Penalized Spline of Propensity Methods for Treatment Comparison

Supplementary Materials

Tingting Zhou*

Michael R. Elliott †‡
and
Roderick J.A. Little §

August 8, 2018

*Department of Biostatistics, School of Public Health, University of Michigan, 1415 Washington Heights, Ann Arbor, MI 48109 (tkzhou@umich.edu)

†Department of Biostatistics, School of Public Health, University of Michigan, 1415 Washington Heights, Ann Arbor, MI 48109

‡Survey Research Center, Institute for Social Research, 426 Thompson Street Ann Arbor, MI 48104-2321 (mrelliot@umich.edu)

§Department of Biostatistics, School of Public Health, University of Michigan, 1415 Washington Heights, Ann Arbor, MI 48109 (rlittle@umich.edu)

1 Double Robustness of PENCOMP

1.1 Single Time Point Treatment Assignment

Let X_1 denote the baseline covariates that affect treatment assignment Z_1. Suppose $Z_1 \in \{0, 1\}$ denotes assignment to control (0) or treatment (1). Let Y^{Z_1} denotes the potential outcome associated with treatment Z_1.

Result 1: The ignorable treatment assignment implies that $(Y^1, Y^0) \perp\!\!\!\perp Z_1 | P_{z_1}(X_1)$ (Rosenbaum and Rubin 1983), where $P_{z_1}(X_1) = Pr(Z_1 = z_1|X_1)$ denotes the propensity of being assigned z_1.

In the single time point treatment setting, suppose Y^0 is observed only for subjects $i = 1, \cdots, n_0$, while Y^1 is observed only for subjects $i = n_0 + 1, \cdots, n$. We are interested in estimating the causal effect $\Delta = E(Y^1 - Y^0)$. Under SUTVA, ignorability and positivity assumptions, we can estimate causal effects from the regression models on covariates X_1: $E(Y|X_1, Z_1 = 1)$ and $E(Y|X_1, Z_1 = 0)$, or from regression models on a summary measure of the covariates-propensity score $P_{z_1}(X_1)$: $E(Y|P_{z_1}(X_1), Z_1 = 1)$ and $E(Y|P_{z_1}(X_1), Z_1 = 0)$.

$$
\begin{aligned}
E(Y^1 - Y^0) &= E(E(Y^1 - Y^0|X_1)) \\
&= E(E(Y^1|X_1)) - E(E(Y^0|X_1)) \\
&= E\Big(E(Y|X_1, Z_1 = 1)\Big) - E\Big(E(Y|X_1, Z_1 = 0)\Big) \text{ by ignorability} \\
&= E\Big(E(Y|P_{z_1}(X_1), Z_1 = 1)\Big) - E\Big(E(Y|P_{z_1}(X_1), Z_1 = 0)\Big) \text{by Result 1}
\end{aligned}
$$

Alternatively, the mean $E(Y^1)$ can be written as $E(Y^1) = P(Z_1 = 1)E(Y^1|Z_1 = 1) + P(Z_1 = 0)E(Y^1|Z_1 = 0)$, estimated as:

$$\hat{E}(Y^1) = \frac{n_0}{n} * \frac{1}{n_0}\sum_{i=1}^{n_0}\hat{Y}_i^1 + \frac{n_1}{n} * \frac{1}{n_1}\sum_{i=(n_0+1)}^{n} Y_i^{obs}$$
$$= \frac{1}{n} * \left(\sum_{i=1}^{n_0}\hat{Y}_i^1 + \sum_{i=n_0+1}^{n} Y_i^{obs}\right)$$

where $E(Y^1|Z_1 = 1) = Y^{obs}$ and $E(Y^1|P_{z_1}(x_1), Z_1 = 0) = \hat{Y}^1$.

PENCOMP imputes the missing potential outcomes $Y^{z_1=1}$ for subjects $i = 1, \cdots, n_0$ from the mean model $E(Y^{z_1}|X_1, Z_1 = z_1, \theta_{z_1}, \beta_{z_1}) = s(\hat{P^*}_{z_1}; \theta_{z_1}) + g_{z_1}(\hat{P^*}_{z_1}, X_1; \beta_{z_1})$, where $\hat{P^*}_{z_1} = \log\,[\hat{P}_{z_1}(X_1)/(1 - \hat{P}_{z_1}(X_1))]$. Zhang and Little (2009) showed that this imputation model is equivalent to a centered version of the form $E(Y^{z_1}|X_1, Z_1 = z_1, \theta_{z_1}, \beta_{z_1}) = s(\hat{P^*}_{z_1}; \theta_{z_1}) + g_{z_1}(\hat{P^*}_{z_1}, X_1 - s_{x_1}(\hat{P^*}_{z_1}; \omega_{z_1}); \beta_{z_1})$, where $s_{x_1}(\hat{P^*}_{z_1}; \omega_{z_1}) = E(X_1|\hat{P^*}_{z_1})$ is the spline of X_1 on the logit of the propensity score, denoted as, $\hat{P^*}_{z_1}$ as shown in Little and An (2004). Specifically, in the centered version, the residuals from the spline regressions of covariates X_1 on $\hat{P^*}_{z_1}$ enter the parametric g function. Both Zhang and Little (2009) and Little and An (2004) showed that both imputation models in the missing data context yields a consistent estimate for $E(Y^1)$. Here we show the double robustness property of PENCOMP using the centered version for simplicity.

a) When the mean model of Y^1 given $(\hat{P^*}_{z_1}, X_1)$ are correctly specified, the marginal mean of Y^1 from the imputation model is consistent, as a consequence of the properties of a well-defined regression model.

b) When the prediction model given X_1 is misspecified, and the propensity and the spline models are correctly specified, the marginal mean of Y^1 is consistent. Here we prove the case for linear g function. In the case of a nonlinear g function, we can approximate it

using linear terms and the results will still hold.

$$
\begin{aligned}
E\left(\hat{Y}^1|P^*_{z_1}\right) &= s_y\left(P^*_{z_1}\right) + E\left[g\left(P^*_{z_1}, X_1 - s_{x_1}(P^*_{z_1})\right)|P^*_{z_1}\right] \\
&= s_y\left(P^*_{z_1}\right) + g\left(P^*_{z_1}, E\left(X_1 - s_{x_1}(P^*_{z_1})\middle|P^*_{z_1}\right)\right) \\
&\approx s_y\left(P^*_{z_1}\right) + g\left(P^*_{z_1}, 0,\right) \\
&= s_y\left(P^*_{z_1}\right) \\
&= E\left(Y^1|P^*_{z_1}\right) \\
&= E(Y^1|P^*_{z_1}, Z_1 = 1) \\
&= E(Y^1|P^*_{z_1}, Z_1 = 0)
\end{aligned}
$$

where the last two equalities again follow from Result 1.

Thus, for the subjects who actually received controls, the marginal mean of the imputed values $\hat{Y}^1$ from our imputation model is consistent even when the prediction model on covariates is misspecified: $\frac{1}{n_0}\sum_{i=1}^{n_0}\hat{Y}_i^1 \to E(Y^1|Z_1 = 0)$ as $n_0 \to \infty$. Similar approaches can be used to estimate $E(Y^0|Z_1 = 1)$ and thus estimated $E(Y^0)$.

1.2 Longitudinal Treatment Assignments

Suppose treatments are assigned at T discrete time points: $t = 1, \ldots, T$. Let $\bar{X}_t$ and $\bar{Z}_t$ denote the covariate and treatment history, respectively, up to and including time point t. Let $Y^{\bar{z}_T}$ denote the potential outcome under treatment regime $\bar{z}_T = (z_1, \cdots, z_T)$. The final outcome of interest $Y^{\bar{z}_T}$ is measured after time point T. Suppose, each z_t is binary treatment. For a particular treatment regime $\bar{z}_T = (z_1, z_2, \cdots, z_t, z_{t+1}, \cdots, z_T)$, under SUTVA, sequential ignorability and positivity assumptions, for all $t = 1, \cdots, T$, the following results hold.

Result 2: $Y^{\bar{z}_T} \perp\!\!\!\perp I(Z_t = z_t)|P_{z_t}(\bar{X}_t, \bar{z}_{t-1})$, where $I(.)$ is the indicator function, and $P_{z_t}(\bar{X}_t, \bar{z}_{t-1}) = P(Z_t = z_t|\bar{X}_t, \bar{Z}_{t-1})$, as a direct extension of the single time point treatment

(Rosenbaum and Rubin 1983).

Result 3: $Y^{\bar{z}_T} \perp\!\!\!\perp I(\bar{Z}_t = \bar{z}_t)|P_{\bar{z}_t}$, where $I(.)$ is the indicator function, $P_{\bar{z}_t} = \prod_{k=1}^{t} P(Z_k = z_k|\bar{Z}_{k-1} = \bar{z}_{k-1}, \bar{X}_k)$, which is the propensity of being assigned treatment regime $\bar{z}_t$, conditional on the past treatment and covariate history. In other words, the treatment regime $\bar{Z}_t$ up to and including time point t is independent of potential outcomes $Y^{\bar{z}_T}$ given the propensity of receiving that treatment regime $\bar{Z}_t$, for all $t = 1, \cdots, T$. The proof is outline here.

