Data Generative Model Design for Repeated Outcomes from a Clustered SMART

Gabriel Durham

Note: This document is a modified appendix from the upcoming manuscript: *Three Level Analyses of Clustered SMARTs: With Application to Public Policy.* Development of the data generative model described below was motivated by the simulation study in this manuscript. When referring to this repository in publications please cite the aforementioned manuscript.

In designing the simulation study for this report, we sought to simulate outcomes sequentially, believing this to be more realistic than other data-generative approaches. Simulating outcomes conditional on past outcomes (and response status) necessitated specifying[1] the conditional distribution of the outcomes. Given that our proposed approach fundamentally concerns marginal parameter estimation, we sought to construct a data-generative mechanism consistent with the marginal structural model discussed in the main body of this report. In an attempt to limit the complexity of the simulation model, we limited our simulations to the case of three time points ($t = 0, 1, 2$), where response is determined at time $t^* = 1$. We use $\vec{Y}_0$, $\vec{Y}_1$, and $\vec{Y}_2$ to denote the vector of outcomes for a given cluster at times $t = 0, 1, 2$, respectively.

Implementation code for this data-generative model can be found at `https://github.com/GabrielDurham/three_lvl_cSMART_sim_study`.

1 Target Structural Mean Model

As discussed above, we sought a conditional data-generative framework; however, as discussed in the main body, we propose a marginal modeling framework. Consequently, studying the properties of our estimator requires proper control over the marginal distribution of the generated data.

In the section below, we seek to generate data that mimics the marginal mean model for a prototypical SMART proposed in the main body:

$$\begin{aligned}\mathbb{E}\left[\vec{Y}_{i,t}^{(a_1,a_{2NR})} \mid X_i\right] &= \eta X_i + \gamma_0^M \vec{\mathbf{1}}_{n_i} + \mathbb{1}_{t\leq 1}\left(\gamma_1^M t + \gamma_2^M a_1 t\right)\vec{\mathbf{1}}_{n_i} \\ &\quad + \mathbb{1}_{t>1}\left(\gamma_1^M + \gamma_2^M a_1 + \gamma_3^M (t-1) + \gamma_4^M (t-1) a_1 + \gamma_5^M (t-1) a_{2NR} + \gamma_6^M (t-1) a_1 a_{2NR}\right)\vec{\mathbf{1}}_{n_i},\end{aligned}$$

where we employ the superscript M to denote marginal parameters. We note that the model above is fully saturated in that it allows the specification of all potential expectations: $\mathbb{E}\left[\vec{Y}_{i,0}^{(.,.)}\right]$, $\mathbb{E}\left[\vec{Y}_{i,1}^{(1,.)}\right]$, $\mathbb{E}\left[\vec{Y}_{i,1}^{(-1,.)}\right]$, $\mathbb{E}\left[\vec{Y}_{i,2}^{(1,1)}\right]$, $\mathbb{E}\left[\vec{Y}_{i,2}^{(1,-1)}\right]$, $\mathbb{E}\left[\vec{Y}_{i,2}^{(-1,1)}\right]$, and $\mathbb{E}\left[\vec{Y}_{i,2}^{(-1,-1)}\right]$. For brevity, we use $\mu_{t,d}$ to denote $\mathbb{E}\left[\left(\vec{Y}_{i,2}^{(d)}\right)_j\right]$. I.e., choice of the aforementioned values induces γ^M parameters via:

$$\begin{bmatrix}\gamma_0^M\\ \gamma_1^M\\ \gamma_2^M\\ \gamma_3^M\\ \gamma_4^M\\ \gamma_5^M\\ \gamma_6^M\end{bmatrix} = \begin{bmatrix}1 & 0 & 0 & 0 & 0 & 0 & 0\\ 1 & 1 & 1 & 0 & 0 & 0 & 0\\ 1 & 1 & -1 & 0 & 0 & 0 & 0\\ 1 & 1 & 1 & 1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1 & 1 & -1 & -1\\ 1 & 1 & -1 & 1 & -1 & 1 & -1\\ 1 & 1 & -1 & 1 & -1 & -1 & 1\end{bmatrix}^{-1}\begin{bmatrix}\mu_{0,(.,.)}\\ \mu_{1,(1,.)}\\ \mu_{1,(-1,.)}\\ \mu_{2,(1,1)}\\ \mu_{2,(1,-1)}\\ \mu_{2,(-1,1)}\\ \mu_{2,(-1,-1)}\end{bmatrix}.$$

Furthermore, we wish to control the corresponding variance $\mathbb{V}\left[\vec{Y}_{i,t}^{(a_1,a_{2NR})} \mid X_i\right]$. We wish to generate a variance structure that is exchangeable in individuals and unstructured in time. I.e.,

- $\mathbb{V}\left[\vec{Y}_{i,t}{}^{(d)} \middle| X = X_i\right] = \left(\sigma_{t,d}^2 - \rho_{t,d}\right) I_{n_i} + \rho_{t,d}\vec{\mathbf{1}}_{n_i}\vec{\mathbf{1}}_{n_i}^T$ and
- $\mathrm{Cov}\left(\vec{Y}_{i,t_1}^{(d_i)}, \vec{Y}_{i,t_2}^{(d_i)} \mid X = X_i\right) = \left(\phi_{t_1,t_2,d} - \rho_{t_1,t_2,d}\right) I_{n_i} + \rho_{t_1,t_2,d}\vec{\mathbf{1}}_{n_i}\vec{\mathbf{1}}_{n_i}^T$ (for $t_1 \neq t_2$).

Specifying such a variance structure requires selecting the following values for distributions of $Y^{(d)}$ (con-

[1]Implicitly so, via the choice of several conditional parameters, as discussed below.

ditioned on X).

$$
\begin{aligned}
\sigma^2_{0,d} = \sigma^2_0 &- \text{The variance of the outcome at baseline for an individual.}\\
\sigma^2_{1,d} = \sigma^2_{1,a_1} &- \text{The variance of the outcome at time } t=1 \text{ under } A_1 = a_1 \text{ for an individual.}\\
\sigma^2_{2,d} = \sigma^2_{2,a_1,a_{2NR}} &- \text{The variance of the outcome at time } t=2 \text{ under } A_1 = a_1 \text{ and } A_{2NR} = a_{2NR} \text{ for an individual.}\\
\rho_{0,d} = \rho_0 &- \text{The intra-cluster covariance of the outcome at baseline.}\\
\rho_{1,d} = \rho_{1,a_1} &- \text{The intra-cluster covariance of the outcome at time } t=1 \text{ under } A_1 = a_1.\\
\rho_{2,d} = \rho_{2,a_1,a_{2NR}} &- \text{The intra-cluster covariance of the outcome at time } t=2 \text{ under } A_1 = a_1 \text{ and } A_{2NR} = a_{2NR}.\\
\phi_{0,1,d} = \phi_{0,1,a_1} &- \text{The intra-person covariance of the outcome, under } A_1 = a_1\text{, between times } t_1 = 0 \text{ and } t_2 = 1.\\
\phi_{0,2,d} = \phi_{t_1,2,a_1,a_{2NR}} &- \text{The intra-person covariance of the outcome, under } A_1 = a_1 \text{ and } A_{2NR} = a_{2NR},\\
&\quad \text{between times } t_1 = 0 \text{ and } t_2 = 2.\\
\phi_{1,2,d} = \phi_{t_1,2,a_1,a_{2NR}} &- \text{The intra-person covariance of the outcome, under } A_1 = a_1 \text{ and } A_{2NR} = a_{2NR},\\
&\quad \text{between times } t_1 = 1 \text{ and } t_2 = 2.\\
\rho_{0,1,d} = \rho_{0,1,a_1} &- \text{The inter-person covariance of the outcome, under } A_1 = a_1\text{, between times } t_1 = 0 \text{ and } t_2 = 1.\\
\rho_{0,2,d} = \rho_{t_1,2,a_1,a_{2NR}} &- \text{The inter-person covariance of the outcome, under } A_1 = a_1 \text{ and } A_{2NR} = a_{2NR},\\
&\quad \text{between times } t_1 = 0 \text{ and } t_2 = 2.\\
\rho_{1,2,d} = \rho_{t_1,2,a_1,a_{2NR}} &- \text{The inter-person covariance of the outcome, under } A_1 = a_1 \text{ and } A_{2NR} = a_{2NR},\\
&\quad \text{between times } t_1 = 1 \text{ and } t_2 = 2.
\end{aligned}
$$

Note that we require pooling across DTR for certain variance parameters, as implied by the notation above. E.g., $\sigma^2_{0,d} = \sigma^2_0$ implies a shared baseline variance of the outcome across all embedded DTRs (e.g., $\sigma^2_{0,d} = \sigma^2_{0,d'}$ for all pairs d, d'). This pooling is merely a consequence of the causal identifiability assumptions discussed in appendices to the aforementioned manuscript.

Furthermore, for a distribution parameter α, we use the notation $\alpha^{<r>}$ to denote the corresponding variance, conditioned on $R = r$. E.g., $\sigma^{2<1>}_{0,d}$ represents the variance of the outcome Y for an individual (conditioned on X) at time 2, given the unit was in a responding cluster. Similarly, $\sigma^{2<0>}_{0,d}$ represents this variance conditioned on the unit belonging to a nonresponding cluster. We also use $\mu^{<r>}_{t,d}$ to denote conditional mean parameters.

2 Data-Generative Model

In the subsections below, we describe a data-generative model designed to produce data under a specified marginal distribution. Here, we focus simulating data for a prototypical SMART; however, this model can be slightly modified to produce data under alternate SMART structures.

2.1 Arguments to Select

In designing a data-generative mechanism for a prototypical SMART consistent with the marginal structural model discussed above, we wanted to retain control over the following:

1. A target marginal mean and variance structure for the outcome $Y^{(d)}$ for all $d \in \mathcal{D}$, as well as corresponding mean/variance structures conditioned on response status (as discussed above). This requires specifying:

   (a) Pre-response marginal mean and variance structures. I.e., μ_0, μ_{1,a_1}, σ^2_0, σ^2_{1,a_1}, ρ_0, ρ_{1,a_1}, $\phi_{0,1,a_1}$, and $\rho_{0,1,a_1}$ for $a_1 = \pm 1$. As discussed below, these choices will induce corresponding parameters conditional on response (which we derive via Monte-Carlo, as in Seewald et al. [2020]).

