X'\gamma^{*,ipt}\right)(1-D)\right)\right] + o_p(1),$$

and hence (S.29) follows with $\gamma^* = \gamma^{*,ipt}$.

Similarly, the first order condition associated with $\widehat{\beta}_{0,\Delta}^{wls,p}$ is given by

$$\mathbb{E}_n\left[(1-D)\exp\left(X'\widehat{\gamma}^{ipt}\right)\left(\Delta Y - X'\widehat{\beta}_{0,\Delta}^{wls,p}\right)X\right] = 0.$$

Since $\widehat{\gamma}^{ipt} - \gamma^{*,ipt} = O_p\left(n^{-1/2}\right)$ and one can easily show that $\widehat{\beta}_{0,\Delta}^{wls,p} - \beta_{0,\Delta}^{*,wls,p} = O_p\left(n^{-1/2}\right)$, too, we have that by the weak law of large numbers and continuous mapping theorem, as $n\to\infty$,

$$\mathbb{E}_n\left[(1-D)\exp\left(X'\widehat{\gamma}^{ipt}\right)\left(\Delta Y - X'\widehat{\beta}_{0,\Delta}^{wls,p}\right)X\right] \xrightarrow{p} \mathbb{E}\left[\exp\left(X'\gamma^{*,ipt}\right)\left(\Delta Y - X'\beta_{0,\Delta}^{*,wls,p}\right)X\middle|D=0\right] = 0,$$

implying that (S.30) follows with $\gamma^* = \gamma^{*,ipt}$ and $\beta_{0,\Delta}^* = \beta_{0,\Delta}^{*,wls,p}$.

Thus, from the Theorem A.1 it now follows that

$$\sqrt{n}(\widehat{\tau}_{imp}^{dr,p} - \tau_{imp}^{dr,p}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \eta_{imp}^{dr,p}\left(W; \gamma^{*,ipt}, \beta_{0,\Delta}^{*,wls,p}, \tau_{imp}^{dr,p}\right) + o_p(1),$$

which establishes that $\widehat{\tau}_{imp}^{dr,p}$ admits an asymptotic linear representation that is insensitive to first-step estimators, which is the key property for doubly-robust inference procedures. The asymptotic normality now follows directly from the Lindeberg–Lévy central limit theorem, completing the proof. ■

Proof of Theorem 3: The doubly-robust consistency of $\widehat{\tau}_{1,imp}^{dr,rc}$ and $\widehat{\tau}_{2,imp}^{dr,rc}$, and that $\widehat{\tau}_{2,imp}^{dr,rc}$ is locally semiparametric efficient whereas $\widehat{\tau}_{1,imp}^{dr,rc}$ is not follows directly from Theorem A.2. Next, we establish that $\widehat{\tau}_{1,imp}^{dr,rc}$ and $\widehat{\tau}_{2,imp}^{dr,rc}$ admit an asymptotic linear representation that is insensitive to first-step estimators, which, in turn, implies that they are also doubly-robust for inference.

Given that $\eta_{est}^{rc,2}(W_i; \gamma^*, \beta^*) = \eta_{est}^{rc,1}(W_i; \gamma^*, \beta^*)$ a.s., we can write $\eta_{est}^{rc}(W_i; \gamma^*, \beta^*)$. Given the structure of η_{est}^{rc}, we have that, if for $t = 0, 1$,

$$\begin{aligned}
\mathbb{E}\left[\left(w_{1,t}^{rc}(D,T) - w_{0,t}^{rc}(D,T,X;\gamma^*)\right)\dot{\mu}_{0,t}^{rc}\left(X;\beta_{0,t}^{*,rc}\right)\right] &= 0, \\
\mathbb{E}\left[\frac{(1-D)\,1\{T=t\}}{(1-\pi(X;\gamma^*))^2}\left(Y - \mu_{0,t}^{rc}\left(X;\beta_{0,t}^{*,rc}\right)\right)\cdot\dot{\pi}(X;\gamma^*)\right] &= 0, \\
\mathbb{E}\left[\frac{\pi(X;\gamma^*)(1-D)\,1\{T=t\}}{(1-\pi(X;\gamma^*))}\left(Y - \mu_{0,t}^{rc}\left(X;\beta_{0,t}^{*,rc}\right)\right)\right] &= 0,
\end{aligned}$$

