SUPPLEMENTARY MATERIAL TO "NEW G-FORMULA FOR THE SEQUENTIAL CAUSAL EFFECT AND BLIP EFFECT OF TREATMENT IN SEQUENTIAL CAUSAL INFERENCE"

XIAOQIN WANG and LI YIN

SUPPLEMENT I: PROOFS FOR FORMULAS (10A), (10B), (10C), (21), (22)

Proof of (10a): For mixed subregime $\mathbf{G}_t^s = (\mathbf{D}_t^{s-1}, R_s)$ given the history $(\mathbf{z}_1^{t-1}, \mathbf{x}_1^{t-1})$, we have the following factorization

$$\mathrm{P}^{\mathbf{G}_t^s}(\mathbf{z}_t^{s-1}, \mathbf{x}_t^{s-1}, z_s \mid \mathbf{z}_1^{t-1}, \mathbf{x}_1^{t-1}) =$$

$$\mathrm{P}^{G_s}(z_s \mid \mathbf{z}_1^{s-1}, \mathbf{x}_1^{s-1}) \prod_{k=t}^{s-1} \mathrm{P}^{G_k}(\mathbf{x}_k \mid \mathbf{z}_1^{k-1}, \mathbf{x}_1^{k-1}, z_k)\mathrm{P}^{G_k}(z_k \mid \mathbf{z}_1^{k-1}, \mathbf{x}_1^{k-1}).$$

Due to the identifying condition, the first factor $\mathrm{P}^{G_s}(z_s \mid \mathbf{z}_1^{s-1}, \mathbf{x}_1^{s-1}) = \mathrm{P}^{\mathbf{O}}(z_s \mid \mathbf{z}_1^{s-1}, \mathbf{x}_1^{s-1})$ and the second factor $\mathrm{P}^{G_k}(\mathbf{x}_k \mid \mathbf{z}_1^{k-1}, \mathbf{x}_1^{k-1}, z_k) = \mathrm{P}^{\mathbf{O}}(\mathbf{x}_k \mid \mathbf{z}_1^{k-1}, \mathbf{x}_1^{k-1}, z_k)$. Furthermore, the third factor $\mathrm{P}^{G_k}(z_k \mid \mathbf{z}_1^{k-1}, \mathbf{x}_1^{k-1}) = 1$ because z_k is deterministically assigned. Therefore, we obtain

$$\mathrm{P}^{\mathbf{G}_t^s}(\mathbf{z}_t^{s-1}, \mathbf{x}_t^{s-1}, z_s \mid \mathbf{z}_1^{t-1}, \mathbf{x}_1^{t-1}) = \mathrm{P}^{\mathbf{O}}(z_s \mid \mathbf{z}_1^{s-1}, \mathbf{x}_1^{s-1}) \prod_{k=t}^{s-1} \mathrm{P}^{\mathbf{O}}(\mathbf{x}_k \mid \mathbf{z}_1^{k}, \mathbf{x}_1^{k-1}),$$

which is (10a).

Proof of (10b): For mixed subregime $\mathbf{G}_t^s = (D_t, \mathbf{R}_{t+1}^s)$ given $(\mathbf{z}_1^{t-1}, \mathbf{x}_1^{t-1})$, we have the factorization

$$\mathrm{P}^{\mathbf{G}_t^s}(z_t, \mathbf{z}_{t+1}^{s-1}, \mathbf{x}_t^{s-1}, z_s \mid \mathbf{z}_1^{t-1}, \mathbf{x}_1^{t-1}) = \mathrm{P}^{\mathbf{G}_t^s}(\mathbf{z}_{t+1}^{s-1}, \mathbf{x}_t^{s-1}, z_s \mid \mathbf{z}_1^{t-1}, \mathbf{x}_1^{t-1}, z_t)$$

$$\mathrm{P}^{G_t}(z_t \mid \mathbf{z}_1^{t-1}, \mathbf{x}_1^{t-1}).$$

The first factor $\mathrm{P}^{\mathbf{G}_t^s}(\mathbf{z}_{t+1}^{s-1}, \mathbf{x}_t^{s-1}, z_s \mid \mathbf{z}_1^{t-1}, \mathbf{x}_1^{t-1}, z_t) = \mathrm{P}^{\mathbf{O}}(\mathbf{z}_{t+1}^{s-1}, \mathbf{x}_t^{s-1}, z_s \mid \mathbf{z}_1^{t-1}, \mathbf{x}_1^{t-1}, z_t)$ under the identifying condition. The second factor $\mathrm{P}^{G_t}(z_t \mid \mathbf{z}_1^{t-1}, \mathbf{x}_1^{t-1}) = 1$ because z_t is deterministically assigned. Therefore, we obtain

$$\mathrm{P}^{\mathbf{G}_t^s}(z_t, \mathbf{z}_{t+1}^{s-1}, \mathbf{x}_t^{s-1}, z_s \mid \mathbf{z}_1^{t-1}, \mathbf{x}_1^{t-1}) = \mathrm{P}^{\mathbf{O}}(\mathbf{z}_{t+1}^{s-1}, \mathbf{x}_t^{s-1}, z_s \mid \mathbf{z}_1^{t-1}, \mathbf{x}_1^{t-1}, z_t),$$

which is (10b).

Proof of (10c): For deterministic subregime $\mathbf{D}_t^s$ given $(\mathbf{z}_1^{t-1}, \mathbf{x}_1^{t-1})$, we have the following factorization

$$\mathrm{P}^{\mathbf{D}_t^s}(\mathbf{z}_t^{s-1}, \mathbf{x}_t^{s-1}, z_s \mid \mathbf{z}_1^{t-1}, \mathbf{x}_1^{t-1})$$

$$= \mathrm{P}^{D_s}(z_s \mid \mathbf{z}_1^{s-1}, \mathbf{x}_1^{s-1}) \prod_{k=t}^{s-1} \mathrm{P}^{D_k}(\mathbf{x}_k \mid \mathbf{z}_1^{k-1}, \mathbf{x}_1^{k-1}, z_k) \mathrm{P}^{D_k}(z_k \mid \mathbf{z}_1^{k-1}, \mathbf{x}_1^{k-1}).$$

The first factor $\mathrm{P}^{D_s}(z_s \mid \mathbf{z}_1^{s-1}, \mathbf{x}_1^{s-1}) = 1$ and the third factor $\mathrm{P}^{D_k}(z_k \mid \mathbf{z}_1^{k-1}, \mathbf{x}_1^{k-1}) = 1$ because z_s and z_k are deterministically assigned. The second factor $\mathrm{P}^{D_k}(\mathbf{x}_k \mid \mathbf{z}_1^{k-1}, \mathbf{x}_1^{k-1}, z_k) = \mathrm{P}^{\mathbf{O}}(\mathbf{x}_k \mid \mathbf{z}_1^{k-1}, \mathbf{x}_1^{k-1}, z_k)$ under the identifying condition. Therefore, we obtain

$$\mathrm{P}^{\mathbf{D}_t^s}(\mathbf{z}_t^{s-1}, \mathbf{x}_t^{s-1}, z_s \mid \mathbf{z}_1^{t-1}, \mathbf{x}_1^{t-1}) = \prod_{k=t}^{s-1} \mathrm{P}^{\mathbf{O}}\{\mathbf{x}_k \mid \mathbf{z}_1^{k-1}, \mathbf{z}_1^{k-1}, z_k\},$$

which is (10c).

Proof of (21): Recalling from (13), the point observable effect of treatment $z_t = 1$ in stratum $(\mathbf{z}_1^{t-1}, \mathbf{x}_1^{t-1})$, denoted by $\vartheta(\mathbf{z}_1^{t-1}, \mathbf{x}_1^{t-1})$, is

$$\vartheta(\mathbf{z}_1^{t-1}, \mathbf{x}_1^{t-1}) = \mu(\mathbf{z}_1^{t-1}, \mathbf{x}_1^{t-1}, z_t = 1) - \mu(\mathbf{z}_1^{t-1}, \mathbf{x}_1^{t-1}, z_t = 0).$$

We need the following lemma to prove (21).

Lemma *Supposing the same variance σ^2 for Y given any $\{\mathbf{z}_1^T, \mathbf{x}_1^{T-1}\}$, then the covariance between the estimated point observable effects at different times is equal to zero, that is,*

$$\mathrm{cov}\{\hat{\vartheta}(\mathbf{z}_1^{t-1}, \mathbf{x}_1^{t-1}), \hat{\vartheta}(\mathbf{z}_1^{s-1}, \mathbf{x}_1^{s-1})\} = 0, \quad t \neq s.$$

Proof: Without a loss of generality, we assume $t < s$. Let S be the stratum of observations satisfying $(\mathbf{z}_{i1}^{t-1}, \mathbf{x}_{i1}^{t-1}, z_{it}) = (\mathbf{z}_1^{t-1}, \mathbf{x}_1^{t-1}, z_t)$; S_0 for $(\mathbf{z}_1^{s-1}, \mathbf{x}_1^{s-1}, z_s = 0)$; S_1 for $(\mathbf{z}_1^{s-1}, \mathbf{x}_1^{s-1}, z_s = 1)$; and $S_2 = S \setminus (S_1 \cup S_2)$. Noticeably, S_0 and S_1 are disjoint, and S either contains both S_0 and S_1 or contains neither. If S contains neither S_0 nor S_1, then lemma is true. Therefore we only prove the lemma when S contains both S_0 and S_1. Let $n(.)$ be the number of observations in a stratum. Hence, $n(S) = n(S_0) + n(S_1) + n(S_2)$. Rewrite $\mu(\mathbf{z}_1^{t-1}, \mathbf{x}_1^{t-1}, z_t) = \mu(S)$, $\mu(\mathbf{z}_1^{s-1}, \mathbf{x}_1^{s-1}, z_s = 0) = \mu(S_0)$, $\mu(\mathbf{z}_1^{s-1}, \mathbf{x}_1^{s-1}, z_s = 1) = \mu(S_1)$. Additionally, denote the mean of Y in S_2 by $\mu(S_2)$. Then, the mean $\mu(S)$ is equal to

$$\mu(S) = \frac{n(S_0)}{n(S)}\mu(S_0) + \frac{n(S_1)}{n(S)}\mu(S_1) + \frac{n(S_2)}{n(S)}\mu(S_2)$$

and the ML estimate of $\mu(S)$ is

$$\hat{\mu}(S) = \frac{n(S_0)}{n(S)}\hat{\mu}(S_0) + \frac{n(S_1)}{n(S)}\hat{\mu}(S_1) + \frac{n(S_2)}{n(S)}\hat{\mu}(S_2).$$