   (b) Post-response marginal mean and variance structures. I.e., $\mu_{2,d}$, $\sigma^2_{2,d}$, $\rho_{2,d}$, $\phi_{0,2,d}$, $\phi_{1,2,d}$, $\rho_{0,2,d}$, and $\rho_{1,2,d}$ for $d \in \mathcal{D}$.

(c) Post-response conditional mean and variance structures. I.e., $\mu_{2,d}^{<r>}$, $\sigma_{2,d}^{2<r>}$, $\rho_{2,d}^{<r>}$, $\phi_{0,2,d}^{<r>}$, $\phi_{1,2,d}^{<r>}$, $\rho_{0,2,d}^{<r>}$, and $\rho_{1,2,d}^{<r>}$ for $d \in \mathcal{D}$ and $r \in \{0,1\}$. As discussed below, these choices must be consistent with the choice of marginal post-response variance.

2. The probability of response under each first-stage treatment assignment. I.e., $\mathbb{P}\left[R(a_1) = 1 \mid X\right] =: p_{r,a_1}$ for $a_1 = \pm 1$.
3. The treatment assignment probabilities (e.g., $\mathbb{P}\left[A_1 = 1\right]$ and $\mathbb{P}\left[A_{2NR} = 1 \middle| A_1\right]$).[2]
4. The distribution of covariates (provided the distribution has mean zero) and their influence on the outcome (η).
5. The number of clusters, K.
6. The distribution of N_i, local sample size (provided the distribution is over $\mathbb{N}$ and independent of outcomes).

Therefore, the analyst must specify the above quantities, ideally to approximate a real-world scenario relevant to the analysis in question. As discussed below, conditional variance components must be chosen to be consistent with the selected marginal variance components. Additionally, we require that variance parameters are chosen such that all resulting covariance matrices are positive definite.

2.2 Marginal and Conditional Specification Consistency

As discussed above, we wish to simulate data conditionally (i.e., conditioned on response). Doing so more closely resembles real-world processes, and allows the user to specify stark differences between responders and non-responders in order to test estimator performance.

In Step 1 above, the user must select the marginal and conditional distribution of $\vec{Y}^{(d)}$ for all d, as well as the probability of response under $A_1 = 1$ and $A_1 = -1$. However, for a given DTR d (with probability of response p_r), we note that, for any $j = 1, \ldots, N_i$ and $t_1, t_2 \in \{0, 1, 2\}$

$$\mathbb{E}\left[\left(\vec{Y}_{i,t}{}^{(d)}\right)_j \mid X_i\right] = p_r \mathbb{E}\left[\left(\vec{Y}_{i,t}{}^{(d)}\right)_j \mid X_i, R_i = 1\right] + (1-p_r)\mathbb{E}\left[\left(\vec{Y}_{i,t}{}^{(d)}\right)_j \mid X_i, R_i = 0\right]$$

and

$$\begin{aligned}
\mathrm{Cov}\left(\left(\vec{Y}_{i,t_1}\right)_j, \left(\vec{Y}_{i,t_2}\right)_k \mid X_i\right) &= \mathbb{E}\left[\mathrm{Cov}\left(\left(\vec{Y}_{i,t_1}{}^{(d)}\right)_j, \left(\vec{Y}_{i,t_2}{}^{(d)}\right)_k \mid X_i, R_i\right)\right] \\
&\quad + \mathrm{Cov}\left(\mathbb{E}\left[\left(\vec{Y}_{i,t_1}{}^{(d)}\right)_j \middle| R_i = r\right], \mathbb{E}\left[\left(\vec{Y}_{i,t_2}{}^{(d)}\right)_k \mid X_i, R_i\right]\right) \\
&= \mathbb{E}\left[\mathbb{1}_{R_i=1}\mathrm{Cov}\left(\left(\vec{Y}_{i,t_1}{}^{(d)}\right)_j, \left(\vec{Y}_{i,t_2}{}^{(d)}\right)_k \mid X_i, R_i = 1\right)\right] \\
&\quad + \mathbb{E}\left[\mathbb{1}_{R_i=0}\mathrm{Cov}\left(\left(\vec{Y}_{i,t_1}{}^{(d)}\right)_j, \left(\vec{Y}_{i,t_2}{}^{(d)}\right)_k \mid X_i, R_i = 0\right)\right] \\
&\quad + \mathrm{Cov}\left(\mathbb{1}_{R=0}\mu_{t_1,d}^{<0>} + (1 - \mathbb{1}_{R=0})\,\mu_{t_1,d}^{<1>},\ \mathbb{1}_{R=0}\mu_{t_2,d}^{<0>} + (1 - \mathbb{1}_{R=0})\,\mu_{t_2,d}^{<1>}\right) \\
&= p_r \mathrm{Cov}\left(\left(\vec{Y}_{i,t_1}\right)_j, \left(\vec{Y}_{i,t_2}\right)_k \mid X_i, R_i = 1\right) + (1-p_r)\,\mathrm{Cov}\left(\left(\vec{Y}_{i,t_1}\right)_j, \left(\vec{Y}_{i,t_2}\right)_k \mid X_i, R_i = 0\right) \\
&\quad + \left(\mu_{t_1,d}^{<0>} - \mu_{t_1,d}^{<1>}\right)\left(\mu_{t_2,d}^{<0>} - \mu_{t_2,d}^{<1>}\right)\left(p_r\left(1-p_r\right)\right).
\end{aligned}$$

Subsequently, the distributions of $\vec{Y}^{(d)}\big|_X$, $\vec{Y}^{(d)}\big|_{X,R=1}$, and $\vec{Y}^{(d)}\big|_{X,R=0}$ cannot vary freely, as choosing two implies the third. We specified the marginal distribution as well as the distribution conditional on positive

[2]These are usually taken to be 0.5; however, an analyst may want to simulate a SMART under imbalanced randomization probabilities.

response.[3]

Lastly, as discussed below, specifying a marginal pre-response distribution (i.e., the distribution of $\begin{bmatrix} Y_0{}^{(d)} \\ Y_1{}^{(d)} \end{bmatrix} \big|_X$) and a response-generation mechanism induces the conditional distribution of $\begin{bmatrix} Y_0{}^{(d)} \\ Y_1{}^{(d)} \end{bmatrix} \big|_{X,R}$. Analytically deriving this distribution for non-trivial models of response proved intractable, and subsequently we used Monte-Carlo methods to estimate necessary conditional mean and variance parameters, as in Seewald et al. [2020].

2.3 Data-Generative Process

Given the chosen parameters discussed above, we propose the algorithm to produce data which targets the desired marginal structural model. We note the choices above require specification of all $\mu_{\cdot}$, $\mu_{\cdot}^{<r>}$, $\sigma_{\cdot}$, $\sigma_{\cdot}^{<r>}$, $\rho_{\cdot}$, $\rho_{\cdot}^{<r>}$, $\phi_{\cdot}$, and $\phi_{\cdot}^{<r>}$ terms, in addition to all response probabilities and treatment assignment probabilities. Choice of these terms define the simulation parameters used below (as we make explicit later in this section). In the model described below, we use $[r]$ notation to denote simulation parameters that depend on response status (and note that these only apply to post-response parameters). Recall the use of $< r >$ notation to denote parameters regarding the distribution of $\vec{Y}^{(d)}\big|_{R=r}$.

We carry out the algorithm below for $i = 1, \ldots, K$ (taking independent draws for each cluster):

[3]We specified the marginal distribution in order to directly control the estimands we sought to estimate (all of which are marginal). We specified the distribution of responders because this distribution should not depend on the unrealized treatment assignment had the cluster not responded. E.g., suppose one were to specify the conditional distribution of $Y_2{}^{(d)}$ for non-responders. This would induce distributions for $Y_2{}^{(d)}\big|_{R=1}$ for each $d \in \mathcal{D}$. However, this would mean $Y_2^{(a_1,a_{2NR})}\big|_{R=1}$ could have a different distribution than $Y_2^{(a_1,a'_{2NR})}\big|_{R=1}$, even though our consistency assumption would say they must be the same.

Algorithm 1 `Data-Generative Algorithm`

1: Simulate $d_i = (A_{1,i}, A_{2NR,i})$ using the treatment assignment probabilities specified in Argument 3. We use $(a_{1,i}, a_{2NR,i})$ to denote simulated treatment assignments.
2: Simulate $N_i = n_i$, drawing it from the distribution specified in Argument 6.
3: Simulate X_i, drawing it from the distribution specified in Argument 4.
4: Simulate $\begin{bmatrix} \vec{\varepsilon}_{i,0} \\ \vec{\varepsilon}^{(a_{1,i})}_{i,1} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \vec{0}_{n_i} \\ \vec{0}_{n_i} \end{bmatrix}, \ \left[\begin{array}{c|c} \Sigma_0 & \left(\phi_{0,1} - P_{1,a_{1,i}}\sigma_0^2\right) I_{n_i} \\ \hline \left(\phi_{0,1} - P_{1,a_{1,i}}\sigma_0^2\right) I_{n_i} & \Sigma_{1,a_{1,i},n_i} \end{array} \right] \right).$
5: Take $\vec{Y}^{(d_i)}_{i,0}\big|_{X_i} = \eta X_i + \gamma_0 \vec{\mathbf{1}}_{n_i} + \vec{\varepsilon}_{i,0}$.
6: Take $\vec{Y}^{(d_i)}_{i,1}\big|_{X_i, \vec{Y}^{(d_i)}_{i,0}} = \left(1 - P_{1,a_{1,i}}\right)\left(\eta X_i + \gamma_0 \vec{\mathbf{1}}_{n_i}\right) + \gamma_1 \vec{\mathbf{1}}_{n_i} + P_{1,a_{1,i}} \vec{Y}^{(d_i)}_{i,0} + \gamma_2 a_{1,i} \vec{\mathbf{1}}_{n_i} + \vec{\varepsilon}^{(a_{1,i})}_{i,1}$.
7: Simulate $R^{(a_{1,i})}_i$ as a draw from a $B\left(g\left(\begin{bmatrix} \vec{Y}^{(d_i)}_{i,0} \\ \vec{Y}^{(d_i)}_{i,1} \end{bmatrix}\right)\right)$ distribution, taking g to be the response propensity mechanism discussed below. We denote this realized value of response as r_i.
8: Simulate $\vec{\varepsilon}^{(d_i,r_i)}_{i,2} \sim \mathcal{N}\left(\vec{0}_{n_i}, \ \Sigma^{[r_i]}_{2,d_i,n_i}\right)$, with $\vec{\varepsilon}^{(d_i,r_i)}_{i,2} \perp\!\!\!\perp \begin{bmatrix} \vec{\varepsilon}_{i,0} \\ \vec{\varepsilon}^{(a_{1,i})}_{i,1} \end{bmatrix}$.
9: Take