$$\begin{aligned} P\bigg(I(\bar{Z}_t = \bar{z}_t)|Y^{\bar{z}_T}, P_{\bar{z}_t}\bigg) &= P\bigg(I(\bar{Z}_t = \bar{z}_t)|P_{\bar{z}_t}\bigg) \\ &= P_{\bar{z}_t} \end{aligned}$$

$$\begin{aligned} P\bigg(I(\bar{Z}_t = \bar{z}_t)|Y^{\bar{z}_T}, P_{\bar{z}_t}\bigg) &= E\bigg(I(\bar{Z}_t = \bar{z}_t)|Y^{\bar{z}_T}, P_{\bar{z}_t}\bigg) \\ &= E\bigg[E\bigg(I(\bar{Z}_t = \bar{z}_t)|Y^{\bar{z}_T}, \bar{X}_t, \bar{Z}_{t-1}, P_{\bar{z}_t}\bigg)|Y^{\bar{z}_T}, P_{\bar{z}_t}\bigg] \\ &= E\bigg[I(\bar{Z}_{t-1} = \bar{z}_{t-1})E\bigg(I(Z_t = z_t)|\bar{X}_t, \bar{Z}_{t-1}, P_{\bar{z}_t}\bigg)|Y^{\bar{z}_T}, P_{\bar{z}_t}\bigg] \\ &\text{by sequential ignorability assumption} \\ &= E\bigg[I(\bar{Z}_{t-1} = \bar{z}_{t-1})P_{z_t}(\bar{X}_t, \bar{z}_{t-1})|Y^{\bar{z}_T}, P_{\bar{z}_t}\bigg] \\ &= E\bigg[E\bigg(I(\bar{Z}_{t-1} = \bar{z}_{t-1})P_{z_t}(\bar{X}_t, \bar{z}_{t-1})|Y^{\bar{z}_T}, \bar{X}_{t-1}, \bar{Z}_{t-2}, P_{\bar{z}_t}\bigg)\bigg|Y^{\bar{z}_T}, P_{\bar{z}_t}\bigg] \\ &= E\bigg[I(\bar{Z}_{t-2} = \bar{z}_{t-2})P_{z_t}(\bar{X}_t, \bar{z}_{t-1})E\bigg(I(Z_{t-1} = z_{t-1})|Y^{\bar{z}_T}, \bar{X}_{t-1}, \bar{Z}_{t-2}, P_{\bar{z}_t}\bigg) \\ &\bigg|Y^{\bar{z}_T}, P_{\bar{z}_t}\bigg] \\ &= E\bigg[I(\bar{Z}_{t-2} = \bar{z}_{t-2})P_{z_t}(\bar{X}_t, \bar{z}_{t-1})P_{z_{t-1}}(\bar{X}_{t-1}, \bar{z}_{t-2})\bigg|Y^{\bar{z}_T}, P_{\bar{z}_t}\bigg] \\ &= E\bigg[P_{\bar{z}_t}|Y^{\bar{z}_T}, P_{\bar{z}_t}\bigg]\text{by the same argument for each } Z_t \\ &= P_{\bar{z}_t} \end{aligned}$$

By the same argument but without the need for the sequential ignorability assumption, $P\Big(I(\bar{Z}_t = \bar{z}_t)|P_{\bar{z}_t}\Big) = P_{\bar{z}_t}$. Thus, $P\Big(I(\bar{Z}_t = \bar{z}_t)|Y^{\bar{z}_T}, P_{\bar{z}_t}\Big) = p\Big(I(\bar{Z}_t = \bar{z}_t)|P_{\bar{z}_t}\Big)$

Suppose we want to impute the missing potential outcomes X_3^{11} for subjects $1, \cdots, n_0$ and subjects $i = n_0 + 1, \cdots n$ receive treatment combination $(1, 1)$. As shown below, we can build a model for X_3^{11} from the subjects with observed treatment sequence of $(1, 1)$ to impute missing potential outcomes X_3^{11} for other subjects. Similar to single time point treatment, we can estimate causal effects from the regression models on the covariates or on the propensity scores.

$$\begin{aligned} E\Big(X_3^{11}\Big) &= E\Big[E(X_3^{11}|\bar{X}_2)\Big] \\ &= E\Big[E(X_3|\bar{X}_2, Z_1 = 1, Z_2 = 1)\Big] \text{by sequential ignorability} \\ &= E\Big[E(X_3|P_{\bar{z}_2=(11)}, Z_1 = 1, Z_2 = 1)\Big] \text{by result 3} \end{aligned}$$

Alternatively, the mean $E(X_3^{11})$ can be written as $E(X_3^{11}) = P(\bar{Z}_2 = (1,1))E(X_3^{11}|\bar{Z}_2 = (1,1)) + P(\bar{Z}_2 \neq (1,1))E(X_3^{11}|\bar{Z}_2 \neq (1,1))$, estimated as:

$$\begin{aligned} \hat{E}(X_3^{11}) &= \frac{n_{00}}{n} * \frac{1}{n_{00}} \sum_{i=1}^{n_{00}} \hat{X_3}^{11} + \frac{n_{01}}{n} * \frac{1}{n_{01}} \sum_{i=n_{00}+1}^{n_{00}+n_{01}} \hat{X_3}^{11} \\ &\quad + \frac{n_{10}}{n} * \frac{1}{n_{10}} \sum_{i=n_{00}+n_{01}+1}^{n_{00}+n_{01}+n_{10}} \hat{X_3}^{11} + \sum_{i=n_{00}+n_{01}+n_{10}+1}^{n} X_3^{obs} \\ &= \frac{1}{n} * \left(\sum_{i=1}^{n_{00}+n_{01}+n_{10}} \hat{X_3}^{11} + \sum_{i=n_{00}+n_{01}+n_{10}+1}^{n} X_3^{obs} \right) \end{aligned}$$

where $\hat{X}_3^{11} = \hat{E}(X_3^{11}|P_{\bar{z}_2}, Z_1 \neq 1, Z_2 \neq 1)$.

PENCOMP imputes the first missing intermediate outcomes X_2 first, X_3, and continue forward to the final outcome Y. By induction, we can show PENCOMP has double robustness property in longitudinal study. We have shown double robustness property for the base case $t = 1$ as in the single treatment. Suppose PENCOMP has the double robustness property in imputing missing potential outcomes X_t. We want to show that the double

robustness property also holds for the missing potential outcomes X_{t+1}, Suppose we are interested in estimating $X_{t+1}^{\bar{z}_t}$, where $\bar{z}_t = (z_1, \cdots, z_t)$ and subjects $i = 1, \cdots, n_0$ do not treatment sequence $\bar{Z}_t$ that match $\bar{z}_t$. Thus, to impute the missing potential outcomes $X_{t+1}^{\bar{z}_t}$ for the subjects whose treatment sequence did not match $\bar{z}_t$, we draw values from the mean model $E(X_{t+1}^{\bar{z}_t}|\bar{X}_t, \bar{Z}_t = \bar{z}_t, \theta_{\bar{z}_t}, \beta_{\bar{z}_t}, \gamma_{\bar{z}_t}) = s_{x_{t+1}}(\hat{P}^*_{\bar{z}_t}; \theta_{\bar{z}_t}) + g\left[\hat{P}^*_{\bar{z}_t}, X_1, \cdots, X_t; \beta_{\bar{z}_t}\right]$, which is equivalent to the mean model $E(X_{t+1}^{\bar{z}_t}|\bar{X}_t, \bar{Z}_t = \bar{z}_t, \theta_{\bar{z}_t}, \beta_{\bar{z}_t}, \gamma_{\bar{z}_t}) = s_{x_{t+1}}(\hat{P}^*_{\bar{z}_t}; \theta_{\bar{z}_t}) + g\left[\hat{P}^*_{\bar{z}_t}, X_1 - s_{x_1}(\hat{P}^*_{\bar{z}_t}; \omega_{\bar{z}_t^1}), \cdots, X_t - s_{x_t}(\hat{P}^*_{\bar{z}_t}; \omega_{\bar{z}_t^t}); \beta_{\bar{z}_t}\right]$, where $\hat{P}^*_{\bar{z}_t} = \log\left(\hat{P}_{\bar{z}_t}/(1-\hat{P}_{\bar{z}_t})\right)$. Here we need to show the double robustness property of PENCOMP with the centered version.

a) When the mean model of $X_{t+1}^{\bar{z}_t}$ given the covariate history $\bar{X}_t$ are correctly specified, the marginal mean of $X_{t+1}^{\bar{z}_t}$ from the imputation model is consistent, as a consequence of well-defined regression models.

b) When the prediction model given $\bar{X}_t$ is misspecified, and all the propensity models up to and including time point t and the spline models are correctly specified, the marginal mean of $X_{t+1}^{\bar{z}_t}$ is consistent. Again we prove the case for linear g function. We can approximate a nonlinear g function with using linear terms and the results will still hold.

$$
\begin{aligned}
E\left(\hat{X}_{t+1}^{\bar{z}_t}|P^*_{\bar{z}_t}\right) &= s_{x_{t+1}}\left(P^*_{\bar{z}_t}\right) + E\left[g\left(P^*_{\bar{z}_t}, X_1 - s_{x_1}(P^*_{\bar{z}_t}), \cdots, X_t - s_{x_t}(P^*_{\bar{z}_t}\right)|P^*_{\bar{z}_t}\right] \\
&= s_{x_{t+1}}\left(P^*_{\bar{z}_t}\right) + g\left[P^*_{\bar{z}_t}, E\left(X_1 - s_{x_1}(P^*_{\bar{z}_t})|P^*_{\bar{z}_t}\right), \cdots, E\left(X_t - s_{x_t}(P^*_{\bar{z}_t})|P^*_{\bar{z}_t}\right)\right] \\
&\approx s_{x_{t+1}}\left(P^*_{\bar{z}_t}\right) + g\left[P^*_{\bar{z}_t}, 0, \cdots, 0\right] \\
&= s_{x_{t+1}}\left(P^*_{\bar{z}_t}\right) \\
&= E\left(X_{t+1}^{\bar{z}_t}|P^*_{\bar{z}_t}\right) \\
&= E(X_{t+1}^{\bar{z}_t}|P^*_{\bar{z}_t}, \bar{Z}_t \neq \bar{z}_t) \\
&= E(X_{t+1}^{\bar{z}_t}|P^*_{\bar{z}_t}, \bar{Z}_t = \bar{z}_t) \text{by result 4}
\end{aligned}
$$

where the last two equalities follow from Result 3.

Thus, $\frac{1}{n_{0k}}\sum_{i=1}^{n_{0k}} \hat{X}_{k+1,i}^{\bar{z}_k} \to E(X_{k+1}^{\bar{z}_k}|\bar{Z}_k \neq \bar{z}_k)$ as $n_{0k} \to \infty$, where n_{0k} is the sample size

of the observations for which $\bar{Z}_k \neq \bar{z}_k$, and we assume that the observations are ordered that the first n_0 corresponds to the observations for which $\bar{Z}_k \neq \bar{z}_k$. Thus, by induction, PENCOMP has double robustness property in longitudinal study.