Thus,

$$\hat{\mu}(S) - \mu(S) = \frac{n(S_0)}{n(S)}\{\hat{\mu}(S_0) - \mu(S_0)\} + \frac{n(S_1)}{n(S)}\{\hat{\mu}(S_1) - \mu(S_1)\}$$
$$+\frac{n(S_2)}{n(S)}\{\hat{\mu}(S_2) - \mu(S_2)\}.$$

On the other hand, we have $\vartheta(\mathbf{z}_1^{s-1}, \mathbf{x}_1^{s-1}) = \mu(S_1) - \mu(S_0)$ and thus

$$\hat{\vartheta}(\mathbf{z}_1^{s-1}, \mathbf{x}_1^{s-1}) - \vartheta(\mathbf{z}_1^{s-1}, \mathbf{x}_1^{s-1}) = \{\hat{\mu}(S_1) - \mu(S_1)\} - \{\hat{\mu}(S_0) - \mu(S_0)\}.$$

Thus, we have

$$\begin{aligned} &\operatorname{cov}\{\hat{\mu}(S), \hat{\vartheta}(\mathbf{z}_1^{s-1}, \mathbf{x}_1^{s-1})\} \\ &= E\left[\{\hat{\mu}(S) - \mu(S)\}\{\hat{\vartheta}(\mathbf{z}_1^{s-1}, \mathbf{x}_1^{s-1}) - \vartheta(\mathbf{z}_1^{s-1}, \mathbf{x}_1^{s-1})\}\right] \\ &= \frac{n(S_1)}{n(S)}E\{\hat{\mu}(S_1) - \mu(S_1)\}^2 - \frac{n(S_0)}{n(S)}E\{\hat{\mu}(S_0) - \mu(S_0)\}^2, \end{aligned}$$

which is equal to

$$\frac{\sigma^2}{n(S)} - \frac{\sigma^2}{n(S)} = 0.$$

Therefore, we have

$$\operatorname{cov}\{\hat{\mu}(S), \hat{\vartheta}(\mathbf{z}_1^{s-1}, \mathbf{x}_1^{s-1})\} = 0,$$

which is true for $\hat{\mu}(S) = \hat{\mu}(\mathbf{z}_1^{t-1}, \mathbf{x}_1^{t-1}, z_t)$. Noticeably, $\hat{\vartheta}(\mathbf{z}_1^{t-1}, \mathbf{x}_1^{t-1}) = \hat{\mu}(\mathbf{z}_1^{t-1}, \mathbf{x}_1^{t-1}, z_t = 1) - \hat{\mu}(\mathbf{z}_1^{t-1}, \mathbf{x}_1^{t-1}, z_t = 0)$; therefore, we have

$$\operatorname{cov}\{\hat{\vartheta}(\mathbf{z}_1^{t-1}, \mathbf{x}_1^{t-1}), \hat{\vartheta}(\mathbf{z}_1^{s-1}, \mathbf{x}_1^{s-1})\} = 0, \quad t < s,$$

which proves the lemma.

As described in Section 5.2, the assignment of treatment z_t under the first-order Markov process is dependent only on $\mathbf{x}_{t-1}$, which implies

$$\mathrm{P}^{\mathbf{O}}(\mathbf{z}_1^{t-1}, \mathbf{x}_1^{t-2} \mid \mathbf{x}_{t-1}, z_t) = \mathrm{P}^{\mathbf{O}}(\mathbf{z}_1^{t-1}, \mathbf{x}_1^{t-2} \mid \mathbf{x}_{t-1}) \tag{25}$$

which in turn implies

$$\hat{\vartheta}(\mathbf{x}_{t-1}) = \sum_{\mathbf{z}_1^{t-1}, \mathbf{x}_1^{t-2}} \hat{\vartheta}(\mathbf{z}_1^{t-1}, \mathbf{x}_1^{t-1}) \mathrm{P}^{\mathbf{O}}(\mathbf{z}_1^{t-1}, \mathbf{x}_1^{t-2} \mid \mathbf{x}_{t-1}),$$

$$\hat{\vartheta}(\mathbf{x}_{s-1}) = \sum_{\mathbf{z}_1^{s-1}, \mathbf{x}_1^{s-2}} \hat{\vartheta}(\mathbf{z}_1^{s-1}, \mathbf{x}_1^{s-1}) \mathrm{P}^{\mathbf{O}}(\mathbf{z}_1^{s-1}, \mathbf{x}_1^{s-2} \mid \mathbf{x}_{s-1}).$$

These expressions together with the lemma above imply (21).

Proof of (22): Inserting the blip effects φ_1, φ_2 and φ_3 into the new G-formula (17) and noticing

$$\sum_{\mathbf{z}_{t+1}^{s-1},\mathbf{x}_t^{s-1}} \mathrm{P}^{\mathbf{O}}(\mathbf{z}_{t+1}^{s-1},\mathbf{x}_t^{s-1},z_s=1 \mid \mathbf{z}_1^{t-1},\mathbf{x}_1^{t-1},z_t) = \mathrm{P}^{\mathbf{O}}(z_s=1 \mid \mathbf{z}_1^{t-1},\mathbf{x}_1^{t-1},z_t),$$

we obtain

$$\vartheta(\mathbf{z}_1^{t-1},\mathbf{x}_1^{t-1}) = \varphi_1 c^{(1)}(\mathbf{z}_1^{t-1},\mathbf{x}_1^{t-1}) + \varphi_2 c^{(2)}(\mathbf{z}_1^{t-1},\mathbf{x}_1^{t-1}) + \varphi_3 c^{(3)}(\mathbf{z}_1^{t-1},\mathbf{x}_1^{t-1}),$$

where

$$\begin{aligned} c^{(1)}(\mathbf{z}_1^{t-1},\mathbf{x}_1^{t-1}) = \Bigg\{ & \sum_{s=t}^{T-2} \mathrm{P}^{\mathbf{O}}(z_s=1 \mid z_1^{t-1},\mathbf{x}_1^{t-1},z_t=1) \\ & - \sum_{s=t}^{T-2} \mathrm{P}^{\mathbf{O}}(z_s=1 \mid z_1^{t-1},\mathbf{x}_1^{t-1},z_t=0) \Bigg\}, \\ c^{(2)}(\mathbf{z}_1^{t-1},\mathbf{x}_1^{t-1}) = \Big\{ & \mathrm{P}^{\mathbf{O}}(z_{T-1}=1 \mid z_1^{t-1},\mathbf{x}_1^{t-1},z_t=1) \\ & - \mathrm{P}^{\mathbf{O}}(z_{T-1}=1 \mid z_1^{t-1},\mathbf{x}_1^{t-1},z_t=0) \Big\}, \\ c^{(3)}(\mathbf{z}_1^{t-1},\mathbf{x}_1^{t-1}) = \Big\{ & \mathrm{P}^{\mathbf{O}}(z_T=1 \mid z_1^{t-1},\mathbf{x}_1^{t-1},z_t=1) \\ & - \mathrm{P}^{\mathbf{O}}(z_T=1 \mid z_1^{t-1},\mathbf{x}_1^{t-1},z_t=0) \Big\}. \end{aligned}$$

Noticeably, if $t=T-1$, then $c^{(1)}(\mathbf{z}_1^{T-2},\mathbf{x}_1^{T-2})=0$; if $t=T$, then $c^{(1)}(\mathbf{z}_1^{T-1},\mathbf{x}_1^{T-1}) = 0$ and $c^{(2)}(\mathbf{z}_1^{T-1},\mathbf{x}_1^{T-1})=0$.

On the other hand, we have according to (25)

$$\begin{aligned} & \sum_{\mathbf{z}_1^{t-1},\mathbf{x}_1^{t-2}} \mathrm{P}^{\mathbf{O}}(z_s=1 \mid z_1^{t-1},\mathbf{x}_1^{t-1},z_t)\mathrm{P}^{\mathbf{O}}(\mathbf{z}_1^{t-1},\mathbf{x}_1^{t-2} \mid \mathbf{x}_{t-1}) \\ = & \sum_{\mathbf{z}_1^{t-1},\mathbf{x}_1^{t-2}} \mathrm{P}^{\mathbf{O}}(z_s=1 \mid z_1^{t-1},\mathbf{x}_1^{t-1},z_t)\mathrm{P}^{\mathbf{O}}(\mathbf{z}_1^{t-1},\mathbf{x}_1^{t-2} \mid \mathbf{x}_{t-1},z_t) \\ & = \mathrm{P}^{\mathbf{O}}(z_s=1 \mid \mathbf{x}_{t-1},z_t) \end{aligned}$$

and

$$\vartheta(\mathbf{x}_{t-1}) = \sum_{\mathbf{z}_1^{t-1},\mathbf{x}_1^{t-2}} \vartheta(\mathbf{z}_1^{t-1},\mathbf{x}_1^{t-1})\mathrm{P}^{\mathbf{O}}(\mathbf{z}_1^{t-1},\mathbf{x}_1^{t-2} \mid \mathbf{x}_{t-1}).$$

Therefore, averaging $\vartheta(\mathbf{z}_1^{t-1},\mathbf{x}_1^{t-1})$ with respect to $\mathrm{P}^{\mathbf{O}}(\mathbf{z}_1^{t-1},\mathbf{x}_1^{t-2} \mid \mathbf{x}_{t-1})$, we obtain

$$\vartheta(\mathbf{x}_{t-1}) = \varphi_1 c^{(1)}(\mathbf{x}_{t-1}) + \varphi_2 c^{(2)}(\mathbf{x}_{t-1}) + \varphi_3 c^{(3)}(\mathbf{x}_{t-1}),$$

where

$$c^{(1)}(\mathbf{x}_{t-1}) = \sum_{s=t}^{T-2}\{\mathrm{pr}(z_s = 1 \mid \mathbf{x}_{t-1}, z_t = 1) - \mathrm{pr}(z_s = 1 \mid \mathbf{x}_{t-1}, z_t = 0)\},$$

$$c^{(2)}(\mathbf{x}_{t-1}) = \mathrm{pr}(z_{T-1} = 1 \mid \mathbf{x}_{t-1}, z_t = 1) - \mathrm{pr}(z_{T-1} = 1 \mid \mathbf{x}_{t-1}, z_t = 0),$$

$$c^{(3)}(\mathbf{x}_{t-1}) = \mathrm{pr}(z_T = 1 \mid \mathbf{x}_{t-1}, z_t = 1) - \mathrm{pr}(z_T = 1 \mid \mathbf{x}_{t-1}, z_t = 0),$$

which is (22).