$$
\begin{aligned}
\vec{Y}^{(d_i)}_{i,2}\big|_{X_i, \vec{Y}^{(d_i)}_{i,0}, \vec{Y}^{(d_i)}_{i,1}, R^{(a_{1,i})}_i = r_i} = & \left(1 - P^{[r_i]}_{2,d_i,n_i} - P^{[r_i]}_{4,d_i,n_i}\right)\left(\eta X_i + \gamma_0 \vec{\mathbf{1}}_{n_i}\right) \\
& + \left(1 - P^{[r_i]}_{4,d_i,n_i}\right)\left(\gamma_1 + \gamma_2 a_{1,i}\right)\vec{\mathbf{1}}_{n_i} + \left(\gamma_3 + \gamma_4 a_{1,i}\right)\vec{\mathbf{1}}_{n_i} \\
& + P^{[r_i]}_{2,d_i,n_i} \vec{Y}_{i,0} + P^{[r_i]}_{4,d_i,n_i} \vec{Y}_{i,1} \\
& + \left(P^{[r_i]}_{3,d_i,n_i} \frac{1}{n_i} \sum_{j=1}^{n_i} \left(\vec{\varepsilon}_{i,0}\right)_j\right)\vec{\mathbf{1}}_{n_i} + \left(P^{[r_i]}_{5,d_i,n_i} \frac{1}{n_i} \sum_{j=1}^{n_i} \left(\vec{\varepsilon}_{i,1}\right)_j\right)\vec{\mathbf{1}}_{n_i} \\
& + \left(\left(\gamma_5 + \gamma_6 a_{1,i}\right) a_{2NR,i}\right)\left(\frac{1 - r_i}{1 - p_{r_{a_{1,i}}}}\right)\vec{\mathbf{1}}_{n_i} \\
& + \left(\lambda_1 + \lambda_2 a_{1,i}\right)\left(\left(r_i - p_{r_{a_{1,i}}}\right)\vec{\mathbf{1}}_{n_i}\right) \\
& + \vec{\varepsilon}^{(r_i)}_{i,2}.
\end{aligned}
$$

We consider the following model of response:

$$
g\left(\begin{bmatrix} X_i \\ \vec{Y}_{i,0} \\ \vec{Y}_{i,1} \end{bmatrix}\right) = F^{\leftarrow}_{p_{r_{a_1}}}\left(\mathbf{\Phi}\left(\frac{\left(\frac{1}{n}\sum_{j=1}^{n}\left(\vec{Y}_{i,1}\right)_j - \left(\eta X_i + \gamma_0 + \gamma_1 + \gamma_2 a_1\right)\right)}{\sqrt{\frac{1}{n}\left(\sigma_1^2 + (n-1)\rho_{1,a_1}\right)}}\right)\right)
$$

where $F^{\leftarrow}_{p_{r_{a_1}}}$ is the inverse CDF of the Beta$\left(\frac{p_{r_{a_1}}}{1 - p_{r_{a_1}}}, 1\right)$ distribution. Note that we force response to be uncorrelated with baseline covariates.

We note that, by construction of g, response propensity is solely a function of $\begin{bmatrix} \vec{\varepsilon}_0 \\ \vec{\varepsilon}^{(a_1)}_1 \end{bmatrix}$. Therefore, the distribution of $\begin{bmatrix} \vec{Y}^{(d)}{}_0 \\ \vec{Y}^{(d)}{}_1 \end{bmatrix}\big|_R$ can be directly derived via the distribution of $\begin{bmatrix} \vec{\varepsilon}_0 \\ \vec{\varepsilon}^{(a_1)}_1 \end{bmatrix}\big|_R$. As discussed above, we used Monte-Carlo techniques to derive this distribution under various settings. Consequently, we use the following notation:

$$
\begin{aligned}
\epsilon_{0,n}^{<r>} &:= \mathbb{E}\left[\left(\vec{\varepsilon}_0\right)_j \mid R=r, N=n\right] \\
\epsilon_{1,n}^{<r>} &:= \mathbb{E}\left[\left(\vec{\varepsilon}_1\right)_j \mid R=r, N=n\right] \\
s_{0,n}^{2<r>} &:= \mathbb{V}\left[\left(\vec{\varepsilon}_0\right)_j \mid R=r, N=n\right] \\
s_{1,n}^{2<r>} &:= \mathbb{V}\left[\left(\vec{\varepsilon}_1\right)_j \mid R=r, N=n\right] \\
c_{0,n}^{<r>} &:= \operatorname{Cov}\left(\left(\vec{\varepsilon}_0\right)_j, \left(\vec{\varepsilon}_0\right)_k \mid R=r, N=n\right) \\
c_{1,n}^{<r>} &:= \operatorname{Cov}\left(\left(\vec{\varepsilon}_1\right)_j, \left(\vec{\varepsilon}_1\right)_k \mid R=r, N=n\right) \\
s_{0,1,n}^{<r>} &:= \operatorname{Cov}\left(\left(\vec{\varepsilon}_0\right)_j, \left(\vec{\varepsilon}_1\right)_j \mid R=r, N=n\right) \\
c_{0,1,n}^{<r>} &:= \operatorname{Cov}\left(\left(\vec{\varepsilon}_0\right)_j, \left(\vec{\varepsilon}_1\right)_k \mid R=r, N=n\right),
\end{aligned}
$$

where $j,k = 1,\ldots,n$ and $j \neq k$. We note that the exchangeability of the distributions of $\vec{\varepsilon}_0$ and $\vec{\varepsilon}_1^{(a_1)}$ ensure well-posedness of the variables described above (i.e., invariance with respect to choice of j,k).

Here, we take

$$
\begin{aligned}
P_{1,a_1} &:= \begin{cases} \dfrac{\rho_{0,1,a_1}}{\rho_0} & \text{if } \rho_0 \neq 0 \\ 0 & \text{if } \rho_0 = 0, \end{cases} \\
\Sigma_0 &:= \left(\sigma_0^2 - \rho_0\right) I_n + \rho_0 \vec{\mathbf{1}}_{n_i} \vec{\mathbf{1}}_{n_i}^T, \\
\Sigma_{1,a_1,n} &:= \left(\left(\sigma_{1,a_1}^2 - \rho_{1,a_1}\right) I_n + \rho_{1,a_1} \vec{\mathbf{1}}_{n_i} \vec{\mathbf{1}}_{n_i}^T\right) - P_{1,a_1}^2 \Sigma_0 - 2P_{1,a_1}\left(\phi_{0,1,a_1} - P_{1,a_1}\sigma_0^2\right) I_n, \\
\Sigma_{2,d,n}^{[r]} &:= \left(\left(\sigma_{2,d}^{2<r>} - \zeta_{d,n}\right) - \left(\rho_{2,d}^{<r>} - \zeta_{d,n}'\right)\right) I_n + \left(\rho_{2,d}^{<r>} - \zeta_{d,n}'\right) \vec{\mathbf{1}}_{n_i} \vec{\mathbf{1}}_{n_i}^T,
\end{aligned}
$$

and

$$
\begin{bmatrix} P_{2,d,n}^{[r]} \\ P_{3,d,n}^{[r]} \\ P_{4,d,n}^{[r]} \\ P_{5,d,n}^{[r]} \end{bmatrix} := \Upsilon_{a_1,n}^{-1} \begin{bmatrix} \phi_{0,2,d}^{<r>} \\ \phi_{1,2,d}^{<r>} \\ \rho_{0,2,d}^{<r>} \\ \rho_{1,2,d}^{<r>} \end{bmatrix},
$$

where $\zeta_{d,n}, \zeta_{d,n}'$, and $\Upsilon_{a_1,n}$ are defined in Section 4 below. Note that the construction of Σ_0 ensures that all the distribution of $\vec{Y}_0{}^{(d)}$ is identical for all $d \in \mathcal{D}$.

We also note that the γ terms above are related, but not identical, to the marginal γ^M terms in the target structural mean model. We let

$$
\vec{\gamma} := \begin{bmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \\ \gamma_5 \\ \gamma_6 \\ \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 1 & 0 & 0 & (1-p_{r,1}) & (1-p_{r,1}) \\
1 & 1 & 1 & 1 & 1 & \frac{1}{1-p_{r,1}} & \frac{1}{1-p_{r,1}} & -p_{r,1} & -p_{r,1} \\
1 & 1 & 1 & 1 & 1 & -\frac{1}{1-p_{r,1}} & -\frac{1}{1-p_{r,1}} & -p_{r,1} & -p_{r,1} \\
1 & 1 & -1 & 1 & -1 & 0 & 0 & (1-p_{r,-1}) & -(1-p_{r,-1}) \\
1 & 1 & -1 & 1 & -1 & \frac{1}{1-p_{r,-1}} & -\frac{1}{1-p_{r,-1}} & -p_{r,-1} & p_{r,-1} \\
1 & 1 & -1 & 1 & -1 & -\frac{1}{1-p_{r,1}} & \frac{1}{1-p_{r,1}} & -p_{r,-1} & p_{r,-1}
\end{bmatrix}^{-1} \begin{bmatrix}
\mu_{0,(.,.)} \\
\mu_{1,(1,.)} \\
\mu_{1,(-1,.)} \\
\mu_{2,(1,1)}^{<1>} - v_{(1,1),1,n} \\
\mu_{2,(1,1)}^{<0>} - v_{(1,1),0,n} \\
\mu_{2,(1,-1)}^{<0>} - v_{(1,-1),0,n} \\
\mu_{2,(-1,1)}^{<1>} - v_{(-1,1),1,n} \\
\mu_{2,(-1,1)}^{<0>} - v_{(-1,1),0,n} \\
\mu_{2,(-1,-1)}^{<0>} - v_{(-1,-1),0,n}
\end{bmatrix},
$$

where $v_{d,r,n} := P_{2,d,n}^{[r]} \epsilon_{0,n}^{<r>} + P_{3,d,n}^{[r]} \epsilon_{0,n}^{<r>} + P_{4,d,n}^{[r]} \left(P_{1,a_1} \epsilon_{0,n}^{<r>} + \epsilon_{1,n}^{<r>}\right) + P_{5,d,n}^{[r]} \epsilon_{1,n}^{<r>}$.

