A. Proofs

A.1. Proof of Lemma 1

In Step 1 the following expressions for each component of $E\{U(\beta^*,\gamma^*)\}$ are derived for $\beta^* = (\beta_1, \beta_2, \beta_3^*)$. In Step 2 we consider the behaviour of these expressions in each of the two misspecification cases.

$$E\{U_1(\beta^*,\gamma^*)\} = E\left[\{h(Z) - h(Z;\gamma_x^*)\}\{f(Z) - f(Z;\gamma_m^*)\}\right] \tag{25}$$

$$E\{U_2(\beta^*,\gamma^*)\} = E\left[\{f(Z) - f(Z;\gamma_m^*)\}\left\{g(X,Z) - \beta_3^* X - g(Z;\gamma_y^*)\right\}\right] \tag{26}$$

$$E\{U_3(\beta^*,\gamma^*)\} = E\left[\{X - h(Z;\gamma_x^*)\}\left\{g(X,Z) - \beta_3^* X - g(Z;\gamma_y^*)\right\}\right] \tag{27}$$

Step 1: For the first component we use the partial linearity to obtain

$$\begin{aligned} E\{U_1(\beta^*,\gamma^*)|X,Z\} &= \Big\{X - h(Z;\gamma_x^*)\Big\}\Big\{E(M - \beta_1 X|X,Z) - f(Z;\gamma_m^*)\Big\} \\ &= \Big\{X - h(Z;\gamma_x^*)\Big\}\Big\{f(Z) - f(Z;\gamma_m^*)\Big\} \\ E\{U_1(\beta^*,\gamma^*)|Z\} &= \Big\{h(Z) - h(Z;\gamma_x^*)\Big\}\Big\{f(Z) - f(Z;\gamma_m^*)\Big\} \end{aligned}$$

Similarly for the second component,

$$\begin{aligned} E\{U_2(\beta^*,\gamma^*)|M,X,Z\} &= \Big\{M - \beta_1 X - f(Z;\gamma_m^*)\Big\}\Big\{E(Y - \beta_2 M|M,X,Z) - \beta_3^* X - g(Z;\gamma_y^*)\Big\} \\ E\{U_2(\beta^*,\gamma^*)|X,Z\} &= \Big\{E(M - \beta_1 X|X,Z) - f(Z;\gamma_m^*)\Big\}\Big\{g(X,Z) - \beta_3^* X - g(Z;\gamma_y^*)\Big\} \\ &= \Big\{f(Z) - f(Z;\gamma_m^*)\Big\}\Big\{g(X,Z) - \beta_3^* X - g(Z;\gamma_y^*)\Big\} \end{aligned}$$

Finally for the third component,

$$\begin{aligned} E\{U_3(\beta^*,\gamma^*)|M,X,Z\} &= \Big\{X - h(Z;\gamma_x^*)\Big\}\Big\{E(Y - \beta_2 M|M,X,Z) - \beta_3^* X - g(Z;\gamma_y^*)\Big\} \\ E\{U_3(\beta^*,\gamma^*)|X,Z\} &= \Big\{X - h(Z;\gamma_x^*)\Big\}\Big\{g(X,Z) - \beta_3^* X - g(Z;\gamma_y^*)\Big\} \end{aligned}$$

Step 2: We shall consider the cases (i) and (ii) separately. In case (i) the conditions for assumption A2 are met, hence $f(Z) = f(Z;\gamma_m^*)$ and so (25) and (26) are exactly zero. The proof for case (i) is completed by letting β_3^* be the value which solves (27) equal to zero.

For case (ii) the conditions of A1 are met, hence $h(Z) = h(Z;\gamma_x^*)$ so (25) is exactly zero. Also there exists β_3 such that $g(X,Z) = \beta_3 X + g(Z)$ and for $\beta_3^* = \beta_3$ then the conditions in A3 are met so $g(Z) = g(Z;\gamma_y^*)$ and hence (26) and (27) are exactly zero, which completes the proof for case (ii).