2 Implementations of the IPTW and the AIPTW Estimators

2.1 IPTW

Let $O_i = (\bar{X}_{iT}, \bar{Z}_{iT}, Y_i)$ denote the observed data for subject i, where $i = 1, \cdots, n$. The likelihood of the observed data can be factored into two components $P(O) = Q_0 g_0$, where $Q_0 = P(Y|\bar{X}_T, \bar{Z}_T = \bar{z}_T) \prod_{t=1}^{T} P(X_t|\bar{X}_{t-1}, \bar{Z}_{t-1})$ and $g_0 = \prod_{t=1}^{T} P(Z_t|\bar{Z}_{t-1}, \bar{X}_{t-1})$. Denote the MLE of Q_0 and g_0 as Q_n and g_n respectively.

From the IPTW estimating equation $\sum_{i=1}^{n} D_{IPTW}(O_i|\beta, g_n) = 0$, we can obtain $\hat{E}(Y^{z_1}) = \sum_{i=1}^{n} \frac{I(Z_{1i}=z_{1i})}{\hat{P}(Z_{1i}=z_{1i}|X_{1i})})^{-1} \sum_{i=1}^{n} \frac{Z_{1i}Y_i}{\hat{P}(Z_{1i}|X_{1i})}$. Thus, the estimated causal effect $\hat{\Delta}$ in a single time point is

$$\hat{\Delta}^{IPTW} = (\sum_{i=1}^{n} \frac{Z_{1i}}{\hat{P}(Z_{1i}|X_{1i})})^{-1} \sum_{i=1}^{n} \frac{Z_{1i}Y_i}{\hat{P}(Z_{1i}|X_{1i})} - (\sum_{i=1}^{n} \frac{1-Z_{1i}}{1-\hat{P}(Z_{1i}|X_{1i})})^{-1} \sum_{i=1}^{n} \frac{(1-Z_{1i})Y_i}{(1-\hat{P}(Z_{1i}|X_{1i}))}$$

Similarly, in a two time points treatment, the estimated causal effects $\hat{\Delta}_{z_1 z_2}$ are

$$\begin{aligned}\hat{\Delta}_{z_1 z_2}^{IPTW} = &(\sum_{i=1}^{n} \frac{I(Z_{1i}=z_1, Z_{2i}=z_2)}{P(Z_{1i}|X_{1i})P(Z_{2i}|X_{1i}, X_{2i}, Z_{1i})})^{-1} \sum_{i=1}^{n} \frac{I(Z_{1i}=z_1, Z_{2i}=z_2)Y_i}{P(Z_{1i}|X_{1i})P(Z_{2i}|X_{1i}, X_{2i}, Z_{1i})} \\ &- (\sum_{i=1}^{n} \frac{I(Z_{1i}=0, Z_{2i}=0)}{P(Z_{1i}|X_{1i})P(Z_{2i}|X_{1i}, X_{2i}, Z_{1i})})^{-1} \sum_{i=1}^{n} \frac{I(Z_{1i}=0, Z_{2i}=0)Y_i}{P(Z_{1i}|X_{1i})P(Z_{2i}|X_{1i}, X_{2i}, Z_{1i})}\end{aligned}$$

2.2 AIPTW

To solve the estimating equation $\sum_{i=1}^{n} D_{AIPTW}(O_i|\beta, g_n, Q_n) = 0$ in the single treatment assignment setting, we proceeds as follows.

(a) For $d = 1, \cdots, D$, generate a bootstrap sample $S^{(d)}$ from the original data S by sampling units with replacement, stratified on treatment group. Then carry out steps (b)-(h) for each sample d:

(b) Estimate a logistic regression model for the distribution of Z_1 given X_1, with regression parameters γ_{z_1} . Estimate the propensity to be assigned treatment $Z_1 = z_1$ as $\hat{P}(Z_1 = z_1|X_1, \hat{\gamma_{z_1}}^{(d)})$, where $\hat{\gamma_{z_1}}^{(d)}$ is the ML estimate of γ_{z_1}.

(c) For $z_1 = 0, 1$, using the cases assigned to treatment group z_1, estimate the distribution Y given X_1 and Z_1, $\hat{P}(Y|X_1 = x_1, Z_1 = z_1)$, using a normal linear regression with mean $E(Y|X_1, Z_1 = z_1, \beta_{z_1}) = g_{z_1}(X_1; \beta_{z_1})$, where $g_{z_1}()$ represents a parametric function of X_1 and Z_1 indexed by parameters β_{z_1}.

(d) Estimate the distributions of baseline covariates $P(X_1)$ using the empirical distributions from the data, denoted as $\hat{P}(X_1)$.

e) Estimate $\hat{\beta}_n^{mc} = (\hat{\beta}_0^{mc}, \hat{\beta}_1^{mc})$ using the g-computation to generate $10,000$ number of Y_0 and Y_1 from their respective counterfactual reference distributions. Specifically, draw x_1^* from the empirical distribution of X_1, $\hat{P}(X_1)$. Set $Z_1 = z_1$ and generate draws y^* from $\hat{P}(Y|X_1 = x_1^*, Z_1 = z_1)$. Then fit the MSM model $E(Y^{Z_1}) = \beta_0 + \beta_1 Z_1$ to this collection of $(y^*, 1)$ and $(y^*, 0)$ to obtain $\hat{\beta}_n^{mc}$.

f) Using Q_n, g_n and $\hat{\beta}_n^{mc}$, estimate $E_{Q_n,g_n}[D_{IPTW}(O_i|\hat{\beta}_n^{mc}, g_n)|Z_{1i} = z_{1i}, X_{1i} = x_{1i}]$ for each subject i as follows. Given $(Z_{1i} = z_{1i}, X_{1i} = x_{1i})$, generate 2,000 draws of Y_i^{mc} from $\hat{P}(Y|X_{1i} = x_{1i}, Z_{1i} = z_{1i})$ and compute

$$D_i^{mc} = \frac{\hat{h}(Z_{1i})}{\hat{P}(Z_{1i}|X_{1i})}(Y_i^{mc} - (\hat{\beta}_0^{mc} + \hat{\beta}_1^{mc} Z_{1i}))$$

where $\hat{h}(Z_{1i}) = \frac{dE(Y^{Z_{1i}})}{d\beta}\hat{P}(Z_{1i})$. Take the mean of 2000 Monte Carlo values as the estimate.

g) Similarly estimate $E_{Q_n,g_n}[D_{IPTW}(O_i|\hat{\beta}_n^{mc}, g_n)|X_{1i} = x_{1i}]$. Given $X_{1i} = x_{1i}$, first generate draws of z_{1i}^{mc} from $\hat{P}(Z_{1i}|X_{1i} = x_{1i})$, then generate draws of Y_i^{mc} from $\hat{P}(Y|X_{1i} = x_{1i}, Z_{1i} = z_{1i}^{mc})$ and compute D_i^{mc}. Take the mean of 2000 Monte Carlo values D_i^{mc} as the estimate.

h) Let $\hat{\pi}_i = \hat{E}_{Q_n,g_n}[D_{IPTW}(O|\hat{\beta}_n^{mc}, g_n)|Z_{1i}, X_{1i}] - E_{Q_n,g_n}[D_{IPTW}(O|\hat{\beta}_n^{mc}, g_n)|X_{1i}]$. Solve

(β_0, β_1) using Newton Raphson algorithm

$$\sum_{i=1}^{n} D_{AIPTW}(O_i|\beta, g_n, Q_n) = \sum_{i=1}^{n} D_{IPTW}(O_i|\beta, g_n) - \hat{\pi}_i = 0$$

The treatment effect is $\hat{\Delta}^{AIPTW} = \sum_{d=1}^{D} \hat{\beta}_1^{(d)}/D$. Estimate the variance by bootstrap and obtain the 95% confidence interval from the bootstrap samples.

Similarly to solve the AIPTW estimating equation in a two time points treatment, the steps proceeds as follows. Let $\beta = (\beta_0, \beta_1, \beta_2, \beta_3)$.

(a) For $d = 1, \cdots, D$, generate a bootstrap sample $S^{(d)}$ from the original data S by sampling units with replacement, stratified on treatment group. Then carry out steps (b)-(i) for each sample d:

(b) Estimate a logistic regression model for the distribution of Z_1 given X_1, with regression parameters γ_{z_1}. Estimate the propensity to be assigned treatment $Z_1 = z_1$ as $\hat{P}(Z_1 = z_1|X_1, \hat{\gamma}_{z_1}^{(d)})$, where $\hat{\gamma}_{z_1}^{(d)}$ is the ML estimate of γ_{z_1}.

(c) Estimate the distributions of baseline covariates $P(X_1)$ as the empirical distributions from the data, denoted as $\hat{P}(X_1)$.

(d) Using the cases assigned to treatment group $Z_1 = z_1$, estimate $\hat{P}(X_2|X_1, Z_1)$ using a normal linear regression with mean

$$E(X_2^{z_1}|X_1, Z_1 = z_1, \theta_{z_1}, \beta_{z_1}) = g_{z_1}(X_1, Z_1, \beta_{z_1}) \tag{1}$$

where $g_{z_1}()$ represents a parametric function of X_1, and Z_1 indexed by parameters β_{z_1}.

(e) Estimate a logistic regression model for the distribution of Z_2 given $\bar{X}_2, Z_1$, with regression parameters γ_{z_2}. Estimate the propensity to be assigned treatment $Z_2 = z_2$ given $Z_1, \bar{X}_2$ as $\hat{P}(Z_2 = z_2|\bar{X}_2, Z_1, \hat{\gamma}_{z_2}^{(d)})$, where $\hat{\gamma}_{z_2}^{(d)}$ is the ML estimate of γ_{z_2}.

