Causal Inference under Temporal and Spatial Interference

(Appendix)

Appendix A Proofs

This section contains proofs for theorems presented in the main text. I start from the case where the values of nuisance parameters are known. The results are natural generalizations of those in Aronow, Samii and Wang (2020). In the last subsection, I discuss how the results would be modified to account for unknown nuisance parameters.

A1 Identification of the estimands under the sequential ignorability assumption

To account for covariates, let's assume that sequential ignorability holds after conditioning on a set $\mathbf{V}_{it}$: $\mathbf{Z}_t \perp Y_{it}(\mathbf{Z}_t, \mathbf{Z}^{1:T} \setminus \mathbf{Z}_t) | \mathbf{V}_{it}$. $\mathbf{V}_{it}$ equals to $\mathbf{X}_1$ when $t = 1$. When $t > 1$, it includes $Z_{i,t-1}$, $Y_{i,t-1}$, and possibly any other elements in the set of $(\mathbf{Z}_i^{1:(t-1)}, \mathbf{Y}_i^{1:(t-1)}, \mathbf{X}_i^{1:t})$. Denote $W_{it} = \prod_{s=1}^{t} P(Z_{is}|\mathbf{V}_{is})$ and $\mathbf{V}_i^{1:t} = \bigcup_{s=1}^{t} \mathbf{V}_{is}$. It is easy to see that $W_{it} = P(Z_{it}|\mathbf{V}_{it}) W_{i,t-1} = P(Z_{it}|\mathbf{V}_i^{1:t})$. For the average contemporaneous direct effect at period t, we can see that:

$$
\begin{aligned}
\tau_t =& \frac{1}{N}\sum_{i=1}^{N} \tau_{it} \\
=& \frac{1}{N}\sum_{i=1}^{N} \mathrm{E}_{\mathbf{Z}^{1:T}\setminus Z_{it}} \left[Y_{it}(1, \mathbf{Z}^{1:T} \setminus Z_{it}) \right] - \frac{1}{N}\sum_{i=1}^{N} \mathrm{E}_{\mathbf{Z}^{1:T}\setminus Z_{it}} \left[Y_{it}(0, \mathbf{Z}^{1:T} \setminus Z_{it}) \right] \\
=& \frac{1}{N}\sum_{i=1}^{N} \mathrm{E}_{\mathbf{Z}^{1:t}\setminus Z_{it}} \left[Y_{it}(1, \mathbf{Z}^{1:t} \setminus Z_{it}) \right] - \frac{1}{N}\sum_{i=1}^{N} \mathrm{E}_{\mathbf{Z}^{1:t}\setminus Z_{it}} \left[Y_{it}(0, \mathbf{Z}^{1:t} \setminus Z_{it}) \right] \\
=& \frac{1}{N}\sum_{i=1}^{N} \mathrm{E}_{\mathbf{V}_{it}} \left[E_{(\mathbf{Z}^{1:t}\setminus Z_{it})|\mathbf{V}_{it}} \left[Y_{it}(1, \mathbf{Z}_t \setminus Z_{it}, \mathbf{Z}^{1:(t-1)}) | \mathbf{V}_{it} \right] \right] \\
& - \frac{1}{N}\sum_{i=1}^{N} \mathrm{E}_{\mathbf{V}_{it}} \left[E_{(\mathbf{Z}^{1:t}\setminus Z_{it})|\mathbf{V}_{it}} \left[Y_{it}(0, \mathbf{Z}_t \setminus Z_{it}, \mathbf{Z}^{1:(t-1)}) | \mathbf{V}_{it} \right] \right] \\
=& \frac{1}{N}\sum_{i=1}^{N} \mathrm{E}_{\mathbf{V}_{it}} \left[E_{\mathbf{Z}^{1:t}|\mathbf{V}_{it}} \left[\frac{Z_{it} Y_{it}(\mathbf{Z}^{1:t})}{P(Z_{it}=1)} | \mathbf{V}_{it} \right] \right] - \frac{1}{N}\sum_{i=1}^{N} \mathrm{E}_{\mathbf{V}_{it}} \left[E_{\mathbf{Z}^{1:t}|\mathbf{V}_{it}} \left[\frac{(1-Z_{it}) Y_{it}(\mathbf{Z}^{1:t})}{1-P(Z_{it}=1)} | \mathbf{V}_{it} \right] \right] \\
=& \frac{1}{N}\sum_{i=1}^{N} \mathrm{E}_{\mathbf{Z}^{1:T}} \left[\frac{Z_{it} Y_{it}(\mathbf{Z}^{1:T})}{P(Z_{it}=1)} \right] - \frac{1}{N}\sum_{i=1}^{N} \mathrm{E}_{\mathbf{Z}^{1:T}} \left[\frac{(1-Z_{it}) Y_{it}(\mathbf{Z}^{1:T})}{1-P(Z_{it}=1)} \right]
\end{aligned}
.
$$

The third and the last equalities use the no reverse causality assumption (Assumption 1) and the fifth equality uses the sequential ignorability assumption (Assumption 2).

For the average cumulative indirect effect at period t and distance d, the intuition is the same. I prove the result sequentially and start from period s:

$$
\begin{aligned}
&\tau_t(\mathbf{z}^{s:t}, \tilde{\mathbf{z}}^{s:t}; d) = \frac{1}{N}\sum_{i=1}^{N} \mu_i(\tau_{(1:N),t}(\mathbf{z}^{s:t}, \tilde{\mathbf{z}}^{s:t}); d) \\
=&\frac{1}{N}\sum_{i=1}^{N} \mathrm{E}_{\mathbf{Z}^{1:T}\setminus \mathbf{Z}_i^{s:t}}\left[\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t}, \mathbf{Z}^{1:T}\setminus \mathbf{Z}_i^{s:t}); d)\right] - \frac{1}{N}\sum_{i=1}^{N} \mathrm{E}_{\mathbf{Z}^{1:T}\setminus \mathbf{Z}_i^{s:t}}\left[\mu_i(\mathbf{Y}_t(\tilde{\mathbf{z}}^{s:t}, \mathbf{Z}^{1:T}\setminus \mathbf{Z}_i^{s:t}); d)\right] \\
=&\frac{1}{N}\sum_{i=1}^{N} \mathrm{E}_{\mathbf{Z}^{1:t}\setminus \mathbf{Z}_i^{s:t}}\left[\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t}, \mathbf{Z}^{1:t}\setminus \mathbf{Z}_i^{s:t}); d)\right] - \frac{1}{N}\sum_{i=1}^{N} \mathrm{E}_{\mathbf{Z}^{1:t}\setminus \mathbf{Z}_i^{s:t}}\left[\mu_i(\mathbf{Y}_t(\tilde{\mathbf{z}}^{s:t}, \mathbf{Z}^{1:t}\setminus \mathbf{Z}_i^{s:t}); d)\right] \\
=&\frac{1}{N}\sum_{i=1}^{N} \mathrm{E}_{\mathbf{V}_{is}}\left[\mathrm{E}_{\mathbf{Z}^{1:t}\setminus \mathbf{Z}_i^{s:t}|\mathbf{V}_{is}}\left[\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t}, \mathbf{Z}^{1:t}\setminus \mathbf{Z}_i^{s:t}); d)|\mathbf{V}_{is}\right]\right] \\
&- \frac{1}{N}\sum_{i=1}^{N} \mathrm{E}_{\mathbf{V}_{is}}\left[\mathrm{E}_{\mathbf{Z}^{1:t}\setminus \mathbf{Z}_i^{s:t}|\mathbf{V}_{is}}\left[\mu_i(\mathbf{Y}_t(\tilde{\mathbf{z}}^{s:t}, \mathbf{Z}^{1:t}\setminus \mathbf{Z}_i^{s:t}); d)|\mathbf{V}_{is}\right]\right] \\
=&\frac{1}{N}\sum_{i=1}^{N} \mathrm{E}_{\mathbf{V}_{is}}\left[\mathrm{E}_{\mathbf{Z}^{1:t}\setminus \mathbf{Z}_i^{(s+1):t}|\mathbf{V}_{is}}\left[\frac{\mathbf{1}\{Z_{is} = z_s\}\mu_i(\mathbf{Y}_t(\mathbf{z}^{(s+1):t}, \mathbf{Z}^{1:t}\setminus \mathbf{Z}_i^{(s+1):t}); d)}{P(Z_{is} = z_s|\mathbf{V}_{is})}|\mathbf{V}_{is}\right]\right] \\
&- \frac{1}{N}\sum_{i=1}^{N} \mathrm{E}_{\mathbf{V}_{is}}\left[\mathrm{E}_{\mathbf{Z}^{1:t}\setminus \mathbf{Z}_i^{(s+1):t}|\mathbf{V}_{is}}\left[\frac{\mathbf{1}\{Z_{is} = \tilde{z}_s\}\mu_i(\mathbf{Y}_t(\mathbf{z}^{(s+1):t}, \mathbf{Z}^{1:t}\setminus \mathbf{Z}_i^{(s+1):t}); d)}{P(Z_{is} = \tilde{z}_s|\mathbf{V}_{is})}|\mathbf{V}_{is}\right]\right] \\
=&\frac{1}{N}\sum_{i=1}^{N} \mathrm{E}_{\mathbf{Z}^{1:t}\setminus \mathbf{Z}_i^{(s+1):t}}\left[\frac{\mathbf{1}\{Z_{is} = z_s\}\mu_i(\mathbf{Y}_t(\mathbf{z}^{(s+1):t}, \mathbf{Z}^{1:t}\setminus \mathbf{Z}_i^{(s+1):t}); d)}{P(Z_{is} = z_s|\mathbf{V}_{is})}\right] \\
&- \frac{1}{N}\sum_{i=1}^{N} \mathrm{E}_{\mathbf{Z}^{1:t}\setminus \mathbf{Z}_i^{(s+1):t}}\left[\frac{\mathbf{1}\{Z_{is} = \tilde{z}_s\}\mu_i(\mathbf{Y}_t(\mathbf{z}^{(s+1):t}, \mathbf{Z}^{1:t}\setminus \mathbf{Z}_i^{(s+1):t}); d)}{P(Z_{is} = \tilde{z}_s|\mathbf{V}_{is})}\right] \\
=&\frac{1}{N}\sum_{i=1}^{N} \mathrm{E}_{\mathbf{V}_{i,s+1}}\left[\mathrm{E}_{\mathbf{Z}^{1:t}\setminus \mathbf{Z}_i^{(s+1):t}|\mathbf{V}_{i,s+1}}\left[\frac{\mathbf{1}\{Z_{is} = z_s\}\mu_i(\mathbf{Y}_t(\mathbf{z}^{(s+1):t}, \mathbf{Z}^{1:t}\setminus \mathbf{Z}_i^{(s+1):t}); d)}{P(Z_{is} = z_s|\mathbf{V}_{is})}|\mathbf{V}_{i,s+1}\right]\right] \\
&- \frac{1}{N}\sum_{i=1}^{N} \mathrm{E}_{\mathbf{V}_{i,s+1}}\left[\mathrm{E}_{\mathbf{Z}^{1:t}\setminus \mathbf{Z}_i^{(s+1):t}|\mathbf{V}_{i,s+1}}\left[\frac{\mathbf{1}\{Z_{is} = \tilde{z}_s\}\mu_i(\mathbf{Y}_t(\mathbf{z}^{(s+1):t}, \mathbf{Z}^{1:t}\setminus \mathbf{Z}_i^{(s+1):t}); d)}{P(Z_{is} = \tilde{z}_s|\mathbf{V}_{is})}|\mathbf{V}_{i,s+1}\right]\right] \\
=&\frac{1}{N}\sum_{i=1}^{N} \mathrm{E}_{\mathbf{Z}^{1:t}\setminus \mathbf{Z}_i^{(s+2):t}}\left[\frac{\mathbf{1}\{Z_{is} = z_s\}\mathbf{1}\{Z_{i,s+1} = z_{s+1}\}\mu_i(\mathbf{Y}_t(\mathbf{z}^{(s+2):t}, \mathbf{Z}^{1:t}\setminus \mathbf{Z}_i^{(s+2):t}); d)}{P(Z_{is} = z_s|\mathbf{V}_{is})P(Z_{i,s+1} = z_{s+1}|\mathbf{V}_{i,s+1})}\right] \\
&- \frac{1}{N}\sum_{i=1}^{N} \mathrm{E}_{\mathbf{Z}^{1:t}\setminus \mathbf{Z}_i^{(s+2):t}}\left[\frac{\mathbf{1}\{Z_{is} = \tilde{z}_s\}\mathbf{1}\{Z_{i,s+1} = \tilde{z}_{s+1}\}\mu_i(\mathbf{Y}_t(\mathbf{z}^{(s+2):t}, \mathbf{Z}^{1:t}\setminus \mathbf{Z}_i^{(s+2):t}); d)}{P(Z_{is} = \tilde{z}_s|\mathbf{V}_{is})P(Z_{i,s+1} = \tilde{z}_{s+1}|\mathbf{V}_{i,s+1})}\right] \\
=&\dots \\
=&\frac{1}{N}\sum_{i=1}^{N} \mathrm{E}_{\mathbf{Z}^{1:T}}\left[\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t} = \mathbf{z}^{s:t}\}\mu_i(\mathbf{Y}_t(\mathbf{Z}^{1:T}); d)}{W_{it}}\right] - \frac{1}{N}\sum_{i=1}^{N} \mathrm{E}_{\mathbf{Z}^{1:T}}\left[\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t} = \tilde{\mathbf{z}}^{s:t}\}\mu_i(\mathbf{Y}_t(\mathbf{Z}^{1:T}); d)}{\tilde{W}_{it}}\right]
\end{aligned}
$$

The $\dots$ part iterates the same steps for periods $s+2$ to t. $W_{it} = P(\mathbf{Z}_i^{s:t} = \mathbf{z}^{s:t}) = \prod_{t'=s}^{t} P(Z_{it'} = z_{t'}|\mathbf{V}_{it'})$ is the probability for history $\mathbf{z}^{s:t}$ to occur. As the average cumulative indirect effect includes all the other three estimands as its special cases, Theorem 1 is proved.

A2 Variance of the IPTW estimators

I first derive the variance of the Horvitz-Thompson estimators with known propensity scores for the average cumulative indirect effect and its average over periods. I denote the estimator for the effect at period t and its average over periods as $\hat{\tau}_t^*(\mathbf{z}^{s:t}, \tilde{\mathbf{z}}^{s:t}; d)$ and $\hat{\tau}^*(\mathbf{z}^{s:t}, \tilde{\mathbf{z}}^{s:t}; d)$, respectively. The results are summarized by the following lemma:

Lemma 1. *Under Assumptions 1-6 in the main text, the variance of the Horvitz-Thompson estimator* $\hat{\tau}_t^*(\mathbf{z}^{s:t}, \tilde{\mathbf{z}}^{s:t}; d)$ *is bounded as follows when* *the true values of the propensity scores are known:*

$$
\begin{aligned}
&\operatorname{Var}\left(\hat{\tau}_t^*(\mathbf{z}^{s:t},\tilde{\mathbf{z}}^{s:t};d)\right)\\
\le&\frac{1}{N^2}\sum_{i=1}^{N}\frac{\mathrm{E}\left[\mu_i^2(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)\right]}{W_{it}}+\frac{1}{N^2}\sum_{i=1}^{N}\frac{\mathrm{E}\left[\mu_i^2(\mathbf{Y}_t(\tilde{\mathbf{z}}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)\right]}{\tilde{W}_{it}}\\
&+\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j\in\mathcal{B}(i;d)}\sum_{\mathbf{a},\mathbf{b}=\mathbf{z}^{s:t}}^{\tilde{\mathbf{z}}^{s:t}}(-1)^{\mathbf{1}\{\mathbf{a}=\mathbf{b}\}}\operatorname{Cov}\left[\mu_i(\mathbf{Y}_t(\mathbf{a},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d),\mu_j(\mathbf{Y}_t(\mathbf{b},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_j^{s:t});d)\right].
\end{aligned}
$$

And the variance of the estimator $\hat{\tau}^*(\mathbf{z}^{s:t},\tilde{\mathbf{z}}^{s:t};d)$ is bounded as:

$$
\begin{aligned}
&\operatorname{Var}\left(\hat{\tau}^*(\mathbf{z}^{s:t},\tilde{\mathbf{z}}^{s:t};d)\right)\\
\le&\frac{1}{N^2(t-s+1)^2}\sum_{t'=s}^{t}\sum_{i=1}^{N}\frac{\mathrm{E}\left[\mu_i^2(\mathbf{Y}_{t'}(\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)\right]}{W_{it'}}+\frac{1}{N^2(t-s+1)^2}\sum_{t'=s}^{t}\sum_{i=1}^{N}\frac{\mathrm{E}\left[\mu_i^2(\mathbf{Y}_{t'}(\tilde{\mathbf{z}}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)\right]}{\tilde{W}_{it'}}\\
&+\frac{1}{N^2(t-s+1)^2}\sum_{t'=s}^{t}\sum_{i=1}^{N}\sum_{(j,s)\in\mathcal{B}((i,t);d)}\sum_{\mathbf{a},\mathbf{b}=\mathbf{z}^{s:t}}^{\tilde{\mathbf{z}}^{s:t}}(-1)^{\mathbf{1}\{\mathbf{a}=\mathbf{b}\}}\operatorname{Cov}\left[\mu_i(\mathbf{Y}_{t'}(\mathbf{a},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d),\mu_j(\mathbf{Y}_s(\mathbf{b},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_j^{s:t});d)\right].
\end{aligned}
$$