A.2. Proof of Lemma 2

In Step 1 the following expressions for each component of $E\{U(\beta^*, \gamma^*)\}$ are derived for $\beta^* = (\beta_1^*, \beta_2, \beta_3)$. In Step 2 we consider the behaviour of these expressions in each of the two misspecification cases.

$$E\{U_1(\beta^*, \gamma^*)\} = E\left[\left\{X - h(Z; \gamma_x^*)\right\}\left\{f(X, Z) - \beta_1^* X - f(Z; \gamma_m^*)\right\}\right] \tag{28}$$

$$E\{U_2(\beta^*, \gamma^*)\} = E\left[\left\{f(X, Z) - \beta_1^* X - f(Z; \gamma_m^*)\right\}\left\{g(Z) - g(Z; \gamma_y^*)\right\}\right] \tag{29}$$

$$E\{U_3(\beta^*, \gamma^*)\} = E\left[\left\{h(Z) - h(Z; \gamma_x^*)\right\}\left\{g(Z) - g(Z; \gamma_y^*)\right\}\right] \tag{30}$$

Step 1. For the first component,

$$E\{U_1(\beta^*, \gamma^*)|X, Z\} = \left\{X - h(Z; \gamma_x^*)\right\}\left\{f(X, Z) - \beta_1^* X - f(Z; \gamma_m^*)\right\}$$

For the second component we use the partial linearity to obtain

$$\begin{aligned} E\{U_2(\beta^*, \gamma^*)|M, X, Z\} &= \left\{M - \beta_1^* X - f(Z; \gamma_m^*)\right\}\left\{E(Y - \beta_2 M - \beta_3 X|M, X, Z) - g(Z; \gamma_y^*)\right\} \\ &= \left\{M - \beta_1^* X - f(Z; \gamma_m^*)\right\}\left\{g(Z) - g(Z; \gamma_y^*)\right\} \\ E\{U_2(\beta^*, \gamma^*)|X, Z\} &= \left\{f(X, Z) - \beta_1^* X - f(Z; \gamma_m^*)\right\}\left\{g(Z) - g(Z; \gamma_y^*)\right\} \end{aligned}$$

Similarly for the third component,

$$\begin{aligned} E\{U_3(\beta^*, \gamma^*)|M, X, Z\} &= \left\{X - h(Z; \gamma_x^*)\right\}\left\{E(Y - \beta_2 M - \beta_3 X|M, X, Z) - g(Z; \gamma_y^*)\right\} \\ &= \left\{X - h(Z; \gamma_x^*)\right\}\left\{g(Z) - g(Z; \gamma_y^*)\right\} \\ E\{U_3(\beta^*, \gamma^*)|Z\} &= \left\{h(Z) - h(Z; \gamma_x^*)\right\}\left\{g(Z) - g(Z; \gamma_y^*)\right\} \end{aligned}$$

Step 2. We shall consider the cases (i) and (ii) separately. In case (i) the conditions for assumption A3 are met, hence $g(Z) = g(Z; \gamma_y^*)$ so (29) and (30) are exactly zero. Letting β_1^* be the value which solves (28) equal to zero completes the proof for case (i).

For case (ii) the conditions of A1 are met, so $h(Z) = h(Z; \gamma_x^*)$ and so (30) is zero. Also there exists β_1 such that $f(X, Z) = \beta_1 X + f(Z)$ and for $\beta_1^* = \beta_1$ then the conditions in A2 are met so $f(Z) = f(Z; \gamma_y^*)$ and hence (28) and (29) are exactly zero, which completes the proof for case (ii).