SUPPLEMENT II: SIMULATION STUDY IN SECTION 5.3

Here, we provide details about the simulation study in Section 5.3. In Section $II.1$, we describe methods (i) and (ii), which are developed in this article. In Section $II.2$, we describe methods (iii), (iv) and (v), which are available in the literature. In Section $II.3$, we describe how to apply methods (i)–(v) to the simulation. In Section $II.4$, we describe how the standard parameter $\mu(\mathbf{z}_1^3, \mathbf{x}_1^2)$ is constructed, which generates the data. The relevant SAS codes used for the simulation are also included in Supplementary Material.

II.1 Methods (i) and (ii). In the first stage of the procedure, we estimate the point observable effects by ML as follows. Due to the first-order Markov process, we have a total of nine point observable effects $\vartheta(\mathbf{x}_{t-1})$ of $z_t = 1$ given by (20) in Section 5.2: one ϑ of $z_1 = 1$, four $\vartheta(\mathbf{x}_1)$ of $z_2 = 1$ with $\mathbf{x}_1 = (0,0), (0,1), (1,0), (1,1)$, and four $\vartheta(\mathbf{x}_2)$ of $z_3 = 1$ with $\mathbf{x}_2 = (0,0), (0,1), (1,0), (1,1)$. We estimate these point observable effects by ML according to (20), that is,

$$\vartheta(\mathbf{x}_{t-1}) = \mu(\mathbf{x}_{t-1}, z_t = 1) - \mu(\mathbf{x}_{t-1}, z_t = 0).$$

Here, the mean $\mu(\mathbf{x}_{t-1}, z_t)$ is estimated by averaging the outcome y in stratum $(\mathbf{x}_{t-1}, z_t)$.

In the second stage, we estimate the blip effects by ML as follows. Two structural nested mean models are available: SNMM1 and SNMM2. In SNMM1, there are a total of nine blip effects of treatments, one φ of $z_1 = 1$, four $\varphi(\mathbf{x}_1)$ of $z_2 = 1$, and four $\varphi(\mathbf{x}_2)$ of $z_3 = 1$. In SNMM2, it is further required that $\varphi(\mathbf{x}_1) = \varphi(\mathbf{x}_2)$ if $\mathbf{x}_1 = \mathbf{x}_2$; therefore, there are four different blip effects for both $z_2 = 1$ and $z_3 = 1$ in addition to one blip effect for $z_1 = 1$.

With method (i), we impose SNMM1 on point observable effects. By decomposing $\vartheta(\mathbf{x}_{t-1})$ into the blip effects of $z_s = 1$ $(s \geq t)$, we obtain

$$\begin{aligned}\text{(26a)}\qquad \vartheta = \varphi &+ \sum_{t=2,3} \sum_{\mathbf{x}_{t-1}} \varphi(\mathbf{x}_{t-1}) \mathrm{P}^{\mathbf{O}}(\mathbf{x}_{t-1}, z_t = 1 \mid z_1 = 1) \\ &- \sum_{t=2,3} \sum_{\mathbf{x}_{t-1}} \varphi(\mathbf{x}_{t-1}) \mathrm{P}^{\mathbf{O}}(\mathbf{x}_{t-1}, z_t = 1 \mid z_1 = 0)\end{aligned}$$

$$\begin{aligned}\text{(26b)}\qquad \vartheta(\mathbf{x}_1) = \varphi(\mathbf{x}_1) &+ \sum_{\mathbf{x}_2} \varphi(\mathbf{x}_2) \mathrm{P}^{\mathbf{O}}(\mathbf{x}_2, z_3 = 1 \mid \mathbf{x}_1, z_2 = 1) \\ &- \sum_{\mathbf{x}_2} \varphi(\mathbf{x}_2) \mathrm{P}^{\mathbf{O}}(\mathbf{x}_2, z_3 = 1 \mid \mathbf{x}_1, z_2 = 0),\end{aligned}$$

(26c) $$\vartheta(\mathbf{x}_2) = \varphi(\mathbf{x}_2).$$

Interested readers may also derive these results by applying the Markov process to (17). All the probabilities $\mathrm{P}^{\mathbf{O}}(.)$ appearing in (26a), (26b) and (26c) are estimated by the corresponding proportions. We estimate the nine blip effects by regressing the nine estimated point observable effects on the proportions according to (26a), (26b) and (26c).

With method (ii), we impose SNMM2 on the point observable effects by letting $\varphi(\mathbf{x}_1) = \varphi(\mathbf{x}_2)$ for $\mathbf{x}_1 = \mathbf{x}_2$ in (26a), (26b) and (26c); therefore, only five different blip effects are estimated by regressing the nine estimated point observable effects on the proportions according to the updated version of (26a), (26b) and (26c).

In the third stage, we insert the estimated blip effects into the new G-formula (18) to obtain the ML estimate of the sequential causal effect $\mathrm{sce}(\mathbf{A}_1^3; \mathbf{B}_1^3)$ of deterministic regime $\mathbf{A}_1^3$ relative to $\mathbf{B}_1^3$. Two sequential causal effects are considered. The first one compares the static regime $\mathbf{A}_1^3 = (1, 1, 1)$ to $\mathbf{B}_1^3 = (0, 0, 0)$. The second one compares the dynamic regime $\mathbf{A}_1^3 = (1, 0, A_3)$ to the static regime $\mathbf{B}_1^3 = (0, 0, 0)$, where $A_3 = 1$ when $\mathbf{x}_2 = (0, 0)$ or $(0, 1)$ and $A_3 = 0$ otherwise.

II.2 Methods (iii), (iv) and (v). Method (iii) is constructed using the well-known G-formula (11) or (12) (Taubman et al. [11]). The sequential causal effect is according to (11)

$$\mathrm{sce}(\mathbf{A}_1^3; \mathbf{B}_1^3) =$$

$$\sum_{\mathbf{x}_1, \mathbf{x}_2} \mu(\mathbf{a}_1^3, \mathbf{x}_1^2) \prod_{t=1}^{2} \mathrm{P}^{\mathbf{O}}(\mathbf{x}_t \mid \mathbf{a}_1^t, \mathbf{x}_1^{t-1}) - \sum_{\mathbf{x}_1, \mathbf{x}_2} \mu(\mathbf{b}_1^3, \mathbf{x}_1^2) \prod_{t=1}^{2} \mathrm{P}^{\mathbf{O}}(\mathbf{x}_t \mid \mathbf{b}_1^t, \mathbf{x}_1^{t-1}).$$

The blip effect φ of $z_1 = 1$ is equal to $\mathrm{sce}(\mathbf{A}_1^3; \mathbf{B}_1^3)$ with $\mathbf{A}_1^3 = (1, 0, 0)$ and $\mathbf{B}_1^3 = (0, 0, 0)$. The blip effect of $z_2 = 1$ in stratum $(z_1, \mathbf{x}_1)$ is according to (12)

$$\begin{aligned} &\varphi(z_1, \mathbf{x}_1) \\ &= \sum_{\mathbf{x}_2} \mu(z_1, \mathbf{x}_1, z_2 = 1, \mathbf{x}_2, z_3 = 0) \mathrm{P}^{\mathbf{O}}(\mathbf{x}_2 \mid z_1, \mathbf{x}_1, z_2 = 1) \\ &- \sum_{\mathbf{x}_2} \mu(z_1, \mathbf{x}_1, z_2 = 0, \mathbf{x}_2, z_3 = 0) \mathrm{P}^{\mathbf{O}}(\mathbf{x}_2 \mid z_1, \mathbf{x}_1, z_2 = 0). \end{aligned}$$

Taking the average of this result over $\mathrm{P}^{\mathbf{O}}(z_1)$, we obtain $\varphi(\mathbf{x}_1)$ of $z_2 = 1$ in stratum $\mathbf{x}_1$. The blip effect of the last treatment $z_3 = 1$ in stratum $\mathbf{x}_2$ is

$$\varphi(\mathbf{x}_2) = \mu(\mathbf{x}_2, z_3 = 1) - \mu(\mathbf{x}_2, z_3 = 0).$$

Using the formulas above, we estimate nine blip effects and two sequential causal effects by ML via a total of $2^3 * 4^2 = 128$ standard parameters $\mu(\mathbf{z}_1^3, \mathbf{x}_1^2)$. We are not able to impose SNMM1 or SNMM2 on standard parameters to improve the estimation without excessive programming.

Method (iv) is constructed using the marginal structural model based on inverse probability weighting (Robins [6, 7]). Using the stabilized weights

$$w_1(z_1, \mathbf{x}_1, z_2, \mathbf{x}_2, z_3) = \frac{\mathrm{P}^{\mathbf{O}}(z_3 \mid z_1, z_2)\mathrm{P}^{\mathbf{O}}(z_2 \mid z_1)\mathrm{pr}^{\mathbf{O}}(z_1)}{\mathrm{P}^{\mathbf{O}}(z_3 \mid z_1, x_1, z_2, x_2)\mathrm{P}^{\mathbf{O}}(z_2 \mid z_1, x_1)\mathrm{pr}^{\mathbf{O}}(z_1)},$$

we obtain the weighted outcome Y^{w_1}, which satisfies $E(Y^{w_1} \mid z_1, z_2, z_3) = E\{Y(z_1, z_2, z_3)\}$. Hence, the sequential causal effect is

$$\mathrm{sce}(\mathbf{A}_1^3; \mathbf{B}_1^3) = E(Y^{w_1} \mid a_1, a_2, a_3) - E(Y^{w_1} \mid b_1, b_2, b_3).$$

