BIOMETRICS 000, 1–18 DOI: 000
000 0000

Wed-based Supplementary Materials for "Comparing Treatment Policies with Assistance from the Structural Nested Mean Model"

by

Xi Lu, Kevin G. Lynch, David W. Oslin and Susan Murphy

This paper has been submitted for consideration for publication in *Biometrics*

Web Appendix A: Proofs

Proof of Lemma 1

We first write a telescoping sum of the conditional mean of $Y(A_1, A_2)$. Since $A_1|H_1 = h_1$ has a conditional distribution given by $p_1(\cdot|h_1)$ (*) and $A_2|H_2(A_1) = h_2$ has a conditional distribution given by $p_2(\cdot|h_2)$ (**), we have $E[Y(A_1, 0)|H_2(A_1), A_2] = E[Y(A_1, 0)|H_2(A_1)]$ and $E[Y(0, 0)|H_1, A_1] = E[Y(0, 0)|H_1]$. Thus we have:

$$\begin{aligned} E[Y(A_1, A_2)|H_2(A_1), A_2] &= E[Y(A_1, A_2)|H_2(A_1), A_2] - E[Y(A_1, 0)|H_2(A_1), A_2] \\ &+ E[Y(A_1, 0)|H_2(A_1)] - E[E[Y(A_1, 0)|H_2(A_1)]|H_1, A_1] \\ &+ E[Y(A_1, 0)|H_1, A_1] - E[Y(0, 0)|H_1, A_1] \\ &+ E[Y(0, 0)|H_1] \end{aligned}$$

Note that the first line on the right hand side is equal to $\mu_2(H_2(A_1), A_2)$ due to (**) and the third line is equal to $\mu_1(H_1, A_1)$ due to (*); the second line has a conditional mean zero, conditional on (H_1, A_1). Thus we conclude that $E[Y(A_1, A_2) - \mu_2(H_2(A_1), A_2) - \mu_1(H_1, A_1)] = E[Y(0, 0)]$.

For a fixed policy d, the associated potential outcomes are $\{X_1, X_2(d_1), Y(d_1, d_2)\}$. Now let us focus on the telescoping sum of the conditional mean of $Y(d_1, d_2)$. Due to (*), we have $E[Y(A_1, a_2)|X_1, A_1, X_2(A_1)]1_{A_1=a_1} = E[Y(a_1, a_2)|X_1, X_2(a_1)]1_{A_1=a_1}$; this implies $E[Y(A_1, a_2)|X_1, A_1, X_2(A_1)]1_{A_1=d_1(H_1)} = E[Y(d_1, a_2)|X_1, X_2(d_1)]1_{A_1=d_1(H_1)}$ because $d_1(H_1)$ is known given X_1. Moreover, since $d_2(H_2(A_1)) = d_2(H_2(d_1))$ on event $\{A_1 = d_1(H_1)\}$, we have $E[Y(A_1, d_2)|X_1, A_1, X_2(A_1)]1_{A_1=d_1(H_1)} = E[Y(d_1, d_2)|X_1, X_2(d_1)]1_{A_1=d_1(H_1)}$. Now let $p_1(\cdot|h_1)$ be a degenerate distribution, that concentrates on $d_1(h_1)$, we then conclude that $E[Y(d_1, d_2)|X_1, X_2(d_1)] - E[Y(d_1, 0)|X_1, X_2(d_1)] = \mu_2(H_2(a_1), a_2)|_{a_2=d_2(H_2(a_1)), a_1=d_1(H_1)}$. Similarly one can show $E[Y(d_1, 0)|X_1] - E[Y(0, 0)|X_1] = \mu_1(H_1, a_1)|_{a_1=d_1(H_1)}$

Based on the arguments above, we can write:

$$\begin{aligned}E[Y(d_1,d_2)|X_1,X_2(d_1)] &= E[Y(d_1,d_2)|X_1,X_2(d_1)] - E[Y(d_1,0)|X_1,X_2(d_1)]\\ &+ E[Y(d_1,0)|X_1,X_2(d_1)] - E\left[E[Y(d_1,0)|X_1,X_2(d_1)]|X_1\right]\\ &+ E[Y(d_1,0)|X_1] - E[Y(0,0)|X_1]\\ &+ E[Y(0,0)|X_1]\end{aligned}$$

and conclude that $E[Y(d_1,d_2)] = E[\mu_1(H_1,a_1)|_{a_1=d_1(H_1)} + \mu_2(H_2(a_1),a_2)|_{a_2=d_2(H_2(a_1)),a_1=d_1(H_1)} + Y(0,0)]$. Thus $V_d = E\big[Y(A_1,A_2) - \mu_2(H_2(A_1),A_2) - \mu_1(H_1,A_1) + \mu_1(H_1,d_1(H_1)) + \mu_2(H_2(a_1),a_2)|_{a_2=d_2(H_2(a_1)),a_1=d_1(H_1)}\big]$. Finally, consider the transition from the degenerate distribution of A_1 that concentrates on $d_1(H_1)$ to the distribution of A_1 given by $p_1(\cdot|H_1)$, we then can rewrite $E[\mu_2(H_2(a_1),a_2)|_{a_2=d_2(H_2(a_1)),a_1=d_1(H_1)}]$ as $E\left[\frac{I\{A_1=d_1(H_1)\}}{p_1(A_1|H_1)}\mu_2(H_2(A_1),d_2(H_2(A_1)))\right]$. This completes the proof of Lemma 1.

Proof of Lemma 2

First we prove the equality for the second-stage treatment effect. By sequential randomization of A_1, $E[Y(a_1,a_2)|H_2(a_1)=h_2] = E[Y(A_1,a_2)|H_2(A_1)=h_2]$ (note that a_1 is also part of h_2), which is then equal to $E[Y(A_1,A_2)|H_2(A_1)=h_2, A_2=a_2]$ due to sequential randomization of A_2. Finally by consistency assumption, we conclude that $E[Y(a_1,a_2)|H_2(a_1)=h_2] = E[Y|H_2=h_2,A_2=a_2]$, thus the first equalilty for $\mu_2(h_2,a_2)$ holds.

Next we prove the equality for the first-stage treatment effect. By sequential randomization of A_1, $E[Y(a_1,0)|H_1=h_1] = E[Y(a_1,0)|H_1=h_1,A_1=a_1]$, which is then equal to $E[E[Y(a_1,0)|H_2(a_1),A_2=0]|H_1=h_1,A_1=a_1]$ due to sequential randomization of A_2. Re-using sequential randomization of A_1 for the inner conditional mean, this quantity can be written as $E[E[Y(A_1,A_2)|H_2(A_1),A_2=0]|H_1=h_1,A_1=a_1]$. Finally by consistency assumption, we conclude that $E[Y(a_1,0)|H_1=h_1] = E[E[Y|H_2,A_2=0]|H_1=h_1,A_1=a_1]$, thus the equality for $\mu_1(h_1,a_1)$ holds.

The equality of expressing the policy value V_d with the observed data directly follows from Lemma 1, due to the consistency assumption.

Proof of Lemma 3

(a) We know from Lemma 2 that $V_d = E\Big[Y - \mu_2(H_2, A_2) - \mu_1(H_1, A_1) + \mu_1(H_1, d_1(H_1)) + \frac{I\{A_1 = d_1(H_1)\}}{p_1(A_1|H_1)}\mu_2(H_2, d_2(H_2))\Big]$. Again using the transition from the degenerate distribution of A_1 that concentrates on $d_1(H_1)$ to the distribution of A_1 given by $p_1(\cdot|H_1)$, it is obvious to see that that $E[m(H_1, d_1(H_1)) - \frac{I\{A_1=d_1(H_1)\}}{p_1(A_1|H_1)}m(H_1, A_1)] = 0, \forall m$ that satisfies the integrable condition.

(b) The asymptotic variance of $\hat{V}_0(d;\hat{\beta})$ is equal to $Var(f_{\beta_0} + C_\varphi^T\varphi)$, and the asymptotic variance of $\hat{V}_m(d;\hat{\beta})$ is equal to $Var(f_{\beta_0} + C_\varphi^T\varphi + g_m)$, where the term $C_\varphi^T\varphi$ comes from the estimation of paramter β in the SNMM (φ is the influence function for β); $f_\beta(h_2, a_2, y) = y - \mu_1(h_1, a_1; \beta_1) - \mu_2(h_2, a_2; \beta_2) + \mu_1(h_1, d_1(h_1); \beta_1) + \omega_{d_1}(h_1, a_1)\mu_2(h_2, d_2(h_2); \beta_2)$, and $g_m(h_1, a_1) = m(h_1, d_1(h_1)) - \omega_{d_1}(h_1, a_1)m(h_1, a_1)$. Then the difference in asymptotic variance between $\hat{V}_m(d;\hat{\beta})$ and $\hat{V}_0(d;\hat{\beta})$ is equal to $2Cov(f_{\beta_0} + C_\varphi^T\varphi, g_m) + Var(g_m)$.