A.3. Proof of Theorem 1

Here we provide a sketch of the proof. Consider the Taylor Expansion

$$E_n\{U(\hat{\beta},\hat{\gamma})\} = E_n\{U(\beta^*,\gamma^*)\} + E_n\left\{\frac{\partial U(\beta^*,\gamma^*)}{\partial \beta}\right\}(\hat{\beta}-\beta^*) + E_n\left\{\frac{\partial U(\beta^*,\gamma^*)}{\partial \gamma}\right\}(\hat{\gamma}-\gamma^*) + o_p\left(n^{-1/2}\right)$$

Since $E_n\{U(\hat{\beta},\hat{\gamma})\} = 0$ then

$$\hat{\beta}-\beta^* = E_n\left\{-\frac{\partial U(\beta^*,\gamma^*)}{\partial \beta}\right\}^{-1}\left[E_n\{U(\beta^*,\gamma^*)\} + E_n\left\{\frac{\partial U(\beta^*,\gamma^*)}{\partial \gamma}\right\}(\hat{\gamma}-\gamma^*)\right] + o_p\left(n^{-1/2}\right)$$

Using the estimator in (10) and rearranging gives

$$\hat{\beta}-\beta^* = E_n\left(E_n\left\{-\frac{\partial U(\beta^*,\gamma^*)}{\partial \beta}\right\}^{-1}\left[U(\beta^*,\gamma^*) + E_n\left\{\frac{\partial U(\beta^*,\gamma^*)}{\partial \gamma}\right\}\phi(\beta^*,\gamma^*)\right]\right) + o_p\left(n^{-1/2}\right)$$

Applying the weak law of large numbers to the partial derivative terms gives the form of the influence function $\varphi(.)$ in (11). We must further show that $E\{\varphi(\beta^*,\gamma^*)\} = 0$

$$E\{\varphi(\beta^*,\gamma^*)\} = E\left\{-\frac{\partial U(\beta^*,\gamma^*)}{\partial \beta}\right\}^{-1}\left[E\{U(\beta^*,\gamma^*)\} + E\left\{\frac{\partial U(\beta^*,\gamma^*)}{\partial \gamma}\right\}E\{\phi(\beta^*,\gamma^*)\}\right]$$

Since $\phi(.)$ is an influence function, $E\{\phi(\beta^*,\gamma^*)\} = 0$. Therefore provided $E\{U(\beta^*,\gamma^*)\} = 0$ then $E\{\varphi(\beta^*,\gamma^*)\} = 0$ as required.

A.4. Derivation of Equation (12)

By Theorem 1,

$$\hat{\beta}_1 = \beta_1 + E_n\{\varphi_1(\beta^*,\gamma^*)\} + o_p\left(n^{-1/2}\right)$$
$$\hat{\beta}_2 = \beta_2 + E_n\{\varphi_2(\beta^*,\gamma^*)\} + o_p\left(n^{-1/2}\right)$$

Therefore, letting $A = E_n\{\varphi_1(\beta^*,\gamma^*)\}$ and $B = E_n\{\varphi_2(\beta^*,\gamma^*)\}$,

$$\hat{\beta}_1\hat{\beta}_2 - \beta_1\beta_2 = E_n\{\omega(\beta^*,\gamma^*)\} + AB + o_p\left(n^{-1/2}\right)$$

and the desired result follows provided that $AB = o_p\left(n^{-1/2}\right)$. Using Markov's inequality,

$$P(|n^{1/2}AB| \geq \epsilon) = P(n(AB)^2 \geq \epsilon^2) \leq \frac{nE\{(AB)^2\}}{\epsilon^2}$$

Examining the expectation term, we find a sum over four indices

$$E\{(AB)^2\} = n^{-4}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n} E\{\varphi_1^{(i)}(\beta^*,\gamma^*)\varphi_1^{(j)}(\beta^*,\gamma^*)\varphi_2^{(k)}(\beta^*,\gamma^*)\varphi_2^{(l)}(\beta^*,\gamma^*)\}$$

where the superscript (i) denotes that the influence function is evaluated on the ith observation. Since the observations are iid and the influence function has mean zero, the terms of this quadruple sum can only be non-zero when their indices are paired, i.e when ($i = j$ and $k = l$) or ($i = k$ and $j = l$) or ($i = l$ and $j = k$). The number of non-zero terms in the sum is therefore of order n^2, and hence

$$P(|n^{1/2}AB| \geq \epsilon) \leq \mathcal{O}\left(n^{-1}\right)$$

where $\mathcal{O}$ denotes conventional big-O notation, i.e. for sufficiently large n there exists some constant k such that $|\mathcal{O}\left(n^{-1}\right)| \leq kn^{-1}$