Proof. Using the expression of the Horvitz-Thompson estimator, we have:

$$
\begin{aligned}
&\operatorname{Var}\left(\hat{\tau}_t^*(\mathbf{z}^{s:t},\tilde{\mathbf{z}}^{s:t};d)\right)\\
=&\frac{1}{N^2}\operatorname{Var}\left[\sum_{i=1}^{N}\left(\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\mathbf{z}^{s:t}\}}{W_{it}}-\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\tilde{\mathbf{z}}^{s:t}\}}{\tilde{W}_{it}}\right)\mu_i(\mathbf{Y}_t(\mathbf{Z}^{1:T});d)\right]\\
=&\frac{1}{N^2}\sum_{i=1}^{N}\operatorname{Var}\left[\left(\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\mathbf{z}^{s:t}\}}{W_{it}}-\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\tilde{\mathbf{z}}^{s:t}\}}{\tilde{W}_{it}}\right)\mu_i(\mathbf{Y}_t(\mathbf{Z}^{1:T});d)\right]\\
&+\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j\neq i}\operatorname{Cov}\left[\left(\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\mathbf{z}^{s:t}\}}{W_{it}}-\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\tilde{\mathbf{z}}^{s:t}\}}{\tilde{W}_{it}}\right)\mu_i(\mathbf{Y}_t(\mathbf{Z}^{1:T});d),\left(\frac{\mathbf{1}\{\mathbf{Z}_j^{s:t}=\mathbf{z}^{s:t}\}}{W_{jt}}-\frac{\mathbf{1}\{\mathbf{Z}_j^{s:t}=\tilde{\mathbf{z}}^{s:t}\}}{\tilde{W}_{jt}}\right)\mu_j(\mathbf{Y}_t(\mathbf{Z}^{1:T});d)\right]\\
=&\frac{1}{N^2}\sum_{i=1}^{N}\mathrm{E}\left[\left(\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\mathbf{z}^{s:t}\}}{W_{it}}-\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\tilde{\mathbf{z}}^{s:t}\}}{\tilde{W}_{it}}\right)\mu_i(\mathbf{Y}_t(\mathbf{Z}^{1:T});d)\right]^2\\
&-\frac{1}{N^2}\sum_{i=1}^{N}\mathrm{E}^2\left[\left(\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\mathbf{z}^{s:t}\}}{W_{it}}-\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\tilde{\mathbf{z}}^{s:t}\}}{\tilde{W}_{it}}\right)\mu_i(\mathbf{Y}_t(\mathbf{Z}^{1:T});d)\right]\\
&+\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j\neq i}\operatorname{Cov}\left[\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\mathbf{z}^{s:t}\}}{W_{it}}\mu_i(\mathbf{Y}_t(\mathbf{Z}^{1:T});d),\frac{\mathbf{1}\{\mathbf{Z}_j^{s:t}=\mathbf{z}^{s:t}\}}{W_{jt}}\mu_j(\mathbf{Y}_t(\mathbf{Z}^{1:T});d)\right]\\
&-\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j\neq i}\operatorname{Cov}\left[\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\mathbf{z}^{s:t}\}}{W_{it}}\mu_i(\mathbf{Y}_t(\mathbf{Z}^{1:T});d),\frac{\mathbf{1}\{\mathbf{Z}_j^{s:t}=\tilde{\mathbf{z}}^{s:t}\}}{\tilde{W}_{jt}}\mu_j(\mathbf{Y}_t(\mathbf{Z}^{1:T});d)\right]\\
&-\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j\neq i}\operatorname{Cov}\left[\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\tilde{\mathbf{z}}^{s:t}\}}{\tilde{W}_{it}}\mu_i(\mathbf{Y}_t(\mathbf{Z}^{1:T});d),\frac{\mathbf{1}\{\mathbf{Z}_j^{s:t}=\mathbf{z}^{s:t}\}}{W_{jt}}\mu_j(\mathbf{Y}_t(\mathbf{Z}^{1:T});d)\right]\\
&+\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j\neq i}\operatorname{Cov}\left[\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\tilde{\mathbf{z}}^{s:t}\}}{\tilde{W}_{it}}\mu_i(\mathbf{Y}_t(\mathbf{Z}^{1:T});d),\frac{\mathbf{1}\{\mathbf{Z}_j^{s:t}=\tilde{\mathbf{z}}^{s:t}\}}{\tilde{W}_{jt}}\mu_j(\mathbf{Y}_t(\mathbf{Z}^{1:T});d)\right].
\end{aligned}
$$

The first two terms in the above expression can be further expanded as:

$$
\begin{aligned}
&\frac{1}{N^2}\sum_{i=1}^{N}\mathrm{E}\left[\left(\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\mathbf{z}^{s:t}\}}{W_{it}}-\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\tilde{\mathbf{z}}^{s:t}\}}{\tilde{W}_{it}}\right)\mu_i(\mathbf{Y}_t(\mathbf{Z}^{1:T});d)\right]^2-\frac{1}{N^2}\sum_{i=1}^{N}\mathrm{E}^2\left[\left(\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\mathbf{z}^{s:t}\}}{W_{it}}-\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\tilde{\mathbf{z}}^{s:t}\}}{\tilde{W}_{it}}\right)\mu_i(\mathbf{Y}_t(\mathbf{Z}^{1:T});d)\right]\\
=&\frac{1}{N^2}\sum_{i=1}^{N}\mathrm{E}\left[\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\mathbf{z}^{s:t}\}}{W_{it}^2}\mu_i^2(\mathbf{Y}_t(\mathbf{Z}^{1:T});d)\right]+\frac{1}{N^2}\sum_{i=1}^{N}\mathrm{E}\left[\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\tilde{\mathbf{z}}^{s:t}\}}{\tilde{W}_{it}^2}\mu_i^2(\mathbf{Y}_t(\mathbf{Z}^{1:T});d)\right]\\
&-\frac{1}{N^2}\sum_{i=1}^{N}\mathrm{E}^2\left[\left(\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\mathbf{z}^{s:t}\}}{W_{it}}-\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\tilde{\mathbf{z}}^{s:t}\}}{\tilde{W}_{it}}\right)\mu_i(\mathbf{Y}_t(\mathbf{Z}^{1:T});d)\right]\\
=&\frac{1}{N^2}\sum_{i=1}^{N}\frac{\mathrm{E}\left[\mu_i^2(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)\right]}{W_{it}}+\frac{1}{N^2}\sum_{i=1}^{N}\frac{\mathrm{E}\left[\mu_i^2(\mathbf{Y}_t(\tilde{\mathbf{z}}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)\right]}{\tilde{W}_{it}}\\
&-\frac{1}{N^2}\sum_{i=1}^{N}\mathrm{E}^2\left[\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)-\mu_i(\mathbf{Y}_t(\tilde{\mathbf{z}}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)\right]\\
\leq&\frac{1}{N^2}\sum_{i=1}^{N}\frac{\mathrm{E}\left[\mu_i^2(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)\right]}{W_{it}}+\frac{1}{N^2}\sum_{i=1}^{N}\frac{\mathrm{E}\left[\mu_i^2(\mathbf{Y}_t(\tilde{\mathbf{z}}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)\right]}{\tilde{W}_{it}}.
\end{aligned}
$$

And the first covariance term equals to:

$$
\begin{aligned}
&\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j\neq i}\mathrm{Cov}\left[\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\mathbf{z}^{s:t}\}}{W_{it}}\mu_i(\mathbf{Y}_t(\mathbf{Z}^{1:T});d),\frac{\mathbf{1}\{\mathbf{Z}_j^{s:t}=\mathbf{z}^{s:t}\}}{W_{jt}}\mu_j(\mathbf{Y}_t(\mathbf{Z}^{1:T});d)\right]\\
=&\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j\neq i}\mathrm{E}\left[\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\mathbf{z}^{s:t}\}\mathbf{1}\{\mathbf{Z}_j^{s:t}=\mathbf{z}^{s:t}\}}{W_{it}W_{jt}}\mu_i(\mathbf{Y}_t(\mathbf{Z}^{1:T});d)\mu_j(\mathbf{Y}_t(\mathbf{Z}^{1:T});d)\right]\\
&-\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j\neq i}\mathrm{E}\left[\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\mathbf{z}^{s:t}\}}{W_{it}}\mu_i(\mathbf{Y}_t(\mathbf{Z}^{1:T});d)\right]\mathrm{E}\left[\frac{\mathbf{1}\{\mathbf{Z}_j^{s:t}=\mathbf{z}^{s:t}\}}{W_{jt}}\mu_j(\mathbf{Y}_t(\mathbf{Z}^{1:T});d)\right]\\
=&\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j\neq i}\mathrm{E}\left[\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_{i,j}^{s:t});d)\mu_j(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_{i,j}^{s:t});d)\right]\\
&-\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j\neq i}\mathrm{E}\left[\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)\right]\mathrm{E}\left[\mu_j(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_j^{s:t});d)\right]\\
=&\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j\neq i}\mathrm{Cov}\left[\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d),\mu_j(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_j^{s:t});d)\right]\\
=&\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j\in\mathcal{B}(i;d)}\mathrm{Cov}\left[\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d),\mu_j(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_j^{s:t});d)\right].
\end{aligned}
$$

The penultimate equality is due to the definition of covariance, and the last equality uses the definition of $\mathcal{B}(i;d)$. Other covariance terms have a similar form. We obtain the final result by combining these terms together. The derivation of $\hat{\tau}(\mathbf{z}^{s:t},\tilde{\mathbf{z}}^{s:t};d)$'s variance proceeds similarly thus is omitted. The only difference is that the expression includes the covariance of observations caused by temporal interference. □

Next, I derive the limiting variance of the Hajek estimator with known propensity scores using linearization. The estimator for the effect at period t and its average over periods are denoted as $\hat{\tau}^*_{t,HA}(\mathbf{z}^{s:t},\tilde{\mathbf{z}}^{s:t};d)$ and $\hat{\tau}^*_{HA}(\mathbf{z}^{s:t},\tilde{\mathbf{z}}^{s:t};d)$, respectively.

Lemma 2. *Define*

$$
\begin{aligned}
\tilde{V}_d=&\frac{1}{N^2}\sum_{i=1}^{N}\frac{\mathrm{E}\left[\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)-\bar{\mu}_t^1\right]^2}{W_{it}}+\frac{1}{N^2}\sum_{i=1}^{N}\frac{\mathrm{E}\left[\mu_i(\mathbf{Y}_t(\tilde{\mathbf{z}}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)-\bar{\mu}_t^0\right]^2}{\tilde{W}_{it}}\\
&+\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j\in\mathcal{B}(i;d)}\sum_{\mathbf{a},\mathbf{b}=\mathbf{z}^{s:t}}^{\tilde{\mathbf{z}}^{s:t}}(-1)^{\mathbf{1}\{\mathbf{a}=\mathbf{b}\}}\mathrm{Cov}\left[\mu_i(\mathbf{Y}_t(\mathbf{a},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)-\bar{\mu}_t^{\mathbf{1}\{\mathbf{a}=\mathbf{z}^{s:t}\}},\mu_j(\mathbf{Y}_t(\mathbf{b},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_j^{s:t});d)-\bar{\mu}_t^{\mathbf{1}\{\mathbf{b}=\mathbf{z}^{s:t}\}}\right].
\end{aligned}
$$

and

$$\tilde{V}_d^* = \frac{1}{N^2(t-s+1)^2}\sum_{t'=s}^{t}\sum_{i=1}^{N}\frac{\mathrm{E}\left[\mu_i(\mathbf{Y}_{t'}(\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)-\bar{\mu}^1\right]^2}{W_{it'}} + \frac{1}{N^2(t-s+1)^2}\sum_{t'=s}^{t}\sum_{i=1}^{N}\frac{\mathrm{E}\left[\mu_i(\mathbf{Y}_{t'}(\tilde{\mathbf{z}}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)-\bar{\mu}^0\right]^2}{\tilde{W}_{it'}}$$
$$+\frac{1}{N^2(t-s+1)^2}\sum_{t'=s}^{t}\sum_{i=1}^{N}\sum_{(j,s)\in\mathcal{B}((i,t);d)}\sum_{\mathbf{a},\mathbf{b}=\mathbf{z}^{s:t}}^{\tilde{\mathbf{z}}^{s:t}}(-1)^{\mathbf{1}\{\mathbf{a}=\mathbf{b}\}}\mathrm{Cov}\left[\mu_i(\mathbf{Y}_{t'}(\mathbf{a},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)-\bar{\mu}^{\mathbf{1}\{\mathbf{a}=\mathbf{z}^{s:t}\}},\mu_j(\mathbf{Y}_s(\mathbf{b},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_j^{s:t});d)-\bar{\mu}^{\mathbf{1}\{\mathbf{b}=\mathbf{z}^{s:t}\}}\right].$$

where $\bar{\mu}_t^1 = \frac{1}{N}\sum_{i=1}^N \mathrm{E}\left[\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)\right]$, $\bar{\mu}_t^0 = \frac{1}{N}\sum_{i=1}^N \mathrm{E}\left[\mu_i(\mathbf{Y}_t(\tilde{\mathbf{z}}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)\right]$, $\bar{\mu}^1 = \frac{1}{N(t-s+1)}\sum_{t'=s}^{t}\sum_{i=1}^N \mathrm{E}\left[\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)\right]$, *and* $\bar{\mu}^0 = \frac{1}{N(t-s+1)}\sum_{t'=s}^{t}\sum_{i=1}^N \mathrm{E}\left[\mu_i(\mathbf{Y}_t(\tilde{\mathbf{z}}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)\right]$. *Under Assumptions 1-6 in the main text, as* $N\to\infty$, *we have:*

$$\Pr\left(\frac{N}{\tilde{b}}\mathrm{Var}\left(\hat{\tau}^*_{HA,t}(d)\right)\le\frac{N}{\tilde{b}}\tilde{V}_d\right)\to 1.$$

and as $N*T\to\infty$, *we have:*

$$\Pr\left(\frac{NT}{\tilde{b}^*}\mathrm{Var}\left(\hat{\tau}^*_{HA}(d)\right)\le\frac{NT}{\tilde{b}^*}\tilde{V}_d^*\right)\to 1.$$

where $1\le\tilde{b}\le\max_{i\in\{1,2,\ldots,N\}}b_{i;d}$ *and* $1\le\tilde{b}^*\le\max_{i\in\{1,2,\ldots,N\},t\in\{s,(s+1),\ldots,t\}}b_{(i,t);d}$

Proof. Denote $\frac{1}{N}\sum_{i=1}^N\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\mathbf{z}^{s:t}\}\mu_i(\mathbf{Y}_t(\mathbf{Z}^{1:t});d)}{W_{it}}$ as $\tilde{\mu}_t^1$, $\frac{1}{N}\sum_{i=1}^N\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\tilde{\mathbf{z}}^{s:t}\}\mu_i(\mathbf{Y}_t(\mathbf{Z}^{1:t});d)}{\tilde{W}_{it}}$ as $\tilde{\mu}_t^0$, $\frac{1}{N}\sum_{i=1}^N\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\mathbf{z}^{s:t}\}}{W_{it}}$ as $\tilde{N}_1$, $\frac{1}{N}\sum_{i=1}^N\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\tilde{\mathbf{z}}^{s:t}\}}{\tilde{W}_{it}}$ as $\tilde{N}_0$, and $\mathbf{W}=(\tilde{\mu}^1(d),\tilde{\mu}^0(d),\tilde{N}_1,\tilde{N}_0)$. Define $f(a,b,c,d)=\frac{a}{c}-\frac{b}{d}$. Then the Hajek estimator can be written as $f(\tilde{\mu}^1(d),\tilde{\mu}^0(d),\tilde{N}_1,\tilde{N}_0)=f(\mathbf{W})$. We know that $\mathrm{E}\left[\tilde{\mu}^1(d)\right]=\bar{\mu}^1(d)$, $\mathrm{E}\left[\tilde{\mu}^0(d)\right]=\bar{\mu}^0(d)$, $\mathrm{E}\left[\tilde{N}_1\right]=\mathrm{E}\left[\tilde{N}_0\right]=1$. Thus, $\mathrm{E}\left[\mathbf{W}\right]=(\bar{\mu}^1(d),\bar{\mu}^0(d)),1,1)$.