The blip effect φ of $z_1 = 1$ is equal to $\mathrm{sce}(\mathbf{A}_1^3; \mathbf{B}_1^3)$ with $\mathbf{A}_1^3 = (1, 0, 0)$ and $\mathbf{B}_1^3 = (0, 0, 0)$. Using the stabilized weights

$$w_2(\mathbf{x}_1, z_2, \mathbf{x}_2, z_3) = \frac{\mathrm{pr}^{\mathbf{O}}(z_3 \mid \mathbf{x}_1, z_2)\mathrm{pr}^{\mathbf{O}}(z_2 \mid \mathbf{x}_1)}{\mathrm{pr}^{\mathbf{O}}(z_3 \mid \mathbf{x}_1, z_2, \mathbf{x}_2)\mathrm{pr}^{\mathbf{O}}(z_2 \mid \mathbf{x}_1)},$$

we obtain the weighted outcome Y^{w_2}, which satisfies $E(Y^{w_2} \mid \mathbf{x}_1, z_2, z_3) = E\{Y(z_2, z_3) \mid \mathbf{x}_1\}$. Hence, the blip effect of $z_2 = 1$ in stratum $\mathbf{x}_1$ is

$$\varphi(\mathbf{x}_1) = E(Y^{w_2} \mid \mathbf{x}_1, z_2 = 1, z_3 = 0) - E(Y^{w_2} \mid \mathbf{x}_1, z_2 = 0, z_3 = 0).$$

The blip effect of the last treatment $z_3 = 1$ in stratum $\mathbf{x}_2$ is

$$\varphi(\mathbf{x}_2) = \mu(\mathbf{x}_2, z_3 = 1) - \mu(\mathbf{x}_2, z_3 = 0).$$

Using these formulas, we estimate nine blip effects and one sequential causal effect of the static regime via a total of 32 parameters, which are $2^3 = 8$ weighted means $E(Y^{w_1} \mid z_1, z_2, z_3)$ and $4 * 2 * 2 = 16$ weighted means $E(Y^{w_2} \mid \mathbf{x}_1, z_2, z_3)$ and $4 * 2 = 8$ usual means $\mu(\mathbf{x}_2, z_3)$. We are not able to use the same weighting system to estimate the sequential causal effect of a dynamic regime. We are also not able to impose SNMM1 or SNMM2 on the weighted means of the outcome to improve the estimation without excessive programming.

Method (v) is constructed using the G-estimation based SNMM1 (Robins [5, 7]). The blip effect of the last treatment $z_3 = 1$ in stratum $\mathbf{x}_2$ is

$$\varphi(\mathbf{x}_2) = \mu(\mathbf{x}_2, z_3 = 1) - \mu(\mathbf{x}_2, z_3 = 0). \quad \text{(27a)}$$

Let the pseudo outcome after z_2 be $\tilde{Y}_2 = Y - \varphi(\mathbf{x}_2)$ when $z_3 = 1$ and $\tilde{Y}_2 = Y$ when $z_3 = 0$; therefore, the mean of the pseudo outcome $E(\tilde{Y}_2 \mid \mathbf{x}_1, z_2) = E\{Y(z_2, z_3 = 0) \mid \mathbf{x}_1\}$. Hence, the blip effect of $z_2 = 1$ in stratum $\mathbf{x}_1$ is

$$\varphi(\mathbf{x}_1) = E(\tilde{Y}_2 \mid \mathbf{x}_1, z_2 = 1) - E(\tilde{Y}_2 \mid \mathbf{x}_1, z_2 = 0). \tag{27b}$$

Let the pseudo outcome after z_1 be $\tilde{Y}_1 = \tilde{Y}_2 - \varphi(\mathbf{x}_1)$ when $z_2 = 1$ and $\tilde{Y}_1 = \tilde{Y}_2$ when $z_2 = 0$; therefore, the mean of the pseudo outcome $E(\tilde{Y}_1 \mid z_1) = E\{Y(z_1, z_2 = 0, z_3 = 0)\}$. Hence, the blip effect of $z_1 = 1$ is

$$\varphi = E(\tilde{Y}_1 \mid z_1 = 1) - E(\tilde{Y}_1 \mid z_1 = 0). \tag{27c}$$

Using these formulas, we estimate nine blip effects via a total of 18 parameters, which are $4 * 2 = 8$ $\mu(\mathbf{x}_2, z_3)$, $4 * 2 = 8$ $E(\tilde{Y}_2 \mid \mathbf{x}_1, z_2)$ and 2 $E(\tilde{Y}_1 \mid z_1)$. Due to the variability occuring when estimating $\varphi(\mathbf{x}_2)$, the pseudo outcome $\tilde{Y}_2$ has a different variability from Y. Similarly, the pseudo outcome $\tilde{Y}_1$ has a different variability from $\tilde{Y}_2$ and Y. In this simulation, we are not able to model these variabilities in a simple way and instead treat Y, $\tilde{Y}_2$ and $\tilde{Y}_1$ with equal variabilities. Furthermore, we are not able to estimate sequential causal effects and to impose SNMM2 on the means of the pseudo outcomes to improve the estimation without excessive programming.

With a correctly specified model for the variabilities of pseudo outcomes, we can prove that formula (27a) is equivalent to (26c), formula (27b) to (26b), and (27c) to (26a). For instance, inserting $\tilde{Y}_2$ into $E(\tilde{Y}_2 \mid \mathbf{x}_1, z_2)$ and noticing $\varphi(\mathbf{x}_2) = \varphi(\mathbf{x}_2; z_3 = 1)$ and $\varphi(\mathbf{x}_2; z_3 = 0) = 0$, we obtain

$$E(\tilde{Y}_2 \mid \mathbf{x}_1, z_2) = \mu(\mathbf{x}_1, z_2) - \sum_{\mathbf{x}_2} \varphi(\mathbf{x}_2) \mathrm{P}^{\mathbf{O}}(\mathbf{x}_2, z_3 = 1 \mid \mathbf{x}_1, z_2).$$

Additionally, using $\vartheta(\mathbf{x}_1) = \mu(\mathbf{x}_1, z_2 = 1) - \mu(\mathbf{x}_1, z_2 = 0)$, we see equivalence between (27b) and (26b). Therefore, method (v) is equivalent to method (i) under a correctly specified model for these variabilities. However, method (i) does not need a specification for these variabilities.

II.3 Application of methods (i)–(v) to the simulation. Using the standard parameters obtained in the next subsection, three data-generating mechanisms are constructed for normal, dichotomous and Poisson outcomes. From each of the three data-generating mechanisms, we generate a total of 1000 data sets, each having 400 independent observations on $(\mathbf{Z}_1^3, \mathbf{X}_1^2, Y)$.

We apply methods (i)–(v) to 1000 data sets for the normal, dichotomous and Poisson outcomes. For every method, we obtain 1000 estimates for each of the blip and sequential causal effects, from which we obtain one average estimate and one variance estimate for the causal effect. To obtain the confidence interval, we use the bootstrap method in which 1000 data sets are resampled with replacement from each data set. For every method, we obtain 1000 95 % bootstrap percentile confidence intervals for each of these causal effects, from which we obtain the actual coverage probability for the causal effect.

Both methods (i) and (ii) only need nine point observable effects to estimate all blip and sequential causal effects. Method (iii) needs 128 standard parameters to estimate these causal effects. Method (iv) needs 32 parameters and a specification of the stabilized weights to estimate the blip effects and only sequential causal effects of the static regimes without excessive programming. Method (v) needs 18 parameters and a specification of the variability of pseudo outcomes to estimate only the blip effects without excessive programming. The results of the simulation are presented in Table 1, from which the following comparisons can be made between these methods.

Regarding the estimates for the blip and sequential causal effects, the following observations can be made. First, all methods yield unbiased estimates for the blip effects of $z_3 = 1$ (columns (**f**)–(**i**) in Table 1). Noticeably, methods (i), (iii), (iv) and (v) yield nearly the same estimates due to the setting of the simulation. Second, methods (i), (ii), (iv) and (v) yield unbiased estimates for the blip effects of $z_1 = 1$ and $z_2 = 1$ (columns (**a**)–(**e**) in Table 1) and two sequential causal effects (columns (**j**) and (**k**) in Table 1), whereas method (iii) yields estimates of certain biases for these causal effects indicating slow convergence of these estimates to the true values.

The following interesting observations can be made for variances in the estimates of the blip and sequential causal effects as well as actual coverage probabilities for the confidence intervals of these causal effects. First, methods (i), (iii), (iv) and (v) yield nearly the same variances and coverage probabilities for the blip effects of $z_3 = 1$ (columns (**f**)–(**i**) in Table 1) due to the setting of the simulation. Second, method (i) achieves considerably smaller variances and more accurate coverage probabilities than methods (iii) and (iv) do for the blip effects of $z_1 = 1$ and $z_2 = 1$ (columns (**a**)–(**e**) in Table 1) and two sequential causal effects (columns (**j**) and (**k**) in Table 1). Third, method (ii) achieves the smallest variances of all methods and nearly the nominal coverage probabilities, which implies that an unsaturated model may improve the estimation without causing biases. Fourth, method (v) has even smaller variances than method (i), although method (i) is based on ML and theoretically yields the variance bounds under SNMM1. As shown in the previous subsection of Supplementary II, this result is due to an inadequate evaluation of the variability of pseudo outcomes in method (v). With a correct specification of this variability, method (v) is equivalent to method (i). Method (i), however, does not need a specification of this variability.

Additionally, we have constructed a misspecified model for the standard parameters $\mu(\mathbf{z}_1^3, \mathbf{x}_1^2)$ in which equalities are imposed between these parameters and then applied method (iii) to estimate the blip and sequential causal effects by ML in the simulation. We have observed the null paradox, that is, the biases for the ML estimates of these causal effects (see discussion in the introduction and references therein). The results are not shown in the paper, although the SAS code is given in Supplementary Material for interested readers.

II.4 Construction of the standard parameters for the normal, dichotomous and Poisson outcomes. The probabilities of treatments and covariates are the same for the three outcomes and presented in Table $II1$. Here, we will use the point observable effects of treatments, the point observable effects of covariates and the grand mean to construct the standard parameters for the conditional distribution of the outcome given all treatments and covariates, which correspond to true values of the blip effects (Wang and Yin [12]).