A.5. Proof of Theorem 2

Here we adapt the proof from Section 5.1 of Dufour et al. (2017) to allow for orthogonal nuisance parameter estimation. We prove the results for the CUE estimator, however they are equally applicable to the two-step estimator. Our extension to the original results relies on three orthogonality-like derivative results for the test statistic of interest. These are derived assuming that the nuisance parameter estimator is orthogonal to the moment conditions in the sense that (13) holds. This may either be because all models are correctly specified or because a bias-reduced strategy is used to estimate nuisance parameters, as described below. We begin by defining the CUE objective function, which we denote by M_n as in the original notation of Dufour et al. (2017),

$$M_n(\beta,\gamma) = D_n^{\top}(\beta,\gamma)I_n^{-1}(\beta,\gamma)D_n(\beta,\gamma)$$

where, for a target parameter moment function $U(\beta,\gamma)$,

$$\begin{aligned} D_n(\beta,\gamma) &= E_n[U(\beta,\gamma)] \\ C_n(\beta,\gamma) &= E_n\left[\frac{\partial U(\beta,\gamma)}{\partial\gamma}\right] = \frac{\partial D_n(\beta,\gamma)}{\partial\gamma} \\ I_n(\beta,\gamma) &= E_n[U(\beta,\gamma)U(\beta,\gamma)^{\top}] \end{aligned}$$

Theorem 2 in the text considers the exactly specified setting, i.e. $\text{Dim}(\beta) = \text{Dim}(U(\beta, \gamma))$. For the current proof, however, we consider the over-specified setting, i.e. $\text{Dim}(\beta) \leq \text{Dim}(U(\beta, \gamma))$. Also define the probability limits, β^*, γ^* as the (assumed to be) unique values such that

$$\frac{\partial M_n(\beta^*, \gamma^*)}{\partial \beta} \xrightarrow{p} 0$$
$$C_n(\beta^*, \gamma^*) \xrightarrow{p} 0$$

The first of these is equivalent to $D_n(\beta^*, \gamma^*) \xrightarrow{p} 0$ in the exactly specified setting. By the central limit theorem,

$$\sqrt{n} D_n(\beta^*, \gamma^*) \xrightarrow{d} \mathcal{N}(0, I_0)$$
$$I_n(\beta^*, \gamma^*) \xrightarrow{p} I_0 = E[U(\beta^*, \gamma^*) U(\beta^*, \gamma^*)^\top]$$

The unconstrained estimated values $\hat{\beta}, \hat{\gamma}$ are those which solve

$$\frac{\partial M_n(\hat{\beta}, \hat{\gamma})}{\partial \beta} = 0 \tag{31}$$
$$C_n(\hat{\beta}, \hat{\gamma}) = 0$$

Again, the first of these is equivalent to $D_n(\hat{\beta}, \hat{\gamma}) = 0$ in the exactly specified setting. The constrained estimated values $\hat{\beta}_\psi, \hat{\gamma}_\psi$ are those which solve

$$\frac{\partial M_n(\hat{\beta}_\psi, \hat{\gamma}_\psi)}{\partial \beta} - \frac{\partial \psi(\hat{\beta}_\psi)}{\partial \beta} \lambda = 0 \tag{32}$$
$$C_n(\hat{\beta}_\psi, \hat{\gamma}_\psi) = 0$$
$$\psi(\hat{\beta}_\psi) = 0 \tag{33}$$

for a constraint function ψ and where λ is a Lagrange multiplier. The statement that we intend to prove is that

$$n[M_n(\hat{\beta}_\psi, \hat{\gamma}_\psi) - M_n(\hat{\beta}, \hat{\gamma})] \xrightarrow{d} \chi^2_r \tag{34}$$

where r is the rank of $\partial \psi(\beta)/\partial \beta$ in a neighbourhood of β^*. In the exactly specified setting, $M_n(\hat{\beta}, \hat{\gamma}) = 0$.