Using Taylor expansion, we have:

$$\sqrt{\frac{N}{\tilde{b}}}\hat{\tau}^*_{t,HA}(d)=f(\tilde{\mu}^1(d),\tilde{\mu}^0(d),\tilde{N}_1,\tilde{N}_0)=\sqrt{\frac{N}{\tilde{b}}}f(\mathbf{W})$$
$$=\sqrt{\frac{N}{\tilde{b}}}f(\mathrm{E}\left[\mathbf{W}\right])+\sqrt{\frac{N}{\tilde{b}}}\frac{\partial f}{\partial\mathbf{W}}|_{\mathrm{E}\left[\mathbf{W}\right]}(\mathbf{W}-\mathrm{E}\left[\mathbf{W}\right])+o_P(\sqrt{\frac{N}{\tilde{b}}}||\mathbf{W}-\mathrm{E}\left[\mathbf{W}\right]||).$$

We know that $\sqrt{\frac{N}{\tilde{b}}}||\mathbf{W}-\mathrm{E}\left[\mathbf{W}\right]||$ is bounded from Lemma 1. Hence, $o_P(\sqrt{\frac{N}{\tilde{b}}}||\mathbf{W}-\mathrm{E}\left[\mathbf{W}\right]||)=o_P(1)$, and,

$$\begin{aligned}&\frac{N}{\tilde{b}}\mathrm{Var}\left(\hat{\tau}^*_{t,HA}(d)\right)\\ =&\frac{N}{\tilde{b}}\mathrm{E}\left[f(\mathbf{W})-f(\mathrm{E}\left[\mathbf{W}\right])\right]^2\\ =&\frac{N}{\tilde{b}}\left(\frac{\partial f}{\partial\mathbf{W}}|_{\mathrm{E}\left[\mathbf{W}\right]}\right)'\mathrm{E}\left[\mathbf{W}-\mathrm{E}\left[\mathbf{W}\right]\right]^2\frac{\partial f}{\partial\mathbf{W}}|_{\mathrm{E}\left[\mathbf{W}\right]}+o_P(1)\\ =&\frac{N}{\tilde{b}}\left(\frac{\partial f}{\partial\mathbf{W}}|_{\mathrm{E}\left[\mathbf{W}\right]}\right)'\mathrm{Var}\left[\mathbf{W}\right]\frac{\partial f}{\partial\mathbf{W}}|_{\mathrm{E}\left[\mathbf{W}\right]}+o_P(1).\end{aligned}$$

It is easy to see that $\frac{\partial f}{\partial\mathbf{W}}|_{\mathrm{E}\left[\mathbf{W}\right]}=(1,-1,-\bar{\mu}^1(d),\bar{\mu}^0(d))'$, and

$$\mathrm{Var}\left[\mathbf{W}\right]=$$
$$\begin{pmatrix}\mathrm{Var}\left[\tilde{\mu}^1(d)\right] & \mathrm{Cov}\left[\tilde{\mu}^1(d),\tilde{\mu}^0(d)\right] & \mathrm{Cov}\left[\tilde{\mu}^1(d),\tilde{N}_1\right] & \mathrm{Cov}\left[\tilde{\mu}^1(d),\tilde{N}_0\right]\\ \mathrm{Cov}\left[\tilde{\mu}^1(d),\tilde{\mu}^0(d)\right] & \mathrm{Var}\left[\tilde{\mu}^0(d)\right] & \mathrm{Cov}\left[\tilde{\mu}^0(d),\tilde{N}_1\right] & \mathrm{Cov}\left[\tilde{\mu}^0(d),\tilde{N}_0\right]\\ \mathrm{Cov}\left[\tilde{\mu}^1(d),\tilde{N}_1\right] & \mathrm{Cov}\left[\tilde{\mu}^0(d),\tilde{N}_1\right] & \mathrm{Var}\left[\tilde{N}_1\right] & \mathrm{Cov}\left[\tilde{N}_1,\tilde{N}_0\right]\\ \mathrm{Cov}\left[\tilde{\mu}^1(d),\tilde{N}_0\right] & \mathrm{Cov}\left[\tilde{\mu}^0(d),\tilde{N}_0\right] & \mathrm{Cov}\left[\tilde{N}_1,\tilde{N}_0\right] & \mathrm{Var}\left[\tilde{N}_0\right]\end{pmatrix}$$

Therefore,

$$
\begin{aligned}
&\frac{N}{\tilde{b}}\left(\frac{\partial f}{\partial \mathbf{W}}|_{\mathrm{E}[\mathbf{W}]}\right)'\mathrm{Var}\left[\mathbf{W}\right]\frac{\partial f}{\partial \mathbf{W}}|_{\mathrm{E}[\mathbf{W}]}\\
=&\frac{N}{\tilde{b}}\mathrm{Var}\left[\tilde{\mu}^1(d)\right]+\frac{N}{\tilde{b}}\mathrm{Var}\left[\tilde{\mu}^0(d)\right]+\frac{N}{\tilde{b}}\left(\bar{\mu}^1(d)\right)^2\mathrm{Var}\left[\tilde{N}_1\right]+\frac{N}{\tilde{b}}\left(\bar{\mu}^0(d)\right)^2\mathrm{Var}\left[\tilde{N}_0\right]\\
&-2\frac{N}{\tilde{b}}\mathrm{Cov}\left[\tilde{\mu}^1(d),\tilde{\mu}^0(d)\right]-2\frac{N}{\tilde{b}}\bar{\mu}^1(d)\mathrm{Cov}\left[\tilde{\mu}^1(d),\tilde{N}_1\right]+2\frac{N}{\tilde{b}}\bar{\mu}^0(d)\mathrm{Cov}\left[\tilde{\mu}^1(d),\tilde{N}_0\right]\\
&+2\frac{N}{\tilde{b}}\bar{\mu}^1(d)\mathrm{Cov}\left[\tilde{\mu}^0(d),\tilde{N}_1\right]-2\frac{N}{\tilde{b}}\bar{\mu}^0(d)\mathrm{Cov}\left[\tilde{\mu}^0(d),\tilde{N}_0\right]-2\frac{N}{\tilde{b}}\left(\bar{\mu}^1(d)\bar{\mu}^0(d)\right)\mathrm{Cov}\left[\tilde{N}_1,\tilde{N}_0\right]\\
=&\frac{N}{\tilde{b}}\mathrm{Var}\left[\tilde{\mu}^1(d)-\bar{\mu}^1(d)\tilde{N}_1\right]+\frac{N}{\tilde{b}}\mathrm{Var}\left[\tilde{\mu}^0(d)-\bar{\mu}^0(d)\tilde{N}_0\right]\\
&-2\frac{N}{\tilde{b}}\mathrm{Cov}\left[\tilde{\mu}^1(d)-\bar{\mu}^1(d)\tilde{N}_1,\tilde{\mu}^0(d)-\bar{\mu}^0(d)\tilde{N}_0\right].
\end{aligned}
$$

For the first term, we have:

$$
\begin{aligned}
&\frac{N}{\tilde{b}}\mathrm{Var}\left[\tilde{\mu}^1(d)-\bar{\mu}^1(d)\tilde{N}_1\right]=\frac{1}{N\tilde{b}}\mathrm{Var}\left[\sum_{i=1}^N\mathbf{1}\{\mathbf{Z}_i^{s:t}=\mathbf{z}^{s:t}\}\frac{\mu_i(\mathbf{Y}_t(\mathbf{Z}^{1:t});d)-\bar{\mu}^1(d)}{W_{it}}\right]\\
=&\frac{1}{N\tilde{b}}\sum_{i=1}^N\mathrm{Var}\left[\mathbf{1}\{\mathbf{Z}_i^{s:t}=\mathbf{z}^{s:t}\}\frac{\mu_i(\mathbf{Y}_t(\mathbf{Z}^{1:t});d)-\bar{\mu}^1(d)}{W_{it}}\right]\\
&+\frac{1}{N\tilde{b}}\sum_{i=1}^N\sum_{j\neq i}\mathrm{Cov}\left[\mathbf{1}\{\mathbf{Z}_i^{s:t}=\mathbf{z}^{s:t}\}\frac{\mu_i(\mathbf{Y}_t(\mathbf{Z}^{1:t});d)-\bar{\mu}^1(d)}{W_{it}},\mathbf{1}\{\mathbf{Z}_j^{s:t}=\mathbf{z}^{s:t}\}\frac{\mu_j(\mathbf{Y}_t(\mathbf{Z}^{1:t});d)-\bar{\mu}^1(d)}{W_{jt}}\right]\\
=&\frac{1}{N\tilde{b}}\sum_{i=1}^N\mathrm{E}\left[\mathbf{1}\{\mathbf{Z}_i^{s:t}=\mathbf{z}^{s:t}\}\frac{\mu_i(\mathbf{Y}_t(\mathbf{Z}^{1:t});d)-\bar{\mu}^1(d)}{W_{it}}\right]^2-\frac{1}{N\tilde{b}}\sum_{i=1}^N\mathrm{E}^2\left[\mathbf{1}\{\mathbf{Z}_i^{s:t}=\mathbf{z}^{s:t}\}\frac{\mu_i(\mathbf{Y}_t(\mathbf{Z}^{1:t});d)-\bar{\mu}^1(d)}{W_{it}}\right]\\
&+\frac{1}{N\tilde{b}}\sum_{i=1}^N\sum_{j\neq i}\mathrm{E}\left[\mathbf{1}\{\mathbf{Z}_i^{s:t}=\mathbf{z}^{s:t}\}\mathbf{1}\{\mathbf{Z}_j^{s:t}=\mathbf{z}^{s:t}\}\frac{\mu_i(\mathbf{Y}_t(\mathbf{Z}^{1:t});d)-\bar{\mu}^1(d)}{W_{it}}\frac{\mu_j(\mathbf{Y}_t(\mathbf{Z}^{1:t});d)-\bar{\mu}^1(d)}{W_{jt}}\right]\\
&-\frac{1}{N\tilde{b}}\sum_{i=1}^N\sum_{j\neq i}\mathrm{E}\left[\mathbf{1}\{\mathbf{Z}_i^{s:t}=\mathbf{z}^{s:t}\}\frac{\mu_i(\mathbf{Y}_t(\mathbf{Z}^{1:t});d)-\bar{\mu}^1(d)}{W_{it}}\right]\mathrm{E}\left[\mathbf{1}\{\mathbf{Z}_j^{s:t}=\mathbf{z}^{s:t}\}\frac{\mu_j(\mathbf{Y}_t(\mathbf{Z}^{1:t});d)-\bar{\mu}^1(d)}{W_{jt}}\right]\\
=&\frac{1}{N\tilde{b}}\sum_{i=1}^N\frac{\mathrm{E}\left[\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)-\bar{\mu}^1(d)\right]^2}{W_{it}}-\frac{1}{N\tilde{b}}\sum_{i=1}^N\mathrm{E}^2\left[\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)-\bar{\mu}^1(d)\right]\\
&+\frac{1}{N\tilde{b}}\sum_{i=1}^N\sum_{j\in\mathcal{B}(i;d)}\mathrm{E}\left[\left(\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_{i,j}^{s:t});d)-\bar{\mu}^1(d)\right)\left(\mu_j(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_{i,j}^{s:t});d)-\bar{\mu}^1(d)\right)\right]\\
&-\frac{1}{N\tilde{b}}\sum_{i=1}^N\sum_{j\in\mathcal{B}(i;d)}\mathrm{E}\left[\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)-\bar{\mu}^1(d)\right]\mathrm{E}\left[\mu_j(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_j^{s:t});d)-\bar{\mu}^1(d)\right]\\
=&\frac{1}{N\tilde{b}}\sum_{i=1}^N\frac{\mathrm{E}\left[\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)-\bar{\mu}^1(d)\right]^2}{W_{it}}-\frac{1}{N\tilde{b}}\sum_{i=1}^N\mathrm{E}^2\left[\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)-\bar{\mu}^1(d)\right]\\
&+\frac{1}{N\tilde{b}}\sum_{i=1}^N\sum_{j\in\mathcal{B}(i;d)}\mathrm{Cov}\left[\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)-\bar{\mu}^1(d),\mu_j(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_j^{s:t});d)-\bar{\mu}^1(d)\right].
\end{aligned}
$$

The penultimate equality uses the fact that under Assumption 5,

$$
\begin{aligned}
&\frac{1}{N\tilde{b}}\sum_{i=1}^N\sum_{j\notin\mathcal{B}(i;d)}\mathrm{E}\left[\left(\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_{i,j}^{s:t});d)-\bar{\mu}^1(d)\right)\left(\mu_j(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_{i,j}^{s:t});d)-\bar{\mu}^1(d)\right)\right]\\
=&\frac{1}{N}\sum_{i=1}^N\sum_{j\notin\mathcal{B}(i;d)}\mathrm{E}\left[\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)-\bar{\mu}^1(d)\right]\mathrm{E}\left[\mu_j(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_j^{s:t});d)-\bar{\mu}^1(d)\right].
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
&\frac{N}{\tilde{b}}\mathrm{Var}\left[\tilde{\mu}^0(d)-\bar{\mu}^0(d)\tilde{N}_1\right]\\
=&\frac{1}{N\tilde{b}}\sum_{i=1}^N\frac{\mathrm{E}\left[\mu_i(\mathbf{Y}_t(\tilde{\mathbf{z}}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)-\bar{\mu}^0(d)\right]^2}{W_{it}}-\frac{1}{N\tilde{b}}\sum_{i=1}^N\mathrm{E}^2\left[\mu_i(\mathbf{Y}_t(\tilde{\mathbf{z}}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)-\bar{\mu}^0(d)\right]\\
&+\frac{1}{N\tilde{b}}\sum_{i=1}^N\sum_{j\in\mathcal{B}(i;d)}\mathrm{Cov}\left[\mu_i(\mathbf{Y}_t(\tilde{\mathbf{z}}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)-\bar{\mu}^0(d),\mu_j(\mathbf{Y}_t(\tilde{\mathbf{z}}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_j^{s:t});d)-\bar{\mu}^0(d)\right].
\end{aligned}
$$

$$
\begin{aligned}
&\frac{N}{\tilde{b}}\text{Cov}\left[\tilde{\mu}^1(d) - \bar{\mu}^1(d)\tilde{N}_1, \tilde{\mu}^0(d) - \bar{\mu}^0(d)\tilde{N}_0\right] \\
=&\frac{1}{N\tilde{b}}\sum_{i=1}^{N}\sum_{j\in\mathcal{B}(i;d)}\text{Cov}\left[\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t}, \mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t}); d) - \bar{\mu}^1(d), \mu_j(\mathbf{Y}_t(\tilde{\mathbf{z}}^{s:t}, \mathbf{Z}^{1:T}\setminus\mathbf{Z}_j^{s:t}); d) - \bar{\mu}^0(d)\right].
\end{aligned}
$$

To summarize,

$$
\begin{aligned}
&\frac{N}{\tilde{b}}\text{Var}\left(\widehat{\tau}^*_{t,HA}(d)\right) \\
=&\frac{1}{N\tilde{b}}\sum_{i=1}^{N}\frac{\text{E}\left[\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t}, \mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t}); d) - \bar{\mu}^1(d)\right]^2}{W_{it}} + \frac{1}{N\tilde{b}}\sum_{i=1}^{N}\frac{\text{E}\left[\mu_i(\mathbf{Y}_t(\tilde{\mathbf{z}}^{s:t}, \mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t}); d) - \bar{\mu}^0(d)\right]^2}{W_{it}} \\
&-\frac{1}{N}\sum_{i=1}^{N}\left(\text{E}\left[\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t}, \mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t}); d) - \bar{\mu}^1(d)\right] - \text{E}\left[\mu_i(\mathbf{Y}_t(\tilde{\mathbf{z}}^{s:t}, \mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t}); d) - \bar{\mu}^0(d)\right]\right)^2 \\
&+\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j\in\mathcal{B}(i;d)}\sum_{\mathbf{a},\mathbf{b}=\mathbf{z}^{s:t}}^{\tilde{\mathbf{z}}^{s:t}}(-1)^{\mathbf{1}\{\mathbf{a}=\mathbf{b}\}}\text{Cov}\left[\mu_i(\mathbf{Y}_t(\mathbf{a}, \mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t}); d) - \bar{\mu}_t^{\mathbf{1}\{\mathbf{a}=\mathbf{z}^{s:t}\}}, \mu_j(\mathbf{Y}_t(\mathbf{b}, \mathbf{Z}^{1:T}\setminus\mathbf{Z}_j^{s:t}); d) - \bar{\mu}_t^{\mathbf{1}\{\mathbf{b}=\mathbf{z}^{s:t}\}}\right] + o_P(1) \\
\leq& N\tilde{V}_d + o_P(1).
\end{aligned}
$$

Hence the lemma holds. Again, the derivation for Var $\left(\widehat{\tau}^*_{HA}(d)\right)$ is similar and omitted. Notice that when $\tilde{b}$ increases with N, the difference between the Horvitz-Thompson estimator and the Hajek estimator gradually diminishes as $\tilde{N}_1$ and $\tilde{N}_0$ converge at the rate of $\sqrt{N}$. But in any finite sample, the Hajek estimator is still more efficient. It is also common to assume that $\tilde{b}$ does not change with N in the spatial setting. That's why I derive the variances for both the Horvitz-Thompson estimator and the Hajek estimator. □