We note that for $\hat{\beta}$ in subclass $\mathscr{B}$ defined in the appendix, $Cov(C_\varphi^T\varphi, g_m) = 0$. More specifically, when $\hat{\beta}$ belongs to the subclass $\mathscr{B}$, it is the solution to an estimating equation with the nuisance function $q_1(h_1;\xi)$ chosen optimally (see the review of g-estimators in the appendix), and one can show that $E[\varphi \cdot g_m] = 0$. Thus for those $\hat{\beta}$'s, $Cov(C_\varphi^T\varphi, g_m) = 0$, and we only need to focus on $2Cov(f_{\beta_0}, g_m) + Var(g_m)$; i.e., the derivation of the optimal m function is the same as the arguments under a known β. For more general $\hat{\beta}$'s, $Cov(C_\varphi^T\varphi, g_m)$ would depend on the estimating equation that produces $\hat{\beta}$ as well as the policy d in a complicated way, thus affecting the choice of optimal m function; for simplicity, in this lemma we assume that $\hat{\beta}$ belongs to $\mathscr{B}$.

In addition, note that $E[(Y - \mu_1(H_1, A_1) - \mu_2(H_2, A_2)) \cdot g_m] = 0$ by taking the conditional mean with respect to (H_1, A_1). As a result, the optimal choice of m remains the same whether the estimator is for the value V_d, or for the contrast between policy d and a static policy that always assigns treatment 0.

Denote $m^*(h_1, a_1) \equiv E[\mu_2(H_2, d_2(H_2))|H_1 = h_1, A_1 = a_1]$; for simplicity, we write ω_{d_1}, m and m^* in short for $\omega_{d_1}(H_1, A_1), m(H_1, A_1)$ and $m^*(H_1, A_1)$, and write $m \circ d_1$ and $m^* \circ d_1$ in short for $m(H_1, d_1(H_1))$ and $m^*(H_1, d_1(H_1))$. Then, since $E[g_m(H_1, A_1)|H_1 = h_1] \equiv 0$, we could derive that $2Cov(f_{\beta_0}, g_m) + Var(g_m) = E[\omega_{d_1}(2m^* - m)(m \circ d_1 - \omega_{d_1} m)]$. Re-using the fact that $E[m \circ d_1 - \omega_{d_1} m|H_1 = h_1] \equiv 0$ for arbitrary function m, we have $2Cov(f_{\beta_0}, g_m) + Var(g_m) = E[\{(m - m^*) \circ d_1 - \omega_{d_1}(m - m^*)\}^2] - E[(m^* \circ d_1 - \omega_{d_1} m^*)^2]$. Thus the use of function m in the assisted estimator leads to efficiency improvement when $E[\{(m - m^*) \circ d_1 - \omega_{d_1}(m - m^*)\}^2] < E[(m^* \circ d_1 - \omega_{d_1} m^*)^2]$; in particular, the largest efficiency improvement is achieved when $m(h_1, d_1(h_1)) \equiv m^*(h_1, d_1(h_1))$. Note that, the values of function m at only $(h_1, d_1(h_1))$ have an impact on $\hat{V}_m(d; \hat{\beta})$.

Remark: To get more intuition about when the efficiency improvement that is achieved by using function m can be large, consider a simple scenario where treatments are binary and equally randomized in the data. Then $E[(m^* \circ d_1 - \omega_{d_1} m^*)^2]$, the maximal amount of variance reduction, is equal to $E[m^*(H_1, d_1(H_1))^2]$. This quantity can be large if, under the circumstance that d_1 is followed in stage one, on average the treatment recommended by d_2 at stage two has a large treatment effect.

Proof of Theorem 1

Under regularity conditions, the following class of functions is Glivenko-Cantelli:

$$\begin{aligned}&\{y - \mu_2(h_2, a_2; \beta_2) - \mu_1(h_1, a_1; \beta_1) + \mu_1(h_1, d_1(h_1); \beta_1)\\ &\quad + \frac{I\{a_1 = d_1(h_1)\}}{p_1(a_1|h_1)}\left(\mu_2(h_2, d_2(h_2); \beta_2) - m(h_1, a_1; \alpha_m)\right) + m(h_1, d_1(h_1); \alpha_m) :\\ &\quad \|\beta_1 - \beta_{10}\| \leqslant \delta, \|\beta_2 - \beta_{20}\| \leqslant \delta, \|\alpha_m - \alpha_m^+\| \leqslant \delta\}\end{aligned}$$

The theorem then follows from Lemma 2 and Lemma 3 by applying Glivenko-Cantelli Theorem to this function class.

Proof of Theorem 2

Since $\hat{V}_0(d;\hat{\beta})$ is a special case of $\hat{V}_{\hat{m}}(d;\hat{\beta})$, we only prove the asymptotic normality of the latter one. We write $\hat{V}_d$ in short for $\hat{V}_{\hat{m}}(d;\hat{\beta})$ and V_d in short for the true policy value of d. For ease of notation, define $\omega_{d_1}(H_1,A_1)=\frac{I\{A_1=d_1(H_1)\}}{p_1(A_1|H_1)}$. Then

$$
\begin{aligned}
\sqrt{n}(\hat{V}_d-V_d)=&\sqrt{n}\mathbb{P}_n\big\{Y-\mu_2(H_2,A_2;\hat{\beta}_2)-\mu_1(H_1,A_1;\hat{\beta}_1)+\mu_1(H_1,d_1(H_1);\hat{\beta}_1)\\
&\qquad+\omega_{d_1}(H_1,A_1)\mu_2(H_2,d_2(H_2);\hat{\beta}_2)\big\}\\
&-\sqrt{n}P\big\{Y-\mu_2(H_2,A_2;\beta_{20})-\mu_1(H_1,A_1;\beta_{10})+\mu_1(H_1,d_1(H_1);\beta_{10})\\
&\qquad+\omega_{d_1}(H_1,A_1)\mu_2(H_2,d_2(H_2);\beta_{20})\big\}\\
&+\sqrt{n}\mathbb{P}_n\big\{m(H_1,d_1(H_1);\hat{\alpha}_m)-\omega_{d_1}(H_1,A_1)m(H_1,A_1;\hat{\alpha}_m)\big\}\\
&-\sqrt{n}P\big\{m(H_1,d_1(H_1);\alpha_m^+)-\omega_{d_1}(H_1,A_1)m(H_1,A_1;\alpha_m^+)\big\}.
\end{aligned}
$$

Under the regularity conditions specified in the theorem, $P\mu_1(H_1,A_1;\beta_1)$, as a function of β_1, is differentiable, and the order of differentiation and integration can be interchanged; moreover $P\dot{\mu}_1(H_1,A_1;\beta_1)$ is continuous in β_1 in a neighborhood of β_{10}. Combined with the fact that $\hat{\beta}_1$ converges in probability to β_{10}, we have: $\sqrt{n}P\mu_1(H_1,A_1;\hat{\beta}_1)-\sqrt{n}P\mu_1(H_1,A_1;\beta_{10})=(P\dot{\mu}_1(H_1,A_1;\beta_{10})+o_p(1))\sqrt{n}(\hat{\beta}_1-\beta_{10})$. By similar arguments and the assumptions that

$\sqrt{n}(\hat{\beta}_1 - \beta_{10}) = O_p(1)$, $\sqrt{n}(\hat{\beta}_2 - \beta_{20}) = O_p(1)$, we can get:

$$\begin{aligned}\sqrt{n}(\hat{V}_d - V_d) =& \sqrt{n}(\mathbb{P}_n - P)\big\{Y - \mu_2(H_2, A_2; \hat{\beta}_2) - \mu_1(H_1, A_1; \hat{\beta}_1) + \mu_1(H_1, d_1(H_1); \hat{\beta}_1)\\ &\quad + \omega_{d_1}(H_1, A_1)\mu_2(H_2, d_2(H_2); \hat{\beta}_2)\big\}\\ &+ P\big[\omega_{d_1}(H_1, A_1)\dot{\mu}_2(H_2, d_2(H_2); \beta_{20}) - \dot{\mu}_2(H_2, A_2; \beta_{20})\big]\sqrt{n}(\hat{\beta}_2 - \beta_{20})\\ &+ P\big[\dot{\mu}_1(H_1, d_1(H_1); \beta_{10}) - \dot{\mu}_1(H_1, A_1; \beta_{10})\big]\sqrt{n}(\hat{\beta}_1 - \beta_{10})\\ &+ \sqrt{n}\mathbb{P}_n\big\{m(H_1, d_1(H_1); \hat{\alpha}_m) - \omega_{d_1}(H_1, A_1)m(H_1, A_1; \hat{\alpha}_m)\big\}\\ &- \sqrt{n}P\big\{m(H_1, d_1(H_1); \alpha_m^+) - \omega_{d_1}(H_1, A_1)m(H_1, A_1; \alpha_m^+)\big\} + o_p(1).\end{aligned}$$