As described in Section 5.3, the blip effects are $\phi(z_1 = 1) = \varphi$, $\phi(z_1, \mathbf{x}_1; z_2 = 1) = \varphi(\mathbf{x}_1)$ and $\phi(z_1, \mathbf{x}_1, z_2, \mathbf{x}_2; z_3 = 1) = \varphi(\mathbf{x}_2)$. Inserting these blip effects into formula (17), we obtain the formula for calculating the point observable effect of treatment

$$\begin{cases} \vartheta(z_1 = 1) = \varphi + \sum_{t=2,3} \sum_{\mathbf{x}_{t-1}} \varphi(\mathbf{x}_{t-1}) \mathrm{P}^{\mathbf{O}}(\mathbf{x}_{t-1}, z_t = 1 \mid z_1 = 1) \\ \qquad\qquad - \sum_{t=2,3} \sum_{\mathbf{x}_{t-1}} \varphi(\mathbf{x}_{t-1}) \mathrm{P}^{\mathbf{O}}(\mathbf{x}_{t-1}, z_t = 1 \mid z_1 = 0), \\ \vartheta(z_1, \mathbf{x}_1; z_2 = 1) = \varphi(\mathbf{x}_1) + \sum_{\mathbf{x}_2} \varphi(\mathbf{x}_2) \mathrm{P}^{\mathbf{O}}(\mathbf{x}_2, z_3 = 1 \mid z_1, \mathbf{x}_1, z_2 = 1) \\ \qquad\qquad - \sum_{\mathbf{x}_2} \varphi(\mathbf{x}_2) \mathrm{P}^{\mathbf{O}}(\mathbf{x}_2, z_3 = 1 \mid z_1, \mathbf{x}_1, z_2 = 0), \\ \vartheta(z_1, \mathbf{x}_1, z_2, \mathbf{x}_2; z_3 = 1) = \varphi(\mathbf{x}_2). \end{cases}$$

Using the true values of the blip effects and the probabilities of treatments and covariates given in Table $II1$, we calculate these point observable effects of the treatments.

The point observable effect of covariate $\mathbf{x}_{t-1} \neq \mathbf{0}$ is

$$\zeta(\mathbf{z}_1^{t-1}, \mathbf{x}_1^{t-2}; \mathbf{x}_{t-1}) = \mu(\mathbf{z}_1^{t-1}, \mathbf{x}_1^{t-2}, \mathbf{x}_{t-1}) - \mu(\mathbf{z}_1^{t-1}, \mathbf{x}_1^{t-2}, \mathbf{x}_{t-1} = \mathbf{0}),$$

where $\mu(\mathbf{z}_1^{t-1}, \mathbf{x}_1^{t-1}) = E(y \mid \mathbf{z}_1^{t-1}, \mathbf{x}_1^{t-1})$. The grand mean is $\mu = E(y)$. According to formula (15), the blip effects are only functions of the point observable effects of treatments; therefore, the point observable effect $\zeta(\mathbf{z}_1^{t-1}, \mathbf{x}_1^{t-2}; \mathbf{x}_{t-1})$ of covariate and the grand mean μ can be arbitrarily chosen to yield the same blip effects. However, the choice should allow for an appropriate mean of the distribution, e.g., the mean must have a range of $(0,1)$ for a dichotomous outcome.

For the normal distribution, we choose the point effects of covariates

$$\zeta(z_1;\mathbf{x}_1)=\begin{cases}10+5z_1, & \mathbf{x}_1=(0,1)\\ 12+5z_1, & \mathbf{x}_1=(1,0)\\ 13+5z_1, & \mathbf{x}_1=(1,1)\end{cases}$$

for $z_1=0,1$, and

$$\zeta(z_1,\mathbf{x}_1,z_2;\mathbf{x}_2)=\begin{cases}10-5z_1-2z_2+f(\mathbf{x}_1), & \mathbf{x}_2=(0,1)\\ 12-5z_1-2z_2+f(\mathbf{x}_1), & \mathbf{x}_2=(1,0)\\ 10-5z_1-3z_2+f(\mathbf{x}_1), & \mathbf{x}_2=(1,1)\end{cases}$$

for $z_1=0,1$, $z_2=0,1$ and $f(\mathbf{x}_1)=0,3,6,9$ when $\mathbf{x}_1=(0,0),(0,1),(1,0),(1,1)$ respectively. We choose the grand mean as $\mu=15$.

For the dichotomous outcome, we choose the point effects of covariates

$$\zeta(z_1;\mathbf{x}_1)=\begin{cases}0.1z_1, & \mathbf{x}_1=(0,1)\\ 0.2z_1, & \mathbf{x}_1=(1,0)\\ -0.1z_1, & \mathbf{x}_1=(1,1)\end{cases}$$

for $z_1=0,1$, and

$$\zeta(z_1,\mathbf{x}_1,z_2;\mathbf{x}_2)=\begin{cases}-0.1z_2, & \mathbf{x}_2=(0,1)\\ -0.2z_2, & \mathbf{x}_2=(1,0)\\ 0.1z_2, & \mathbf{x}_2=(1,1)\end{cases}$$

for $z_1=0,1$, $z_2=0,1$ and $\mathbf{x}_1=(0,0),(0,1),(1,0),(1,1)$. We choose the grand mean as $\mu=0.5$.

For the Poisson outcome, we choose the same point effects of covariates as those for the dichotomous outcome. As for the grand mean, we choose $\mu=10$, which allows for a Poisson distribution far from the normal one.

Finally, we use the obtained $\vartheta(\mathbf{z}_1^{t-1},\mathbf{x}_1^{t-1};z_t)$, $\zeta(\mathbf{z}_1^{t},\mathbf{x}_1^{t-1};\mathbf{x}_t)$ and μ to construct the standard parameter $\mu(z_1,\mathbf{x}_1,z_2,\mathbf{x}_2,z_3)$ by applying formula (16) of Wang and Yin [12], that is,

$$\mu(z_1,\mathbf{x}_1,z_2,\mathbf{x}_2,z_3)=-\sum_{t=1}^{3}\vartheta(\mathbf{z}_1^{t-1},\mathbf{x}_1^{t-1};z_t)\{\mathrm{P}^{\mathbf{O}}(z_t^*=1\mid\mathbf{z}_1^{t-1},\mathbf{x}_1^{t-1})-I(z_t)\}$$
$$-\sum_{t=1}^{2}\left\{\sum_{\mathbf{x}_t^*\neq\mathbf{0}}\zeta(\mathbf{z}_1^{t},\mathbf{x}_1^{t-1};\mathbf{x}_t^*)\mathrm{P}^{\mathbf{O}}(\mathbf{x}_t^*\mid\mathbf{z}_1^{t},\mathbf{x}_1^{t-1})-\zeta(\mathbf{z}_1^{t},\mathbf{x}_1^{t-1};\mathbf{x}_t)\right\}+\mu,$$

where $I(z_t)$ equals one when $z_t=1$ and zero otherwise. The obtained standard parameters are presented in Tables $II2$-$II4$ for the normal, Bernoulli and Poisson distributions.

SUPPLEMENT III: MEDICAL EXAMPLE IN SECTION 5.4

Here, we provide details about the medical study in Section 5.4. The data and relevant SAS codes used in the analysis are also included in Supplementary Material.

III.1 Medical background and the data. In this medical study (Zeger and Diggle [14]), each participant was required to visit medical centers and provide information about the physical conditions and the use of various drugs including recreational drugs between the current and previous visits. Suppose that the drug use occurred prior to the physical conditions. The treatment variables are the drug use Z_t at time $t = 0, 1, 2$. The physical conditions between Z_t and Z_{t+1} include the CD4 count (X_{t1}), the number of packs of cigarettes smoked daily (X_{t2}), the number of sexual partners (X_{t3}) and a mental illness score (X_{t4}). Age (X_{05}) is also included as a stationary covariate at visit $t = 0$. Consequently, the temporal order of these variables is $\{Z_0, (X_{01}, X_{02}, X_{03}, X_{04}, X_{05}), Z_1, (X_{11}, X_{12}, X_{13}, X_{14}), Z_2, (X_{21}, X_{22}, X_{23}, X_{24})\}$.

Let $\mathbf{X}_0 = (Z_0, X_{01}, X_{02}, X_{03}, X_{04}, X_{05})$ be the stationary covariate vector, and $\mathbf{X}_1 = (X_{11}, X_{12}, X_{13}, X_{14})$ be the time-dependent covariate vector between drug uses Z_1 and Z_2. The treatment variables of interest are the drug uses Z_1 and Z_2. The outcome is the logarithm of CD4 count at $t = 2$, that is, $Y = \log(X_{21})$. All variables prior to Y are dichotomized, with ones implying 'yes' or 'high' and zeros 'no' or 'low'. We assume that the outcome Y is normally distributed. We also assume that the identifying condition described in Section 4.1 is satisfied.

We wish to estimate the blip and sequential causal effects. Wang and Yin [12] estimated the blip effect by incorporating variability of only the outcome. Now, we estimate both the blip and sequential causal effects by incorporating variability of not only the outcome but also the treatments and covariates.

We will apply our method and three methods in the literature to this study. These methods estimate the blip effects separately at different times and are similar to methods (i), (iii) and (iv) and (v) in the simulation study of Section 5.3 or Supplement II of Supplementary Material; therefore, we still refer to them as methods (i), (iii), (iv) and (v). That is, method (i) is constructed using our method described in Section 5.1; method (iii) is constructed using the well-known G-formula (11) or (12) (Taubman et al. [11]); method (iv) is constructed using the marginal structural model based on inverse probability weighting (Robins [6, 7]); method (v) is constructed using the G-estimation based on SNMM (Robins [5, 7]).

III.2 Method (i). Based on the data of this medical study, we have not found any sensible assignment conditions for drug uses Z_1 and Z_2. By the likelihood ratio-based significance test of the covariates at the significance level of 10 %, we find the model for $\mu(x_{01}, x_{11}, z_2)$

$$\mu(x_{01}, x_{11}, z_2) = \alpha_2 + z_2\vartheta_2 + x_{01}\gamma_2 + x_{11}\gamma_3, \tag{28}$$

where $\vartheta_2 = \vartheta(x_{01}, x_{11})$ is the point observable effect of $z_2 = 1$, which is the same for all (x_{01}, x_{11}) as seen from (28). We also obtain the model for $\mu(x_{01}, z_1)$

$$\mu(x_{01}, z_1) = \alpha_1 + z_1\vartheta_1 + x_{01}\gamma_1, \tag{29}$$

where $\vartheta_1 = \vartheta(x_{01})$ is the point observable effect of $z_1 = 1$, which is the same for all x_{01} as seen from (29). Here, the variance of the ML estimate $\hat{\vartheta}_1$ is estimated by incorporating the variability of covariate x_{11}, treatment z_2 and the outcome y. Noticeably, covariate x_{01} and treatment z_1 are ancillary for the model parameters and thus their variability has little influence on the variance of $\hat{\vartheta}_1$. From (28) and (29), we see that the covariates relevant to the blip and sequential causal effects are x_{01} and x_{11}.