Three necessary derivative results

In this subsection we show that, since $C_n(\beta^*,\gamma^*) = o_p(1), D_n(\beta^*,\gamma^*) = o_p(1)$, and $\sqrt{n}D_n(\beta^*,\gamma^*) = O_p(1)$ then

$$\sqrt{n}\frac{\partial M_n(\beta^*,\gamma^*)}{\partial\gamma} = o_p(1) \tag{35}$$
$$\frac{\partial^2 M_n(\beta^*,\gamma^*)}{\partial\gamma\partial\beta} = o_p(1) \tag{36}$$
$$\frac{\partial^2 M_n(\beta^*,\gamma^*)}{\partial\gamma\partial\gamma} = o_p(1) \tag{37}$$

To do so it is easier to work in an index notation where D_i is the ith component of $D_n(\beta,\gamma)$, and I_{ij}^{-1} is the i,jth component of $I_n^{-1}(\beta,\gamma)$ and all quantities are evaluated at $(\beta,\gamma) = (\beta^*,\gamma^*)$. For example, letting $q = \text{Dim}(U(\beta,\gamma))$, then

$$M_n(\beta^*,\gamma^*) = \sum_{i=1}^{q}\sum_{j=1}^{q} D_i I_{ij}^{-1} D_j$$

For the first derivative term of interest,

$$\sqrt{n}\frac{\partial M_n(\beta^*,\gamma^*)}{\partial\gamma} = \sum_{i=1}^{q}\sum_{j=1}^{q}\left\{\left(2\frac{\partial D_i}{\partial\gamma}I_{ij}^{-1} + D_i\frac{\partial I_{ij}^{-1}}{\partial\gamma}\right)\sqrt{n}D_j\right\}$$

and for the second derivative term of interest,

$$\frac{\partial^2 M_n(\beta^*,\gamma^*)}{\partial\gamma\partial\beta} = \sum_{i=1}^{q}\sum_{j=1}^{q}\left\{2\frac{\partial D_i}{\partial\beta}I_{ij}^{-1}\frac{\partial D_j}{\partial\gamma} + \left(2\frac{\partial^2 D_i}{\partial\gamma\partial\beta}I_{ij}^{-1} + 2\frac{\partial D_i}{\partial\beta}\frac{\partial I_{ij}^{-1}}{\partial\gamma} + 2\frac{\partial D_i}{\partial\gamma}\frac{\partial I_{ij}^{-1}}{\partial\beta} + D_i\frac{\partial^2 I_{ij}^{-1}}{\partial\gamma\partial\beta}\right)D_j\right\}$$

For the third,

$$\frac{\partial^2 M_n(\beta^*,\gamma^*)}{\partial\gamma\partial\gamma} = \sum_{i=1}^{q}\sum_{j=1}^{q}\left\{2\frac{\partial D_i}{\partial\gamma}I_{ij}^{-1}\frac{\partial D_j}{\partial\gamma} + \left(2\frac{\partial^2 D_i}{\partial\gamma\partial\gamma}I_{ij}^{-1} + 2\frac{\partial D_i}{\partial\beta}\frac{\partial I_{ij}^{-1}}{\partial\gamma} + 2\frac{\partial D_i}{\partial\gamma}\frac{\partial I_{ij}^{-1}}{\partial\gamma} + D_i\frac{\partial^2 I_{ij}^{-1}}{\partial\gamma\partial\gamma}\right)D_j\right\}$$

By the orthogonality of the nuisance parameter estimator, $\partial D_j/\partial\gamma = o_p(1)$, and since $D_j = o_p(1)$, and $\sqrt{n}D_j = O_p(1)$ then the results follow.