It is then obvious that that

$$
\begin{aligned}
V_{1t}(d) =& \frac{1}{N^2}\sum_{i=1}^{N}\frac{\text{E}\left[\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t}, \mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t}); d) - \bar{\mu}_t^1\right]^2}{W_{it}} + \frac{1}{N^2}\sum_{i=1}^{N}\frac{\text{E}\left[\mu_i(\mathbf{Y}_t(\tilde{\mathbf{z}}^{s:t}, \mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t}); d) - \bar{\mu}_t^0\right]^2}{\tilde{W}_{it}} \\
&-\frac{1}{N^2}\sum_{i=1}^{N}\text{E}^2\left[\left(\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t}, \mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t}); d) - \bar{\mu}_t^1\right) - \left(\mu_i(\mathbf{Y}_t(\tilde{\mathbf{z}}^{s:t}, \mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t}); d) - \bar{\mu}_t^0\right)\right]
\end{aligned}
$$

and

$$
\begin{aligned}
&V_{2t}(d) \\
=&\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j\in\mathcal{B}(i;d)}\sum_{\mathbf{a},\mathbf{b}=\mathbf{z}^{s:t}}^{\tilde{\mathbf{z}}^{s:t}}(-1)^{\mathbf{1}\{\mathbf{a}=\mathbf{b}\}}\text{Cov}\left[\mu_i(\mathbf{Y}_t(\mathbf{a}, \mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t}); d) - \bar{\mu}_t^{\mathbf{1}\{\mathbf{a}=\mathbf{z}^{s:t}\}}, \mu_j(\mathbf{Y}_t(\mathbf{b}, \mathbf{Z}^{1:T}\setminus\mathbf{Z}_j^{s:t}); d) - \bar{\mu}_t^{\mathbf{1}\{\mathbf{b}=\mathbf{z}^{s:t}\}}\right].
\end{aligned}
$$

The variance expression can be further simplified with an extra assumption:

Assumption (Homophily in treatment effects)**.** *For any d, define*

$$
\tau_{it}(\mathbf{z}^{s:t}, \tilde{\mathbf{z}}^{s:t}; d) = \text{E}\left[\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t}, \mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t}); d)\right] - \text{E}\left[\mu_i(\mathbf{Y}_t(\tilde{\mathbf{z}}^{s:t}, \mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t}); d)\right],
$$

and

$$
\bar{\tau}_t(d) = \frac{1}{N}\sum_{i=1}^{N}\tau_{it}(d),
$$

then $\frac{1}{N}\sum_{i=1}^{N}(\tau_{it}(d) - \bar{\tau}_t(d))\sum_{j\in\mathcal{B}(i)}(\tau_{jt}(d) - \bar{\tau}_t(d)) \geq 0.$

The assumption means that the expected treatment effect generated by point i at distance d, period t is positively correlated with those effects generated by its neighbors. There is "homophily" in treatment effects on the space $\mathcal{X}$: points that generate larger-than-average effects reside close to each other. It is often the case in reality.

Lemma 3. *Under Assumptions 1-6 and the extra assumption, $\tilde{V}_d$ defined in lemma 2 is bounded by:*

$$\tilde{V}_d = \frac{1}{N^2}\sum_{i=1}^{N}\frac{\mathrm{E}\left[\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)-\bar{\mu}_t^1\right]^2}{W_{it}} + \frac{1}{N^2}\sum_{i=1}^{N}\frac{\mathrm{E}\left[\mu_i(\mathbf{Y}_t(\tilde{\mathbf{z}}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)-\bar{\mu}_t^0\right]^2}{\tilde{W}_{it}}$$
$$+\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j\in\mathcal{B}(i;d)}\sum_{\mathbf{a},\mathbf{b}=\mathbf{z}^{s:t}}^{\tilde{\mathbf{z}}^{s:t}}(-1)^{\mathbf{1}\{\mathbf{a}=\mathbf{b}\}}\mathrm{E}\left[\left(\mu_i(\mathbf{Y}_t(\mathbf{a},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)-\bar{\mu}_t^{\mathbf{1}\{\mathbf{a}=\mathbf{z}^{s:t}\}}\right),\left(\mu_j(\mathbf{Y}_t(\mathbf{b},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_j^{s:t});d)-\bar{\mu}_t^{\mathbf{1}\{\mathbf{b}=\mathbf{z}^{s:t}\}}\right)\right].$$

Proof. Since

$$\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j\in\mathcal{B}(i;d)}\sum_{\mathbf{a},\mathbf{b}=\mathbf{z}^{s:t}}^{\tilde{\mathbf{z}}^{s:t}}(-1)^{\mathbf{1}\{\mathbf{a}=\mathbf{b}\}}\mathrm{Cov}\left[\mu_i(\mathbf{Y}_t(\mathbf{a},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)-\bar{\mu}_t^{\mathbf{1}\{\mathbf{a}=\mathbf{z}^{s:t}\}},\mu_j(\mathbf{Y}_t(\mathbf{b},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_j^{s:t});d)-\bar{\mu}_t^{\mathbf{1}\{\mathbf{b}=\mathbf{z}^{s:t}\}}\right]$$
$$=\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j\in\mathcal{B}(i;d)}\sum_{\mathbf{a},\mathbf{b}=\mathbf{z}^{s:t}}^{\tilde{\mathbf{z}}^{s:t}}(-1)^{\mathbf{1}\{\mathbf{a}=\mathbf{b}\}}\mathrm{E}\left[\left(\mu_i(\mathbf{Y}_t(\mathbf{a},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)-\bar{\mu}_t^{\mathbf{1}\{\mathbf{a}=\mathbf{z}^{s:t}\}}\right)\left(\mu_j(\mathbf{Y}_t(\mathbf{b},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_j^{s:t});d)-\bar{\mu}_t^{\mathbf{1}\{\mathbf{b}=\mathbf{z}^{s:t}\}}\right)\right]$$
$$-\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j\in\mathcal{B}(i;d)}\sum_{\mathbf{a},\mathbf{b}=\mathbf{z}^{s:t}}^{\tilde{\mathbf{z}}^{s:t}}(-1)^{\mathbf{1}\{\mathbf{a}=\mathbf{b}\}}\mathrm{E}\left[\mu_i(\mathbf{Y}_t(\mathbf{a},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)-\bar{\mu}_t^{\mathbf{1}\{\mathbf{a}=\mathbf{z}^{s:t}\}}\right]\mathrm{E}\left[\mu_j(\mathbf{Y}_t(\mathbf{b},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_j^{s:t});d)-\bar{\mu}_t^{\mathbf{1}\{\mathbf{b}=\mathbf{z}^{s:t}\}}\right].$$

All that need to be shown is that the second term is smaller than or equal to zero in the limit under the extra assumption. We know that,

$$\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j\in\mathcal{B}(i;d)}\sum_{\mathbf{a},\mathbf{b}=\mathbf{z}^{s:t}}^{\tilde{\mathbf{z}}^{s:t}}(-1)^{\mathbf{1}\{\mathbf{a}=\mathbf{b}\}}\mathrm{E}\left[\mu_i(\mathbf{Y}_t(\mathbf{a},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)-\bar{\mu}_t^{\mathbf{1}\{\mathbf{a}=\mathbf{z}^{s:t}\}}\right]\mathrm{E}\left[\mu_j(\mathbf{Y}_t(\mathbf{b},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_j^{s:t});d)-\bar{\mu}_t^{\mathbf{1}\{\mathbf{b}=\mathbf{z}^{s:t}\}}\right]$$
$$=\frac{1}{N}\sum_{i=1}^{N}\sum_{j\in\mathcal{B}(i;d)}\mathrm{E}\left[\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)-\bar{\mu}_t^{\mathbf{1}\{\mathbf{a}=\mathbf{z}^{s:t}\}}(d)\right]*$$
$$\mathrm{E}\left[\mu_j(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_j^{s:t});d)-\mu_j(\mathbf{Y}_t(\tilde{\mathbf{z}}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_j^{s:t});d)-(\bar{\mu}^1(d)-\bar{\mu}^0(d))\right]$$
$$-\frac{1}{N}\sum_{i=1}^{N}\sum_{j\in\mathcal{B}(i;d)}\mathrm{E}\left[\mu_j(\mathbf{Y}(0,\mathbf{Z}_{-j});d)-\bar{\mu}^0(d)\right]*$$
$$\mathrm{E}\left[\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)-\mu_i(\mathbf{Y}_t(\tilde{\mathbf{z}}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)-(\bar{\mu}^1(d)-\bar{\mu}^0(d))\right]$$
$$=\frac{1}{N}\sum_{i=1}^{N}(\tau_{it}(d)-\bar{\tau}_t(d))\sum_{j\in\mathcal{B}(i;d)}(\tau_{jt}(d)-\bar{\tau}_t(d)).$$

Obviously, the requirement is satisfied under the extra assumption. The lemma is proved. □

A3 Consistency of the IPTW estimators

All we need to show is that the variance of each estimator converges to zero as $N\to\infty$. For $V_t(d)$, we know that $V_t(d)=V_{1t}(d)+V_{2t}(d)+V_{3t}(d)$. Since all the moments of $\mu(;d)$ are bounded, it is easy to see $V_{1t}(d)\to 0$ as N rises. The convergence of $V_{3t}(d)$ to zero is guaranteed by the consistency of the selected propensity score model. For $V_{2t}(d)$, notice that $V_{2t}(d)\le\frac{4}{N^2}\sum_{i=1}^{N}\sum_{j\in\mathcal{B}(i;d)}\tilde{y}^2=\frac{4}{N^2}\sum_{i=1}^{N}b_{i;d}\tilde{y}^2$. From Assumption 5, we know that $\frac{b_{i;d}}{N}\to 0$ for any i. Hence, $V_{2t}(d)\to 0$ as $N\to\infty$. The consistency of other estimators can be shown in a similar way. Notice that consistency actually only requires $\max_{i\in\{1,2,\ldots,N\}}b_{i;d}=o_P(N)$. The conclusion has been drawn in Sävje, Aronow and Hudgens (2021).

A4 Asymptotic distribution of the IPTW estimators

The asymptotic Normality of the Horvitz-Thompson estimators can be derived using central limit theorems for dependent random variables based on Stein's method (Ross, 2011, Ogburn et al. (2020)). The Hajek estimators' asymptotic distribution can be then obtained via the Delta method. I first restate the key lemmas in Ogburn et al. (2020) using terms defined in this paper.

Lemma 4. *(Ogburn et al. (2020), Lemma 1 and 2) Consider a set of N units. Let $U_1,\ldots,U_N$ be bounded mean-zero random variables with finite* *fourth moments and dependency neighborhoods $\mathcal{B}(i;d)$. If $b_{i;d} \leq \tilde{b}$ for all i and $\frac{\tilde{b}^2}{N} \to 0$, then*

$$\frac{\sum_{i=1}^N U_i}{\sqrt{\text{Var}\left(\sum_{i=1}^N U_i\right)}} \to N(0,1).$$

Now, I can prove the asymptotic Normality of our estimators using the above lemma.

Proof. Remember that $\tau_{it}(\mathbf{z}^{s:t}, \tilde{\mathbf{z}}^{s:t}; d) = \text{E}\left[\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t}, \mathbf{Z}^{1:T} \setminus \mathbf{Z}_i^{s:t}); d)\right] - \text{E}\left[\mu_i(\mathbf{Y}_t(\tilde{\mathbf{z}}^{s:t}, \mathbf{Z}^{1:T} \setminus \mathbf{Z}_i^{s:t}); d)\right]$, let's define U_i as

$$\sqrt{\frac{N}{\tilde{b}_d}}\left(\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t} = \mathbf{z}^{s:t}\}\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t}, \mathbf{Z}^{1:T} \setminus \mathbf{Z}_i^{s:t}); d)}{NW_{it}} - \frac{\mathbf{1}\{\mathbf{Z}_i^{s:t} = \tilde{\mathbf{z}}^{s:t}\}\mu_i(\mathbf{Y}_t(\tilde{\mathbf{z}}^{s:t}, \mathbf{Z}^{1:T} \setminus \mathbf{Z}_i^{s:t}); d)}{N\tilde{W}_{it}} - \frac{\tau_i}{N}\right).$$

Then, $\sum_{i=1}^N U_i = \sqrt{\frac{N}{\tilde{b}_d}}\left(\hat{\tau}_t^*(\mathbf{z}^{s:t}, \tilde{\mathbf{z}}^{s:t}; d) - \tau_t(\mathbf{z}^{s:t}, \tilde{\mathbf{z}}^{s:t}; d)\right)$ and $\text{E}[U_i] = 0$. We know that U_i has finite fourth moments as all the outcomes are bounded, and $\text{Var}\left(\sum_{i=1}^N U_i\right) = \frac{N}{\tilde{b}_d}\text{Var}\left(\hat{\tau}_t(\mathbf{z}^{s:t}, \tilde{\mathbf{z}}^{s:t}; d)\right)$ is also finite. Assumption 5 stipulates that $\tilde{b} = \max_{i \in \{1,2,\ldots,N\}} b_{i;d}$ and $\tilde{b}^2/N \to 0$. From Lemma 4, we know that $\frac{\sqrt{\frac{N}{\tilde{b}_d}}\left(\hat{\tau}_t(\mathbf{z}^{s:t}, \tilde{\mathbf{z}}^{s:t}; d) - \tau_t(\mathbf{z}^{s:t}, \tilde{\mathbf{z}}^{s:t}; d)\right)}{\sqrt{\text{Var}\left[\sqrt{\frac{N}{\tilde{b}_d}}\left(\hat{\tau}_t(\mathbf{z}^{s:t}, \tilde{\mathbf{z}}^{s:t}; d) - \tau_t(\mathbf{z}^{s:t}, \tilde{\mathbf{z}}^{s:t}; d)\right)\right]}} \to N(0,1)$. Therefore, $\sqrt{\frac{N}{\tilde{b}_d}}\left(\hat{\tau}_t(\mathbf{z}^{s:t}, \tilde{\mathbf{z}}^{s:t}; d) - \tau_t(\mathbf{z}^{s:t}, \tilde{\mathbf{z}}^{s:t}; d)\right) \to N(0, V_t^*(d))$. It is easy to see that both $\sqrt{N}(\tilde{N}_1 - 1)$ and $\sqrt{N}(\tilde{N}_0 - 1)$ are asymptotically Normal. By linearization, we know that,

$$\begin{aligned}
&\sqrt{\frac{N}{\tilde{b}_d}}\left(\hat{\tau}_{t,HA}(\mathbf{z}^{s:t}, \tilde{\mathbf{z}}^{s:t}; d) - \tau_t(\mathbf{z}^{s:t}, \tilde{\mathbf{z}}^{s:t}; d)\right) \\
=&\sqrt{\frac{N}{\tilde{b}_d}}\left(\hat{\tau}_t^*(\mathbf{z}^{s:t}, \tilde{\mathbf{z}}^{s:t}; d) - \tau_t(\mathbf{z}^{s:t}, \tilde{\mathbf{z}}^{s:t}; d)\right) + \sqrt{\frac{N}{\tilde{b}_d}}\bar{\mu}^0(d)(\tilde{N}_1 - 1) \\
&+ \sqrt{\frac{N}{\tilde{b}_d}}\bar{\mu}^1(d)(\tilde{N}_0 - 1) + \sqrt{\frac{N}{\tilde{b}_d}}\left(\frac{\partial \hat{\tau}_{t,HA}^*(d)}{\partial \mathbf{W}}\right)^T(\hat{\mathbf{W}} - \mathbf{W}) + o_P(1).
\end{aligned}$$

Since each of the terms is asymptotically Normal, $\sqrt{\frac{N}{\tilde{b}_d}}(\hat{\tau}_{HA}(d) - \tau(d;\alpha))$ converges to a Normal distribution as well.[1] □

A5 Statistical properties of the augmented estimators

The proposed augmented estimators belong to the class of doubly-robust estimators in statistics. Again, I first assume that the true values of the nuisance parameters are known. When either the propensity score or the response surface is correctly specified, we have the following lemma:

Lemma 5. *Under Assumptions 1-6, the augmented estimator is unbiased and consistent when either the propensity scores or the outcome values are fixed at their true values.*

[1] Notice that $\sqrt{\frac{\tilde{b}_d}{N}} \leq \sqrt{\frac{\tilde{b}_d^*}{N(t-s+1)}}$. So, it also ensures the convergence of $\tau_{HA}(\mathbf{z}^{s:t}, \tilde{\mathbf{z}}^{s:t}; d)$.