We briefly note our introduction of cross-temporal cluster spillover into $\vec{Y}_{i,2}^{(d_i)}$ via the inclusion of the $P_{3,d,n}^{[r_i]} \frac{1}{n} \sum_{j=1}^{n} (\vec{\varepsilon}_{i,0})_j$ and $P_{5,d,n}^{[r_i]} \frac{1}{n} \sum_{j=1}^{n} (\vec{\varepsilon}_{i,1})_j$ terms. We do so in order to control cross-temporal intra-cluster covariances. Simulations under which $P_{3,d,n}^{[r_i]} = 0$ and/or $P_{5,d,n}^{[r_i]} = 0$ correspond to situations in which there is no direct cross-temporal intra-cluster spillover effects.

Proof of Marginal Consistency

We will consider the simulation of the outcome trajectory of a cluster of size $N = n$, with covariate data X, under a given DTR $d = (A_1 = a_1, A_{2NR} = a_{2NR})$, where N, X, and (A_1, A_{2NR}) are simulated according to the data-generative algorithm laid out above. In this subsection, we prove that the distribution of such an outcome trajectory has the distribution specified above. This subsection will be laid out as follows:

- Remarks on notational conventions henceforth adopted to promote conciseness.
- Proof that the pre-response marginal distribution of the data-generative model matches that of the target distribution.
- Proof that the response probability of the data-generative model matches that of the target model.
- Proof that the post-response conditional distributions of the data-generative model match those of the target distribution.

We note that the post-response marginal distribution of the data-generative model is implied by the response probability and post-response conditional distributions of the data generative model, as discussed above. Therefore, proving the three characteristics listed above will guarantee that the marginal distribution of the data-generative model matches the specified target distribution.

Notational Remarks

The following arguments employ heavy use of sub/superscripts. Given that we are restricting consideration to one sample size (n) and one embedded DTR $d = (a_1, a_{2NR})$, in order to promote readability in this section we will henceforth omit subscripts identifying parameters that denote dependence on sample size and DTR assignment.

I.e., we are removing subscripts signifying the dependence of various parameters on treatment assignment and local sample size. This is purely for the purpose of conciseness, and the reader should note that these parameters are still indexed by treatment assignment and local sample size. We also note that the independence of N from other random values in our simulation allows us to discuss the results below without explicitly conditioning on N. For example, we will use $\phi_{0,2}$ rather than $\phi_{0,2,d}$ to denote the user-specified value for $\mathrm{Cov}\left(\left(\vec{Y}_0{}^{(d)}\right)_j, \left(\vec{Y}_2{}^{(d)}\right)_j \mid X\right)$. Similarly, we use $P_5^{[r]}$ as a shorthand for $P_{5,d,n}^{[r]}$. We will also note that we are simulating the potential outcomes under the DTR d, but will use $\vec{Y}_{\cdot}$ and R in lieu of $\vec{Y}_{\cdot}{}^{(d)}$, $R_i^{(a_{1,i})}$.

Marginal Distribution of Chosen Data-Generative Model - Pre-Response Outcomes

We recall that we began by simulating $\vec{Y}_0$ and $\vec{Y}_1|_{\vec{Y}_0}$ as follows:

$$\begin{aligned}
\vec{Y}_0\big|_X &= \eta X + \gamma_0 \vec{\mathbf{1}}_n + \vec{\varepsilon}_0 \\
\vec{Y}_1\big|_{X,\vec{Y}_0} &= (1 - P_1)\left(\eta X + \gamma_0 \vec{\mathbf{1}}_n\right) + \gamma_1 \vec{\mathbf{1}}_n + P_1 \vec{Y}_0 + \gamma_2 a_1 \vec{\mathbf{1}}_n + \vec{\varepsilon}_1,
\end{aligned}$$

where

$$P_1 = \begin{cases} \dfrac{\rho_{0,1}}{\rho_0} & \text{if } \rho_0 \neq 0 \\ 0 & \text{if } \rho_0 = 0. \end{cases}$$

We generated $\vec{\varepsilon}_0$ and $\vec{\varepsilon}_1$ jointly, with:

$$\begin{bmatrix}\vec{\varepsilon}_0 \\ \vec{\varepsilon}_1\end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix}\vec{0}_n \\ \vec{0}_n\end{bmatrix},\ \left[\begin{array}{c|c}\Sigma_0 & \left(\phi_{0,1} - P_1\sigma_0^2\right) I_n \\ \hline \left(\phi_{0,1} - P_1\sigma_0^2\right) I_n & \Sigma_1\end{array}\right]\right),$$

where

$$\Sigma_0 = \left(\sigma_0^2 - \rho_0\right) I_n + \rho_0 \vec{\mathbf{1}}_n \vec{\mathbf{1}}_n^T \text{ and}$$
$$\Sigma_1 = \left(\left(\sigma_1^2 - \rho_1\right) I_n + \rho_1 \vec{\mathbf{1}}_n \vec{\mathbf{1}}_n^T\right) - P_1^2 \Sigma_0 - 2P_1\left(\phi_{0,1} - P_1\sigma_0^2\right) I_n.$$

We immediately observe that $\mathbb{E}\left[\vec{Y}_0 \mid X\right] = \eta X + \gamma_0 \vec{\mathbf{1}}_n$ (and thus $\mathbb{E}\left[\vec{Y}_0\right] = \gamma_0 \vec{\mathbf{1}}_n = \mu_{0,(.,.)} \vec{\mathbf{1}}_n$) and $\mathbb{V}\left[\vec{Y}_0 \mid X\right] = \left(\sigma_0^2 - \rho_0\right) I_n + \rho_0 \vec{\mathbf{1}}_n \vec{\mathbf{1}}_n^T$, satisfying the requirements for the marginal distribution of $\vec{Y}_0$. Additionally,

$$\begin{aligned}\mathbb{E}\left[\vec{Y}_1 \mid X\right] &= (1 - P_1)\left(\eta X + \gamma_0 \vec{\mathbf{1}}_n\right) + \gamma_1 \vec{\mathbf{1}}_n + P_1 \mathbb{E}\left[\vec{Y}_0 \mid X\right] + \gamma_2 a_1 \vec{\mathbf{1}}_n \\ &= \eta X + \left(\gamma_0 + \gamma_1 + \gamma_2 a_1\right) \vec{\mathbf{1}}_n,\end{aligned}$$

so

$$\begin{aligned}\mathbb{E}\left[\vec{Y}_1\right] &= \left(\gamma_0 + \gamma_1 + \gamma_2 a_1\right) \vec{\mathbf{1}}_n \\ &= \mu_{1,(a_1,.)} \vec{\mathbf{1}}_n \text{ (by construction of } \vec{\boldsymbol{\gamma}}).\end{aligned}$$

Furthermore,

$$\begin{aligned}\mathbb{V}\left[\vec{Y}_1 \mid X\right] &= \mathbb{V}\left[P_1 \vec{\varepsilon}_0 + \vec{\varepsilon}_1 \mid X\right] \\ &= P_1^2 \mathbb{V}\left[\vec{Y}_0 \mid X\right] + 2P_1 \operatorname{Cov}\left(\vec{\varepsilon}_0, \vec{\varepsilon}_1\right) + \mathbb{V}\left[\vec{\varepsilon}_1 \mid X\right] \\ &= P_1^2 \Sigma_0 + 2P_1\left(\phi_{0,1} - P_1\sigma_0^2\right) I_n + \Sigma_1 \\ &= P_1^2 \Sigma_0 + 2P_1\left(\phi_{0,1} - P_1\sigma_0^2\right) I_n + \left(\left(\left(\sigma_1^2 - \rho_1\right) I_n + \rho_1 \vec{\mathbf{1}}_n \vec{\mathbf{1}}_n^T\right) - P_1^2 \Sigma_0 - 2P_1\left(\phi_{0,1} - P_1\sigma_0^2\right) I_n\right) \\ &= \left(\sigma_1^2 - \rho_1\right) I_n + \rho_1 \vec{\mathbf{1}}_n \vec{\mathbf{1}}_n^T,\end{aligned}$$

satisfying the requirements for the marginal distribution of $\vec{Y}_1$. Finally,

$$\begin{aligned}\operatorname{Cov}\left(\vec{Y}_0, \vec{Y}_1 \mid X\right) &= \operatorname{Cov}\left(\vec{\varepsilon}_0, P_1 \vec{\varepsilon}_0 + \vec{\varepsilon}_1\right) \\ &= P_1 \Sigma_0 + \left(\phi_{0,1} - P_1\sigma_0^2\right) I_n \\ &= P_1\left(\left(\sigma_0^2 - \rho_0\right) I_n + \rho_0 \vec{\mathbf{1}}_n \vec{\mathbf{1}}_n^T\right) + \left(\phi_{0,1} - P_1\sigma_0^2\right) I_n \\ &= \left(\phi_{0,1} - P_1\rho_0\right) I_n + P_1\rho_0 \vec{\mathbf{1}}_n \vec{\mathbf{1}}_n^T.\end{aligned}$$

Therefore, by construction of P_1, $\operatorname{Cov}\left(\vec{Y}_0, \vec{Y}_1 \mid X\right) = \left(\phi_{0,1} - \rho_{0,1}\right) I_n + \rho_{0,1} \vec{\mathbf{1}}_n \vec{\mathbf{1}}_n^T$, satisfying our requirements for the temporal covariance between $\vec{Y}_0$ and $\vec{Y}_1$.