then η_{est}^{rc} is asymptotically negligible. Since the first component of X is a constant, and we adopt the working models (3.8), it follows that these three vector of moments reduces to

$$\begin{aligned}
\mathbb{E}\left[\left(\frac{D}{\mathbb{E}[D]} - \frac{\exp(X'\gamma^*)(1-D)}{\mathbb{E}[\exp(X'\gamma^*)(1-D)]}\right)X \middle| T=t\right] &= 0 \\
\mathbb{E}\left[\exp(X'\gamma^*)\left(Y - X'\beta_{0,t}^*\right)X \middle| D=0, T=t\right] &= 0.
\end{aligned} \tag{S.31}$$

In turn, given that (D,X) are stationary, (S.31) reduces to

$$\mathbb{E}\left[\left(\frac{D}{\mathbb{E}[D]} - \frac{\exp(X'\gamma^*)(1-D)}{\mathbb{E}[\exp(X'\gamma^*)(1-D)]}\right)X\right] = 0,$$

and the proof now follows the same steps as in the proof of Theorem 2, which are omitted to avoid repetition.■

Proof of Corollary 2: Let $p = \mathbb{E}[D]$ and $\lambda = \mathbb{E}[T]$. Under the conditions of Theorem A.2—which are implied by the conditions of Corollary 2—we can express V_2^{rc} as

$$\begin{aligned}
V_2^{rc} &= \mathbb{E}\left[\frac{DT}{p^2\lambda^2}\left(Y - m_{1,1}^{rc}(X)\right)^2 + \frac{D(1-T)}{p^2(1-\lambda)^2}\left(Y - m_{1,0}^{rc}(X)\right)^2\right] \\
&+ \mathbb{E}\left[\frac{D}{p^2}\left(m_{1,1}^{rc}(X) - m_{1,0}^{rc}(X) - m_{0,1}^{rc}(X) + m_{0,0}^{rc}(X) - \tau\right)^2\right] \\
&+ \mathbb{E}\left[\frac{(1-D)\,p(X)^2 T}{(1-p(X))^2 p^2\lambda^2}(Y - m_{0,1}^{rc}(X))^2 + \frac{(1-D)\,p(X)^2(1-T)}{(1-p(X))^2 p^2(1-\lambda)^2}(Y - m_{0,0}^{rc}(X))^2\right]
\end{aligned}$$

$$
\begin{aligned}
&= \mathbb{E}\left[\frac{1}{p\lambda}\left(Y-m^{rc}_{1,1}(X)\right)^2\middle|D=1,T=1\right]+\mathbb{E}\left[\frac{1}{p(1-\lambda)}\left(Y-m^{rc}_{1,0}(X)\right)^2\middle|D=1,T=0\right] \\
&\quad +\mathbb{E}\left[\frac{1}{p}\left(m^{rc}_{1,1}(X)-m^{rc}_{0,1}(X)-\left(m^{rc}_{1,0}(X)-m^{rc}_{0,0}(X)\right)-\tau\right)^2\middle|D=1\right] \\
&\quad +\mathbb{E}\left[\frac{(1-D)\,p(X)^2T}{(1-p(X))^2p^2\lambda^2}(Y-m^{rc}_{0,1}(X))^2+\frac{(1-D)\,p(X)^2(1-T)}{(1-p(X))^2p^2(1-\lambda)^2}(Y-m^{rc}_{0,0}(X))^2\right] \\
&= \mathbb{E}\left[\frac{1}{p\lambda}\left(Y-m^{rc}_{1,1}(X)\right)^2\middle|D=1,T=1\right]+\mathbb{E}\left[\frac{1}{p(1-\lambda)}\left(Y-m^{rc}_{1,0}(X)\right)^2\middle|D=1,T=0\right] \\
&\quad +\frac{1}{p}Var\left[m^{rc}_{1,1}(X)-m^{rc}_{0,1}(X)-\left(m^{rc}_{1,0}(X)-m^{rc}_{0,0}(X)\right)\middle|D=1\right] \\
&\quad +\mathbb{E}\left[\frac{(1-D)\,p(X)^2T}{(1-p(X))^2p^2\lambda^2}(Y-m^{rc}_{0,1}(X))^2+\frac{(1-D)\,p(X)^2(1-T)}{(1-p(X))^2p^2(1-\lambda)^2}(Y-m^{rc}_{0,0}(X))^2\right]
\end{aligned}
$$