Proof. First, let's assume the propensity scores take their true values but the diffusion model is inaccurate.

$$
\begin{aligned}
&\mathrm{E}\left[\hat{\tau}^*_{t,Aug}(\mathbf{z}^{s:t},\tilde{\mathbf{z}}^{s:t};d)\right]\\
=&\frac{1}{N}\sum_{i=1}^{N}\mathrm{E}\left[\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\mathbf{z}^{s:t}\}\left[\mu_i(Y_t(\mathbf{Z}^{1:t});d)-\mu_i(\hat{Y}_t(\mathbf{Z}^{1:t});d)\right]}{P(\mathbf{Z}_i^{s:t}=\mathbf{z}^{s:t})}+\mu_i(\hat{Y}_{t;i}(\mathbf{z}^{s:t});d)\right]\\
&-\frac{1}{N}\sum_{i=1}^{N}\mathrm{E}\left[\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\tilde{\mathbf{z}}^{s:t}\}\left[\mu_i(Y_t(\mathbf{Z}^{1:t});d)-\mu_i(\hat{Y}_t(\mathbf{Z}^{1:t});d)\right]}{P(\mathbf{Z}_i^{s:t}=\tilde{\mathbf{z}}^{s:t})}+\mu_i(\hat{Y}_{t;i}(\tilde{\mathbf{z}}^{s:t});d)\right]\\
=&\frac{1}{N}\sum_{i=1}^{N}\mathrm{E}\left[\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\mathbf{z}_i^{s:t}\}\mu_i(Y_t(\mathbf{Z}^{1:t});d)}{P(\mathbf{Z}_i^{s:t}=\mathbf{z}^{s:t})}-\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\tilde{\mathbf{z}}^{s:t}\}\mu_i(Y_t(\mathbf{Z}^{1:t});d)}{P(\mathbf{Z}_i^{s:t}=\tilde{\mathbf{z}}^{s:t})}\right]\\
&-\frac{1}{N}\sum_{i=1}^{N}\mathrm{E}\left[\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\mathbf{z}^{s:t}\}\mu_i(\hat{Y}_t(\mathbf{Z}^{1:t});d)}{P(\mathbf{Z}_i^{s:t}=\mathbf{z}^{s:t})}+\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\tilde{\mathbf{z}}^{s:t}\}\mu_i(\hat{Y}_t(\mathbf{Z}^{1:t});d)}{P(\mathbf{Z}_i^{s:t}=\tilde{\mathbf{z}}^{s:t})}\right]\\
&+\frac{1}{N}\sum_{i=1}^{N}\mathrm{E}\left[\mu_i(\hat{Y}_{t;i}(\mathbf{z}^{s:t});d)-\mu_i(\hat{Y}_{t;i}(\tilde{\mathbf{z}}^{s:t});d)\right]\\
=&\tau^*_{t,Aug}(\mathbf{z}^{s:t},\tilde{\mathbf{z}}^{s:t};d)-\frac{1}{N}\sum_{i=1}^{N}\mathrm{E}\left[\mu_i(\hat{Y}_{t;i}(\mathbf{z}^{s:t});d)-\mu_i(\hat{Y}_{t;i}(\tilde{\mathbf{z}}^{s:t});d)\right]\\
&+\frac{1}{N}\sum_{i=1}^{N}\mathrm{E}\left[\mu_i(\hat{Y}_{t;i}(\mathbf{z}^{s:t});d)-\mu_i(\hat{Y}_{t;i}(\tilde{\mathbf{z}}^{s:t});d)\right]\\
=&\tau^*_{t,Aug}(\mathbf{z}^{s:t},\tilde{\mathbf{z}}^{s:t};d)
\end{aligned}
$$

The second to last equality holds due to the definition of $\mu_i(\hat{Y}_{t;i}(\mathbf{z}^{s:t});d)$. The expression indicates that the augmented estimator can be seen as the combination of the Horvitz-Thompson estimator and the weighted difference between two marginalized prediction values. The first term has been shown to be unbiased and consistent. The second term cancels out as long as the propensity scores are accurate. In addition, the variance of $\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\tilde{\mathbf{z}}^{s:t}\}}{P(\mathbf{Z}_i^{s:t}=\tilde{\mathbf{z}}^{s:t})}$ is bounded. Therefore, the variance of the second term converges to zero in the limit and consistency is ensured.

Now, suppose the response surfaces are accurate but the propensity scores are not. Then, $Y_{it}-\hat{Y}_{it}=\hat{e}_{it}$ and $\mathrm{E}\left[\hat{e}_{it}\right]=0$. We can see that

$$
\begin{aligned}
&\mathrm{E}\left[\hat{\tau}^*_{t,Aug}(\mathbf{z}^{s:t},\tilde{\mathbf{z}}^{s:t};d)\right]\\
=&\frac{1}{N}\sum_{i=1}^{N}\mathrm{E}\left[\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\mathbf{z}^{s:t}\}\left[\mu_i(Y_t(\mathbf{Z}^{1:t});d)-\mu_i(\hat{Y}_t(\mathbf{Z}^{1:t});d)\right]}{P(\mathbf{Z}_i^{s:t}=\mathbf{z}^{s:t})}+\mu_i(\hat{Y}_{t;i}(\mathbf{z}^{s:t});d)\right]\\
&-\frac{1}{N}\sum_{i=1}^{N}\mathrm{E}\left[\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\tilde{\mathbf{z}}^{s:t}\}\left[\mu_i(Y_t(\mathbf{Z}^{1:t});d)-\mu_i(\hat{Y}^t(\mathbf{Z}^{1:t});d)\right]}{P(\mathbf{Z}_i^{s:t}=\tilde{\mathbf{z}}^{s:t})}+\mu_i(\hat{Y}_{t;i}(\tilde{\mathbf{z}}^{s:t});d)\right]\\
=&\frac{1}{N}\sum_{i=1}^{N}\mathrm{E}\left[\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\mathbf{z}^{s:t}\}\mu_i(\hat{e}_{it};d)}{P(\mathbf{Z}_i^{s:t}=\mathbf{z}^{s:t})}\right]-\frac{1}{N}\sum_{i=1}^{N}\mathrm{E}\left[\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\tilde{\mathbf{z}}^{s:t}\}\mu_i(\hat{e}_{it};d)}{P(\mathbf{Z}_i^{s:t}=\tilde{\mathbf{z}}^{s:t})}\right]\\
&+\frac{1}{N}\sum_{i=1}^{N}\mu_i(\hat{Y}_{t;i}(\mathbf{z}^{s:t});d)-\frac{1}{N}\sum_{i=1}^{N}\mu_i(\hat{Y}_{t;i}(\tilde{\mathbf{z}}^{s:t});d)\\
=&\tau^*_{t,Aug}(\mathbf{z}^{s:t},\tilde{\mathbf{z}}^{s:t};d).
\end{aligned}
$$

The consistency of the estimator holds as the propensity scores are bounded and the response surface converges to the corresponding conditional expectations. □

From the proof of lemma 5, we have seen that the augmented estimator is a combination of the IPTW estimator and the difference between weighted predicted values. The consistency of the former has been shown. The latter's consistency comes directly from the consistency of the estimated propensity scores and the boundedness of the diffusion model. Hence, Theorem 2 for the augmented estimators is proved. Similarly, if we have shown the asymptotic Normality of the former, then the estimator is also asymptotically Normal.

A6 The DID estimator as an augmented estimator

When the dataset has a generalized DID or staggered adoption structure, a simple choice for the augmented estimator is to set $\hat{Y}_{it}$ to be the average of Y_{is} in the pre-treatment periods. It is a diffusion model that assumes null treatment effect. Let's denote the last period before the onset of unit i's treatment as K_i. Units with the same value of K_i are called a cohort and all the units with $K_i = \infty$ define the control group. Notice that K_i is a sufficient statistic of the assignment history $\mathbf{Z}_i^{1:t}$ for any period t, thus $P(\mathbf{Z}_i^{1:t} = \mathbf{z}^{1:t})$ equals to $P(K_i = k; t)$ and the outcome can be written as $\mathbf{Y}^t(\mathbf{K})$, where $\mathbf{K} = (K_1, K_2, \ldots, K_N)$ represents the cohorts of all the units. The estimator $\tau_{t,Aug}(\mathbf{z}^{s:t}, \tilde{\mathbf{z}}^{s:t}; d)$ can be presented as $\tau_{t,Aug}(k_1, k_2; d)$. Moreover, the prediction for Y_{it} is $\hat{Y}_{it}(K_i) = \frac{\mathbf{1}\{K_i = k_i\}}{k} \sum_{s=1}^{k} \hat{Y}_{is}$ where k_i is the actual cohort i belongs to. In other words, the prediction equals to the average pre-treatment outcomes if i has its actual treatment history and zero otherwise. Due to its linearity, $\mu_i(\hat{Y}_t) = \frac{1}{k} \sum_{s=1}^{k} \hat{\mathbf{Y}}_s(k_i)$. Therefore, the augmented estimator for the average cumulative indirect effect relative to the control group has the following expression:

$$
\begin{aligned}
&\hat{\tau}_{t,Aug}(k, \infty; d) \\
=&\frac{1}{N} \sum_{i=1}^{N} \frac{\mathbf{1}\{K_i = k\}}{\hat{P}(K_i = k; t)} \left[\mu_i(Y_t(\mathbf{K}); d) - \frac{1}{k} \sum_{s=1}^{k} \mu_i(\hat{Y}_s(\mathbf{k}); d) \right] \\
&- \frac{1}{N} \sum_{i=1}^{N} \frac{\mathbf{1}\{K_i \geq k\}}{P(K_i \geq k; t)} \left[\mu_i(Y_t(\mathbf{K}); d) - \frac{1}{k} \sum_{s=1}^{k} \mu_i(\hat{Y}_s(\mathbf{k}); d) \right] \\
&+ \frac{1}{N} \sum_{i=1}^{N} \frac{\mathbf{1}\{K_i = k\}}{k} \sum_{s=1}^{k} \mu_i(Y_s(\mathbf{k}); d) - \frac{1}{N} \sum_{i=1}^{N} \frac{\mathbf{1}\{K_i \geq k\}}{k} \sum_{s=1}^{k} \mu_i(Y_s(\mathbf{k}); d).
\end{aligned}
$$

The estimator has a similar form as the one proposed by Strezhnev (2018) under the conditional parallel trends assumption, with the outcomes replaced by spillover mappings and the extra two terms in the end. This similarity indicates that estimates from the weighted DID estimator should be close to that from the augmented estimator when the average pre-treatment outcome does not differ significantly over cohorts. This result holds no matter whether interference exists. But we should keep in mind that the two estimators are built on different assumptions and their results have different interpretations (cohort ATT and marginalized ATE). The augmented estimator justifies the usage of past outcomes and treatments when estimating the propensity scores.

A7 Variance estimation

I first show that the spatial HAC variance from the regression representation of the Hajek estimator $\hat{\tau}_{t,HA}(\mathbf{z}^{s:t}, \tilde{\mathbf{z}}^{s:t}; d)$ is asymptotically valid for the first two parts of its asymptotic variance $V_t(d)$ under the assumption on homophily in treatment effects. Let's denote the diagonal weighting matrix $\left\{ \frac{\mathbf{1}\{\mathbf{Z}_i^{s:t} = \mathbf{z}^{s:t}\} P(\mathbf{Z}_i^{s:t} = \tilde{\mathbf{z}}^{s:t}) - \mathbf{1}\{\mathbf{Z}_i^{s:t} = \tilde{\mathbf{z}}^{s:t}\} P(\mathbf{Z}_i^{s:t} = \mathbf{z}^{s:t})}{P(\mathbf{Z}_i^{s:t} = \tilde{\mathbf{z}}^{s:t}) P(\mathbf{Z}_i^{s:t} = \mathbf{z}^{s:t})} \right\}_{i=1,2,\ldots,N}$ as $\tilde{\mathbf{M}}_t = \{\tilde{M}_{it}\}_{i=1,2,\ldots,N}$. Define $\tilde{\mathbf{X}}_t$ as $\begin{pmatrix} 1, \mathbf{1}\{\mathbf{Z}_1^{s:t} = \mathbf{z}^{s:t}\} \\ 1, \mathbf{1}\{\mathbf{Z}_2^{s:t} = \mathbf{z}^{s:t}\} \\ \cdots \\ 1, \mathbf{1}\{\mathbf{Z}_N^{s:t} = \mathbf{z}^{s:t}\} \end{pmatrix}$ and $\tilde{\mathbf{Y}}_t$ as $\begin{pmatrix} \mu_1(\mathbf{Y}_t(\mathbf{Z}^{1:t}); d) \\ \mu_2(\mathbf{Y}_t(\mathbf{Z}^{1:t}); d) \\ \cdots \\ \mu_N(\mathbf{Y}_t(\mathbf{Z}^{1:t}); d) \end{pmatrix}$. Then the regression representation of the Hajek estimator has the solution $\begin{pmatrix} \hat{\alpha}(d) \\ \hat{\tau}_{t,OLS}(\mathbf{z}^{s:t}, \tilde{\mathbf{z}}^{s:t}; d) \end{pmatrix} = \left(\tilde{\mathbf{X}}_t' \tilde{\mathbf{M}}_t \tilde{\mathbf{X}}_t \right)^{-1} \left(\tilde{\mathbf{X}}_t' \tilde{\mathbf{M}}_t \tilde{\mathbf{Y}}_t \right)$.

It is easy to show that $\hat{\alpha}(d) = \frac{\sum_{i=1}^{N} \mathbf{1}\{\mathbf{Z}_i^{s:t} = \tilde{\mathbf{z}}^{s:t}\} \mu_i(\mathbf{Y}_t(\mathbf{Z}^{1:t}); d) / \hat{P}(\mathbf{Z}_i^{s:t} = \tilde{\mathbf{z}}^{s:t})}{\sum_{i=1}^{N} \mathbf{1}\{\mathbf{Z}_i^{s:t} = \tilde{\mathbf{z}}^{s:t}\} / \hat{P}(\mathbf{Z}_i^{s:t} = \tilde{\mathbf{z}}^{s:t})} = \hat{\mu}^0$ and $\hat{\tau}_{t,OLS}(d) = \hat{\tau}_{t,HA}(\mathbf{z}^{s:t}, \tilde{\mathbf{z}}^{s:t}; d) = \hat{\mu}^1 - \hat{\mu}^0$. I denote the residual for each observation as $\hat{e}_{it}(d)$ and the variance-covariance matrix of $\{\hat{e}_{it}(d)\}_{i=1,2,\ldots,N}$ as Σ_t. It is worth noting that the covariance of $\hat{e}_{it}(d)$ and $\hat{e}_{jt}(d)$ does not equal to zero if and only if $j \in \mathcal{B}(i; d)$. Moreover, I denote $\sum_{i=1}^{N} \tilde{M}_{it}$ as K, $\sum_{i=1}^{N} \tilde{M}_{it} \mathbf{1}\{\mathbf{Z}_i^{s:t} = \mathbf{z}^{s:t}\}$ as K_1, and $K_0 = K - K_1$. We know the spatial HAC variance of $\begin{pmatrix} \hat{\alpha}(d) \\ \hat{\tau}_{t,OLS}(\mathbf{z}^{s:t}, \tilde{\mathbf{z}}^{s:t}; d) \end{pmatrix}$ can be expressed as:

$$
\begin{aligned}
&\widehat{\operatorname{Var}}\begin{pmatrix} \hat{\alpha}(d) \\ \hat{\tau}_{t,OLS}(\mathbf{z}^{s:t}, \tilde{\mathbf{z}}^{s:t}; d) \end{pmatrix} \\
=&(\tilde{\mathbf{X}}'\tilde{\mathbf{M}}\tilde{\mathbf{X}})^{-1}(\tilde{\mathbf{X}}'\tilde{\mathbf{M}}\Sigma\tilde{\mathbf{M}}'\mathbf{X})(\tilde{\mathbf{X}}'\tilde{\mathbf{M}}\tilde{\mathbf{X}})^{-1} \\
=&\begin{pmatrix} K, K_1 \\ K_1, K_1 \end{pmatrix}^{-1} \left(\sum_{i=1}^{N}\sum_{j=1}^{N} \begin{pmatrix} \tilde{M}_{it}\tilde{M}_{jt} & \tilde{M}_{it}\tilde{M}_{jt}\mathbf{1}\{\mathbf{Z}_j^{s:t} = \mathbf{z}^{s:t}\} \\ \tilde{M}_{it}\tilde{M}_{jt}\mathbf{1}\{\mathbf{Z}_i^{s:t} = \mathbf{z}^{s:t}\} & \tilde{M}_{it}\tilde{W}_{jt}\mathbf{1}\{\mathbf{Z}_i^{s:t} = \mathbf{z}^{s:t}, \mathbf{Z}_j^{s:t} = \mathbf{z}^{s:t}\} \end{pmatrix} \hat{e}_{it}\hat{e}_{jt}\mathbf{1}\{j \in \mathcal{B}(i;d)\} \right) \begin{pmatrix} K, K_1 \\ K_1, K_1 \end{pmatrix}^{-1} \\
=&\frac{1}{K_1^2K_0^2}\begin{pmatrix} K_1, -K_1 \\ -K_1, K \end{pmatrix} \left(\sum_{i=1}^{N}\sum_{j=1}^{N} \begin{pmatrix} \tilde{M}_{it}\tilde{M}_{jt} & \tilde{M}_{it}\tilde{M}_{jt}\mathbf{1}\{\mathbf{Z}_j^{s:t} = \mathbf{z}^{s:t}\} \\ \tilde{M}_{it}\tilde{M}_{jt}\mathbf{1}\{\mathbf{Z}_i^{s:t} = \mathbf{z}^{s:t}\} & \tilde{M}_{it}\tilde{M}_{jt}\mathbf{1}\{\mathbf{Z}_i^{s:t} = \mathbf{z}^{s:t}, \mathbf{Z}_j^{s:t} = \mathbf{z}^{s:t}\} \end{pmatrix} \hat{e}_{it}\hat{e}_{jt}\mathbf{1}\{j \in \mathcal{B}(i;d)\} \right) \begin{pmatrix} K_1, -K_1 \\ -K_1, K \end{pmatrix}
\end{aligned}
$$