Marginal Distribution of Chosen Data-Generative Model - Response

In response model g_1 above, we centered and standardized $\bar{\vec{Y}}_1$ with respect to the (known) true mean and standard deviation. We then applied a inverse-CDF transform to simulate response likelihood via a $\operatorname{Beta}\left(\frac{p_r}{1 - p_r}, 1\right)$ distribution, guaranteeing a marginal likelihood of response of p_r.[4] I.e., we have satisfied $\mathbb{E}\left[R\right] = p_r$.

[4]And ensuring that $g_1\left(\begin{bmatrix}\vec{Y}_0 \\ \vec{Y}_1\end{bmatrix}\right) \in (0, 1)$.

Conditional Distributions of Chosen Data-Generative Model - Post-Response Outcome Simulation

Recall that we generated $\vec{Y}_2$ according to the following data-generative procedure:

$$\begin{aligned}\vec{Y}_2\big|_{X,\vec{Y}_0,\vec{Y}_1,R=r} = {} & \left(1 - P_2^{[r]} - P_4^{[r]}\right)\left(\eta X + \gamma_0 \vec{\mathbf{1}}_n\right) + \left(1 - P_4^{[r]}\right)(\gamma_1 + \gamma_2 a_1)\,\vec{\mathbf{1}}_n + (\gamma_3 + \gamma_4 a_1)\,\vec{\mathbf{1}}_n \\ & + P_2^{[r]}\vec{Y}_0 + P_3^{[r]}\frac{1}{n}\sum_{j=1}^{n}(\vec{\varepsilon}_0)_j\,\vec{\mathbf{1}}_n + P_4^{[r]}\vec{Y}_1 + P_5^{[r]}\frac{1}{n}\sum_{j=1}^{n}(\vec{\varepsilon}_1)_j\,\vec{\mathbf{1}}_n \\ & + ((\gamma_5 + \gamma_6 a_1)\,a_{2NR})\left(\frac{1-r}{1-p_r}\right)\vec{\mathbf{1}}_n \\ & + (\lambda_1 + \lambda_2 a_1)\left((r - p_r)\,\vec{\mathbf{1}}_n\right) \\ & + \vec{\varepsilon}_2^{(r)},\end{aligned}$$

where $\vec{\varepsilon}_2^{(r)} \perp\!\!\!\perp \begin{bmatrix}\vec{\varepsilon}_0\\ \vec{\varepsilon}_1\end{bmatrix}\big|_{R=r}$ and

$$\vec{\varepsilon}_2^{(r)} \sim \mathcal{N}\left(\vec{0}_n,\ \Sigma_2^{[r]}\right).$$

This setup ensures

$$\begin{aligned}\mathbb{E}\left[\vec{Y}_2 \mid X, R=r\right] = {} & \left(1 - P_2^{[r]} - P_4^{[r]}\right)\left(\eta X + \gamma_0 \vec{\mathbf{1}}_n\right) + \left(1 - P_4^{[r]}\right)(\gamma_1 + \gamma_2 a_1)\,\vec{\mathbf{1}}_n + (\gamma_3 + \gamma_4 a_1)\,\vec{\mathbf{1}}_n \\ & + P_2^{[r]}\mathbb{E}\left[\vec{Y}_0 \mid X, R=r\right] + P_3^{[r]}\mathbb{E}\left[\vec{\varepsilon}_0 \mid X, R=r\right] \\ & + P_4^{[r]}\mathbb{E}\left[\vec{Y}_1 \mid X, R=r\right] + P_5^{[r]}\mathbb{E}\left[\vec{\varepsilon}_1 \mid X, R=r\right] \\ & + ((\gamma_5 + \gamma_6 a_1)\,a_{2NR})\left(\frac{1-r}{1-p_r}\right)\vec{\mathbf{1}}_n \\ & + (\lambda_1 + \lambda_2 a_1)\left((r - p_r)\,\vec{\mathbf{1}}_n\right) \\ = {} & \eta X + (\gamma_0 + \gamma_1 + \gamma_2 a_1 + \gamma_3 + \gamma_4 a_1)\,\vec{\mathbf{1}}_n \\ & + P_2^{[r]}\mathbb{E}\left[\vec{\varepsilon}_0 \mid R=r\right] + P_3^{[r]}\mathbb{E}\left[\vec{\varepsilon}_0 \mid R=r\right] \\ & + P_4^{[r]}\mathbb{E}\left[P_1\vec{\varepsilon}_0 + \vec{\varepsilon}_1 \mid R=r\right] + P_5^{[r]}\mathbb{E}\left[\vec{\varepsilon}_1 \mid R=r\right] \\ & + ((\gamma_5 + \gamma_6 a_1)\,a_{2NR})\left(\frac{1-r}{1-p_r}\right)\vec{\mathbf{1}}_n \\ & + (\lambda_1 + \lambda_2 a_1)\left((r - p_r)\,\vec{\mathbf{1}}_n\right) \\ = {} & \eta X + (\gamma_0 + \gamma_1 + \gamma_2 a_1 + \gamma_3 + \gamma_4 a_1)\,\vec{\mathbf{1}}_n \\ & + \left((\gamma_5 + \gamma_6 a_1)\,a_{2NR}\left(\frac{1-r}{1-p_r}\right) + (\lambda_1 + \lambda_2 a_1)(r - p_r)\right)\vec{\mathbf{1}}_n \\ & + \left(P_2^{[r]}\epsilon_0^{<r>} + P_3^{[r]}\epsilon_0^{<r>} + P_4^{[r]}\left(P_1\epsilon_0^{<r>} + \epsilon_1^{<r>}\right) + P_5^{[r]}\epsilon_1^{<r>}\right).\end{aligned}$$

By construction of $\vec{\boldsymbol{\gamma}}$, the above resolves to

$$\begin{aligned}\mathbb{E}\left[\vec{Y}_2 \mid X, R=r\right] = {} & \eta X + \left(\mu_{2,d}^{<r>} - v_r\right)\vec{\mathbf{1}}_n \\ & + \underbrace{\left(P_2^{[r]}\epsilon_0^{<r>} + P_3^{[r]}\epsilon_0^{<r>} + P_4^{[r]}\left(P_1\epsilon_0^{<r>} + \epsilon_1^{<r>}\right) + P_5^{[r]}\epsilon_1^{<r>}\right)}_{=:v_r}\vec{\mathbf{1}}_n.\end{aligned}$$

Therefore,

$$\mathbb{E}\left[\vec{Y}_2 \mid X, R=r\right] = \eta X + \mu_{2,d}^{<r>}\vec{\mathbf{1}}_n$$

and

$$\mathbb{E}\left[\vec{Y}_2 \mid R=r\right] = \mu_{2,d}^{<r>}\vec{\mathbf{1}}_n.$$

The above result guarantees that the response-conditional means of the generated $\vec{Y}_2$ match those specified in the target distribution. This result, combined with the consistency of the data-generative and target response probabilities, guarantees that the data-generative model produces trajectories with the same marginal mean as the specified target distribution.

We now turn to the response-conditional variance components. We first examine the cross-response conditional variance components (i.e., $\rho_{0,2}^{<r>}$, $\rho_{1,2}^{<r>}$, $\phi_{0,2}^{<r>}$, and $\phi_{1,2}^{<r>}$). Recall that we obtained the following parameters via Monte-Carlo simulation:

$$\mathbb{V}\left[\begin{bmatrix}\vec{\varepsilon}_0\\ \vec{\varepsilon}_1\end{bmatrix} \mid X, R=r\right] = \left[\begin{array}{c|c}\left(s_0^{2<r>} - c_0^{<r>}\right) I_n + c_0^{<r>}\vec{\mathbf{1}}_n\vec{\mathbf{1}}_n^T & \left(s_{0,1}^{<r>} - c_{0,1}^{<r>}\right) I_n + c_{0,1}^{<r>}\vec{\mathbf{1}}_n\vec{\mathbf{1}}_n^T \\ \hline \left(s_{0,1}^{<r>} - c_{0,1}^{<r>}\right) I_n + c_{0,1}^{<r>}\vec{\mathbf{1}}_n\vec{\mathbf{1}}_n^T & \left(s_1^{2<r>} - c_1^{<r>}\right) I_n + c_1^{<r>}\vec{\mathbf{1}}_n\vec{\mathbf{1}}_n^T\end{array}\right],$$