Applying the derivative results

Consider the test statistic

$$\xi = M_n(\hat\beta_\psi,\hat\gamma_\psi) - M_n(\hat\beta,\hat\gamma)$$

Under standard regularity conditions, and the rank condition in (20) (see Dufour et al. (2017) for details), $\hat{\gamma}, \hat{\gamma}_\psi, \hat{\beta}$ and $\hat{\beta}_\psi$ are CAN, hence expanding this test statistics to second order gives

$$\begin{aligned}
n\xi &= \sqrt{n}\frac{\partial M_n(\beta^*,\gamma^*)}{\partial \beta^\top}\left\{\sqrt{n}(\hat{\beta}_\psi - \beta^*) - \sqrt{n}(\hat{\beta} - \beta^*)\right\} \\
&+ \sqrt{n}\frac{\partial M_n(\beta^*,\gamma^*)}{\partial \gamma^\top}\left\{\sqrt{n}(\hat{\gamma}_\psi - \gamma^*) - \sqrt{n}(\hat{\gamma} - \gamma^*)\right\} \\
&+ \frac{1}{2}\left\{\sqrt{n}(\hat{\beta}_\psi - \beta^*)^\top \frac{\partial^2 M_n(\beta^*,\gamma^*)}{\partial\beta\partial\beta^\top}\sqrt{n}(\hat{\beta}_\psi - \beta^*)^\top - \sqrt{n}(\hat{\beta} - \beta^*)^\top \frac{\partial^2 M_n(\beta^*,\gamma^*)}{\partial\beta\partial\beta^\top}\sqrt{n}(\hat{\beta} - \beta^*)^\top\right\} \\
&+ \frac{1}{2}\left\{\sqrt{n}(\hat{\gamma}_\psi - \gamma^*)^\top \frac{\partial^2 M_n(\beta^*,\gamma^*)}{\partial\gamma\partial\gamma^\top}\sqrt{n}(\hat{\gamma}_\psi - \gamma^*)^\top - \sqrt{n}(\hat{\gamma} - \gamma^*)^\top \frac{\partial^2 M_n(\beta^*,\gamma^*)}{\partial\gamma\partial\gamma^\top}\sqrt{n}(\hat{\gamma} - \gamma^*)^\top\right\} \\
&+ \left\{\sqrt{n}(\hat{\gamma}_\psi - \gamma^*)^\top \frac{\partial^2 M_n(\beta^*,\gamma^*)}{\partial\gamma\partial\beta^\top}\sqrt{n}(\hat{\beta}_\psi - \beta^*)^\top - \sqrt{n}(\hat{\gamma} - \gamma^*)^\top \frac{\partial^2 M_n(\beta^*,\gamma^*)}{\partial\gamma\partial\beta^\top}\sqrt{n}(\hat{\beta} - \beta^*)^\top\right\} \\
&+ o_p(1)
\end{aligned}$$

Using the derivative results (35) to (37) our expansion reduces to

$$\begin{aligned}
n\xi &= \sqrt{n}\frac{\partial M_n(\beta^*,\gamma^*)}{\partial \beta^\top}\left\{\sqrt{n}(\hat{\beta}_\psi - \beta^*) - \sqrt{n}(\hat{\beta} - \beta^*)\right\} \\
&+ \frac{1}{2}\left\{\sqrt{n}(\hat{\beta}_\psi - \beta^*)^\top \frac{\partial^2 M_n(\beta^*,\gamma^*)}{\partial\beta\partial\beta^\top}\sqrt{n}(\hat{\beta}_\psi - \beta^*)^\top - \sqrt{n}(\hat{\beta} - \beta^*)^\top \frac{\partial^2 M_n(\beta^*,\gamma^*)}{\partial\beta\partial\beta^\top}\sqrt{n}(\hat{\beta} - \beta^*)^\top\right\} \\
&+ o_p(1)
\end{aligned}$$