In the first stage, we estimate the point observable effects by ML as follows. We assume a unit variance for Y given $(x_{01}, z_1, x_{11}, z_2)$. Based on the obtained models (28) and (29), we calculate the ML estimate $\hat{\vartheta}_2$ and $\hat{\vartheta}_1$ and their variances.

In the second stage, we estimate the blip effects by ML as follows. Given the small sample, it is reasonable to assume that the blip effects follow a simple SNMM: the blip effects $\phi(x_{01})$ of $z_1 = 1$ are the same for all x_{01} and equal to φ_1; the blip effects $\phi(x_{01}, z_1, x_{11})$ of $z_2 = 1$ are the same for all (x_{01}, z_1, x_{11}) and equal to φ_2. Decomposing $\vartheta_1 = \vartheta(x_{01})$ and $\vartheta_2 = \vartheta(x_{01}, x_{11})$ into the blip effects φ_1 and φ_2, we obtain the following model for the point observable effects

$$\begin{cases} \vartheta_1 = \varphi_1 + \varphi_2\{\mathrm{P}^{\mathbf{O}}(z_2 = 1 \mid z_1 = 1) - \mathrm{P}^{\mathbf{O}}(z_2 = 1 \mid z_1 = 0)\}, \\ \vartheta_2 = \varphi_2. \end{cases} \tag{30}$$

The probabilities $\mathrm{P}^{\mathbf{O}}(.)$ in (30) are estimated by the corresponding proportions $\hat{\mathrm{P}}^{\mathbf{O}}(.)$. We obtain the ML estimates $\hat{\varphi}_1$ and $\hat{\varphi}_2$ by a regression of the ML estimates $\hat{\vartheta}_1$ and $\hat{\vartheta}_2$ on the proportions $\hat{\mathrm{P}}^{\mathbf{O}}(.)$ based on (30).

In the third stage, we estimate the sequential causal effect $\mathrm{sce}(\mathbf{A}_1^2; \mathbf{B}_1^2)$ of deterministic regimes $\mathbf{A}_1^2$ relative to $\mathbf{B}_1^2$ by ML as follows. By replacing the blip effects by $\hat{\varphi}_1$ and $\hat{\varphi}_2$ in the new G-formula (18), we obtain the ML estimate $\widehat{\mathrm{sce}}(\mathbf{A}_1^2; \mathbf{B}_1^2; x_{01})$ in subpopulation x_{01}. Taking the average of the ML estimate with respect to the probability of x_{01} estimated by the corresponding proportion, we obtain the ML estimate $\widehat{\text{sce}}(\mathbf{A}_1^2; \mathbf{B}_1^2)$ in the population. Two sequential causal effects are considered. The first one compares the static regime $\mathbf{A}_1^2 = (1, 1)$ to $\mathbf{B}_1^2 = (0, 0)$. The second one compares the dynamic regime $\mathbf{A}_1^2 = (1, A_2)$ to the static regime $\mathbf{B}_1^2 = (0, 0)$, where $A_2 = 1$ when $\mathbf{x}_{11} = 0$ and $A_2 = 0$ otherwise.

In the fourth stage, the variances and confidence intervals for the blip and sequential causal effects are obtained by the bootstrap procedure, where 1000 bootstrap data sets are generated. The results are presented in Table 2.

If $(\mathbf{x}_0^{t-1}, \mathbf{z}_0^{t-1})$ does not sufficiently describe the population before time t, then Z_t has unmeasured confounders and the identifying condition is not satisfied. In this case, φ_1 and φ_2 do not describe the blip effects nor does $\text{sce}(\mathbf{A}_1^2; \mathbf{B}_1^2)$ describe the sequential causal effect, although these parameters reflect certain associations between the treatments and outcome. A sensitivity analysis can be carried out using the same estimation method to investigate the influences of possible unmeasured confounders on the estimation of the blip and sequential causal effects.

III.3 Methods (iii), (iv) and (v). The analysis above shows that the covariates relevant to the causal effects are X_{01} and X_{11}. With these two covariates as well as the two treatment variables Z_1 and Z_2, we can readily apply methods (iii), (iv) and (v) to the data. The results are presented in Table 2.

III.4 Discussion. Table 2 shows that method (i) estimates all the blip and sequential causal effects while achieving the smallest variance. Furthermore, method (i) estimates these causal effects via only two point observable effects of $z_1 = 1$ and $z_2 = 1$.

Table 2 also shows that the results obtained by all methods indicate that the recreational drug has a decreasing distant effect on the CD4 count compared with the increasing short-term effect. This finding has not been medically reported to our knowledge.

Table II1 The probabilities for treatments and covariates $(z_1, \boldsymbol{x}_1, z_2, \boldsymbol{x}_2, z_3)$, where $\boldsymbol{x}_1 = (x_{11}, x_{12})$ and $\boldsymbol{x}_2 = (x_{21}, x_{22})$. It is the same for all the three data-generating mechanisms corresponding to normal, dichotomous and Poisson outcomes.

Table II2 The conditional means $\mu(\boldsymbol{z}_1^3, \boldsymbol{x}_1^2)$ of the normal outcome *Y* given the treatments and covariates $(\boldsymbol{z}_1^3, \boldsymbol{x}_1^2)$ which correspond to the blip effects shown in Table 1. The construction of $\mu(\boldsymbol{z}_1^3, \boldsymbol{x}_1^2)$ is explicated in Section II.4 of this supplementary material. The SAS codes yielding the means are included in the SAS codes library of this supplementary material.

Table II3 The conditional means $\mu(\boldsymbol{z}_1^3, \boldsymbol{x}_1^2)$ of the binomial outcome *Y* given the treatments and covariates $(\boldsymbol{z}_1^3, \boldsymbol{x}_1^2)$ which correspond to the blip effects shown in Table 1. The construction of $\mu(\boldsymbol{z}_1^3, \boldsymbol{x}_1^2)$ is explicated in Section II.4 of this supplementary material. The SAS code yielding the means is included in the SAS codes library of this supplementary material.

Table II4 The conditional means $\mu(\boldsymbol{z}_1^3, \boldsymbol{x}_1^2)$ of the Poisson outcome *Y* given the treatments and covariates $(\boldsymbol{z}_1^3, \boldsymbol{x}_1^2)$ which correspond to the blip effects shown in Table 1. The construction of $\mu(\boldsymbol{z}_1^3, \boldsymbol{x}_1^2)$ is explicated in Section II.4 file of this supplementary material. The SAS code yielding the means is included in the SAS codes library of this supplementary material.

Table II1

$P^{\mathbf{O}}(z_1 = 1)$	1/2
$P^{\mathbf{O}}(\boldsymbol{x}_1 \mid z_1)$	$P^{\mathbf{O}}(0,0 \mid 0) = 1/6$, $P^{\mathbf{O}}(0,1 \mid 0) = 2/6$, $P^{\mathbf{O}}(1,0 \mid 0) = 2/6$, $P^{\mathbf{O}}(0,0 \mid 1) = 2/6$, $P^{\mathbf{O}}(0,1 \mid 1) = 1/6$, $P^{\mathbf{O}}(1,0 \mid 1) = 1/6$
$P^{\mathbf{O}}(z_2 \mid \boldsymbol{x}_1) = P^{\mathbf{O}}(z_2 \mid z_1, \boldsymbol{x}_1)$	$P^{\mathbf{O}}(1 \mid 0,0) = 1/3$, $P^{\mathbf{O}}(1 \mid 0,1) = 2/3$, $P^{\mathbf{O}}(1 \mid 1,0) = 2/3$, $P^{\mathbf{O}}(1 \mid 1,1) = 1/3$
$P^{\mathbf{O}}(\boldsymbol{x}_2 \mid z_2) = P^{\mathbf{O}}(\boldsymbol{x}_2 \mid z_1, \boldsymbol{x}_1, z_2)$	$P^{\mathbf{O}}(0,0 \mid 0) = 1/6$, $P^{\mathbf{O}}(0,1 \mid 0) = 2/6$, $P^{\mathbf{O}}(1,0 \mid 0) = 2/6$, $P^{\mathbf{O}}(0,0 \mid 1) = 2/6$, $P^{\mathbf{O}}(0,1 \mid 1) = 1/6$, $P^{\mathbf{O}}(1,0 \mid 1) = 1/6$
$P^{\mathbf{O}}(z_3 \mid \boldsymbol{x}_2) = P^{\mathbf{O}}(z_3 \mid z_1, \boldsymbol{x}_1, z_2, \boldsymbol{x}_2)$	$P^{\mathbf{O}}(1 \mid 0,0) = 1/3$, $P^{\mathbf{O}}(1 \mid 0,1) = 2/3$, $P^{\mathbf{O}}(1 \mid 1,0) = 2/3$, $P^{\mathbf{O}}(1 \mid 1,1) = 1/3$