We are interested in entry $(2,2)$ of the above expression. Rearranging the observations such that those with $\mathbf{Z}_i^{s:t} = \mathbf{z}^{s:t}$ rank before those with $\mathbf{Z}_i^{s:t} = \tilde{\mathbf{z}}^{s:t}$, the quantity of interest can be simplified as:

$$
\begin{aligned}
&\widehat{\operatorname{Var}}\left(\hat{\tau}_{t,OLS}(\mathbf{z}^{s:t}, \tilde{\mathbf{z}}^{s:t}; d)\right) \\
=&\frac{1}{K_1^2}\sum_{i=1}^{N}\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t} = \mathbf{z}^{s:t}\}}{W_{it}^2}\hat{e}_i^2 + \frac{1}{K_0^2}\sum_{i=1}^{N}\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t} = \tilde{\mathbf{z}}^{s:t}\}}{\tilde{W}_{it}^2}\hat{e}_i^2 + \frac{1}{K_1^2}\sum_{i=1}^{N}\sum_{j\in\mathcal{B}(i;d)}\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t} = \mathbf{z}^{s:t}, \mathbf{Z}_j^{s:t} = \mathbf{z}^{s:t}\}}{W_{it}W_{jt}}\hat{e}_i\hat{e}_j \\
-&\frac{2}{K_1K_0}\sum_{i=1}^{N}\sum_{j\in\mathcal{B}(i;d)}\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t} = \mathbf{z}^{s:t}, \mathbf{Z}_j^{s:t} = \tilde{\mathbf{z}}^{s:t}\}}{W_{it}\tilde{W}_{jt}}\hat{e}_i\hat{e}_j + \frac{1}{K_0^2}\sum_{i=1}^{N}\sum_{j\in\mathcal{B}(i;d)}\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t} = \tilde{\mathbf{z}}^{s:t}, \mathbf{Z}_j^{s:t} = \tilde{\mathbf{z}}^{s:t}\}}{\tilde{W}_{it}\tilde{W}_{jt}}\hat{e}_i\hat{e}_j
\end{aligned}
$$

The last step is to show that the variance estimate $N\widehat{\operatorname{Var}}\left(\hat{\tau}_{t,OLS}(d)\right)$ is asymptotically valid for the variance of the Hajek estimator $\operatorname{Var}\left(\hat{\tau}_{t,HA}(d)\right)$. It means that they are of the same order in convergence and the former is conservative for the latter after rescaling. In the proof, we keep using the fact that the average covariance of the observations/units has the order of $O_p(\tilde{b}_d)$. Hence, after scaling by $\tilde{b}_d$, the average covariance is bounded.

Proof. Taking expectation on the first term of the variance estimate, we have:

$$
\begin{aligned}
&\frac{N}{\tilde{b}_d}\mathrm{E}\left[\frac{1}{K_1^2}\sum_{i=1}^{N}\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t} = \mathbf{z}^{s:t}\}}{W_{it}^2}\hat{e}_i^2\right] \\
=&\frac{N}{\tilde{b}_d}\mathrm{E}\left[\frac{1}{K_1^2}\sum_{i=1}^{N}\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t} = \mathbf{z}^{s:t}\}}{W_{it}^2}\left(\mu_i(\mathbf{Y}_t(\mathbf{Z}^{1:t}); d) - \hat{\mu}_t^1\right)^2\right] \\
=&\frac{N}{\tilde{b}_d}\mathrm{E}\left[\frac{1}{K_1^2}\sum_{i=1}^{N}\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t} = \mathbf{z}^{s:t}\}}{W_{it}^2}\left(\mu_i^2(\mathbf{Y}_t(\mathbf{Z}^{1:t}); d) - 2\hat{\mu}^1\mu_i(\mathbf{Y}_t(\mathbf{Z}^{1:t}); d) + (\hat{\mu}^1(d))^2\right)\right] \\
=&\frac{1}{N\tilde{b}_d}\sum_{i=1}^{N}\frac{\mathrm{E}\left[\mathbf{1}\{\mathbf{Z}_i^{s:t} = \mathbf{z}^{s:t}\}\mu_i^2(\mathbf{Y}_t(\mathbf{Z}^{1:t}); d)\right]}{W_{it}^2} \\
&-\frac{2}{N^2\tilde{b}_d}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathrm{E}\left[\frac{\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t}, \mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t}); d)}{W_{it}}\mu_j(\mathbf{Y}_t(\mathbf{z}^{s:t}, \mathbf{Z}^{1:T}\setminus\mathbf{Z}_j^{s:t}); d)\right] \\
&+\frac{1}{N^2\tilde{b}_d}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathrm{E}\left[\frac{\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t}, \mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t}); d)}{W_{it}}\mu_j(\mathbf{Y}_t(\mathbf{z}^{s:t}, \mathbf{Z}^{1:T}\setminus\mathbf{Z}_j^{s:t}); d)\right] + o_P(1) \\
=&\frac{1}{N\tilde{b}_d}\sum_{i=1}^{N}\frac{\mathrm{E}\left[\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t}, \mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t}); d) - \bar{\mu}_t^1\right]^2}{W_{it}} + o_P(1).
\end{aligned}
$$

The second term of the variance estimate can be shown to be asymptotically unbiased in the same way. Now consider the first covariance part in the variance estimate, we have:

$$
\begin{aligned}
&\frac{N}{\tilde{b}_d}\mathrm{E}\left[\frac{1}{K_1^2}\sum_{i=1}^{N}\sum_{j\in\mathcal{B}(i;d)}\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\mathbf{z}^{s:t},\mathbf{Z}_j^{s:t}=\mathbf{z}^{s:t}\}}{W_{it}W_{jt}}\hat{e}_i\hat{e}_j\right]\\
&\frac{N}{\tilde{b}_d}\mathrm{E}\left[\frac{1}{K_1^2}\sum_{i=1}^{N}\sum_{j\in\mathcal{B}(i;d)}\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\mathbf{z}^{s:t},\mathbf{Z}_j^{s:t}=\mathbf{z}^{s:t}\}}{W_{it}W_{jt}}\left(\mu_i(\mathbf{Y}_t(\mathbf{Z}^{1:t});d)-\hat{\mu}_t^1\right)\left(\mu_j(\mathbf{Y}_t(\mathbf{Z}^{1:t});d)-\hat{\mu}_t^1\right)\right]\\
=&\frac{1}{N\tilde{b}_d}\sum_{i=1}^{N}\sum_{j\in\mathcal{B}(i;d)}\mathrm{E}\left[\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\mathbf{z}^{s:t},\mathbf{Z}_j^{s:t}=\mathbf{z}^{s:t}\}}{W_{it}W_{jt}}\mu_i(\mathbf{Y}_t(\mathbf{Z}^{1:t});d)\mu_j(\mathbf{Y}_t(\mathbf{Z}^{1:t});d)\right]\\
&-\frac{1}{N\tilde{b}_d}\mathrm{E}\left[\sum_{i=1}^{N}\sum_{j\in\mathcal{B}(i;d)}\frac{\mathbf{1}\{\mathbf{Z}_i^{s:t}=\mathbf{z}^{s:t},\mathbf{Z}_j^{s:t}=\mathbf{z}^{s:t}\}}{W_{it}W_{jt}}\hat{\mu}_t^1\left(\mu_i(\mathbf{Y}_t(\mathbf{Z}^{1:t});d)-\hat{\mu}_t^1+\mu_j(\mathbf{Y}_t(\mathbf{Z}^{1:t});d)\right)\right]+o_P(1)\\
=&\frac{1}{N\tilde{b}_d}\sum_{i=1}^{N}\sum_{j\in\mathcal{B}(i)}\mathrm{E}\left[\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_{i,j}^{s:t});d)\mu_j(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_{i,j}^{s:t});d)\right]\\
&-\frac{1}{N^2\tilde{b}_d}\sum_{i=1}^{N}\sum_{j\in\mathcal{B}(i;d)}\sum_{k=1}^{N}\mathrm{E}\left[\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_{i,j}^{s:t});d)\mu_k(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_{k}^{s:t});d)\right]\\
&-\frac{1}{N^2\tilde{b}_d}\sum_{i=1}^{N}\sum_{j\in\mathcal{B}(i;d)}\sum_{k=1}^{N}\mathrm{E}\left[\mu_j(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_{i,j}^{s:t});d)\mu_k(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_{k}^{s:t});d)\right]\\
&+\frac{1}{N^2\tilde{b}_d}\mathrm{E}\left(\sum_{k=1}^{N}\mu_k(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_{k}^{s:t});d)\right)^2\\
=&\frac{1}{N\tilde{b}_d}\sum_{i=1}^{N}\sum_{j\in\mathcal{B}(i;d)}\mathrm{E}\left[\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_{i,j}^{s:t});d)\mu_j(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_{i,j}^{s:t});d)\right]\\
&-\frac{1}{N^2\tilde{b}_d}\left(\sum_{i=1}^{N}\sum_{j\in\mathcal{B}(i;d)}\mathrm{E}\left[\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_{i,j}^{s:t});d)\right]\right)\sum_{k=1}^{N}\mathrm{E}\left[\mu_k(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_{k}^{s:t});d)\right]\\
&-\frac{1}{N^2\tilde{b}_d}\left(\sum_{i=1}^{N}\sum_{j\in\mathcal{B}(i;d)}\mathrm{E}\left[\mu_j(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_{i,j}^{s:t});d)\right]\right)\sum_{k=1}^{N}\mathrm{E}\left[\mu_k(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_{k}^{s:t});d)\right]\\
&+\frac{1}{N^2\tilde{b}_d}\left(\sum_{k=1}^{N}\mathrm{E}\left[\mu_k(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_{k}^{s:t});d)\right]\right)^2+o_P(1)\\
=&\frac{1}{N}\sum_{i=1}^{N}\sum_{j\in\mathcal{B}(i;d)}\mathrm{E}\left[\left(\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_{i}^{s:t});d)-\bar{\mu}^1\right)\left(\mu_j(\mathbf{Y}(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_{j}^{s:t});d)-\bar{\mu}^1\right)\right]+o_P(1).
\end{aligned}
$$

The asymptotic unbiasedness of other covariance terms can be similarly shown. Hence,

$$
\begin{aligned}
&\mathrm{E}\left[\frac{N}{\tilde{b}_d}\widehat{\mathrm{Var}}\left(\hat{\tau}_{t,OLS}(d)\right)\right]\\
=&\frac{1}{N\tilde{b}_d}\sum_{i=1}^{N}\frac{\mathrm{E}\left[\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)-\bar{\mu}_t^1\right]^2}{W_{it}}+\frac{1}{N\tilde{b}_d}\sum_{i=1}^{N}\frac{\mathrm{E}\left[\mu_i(\mathbf{Y}_t(\tilde{\mathbf{z}}^{s:t},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)-\bar{\mu}_t^0\right]^2}{\tilde{W}_{it}}\\
&+\frac{1}{N\tilde{b}_d}\sum_{i=1}^{N}\sum_{j\in\mathcal{B}(i;d)}\sum_{\mathbf{a},\mathbf{b}=\mathbf{z}^{s:t}}^{\tilde{\mathbf{z}}^{s:t}}(-1)^{\mathbf{1}\{\mathbf{a}=\mathbf{b}\}}\mathrm{E}\left[\left(\mu_i(\mathbf{Y}_t(\mathbf{a},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_i^{s:t});d)-\bar{\mu}_t^{\mathbf{1}\{\mathbf{a}=\mathbf{z}^{s:t}\}}\right),\left(\mu_j(\mathbf{Y}_t(\mathbf{b},\mathbf{Z}^{1:T}\setminus\mathbf{Z}_j^{s:t});d)-\bar{\mu}_t^{\mathbf{1}\{\mathbf{b}=\mathbf{z}^{s:t}\}}\right)\right]o_P(1)\\
\geq&\frac{N}{\tilde{b}_d}\mathrm{Var}\left[\hat{\tau}_{HA}(d)\right].
\end{aligned}
$$

□

A8 Estimate the nuisance parameters

So far, I have assumed that the values of the nuisance parameters are known. Now, let's consider the scenario where these values have to be estimated via models whose convergence rate satisfies our requirement imposed in the main text. I start from the Hajek estimator

$\widehat{\tau}_{t,HA}(d)$. Define $\mathbf{W}$ as $(W_{11}, \ldots, W_{1T}, \ldots, W_{N1}, \ldots, W_{NT})$ and $\hat{\mathbf{W}}$ its estimate. By Taylor expansion, we have,

$$\sqrt{\frac{N}{\tilde{b}}}\left(\widehat{\tau}_{t,HA}(d) - \widehat{\tau}^*_{t,HA}(d)\right) = \sqrt{\frac{N}{\tilde{b}}}\left(\frac{\partial \widehat{\tau}^*_{t,HA}(d)}{\partial \mathbf{W}}\right)^T (\hat{\mathbf{W}} - \mathbf{W}) + o_p\left(\sqrt{\frac{N}{\tilde{b}}}||\hat{\mathbf{W}} - \mathbf{W}||\right)$$

Then, we can see that,

$$\begin{aligned}\frac{N}{\tilde{b}}\text{Var}\left(\widehat{\tau}_{t,HA}(d)\right) =& \frac{N}{\tilde{b}}\text{Var}\left(\widehat{\tau}^*_{t,HA}(d)\right) + \frac{N}{\tilde{b}}\left(\frac{\partial \widehat{\tau}^*_{t,HA}(d)}{\partial \mathbf{W}}\right)^T \text{Var}\left(\mathbf{W}\right)\left(\frac{\partial \widehat{\tau}^*_{t,HA}(d)}{\partial \mathbf{W}}\right) \\ & - 2\frac{N}{\tilde{b}}\text{Cov}\left(\widehat{\tau}^*_{t,HA}(d), \left(\frac{\partial \widehat{\tau}^*_{t,HA}(d)}{\partial \mathbf{W}}\right)^T \hat{\mathbf{W}}\right) + o_p\left(\frac{N}{\tilde{b}}||\hat{\mathbf{W}} - \mathbf{W}||^2\right)\end{aligned}$$

We know that $\frac{N}{\tilde{b}}\text{Var}\left(\widehat{\tau}^*_{t,HA}(d)\right) = V_{1t}(d) + V_{2t}(d)$, then the remaining parts in the expression above equals to $V_{3t}(d)$. It can be further simplified based on the specific model. In Lunceford and Davidian (2004), for examples, the authors study classic parametric models for the propensity scores: $W = W(\mathbf{V}, \beta)$, where $\mathbf{V}$ denotes all the confounders and β is the finite-dimensional parameter to be estimated. Suppose the model is correctly specified, then for the Hajek estimator, $V_{3t}(d) = \mathbf{H}'\Sigma_W^{-1}\mathbf{H}$, where

$$\begin{aligned}\mathbf{H} &= \frac{1}{N}\sum_{i=1}^{N}\left\{\frac{\text{E}\left[\mu_i(\mathbf{Y}_t(\mathbf{z}^{s:t}, \mathbf{Z}^{1:T} \setminus \mathbf{Z}_i^{s:t}); d) - \bar{\mu}_t^1\right]}{W_{it}} + \frac{\text{E}\left[\mu_i(\mathbf{Y}_t(\tilde{\mathbf{z}}^{s:t}, \mathbf{Z}^{1:T} \setminus \mathbf{Z}_i^{s:t}); d) - \bar{\mu}_t^0\right]}{\tilde{W}_{it}}\right\}W_{it,\beta}, \\ \Sigma_W &= \frac{1}{N}\sum_{i=1}^{N}\frac{W_{it,\beta}W_{it,\beta}'}{W_{it}\tilde{W}_{it}}, W_{it,\beta} = \frac{\partial W(\mathbf{V}_{it}, \beta)}{\partial \beta}.\end{aligned}$$

Hirano, Imbens and Ridder (2003) suggest that we can rely on the sieve estimator to estimate the propensity scores. Under conditions stated in the paper, the estimated propensity scores won't affect our main results. See their paper for details of the variance estimation.

From the Taylor expansion, we also see that as long as $\sqrt{\frac{N}{\tilde{b}}}(\hat{\mathbf{W}} - \mathbf{W}) = O_P(1)$, the difference between $\widehat{\tau}_{t,HA}(d)$ and $\widehat{\tau}^*_{t,HA}(d)$ is bounded and the estimate is both consistent and asymptotically Normal.