Under regularity conditions on $m(h_1, a_1; \alpha_m)$, we can derive

$$\begin{aligned}&\sqrt{n}P\left(m(H_1, d_1(H_1); \hat{\alpha}_m) - \omega_{d_1}(H_1, A_1)m(H_1, A_1; \hat{\alpha}_m)\right)\\ &\quad - \sqrt{n}P\left(m(H_1, d_1(H_1); \alpha_m^+) - \omega_{d_1}(H_1, A_1)m(H_1, A_1; \alpha_m^+)\right)\\ &\quad = \left(P\dot{m}(H_1, d_1(H_1); \alpha_m^+) - P\omega_{d_1}(H_1, A_1)\dot{m}(H_1, A_1; \alpha_m^+) + o_p(1)\right)\sqrt{n}(\hat{\alpha}_m - \alpha_m^+).\end{aligned}$$

Since $P\left[m(H_1, d_1(H_1); \alpha_m) - \omega_{d_1}(H_1, A_1)m(H_1, A_1; \alpha_m)\right] \equiv 0$, for all α_m, we derive the following equality, as long as $\sqrt{n}(\hat{\alpha}_m - \alpha_m^+) = O_p(1)$:

$$\begin{aligned}\sqrt{n}(\hat{V}_d - V_d) =& \sqrt{n}(\mathbb{P}_n - P)\big\{Y - \mu_2(H_2, A_2; \hat{\beta}_2) - \mu_1(H_1, A_1; \hat{\beta}_1) + \mu_1(H_1, d_1(H_1); \hat{\beta}_1)\\ &\quad + \omega_{d_1}(H_1, A_1)\mu_2(H_2, d_2(H_2); \hat{\beta}_2)\big\}\\ &+ P\left[\omega_{d_1}(H_1, A_1)\dot{\mu}_2(H_2, d_2(H_2); \beta_{20}) - \dot{\mu}_2(H_2, A_2; \beta_{20})\right]\sqrt{n}(\hat{\beta}_2 - \beta_{20})\\ &+ P\left[\dot{\mu}_1(H_1, d_1(H_1); \beta_{10}) - \dot{\mu}_1(H_1, A_1; \beta_{10})\right]\sqrt{n}(\hat{\beta}_1 - \beta_{10})\\ &+ \sqrt{n}(\mathbb{P}_n - P)\left\{m(H_1, d_1(H_1); \hat{\alpha}_m) - \omega_{d_1}(H_1, A_1)m(H_1, A_1; \hat{\alpha}_m)\right\} + o_p(1)\end{aligned}$$

Define functions indexed by $\beta_1, \beta_2, \alpha_m$ as $f_{\beta_1,\beta_2,\alpha_m}(x_1, a_1, x_2, a_2, y) = y - \mu_2(h_2, a_2; \beta_2) - \mu_1(h_1, a_1; \beta_1) + \mu_1(h_1, d_1(h_1); \beta_1) + \omega_{d_1}(h_1, a_1)\mu_2(h_2, d_2(h_2); \beta_2) + m(h_1, d_1(h_1); \alpha_m) - \omega_{d_1}(h_1, a_1)m(h_1, a_1; \alpha_m)$ and function class

$\mathscr{F}_\delta = \left\{\tilde{f}_{\beta_1,\beta_2,\alpha_m} := 1_{\|\beta_1-\beta_{10}\|\leqslant\delta, \|\beta_2-\beta_{20}\|\leqslant\delta, \|\alpha_m-\alpha_m^+\|\leqslant\delta} f_{\beta_1,\beta_2,\alpha_m}\right\}$. Since $\hat{\beta} \xrightarrow{p} \beta_0$ and $\hat{\alpha}_m \xrightarrow{p} \alpha_m^+$, $\sqrt{n}(\mathbb{P}_n - P)f_{\hat{\beta}_1,\hat{\beta}_2,\hat{\alpha}_m} = \sqrt{n}(\mathbb{P}_n - P)\tilde{f}_{\hat{\beta}_1,\hat{\beta}_2,\hat{\alpha}_m} + o_p(1)$. Under regularity conditions, $P\sup|\tilde{f}_{\beta_1,\beta_2,\alpha_m}|^2 <$

∞, thus $P[\tilde{f}_{\hat{\beta}_1,\hat{\beta}_2,\hat{\alpha}_m} - \tilde{f}_{\beta_{10},\beta_{20},\alpha_m^+}]^2 \xrightarrow{p} 0$. By assuming $\sum_{a_1} P \sup_{\|\beta_1-\beta_{10}\|\leqslant\delta} |\dot{\mu}_1(H_1,a_1;\beta_1)|^2 + |\mu_1(H_1,a_1;\beta_1)|^2 < \infty$, $\sum_{a_2} P \sup_{\|\beta_2-\beta_{20}\|\leqslant\delta} |\dot{\mu}_2(H_2,a_2;\beta_2)|^2 + |\mu_2(H_2,a_2;\beta_2)|^2 < \infty$ and $\sum_{a_1} P \sup_{\|\alpha_m-\alpha_m^+\|\leqslant\delta} |\dot{m}_1(H_1,a_1;\alpha_m)|^2 + |m_1(H_1,a_1;\alpha_m)|^2 < \infty$, it can be shown that $\mathscr{F}_\delta$ is a $P-$Donsker class. By Lemma 19.24 in Van der Vaart (2000), $\sqrt{n}(\mathbb{P}_n - P)\tilde{f}_{\hat{\beta}_1,\hat{\beta}_2,\hat{\alpha}_m} = \sqrt{n}(\mathbb{P}_n - P)f_{\beta_{10},\beta_{20},\alpha_m^+} + o_p(1)$. Hence we have shown that

$$\begin{aligned}\sqrt{n}(\hat{V}_d - V_d) = {} & \sqrt{n}(\mathbb{P}_n - P)f_{\beta_{10},\beta_{20},\alpha_m^+} \\ & + P\left[\omega_{d_1}(H_1,A_1)\dot{\mu}_2(H_2,d_2(H_2);\beta_{20}) - \dot{\mu}_2(H_2,A_2;\beta_{20})\right]\sqrt{n}(\hat{\beta}_2 - \beta_{20}) \\ & + P\left[\dot{\mu}_1(H_1,d_1(H_1);\beta_{10}) - \dot{\mu}_1(H_1,A_1;\beta_{10})\right]\sqrt{n}(\hat{\beta}_1 - \beta_{10}) + o_p(1).\end{aligned}$$

This combined with the assumption that $\hat{\beta}$ is an asymptotically normal estimator for the parameter β in the SNMM, yields that $\hat{V}_d$ is an asymptotically normal estimator for V_d.

Therefore, the asymptotic variance of $\sqrt{n}\left(\hat{V}_d - V_d\right)$ is equal to $E\big(f_{\beta_{10},\beta_{20},\alpha_m^+}(X_1,A_1,X_2,A_2,Y) + P\left[\dot{\mu}_1(H_1,d_1(H_1);\beta_{10}) - \dot{\mu}_1(H_1,A_1;\beta_{10})\right]\varphi_1 + P\left[\omega_{d_1}(H_1,A_1)\dot{\mu}_2(H_2,d_2(H_2);\beta_{20}) - \dot{\mu}_2(H_2,A_2;\beta_{20})\right]\varphi_2\big)^2$, where φ_1,φ_2 are the influence functions for $\hat{\beta}_1$ and $\hat{\beta}_2$.

Proof of Lemma 4

With arguments similar to part (b) in Lemma 3, under the assumption that $\hat{\beta}$ belongs to the subclass $\mathscr{B}$ of g-estimators, the difference in asymptotic variance between the estimators with function m and without function m is equal to $2Cov(f_{\tilde{d},\beta_0} - f_{d,\beta_0}, g_{m_{\tilde{d}}} - g_{m_d}) + Var(g_{m_{\tilde{d}}} - g_{m_d})$, where $f_{d,\beta}(x_1,a_1,x_2,a_2,y) = \mu_1(h_1,d_1(h_1);\beta_1) + \omega_{d_1}(h_1,a_1)\mu_2(h_2,d_2(h_2);\beta_2)$, and $g_{m_d}(x_1,a_1) = m_d(h_1,d_1(h_1)) - \omega_{d_1}(h_1,a_1)m_d(h_1,a_1)$.