For the V^{rc}_1, we have that

$$
\begin{aligned}
V^{rc}_1 &= \mathbb{E}\left[\frac{DT}{p^2\lambda^2}\left(Y-m^{rc}_{0,1}(X)-\mathbb{E}\left[w^{rc}_{1,1}(D,T)\left(Y-m^{rc}_{0,1}(X)\right)\right]\right)^2\right] \\
&\quad +\mathbb{E}\left[\frac{D(1-T)}{p^2(1-\lambda)^2}\left(Y-m^{rc}_{0,0}(X)-\mathbb{E}\left[w^{rc}_{1,0}(D,T)\left(Y-m^{rc}_{0,0}(X)\right)\right]\right)^2\right] \\
&\quad +\mathbb{E}\left[\frac{(1-D)\,p(X)^2T}{(1-p(X))^2p^2\lambda^2}(Y-m^{rc}_{0,1}(X))^2+\frac{(1-D)\,p(X)^2(1-T)}{(1-p(X))^2p^2(1-\lambda)^2}(Y-m^{rc}_{0,0}(X))^2\right] \\
&= \mathbb{E}\left[\frac{1}{p\lambda}\left(Y-m^{rc}_{0,1}(X)-\mathbb{E}\left[Y-m^{rc}_{0,1}(X)\middle|D=1,T=1\right]\right)^2\middle|D=1,T=1\right] \\
&\quad +\mathbb{E}\left[\frac{1}{p(1-\lambda)}\left(Y-m^{rc}_{0,0}(X)-\mathbb{E}\left[Y-m^{rc}_{0,0}(X)\middle|D=1,T=0\right]\right)^2\middle|D=1,T=0\right] \\
&\quad +\mathbb{E}\left[\frac{(1-D)\,p(X)^2T}{(1-p(X))^2p^2\lambda^2}(Y-m^{rc}_{0,1}(X))^2+\frac{(1-D)\,p(X)^2(1-T)}{(1-p(X))^2p^2(1-\lambda)^2}(Y-m^{rc}_{0,0}(X))^2\right].
\end{aligned}
$$

Now, with a bit of manipulation, repeated application of the law of iterated expectations, and the stationarity condition,

$$
\begin{aligned}
&\mathbb{E}\left[\left(Y-m^{rc}_{0,1}(X)-\mathbb{E}\left[Y-m^{rc}_{0,1}(X)\middle|D=1,T=1\right]\right)^2\middle|D=1,T=1\right] \\
&= \mathbb{E}\left[\left(Y-m^{rc}_{0,1}(X)\right)^2\middle|D=1,T=1\right]-\mathbb{E}\left[Y-m^{rc}_{0,1}(X)\middle|D=1,T=1\right]^2 \\
&= \mathbb{E}\left[\left(\left(Y-m^{rc}_{1,1}(X)\right)+\left(m^{rc}_{1,1}(X)-m^{rc}_{0,1}(X)\right)\right)^2\middle|D=1,T=1\right]-\mathbb{E}\left[m^{rc}_{1,1}(X)-m^{rc}_{0,1}(X)\middle|D=1,T=1\right]^2 \\
&= \mathbb{E}\left[\left(Y-m^{rc}_{1,1}(X)\right)^2\middle|D=1,T=1\right]+\mathbb{E}\left[\left(m^{rc}_{1,1}(X)-m^{rc}_{0,1}(X)\right)^2\middle|D=1\right] \\
&\quad -\mathbb{E}\left[m^{rc}_{1,1}(X)-m^{rc}_{0,1}(X)\middle|D=1\right]^2 \\
&= \mathbb{E}\left[\left(Y-m^{rc}_{1,1}(X)\right)^2\middle|D=1,T=1\right]+Var\left[m^{rc}_{1,1}(X)-m^{rc}_{0,1}(X)\middle|D=1\right].
\end{aligned}
$$