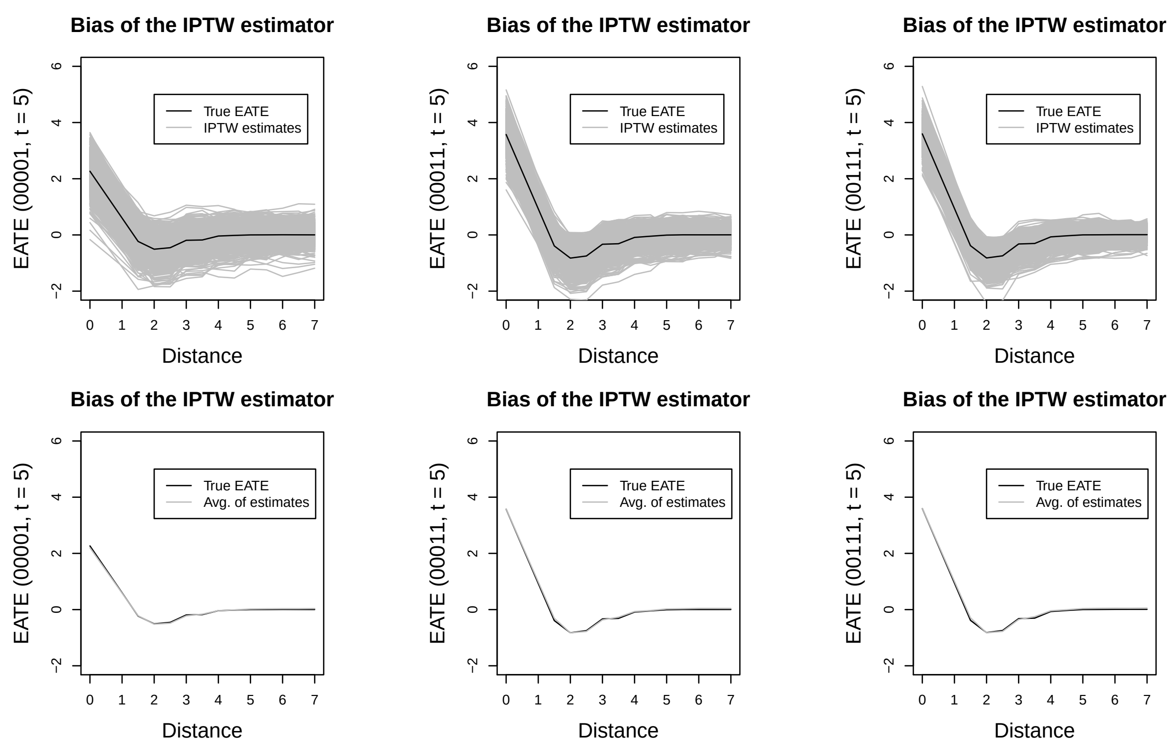
Appendix B Extra results from simulation and applications

This section presents extra results from both simulation and our empirical examples.

B1 Bias of the IPTW estimators under the sequential ignorability assumption and non-monotonic effect function

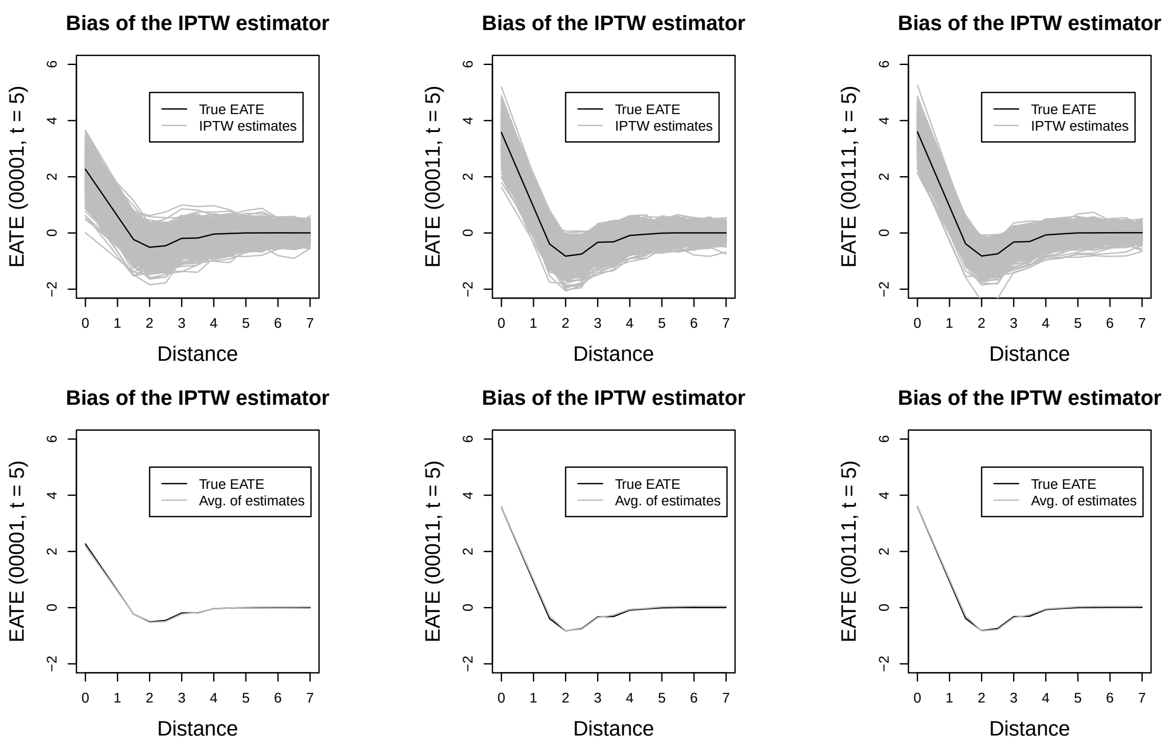
I first show the bias of both the Horvitz-Thompson estimator and the Hajek estimator when sequential ignorability holds and the effect function is non-monotonic. To be more specific, the effect function first declines as distance rises and then increases to zero. From Figure 1 and Figure 2, we can see that both estimators are unbiased.

Figure 1: Bias of the Horvitz-Thompson estimator



Notes: Figures on the top show the estimates from the Horvitz-Thompson estimator for all the 1000 assignments and figures on the bottom compare the averages of estimates against the EATEs. The difference between the gray and black curves on the bottom is the bias of the estimator. The effect function is non-monotonic and sequential ignorability holds.

Figure 2: Bias of the Hajek estimator

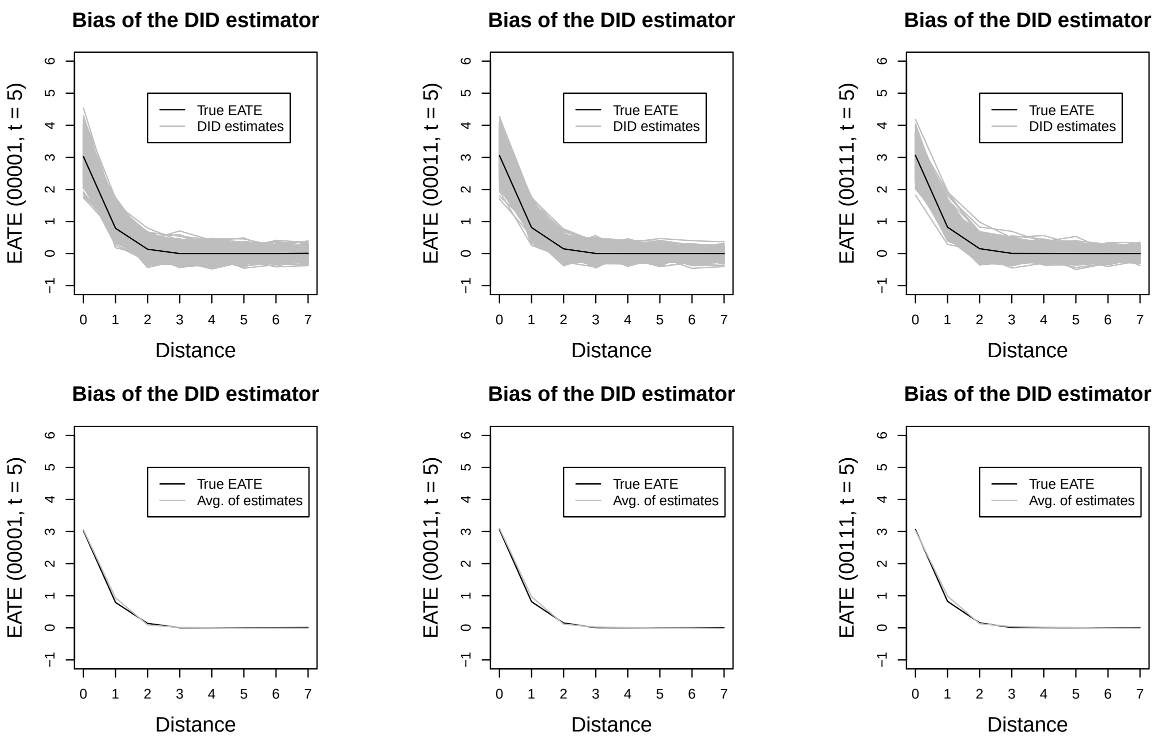


Notes: Figures on the top show the estimates from the Hajek estimator for all the 1000 assignments and figures on the bottom compare the averages of estimates against the EATEs. The difference between the gray and black curves on the bottom is the bias of the estimator. The effect function is non-monotonic and sequential ignorability holds.

B2 Bias of the DID estimator under homogeneous treatment effect

Figure 3 shows the bias of the DID estimator under the same data generating process we have seen in the main text—sequential ignorability does not hold and unit fixed effects are confounders. The only difference is that now the effects are homogeneous across units and not cumulative over periods. As predicted by our theory, now the DID estimator is unbiased for the EATEs.

Figure 3: Bias of the DID estimator under homogeneous treatment effect



Notes: Figures on the top show the estimates from the DID estimator for all the 1000 assignments and figures on the bottom compare the averages of estimates against the EATEs. The difference between the gray and black curves on the bottom is the bias of the estimator. The effect function is monotonic and homogeneous. Sequential ignorability does not hold.

B3 Bias of the augmented estimator under the sequential ignorability assumption

In Figures 4 and 5, I use the same data generating process for the IPTW estimators as in the main text. Sequential ignorability holds and the effects are heterogeneous. I rely on the following model to predict the value of each Y_{it}:

$$Y_{it} = \sum_{d=d_1}^{d_{\bar{d}}} \alpha_d \sum_{j=1}^{N} \mathbf{1}\{d_{ij} = d\} + \sum_{d=d_1}^{d_{\bar{d}}} \beta_d \sum_{j=1}^{N} Z_{jt} \mathbf{1}\{d_{ij} = d\} + \varepsilon_{it}$$

The model can be estimated via OLS. Then, we have the predicted value for each Y_{it}:

$$\hat{Y}_{it} = \sum_{d=d_1}^{d_{\bar{d}}} \hat{\alpha}_d \sum_{j=1}^{N} \mathbf{1}\{d_{ij} = d\} + \sum_{d=d_1}^{d_{\bar{d}}} \hat{\beta}_d \sum_{j=1}^{N} Z_{jt} \mathbf{1}\{d_{ij} = d\}$$

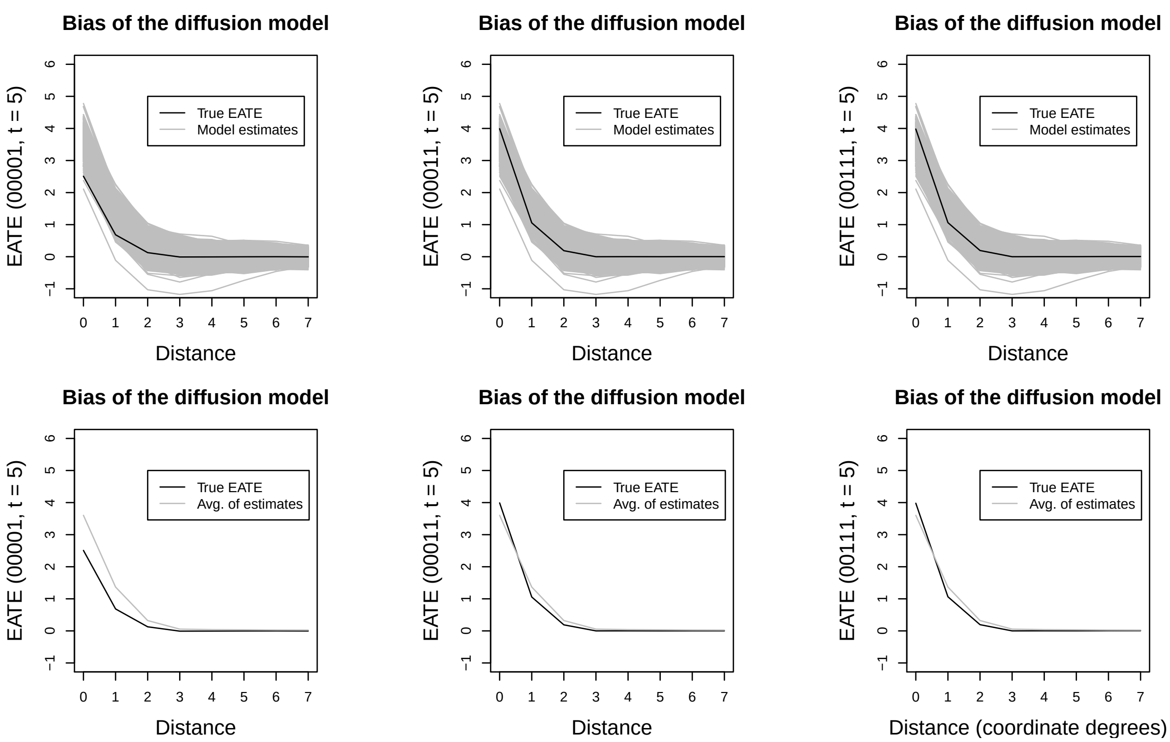
We can further estimate the marginalized outcomes by:

$$\hat{Y}_{it;k}(\mathbf{z}^{s:t}) = \widehat{\mathrm{E}}\left[Y_{it} | \mathbf{z}^{s:t}, \mathbf{Z}^{s:t} \setminus \mathbf{Z}_k^{s:t}\right] = \sum_{d=d_1}^{d_{\bar{d}}} \hat{\alpha}_d \sum_{j=1}^{N} \mathbf{1}\{d_{ij} = d\} + \sum_{d=d_1}^{d_{\bar{d}}} \hat{\beta}_d \left(z_{kt} \mathbf{1}\{d_{ik} = d\} + \sum_{j \neq k}^{N} P(Z_{jt} = 1) \mathbf{1}\{d_{ij} = d\} \right)$$

$\hat{\mu}_{it}$s and the marginalized ones are obtained by plugging in the corresponding $\hat{Y}_{it}$s and $\hat{Y}_{it;k}(\mathbf{z}^{s:t})$s. Clearly, the model ignores the lagged effects as well as the heterogeneity in the effects. From Figure 4, we can see that estimates from the diffusion model ($\hat{\beta}_d$s) are biased, especially for the first cohort which is not influenced by lagged effects.[2] Figure 5 shows that when I apply the augmented estimator, which combines the model with the weights in the IPTW estimators, the bias disappears and the estimates have smaller variances (compared with the Hajek estimator from the main text).

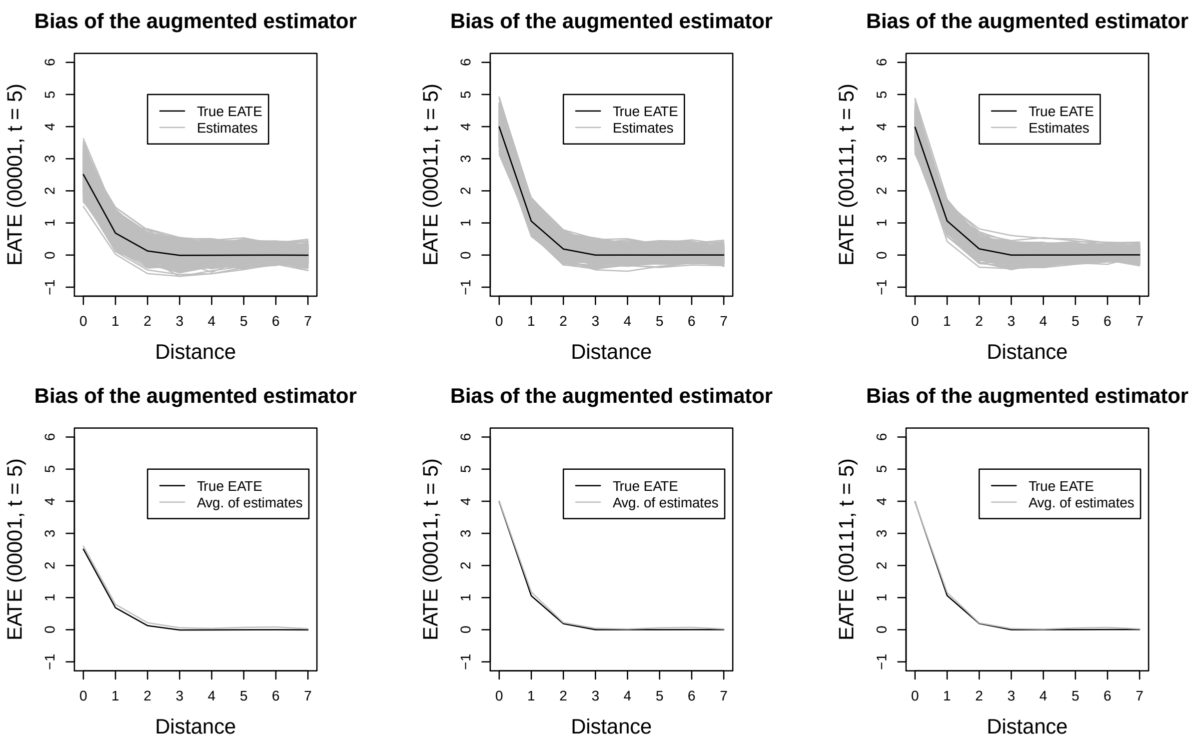
[2] The estimates are biased by the lagged effects hence the bias is larger where there is no lagged effect.

