Presence of Time-Varying confounders in TT design

We show what happens in the presence of time-varying unmeasured confounders when using TT design. We first stratify the entire population into g strata according to the quintiles of the estimated CPE, which is estimated based on the observed covariates via the logistic regression. For each subgroup G and each time point t, we aggregate the individual-level data to obtain quantities in the following 2×2 table

	outcome $Y_i^t = 1$	outcome $Y_i^t = 0$	Total
Exposure $Z_i^t = 1$	n_{11}^t	n_{10}^t	n_1^t
Exposure $Z_i^t = 0$	n_{01}^t	n_{00}^t	n_0^t

Because h is the logit function, we have:

$$\begin{aligned} E(Y_i^t|Z_i^t, G) &= E(E(Y_i^t|Z_i^t, G, X_i^t)) \\ &= \int \frac{\exp(\beta_0 + \beta_1 Z_i^t + \beta_2 t + \gamma^T X_i)}{1 + \exp(\beta_0 + \beta_1 Z_i^t + \beta_2 t + \gamma^T X_i^t)} dF(X_i^t|Z_i^t, G), \end{aligned} \tag{1}$$

where X^t is the vector of measured/unmeasured confounders. We first show that the treatment effect wont be identifiable under the standard TT assumptions that are

1. Covariates and time have multiplicative effects on being exposed. i.e. $P(Z_i^t|X_i^t) = h_1(X_i^t)h_2(t)$.
2. Covariates for all individuals in any subgroup G are random variables from an unknown distribution. i.e., $p(X_i^t|G) = f_G^t$.
3. The outcome is a rare event, and therefore we can omit the denominator of the integrand in equation (1).

With these assumptions, we have:

$$\begin{aligned} E(Y_i^t|Z_i^t, G) &\approx \int exp(\beta_0 + \beta_1 Z_i^T + \beta_2 t + \gamma^T X_i^t) dF(X_i^t|Z_i^t, G) \\ &= exp(\beta_0 + \beta_1 Z_i^T + \beta_2 t) E(exp\{\gamma^T X_i^t\}|Z_i^t, G) \end{aligned} \tag{2}$$

In order to expand $E(\gamma^T X_i^t | Z_i^t, G)$, we compute the conditional distribution of covariates X_i given Z_i^T and G using the Bayes rule:

$$\begin{aligned}
p(X_i^t|Z_i^t = 1, G) &= \frac{p(Z_i^t = 1, X_i^t|G)}{p(Z_i^t = 1|G)} = \frac{p(Z_i^t = 1|X_i^t, G)p(X_i^t|G)}{p(Z_i^t = 1|G)} \\
&= \frac{p(Z_i^t = 1|X_i^t)p(X_i^t|G)}{p(Z_i^t = 1|G)} = \frac{h_1(X_i^t)h_2(t)f_G^t}{p(Z_i^t = 1|G)} \\
p(X_i^t|Z_i^t = 0, G) &= \frac{p(Z_i^t = 0, X_i^t|G)}{p(Z_i^t = 0|G)} = \frac{p(Z_i^t = 0|X_i^t, G)p(X_i^t|G)}{p(Z_i^t = 0|G)} \\
&= \frac{p(Z_i^t = 0|X_i^t)p(X_i^t|G)}{p(Z_i^t = 0|G)} = \frac{f_G^t - h_1(X_i^t)h_2(t)f_G^t}{p(Z_i^t = 0|G)}
\end{aligned}$$

Therefore,

$$\begin{aligned}
p(X_i^t|Z_i^t = 1, G) &= \frac{h_1(X_i^t)h_2(t)f_G^t}{p(Z_i^t = 1|G)} \\
p(X_i^t|Z_i^t = 0, G) &= \frac{f_G^t - h_1(X_i^t)h_2(t)f_G^t}{p(Z_i^t = 0|G)}
\end{aligned}$$

Define the following constants which depend on both G and t

$$\begin{aligned}
C_{1G}^t &:= \int \exp(\gamma^T X_i^t)h_1(X_i^t)f_G^t dX_i^t \\
C_{2G}^t &:= \int \exp(\gamma^T X_i^t)f_G^t dX_i^t \\
C_{3G}^t &:= \int h_1(X_i^t)f_G^t dX_i^t
\end{aligned}$$

The marginal expectation $E(Y_i^t|Z_i^t, G)$ now becomes:

$$\begin{aligned}
E(Y_i^t|Z_i^t = 1, G) &= \exp(\beta_0 + \beta_1 + \beta_2 t)\frac{C_{1G}^t}{C_{3G}^t} \\
E(Y_i^t|Z_i^t = 0, G) &= \exp(\beta_0 + \beta_2 t)\frac{C_{2G}^t - C_{1G}^t h_2(t)}{1 - C_{3G}^t h_2(t)}
\end{aligned}$$

The two equations above are covariates-free. Thus, the marginal expectation of outcome is the same across treated/control individuals within the same subgroup and time.