We observe that for $\operatorname{Cov}\left(\left(\vec{Y}_t\right)_j, \left(\vec{Y}_2\right)_k \mid X, R=r\right)$ $(j \neq k)$ in the data-generative model to match the user specified value $\rho_{t,2}^{<r>}$, we must have

$$\begin{aligned}\rho_{0,2}^{<r>} &= \operatorname{Cov}\left(\left(\vec{\varepsilon}_0\right)_j, P_2^{[r]}\left(\vec{\varepsilon}_0\right)_k + P_3^{[r]}\frac{1}{n}\sum_{m=1}^{n}\left(\vec{\varepsilon}_0\right)_m + P_4^{[r]}\left(P_1\left(\vec{\varepsilon}_0\right)_k + \left(\vec{\varepsilon}_1\right)_k\right) + P_5^{[r]}\frac{1}{n}\sum_{m'=1}^{n}\left(\vec{\varepsilon}_1\right)_{m'} \middle| R=r\right)\\ &= P_2^{[r]}\left(c_0^{<r>}\right) + P_3^{[r]}\left(\frac{s_0^{2<r>}}{n} + \frac{n-1}{n}c_0^{<r>}\right) + P_4^{[r]}\left(P_1 c_0^{<r>} + c_{0,1}^{<r>}\right) + P_5^{[r]}\left(\frac{s_{0,1}^{<r>}}{n} + \frac{n-1}{n}c_{0,1}^{<r>}\right),\end{aligned}$$

and

$$\begin{aligned}\rho_{1,2}^{<r>} &= \operatorname{Cov}\left(P_1\left(\vec{\varepsilon}_0\right)_j + \left(\vec{\varepsilon}_1\right)_j, P_2^{[r]}\left(\vec{\varepsilon}_0\right)_k + P_3^{[r]}\frac{1}{n}\sum_{m=1}^{n}\left(\vec{\varepsilon}_0\right)_m + P_4^{[r]}\left(P_1\left(\vec{\varepsilon}_0\right)_k + \left(\vec{\varepsilon}_1\right)_k\right) + P_5^{[r]}\frac{1}{n}\sum_{m'=1}^{n}\left(\vec{\varepsilon}_1\right)_{m'} \middle| R=r\right)\\ &= P_1\left(P_2^{[r]}\left(c_0^{<r>}\right) + P_3^{[r]}\left(\left(\frac{s_0^{2<r>}}{n} + \frac{n-1}{n}c_0^{<r>}\right)\right) + P_4^{[r]}\left(P_1 c_0^{<r>} + c_{0,1}^{<r>}\right) + P_5^{[r]}\left(\frac{s_{0,1}^{<r>}}{n} + \frac{n-1}{n}c_{0,1}^{<r>}\right)\right)\\ &\quad + P_2^{[r]}\left(c_{0,1}^{<r>}\right) + P_3^{[r]}\left(\frac{s_{0,1}^{<r>}}{n} + \frac{n-1}{n}c_{0,1}^{<r>}\right) + P_4^{[r]}\left(P_1 c_{0,1}^{<r>} + c_1^{<r>}\right) + P_5^{[r]}\left(\left(\frac{s_1^{2<r>}}{n} + \frac{n-1}{n}c_1^{<r>}\right)\right)\\ &= P_2^{[r]}\left(P_1 c_0^{<r>} + c_{0,1}^{<r>}\right) + P_3^{[r]}\left(\frac{P_1 s_0^{2<r>} + s_{0,1}^{<r>}}{n} + \frac{n-1}{n}\left(P_1 c_0^{<r>} + c_{0,1}^{<r>}\right)\right)\\ &\quad + P_4^{[r]}\left(P_1^2 c_0^{<r>} + 2P_1 c_{0,1}^{<r>} + c_1^{<r>}\right) + P_5^{[r]}\left(\frac{P_1 s_{0,1}^{<r>} + s_1^{2<r>}}{n} + \frac{n-1}{n}\left(P_1 c_{0,1}^{<r>} + c_1^{<r>}\right)\right).\end{aligned}$$

Similarly, for $\operatorname{Cov}\left(\left(\vec{Y}_t\right)_j, \left(\vec{Y}_2\right)_j \mid X, R=r\right)$ in the data-generative model to match the target value $\phi_{t,2}^{<r>}$, we must have

$$\phi_{0,2}^{<r>} = P_2^{[r]}\left(s_0^{2<r>}\right) + P_3^{[r]}\left(\frac{s_0^{2<r>}}{n} + \frac{n-1}{n}c_0^{<r>}\right) + P_4^{[r]}\left(P_1 s_0^{2<r>} + s_{0,1}^{<r>}\right) + P_5^{[r]}\left(\frac{s_{0,1}^{<r>}}{n} + \frac{n-1}{n}c_{0,1}^{<r>}\right)$$

and

$$\begin{aligned}\phi_{1,2}^{<r>} &= P_2^{[r]}\left(P_1 s_0^{2<r>} + s_{0,1}^{<r>}\right) + P_3^{[r]}\left(\frac{P_1 s_0^{2<r>} + s_{0,1}^{<r>}}{n} + \frac{n-1}{n}\left(P_1 c_0^{<r>} + c_{0,1}^{<r>}\right)\right)\\ &\quad + P_4^{[r]}\left(P_1^2 s_0^{2<r>} + 2P_1 s_{0,1}^{<r>} + s_1^{2<r>}\right) + P_5^{[r]}\left(\frac{P_1 s_{0,1}^{<r>} + s_1^{2<r>}}{n} + \frac{n-1}{n}\left(P_1 c_{0,1}^{<r>} + c_1^{<r>}\right)\right).\end{aligned}$$

Therefore, letting

$$\Upsilon := \begin{bmatrix} s_0^{2<r>} & \frac{s_0^{2<r>}}{n} + \frac{n-1}{n} c_0^{<r>} & P_1 s_0^{2<r>} + s_{0,1}^{<r>} & \frac{s_{0,1}^{<r>}}{n} + \frac{n-1}{n} c_{0,1}^{<r>} \\ P_1 s_0^{2<r>} + s_{0,1}^{<r>} & \frac{P_1 s_0^{2<r>} + s_{0,1}^{<r>}}{n} + \frac{n-1}{n}\left(P_1 c_0^{<r>} + c_{0,1}^{<r>}\right) & P_1^2 s_0^{2<r>} + 2P_1 s_{0,1}^{<r>} + s_1^{2<r>} & \frac{P_1 s_{0,1}^{<r>} + s_1^{2<r>}}{n} + \frac{n-1}{n}\left(P_1 c_{0,1}^{<r>} + c_1^{<r>}\right) \\ c_0^{<r>} & \frac{s_0^{2<r>}}{n} + \frac{n-1}{n} c_0^{<r>} & P_1 c_0^{<r>} + c_{0,1}^{<r>} & \frac{s_{0,1}^{<r>}}{n} + \frac{n-1}{n} c_{0,1}^{<r>} \\ P_1 c_0^{<r>} + c_{0,1}^{<r>} & \frac{P_1 s_0^{2<r>} + s_{0,1}^{<r>}}{n} + \frac{n-1}{n}\left(P_1 c_0^{<r>} + c_{0,1}^{<r>}\right) & P_1^2 c_0^{<r>} + 2P_1 c_{0,1}^{<r>} + c_1^{<r>} & \frac{P_1 s_{0,1}^{<r>} + s_1^{2<r>}}{n} + \frac{n-1}{n}\left(P_1 c_{0,1}^{<r>} + c_1^{<r>}\right) \end{bmatrix},$$

we see the following is a necessary and sufficient condition for our pre-response/post-response conditional variance requirements to hold:

$$\begin{bmatrix} \phi_{0,2}^{<r>} \\ \phi_{1,2}^{<r>} \\ \rho_{0,2}^{<r>} \\ \rho_{1,2}^{<r>} \end{bmatrix} = \Upsilon \begin{bmatrix} P_2^{[r]} \\ P_3^{[r]} \\ P_4^{[r]} \\ P_5^{[r]} \end{bmatrix}.$$

Recalling that $\begin{bmatrix} P_2^{[r]} \\ P_3^{[r]} \\ P_4^{[r]} \\ P_5^{[r]} \end{bmatrix} = \Upsilon^{-1} \begin{bmatrix} \phi_{0,2}^{<r>} \\ \phi_{1,2}^{<r>} \\ \rho_{0,2}^{<r>} \\ \rho_{1,2}^{<r>} \end{bmatrix}$, we observe that our data-generative model satisfies our cross-response conditional covariance requirements.

Finally, we check our requirements on $\mathbb{V}\left[\vec{Y}_2 \mid X, R = r\right]$. Recalling that we constructed $\vec{\varepsilon}_2^{(r)} \perp\!\!\!\perp \begin{bmatrix} \vec{\varepsilon}_0 \\ \vec{\varepsilon}_1 \end{bmatrix} \big|_{R=r}$, we see that the following are necessary and sufficient conditions for $\mathrm{Cov}\left(\left(\vec{Y}_2\right)_j, \left(\vec{Y}_2\right)_k \mid X, R = r\right)$ and $\mathbb{V}\left[\left(\vec{Y}_2\right)_j \mid X, R = r\right]$ to match their respective target values, $\rho_2^{<r>}$ and $\sigma_2^{2<r>}$ (for any $j, k = 1, \ldots, n,\ j \neq k$):

$$\begin{aligned} \rho_2^{<r>} = \mathrm{Cov}\Bigg(& P_2^{[r]} \left(\vec{\varepsilon}_0\right)_j + P_3^{[r]} \frac{1}{n} \sum_{m=1}^{n} \left(\vec{\varepsilon}_0\right)_m + P_4^{[r]} \left(P_1 \left(\vec{\varepsilon}_0\right)_j + \left(\vec{\varepsilon}_1\right)_j\right) + P_5^{[r]} \frac{1}{n} \sum_{m'=1}^{n} \left(\vec{\varepsilon}_1\right)_{m'} + \left(\vec{\varepsilon}_2^{(r)}\right)_j, \\ & P_2^{[r]} \left(\vec{\varepsilon}_0\right)_k + P_3^{[r]} \frac{1}{n} \sum_{m=1}^{n} \left(\vec{\varepsilon}_0\right)_m + P_4^{[r]} \left(P_1 \left(\vec{\varepsilon}_0\right)_k + \left(\vec{\varepsilon}_1\right)_k\right) + P_5^{[r]} \frac{1}{n} \sum_{m'=1}^{n} \left(\vec{\varepsilon}_1\right)_{m'} + \left(\vec{\varepsilon}_2^{(r)}\right)_k \big| R = r \Bigg) \\ = \mathrm{Cov}\Bigg(& \left(P_2^{[r]} + \frac{P_3^{[r]}}{n} + P_4^{[r]} P_1\right) \left(\vec{\varepsilon}_{0,i}\right)_j + \frac{P_3^{[r]}}{n} \left(\vec{\varepsilon}_{0,i}\right)_k + \frac{P_3^{[r]}}{n} \sum_{m \neq j,k} \left(\vec{\varepsilon}_{0,i}\right)_m + \\ & \left(P_4^{[r]} + \frac{P_5^{[r]}}{n}\right) \left(\vec{\varepsilon}_{1,i}\right)_j + \frac{P_5^{[r]}}{n} \left(\vec{\varepsilon}_{1,i}\right)_k + \frac{P_5^{[r]}}{n} \sum_{m \neq j,k} \left(\vec{\varepsilon}_{1,i}\right)_m + \left(\vec{\varepsilon}_2^{(r)}\right)_j, \\ & \left(P_2^{[r]} + \frac{P_3^{[r]}}{n} + P_4^{[r]} P_1\right) \left(\vec{\varepsilon}_{0,i}\right)_k + \frac{P_3^{[r]}}{n} \left(\vec{\varepsilon}_{0,i}\right)_j + \frac{P_3^{[r]}}{n} \sum_{m \neq j,k} \left(\vec{\varepsilon}_{0,i}\right)_m + \\ & \left(P_4^{[r]} + \frac{P_5^{[r]}}{n}\right) \left(\vec{\varepsilon}_{1,i}\right)_k + \frac{P_5^{[r]}}{n} \left(\vec{\varepsilon}_{1,i}\right)_j + \frac{P_5^{[r]}}{n} \sum_{m \neq j,k} \left(\vec{\varepsilon}_{1,i}\right)_m + \left(\vec{\varepsilon}_2^{(r)}\right)_k \big| R = r \Bigg) \end{aligned}$$