Figure 4: Bias of the diffusion model



Notes: Figures on the top show the estimates from the diffusion model for all the 1000 assignments and figures on the bottom compare the averages of estimates against the EATEs. The difference between the gray and black curves on the bottom is the bias of the estimator. The effect function is monotonic and sequential ignorability holds.

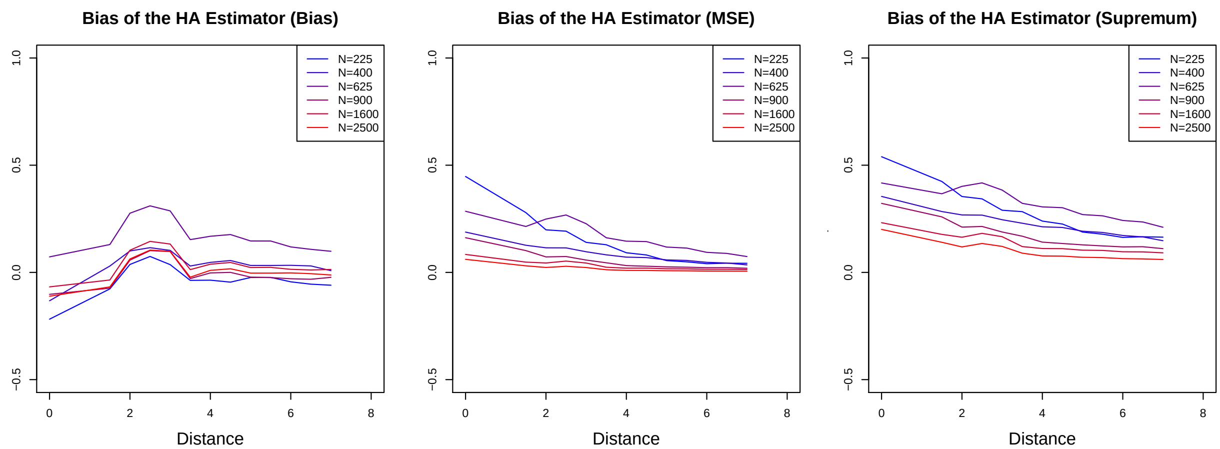
Figure 5: Bias of the augmented estimator



Notes: Figures on the top show the estimates from the augmented estimator for all the 1000 assignments and figures on the bottom compare the averages of estimates against the EATEs. The difference between the gray and black curves on the bottom is the bias of the estimator. The effect function is monotonic and sequential ignorability holds.

B4 Consistency of the IPTW estimators under the sequential ignorability assumption

Figure 6: Consistency of the Hajek estimator

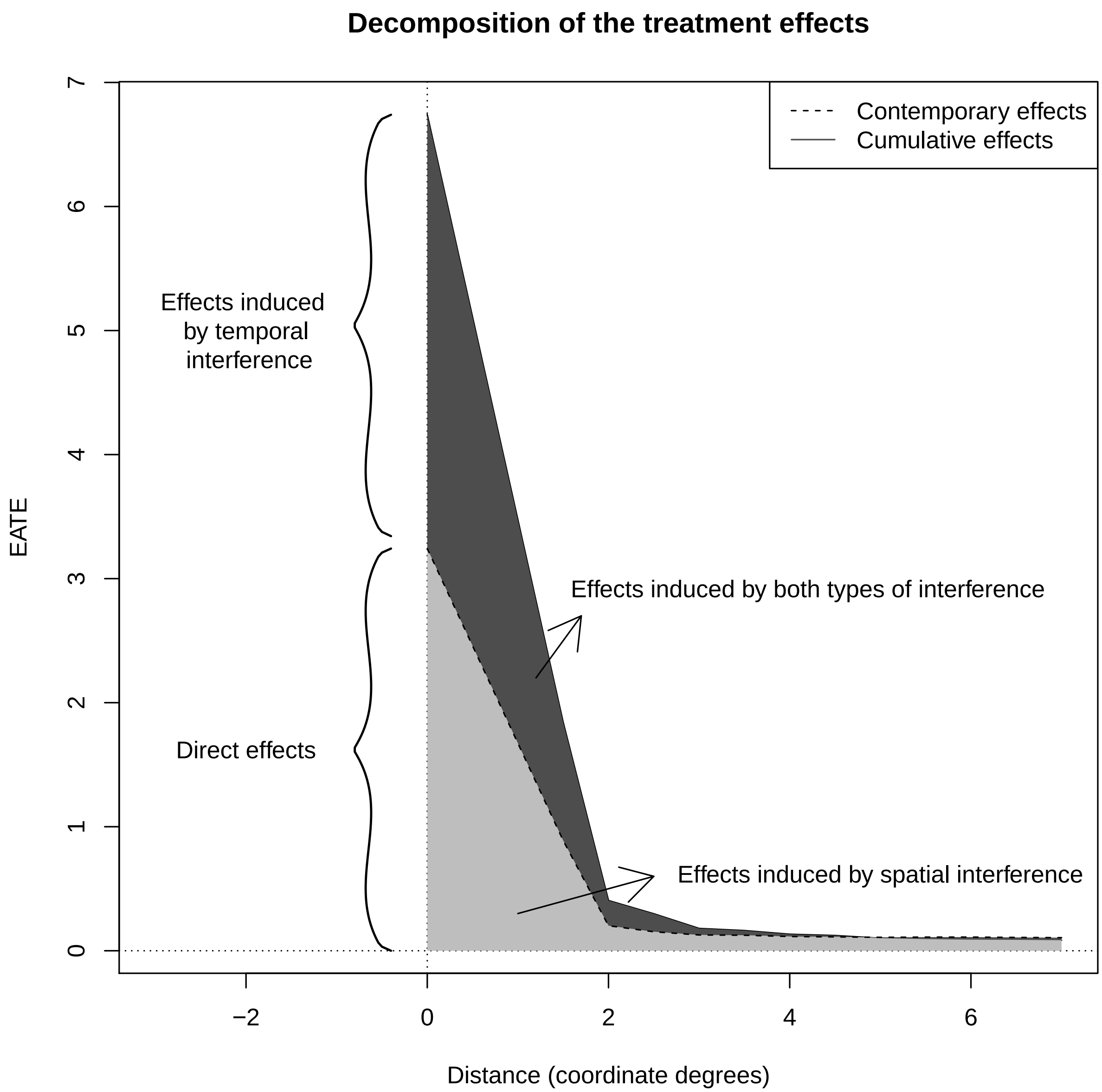


Notes: Figures from left to right show the average bias, average MSE, and supreme bias from 1,000 assignments using the Hajek estimator. Different colors indicate varying sample sizes.

B5 Decomposition of the treatment effects

I show how to decompose the total effect into different parts, in the spirit of Hu, Li and Wager (2021). I use the same data generating process for the IPTW estimators as in the main text. Sequential ignorability holds and the effects are heterogeneous. Remember that the treatment assignment has a structure of staggered adoption. I thus can estimate the contemporaneous effect at each d for each cohort K_i in period $K_i + 1$ (the first period when the treatment kicks off for units in K_i). Since these units are untreated in all the previous periods, the estimates appropriately capture the treatment effect that is not induced by temporal interference. In the figure below, I plot the average of these estimates across all the cohorts and how they vary with d. In addition, I use the effect generated by cohort 2 (with history $(0, 0, 1, 1, 1)$) to measure the cumulative effect.[3] Now, the average contemporaneous effect at $d = 0$ is the direct effect generated by the treatment. When $d > 0$, the estimate reflects the impact of only spatial interference. The difference between the average contemporaneous effect and the cumulative effect at $d = 0$ is induced by temporal interference, and their difference at any $d > 0$ is non-zero only when we have both types of interference. By choosing a density function $s(d)$ properly, I can further calculate the total effect generated by spatial interference via the area below both curves. One example of $s(d)$ is the average number of units that are passed through by a circle with radius d.

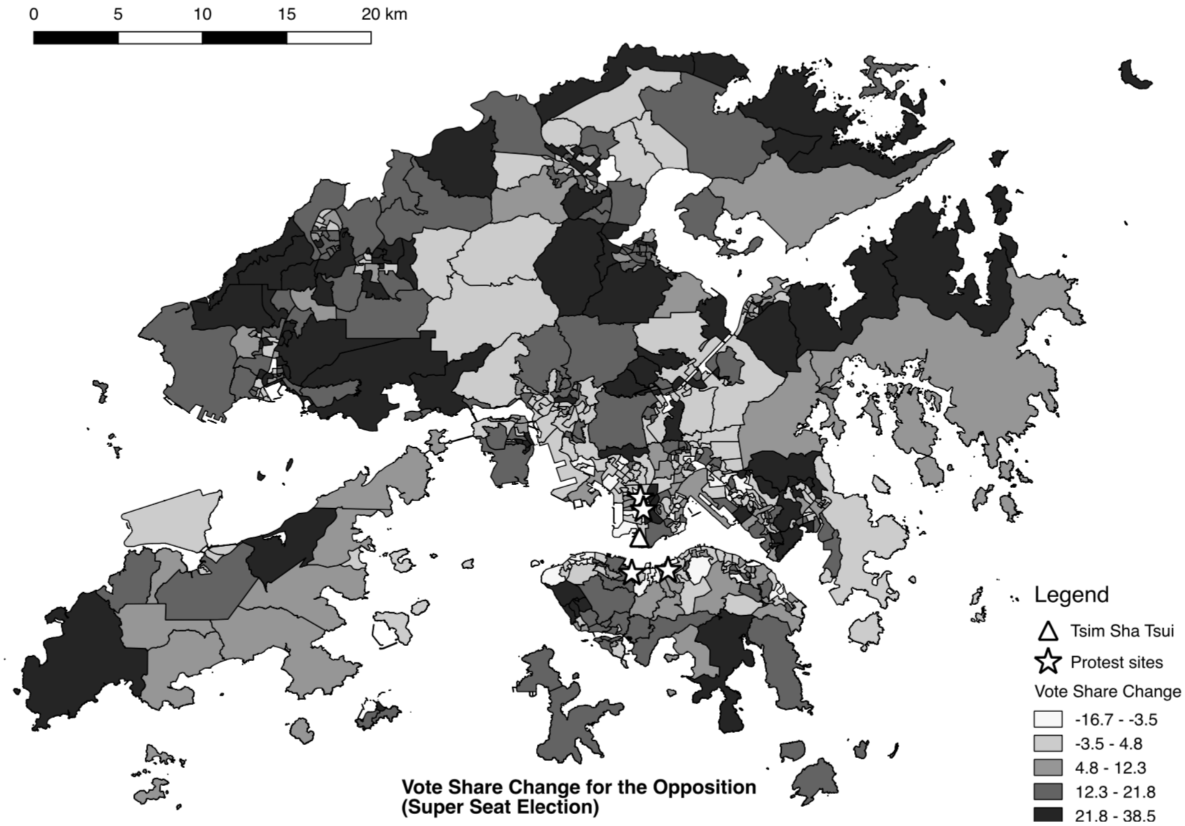
[3] We can similarly take an average over all the estimates of the cumulative effects based on their relative periods to the treatment's onset.



Notes: The figure shows both the curve for the average contemporaneous effect across all the cohorts and the cumulative effect generated by cohort 2. See the paragraph above to see the interpretation of different parts of the figure.

B6 Map of Hong Kong constituencies

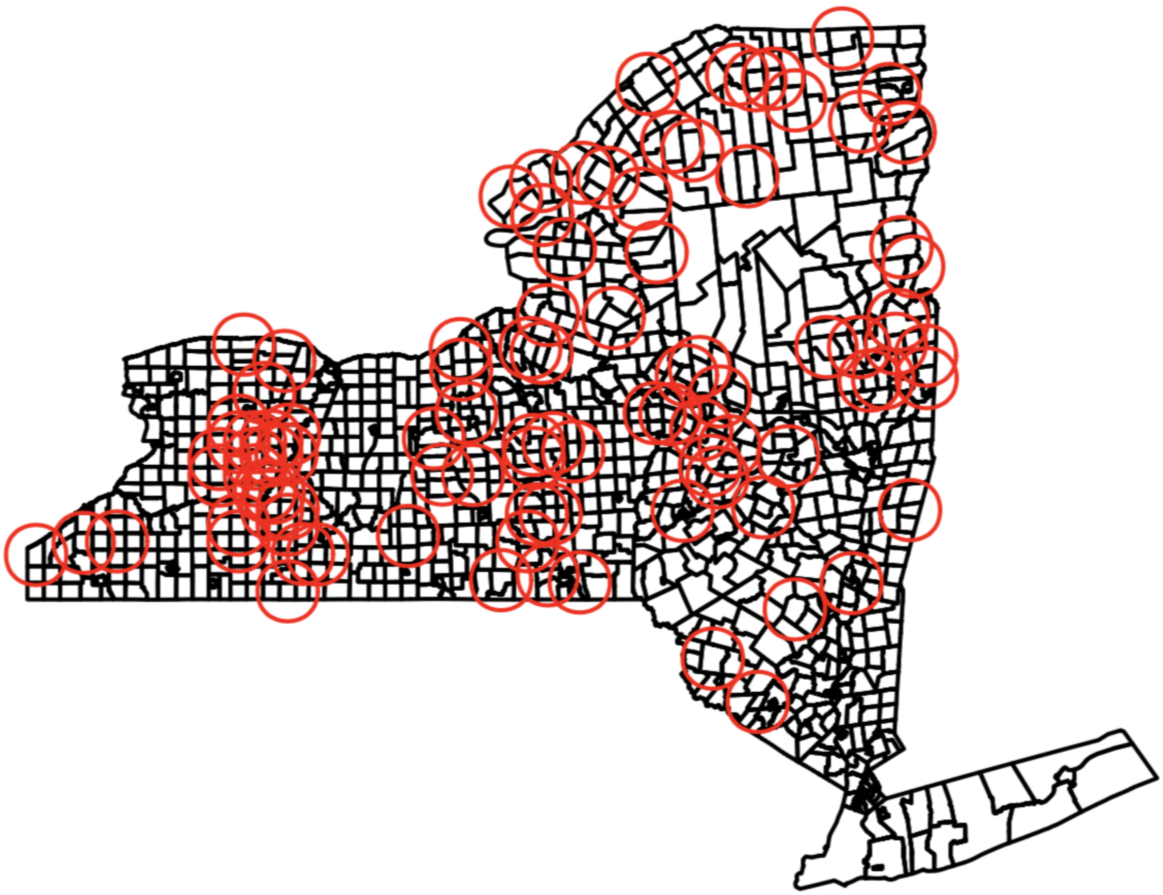
Figure 7: Map of Hong Kong constituencies



Notes: The figure shows district council constituencies in Hong Kong. The color of each cell represents the vote share change of the opposition from the 2012 election to the 2016 election.

B7 Map of towns in New York State

Figure 8: Map of towns in New York State



Notes: The figure shows towns in the State of New York. Towns surrounded by red circles belong to cohort 3, the treatment group of interest in the analysis.

References

Aronow, Peter M, Cyrus Samii and Ye Wang. 2020. "Design-based inference for spatial experiments with interference." *arXiv preprint arXiv:2010.13599* .

Hirano, Keisuke, Guido W Imbens and Geert Ridder. 2003. "Efficient estimation of average treatment effects using the estimated propensity score." *Econometrica* 71(4):1161–1189.

Hu, Yuchen, Shuangning Li and Stefan Wager. 2021. "Average Treatment Effects in the Presence of Interference." *arXiv preprint arXiv:2104.03802* .

Lunceford, Jared K and Marie Davidian. 2004. “Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study.” *Statistics in medicine* 23(19):2937–2960.

Ogburn, Elizabeth L, Oleg Sofrygin, Ivan Diaz and Mark J van der Laan. 2020. “Causal inference for social network data.” *arXiv preprint arXiv:1705.08527* .

Ross, Nathan. 2011. “Fundamentals of Stein?s method.” *Probability Surveys* 8:210–293.

Sävje, Fredrik, Peter M Aronow and Michael G Hudgens. 2021. “Average treatment effects in the presence of unknown interference.” *The Annals of Statistics* 49(2):673–701.

Strezhnev, Anton. 2018. “Semiparametric weighting estimators for multi-period difference-in-differences designs.” Working Paper, Harvard University.