Define $m_d^*(h_1,a_1) \equiv E[\mu_2(H_2,d_2(H_2))|H_1 = h_1, A_1 = a_1]$ and further define $\Delta_d = m_d - m_d^*$. It can be derived that $2Cov(f_{\tilde{d},\beta_0} - f_{d,\beta_0}, g_{m_{\tilde{d}}} - g_{m_d}) + Var(g_{m_{\tilde{d}}} - g_{m_d}) = E\big[\{(\Delta_{\tilde{d}} \circ \tilde{d}_1 - \omega_{\tilde{d}_1}\Delta_{\tilde{d}}) - (\Delta_d \circ d_1 - \omega_{d_1}\Delta_d)\}^2\big] - E\big[\{(m_{\tilde{d}}^* \circ \tilde{d}_1 - \omega_{\tilde{d}_1}m_{\tilde{d}}^*) - (m_d^* \circ d_1 - \omega_{d_1}m_d^*)\}^2\big]$. The second term in the previous formula is not dependent on m_d or $m_{\tilde{d}}$, and thus the lowest asymptotic variance is obtained when $(\Delta_{\tilde{d}} \circ \tilde{d}_1 - \omega_{\tilde{d}_1}\Delta_{\tilde{d}}) = (\Delta_d \circ d_1 - \omega_{d_1}\Delta_d)$, a.s.. The conclusions of the lemma are implied by this equality.

Proof of Lemma 5

Define $\hat{\Delta}(d, \tilde{d}) := \hat{V}_{\hat{m}_{\tilde{d}}}(\tilde{d}; \hat{\beta}) - \hat{V}_{\hat{m}_d}(d; \hat{\beta})$; since each assisted estimator for the value is asymptotically normal, $\hat{\Delta}(d, \tilde{d})$ is also asymptotically normal. For notational simplicity, assume that the treatment effect functions can be modeled as linear in unknown parameters, i.e., $\mu_1(h_1, a_1; \beta_1) = \phi_1(h_1, a_1)^T\beta_1$ and $\mu_2(h_2, a_2; \beta_2) = \phi_2(h_2, a_2)^T\beta_2$, where ϕ_t is some feature of (h_t, a_t). Denote $\Delta(d, \tilde{d}) := V_{\tilde{d}} - V_d$.

We first write the estimated value of m functions for each individual explicitly, assuming that m is a working estimate of $E[\mu_2(H_2, d_2(H_2))|H_1 = h_1, A_1 = a_1]$ obtained from least-squares. This assumption is made only for notational simplicity; in practice, more complicated approach can be taken to estimate m if considered necessary. Denote the predictors that are used to estimate m as $D_m = D_m(H_1, A_1)$, then the fitted value of m function for an individual with $(H_1, A_1) = (h_1, a_1)$ would be equal to:

$$m(h_1, a_1; \hat{\alpha}_m) = D_m(h_1, a_1)^T \left(\mathbb{P}_n D_m D_m^T\right)^{-1} \mathbb{P}_n D_m \phi_2(H_2, d_2(H_2))^T \hat{\beta}_2.$$

To simplify the notation, define $\hat{D} := \mathbb{P}_n D_m D_m^T$, $\hat{Z}_d := \mathbb{P}_n D_m \phi_2(H_2, d_2(H_2))$, then under the specified regularity conditions, we have:

$$\begin{aligned}
&\sqrt{n}(\hat{\Delta}(d, \tilde{d}) - \Delta(d, \tilde{d})) \\
&= \sqrt{n}(\mathbb{P}_n - P) f_{d,\tilde{d},\beta_{10},\beta_{20}} \\
&+ P\left[\phi_1(H_1, \tilde{d}_1(H_1)) - \phi_1(H_1, d_1(H_1))\right]^T \sqrt{n}(\hat{\beta}_1 - \beta_{10}) \\
&+ P\left[\omega_{\tilde{d}_1}(H_1, A_1)\phi_2(H_2, \tilde{d}_2(H_2)) - \omega_{d_1}(H_1, A_1)\phi_2(H_2, d_2(H_2))\right]^T \sqrt{n}(\hat{\beta}_2 - \beta_{20}) \\
&+ \sqrt{n}\mathbb{P}_n g_{\tilde{d}_1}^T P[D_m D_m^T]^{-1} P[D_m \phi_2(H_2, \tilde{d}_2(H_2))^T \beta_{20}] \\
&- \sqrt{n}\mathbb{P}_n g_{d_1}^T P[D_m D_m^T]^{-1} P[D_m \phi_2(H_2, d_2(H_2))^T \beta_{20}] + o_p(1),
\end{aligned}$$

in which $f_{d,\tilde{d},\beta_1,\beta_2}(h_2,a_2) = \Big(\phi_1(h_1,\tilde{d}_1(h_1)) - \phi_1(h_1,d_1(h_1))\Big)^T \beta_1 + \omega_{\tilde{d}_1}(h_1,a_1)\phi_2(h_2,\tilde{d}_2(h_2))^T\beta_2 - \omega_{d_1}(h_1,a_1)\phi_2(h_2,d_2(h_2))^T\beta_2$, and $g_{d_1}(h_1,a_1) = D_m(h_1,d_1(h_1)) - \omega_{d_1}(h_1,a_1)D_m(h_1,a_1)$.

Thus if we denote the influence function of the estimator for parameters in the SNMM by (φ_1,φ_2), namely if $\sqrt{n}(\hat{\beta}_1 - \beta_{10}) = \sqrt{n}\mathbb{P}_n\varphi_1 + o_p(1)$, and $\sqrt{n}(\hat{\beta}_2 - \beta_{20}) = \sqrt{n}\mathbb{P}_n\varphi_2 + o_p(1)$, then the asymptotic variance of $\hat{\Delta}(d,\tilde{d})$ is equal to

$$
\begin{aligned}
\Sigma_\Delta = Var\Big(f_{d,\tilde{d},\beta_{10},\beta_{20}} &+ P\left[\phi_1(H_1,\tilde{d}_1(H_1)) - \phi_1(H_1,d_1(H_1))\right]^T \varphi_1 \\
&+ P\left[\omega_{\tilde{d}_1}(H_1,A_1)\phi_2(H_2,\tilde{d}_2(H_2)) - \omega_{d_1}(H_1,A_1)\phi_2(H_2,d_2(H_2))\right]^T \varphi_2 \\
&+ P[D_m\phi_2(H_2,\tilde{d}_2(H_2))^T\beta_{20}]^T P[D_m D_m^T]^{-1} g_{\tilde{d}_1} \\
&- P[D_m\phi_2(H_2,d_2(H_2))^T\beta_{20}]^T P[D_m D_m^T]^{-1} g_{d_1}\Big).
\end{aligned}
$$

Next we provide the form of the plug-in estimator $\hat{\Sigma}_\Delta$ for Σ_Δ. Suppose we are able to estimate the influence function of $(\hat{\beta}_1,\hat{\beta}_2)$ evaluated at each data point by $(\hat{\varphi}_1,\hat{\varphi}_2) = (\varphi_1(Z;\hat{\beta},\hat{\xi}),\varphi_2(Z;\hat{\beta},\hat{\xi}))$ (Z includes all the observables from one individual; ξ is the nuisance parameter in estimating SNMM). Define $\hat{\Sigma}_\Delta =$

$$
\begin{aligned}
\mathbb{P}_n\Big(f_{d,\tilde{d},\hat{\beta}_1,\hat{\beta}_2} &+ \mathbb{P}_n\left[\phi_1(H_1,\tilde{d}_1(H_1)) - \phi_1(H_1,d_1(H_1))\right]^T \hat{\varphi}_1 \qquad (1) \\
&+ \mathbb{P}_n\left[\omega_{\tilde{d}_1}(H_1,A_1)\phi_2(H_2,\tilde{d}_2(H_2)) - \omega_{d_1}(H_1,A_1)\phi_2(H_2,d_2(H_2))\right]^T \hat{\varphi}_2 \\
&+ \mathbb{P}_n[D_m\phi_2(H_2,\tilde{d}_2(H_2))^T\hat{\beta}_2]^T \mathbb{P}_n[D_m D_m^T]^{-1} g_{\tilde{d}_1} \\
&- \mathbb{P}_n[D_m\phi_2(H_2,d_2(H_2))^T\hat{\beta}_2]^T \mathbb{P}_n[D_m D_m^T]^{-1} g_{d_1}\Big)^2.
\end{aligned}
$$

To show that $\hat{\Sigma}_\Delta$ converges in probability to Σ_Δ, we may use the result that the class of functions involved is a Glivenko-Cantelli class using arguments similar to the proof of Theorem 1.