$$
\begin{aligned}
&= \left(P_2^{[r]} + \frac{P_3^{[r]}}{n} + P_4^{[r]}P_1\right)\left(\left(P_2^{[r]} + \frac{P_3^{[r]}}{n}(n-1) + P_4^{[r]}P_1\right)c_0^{<r>} + \frac{P_3^{[r]}}{n}s_0^{2<r>}\right)\\
&\qquad + \left(P_2^{[r]} + \frac{P_3^{[r]}}{n} + P_4^{[r]}P_1\right)\left(\left(P_4^{[r]} + \frac{P_5^{[r]}}{n}(n-1)\right)c_{0,1}^{<r>} + \frac{P_5^{[r]}}{n}s_{0,1}^{<r>}\right)\\
&+ \frac{P_3^{[r]}}{n}\left(P_2^{[r]} + \frac{P_3^{[r]}}{n} + P_4^{[r]}P_1\right)s_0^{2<r>} + \left(\frac{P_3^{[r]}}{n}\right)^2(n-1)\,c_0^{<r>}\\
&\qquad + \frac{P_3^{[r]}}{n}\left(P_4^{[r]} + \frac{P_5^{[r]}}{n}\right)s_{0,1}^{<r>} + \frac{P_3^{[r]}P_5^{[r]}}{n^2}(n-1)\,c_{0,1}^{<r>}\\
&+ \frac{P_3^{[r]}}{n}(n-2)\left(P_2^{[r]} + 2\frac{P_3^{[r]}}{n} + P_4^{[r]}P_1\right)c_0^{<r>} + \left(\frac{P_3^{[r]}}{n}\right)^2(n-2)\left[s_0^{2<r>} + ((n-2)-1)\,c_0^{<r>}\right]\\
&\qquad + \frac{P_3^{[r]}}{n}(n-2)\left(P_4^{[r]} + 2\frac{P_5^{[r]}}{n}\right)c_{0,1}^{<r>} + \frac{P_3^{[r]}P_5^{[r]}}{n^2}(n-2)\left[s_{0,1}^{<r>} + ((n-2)-1)\,c_{0,1}^{<r>}\right]\\
&+ \left(P_4^{[r]} + \frac{P_5^{[r]}}{n}\right)\left(\left(P_2^{[r]} + \frac{P_3^{[r]}}{n}(n-1) + P_4^{[r]}P_1\right)c_{0,1}^{<r>} + \frac{P_3^{[r]}}{n}s_{0,1}^{<r>}\right)\\
&\qquad + \left(P_4^{[r]} + \frac{P_5^{[r]}}{n}\right)\left(\left(P_4^{[r]} + \frac{P_5^{[r]}}{n}(n-1)\right)c_1^{<r>} + \frac{P_5^{[r]}}{n}s_1^{2<r>}\right)\\
&+ \frac{P_5^{[r]}}{n}\left(P_2^{[r]} + \frac{P_3^{[r]}}{n} + P_4^{[r]}P_1\right)s_{0,1}^{<r>} + \frac{P_3^{[r]}P_5^{[r]}}{n^2}(n-1)\,c_{0,1}^{<r>}\\
&\qquad + \frac{P_5^{[r]}}{n}\left(P_4^{[r]} + \frac{P_5^{[r]}}{n}\right)s_1^{2<r>} + \left(\frac{P_5^{[r]}}{n}\right)^2(n-1)\,c_1^{<r>}\\
&+ \frac{P_5^{[r]}}{n}(n-2)\left(P_2^{[r]} + 2\frac{P_3^{[r]}}{n} + P_4^{[r]}P_1\right)c_{0,1}^{<r>} + \left(\frac{P_3^{[r]}P_5^{[r]}}{n^2}\right)(n-2)\left[s_{0,1}^{<r>} + ((n-2)-1)\,c_{0,1}^{<r>}\right]\\
&\qquad + \frac{P_5^{[r]}}{n}(n-2)\left(P_4^{[r]} + 2\frac{P_5^{[r]}}{n}\right)c_1^{<r>} + \left(\frac{P_5^{[r]}}{n}\right)^2(n-2)\left[s_1^{2<r>} + ((n-2)-1)\,c_1^{<r>}\right]\\
&+ \left(\Sigma_2^{[r]}\right)_{j,k}\\
&=: \zeta' + \left(\Sigma_2^{[r]}\right)_{j,k},
\end{aligned}
$$

and

$$\sigma_2^{2<r>} = \mathbb{V}\left[\left(P_2^{[r]} + \frac{P_3^{[r]}}{n} + P_4^{[r]}P_1\right)(\vec{\varepsilon}_0)_j + \frac{P_3^{[r]}}{n}\sum_{k\neq j}(\vec{\varepsilon}_{0,i})_k + \left(P_4^{[r]} + \frac{P_5^{[r]}}{n}\right)(\vec{\varepsilon}_{1,i})_j + \frac{P_5^{[r]}}{n}\sum_{k\neq j}(\vec{\varepsilon}_{1,i})_k + \left(\vec{\varepsilon}_2^{(r)}\right)_j \middle| R = r\right]$$

$$\begin{aligned}
&= \left(P_2^{[r]} + \frac{P_3^{[r]}}{n} + P_4^{[r]}P_1\right)^2 s_0^{2<r>} + \left(\frac{P_3^{[r]}}{n}\right)^2 (n-1)\left[s_0^{2<r>} + ((n-1)-1)\, c_0^{<r>}\right] \\
&\qquad + \left(P_4^{[r]} + \frac{P_5^{[r]}}{n}\right)^2 s_1^{2<r>} + \left(\frac{P_5^{[r]}}{n}\right)^2 (n-1)\left[s_1^{2<r>} + ((n-1)-1)\, c_1^{<r>}\right] + \left(\Sigma_2^{[r]}\right)_{j,j} \\
&+ 2\left(P_2^{[r]} + \frac{P_3^{[r]}}{n} + P_4^{[r]}P_1\right)\left(\frac{P_3^{[r]}}{n}(n-1)\, c_0^{<r>} + \left(P_4^{[r]} + \frac{P_5^{[r]}}{n}\right) s_{0,1}^{<r>} + \frac{P_5^{[r]}}{n}(n-1)\, c_{0,1}^{<r>}\right) \\
&\qquad + 2\frac{P_3^{[r]}}{n}(n-1)\left(\left(P_4^{[r]} + \frac{P_5^{[r]}}{n}\right) c_{0,1}^{<r>} + \frac{P_5^{[r]}}{n}\left[s_{0,1}^{<r>} + ((n-1)-1)\, c_{0,1}^{<r>}\right]\right) \\
&+ 2\left(P_4^{[r]} + \frac{P_5^{[r]}}{n}\right)\left(\frac{P_5^{[r]}}{n}\right)(n-1)\, c_1^{<r>} \\
&=: \zeta + \left(\Sigma_2^{[r]}\right)_{j,j}.
\end{aligned}$$

Given that we generated $\vec{\varepsilon}_2 \sim \mathcal{N}\left(\vec{0}_n,\ \Sigma_2^{[r]}\right)$, where $\Sigma_2^{[r]} = \left(\left(\sigma_2^{2<r>} - \zeta\right) - \left(\rho_2^{<r>} - \zeta'\right)\right) I_n + \left(\rho_2^{<r>} - \zeta'\right)\vec{\mathbf{1}}_n\vec{\mathbf{1}}_n^T$, our construction of $\Sigma_2^{[r]}$ guarantees that the overall data-generating model satisfies the requirements on $\mathbb{V}\left[\vec{Y}_2 \mid X, R = r\right]$.

3 Parameter Selection

Section 2 discussed the arguments to select in order to fully specify the data generative model. The parameters defining the simulation settings in the main body of the report were derived via analysis of the ASIC data. For example, in the report we considered different working variance model-fitting assumptions when the data had been generated using various variance settings. To obtain marginal mean/variance data-generative parameters under a given variance setting, we used the method described in the main body (along with the example marginal mean model), assuming the working variance structure in question, to model weekly delivery of CBT in the ASIC data.[5]

As noted above, we also needed to specify the conditional mean/variance parameters for responding clusters - we did so by analyzing the non-responding clusters in ASIC (separately for each embedded DTR). We then averaged non-responding parameters across first-stage treatment arms and used these, along with the marginal distribution parameters, to derive responder parameters.[6][7]

As recommended in Pan et al. [2024+], we assumed all covariance terms were non-negative when fitting the marginal model. While this assumption is reasonable in ASIC-type scenarios and facilitates estimator stability, it is not necessarily justifiable when considering conditional covariance terms.[8] Therefore, we fit the conditional models using unrestricted correlation estimation.

Lastly, as discussed in Section 2.3, we simulated DTR assignment prior to any outcome/covariates.[9]

[5] Note that, for this purpose, we restricted the ASIC data to times $t_0, t^*, and t_T$ in order to provide parameters to specify a data-generative model with three time points.

[6] We did so because ASIC had a low response rate, which may lead to unstable conditional distribution parameters for responding clusters.