z1	x1	z2	x2	z3	mean
0	(0,0)	0	(0,0)	0	-2,333333333
0	(0,0)	0	(0,0)	1	-6,333333333
0	(0,0)	0	(0,1)	0	9,666666667
0	(0,0)	0	(0,1)	1	4,666666667
0	(0,0)	0	(1,0)	0	5
0	(0,0)	0	(1,0)	1	10
0	(0,0)	0	(1,1)	0	5
0	(0,0)	0	(1,1)	1	9
0	(0,0)	1	(0,0)	0	-2,666666667
0	(0,0)	1	(0,0)	1	-6,666666667
0	(0,0)	1	(0,1)	0	7,333333333
0	(0,0)	1	(0,1)	1	2,333333333
0	(0,0)	1	(1,0)	0	2,666666667
0	(0,0)	1	(1,0)	1	7,666666667
0	(0,0)	1	(1,1)	0	1,666666667
0	(0,0)	1	(1,1)	1	5,666666667
0	(0,1)	0	(0,0)	0	7,166666667
0	(0,1)	0	(0,0)	1	3,166666667
0	(0,1)	0	(0,1)	0	22,16666667
0	(0,1)	0	(0,1)	1	17,16666667
0	(0,1)	0	(1,0)	0	17,5
0	(0,1)	0	(1,0)	1	22,5
0	(0,1)	0	(1,1)	0	17,5
0	(0,1)	0	(1,1)	1	21,5
0	(0,1)	1	(0,0)	0	6,333333333
0	(0,1)	1	(0,0)	1	2,333333333
0	(0,1)	1	(0,1)	0	19,33333333
0	(0,1)	1	(0,1)	1	14,33333333
0	(0,1)	1	(1,0)	0	14,66666667
0	(0,1)	1	(1,0)	1	19,66666667
0	(0,1)	1	(1,1)	0	13,66666667
0	(0,1)	1	(1,1)	1	17,66666667
0	(1,0)	0	(0,0)	0	1,77636E-15
0	(1,0)	0	(0,0)	1	-4
0	(1,0)	0	(0,1)	0	18
0	(1,0)	0	(0,1)	1	13
0	(1,0)	0	(1,0)	0	13,33333333
0	(1,0)	0	(1,0)	1	18,33333333
0	(1,0)	0	(1,1)	0	13,33333333
0	(1,0)	0	(1,1)	1	17,33333333
0	(1,0)	1	(0,0)	0	9,666666667
0	(1,0)	1	(0,0)	1	5,666666667
0	(1,0)	1	(0,1)	0	25,66666667
0	(1,0)	1	(0,1)	1	20,66666667
0	(1,0)	1	(1,0)	0	21
0	(1,0)	1	(1,0)	1	26
0	(1,0)	1	(1,1)	0	20
0	(1,0)	1	(1,1)	1	24
0	(1,1)	0	(0,0)	0	0,5
0	(1,1)	0	(0,0)	1	-3,5
0	(1,1)	0	(0,1)	0	21,5
0	(1,1)	0	(0,1)	1	16,5
0	(1,1)	0	(1,0)	0	16,83333333
0	(1,1)	0	(1,0)	1	21,83333333
0	(1,1)	0	(1,1)	0	16,83333333

0	(1,1)	0	(1,1)	1	20,83333333
0	(1,1)	1	(0,0)	0	9,666666667
0	(1,1)	1	(0,0)	1	5,666666667
0	(1,1)	1	(0,1)	0	28,66666667
0	(1,1)	1	(0,1)	1	23,66666667
0	(1,1)	1	(1,0)	0	24
0	(1,1)	1	(1,0)	1	29
0	(1,1)	1	(1,1)	0	23
0	(1,1)	1	(1,1)	1	27
1	(0,0)	0	(0,0)	0	3
1	(0,0)	0	(0,0)	1	-1
1	(0,0)	0	(0,1)	0	10
1	(0,0)	0	(0,1)	1	5
1	(0,0)	0	(1,0)	0	5,333333333
1	(0,0)	0	(1,0)	1	10,33333333
1	(0,0)	0	(1,1)	0	5,333333333
1	(0,0)	0	(1,1)	1	9,333333333
1	(0,0)	1	(0,0)	0	1,833333333
1	(0,0)	1	(0,0)	1	-2,166666667
1	(0,0)	1	(0,1)	0	6,833333333
1	(0,0)	1	(0,1)	1	1,833333333
1	(0,0)	1	(1,0)	0	2,166666667
1	(0,0)	1	(1,0)	1	7,166666667
1	(0,0)	1	(1,1)	0	1,166666667
1	(0,0)	1	(1,1)	1	5,166666667
1	(0,1)	0	(0,0)	0	17,5
1	(0,1)	0	(0,0)	1	13,5
1	(0,1)	0	(0,1)	0	27,5
1	(0,1)	0	(0,1)	1	22,5
1	(0,1)	0	(1,0)	0	22,83333333
1	(0,1)	0	(1,0)	1	27,83333333
1	(0,1)	0	(1,1)	0	22,83333333
1	(0,1)	0	(1,1)	1	26,83333333
1	(0,1)	1	(0,0)	0	15,83333333
1	(0,1)	1	(0,0)	1	11,83333333
1	(0,1)	1	(0,1)	0	23,83333333
1	(0,1)	1	(0,1)	1	18,83333333
1	(0,1)	1	(1,0)	0	19,16666667
1	(0,1)	1	(1,0)	1	24,16666667
1	(0,1)	1	(1,1)	0	18,16666667
1	(0,1)	1	(1,1)	1	22,16666667
1	(1,0)	0	(0,0)	0	10,33333333
1	(1,0)	0	(0,0)	1	6,333333333
1	(1,0)	0	(0,1)	0	23,33333333
1	(1,0)	0	(0,1)	1	18,33333333
1	(1,0)	0	(1,0)	0	18,66666667
1	(1,0)	0	(1,0)	1	23,66666667
1	(1,0)	0	(1,1)	0	18,66666667
1	(1,0)	0	(1,1)	1	22,66666667
1	(1,0)	1	(0,0)	0	19,16666667
1	(1,0)	1	(0,0)	1	15,16666667
1	(1,0)	1	(0,1)	0	30,16666667
1	(1,0)	1	(0,1)	1	25,16666667
1	(1,0)	1	(1,0)	0	25,5
1	(1,0)	1	(1,0)	1	30,5
1	(1,0)	1	(1,1)	0	24,5

1	(1,0)	1	(1,1)	1	28,5
1	(1,1)	0	(0,0)	0	10,83333333
1	(1,1)	0	(0,0)	1	6,833333333
1	(1,1)	0	(0,1)	0	26,83333333
1	(1,1)	0	(0,1)	1	21,83333333
1	(1,1)	0	(1,0)	0	22,16666667
1	(1,1)	0	(1,0)	1	27,16666667
1	(1,1)	0	(1,1)	0	22,16666667
1	(1,1)	0	(1,1)	1	26,16666667
1	(1,1)	1	(0,0)	0	19,16666667
1	(1,1)	1	(0,0)	1	15,16666667
1	(1,1)	1	(0,1)	0	33,16666667
1	(1,1)	1	(0,1)	1	28,16666667
1	(1,1)	1	(1,0)	0	28,5
1	(1,1)	1	(1,0)	1	33,5
1	(1,1)	1	(1,1)	0	27,5
1	(1,1)	1	(1,1)	1	31,5

z1	x1	z2	x2	z3	mean
0	(0,0)	0	(0,0)	0	0,532407407
0	(0,0)	0	(0,0)	1	0,382407407
0	(0,0)	0	(0,1)	0	0,382407407
0	(0,0)	0	(0,1)	1	0,532407407
0	(0,0)	0	(1,0)	0	0,415740741
0	(0,0)	0	(1,0)	1	0,515740741
0	(0,0)	0	(1,1)	0	0,515740741
0	(0,0)	0	(1,1)	1	0,415740741
0	(0,0)	1	(0,0)	0	0,357407407
0	(0,0)	1	(0,0)	1	0,207407407
0	(0,0)	1	(0,1)	0	0,107407407
0	(0,0)	1	(0,1)	1	0,257407407
0	(0,0)	1	(1,0)	0	0,040740741
0	(0,0)	1	(1,0)	1	0,140740741
0	(0,0)	1	(1,1)	0	0,440740741
0	(0,0)	1	(1,1)	1	0,340740741
0	(0,1)	0	(0,0)	0	0,396296296
0	(0,1)	0	(0,0)	1	0,246296296
0	(0,1)	0	(0,1)	0	0,246296296
0	(0,1)	0	(0,1)	1	0,396296296
0	(0,1)	0	(1,0)	0	0,27962963
0	(0,1)	0	(1,0)	1	0,37962963
0	(0,1)	0	(1,1)	0	0,37962963
0	(0,1)	0	(1,1)	1	0,27962963
0	(0,1)	1	(0,0)	0	0,521296296
0	(0,1)	1	(0,0)	1	0,371296296
0	(0,1)	1	(0,1)	0	0,271296296
0	(0,1)	1	(0,1)	1	0,421296296
0	(0,1)	1	(1,0)	0	0,20462963
0	(0,1)	1	(1,0)	1	0,30462963
0	(0,1)	1	(1,1)	0	0,60462963
0	(0,1)	1	(1,1)	1	0,50462963
0	(1,0)	0	(0,0)	0	0,42962963
0	(1,0)	0	(0,0)	1	0,27962963
0	(1,0)	0	(0,1)	0	0,27962963
0	(1,0)	0	(0,1)	1	0,42962963
0	(1,0)	0	(1,0)	0	0,312962963
0	(1,0)	0	(1,0)	1	0,412962963
0	(1,0)	0	(1,1)	0	0,412962963
0	(1,0)	0	(1,1)	1	0,312962963
0	(1,0)	1	(0,0)	0	0,50462963
0	(1,0)	1	(0,0)	1	0,35462963
0	(1,0)	1	(0,1)	0	0,25462963
0	(1,0)	1	(0,1)	1	0,40462963
0	(1,0)	1	(1,0)	0	0,187962963
0	(1,0)	1	(1,0)	1	0,287962963
0	(1,0)	1	(1,1)	0	0,587962963
0	(1,0)	1	(1,1)	1	0,487962963
0	(1,1)	0	(0,0)	0	0,515740741
0	(1,1)	0	(0,0)	1	0,365740741
0	(1,1)	0	(0,1)	0	0,365740741
0	(1,1)	0	(0,1)	1	0,515740741
0	(1,1)	0	(1,0)	0	0,399074074
0	(1,1)	0	(1,0)	1	0,499074074
0	(1,1)	0	(1,1)	0	0,499074074