Web Appendix B: Technical Details

Review: Robins' G-Estimators for SNMM

Here we give a brief review of Robins' class of g-estimating equations (Robins, 1994) and the semiparametric locally efficient g-estimator. Assume that the SNMM is correctly specified. A class of estimating equations which can be used to solve for consistent estimators for β is:
$\mathbb{P}_n \left\{r_1(H_1, A_1)\left(Y - \mu_2(H_2, A_2; \beta_2) - \mu_1(H_1, A_1; \beta_1) - q_1(H_1)\right) + r_2(H_2, A_2)\left(Y - \mu_2(H_2, A_2; \beta_2) - q_2(H_2)\right)\right\}$
$= 0$, where r_1, r_2 are arbitrary functions, both of the same dimension as the length of (β_1^T, β_2^T), that satisfy $E[r_1(H_1, A_1)|H_1] \equiv 0, E[r_2(H_2, A_2)|H_2] \equiv 0$; q_1, q_2 are arbitrary functions.

Assume that $Var(Y-\mu_2(H_2, A_2)-\mu_1(H_1, A_1)|H_1, A_1) \equiv Var(Y-\mu_2(H_2, A_2)-\mu_1(H_1, A_1)|H_1)$, which we will denote as $\sigma_1^2(H_1)$, and that $Var(Y-\mu_2(H_2, A_2)|H_2, A_2) \equiv Var(Y-\mu_2(H_2, A_2)|H_2)$, which we will denote as $\sigma_2^2(H_2)$. Robins provides r_1, r_2, q_1, q_2 functions that make the estimating equation semiparametric locally efficient; in particular the semiparametric locally efficient estimating equation is obtained by setting $q_1^*(h_1) = E[Y-\mu_2(H_2, A_2; \beta_{20})-\mu_1(H_1, A_1; \beta_{10})|H_1 = h_1]$, $q_2^*(h_2) = E[Y - \mu_2(H_2, A_2; \beta_{20})|H_2 = h_2]$,

$$r_1^*(h_1, a_1) = \sigma_1^{-2}(h_1)\begin{pmatrix} \dot{\mu}_1(h_1, a_1; \beta_{10}) - E[\dot{\mu}_1(H_1, A_1; \beta_{10})|H_1 = h_1] \\ E[\dot{\mu}_2(H_2, A_2; \beta_{20})|H_1 = h_1, A_1 = a_1] - E[\dot{\mu}_2(H_2, A_2; \beta_{20})|H_1 = h_1] \end{pmatrix}$$

and

$$r_2^*(h_2, a_2) = \sigma_2^{-2}(h_2)\begin{pmatrix} 0 \\ \dot{\mu}_2(h_2, a_2; \beta_{20}) - E[\dot{\mu}_2(H_2, A_2; \beta_{20})|H_2 = h_2] \end{pmatrix}.$$

Consider models for $r_1(\cdot), r_2(\cdot), q_1(\cdot), q_2(\cdot)$, namely $r_1(\cdot; \eta), r_2(\cdot; \eta), q_1(\cdot; \xi), q_2(\cdot; \xi)$. If the parametric models specified for r_1, r_2, q_1, q_2 contain the truth (i.e., $r_1^*, r_2^*, q_1^*, q_2^*$), the estimator for β is then semiparametric efficient.

Definition of $\mathscr{B}$: The subclass $\mathscr{B}$ of g-estimators is defined as the collection of g-estimators in which $q_1(h_1; \xi)$ is a correctly specified model for $q_1^*(h_1)$. In Lemma 3 in Section 2.1, we show that the optimal m function in the assisted estimator can be identified if $\hat{\beta}$ belongs to this subclass. Note that the semiparametric efficient estimator belongs to this subclass.

Regression-Type Implementation of the G-Estimator

It turns out that for particular models of the nuisance functions (i.e., r_1, r_2, q_1, q_2) in the g-estimating equation, one can estimate both the nuisance functions and the β's simultaneously via least-squares. We use this approach to estimate the β parameters in the intermediate treatment effects in our simulations. We assume that the treatment effect functions are linear in the unknown parameters: $\mu_1(h_1, a_1; \beta_1) = \phi_1(h_1, a_1)^T\beta_1$ and $\mu_2(h_2, a_2; \beta_2) = \phi_2(h_2, a_2)^T\beta_2$, where ϕ_t is some feature of (h_t, a_t). The estimation is as follows:

(1) First solve a linear regression of Y on $(\phi_2(H_2, A_2) - E[\phi_2(H_2, A_2)|H_2], M_2)$, in which M_2 is a summary of the history H_2. Note that in the setting of a randomized trial, the distribution of A_2 is known; thus $E[\phi_2(H_2, A_2)|H_2]$ can be calculated. Put $\hat{\beta}_2$ equal to the vector of the estimated coefficients for $\phi_2(H_2, A_2) - E[\phi_2(H_2, A_2)|H_2]$.

(2) Second solve a linear regression of $Y - \phi_2(H_2, A_2)^T\hat{\beta}_2$ on $(\phi_1(H_1, A_1) - E[\phi_1(H_1, A_1)|H_1], M_1)$, in which M_1 is a summary of the history H_1. Again since the distribution of A_1 is known, $E[\phi_1(H_1, A_1)|H_1]$ can be calculated. Put $\hat{\beta}_1$ equal to the vector of the estimated coefficients for $\phi_1(H_1, A_1) - E[\phi_1(H_1, A_1)|H_1]$.

$\hat{\beta}$ obtained from this least-squares implementation is equivalent to an estimating equation with the following choice of nuisance functions: $r_1(H_1, A_1) = \tilde{\phi}_1(H_1, A_1)$, $r_2(H_2, A_2) = \tilde{\phi}_2(H_2, A_2)$, $q_1(H_1) = M_1^T\kappa_1^+ - E[\phi_1(H_1, A_1)|H_1]^T\beta_{10}$, $q_2(H_2) = M_2^T\kappa_2^+ - E[\phi_2(H_2, A_2)|H_2]^T\beta_{20}$, where $\tilde{\phi}_1 \equiv \tilde{\phi}_1(H_1, A_1) = \phi_1(H_1, A_1) - E[\phi_1(H_1, A_1)|H_1]$ and $\tilde{\phi}_2 \equiv \tilde{\phi}_2(H_2, A_2) = \phi_2(H_2, A_2) - E[\phi_2(H_2, A_2)|H_2]$; κ_1^+ and κ_2^+ denote the probabilistic limits of the estimated coefficients of M_1 and M_2 in the least-squares procedure.

Each member of the class of g-estimators is consistent and asymptotically normal. In particular, the asymptotic distribution of $\sqrt{n}(\hat{\beta} - \beta_0)$ is a multivariate normal with mean zero and var-covariance matrix $B^{-1}\Sigma B^{-1,T}$ where

$$B = \begin{pmatrix} E[\tilde{\phi}_1\tilde{\phi}_1^T] & E[\tilde{\phi}_1\phi_2^T] \\ 0 & E[\tilde{\phi}_2\tilde{\phi}_2^T] \end{pmatrix}$$

and $\Sigma = E\left(\left((Y - \phi_2^T\beta_{20} - \tilde{\phi}_1^T\beta_{10} - M_1^T\kappa_1^+)\tilde{\phi}_1^T, (Y - \tilde{\phi}_2^T\beta_{20} - M_2^T\kappa_2^+)\tilde{\phi}_2^T\right)^T\right)^{\otimes 2}$, where $V^{\otimes 2} = VV^T$. Plug-in estimates $\hat{B}$ and $\hat{\Sigma}$ can be obtained by replacing population expectation in B and Σ with sample mean, and replacing β, κ by the estimates from the series of least squares.

Remark: $\hat{\beta}$ obtained from this least-squares implementation belongs to the subclass $\mathscr{B}$ defined previously, provided that $M_1^T\kappa_1$ is a correct model for $q_1^*(H_1)+E[\phi_1(H_1, A_1)|H_1]^T\beta_{10}$.

Estimator Arising from the Marginal Mean Models

Here we briefly review the estimator of the policy value, based on the marginal mean models (Murphy et al., 2001); we implement this estimator in both simulation studies and data analysis.

Assume that the data take the format of (X_1, A_1, X_2, A_2, Y) and the two-stage policy to be evaluated is $d = (d_1, d_2)$. The marginal-mean-model-based estimator for the policy value requires for modeling and estimating $g_t(\bar{x}_t, \bar{a}_t; \alpha_g)$ ($\bar{Z}_t$ denotes the history of a covariate Z up to time t; $\bar{z}_t$ denotes its realization), where $g_2(\bar{x}_2, \bar{a}_2; \alpha_g)$ models $E[Y|\bar{X}_2 = \bar{x}_2, \bar{A}_2 = \bar{a}_2]$ and $g_1(x_1, a_1; \alpha_g)$ models $E[g_2(X_1, A_1, X_2, d_2(X_1, A_1, X_2))|X_1 = x_1, A_1 = a_1]$.