Next we consider the first order Taylor expansions of the estimating equations (31) to (33), taken about the probability limit values. Again, since $\hat{\gamma}, \hat{\gamma}_\psi, \hat{\beta}$ and $\hat{\beta}_\psi$ are CAN,

$$\begin{aligned}
0 &= \frac{\partial M_n(\beta^*,\gamma^*)}{\partial\beta} + \frac{\partial^2 M_n(\beta^*,\gamma^*)}{\partial\beta\partial\beta}(\hat{\beta} - \beta^*) + \frac{\partial^2 M_n(\beta^*,\gamma^*)}{\partial\gamma\partial\beta}(\hat{\gamma} - \gamma^*) + o_p(n^{-1/2}) \\
0 &= \frac{\partial M_n(\beta^*,\gamma^*)}{\partial\beta} + \frac{\partial^2 M_n(\beta^*,\gamma^*)}{\partial\beta\partial\beta}(\hat{\beta}_\psi - \beta^*) + \frac{\partial^2 M_n(\beta^*,\gamma^*)}{\partial\gamma\partial\beta}(\hat{\gamma}_\psi - \gamma^*) - \frac{\partial\psi(\hat{\beta}_\psi)}{\partial\beta}\lambda + o_p(n^{-1/2}) \\
0 &= \psi(\beta^*) + \frac{\partial\psi(\beta^*)}{\partial\beta}(\hat{\beta}_\psi - \beta^*) + o_p(n^{-1/2})
\end{aligned}$$

Under the null, $\psi(\beta^*) = 0$, application of (36) gives

$$\begin{aligned}
0 &= X_n + V_0(\hat{\beta} - \beta^*) + o_p(n^{-1/2}) \\
0 &= X_n + V_0(\hat{\beta}_\psi - \beta^*) - P_0\lambda + o_p(n^{-1/2}) \\
0 &= P_0(\hat{\beta}_\psi - \beta^*) + o_p(n^{-1/2})
\end{aligned}$$

where,

$$\frac{\partial \psi(\beta^*)}{\partial \beta} = P_0$$
$$\frac{\partial M_n(\beta^*, \gamma^*)}{\partial \beta} = X_n$$
$$\frac{\partial^2 M_n(\beta^*, \gamma^*)}{\partial \beta \partial \beta} \xrightarrow{p} V_0$$

It follows immediately from the original proof in Dufour et al. (2017) that

$$\xi = \frac{1}{2} X_n^\top V_0^{-1} P_0^\top (P_0 V_0^{-1} P_0^\top)^{-1} P_0 V_0^{-1} X_n + o_p(n^{-1})$$

The final result follows when $\sqrt{n} X_n \xrightarrow{d} \mathcal{N}(0, 2V_0)$. This can be shown using the same derivative methods as used to show (35) to (37).

A.6. Proof of Equations (21) *and* (22)

We will prove (21) with the result for (22) proceeding in a similar fashion.

Consider Theorem 2 under the null hypothesis $\psi^{(0)}(\beta^*) = \beta_1 \beta_2 = 0$. With the null parameter space given by $B_0 = \{\beta | \psi^{(0)}(\beta) = 0\}$, with

$$\text{Rank}\left(\frac{\partial \psi^{(0)}(\beta)}{\partial \beta}\right) = \begin{cases} 0 & \text{for } \beta_1 = \beta_2 = 0 \\ 1 & \text{otherwise} \end{cases}$$