(f) Using the cases assigned to treatment regime $\bar{Z}_2 = \bar{z}_2$, estimate $\hat{P}(Y|\bar{X}_2, \bar{Z}_2)$ using

a normal linear regression with mean

$$E(Y^{\bar{z}_2}|\bar{X}_2, \bar{Z}_2 = \bar{z}_2, \beta_{\bar{z}_2}) = g_{z_1 z_2}(\bar{X}_2, \bar{Z}_2; \beta_{\bar{z}_2})$$

where $g_{\bar{z}_2}()$ represents a parametric function indexed by parameters $\beta_{\bar{z}_2}$.

g) Estimate $\hat{\beta}_n^{mc} = (\hat{\beta}_0^{mc}, \hat{\beta}_1^{mc}, \hat{\beta}_2^{mc}, \hat{\beta}_3^{mc})$ using the g-computation to generate $10,000$ draws of the potential outcomes Y^{00},Y^{01},Y^{11},Y^{10} from their respective counterfactual distributions. Specifically, first generate a draw x_1^* from the empirical distribution $\hat{P}(X_1)$. Set $Z_1 = z_1$ and generate a draw x_2^* from $\hat{P}(X_2|X_1 = x_1^*, Z_1 = z_1)$. Then set $Z_2 = z_2$ and generate draws y^* from $\hat{P}(Y|X_1 = x_1^*, Z_1 = z_1, X_2 = x_2^*, Z_2 = z_2)$. Then fit the model $E(Y^{\bar{Z}_2}) = \beta_0 + \beta_1 Z_1 + \beta_2 Z_2 + \beta_3 Z_1 Z_2$ to this collection of $(y^*, 0, 0)$, $(y^*, 1, 0)$, $(y^*, 0, 1)$ and $(y^*, 1, 1)$ to obtain $\hat{\beta}_n^{mc}$.

h) Using Q_n, g_n and $\hat{\beta}_n^{mc}$, estimate $E_{Q_n,g_n}[D_{IPTW}(O_i|\hat{\beta}_n^{mc}, g_n)|\bar{Z}_{2i} = \bar{z}_{2i}, \bar{X}_{2i} = \bar{x}_{2i}]$ for each subject i as follows. Given $(\bar{Z}_{2i} = \bar{z}_{2i}, \bar{X}_{2i} = \bar{x}_{2i})$, generate 2,000 draws of Y_i^{mc} from $P(Y|\bar{Z}_{2i} = \bar{z}_{2i}, \bar{X}_{2i} = \bar{x}_{2i})$ and compute D_i^{mc}. Take the mean of the 2,000 Monte Carlo values as the estimate.

$$D_i^{mc} = \frac{\hat{h}(\bar{Z}_{2i})}{\hat{P}(Z_{1i}|X_{1i})\hat{P}(Z_{2i}|Z_{1i}, \bar{X}_{2i})}(Y_i^{mc} - (\hat{\beta}_0^{mc} + \hat{\beta}_1^{mc} Z_{1i} + \hat{\beta}_2^{mc} Z_{2i} + \hat{\beta}_3^{mc} Z_{1i} Z_{2i}))$$

where $\hat{h}(\bar{Z}_{2i}) = \frac{dE(Y^{\bar{Z}_{2i}})}{d\beta}\hat{P}(Z_{1i})\hat{P}(Z_{2i}|Z_{1i})$. Follow the similar procedures to estimate the other three conditional expectations.

i) Solve the estimating equation using Newton Raphson algorithm

$$\sum_{i=1}^{n} D_{AIPTW}(O_i|\beta, g_n, Q_n) = \sum_{i=1}^{n} D_{IPTW}(O_i|\beta, g_n) - \hat{\pi}_i = 0 \tag{2}$$

where,

$$\hat{\pi}^i = \sum_{j=1}^{j=2} E_{Q_n,g_n}[D_{IPTW}(O|\hat{\beta}^{mc}, g_n)|\bar{Z}_j, \bar{X}_j] - E_{Q_n,g_n}[D_{IPTW}(O|\hat{\beta}^{mc}, g_n)|\bar{X}_j]$$

The treatment effects are $\hat{\Delta}_{11}^{AIPTW} = \sum_{d=1}^{D} \hat{\Delta}_{11}^{AIPTW(d)}/D$, $\hat{\Delta}_{10}^{AIPTW} = \sum_{d=1}^{D} \hat{\Delta}_{10}^{AIPTW(d)}/D$,

and $\hat{\Delta}_{01}^{AIPTW} = \sum_{d=1}^{D} \hat{\Delta}_{01}^{AIPTW(d)}/D$, where $\hat{\Delta}_{11}^{AIPTW(d)} = \hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3$; $\hat{\Delta}_{10}^{AIPTW(d)} = \hat{\beta}_1$; $\hat{\Delta}_{01}^{AIPTW(d)} = \hat{\beta}_2$. Estimate the variance and obtain the 95% confidence interval from D bootstrap samples.

3 Supplemental Tables from the Simulation Study

Table 5: 100 * Ratio of bias over RMSE of IPTW (A), under (A) correctly-specified propensity and prediction models; (B) a correctly-specified propensity model only; (C) a correctly-specified prediction model only, based on 1000 simulations with sample size of 200. The treatment effects Δs under linear and nonlinear outcome models were 5 and 9, respectively.

$\Delta = E(Y^1) - E(Y^0)$						
100 * Empirical Bias / RMSE IPTW(A)						
	Linear Outcome			NonLinear Outcome		
Method	Low	Mod	High	Low	Mod	High
IPTW(A)	1	14	25	4	3	4
g-computation(A)	-0	-0	-1	6	4	3
AIPTW(A)	0	0	2	2	2	2
PENCOMP(A)	0	2	2	2	3	3
IPTW(A)	1	14	25	4	3	4
g-computation(B)	79	357	303	41	225	225
AIPTW(B)	0	18	29	-1	3	5
PENCOMP(B)	11	2	5	-1	-37	-52
IPTW(C)	82	375	340	46	250	273
g-computation(A)	-0	-0	-1	6	4	3
AIPTW(C)	0	0	0	2	2	1
PENCOMP(C)	-0	0	0	2	2	1

Table 6: 100*Ratio of empirical RMSE over RMSE of IPTW (A), denoted as RMSE/RMSE IPTW(A), under (A) correctly-specified propensity and prediction models; (B) a correctly-specified propensity model only; (C) a correctly- specified prediction model only, based on 1000 simulations with sample size of 200.

$\Delta = E(Y^1) - E(Y^0)$						
100 * RMSE / RMSE IPTW(A)						
	Linear Outcome			NonLinear Outcome		
Method	Low	Mod	High	Low	Mod	High
IPTW(A)	100	100	100	100	100	100
g-computation(A)	80	51	37	75	61	51
AIPTW(A)	78	59	47	73	63	55
PENCOMP(A)	78	57	46	73	62	54
IPTW(A)	100	100	100	100	100	100
g-computation(B)	168	367	307	124	246	240
AIPTW(B)	81	89	90	95	99	97
PENCOMP(B)	83	65	56	91	97	102
IPTW(C)	181	389	347	130	273	290
g-computation(A)	80	51	37	75	61	51
AIPTW(C)	78	53	39	73	60	51
PENCOMP(C)	78	54	40	73	61	51

.

Table 7: Empirical 95% non-coverage rate*100 (nominal noncoverage of 5), under (A) correctly-specified propensity and prediction models; (B) a correctly-specified propensity model only; (C) a correctly-specified prediction model only, based on 1000 simulations with sample size of 200.

$\Delta = E(Y^1) - E(Y^0)$						
100 * 95% Non-coverage Rate						
	Linear Outcome			NonLinear Outcome		
Method	Low	Mod	High	Low	Mod	High
IPTW(A)	5	6	12	5	6	7
g-computation(A)	4	5	6	5	6	5
AIPTW(A)	4	6	6	5	6	5
PENCOMP(A)	4	3	3	4	5	3
IPTW(A)	5	6	12	5	6	7
g-computation(B)	10	99	100	6	64	81
AIPTW(B)	3	8	13	5	6	7
PENCOMP(B)	0	0	1	2	5	6
IPTW(C)	10	96	99	6	63	82
g-computation(A)	4	5	6	5	6	5
AIPTW(C)	4	5	7	5	6	6
PENCOMP(C)	4	4	5	5	5	5

Table 8: 100 * Ratio of empirical mean 95% confidence interval width to that of IPTW (A), denoted as mean 95% interval width/mean 95% interval width IPTW(A), under (A) correctly-specified propensity and prediction models; (B) a correctly-specified propensity model only; (C) a correctly-specified prediction model only, based on 1000 simulations with sample size of 200.

$\Delta = E(Y^1) - E(Y^0)$						
100 * mean 95% interval width/mean 95% interval width IPTW(A)						
	Linear Outcome			NonLinear Outcome		
Method	Low	Mod	High	Low	Mod	High
IPTW(A)	100	100	100	100	100	100
g-computation(A)	80	53	42	74	64	57
AIPTW(A)	79	60	60	73	66	67
PENCOMP(A)	80	69	68	74	70	72
IPTW(A)	100	100	100	100	100	100
g-computation(B)	136	84	64	114	105	95
AIPTW(B)	84	89	93	96	97	100
PENCOMP(B)	124	102	102	113	120	130
IPTW(C)	147	103	82	118	117	113
g-computation(A)	80	53	42	74	64	57
AIPTW(C)	79	55	44	73	64	58
PENCOMP(C)	79	58	50	73	65	61

Table 9: 100 * Ratio of bias over RMSE of IPTW (A), under (A) correctly-specified propensity and prediction models; (B) a correctly-specified propensity model only; (C) a correctly-specified prediction model only, based on 1000 simulations with sample size of 500. The treatment effects Δs under linear and nonlinear outcome models were 5 and 9, respectively.

$\Delta = E(Y^1) - E(Y^0)$						
100 * Empirical Bias / RMSE IPTW (A)						
	Linear Outcome			NonLinear Outcome		
Method	Low	Mod	High	Low	Mod	High
IPTW(A)	1	4	16	-5	-8	-1
g-computation(A)	0	-0	-1	-0	-0	-0
AIPTW(A)	-0	-2	0	-2	-3	-1
PENCOMP(A)	-0	-0	0	-3	-2	-1
IPTW(A)	1	4	16	-5	-8	-1
g-computation(B)	123	482	406	58	333	327
AIPTW(B)	1	6	16	-8	-9	-4
PENCOMP(B)	15	-1	2	-2	-46	-67
IPTW(C)	125	510	458	62	367	396
g-computation(A)	0	-0	-1	-0	-0	-0
AIPTW(C)	-0	0	-0	-2	-2	-1
PENCOMP(C)	-1	-0	-0	-3	-2	-1

Table 10: 100*Ratio of empirical RMSE over RMSE of IPTW (A), denoted as RMSE/RMSE IPTW(A), under (A) correctly-specified propensity and prediction models; (B) a correctly-specified propensity model only; (C) a correctly- specified prediction model only, based on 1000 simulations with sample size of 500.