With an estimator for α_g, the estimator for a two-stage policy, d, arising from the marginal mean model, is written as

$$\begin{aligned}\hat{V}_{MM}(d; \hat{\alpha}_g) =& \mathbb{P}_n\Big\{\frac{I\{A_1 = d_1(X_1)\}}{p_1(A_1|H_1)} \cdot \frac{I\{A_2 = d_2(X_1, A_1, X_2)\}}{p_2(A_2|H_2)} (Y - g_2(X_1, A_1, X_2, A_2; \hat{\alpha}_g)) \\ &+ \frac{I\{A_1 = d_1(X_1)\}}{p_1(A_1|H_1)} (g_2(X_1, A_1, X_2, d_2(X_1, A_1, X_2); \hat{\alpha}_g) - g_1(X_1, A_1; \hat{\alpha}_g)) \\ &+ g_1(X_1, d_1(X_1); \hat{\alpha}_g)\Big\}\end{aligned}$$

The model specification for $g_t(\cdot)$ does not have an impact on the consistency of the estimator $\hat{V}_{MM}(d; \hat{\alpha}_g)$. According to Murphy et al. (2001), to guarantee that the models for

g_t are consistent with each other under the null, we model g_t as linear in $\bar{x}_t$ and independent of $\bar{a}_t$. In particular, we estimate $g_2(X_1, A_1, X_2, A_2; \hat{\alpha}_g)$ by regressing Y on intercept and $\bar{X}_2$, then regress the fitted values on intercept and X_1 to obtain $g_1(X_1, A_1; \hat{\alpha}_g)$.

Equivalence between Estimators in Zhang et al. (2013) and Murphy et al. (2001)

Zhang et al. (2013) presents a robust augmented inverse probability weighted estimator for the values of a restricted class of treatment policies. In their paper the problem of policy value estimation is cast as one of monotone coarsening. With some calculation one can show that the general class of estimators proposed in this paper are equivalent to the estimators arising from the marginal mean model in Murphy et al. (2001). Here we briefly present the equivalence in the case of a two-stage problem (i.e., the data is (X_1, A_1, X_2, A_2, Y), and we are interested in evaluating a two-stage policy $d = (d_1, d_2)$).

For each two-stage policy $d = (d_1, d_2)$, conceptualize the complete data to be the potential outcomes associated with d: $(X_1, X_2(d_1), Y(d_1, d_2))$. Then a coarsening variable C_d can be defined for the complete data as below: If $A_1 \neq d_1(H_1)$, then $C_d = 1$. If $A_1 = d_1(H_1)$ and $A_2 \neq d_2(H_2)$, then $C_d = 2$. If $A_1 = d_1(H_1)$ and $A_2 = d_2(H_2)$, then $C_d = \infty$. Then define the hazard functions for this coarsening variable C_d as follows (coarsening at random is assumed, and in the scenario of sequential randomized trials this assumption is naturally satisfied): $\lambda_{d,1}(X_1) = Pr(C_d = 1|X_1)$, and $\lambda_{d,2}(X_1, X_2) = Pr(C_d = 2|C_d \geqslant 2, X_1, X_2)$. Then the class of estimators (indexed by the functions $L_1(x_1)$ and $L_2(x_1, a_1, x_2)$) proposed in Zhang et al. (2013) can be written as:

$$\mathbb{P}_n\{\frac{I\{C_d = \infty\}}{(1-\lambda_{d,1})(1-\lambda_{d,2})}Y + \frac{I\{C_d = 1\} - \lambda_{d,1}}{1-\lambda_{d,1}}L_1(X_1) + \frac{I\{C_d = 2\} - \lambda_{d,2}I\{C_d \geqslant 2\}}{(1-\lambda_{d,1})(1-\lambda_{d,2})}L_2(X_1, A_1, X_2)\},$$

in which $\lambda_{d,1} = \lambda_{d,1}(X_1), \lambda_{d,2} = \lambda_{d,2}(X_1, X_2)$. The consistency of any estimator in this class is guaranteed, regardless of the choices of L_1, L_2.

On the other hand, the class of estimators (indexed by the functions $g_1(x_1, a_1)$ and $g_2(x_1, a_1, x_2, a_2)$) arising from the marginal mean model can be written in the following form:

$$\begin{aligned}\mathbb{P}_n\Big\{&\frac{I\{A_1 = d_1(X_1)\}}{p_1(A_1|H_1)} \cdot \frac{I\{A_2 = d_2(X_1, A_1, X_2)\}}{p_2(A_2|H_2)} (Y - g_2(X_1, A_1, X_2, A_2)) \\ &+ \frac{I\{A_1 = d_1(X_1)\}}{p_1(A_1|H_1)} (g_2(X_1, A_1, X_2, d_2(X_1, A_1, X_2)) - g_1(X_1, A_1)) \\ &+ g_1(X_1, d_1(X_1))\Big\}.\end{aligned}$$

Equivalency can be established by setting $L_1(X_1) = g_1(X_1, d_1(X_1))$ and $L_2(X_1, A_1, X_2) = g_2(\bar{X}_2, A_1, d_2(\bar{X}_2, A_1))$. Moreover, in Zhang et al. (2013), a practical choice for the functions L_1, L_2 is proposed to be obtained from Q-learning.

Assisted Estimator with Missingness in the Outcome

Real data arising from SMART studies normally contains some missing data, due to participants' dropouts or missing some intermediate treatment sessions or research outcome measurement sesssions for various reasons. In this section we describe an approach to the adjustment of the proposed assisted estimator, in the simplified scenario where only the primary outcome variable Y contains missing values. In particular, this requires that patients do not leave the study before the second randomization.

First we denote our data from each participant as $(X_1, A_1, X_2, A_2, R_\pi, R_\pi Y)$, where R_π is an indicator of whether ($R_\pi = 1$) or not ($R_\pi = 0$) the outcome variable Y is observed for this participant. Let $\pi(h_2, a_2) = Pr[R_\pi = 1|H_2 = h_2, A_2 = a_2]$ be the conditional probability of observing Y given history (h_2, a_2). Estimator for the parameters in SNMM can be obtained following a similar least-squares procedure as the one introduced in the paper:

(1) Generalized linear regression to obtain $\pi(H_2, A_2; \hat{\alpha}_\pi)$ as an estimator for $\pi(H_2, A_2)$.

(2) Weighted linear regression of Y on $(\phi_2(H_2, A_2) - E[\phi_2(H_2, A_2)|H_2], M_2)$ with weights $R_\pi/\pi(H_2, A_2; \hat{\alpha}_\pi)$ (note that only those observations with non-missing Y get non-zero

weights); this regression outputs $\hat{\beta}_2$, which is the vector of the estimated coefficients for $\phi_2(H_2, A_2) - E[\phi_2(H_2, A_2)|H_2]$.

(3) Weighted linear regression of $Y - \phi_2(H_2, A_2)^T\hat{\beta}_2$ on $(\phi_1(H_1, A_1) - E[\phi_1(H_1, A_1)|H_1], M_1)$ with weights $R_\pi/\pi(H_2, A_2; \hat{\alpha}_\pi)$ (again only those observations with non-missing Y get non-zero weights); this regression outputs $\hat{\beta}_1$, which is the vector of the estimated coefficients for $\phi_1(H_1, A_1) - E[\phi_1(H_1, A_1)|H_1]$.

Then one can use the following assisted estimator for the policy value:

$$\hat{V}_m(d;\hat{\beta}) = \mathbb{P}_n\Big\{\frac{R_\pi}{\pi(H_2, A_2; \hat{\alpha}_\pi)}Y - \mu_2(H_2, A_2; \hat{\beta}_2) - \mu_1(H_1, A_1; \hat{\beta}_1) + \mu_1(H_1, d_1(H_1); \hat{\beta}_1)$$
$$+ \frac{I\{A_1 = d_1(H_1)\}}{p_1(A_1|H_1)}\Big(\mu_2(H_2, d_2(H_2); \hat{\beta}_2) - m(H_1, A_1; \hat{\alpha}_m)\Big) + m(H_1, d_1(H_1); \hat{\alpha}_m)\Big\}.$$

Web Appendix C: Details on Simulation Studies

Further Details about the Generative Model in Simulation

Here we provide more details about the generative model used in the simulation experiments. $\eta_0(\cdot), \eta_1(\cdot)$ and the variance of ϵ that we use are all based on the by-products of estimating the SNMM with the ExTENd data, using PACS as the primary outcome. More specifically, $\eta_0(\cdot)$ is the main effect of X_1, and it is set to $\eta_0(X_1) = (1, X_{11}, X_{12}, X_{13}, X_{11}X_{12}, X_{11}X_{13}, X_{12}X_{13}, X_{11}^2, X_{12}^2, X_{13}^2)\alpha_0$ where $\alpha_0 = (11.23, 0.3, 2.28, -0.25, 0.24, 0.73, 0.3, -0.74, -0.53, -0.47)$. $\eta_1(\cdot)$ is the main effect of X_2 conditional on (X_1, A_1), and it is set to $\eta_1(X_1, A_1, X_2) = 2(X_{21} - E[X_{21}|X_1, A_1]) - 2(X_{22} - E[X_{22}|X_1, A_1])$. The standard deviation of ϵ is set to be 5.54.