0	(1,1)	0	(1,1)	1	0,399074074
0	(1,1)	1	(0,0)	0	0,390740741
0	(1,1)	1	(0,0)	1	0,240740741
0	(1,1)	1	(0,1)	0	0,140740741
0	(1,1)	1	(0,1)	1	0,290740741
0	(1,1)	1	(1,0)	0	0,074074074
0	(1,1)	1	(1,0)	1	0,174074074
0	(1,1)	1	(1,1)	0	0,474074074
0	(1,1)	1	(1,1)	1	0,374074074
1	(0,0)	0	(0,0)	0	0,678703704
1	(0,0)	0	(0,0)	1	0,528703704
1	(0,0)	0	(0,1)	0	0,528703704
1	(0,0)	0	(0,1)	1	0,678703704
1	(0,0)	0	(1,0)	0	0,562037037
1	(0,0)	0	(1,0)	1	0,662037037
1	(0,0)	0	(1,1)	0	0,662037037
1	(0,0)	0	(1,1)	1	0,562037037
1	(0,0)	1	(0,0)	0	0,503703704
1	(0,0)	1	(0,0)	1	0,353703704
1	(0,0)	1	(0,1)	0	0,253703704
1	(0,0)	1	(0,1)	1	0,403703704
1	(0,0)	1	(1,0)	0	0,187037037
1	(0,0)	1	(1,0)	1	0,287037037
1	(0,0)	1	(1,1)	0	0,587037037
1	(0,0)	1	(1,1)	1	0,487037037
1	(0,1)	0	(0,0)	0	0,642592593
1	(0,1)	0	(0,0)	1	0,492592593
1	(0,1)	0	(0,1)	0	0,492592593
1	(0,1)	0	(0,1)	1	0,642592593
1	(0,1)	0	(1,0)	0	0,525925926
1	(0,1)	0	(1,0)	1	0,625925926
1	(0,1)	0	(1,1)	0	0,625925926
1	(0,1)	0	(1,1)	1	0,525925926
1	(0,1)	1	(0,0)	0	0,767592593
1	(0,1)	1	(0,0)	1	0,617592593
1	(0,1)	1	(0,1)	0	0,517592593
1	(0,1)	1	(0,1)	1	0,667592593
1	(0,1)	1	(1,0)	0	0,450925926
1	(0,1)	1	(1,0)	1	0,550925926
1	(0,1)	1	(1,1)	0	0,850925926
1	(0,1)	1	(1,1)	1	0,750925926
1	(1,0)	0	(0,0)	0	0,775925926
1	(1,0)	0	(0,0)	1	0,625925926
1	(1,0)	0	(0,1)	0	0,625925926
1	(1,0)	0	(0,1)	1	0,775925926
1	(1,0)	0	(1,0)	0	0,659259259
1	(1,0)	0	(1,0)	1	0,759259259
1	(1,0)	0	(1,1)	0	0,759259259
1	(1,0)	0	(1,1)	1	0,659259259
1	(1,0)	1	(0,0)	0	0,850925926
1	(1,0)	1	(0,0)	1	0,700925926
1	(1,0)	1	(0,1)	0	0,600925926
1	(1,0)	1	(0,1)	1	0,750925926
1	(1,0)	1	(1,0)	0	0,534259259
1	(1,0)	1	(1,0)	1	0,634259259
1	(1,0)	1	(1,1)	0	0,934259259

1	(1,0)	1	(1,1)	1	0,834259259
1	(1,1)	0	(0,0)	0	0,562037037
1	(1,1)	0	(0,0)	1	0,412037037
1	(1,1)	0	(0,1)	0	0,412037037
1	(1,1)	0	(0,1)	1	0,562037037
1	(1,1)	0	(1,0)	0	0,44537037
1	(1,1)	0	(1,0)	1	0,54537037
1	(1,1)	0	(1,1)	0	0,54537037
1	(1,1)	0	(1,1)	1	0,44537037
1	(1,1)	1	(0,0)	0	0,437037037
1	(1,1)	1	(0,0)	1	0,287037037
1	(1,1)	1	(0,1)	0	0,187037037
1	(1,1)	1	(0,1)	1	0,337037037
1	(1,1)	1	(1,0)	0	0,12037037
1	(1,1)	1	(1,0)	1	0,22037037
1	(1,1)	1	(1,1)	0	0,52037037
1	(1,1)	1	(1,1)	1	0,42037037

z1	x1	z2	x2	z3	mean
0	(0,0)	0	(0,0)	0	6,333333333
0	(0,0)	0	(0,0)	1	10,33333333
0	(0,0)	0	(0,1)	0	5,666666667
0	(0,0)	0	(0,1)	1	8,666666667
0	(0,0)	0	(1,0)	0	9,666666667
0	(0,0)	0	(1,0)	1	6,666666667
0	(0,0)	0	(1,1)	0	9
0	(0,0)	0	(1,1)	1	5
0	(0,0)	1	(0,0)	0	10,35
0	(0,0)	1	(0,0)	1	14,35
0	(0,0)	1	(0,1)	0	9,583333333
0	(0,0)	1	(0,1)	1	12,58333333
0	(0,0)	1	(1,0)	0	13,48333333
0	(0,0)	1	(1,0)	1	10,48333333
0	(0,0)	1	(1,1)	0	13,11666667
0	(0,0)	1	(1,1)	1	9,116666667
0	(0,1)	0	(0,0)	0	5,666666667
0	(0,1)	0	(0,0)	1	9,666666667
0	(0,1)	0	(0,1)	0	5
0	(0,1)	0	(0,1)	1	8
0	(0,1)	0	(1,0)	0	9
0	(0,1)	0	(1,0)	1	6
0	(0,1)	0	(1,1)	0	8,333333333
0	(0,1)	0	(1,1)	1	4,333333333
0	(0,1)	1	(0,0)	0	8,683333333
0	(0,1)	1	(0,0)	1	12,68333333
0	(0,1)	1	(0,1)	0	7,916666667
0	(0,1)	1	(0,1)	1	10,91666667
0	(0,1)	1	(1,0)	0	11,81666667
0	(0,1)	1	(1,0)	1	8,816666667
0	(0,1)	1	(1,1)	0	11,45
0	(0,1)	1	(1,1)	1	7,45
0	(1,0)	0	(0,0)	0	9,666666667
0	(1,0)	0	(0,0)	1	13,66666667
0	(1,0)	0	(0,1)	0	9
0	(1,0)	0	(0,1)	1	12
0	(1,0)	0	(1,0)	0	13
0	(1,0)	0	(1,0)	1	10
0	(1,0)	0	(1,1)	0	12,33333333
0	(1,0)	0	(1,1)	1	8,333333333
0	(1,0)	1	(0,0)	0	6,683333333
0	(1,0)	1	(0,0)	1	10,68333333
0	(1,0)	1	(0,1)	0	5,916666667
0	(1,0)	1	(0,1)	1	8,916666667
0	(1,0)	1	(1,0)	0	9,816666667
0	(1,0)	1	(1,0)	1	6,816666667
0	(1,0)	1	(1,1)	0	9,45
0	(1,0)	1	(1,1)	1	5,45
0	(1,1)	0	(0,0)	0	9
0	(1,1)	0	(0,0)	1	13
0	(1,1)	0	(0,1)	0	8,333333333
0	(1,1)	0	(0,1)	1	11,33333333
0	(1,1)	0	(1,0)	0	12,33333333
0	(1,1)	0	(1,0)	1	9,333333333
0	(1,1)	0	(1,1)	0	11,66666667

0	(1,1)	0	(1,1)	1	7,666666667
0	(1,1)	1	(0,0)	0	5,016666667
0	(1,1)	1	(0,0)	1	9,016666667
0	(1,1)	1	(0,1)	0	4,25
0	(1,1)	1	(0,1)	1	7,25
0	(1,1)	1	(1,0)	0	8,15
0	(1,1)	1	(1,0)	1	5,15
0	(1,1)	1	(1,1)	0	7,783333333
0	(1,1)	1	(1,1)	1	3,783333333
1	(0,0)	0	(0,0)	0	8,316666667
1	(0,0)	0	(0,0)	1	12,31666667
1	(0,0)	0	(0,1)	0	7,65
1	(0,0)	0	(0,1)	1	10,65
1	(0,0)	0	(1,0)	0	11,65
1	(0,0)	0	(1,0)	1	8,65
1	(0,0)	0	(1,1)	0	10,98333333
1	(0,0)	0	(1,1)	1	6,983333333
1	(0,0)	1	(0,0)	0	12,33333333
1	(0,0)	1	(0,0)	1	16,33333333
1	(0,0)	1	(0,1)	0	11,56666667
1	(0,0)	1	(0,1)	1	14,56666667
1	(0,0)	1	(1,0)	0	15,46666667
1	(0,0)	1	(1,0)	1	12,46666667
1	(0,0)	1	(1,1)	0	15,1
1	(0,0)	1	(1,1)	1	11,1
1	(0,1)	0	(0,0)	0	7,75
1	(0,1)	0	(0,0)	1	11,75
1	(0,1)	0	(0,1)	0	7,083333333
1	(0,1)	0	(0,1)	1	10,08333333
1	(0,1)	0	(1,0)	0	11,08333333
1	(0,1)	0	(1,0)	1	8,083333333
1	(0,1)	0	(1,1)	0	10,41666667
1	(0,1)	0	(1,1)	1	6,416666667
1	(0,1)	1	(0,0)	0	10,76666667
1	(0,1)	1	(0,0)	1	14,76666667
1	(0,1)	1	(0,1)	0	10
1	(0,1)	1	(0,1)	1	13
1	(0,1)	1	(1,0)	0	13,9
1	(0,1)	1	(1,0)	1	10,9
1	(0,1)	1	(1,1)	0	13,53333333
1	(0,1)	1	(1,1)	1	9,533333333
1	(1,0)	0	(0,0)	0	11,85
1	(1,0)	0	(0,0)	1	15,85
1	(1,0)	0	(0,1)	0	11,18333333
1	(1,0)	0	(0,1)	1	14,18333333
1	(1,0)	0	(1,0)	0	15,18333333
1	(1,0)	0	(1,0)	1	12,18333333
1	(1,0)	0	(1,1)	0	14,51666667
1	(1,0)	0	(1,1)	1	10,51666667
1	(1,0)	1	(0,0)	0	8,866666667
1	(1,0)	1	(0,0)	1	12,86666667
1	(1,0)	1	(0,1)	0	8,1
1	(1,0)	1	(0,1)	1	11,1
1	(1,0)	1	(1,0)	0	12
1	(1,0)	1	(1,0)	1	9
1	(1,0)	1	(1,1)	0	11,63333333

1	(1,0)	1	(1,1)	1	7,633333333
1	(1,1)	0	(0,0)	0	10,88333333
1	(1,1)	0	(0,0)	1	14,88333333
1	(1,1)	0	(0,1)	0	10,21666667
1	(1,1)	0	(0,1)	1	13,21666667
1	(1,1)	0	(1,0)	0	14,21666667
1	(1,1)	0	(1,0)	1	11,21666667
1	(1,1)	0	(1,1)	0	13,55
1	(1,1)	0	(1,1)	1	9,55
1	(1,1)	1	(0,0)	0	6,9
1	(1,1)	1	(0,0)	1	10,9
1	(1,1)	1	(0,1)	0	6,133333333
1	(1,1)	1	(0,1)	1	9,133333333
1	(1,1)	1	(1,0)	0	10,03333333
1	(1,1)	1	(1,0)	1	7,033333333
1	(1,1)	1	(1,1)	0	9,666666667
1	(1,1)	1	(1,1)	1	5,666666667