$\Delta = E(Y^1) - E(Y^0)$						
100 * RMSE / RMSE IPTW(A)						
	Linear Outcome			NonLinear Outcome		
Method	Low	Mod	High	Low	Mod	High
IPTW(A)	100	100	100	100	100	100
g-computation(A)	82	45	31	72	54	44
AIPTW(A)	79	54	42	70	56	49
PENCOMP(A)	79	49	38	70	55	47
IPTW(A)	100	100	100	100	100	100
g-computation(B)	184	487	408	128	345	336
AIPTW(B)	79	85	92	94	94	98
PENCOMP(B)	82	59	51	90	93	101
IPTW(C)	193	517	462	134	382	406
g-computation(A)	82	45	31	72	54	44
AIPTW(C)	79	46	32	70	53	44
PENCOMP(C)	79	46	33	70	53	44

.

Table 11: Empirical 95% non-coverage rate*100 (nominal noncoverage of 5), under (A) correctly-specified propensity and prediction models; (B) a correctly-specified propensity model only; (C) a correctly-specified prediction model only, based on 1000 simulations with sample size of 500.

$\Delta = E(Y^1) - E(Y^0)$						
100 * 95% Non-coverage Rate						
	Linear Outcome			NonLinear Outcome		
Method	Low	Mod	High	Low	Mod	High
IPTW(A)	5	6	11	5	6	6
g-computation(A)	5	6	6	4	4	4
AIPTW(A)	4	6	7	3	4	4
PENCOMP(A)	4	4	3	3	3	2
IPTW(A)	5	6	11	5	6	6
g-computation(B)	15	100	100	7	96	100
AIPTW(B)	4	7	13	5	6	6
PENCOMP(B)	0	1	1	3	6	10
IPTW(C)	12	100	100	7	97	100
g-computation(A)	5	6	6	4	4	4
AIPTW(C)	4	6	5	3	4	4
PENCOMP(C)	4	4	4	3	3	3

Table 12: 100 * Ratio of empirical mean 95% confidence interval width to that of IPTW (A), denoted as mean 95% interval width/mean 95% interval width IPTW(A), under (A) correctly-specified propensity and prediction models; (B) a correctly-specified propensity model only; (C) a correctly-specified prediction model only, based on 1000 simulations with sample size of 500.

$\Delta = E(Y^1) - E(Y^0)$						
100 * mean 95% interval width/mean 95% interval width IPTW(A)						
	Linear Outcome			NonLinear Outcome		
Method	Low	Mod	High	Low	Mod	High
IPTW(A)	100	100	100	100	100	100
g-computation(A)	84	52	38	77	61	53
AIPTW(A)	81	58	51	74	63	60
PENCOMP(A)	81	61	54	75	64	61
IPTW(A)	100	100	100	100	100	100
g-computation(B)	141	82	57	116	101	88
AIPTW(B)	82	88	92	95	95	96
PENCOMP(B)	120	92	85	107	108	113
IPTW(C)	151	98	73	119	113	105
g-computation(A)	84	52	38	77	61	53
AIPTW(C)	81	52	39	74	60	53
PENCOMP(C)	81	55	43	74	61	55

Table 13: 100 * Ratio of bias over RMSE of IPTW (A), under (A) correctly-specified propensity and prediction models; (B) a correctly-specified propensity model only; (C) a correctly-specified prediction model only, based on 1000 simulations with sample size of 1000. The treatment effects Δs under linear and nonlinear outcome models were 5 and 9, respectively.

$\Delta = E(Y^1) - E(Y^0)$						
100 * Empirical Bias / RMSE IPTW (A)						
	Linear Outcome			NonLinear Outcome		
Method	Low	Mod	High	Low	Mod	High
IPTW(A)	-1	3	11	-2	-3	-1
g-computation(A)	2	2	1	-2	-1	-0
AIPTW(A)	2	1	0	-3	-2	-2
PENCOMP(A)	3	2	1	-2	-1	-1
IPTW(A)	-1	3	11	-2	-3	-1
g-computation(B)	182	674	517	92	459	420
AIPTW(B)	2	7	14	-2	-1	-0
PENCOMP(B)	21	1	3	6	-36	-61
IPTW(C)	181	706	578	95	502	505
g-computation(A)	2	2	1	-2	-1	-0
AIPTW(C)	1	1	0	-3	-2	-1
PENCOMP(C)	3	2	1	-2	-1	-1

Table 14: 100*Ratio of empirical RMSE over RMSE of IPTW (A), denoted as RMSE/RMSE IPTW(A), under (A) correctly-specified propensity and prediction models; (B) a correctly-specified propensity model only; (C) a correctly- specified prediction model only, based on 1000 simulations with sample size of 1000.

$\Delta = E(Y^1) - E(Y^0)$						
100 * RMSE / RMSE IPTW(A)						
	Linear Outcome			NonLinear Outcome		
Method	Low	Mod	High	Low	Mod	High
IPTW(A)	100	100	100	100	100	100
g-computation(A)	85	45	29	79	58	44
AIPTW(A)	80	53	45	74	59	52
PENCOMP(A)	80	49	36	74	56	45
IPTW(A)	100	100	100	100	100	100
g-computation(B)	233	678	518	151	468	427
AIPTW(B)	81	90	94	96	94	94
PENCOMP(B)	85	59	50	92	88	96
IPTW(C)	238	711	581	153	513	512
g-computation(A)	85	45	29	79	58	44
AIPTW(C)	80	45	30	74	54	42
PENCOMP(C)	80	45	30	74	54	42

.

Table 15: Empirical 95% non-coverage rate*100 (nominal noncoverage of 5), under (A) correctly-specified propensity and prediction models; (B) a correctly-specified propensity model only; (C) a correctly-specified prediction model only, based on 1000 simulations with sample size of 1000.

$\Delta = E(Y^1) - E(Y^0)$						
100 * 95% Non-coverage Rate						
	Linear Outcome			NonLinear Outcome		
Method	Low	Mod	High	Low	Mod	High
IPTW(A)	5	6	12	4	6	6
g-computation(A)	6	5	5	6	5	5
AIPTW(A)	5	5	6	5	6	6
PENCOMP(A)	5	5	4	5	5	5
IPTW(A)	5	6	12	4	6	6
g-computation(B)	26	100	100	12	100	100
AIPTW(B)	4	7	13	6	6	6
PENCOMP(B)	1	1	2	3	4	8
IPTW(C)	24	100	100	11	100	100
g-computation(A)	6	5	5	6	5	5
AIPTW(C)	5	5	6	5	5	5
PENCOMP(C)	5	4	4	5	5	5

Table 16: 100 * Ratio of empirical mean 95% confidence interval width to that of IPTW (A), denoted as mean 95% interval width/mean 95% interval width IPTW(A), under (A) correctly-specified propensity and prediction models; (B) a correctly-specified propensity model only; (C) a correctly-specified prediction model only, based on 1000 simulations with sample size of 1000.

$\Delta = E(Y^1) - E(Y^0)$						
100 * mean 95% interval width/mean 95% interval width IPTW(A)						
	Linear Outcome			NonLinear Outcome		
Method	Low	Mod	High	Low	Mod	High
IPTW(A)	100	100	100	100	100	100
g-computation(A)	88	54	36	79	63	51
AIPTW(A)	81	59	50	74	63	57
PENCOMP(A)	82	60	49	74	63	56
IPTW(A)	100	100	100	100	100	100
g-computation(B)	144	83	54	118	103	84
AIPTW(B)	82	87	90	95	95	95
PENCOMP(B)	119	90	79	105	106	106
IPTW(C)	152	99	69	120	113	98
g-computation(A)	88	54	36	79	63	51
AIPTW(C)	81	52	37	74	60	50
PENCOMP(C)	81	54	39	74	61	51

Table 17: 100 * Empirical bias over RMSE of IPTW (A), under (A) correctly-specified propensity and prediction models; (B) a correctly-specified propensity model only; (C) a correctly-specified prediction model only, based on 500 simulations with sample size of 200. Under the linear outcome model, $(\Delta_{11}, \Delta_{10}, \Delta_{01})$ were $(22.35, 11.17, 10.45)$, respectively. Under the nonlinear outcome model, $(\Delta_{11}, \Delta_{10}, \Delta_{01})$ were $(25.31, 12.69, 10.57)$, respectively.

100 * Empirical Bias / RMSE IPTW(A)																		
	Linear Outcome									Nonlinear Outcome								
	Δ_{11}			Δ_{10}			Δ_{01}			Δ_{11}			Δ_{10}			Δ_{01}		
Method	Low	Mod	High	Low	Mod	High	Low	Mod	High	Low	Mod	High	Low	Mod	High	Low	Mod	High
IPTW(A)	-2	7	8	-0	-26	-45	6	5	-1	-6	1	1	2	-25	-47	5	3	-2
g-computation(A)	1	1	-0	2	1	0	2	2	1	1	1	0	2	1	1	3	3	2
AIPTW(A)	-0	0	-1	1	1	-3	2	1	1	0	0	-0	1	0	-4	3	2	2
PENCOMP(A)	1	1	1	2	2	1	3	2	2	2	3	3	2	2	1	4	3	2
IPTW(A)	-2	7	8	-0	-26	-45	6	5	-1	-6	1	1	2	-25	-47	5	3	-2
g-computation(B)	-64	-34	-11	-75	-53	-55	-7	-5	-4	-52	-60	-50	-63	-83	-95	-8	-7	5
AIPTW(B)	-4	-1	2	-3	-5	-13	1	2	3	-8	-9	-9	-2	-22	-40	-1	-0	4
PENCOMP(B)	1	5	8	1	8	9	1	1	3	-4	-4	0	-2	-9	-15	2	-1	5
IPTW(C)	25	82	100	-86	-160	-193	-83	-179	-257	19	58	93	-57	-115	-148	-93	-208	-277
g-computation(A)	1	1	-0	2	1	0	2	2	1	1	1	0	2	1	1	3	3	2
AIPTW(C)	0	0	-1	1	0	-0	3	2	2	0	0	-0	1	1	0	3	3	2
PENCOMP(C)	1	1	0	2	0	-1	3	2	1	2	3	2	2	1	1	4	3	2

Table 18: 100* Ratio of empirical RMSE to RMSE of IPTW(A), denoted as 100 * RMSE/RMSE(IPTW(A)), under (A) correctly-specified propensity and prediction models; (B) a correctly-specified propensity model only; (C) a correctly-specified prediction model only, based on 500 simulations with sample size of 200.