Additional Results from Simulation 1

Here we present the simulation results for Simulation 1 (concerning the basic statistical properties of the assisted estimators) in the manuscript, when $N = 250$ (Table 1). Recall that these experiments are conducted using an estimator $\hat{\beta}$ that belongs to the subclass $\mathscr{B}$. We found that, in general, the oracle estimator and the two assisted estimators have similar MSE; when the treatment effect at stage two is at medium level, the assisted estimator with a working estimate for the optimal m_d has slightly lower MSE than the assisted estimator with $m_d = 0$. Also similar to the $N = 100$ experiments, the confidence intervals based on the asymptotic standard error show good coverage.

[Table 1 about here.]

Moreover, we conduct the same set of simulation experiments using an estimator for β that does not belong to the subclass $\mathscr{B}$ (simulation 1*); that is, the particular nuisance function referred to in the definition of $\mathscr{B}$ is not correctly modeled (in fact, the true nuisance function includes linear terms and second-order terms of X_1; in this simulation, in the estimation of β we only model the nuisance function by the linear terms). Our conjecture is that the assisted estimator $\hat{V}_{\hat{m}_d}(d;\hat{\beta})$ with $\hat{m}_d$ is still slightly more efficient than the assisted estimator with $m_d \equiv 0$. Since $\hat{\beta}$ no longer belongs to $\mathscr{B}$, we do not compare the assisted estimators with an "oracle" estimator. Results are shown in Table 2. We found that, as expected, the resulting assisted estimators are unbiased. The MSEs of the two different assisted estimators are similar; yet the one using a good working estimate $\hat{m}_d$ seems to be slightly more efficient in some cases. In general, the results are very similar to the results from the experiments using a $\hat{\beta}$ that belongs to the subclass $\mathscr{B}$.

[Table 2 about here.]

Simulation of the Relative Efficiency of Assisted Estimators

In this section we further investigate the extent to which the assisted estimator with a working estimate of the optimal m_d improves efficiency over the assisted estimator with $m_d = 0$. We apply the two types of assisted estimators to estimate each of the two policy contrasts: (1) contrast between embedded policies (1, 1, 1) and (0, 0, 0); (2) contrast between embedded policies (1, 1, 0) and (0, 0, 0). Motivated by the remark in the proof of Lemma 3 about the magnitude of the achievable variance reduction by adopting a good choice of m_d, the experiments are conducted with data from a series of generative models, in which the standardized effect size (SES) of the coordinate in β_2 that corresponds to the A_2 main effect varies from 0.0 to 3.0, and all the other coordinates in β_1 and β_2 have an SES equal to 0.2. We focus on the relative mean squared errors of the assisted estimator with a working estimate of the optimal m_d as compared to that with $m_d = 0$, and for both estimands we plot the trend of the relative mean squared error as the A_2 main effect grows.

[Figure 1 about here.]

The simulation results are shown in Figure 1. As expected, the benefit of using a working estimate m_d in the assisted estimator increases when the stage two treatment effect amplifies.

However, under the generative model we consider, the A_2 main effect needs to be as large as having an effect size of 1.5 so that the efficiency improvement is about 20%. In practice, we suspect whether such a huge treatment effect would ever be present in a SMART; thus in general using $m_d = 0$ in the assisted estimator may perform just as well as the assisted estimator with a working estimate of m_d. We also notice that, the extent to which using a working estimate of m_d is more efficient than using $m_d = 0$ varies with the estimand.

Additional Results from Simulation 2

Here we present the simulation results for Simulation 2 (concerning the efficiency gain of using the assisted estimator compared with the marginal mean model) in the manuscript, when $N = 250$ (Table 3). We found that with a larger sample size ($N = 250$ as compared to $N = 100$), the advantage of the assisted estimators in terms of having a lower MSE than the marginal-mean-model-based estimators is more evident. Similar to the $N = 100$ experiments, mis-specifying the SNMM introduces bias in some scenarios, but even in those scenarios the performance of the assisted estimators in terms of the MSE does not worsen, because reduction in the variance dominates the bias-variance tradeoff. We notice that under the most severe mis-specification of SNMM (Assist3), the confidence interval of the contrast between the tailored policy and the policy $(0, 0, 0)$ has noticeable under-coverage. However, we expect that in practice, such severe mis-specification, which fails to use any variable correlated with the variables in the true SNMM, might be unlikely to happen.

[Table 3 about here.]

References

Murphy, S. A., van der Laan, M. J., and Robins, J. M. (2001). Marginal mean models for dynamic regimes. *Journal of the American Statistical Association* **96,** 1410–1423.

Robins, J. M. (1994). Correcting for non-compliance in randomized trials using structural nested mean models. *Communications in Statistics-Theory and methods* **23,** 2379–2412.

Van der Vaart, A. W. (2000). *Asymptotic statistics*, volume 3. Cambridge university press.

Zhang, B., Tsiatis, A. A., Laber, E. B., and Davidian, M. (2013). Robust estimation of optimal dynamic treatment regimes for sequential treatment decisions. *Biometrika* **100,** 681–694.

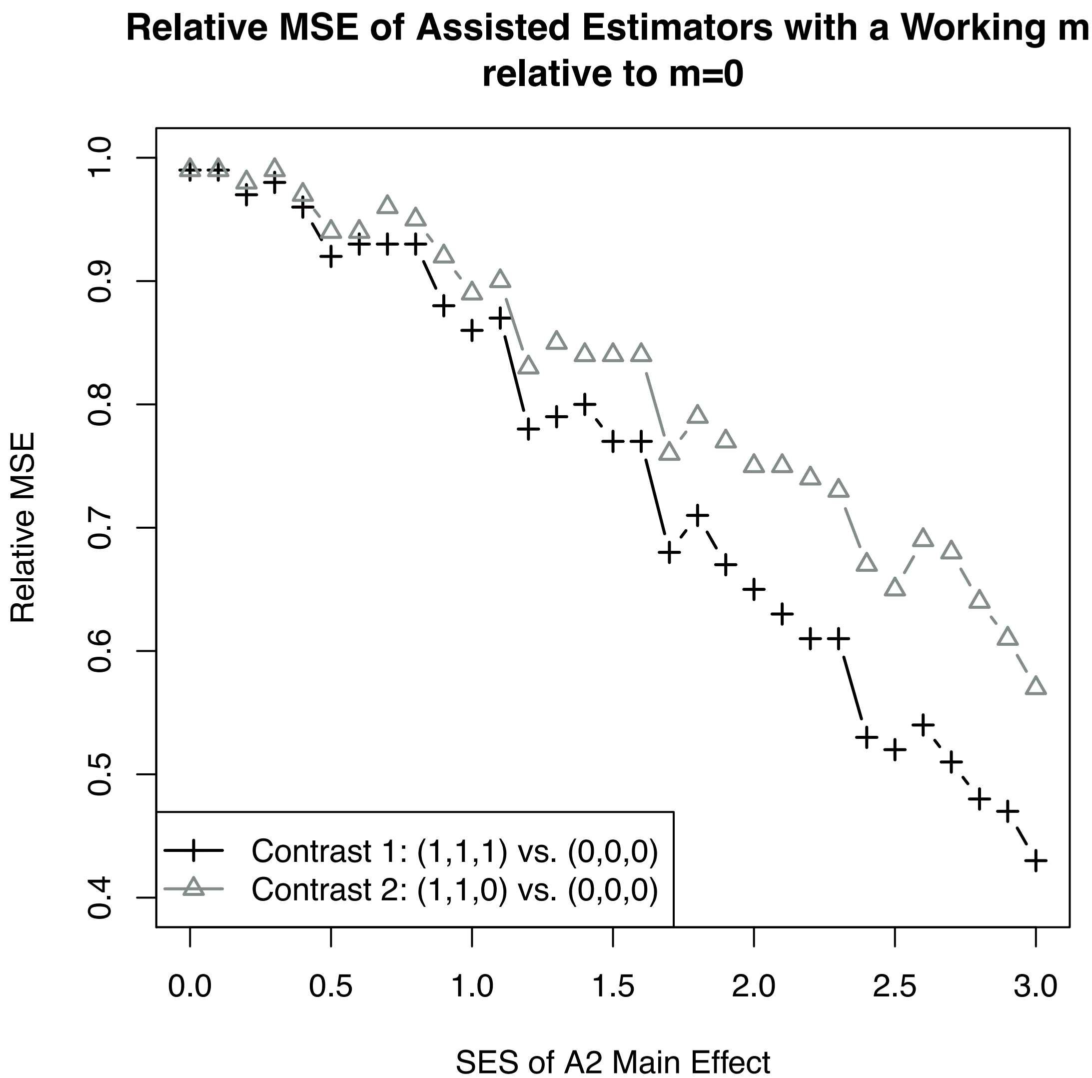

Figure 1. Relative mean squared error of the two assisted estimators, as a function of the SES of A_2 main effect in the generative model.