Because each Y_i^t is a binary variable, aggregating outcomes for the treated and the control yield two binomial distributions. Consequently, we can write down the parametric likelihood for $(n_{11}^t, n_{01}^t, n_{10}^t, n_{00}^t)$:

$$n_{11}^t \sim Binomial(n_{11}^t + n_{10}^t, e^{\beta_0+\beta_1+\beta_2 t}\frac{C_{1G}^t}{C_{3G}^t}) \tag{3}$$

$$n_{01}^t \sim Binomial(n_{01}^t + n_{00}^t, e^{\beta_0+\beta_2 t}\frac{C_{2G}^t - h_2(t)C_{1G}^t}{1 - h_2(t)C_{3G}^t}) \tag{4}$$

where $C_{1G}^t, C_{2G}^t, C_{3G}^t$ are unknown constants that depend on group and time.

The resulted parametric likelihood has more parameters than data points, and thus, the parameters are not identifiable. One way to reduce the number of parameters is to impose two additional assumptions:

1'. Covariates and time have multiplicative effects on being exposed. i.e. $P(Z_i^t|X_i^t) = h_1(X_i^0)h_2(t)$, where X^0 denote the vector of confounders at baseline.

2'. Covariates for all individuals in any subgroup G are random variables from an unknown distribution such that

$$E(exp\{\gamma^T X_i^t\}|Z_i^t, G) = \kappa h_3(t)E(\exp\{\gamma^T X_i^0\}|Z_i^t, G).$$

where κ is a constant and $h_3(t)$ is a function of time. This assumption holds when, for example, $X^t = \omega(t) + X^0 + \epsilon$, where ϵ is a mean zero random noise that is independent of (Z^t, G). Then,

$$E(exp\{\gamma^T X_i^t\}|Z_i^t, G) = \exp\{\gamma^T \omega(t)\}E(\exp\{\gamma^T \epsilon\})E(\exp\{\gamma^T X_i^0\}|Z_i^t, G)$$

Let $f_G^0 = p(X_i^0|G)$. Under these two assumptions,

$$\begin{aligned} p(X_i^t|Z_i^t = 1, G) &= \frac{h_1(X_i^0)h_2(t)f_G^0 h_3(t)}{p(Z_i^t = 1|G)} \\ &= \frac{h_1(X_i^0)h_2(t)f_G^0 h_3(t)}{\int h_1(X_i^0)h_2(t)f_G^0 h_3(t)dX_i^0} = \frac{h_1(X_i^0)f_G^0}{\int h_1(X_i^0)f_G^0 dX_i^0} \\ p(X_i^t|Z_i^t = 0, G) &= \frac{f_G^0 h_3(t) - h_1(X_i^0)h_2(t)f_G^0 h_3(t)}{p(Z_i^t = 0|G)} \\ &= \frac{f_G^0 h_3(t) - h_1(X_i^0)h_2(t)f_G^0 h_3(t)}{1 - \int h_1(X_i^0)h_2(t)f_G^0 h_3(t)dX_i^0} = \frac{f_G^0 h_3(t) - h_1(X_i^0)h_2(t)f_G^0 h_3(t)}{1 - h_2^*(t)\int h_1(X_i^0)f_G^0 dX_i^0}, \end{aligned}$$

where $h_2^*(t) = h_2(t)h_3(t)$. Therefore,

$$\begin{aligned} C_{1G} &:= \int \exp(\gamma^T X_i^t)h_1(X_i^0)f_G^0 dX_i^0 \\ C_{2G} &:= \int \exp(\gamma^T X_i^0)f_G^0 dX_i^0 \\ C_{3G} &:= \int h_1(X_i^0)f_G^0 dX_i^0 \end{aligned}$$

The marginal expectation $E(Y_i^t|Z_i^t, G)$ now becomes:

$$\begin{aligned} E(Y_i^t|Z_i^t = 1, G) &= \exp(\beta_0 + \beta_1 + \beta_2 t)\frac{C_{1G}}{C_{3G}} \\ E(Y_i^t|Z_i^t = 0, G) &= \exp(\beta_0 + \beta_2 t)\frac{h_3(t)C_{2G} - C_{1G}h_2^*(t)}{1 - C_{3G}h_2^*(t)}. \end{aligned}$$

The model above has $3*g+2*T+3$ parameters and $g*T$ data points. Thus as long as $3*g+2*T+3 \leq g*T$, the treatment effect would be identifiable.