100 * RMSE / RMSE IPTW(A)																		
	Linear Outcome									Nonlinear Outcome								
	Δ_{11}			Δ_{10}			Δ_{01}			Δ_{11}			Δ_{10}			Δ_{01}		
Method	Low	Mod	High	Low	Mod	High	Low	Mod	High	Low	Mod	High	Low	Mod	High	Low	Mod	High
IPTW(A)	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
g-computation(A)	58	52	41	45	26	21	34	30	27	53	59	57	39	23	21	34	32	27
AIPTW(A)	56	52	41	46	31	68	33	30	28	53	62	60	40	30	105	33	31	27
PENCOMP(A)	58	52	41	46	29	27	33	30	27	53	61	59	39	24	23	32	31	27
IPTW(A)	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
g-computation(B)	105	80	55	103	67	65	70	65	62	93	95	84	88	91	99	82	79	69
AIPTW(B)	71	65	54	64	45	70	66	62	59	82	84	85	74	72	78	82	78	75
PENCOMP(B)	63	63	55	50	42	51	43	44	46	69	76	85	51	40	52	61	62	62
IPTW(C)	97	116	118	140	175	200	139	213	277	108	121	139	110	128	153	145	236	292
g-computation(A)	58	52	41	45	26	21	34	30	27	53	59	57	39	23	21	34	32	27
AIPTW(C)	57	51	40	45	27	22	33	30	27	53	61	59	40	26	23	32	31	27
PENCOMP(C)	57	52	40	46	29	24	33	30	27	53	60	58	39	25	37	32	31	27

Table 19: Empirical 95% non-coverage rate*100 (nominal noncoverage of 5), under (A) correctly- specified propensity and prediction models; (B) a correctly-specified propensity model only; (C) a correctly-specified prediction model only, based on 500 simulations with sample size of 200.

100 * 95% Non-coverage Rate																		
Linear Outcome										Nonlinear Outcome								
	Δ_{11}			Δ_{10}			Δ_{01}			Δ_{11}			Δ_{10}			Δ_{01}		
Method	Low	Mod	High	Low	Mod	High	Low	Mod	High	Low	Mod	High	Low	Mod	High	Low	Mod	High
IPTW(A)	6	7	8	4	18	29	6	4	6	11	10	10	11	30	47	6	5	4
g-computation(A)	6	7	8	5	6	6	5	4	4	5	7	8	5	6	6	5	4	4
AIPTW(A)	5	8	6	5	4	3	5	4	4	5	8	7	6	5	3	4	4	4
PENCOMP(A)	4	3	2	2	1	1	2	1	1	4	4	3	2	1	1	1	1	1
IPTW(A)	6	7	8	4	18	29	6	4	6	11	10	10	11	30	47	6	5	4
g-computation(B)	15	10	8	20	24	36	5	6	6	18	20	18	22	66	85	5	6	5
AIPTW(B)	5	7	4	5	6	5	5	6	5	11	10	9	9	26	35	5	5	6
PENCOMP(B)	2	1	0	0	0	0	0	0	0	2	2	2	1	1	1	0	0	0
IPTW(C)	6	16	39	15	69	93	11	42	80	10	7	13	20	73	94	14	52	88
g-computation(A)	6	7	8	5	6	6	5	4	4	5	7	8	5	6	6	5	4	4
AIPTW(C)	6	6	7	5	4	6	4	3	5	5	8	8	5	6	7	5	3	3
PENCOMP(C)	4	3	3	1	1	1	2	2	2	3	3	3	2	2	2	2	1	2

Table 20: 100*Ratio of empirical 95% confidence interval width to that of IPTW (A), denoted as mean 95% interval width/that of IPTW(A), under (B) correctly-specified propensity and prediction models; (B) a correctly-specified propensity model only; (C) a correctly-specified prediction model only, based on 500 simulations with sample size of 200.

	100 * mean 95% interval width / mean 95% interval width IPTW(A)																	
	Linear Outcome									Nonlinear Outcome								
	Δ_{11}			Δ_{10}			Δ_{01}			Δ_{11}			Δ_{10}			Δ_{01}		
Method	Low	Mod	High	Low	Mod	High	Low	Mod	High	Low	Mod	High	Low	Mod	High	Low	Mod	High
IPTW(A)	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
g-computation(A)	57	48	40	45	33	32	34	31	31	58	58	55	45	37	41	33	32	30
AIPTW(A)	57	49	47	46	40	78	33	32	31	61	63	61	47	51	130	32	31	30
PENCOMP(A)	63	58	56	53	62	91	39	40	40	66	73	73	53	59	86	38	40	38
IPTW(A)	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
g-computation(B)	79	66	56	72	57	55	67	64	63	81	74	67	72	58	59	79	76	72
AIPTW(B)	72	63	60	66	61	87	64	61	61	84	84	86	78	78	87	79	76	75
PENCOMP(B)	83	92	108	74	108	191	67	73	78	104	119	172	93	135	228	89	99	98
IPTW(C)	90	80	66	113	88	75	107	106	98	105	107	102	101	75	63	106	103	91
g-computation(A)	57	48	40	45	33	32	34	31	31	58	58	55	45	37	41	33	32	30
AIPTW(C)	57	49	41	45	34	33	33	32	31	61	62	58	47	42	48	32	31	30
PENCOMP(C)	63	57	49	54	58	61	39	39	39	66	72	69	54	56	102	38	40	38

Table 21: 100 * Empirical bias over RMSE of IPTW (A), under (A) correctly-specified propensity and prediction models; (B) a correctly-specified propensity model only; (C) a correctly-specified prediction model only, based on 1000 simulations with sample size of 500. Under the linear outcome model, $(\Delta_{11}, \Delta_{10}, \Delta_{01})$ were $(22.35, 11.17, 10.45)$, respectively. Under the nonlinear outcome model, $(\Delta_{11}, \Delta_{10}, \Delta_{01})$ were $(25.31, 12.69, 10.57)$, respectively.

100 * Empirical Bias / RMSE IPTW(A)																		
	Linear Outcome									Nonlinear Outcome								
	Δ_{11}			Δ_{10}			Δ_{01}			Δ_{11}			Δ_{10}			Δ_{01}		
Method	Low	Mod	High	Low	Mod	High	Low	Mod	High	Low	Mod	High	Low	Mod	High	Low	Mod	High
IPTW(A)	-3	-0	2	-6	-17	-31	-5	-5	-5	-2	0	-0	-6	-16	-31	-3	-3	-1
g-computation(A)	2	1	1	1	0	1	-0	-0	0	2	1	1	1	0	0	1	1	1
AIPTW(A)	1	0	1	1	0	1	-1	-1	-0	1	1	1	1	0	1	0	0	1
PENCOMP(A)	3	1	1	1	-1	-0	-1	-1	-1	3	3	3	2	0	0	0	-0	0
IPTW(A)	-3	-0	2	-6	-17	-31	-5	-5	-5	-2	0	-0	-6	-16	-31	-3	-3	-1
g-computation(B)	-102	-51	-18	-132	-75	-67	-14	-14	-12	-77	-81	-72	-117	-104	-101	-10	-10	10
AIPTW(B)	0	0	1	-3	-5	-6	0	-0	1	-0	-2	-4	-5	-14	-25	3	3	7
PENCOMP(B)	3	5	6	1	6	7	0	0	1	-0	-4	-4	-3	-7	-9	1	-0	4
IPTW(C)	59	127	146	-155	-205	-226	-148	-289	-431	49	88	138	-111	-141	-157	-157	-326	-466
g-computation(A)	2	1	1	1	0	1	-0	-0	0	2	1	1	1	0	0	1	1	1
AIPTW(C)	1	1	1	1	0	1	-1	-0	0	1	1	1	1	-0	0	0	0	1
PENCOMP(C)	3	1	1	1	-0	1	-1	-1	-1	3	3	3	2	0	1	0	-0	0

Table 22: 100* Ratio of empirical RMSE to RMSE of IPTW(A), denoted as 100 * RMSE/RMSE(IPTW(A)), under (A) correctly-specified propensity and prediction models; (B) a correctly-specified propensity model only; (C) a correctly-specified prediction model only, based on 1000 simulations with sample size of 500.

100 * RMSE / RMSE IPTW(A)																		
	Linear Outcome									Nonlinear Outcome								
	Δ_{11}			Δ_{10}			Δ_{01}			Δ_{11}			Δ_{10}			Δ_{01}		
Method	Low	Mod	High	Low	Mod	High	Low	Mod	High	Low	Mod	High	Low	Mod	High	Low	Mod	High
IPTW(A)	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
g-computation(A)	61	48	36	48	21	16	36	32	32	55	50	49	43	17	14	36	33	32
AIPTW(A)	58	47	36	46	23	23	34	30	30	56	55	54	44	21	25	32	31	30
PENCOMP(A)	58	47	36	46	23	19	33	30	30	54	52	51	43	19	15	32	31	30
IPTW(A)	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
g-computation(B)	132	83	50	152	83	72	70	63	61	109	103	92	136	109	104	79	75	72
AIPTW(B)	72	59	49	65	42	40	62	57	58	83	80	86	79	76	78	75	73	76
PENCOMP(B)	61	55	47	48	33	39	41	41	44	67	65	74	54	33	40	55	56	59
IPTW(C)	111	151	158	194	214	230	183	310	444	120	140	174	149	147	159	188	342	475
g-computation(A)	61	48	36	48	21	16	36	32	32	55	50	49	43	17	14	36	33	32
AIPTW(C)	58	47	36	46	21	16	34	30	30	55	54	52	44	20	16	32	30	29
PENCOMP(C)	58	46	35	46	22	17	33	30	30	54	52	51	43	18	15	32	30	29

Table 23: Empirical 95% non-coverage rate*100 (nominal noncoverage of 5), under (A) correctly- specified propensity and prediction models; (B) a correctly-specified propensity model only; (C) a correctly-specified prediction model only, based on 1000 simulations with sample size of 500.