Analogously,

$$
\mathbb{E}\left[\left(Y-m^{rc}_{0,0}(X)-\mathbb{E}\left[Y-m^{rc}_{0,0}(X)\middle|D=1,T=0\right]\right)^2\middle|D=1,T=0\right]
$$

$$= \mathbb{E}\left[\left(Y - m^{rc}_{1,0}(X)\right)^2 \middle| D=1, T=0\right] + Var\left[m^{rc}_{1,0}(X) - m^{rc}_{0,0}(X) \middle| D=1\right],$$

implying that

$$\begin{aligned} V^{rc}_1 &= \mathbb{E}\left[\left(Y - m^{rc}_{1,1}(X)\right)^2 \middle| D=1, T=1\right] + Var\left[m^{rc}_{1,1}(X) - m^{rc}_{0,1}(X) \middle| D=1\right] \\ &+ \mathbb{E}\left[\left(Y - m^{rc}_{1,0}(X)\right)^2 \middle| D=1, T=0\right] + Var\left[m^{rc}_{1,0}(X) - m^{rc}_{0,0}(X) \middle| D=1\right] \\ &+ \mathbb{E}\left[\frac{(1-D)\,p(X)^2 T}{(1-p(X))^2 p^2 \lambda^2}(Y - m^{rc}_{0,1}(X))^2 + \frac{(1-D)\,p(X)^2\,(1-T)}{(1-p(X))^2 p^2 (1-\lambda)^2}(Y - m^{rc}_{0,0}(X))^2\right]. \end{aligned}$$

Thus, we have that

$$\begin{aligned} p \cdot (V^{rc}_1 - V^{rc}_2) &= \frac{Var\left[m^{rc}_{1,1}(X) - m^{rc}_{0,1}(X) \middle| D=1\right]}{\lambda} + \frac{Var\left[m^{rc}_{1,0}(X) - m^{rc}_{0,0}(X) \middle| D=1\right]}{1-\lambda} \\ &\quad - Var\left[m^{rc}_{1,1}(X) - m^{rc}_{0,1}(X) - \left(m^{rc}_{1,0}(X) - m^{rc}_{0,0}(X)\right) \middle| D=1\right] \\ &= Var\left[m^{rc}_{1,1}(X) - m^{rc}_{0,1}(X) \middle| D=1\right] \frac{1-\lambda}{\lambda} \\ &\quad + Var\left[m^{rc}_{1,0}(X) - m^{rc}_{0,0}(X) \middle| D=1\right] \frac{\lambda}{1-\lambda} \\ &\quad + 2Cov\left[m^{rc}_{1,1}(X) - m^{rc}_{0,1}(X), m^{rc}_{1,0}(X) - m^{rc}_{0,0}(X) \middle| D=1\right] \\ &= Var\left[\sqrt{\frac{1-\lambda}{\lambda}}\left(m^{rc}_{1,1}(X) - m^{rc}_{0,1}(X)\right) + \sqrt{\frac{\lambda}{1-\lambda}}\left(m^{rc}_{1,0}(X) - m^{rc}_{0,0}(X)\right) \middle| D=1\right] \\ &\geq 0, \end{aligned}$$

concluding the proof.■

References

Abadie, A. (2005), “Semiparametric difference-in-difference estimators,” *Review of Economic Studies*, 72, 1–19.

Chen, X., Hong, H., and Tarozzi, A. (2008), “Semiparametric efficiency in GMM models with auxiliary data,” *The Annals of Statistics*, 36(2), 808–843.

Graham, B., Pinto, C., and Egel, D. (2012), “Inverse Probability Tilting for Moment Condition Models with Missing Data,” *The Review of Economic Studies*, 79(3), 1053–1079.

Hahn, J. (1998), “On the Role of the Propensity Score in Efficient Semiparametric Estimation of Average Treatment Effects,” *Econometrica*, 66(2), 315–331.

Newey, W. K. (1990), “Semiparametric efficiency bounds,” *Journal of Applied Econometrics*, 5(2), 99–135.