Table 1

Simulation 1: Statistical properties of the assisted estimators of the contrast between values of policies (1,1,1) and (0,0,0). Oracle = contrast estimator based on $\hat{V}_{m_d}(d;\hat{\beta})$ with the true optimal m_d. Assist = contrast estimator based on $\hat{V}_{\hat{m}_d}(d;\hat{\beta})$ with a working estimate of the optimal m_d. Assist ($m_d = 0$) = contrast estimator based on $\hat{V}_0(d;\hat{\beta})$. The displayed numbers for confidence interval coverage are the coverage proportion $\times$ 100. An Asterisk indicates that the MSE of Oracle or Assist ($m_d = 0$) is significantly different from MSE of Assist (at 0.05 level).

$N = 250$								
Scenario	Bias / SD			MSE			ASE Coverage	
	Oracle	Assist	Assist ($m_d = 0$)	Oracle	Assist	Assist ($m_d = 0$)	Assist	Assist ($m_d = 0$)
(none,none)	0	0	0	1.3	1.31	1.3	95	95
(none,low)	0.03	0.03	0.03	1.44	1.45	1.47*	95.1	95.1
(none,med)	0.03	0.02	0.02	1.4	1.4	1.48*	94.6	95.6
(low,none)	-0.01	-0.01	-0.01	1.31	1.32	1.31	95.7	95.6
(low,low)	-0.02	-0.02	-0.02	1.69	1.71	1.71	93.2	93.6
(low,med)	0	-0.01	0	1.42	1.42	1.54*	95.3	95.2
(med,none)	-0.03	-0.03	-0.03	1.38	1.38	1.38	95.2	95.1
(med,low)	0	0	0.01	1.64	1.64	1.67	94.8	95
(med,med)	-0.02	-0.02	-0.01	1.55	1.55	1.63*	94.8	94.7

Table 2

Simulation 1: Statistical properties of the assisted estimators of the contrast between values of policies (1,1,1) and (0,0,0), when $\hat{\beta}$ does not belong to $\mathscr{B}$. Assist = contrast estimator based on $\hat{V}_{\hat{m}_d}(d;\hat{\beta})$ with a working estimate of the optimal m_d. Assist ($m_d = 0$) = contrast estimator based on $\hat{V}_0(d;\hat{\beta})$. The displayed numbers for confidence interval coverage are the coverage proportion × 100. An Asterisk indicates that the MSE of Assist ($m_d = 0$) is significantly different from MSE of Assist (at 0.05 level).*

$N = 100$

Scenario	Bias / SD		MSE		ASE Coverage	
	Assist	Assist ($m_d = 0$)	Assist	Assist ($m_d = 0$)	Assist	Assist ($m_d = 0$)
(none,none)	0.04	0.04	3.48	3.54*	95.1	95.2
(none,low)	0.02	0.01	4.33	4.41	94.3	94.6
(none,med)	0.03	0.01	3.89	4.28*	95.6	95.5
(low,none)	-0.02	-0.02	3.38	3.39	95	95.4
(low,low)	0.01	0.01	4.15	4.14	95	95.6
(low,med)	0.03	0.02	3.97	4.13*	95.3	95.6
(med,none)	0.05	0.05	3.93	3.98	95.2	95
(med,low)	-0.02	-0.02	4.42	4.43	94.9	94.7
(med,med)	0	0	4.04	4.25*	94.8	95.5

$N = 250$

Scenario	Bias / SD		MSE		ASE Coverage	
	Assist	Assist ($m_d = 0$)	Assist	Assist ($m_d = 0$)	Assist	Assist ($m_d = 0$)
(none,none)	0.01	0.01	1.36	1.36	93.8	94
(none,low)	0.01	0.02	1.51	1.53*	95.4	95.6
(none,med)	0.03	0.03	1.44	1.53*	94.7	95.2
(low,none)	-0.01	-0.01	1.35	1.35	95.9	95.9
(low,low)	-0.01	-0.01	1.73	1.73	94	94.1
(low,med)	-0.01	-0.01	1.45	1.57*	95.4	94.8
(med,none)	-0.03	-0.03	1.47	1.47	94.9	94.8
(med,low)	0	0.01	1.66	1.68	94.8	95.3
(med,med)	-0.02	-0.01	1.53	1.61*	94.7	95.1

Table 3

Simulation 2: Comparison between the marginal-mean-model-based estimators and the assisted estimators, with respect to the performance in estimating the policy contrasts. $\hat{\beta}$ used in the assisted estimators belongs to the subclass $\mathscr{B}$. MM = Marginal-mean-model-based estimator. Assist1 = Assisted estimator with correctly specified SNMM. Assist2 = Assisted estimator with mis-specified SNMM that excludes X_{11}, X_{21}, RX_{21}. Assist3 = Assisted estimator with mis-specified SNMM that excludes all the covariates interacting with treatments. Bias significantly different from 0, and coverage proportion significantly different from 95%, are marked with an asterisk. Relative MSE is calculated as the ratio of MSE with that of MM.

$N = 250$

Estimation of the first contrast,(1, 1, 1) vs (0,0,0)

Scenario	Bias x 100				Coverage of 95% CI x 100				Relative MSE		
	MM	Assist1	Assist2	Assist3	MM	Assist1	Assist2	Assist3	Assist1	Assist2	Assist3
(none,none)	2.6	4.7	4.4	4.8	93.5*	94.6	94.4	94.4	0.87	0.86	0.89
(none,low)	2	1.5	1.6	2.5	93.8	94.7	94.5	94.6	0.82	0.81	0.87
(none,med)	7.2	-1.2	-1.4	0.3	94.6	94.7	94.9	94.9	0.82	0.83	0.85
(low,none)	-4.6	-2.4	-3.3	-3.8	95.2	95.1	95	95.3	0.83	0.83	0.86
(low,low)	-5.1	-6	-5.8	-6.4	94.5	93.9	93.6*	93.8	0.87	0.87	0.89
(low,med)	6	-0.3	-0.5	1.3	96	95.4	95.4	95.8	0.79	0.8	0.84
(med,none)	-2.3	-1.3	-1.8	-1.1	94.5	94.3	94.3	95.9	0.75	0.76	0.78
(med,low)	9.3*	7.6	7.9	8.1	94.4	94.5	94.3	94.1	0.79	0.79	0.8
(med,med)	20.6*	11.2*	10.6*	14.3*	94.5	94.5	93.7	94.2	0.73	0.74	0.78

Estimation of the second contrast, the tailored policy vs (0,0,0)

Scenario	Bias x 100				Coverage of 95% CI x 100				Relative MSE		
	MM	Assist1	Assist2	Assist3	MM	Assist1	Assist2	Assist3	Assist1	Assist2	Assist3
(none,none)	-0.8	0.2	-0.2	2.3	95.1	93.6*	93.3*	95.4	0.69	0.67	0.48
(none,low)	0.2	0.5	-7.3*	13.2*	93.6*	94.7	94.3	94.9	0.65	0.64	0.46
(none,med)	6.7	0.9	-20*	-39.5*	93.8	94.6	93.7	92.8*	0.69	0.71	0.58
(low,none)	-1.6	0.1	-1	38.5*	95.1	95.1	95.1	92.4*	0.67	0.66	0.6
(low,low)	-3.8	-1.5	-9.2*	49.1*	95.5	94.8	94.9	91.5*	0.68	0.68	0.62
(low,med)	0.7	-0.5	-21.9*	-0.7	95.1	95.2	94.7	96	0.68	0.7	0.45
(med,none)	2	3.6	2.4	46.9*	95.3	94.9	95.4	91.1*	0.6	0.59	0.55
(med,low)	10.4*	7.3*	-0.5	63.7*	94.4	94.6	94.8	88.8*	0.62	0.6	0.65
(med,med)	8.7*	6.7	-15*	15.3*	94.7	95.4	94.9	94.1	0.64	0.62	0.46