100 * 95% Non-coverage Rate																		
Linear Outcome										Nonlinear Outcome								
	Δ_{11}			Δ_{10}			Δ_{01}			Δ_{11}			Δ_{10}			Δ_{01}		
Method	Low	Mod	High	Low	Mod	High	Low	Mod	High	Low	Mod	High	Low	Mod	High	Low	Mod	High
IPTW(A)	5	6	6	6	18	29	5	6	6	8	8	8	9	28	44	5	6	5
g-computation(A)	6	7	8	6	6	6	6	5	6	6	7	7	6	6	6	5	5	6
AIPTW(A)	6	6	8	6	6	4	6	6	6	6	7	7	5	6	4	6	6	6
PENCOMP(A)	4	4	5	5	4	3	5	4	5	5	5	5	4	4	4	5	5	5
IPTW(A)	5	6	6	6	18	29	5	6	6	8	8	8	9	28	44	5	6	5
g-computation(B)	26	15	8	42	56	70	5	6	4	22	31	29	41	91	99	4	6	4
AIPTW(B)	6	6	7	6	8	9	4	5	5	7	8	8	8	26	41	4	5	6
PENCOMP(B)	2	2	2	2	2	2	0	1	1	4	3	2	2	3	3	1	1	1
IPTW(C)	9	40	84	30	91	100	29	80	99	6	11	28	29	87	98	34	89	100
g-computation(A)	6	7	8	6	6	6	6	5	6	6	7	7	6	6	6	5	5	6
AIPTW(C)	6	6	9	5	6	6	6	5	6	6	7	8	5	6	6	6	6	6
PENCOMP(C)	5	4	5	4	4	3	5	4	4	5	5	5	4	4	4	6	5	5

Table 24: 100*Ratio of empirical 95% confidence interval width to that of IPTW (A), denoted as mean 95% interval width/that of IPTW(A), under (B) correctly-specified propensity and prediction models; (B) a correctly-specified propensity model only; (C) a correctly-specified prediction model only, based on 1000 simulations with sample size of 500.

	100 * mean 95% interval width / mean 95% interval width IPTW(A)																	
	Linear Outcome									Nonlinear Outcome								
	Δ_{11}			Δ_{10}			Δ_{01}			Δ_{11}			Δ_{10}			Δ_{01}		
Method	Low	Mod	High	Low	Mod	High	Low	Mod	High	Low	Mod	High	Low	Mod	High	Low	Mod	High
IPTW(A)	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
g-computation(A)	60	49	38	48	29	26	36	32	32	58	56	54	45	28	28	36	33	31
AIPTW(A)	59	49	39	47	32	41	34	31	30	59	61	59	46	34	59	33	31	29
PENCOMP(A)	61	53	44	49	35	40	34	32	32	61	63	62	47	33	36	33	33	31
IPTW(A)	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
g-computation(B)	83	67	53	75	49	44	70	63	63	81	71	65	73	47	44	83	77	74
AIPTW(B)	73	62	53	66	51	58	64	59	59	84	84	87	79	76	80	80	75	76
PENCOMP(B)	75	74	70	60	57	86	57	57	59	84	89	112	71	71	121	77	79	78
IPTW(C)	92	80	62	115	79	63	108	107	100	109	109	106	100	63	48	108	102	91
g-computation(A)	60	49	38	48	29	26	36	32	32	58	56	54	45	28	28	36	33	31
AIPTW(C)	59	49	38	47	29	26	34	31	30	59	60	58	47	32	32	33	31	29
PENCOMP(C)	61	53	43	49	34	32	35	32	32	61	63	62	47	32	33	33	32	31

Table 25: 100 * Empirical bias over RMSE of IPTW (A), under (A) correctly-specified propensity and prediction models; (B) a correctly-specified propensity model only; (C) a correctly-specified prediction model only, based on 500 simulations with sample size of 1000. Under the linear outcome model, $(\Delta_{11}, \Delta_{10}, \Delta_{01})$ were $(22.35, 11.17, 10.45)$, respectively. Under the nonlinear outcome model, $(\Delta_{11}, \Delta_{10}, \Delta_{01})$ were $(25.31, 12.69, 10.57)$, respectively.

100 * Empirical Bias / RMSE IPTW (A)																		
	Linear Outcome									Nonlinear Outcome								
	Δ_{11}			Δ_{10}			Δ_{01}			Δ_{11}			Δ_{10}			Δ_{01}		
Method	Low	Mod	High	Low	Mod	High	Low	Mod	High	Low	Mod	High	Low	Mod	High	Low	Mod	High
IPTW(A)	-3	-0	6	5	-14	-24	-1	2	1	-3	1	2	1	-19	-28	1	3	1
g-computation(A)	6	6	3	4	2	1	-0	-1	-1	3	3	3	3	2	1	2	1	1
AIPTW(A)	5	5	3	2	2	1	-1	-2	-1	3	4	4	2	1	1	1	-1	0
PENCOMP(A)	5	5	3	2	1	1	-2	-3	-2	4	5	4	2	1	1	0	-1	-1
IPTW(A)	-3	-0	6	5	-14	-24	-1	2	1	-3	1	2	1	-19	-28	1	3	1
g-computation(B)	-146	-76	-25	-182	-110	-87	-22	-18	-14	-116	-120	-107	-156	-157	-131	-17	-13	15
AIPTW(B)	1	2	7	1	-2	-5	-3	-3	2	-1	1	0	-2	-17	-24	-2	-2	2
PENCOMP(B)	4	4	8	2	10	8	-3	-4	-1	-1	-2	-5	-3	-5	-7	-3	-5	-0
IPTW(C)	84	192	205	-203	-306	-296	-207	-405	-585	71	134	209	-142	-210	-201	-225	-467	-670
g-computation(A)	6	6	3	4	2	1	-0	-1	-1	3	3	3	3	2	1	2	1	1
AIPTW(C)	5	5	3	3	1	1	-1	-2	-2	3	4	3	2	1	1	1	-1	-0
PENCOMP(C)	5	5	3	2	1	1	-2	-3	-2	4	5	5	2	1	1	0	-1	-1

Table 26: 100* Ratio of empirical RMSE to RMSE of IPTW(A), denoted as 100 * RMSE/RMSE(IPTW(A)), under (A) correctly-specified propensity and prediction models; (B) a correctly-specified propensity model only; (C) a correctly-specified prediction model only, based on 500 simulations with sample size of 1000.

100 * RMSE / RMSE IPTW(A)																		
	Linear Outcome									Nonlinear Outcome								
	Δ_{11}			Δ_{10}			Δ_{01}			Δ_{11}			Δ_{10}			Δ_{01}		
Method	Low	Mod	High	Low	Mod	High	Low	Mod	High	Low	Mod	High	Low	Mod	High	Low	Mod	High
IPTW(A)	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
g-computation(A)	67	56	40	49	24	16	40	35	33	61	58	58	44	20	13	41	38	36
AIPTW(A)	61	52	38	48	26	22	34	31	30	59	60	60	44	23	21	32	31	30
PENCOMP(A)	60	52	37	49	24	18	34	30	30	57	58	57	44	21	14	32	31	30
IPTW(A)	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
g-computation(B)	171	106	56	197	117	91	78	65	62	143	140	127	170	161	132	92	80	79
AIPTW(B)	72	63	47	68	41	37	66	58	59	83	86	104	77	75	78	83	78	83
PENCOMP(B)	61	60	48	51	34	36	42	39	43	70	71	83	53	34	38	60	58	61
IPTW(C)	124	209	213	234	313	298	236	421	595	128	174	237	172	215	202	253	480	677
g-computation(A)	67	56	40	49	24	16	40	35	33	61	58	58	44	20	13	41	38	36
AIPTW(C)	61	52	37	49	24	16	34	32	30	59	60	59	44	22	15	33	32	31
PENCOMP(C)	60	52	37	49	24	16	34	30	29	57	57	57	44	20	14	32	31	30

Table 27: Empirical 95% non-coverage rate*100 (nominal noncoverage of 5), under (A) correctly- specified propensity and prediction models; (B) a correctly-specified propensity model only; (C) a correctly-specified prediction model only, based on 500 simulations with sample size of 1000.

100 * 95% Non-coverage Rate																		
Linear Outcome										Nonlinear Outcome								
	Δ_{11}			Δ_{10}			Δ_{01}			Δ_{11}			Δ_{10}			Δ_{01}		
Method	Low	Mod	High	Low	Mod	High	Low	Mod	High	Low	Mod	High	Low	Mod	High	Low	Mod	High
IPTW(A)	7	5	9	4	12	26	5	5	6	7	8	8	7	21	36	3	4	5
g-computation(A)	8	9	9	5	5	8	6	6	7	8	8	9	7	6	6	7	5	6
AIPTW(A)	8	8	9	5	8	8	5	5	6	7	8	9	7	8	6	4	4	4
PENCOMP(A)	7	7	7	7	4	4	4	4	4	7	7	7	7	6	6	4	3	4
IPTW(A)	7	5	9	4	12	26	5	5	6	7	8	8	7	21	36	3	4	5
g-computation(B)	41	23	11	69	81	94	6	6	5	33	50	48	64	100	100	5	5	6
AIPTW(B)	7	6	8	6	7	9	5	6	6	6	8	10	7	20	37	5	5	6
PENCOMP(B)	4	3	5	4	3	4	1	1	1	5	5	5	4	4	6	2	2	1
IPTW(C)	17	74	98	44	98	100	46	97	100	8	23	58	39	96	100	52	98	100
g-computation(A)	8	9	9	5	5	8	6	6	7	8	8	9	7	6	6	7	5	6
AIPTW(C)	8	8	9	6	7	8	5	5	5	8	8	9	7	8	8	5	4	5
PENCOMP(C)	7	6	6	6	5	6	4	4	4	7	6	6	7	6	5	4	3	4

Table 28: 100*Ratio of empirical 95% confidence interval width to that of IPTW (A), denoted as mean 95% interval width/that of IPTW(A), under (B) correctly-specified propensity and prediction models; (B) a correctly-specified propensity model only; (C) a correctly-specified prediction model only, based on 500 simulations with sample size of 1000.