[7] Occasionally, this approach led to distributions which were incompatible with the data generative model. In these cases, we made slight adjustments to the parameter inputs.

[8] For example, if a responding cluster has two individuals and one of them was well-off at time t^*, then it was likely the case that the other individual sufficiently not well-off as to offset their counterpart and trigger cluster-wide non-response.

[9] Note that this simulation was independent of all other data, and therefore is equivalent to simulating A_1, R, A_2 sequentially.

4 Supplemental Definitions

After using Monte-Carlo simulation to obtain the following s and c parameters:

$$\mathbb{V}\left[\begin{bmatrix}\vec{\varepsilon}_0\\ \vec{\varepsilon}_1\end{bmatrix} \mid X, R=r\right] = \left[\begin{array}{c|c}\left(s_0^{2<r>} - c_0^{<r>}\right) I_n + c_0^{<r>} \vec{\mathbf{1}}_n \vec{\mathbf{1}}_n^T & \left(s_{0,1}^{<r>} - c_{0,1}^{<r>}\right) I_n + c_{0,1}^{<r>} \vec{\mathbf{1}}_n \vec{\mathbf{1}}_n^T \\ \hline \left(s_{0,1}^{<r>} - c_{0,1}^{<r>}\right) I_n + c_{0,1}^{<r>} \vec{\mathbf{1}}_n \vec{\mathbf{1}}_n^T & \left(s_1^{2<r>} - c_1^{<r>}\right) I_n + c_1^{<r>} \vec{\mathbf{1}}_n \vec{\mathbf{1}}_n^T\end{array}\right],$$

we employ the following notation:

$$\Upsilon := \begin{bmatrix} s_0^{2<r>} & \frac{s_0^{2<r>}}{n} + \frac{n-1}{n} c_0^{<r>} & P_1 s_0^{2<r>} + s_{0,1}^{<r>} & \frac{s_{0,1}^{<r>}}{n} + \frac{n-1}{n} c_{0,1}^{<r>} \\ P_1 s_0^{2<r>} + s_{0,1}^{<r>} & \frac{P_1 s_0^{2<r>} + s_{0,1}^{<r>}}{n} + \frac{n-1}{n}\left(P_1 c_0^{<r>} + c_{0,1}^{<r>}\right) & P_1^2 s_0^{2<r>} + 2P_1 s_{0,1}^{<r>} + s_1^{2<r>} & \frac{P_1 s_{0,1}^{<r>} + s_1^{2<r>}}{n} + \frac{n-1}{n}\left(P_1 c_{0,1}^{<r>} + c_1^{<r>}\right) \\ c_0^{<r>} & \frac{s_0^{2<r>}}{n} + \frac{n-1}{n} c_0^{<r>} & P_1 c_0^{<r>} + c_{0,1}^{<r>} & \frac{s_{0,1}^{<r>}}{n} + \frac{n-1}{n} c_{0,1}^{<r>} \\ P_1 c_0^{<r>} + c_{0,1}^{<r>} & \frac{P_1 s_0^{2<r>} + s_{0,1}^{<r>}}{n} + \frac{n-1}{n}\left(P_1 c_0^{<r>} + c_{0,1}^{<r>}\right) & P_1^2 c_0^{<r>} + 2P_1 c_{0,1}^{<r>} + c_1^{<r>} & \frac{P_1 s_{0,1}^{<r>} + s_1^{2<r>}}{n} + \frac{n-1}{n}\left(P_1 c_{0,1}^{<r>} + c_1^{<r>}\right) \end{bmatrix},$$

$$\begin{aligned}\zeta := &\left(P_2^{[r]} + \frac{P_3^{[r]}}{n} + P_4^{[r]} P_1\right)^2 s_0^{2<r>} + \left(\frac{P_3^{[r]}}{n}\right)^2 (n-1)\left[s_0^{2<r>} + ((n-1)-1)\, c_0^{<r>}\right] \\ &\quad + \left(P_4^{[r]} + \frac{P_5^{[r]}}{n}\right)^2 s_1^{2<r>} + \left(\frac{P_5^{[r]}}{n}\right)^2 (n-1)\left[s_1^{2<r>} + ((n-1)-1)\, c_1^{<r>}\right] \\ &+ 2\left(P_2^{[r]} + \frac{P_3^{[r]}}{n} + P_4^{[r]} P_1\right)\left(\frac{P_3^{[r]}}{n}(n-1)\, c_0^{<r>} + \left(P_4^{[r]} + \frac{P_5^{[r]}}{n}\right) s_{0,1}^{<r>} + \frac{P_5^{[r]}}{n}(n-1)\, c_{0,1}^{<r>}\right) \\ &\quad + 2\frac{P_3^{[r]}}{n}(n-1)\left(\left(P_4^{[r]} + \frac{P_5^{[r]}}{n}\right) c_{0,1}^{<r>} + \frac{P_5^{[r]}}{n}\left[s_{0,1}^{<r>} + ((n-1)-1)\, c_{0,1}^{<r>}\right]\right) \\ &+ 2\left(P_4^{[r]} + \frac{P_5^{[r]}}{n}\right)\left(\frac{P_5^{[r]}}{n}\right)(n-1)\, c_1^{<r>},\end{aligned}$$

and

$$
\begin{aligned}
\zeta' := &\left(P_2^{[r]} + \frac{P_3^{[r]}}{n} + P_4^{[r]} P_1\right) \left(\left(\left(P_2^{[r]} + \frac{P_3^{[r]}}{n}(n-1) + P_4^{[r]} P_1\right) c_0^{<r>} + \frac{P_3^{[r]}}{n} s_0^{2<r>}\right)\right. \\
&\qquad\qquad + \left(P_2^{[r]} + \frac{P_3^{[r]}}{n} + P_4^{[r]} P_1\right) \left(\left(\left(P_4^{[r]} + \frac{P_5^{[r]}}{n}(n-1)\right) c_{0,1}^{<r>} + \frac{P_5^{[r]}}{n} s_{0,1}^{<r>}\right)\right. \\
&+ \frac{P_3^{[r]}}{n}\left(P_2^{[r]} + \frac{P_3^{[r]}}{n} + P_4^{[r]} P_1\right) s_0^{2<r>} + \left(\frac{P_3^{[r]}}{n}\right)^2 (n-1)\, c_0^{<r>} \\
&\qquad\qquad + \frac{P_3^{[r]}}{n}\left(P_4^{[r]} + \frac{P_5^{[r]}}{n}\right) s_{0,1}^{<r>} + \frac{P_3^{[r]} P_5^{[r]}}{n^2}(n-1)\, c_{0,1}^{<r>} \\
&+ \frac{P_3^{[r]}}{n}(n-2)\left(P_2^{[r]} + 2\frac{P_3^{[r]}}{n} + P_4^{[r]} P_1\right) c_0^{<r>} + \left(\frac{P_3^{[r]}}{n}\right)^2 (n-2)\left[s_0^{2<r>} + ((n-2)-1)\, c_0^{<r>}\right] \\
&\qquad\qquad + \frac{P_3^{[r]}}{n}(n-2)\left(P_4^{[r]} + 2\frac{P_5^{[r]}}{n}\right) c_{0,1}^{<r>} + \frac{P_3^{[r]} P_5^{[r]}}{n^2}(n-2)\left[s_{0,1}^{<r>} + ((n-2)-1)\, c_{0,1}^{<r>}\right] \\
&+ \left(P_4^{[r]} + \frac{P_5^{[r]}}{n}\right) \left(\left(\left(P_2^{[r]} + \frac{P_3^{[r]}}{n}(n-1) + P_4^{[r]} P_1\right) c_{0,1}^{<r>} + \frac{P_3^{[r]}}{n} s_{0,1}^{<r>}\right)\right. \\
&\qquad\qquad + \left(P_4^{[r]} + \frac{P_5^{[r]}}{n}\right) \left(\left(\left(P_4^{[r]} + \frac{P_5^{[r]}}{n}(n-1)\right) c_1^{<r>} + \frac{P_5^{[r]}}{n} s_1^{2<r>}\right)\right. \\
&+ \frac{P_5^{[r]}}{n}\left(P_2^{[r]} + \frac{P_3^{[r]}}{n} + P_4^{[r]} P_1\right) s_{0,1}^{<r>} + \frac{P_3^{[r]} P_5^{[r]}}{n^2}(n-1)\, c_{0,1}^{<r>} \\
&\qquad\qquad + \frac{P_5^{[r]}}{n}\left(P_4^{[r]} + \frac{P_5^{[r]}}{n}\right) s_1^{2<r>} + \left(\frac{P_5^{[r]}}{n}\right)^2 (n-1)\, c_1^{<r>} \\
&+ \frac{P_5^{[r]}}{n}(n-2)\left(P_2^{[r]} + 2\frac{P_3^{[r]}}{n} + P_4^{[r]} P_1\right) c_{0,1}^{<r>} + \left(\frac{P_3^{[r]} P_5^{[r]}}{n^2}\right)(n-2)\left[s_{0,1}^{<r>} + ((n-2)-1)\, c_{0,1}^{<r>}\right] \\
&\qquad\qquad + \frac{P_5^{[r]}}{n}(n-2)\left(P_4^{[r]} + 2\frac{P_5^{[r]}}{n}\right) c_1^{<r>} + \left(\frac{P_5^{[r]}}{n}\right)^2 (n-2)\left[s_1^{2<r>} + ((n-2)-1)\, c_1^{<r>}\right].
\end{aligned}
$$

References

Wenchu Pan, Daniel Almirall, and Lu Wang. Finite-sample adjustments for comparing clustered adaptive interventions using data from a clustered SMART. *Upcoming*, 2024+.

Nicholas J Seewald, Kelley M Kidwell, Inbal Nahum-Shani, Tianshuang Wu, James R McKay, and Daniel Almirall. Sample size considerations for comparing dynamic treatment regimens in a sequential multiple-assignment randomized trial with a continuous longitudinal outcome. *Statistical Methods in Medical Research*, 29(7):1891–1912, 2020. ISSN 0962-2802. doi: 10/gf85ss.