	100 * mean 95% interval width / mean 95% interval width IPTW(A)																	
	Linear Outcome									Nonlinear Outcome								
	Δ_{11}			Δ_{10}			Δ_{01}			Δ_{11}			Δ_{10}			Δ_{01}		
Method	Low	Mod	High	Low	Mod	High	Low	Mod	High	Low	Mod	High	Low	Mod	High	Low	Mod	High
IPTW(A)	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
g-computation(A)	63	51	39	51	27	22	38	34	33	58	57	55	46	26	23	39	37	34
AIPTW(A)	60	49	38	47	29	30	33	30	30	59	61	60	46	29	36	33	31	29
PENCOMP(A)	61	52	41	48	29	29	34	31	30	59	62	61	46	28	26	33	31	30
IPTW(A)	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
g-computation(B)	86	68	52	77	45	38	71	64	63	82	71	66	72	42	36	85	79	76
AIPTW(B)	74	61	51	66	48	49	64	58	59	84	87	93	79	75	79	80	76	78
PENCOMP(B)	71	68	61	54	46	62	52	51	53	77	83	102	61	54	88	72	74	72
IPTW(C)	92	80	60	115	74	55	108	106	100	109	110	108	100	59	40	108	102	92
g-computation(A)	63	51	39	51	27	22	38	34	33	58	57	55	46	26	23	39	37	34
AIPTW(C)	60	49	37	47	27	22	33	30	30	58	60	58	46	29	26	33	31	29
PENCOMP(C)	61	52	40	48	29	25	34	31	30	59	62	60	46	27	25	33	31	30

4 Supplemental Table from Application

Table 29: The number of subjects with observed treatment regimen $(1,1)$, $(1,0)$, $(0,1)$, and $(0,0)$, denoted as no11, no10, no01, and no00, respectively in each three-visit window, as well as the number of subjects kept in the estimation of Δ_{11}, Δ_{10}, and Δ_{00}, after trimming, denoted as no Δ_{11}, no Δ_{10}, and no Δ_{01}, respectively. The total is the total number of subjects with complete data on blood count measures considered in the models in each window.

	sample size observed				sample size after trimming			
Window	no11	no10	no01	no00	no Δ_{11}	no Δ_{10}	no Δ_{01}	Total
Window1	88	11	82	638	772	731	770	819
Window2	138	16	88	620	770	794	785	862
Window3	178	23	76	550	635	700	721	827
Window4	160	42	134	352	612	459	603	688
Window5	265	13	114	292	509	458	518	684
Window6	390	26	59	348	773	749	756	823
Window7	401	13	41	322	694	648	686	777
Window8	397	12	43	299	717	655	564	751
Window9	389	14	30	281	541	518	544	714
Window10	373	14	48	245	516	462	476	680
Window11	356	21	37	225	590	545	562	639
Window12	310	36	16	217	552	514	532	579
Window13	254	45	24	220	504	395	437	543
Window14	216	31	30	203	420	410	353	480
Window15	197	14	39	182	374	365	373	432

Table 30: Summary of the stabilized weights.

	Stabilized Weights	
Window	Mean(SD)	Minimum/Maximum
Window1	1.091 (1.97)	0.1103 / 40.3
Window2	1.065 (3.20)	0.1026 / 91.5
Window3	6.160 (146.78)	0.2010 / 4220.5
Window4	4.662 (83.11)	0.1391 / 2163.2
Window5	0.966 (1.11)	0.3274 / 15.2
Window6	2.378 (37.86)	0.4039 / 1083.5
Window7	3.052 (59.23)	0.1692 / 1651.0
Window8	23.893 (618.50)	0.1102 / 16949.0
Window9	4.085 (63.72)	0.2095 / 1541.6
Window10	6.937 (106.37)	0.1468 / 2307.3
Window11	1.586 (11.11)	0.2741 / 250.8
Window12	1.731 (12.57)	0.2944 / 266.1
Window13	1.336 (7.32)	0.1705 / 164.7
Window14	1.033 (1.67)	0.1935 / 17.6
Window15	1.046 (2.05)	0.2134 / 32.0

Table 31: Summary of overlap proportions at both time points.

	First Time Point		Second Time Point			
	$\pi_1^{0.95}$	$\pi_0^{0.95}$	$\pi_{11}^{0.95}$	$\pi_{10}^{0.95}$	$\pi_{01}^{0.95}$	$\pi_{00}^{0.95}$
Window1	81	40	73	86	89	51
Window2	60	38	63	96	83	45
Window3	50	37	41	96	86	38
Window4	36	28	34	42	69	39
Window5	45	78	45	79	92	39
Window6	40	31	40	92	84	35
Window7	36	19	34	82	87	17
Window8	21	14	22	85	61	18
Window9	22	12	23	78	70	14
Window10	8	6	12	74	31	15
Window11	27	19	33	79	62	21
Window12	38	25	40	67	51	28
Window13	22	52	34	57	73	49
Window14	42	51	38	85	52	57
Window15	32	51	34	96	82	41

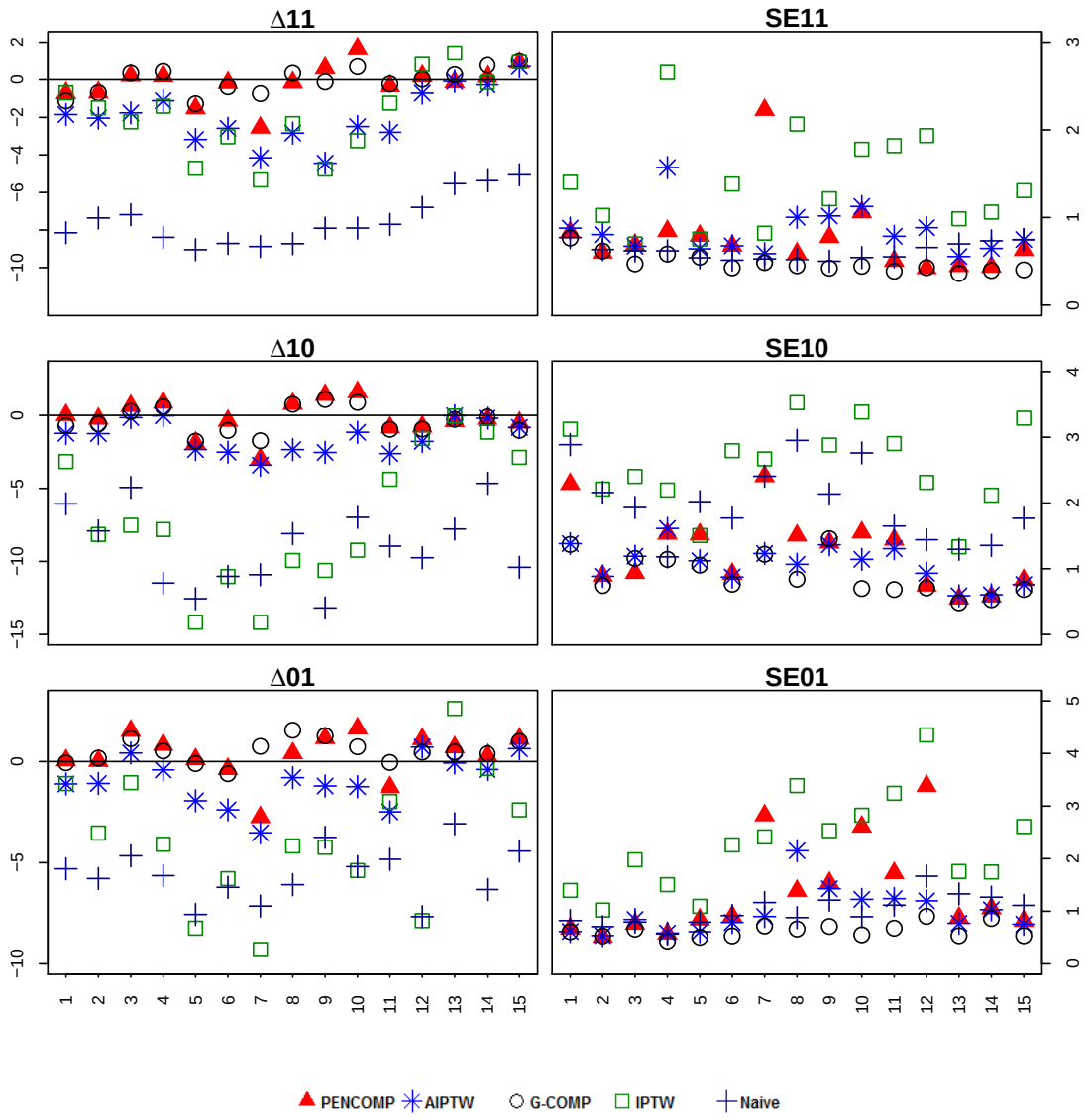

Figure 11: For each of the three-visit windows $1, \cdots, 15$, the estimates and standard errors (SE) of the treatment effects Δ_{11}, Δ_{10}, and Δ_{01} of the four methods: PENCOMP, AIPTW, IPTW, and Naive. Here 1st% and 99th% weight truncation was done for IPTW and AIPTW. PENCOMP estimates were computed on the overlapping regions, as described in Section 2.4. Since the propensity score distributions were very skewed for some windows, restricting to the quantiles $c(\alpha, 1-\alpha)$ (for example $\alpha = 0.025$) of the propensity score distributions can significantly reduce the variances without changing the estimates much (results not shown here). Note the estimands are different.