One may decompose the supremum in (21) as

$$\sup_{\beta_* \in B_0} P_{\beta^*}(S_0 > x) = \max\left\{\sup_{\beta^* \in B_0 \setminus A} P_{\beta^*}(S_0 > x), \sup_{\beta^* \in A} P_{\beta^*}(S_0 > x)\right\} \tag{38}$$

where $A = \{\beta | \beta_1 = \beta_2 = 0\}$. For the first term in the max bracket above, the rank condition of Theorem 2 holds, so for all $\beta^* \in B_0 \setminus A$

$$P_{\beta^*}(S_0 > x) \to 1 - F_{\chi_1^2}(x)$$

Considering the second term, one may decompose the test statistic as

$$P_{\beta^*}(S_0 > x) = P_{\beta^*}(S_1 > x, S_2 > x) \leq P_{\beta^*}(S_1 > x)$$

where (for $j = 1, 2$) $S_j = \min_{\beta \in C_j} S(\beta)$ and $C_j = \{\beta | \beta_j = 0\}$. By Theorem 2, for all β^* in A,

$$P_{\beta^*}(S_1 > x) \to 1 - F_{\chi_1^2}(x)$$

Hence $P_{\beta^*}(S_0 > x)$ is asymptotically bounded from above by $1 - F_{\chi_1^2}(x)$ for all β^* in B_0, so (21) holds.

A.7. G-estimation when outcome model has exposure-mediator interaction

In the following we reason about the NIDE obtained by G-estimation using moment conditions (6) – (8), when one has erroneously excluded an interaction term from the outcome model, but the mediator model, $E(M|X,Z)$ is correctly specified and partially linear, i.e. in truth, (1) and (24) both hold. We define the probability limit as $\beta^* = (\beta_1^*, \beta_2^*, \beta_3^*)$ which solves $E\{U(\beta^*, \gamma^*)\} = 0$ and use assumption A2 as before. Let,

$$\delta_j = \beta_j - \beta_j^*$$
$$\epsilon_x = X - E(X|Z)$$
$$\epsilon_m = M - E(M|X,Z)$$
$$\beta_3 = \frac{E(\epsilon_x g(X,Z))}{E(\epsilon_x X)}$$
$$\Delta_f = f(Z) - f(Z;\gamma_m^*)$$
$$\Delta_g = g(X,Z) - \beta_3 X - g(Z;\gamma_m^*)$$
$$\Delta_h = h(Z) - h(Z;\gamma_m^*)$$

for $j = 1,2,3$. Here β_3 is the least squares coefficient of a regression of $g(X,Z)$ on X. It follows that the expected moment conditions can be written

$$E[U_1(\beta^*, \gamma^*)] = E\{(\epsilon_x + \Delta_h)(\delta_1 X + \Delta_f)\}$$
$$E[U_2(\beta^*, \gamma^*)] = E\{(\epsilon_m + \delta_1 X + \Delta_f)(\delta_2 M + \theta MX + \delta_3 X + \Delta_g)\}$$
$$E[U_3(\beta^*, \gamma^*)] = E\{(\epsilon_x + \Delta_h)(\delta_2 M + \theta MX + \delta_3 X + \Delta_g)\}$$

For the first equation, since $E(\epsilon_x \Delta_f) = 0$, then

$$E[U_1(\beta^*, \gamma^*)] = \delta_1 E[(\epsilon_x + \Delta_h)X] + E(\Delta_h \Delta_f)$$

We assume that $f(z)$ is modelled correctly and when $\delta_1 = 0$ then assumption A2 is satisfied and $\Delta_f = 0$. Using this fact, the second equation becomes

$$E[U_2(\beta^*, \gamma^*)] = \delta_2 E(\epsilon_m M) + \theta E(\epsilon_m MX)$$

where we have used the fact that $E(\epsilon_m X) = E(\epsilon_m \Delta_g) = 0$. Hence,

$$\delta_2 = -\theta \frac{E(\epsilon_m MX)}{E(\epsilon_m M)}$$

which, since $\delta_1 = 0$, gives the result

$$\beta_1^* \beta_2^* = \beta_1 \left(\beta_2 + \theta \bar{x}\right)$$

where

$$\bar{x} = \frac{E(\epsilon_m M X)}{E(\epsilon_m M)} = \frac{E[X \text{var}(M|X, Z)]}{E[\text{var}(M|X, Z)]} \tag{39}$$

can be thought of as a weighted average of X, or as a population least squares regression coefficient from regressing MX on M.

B. Additional Simulation Plots

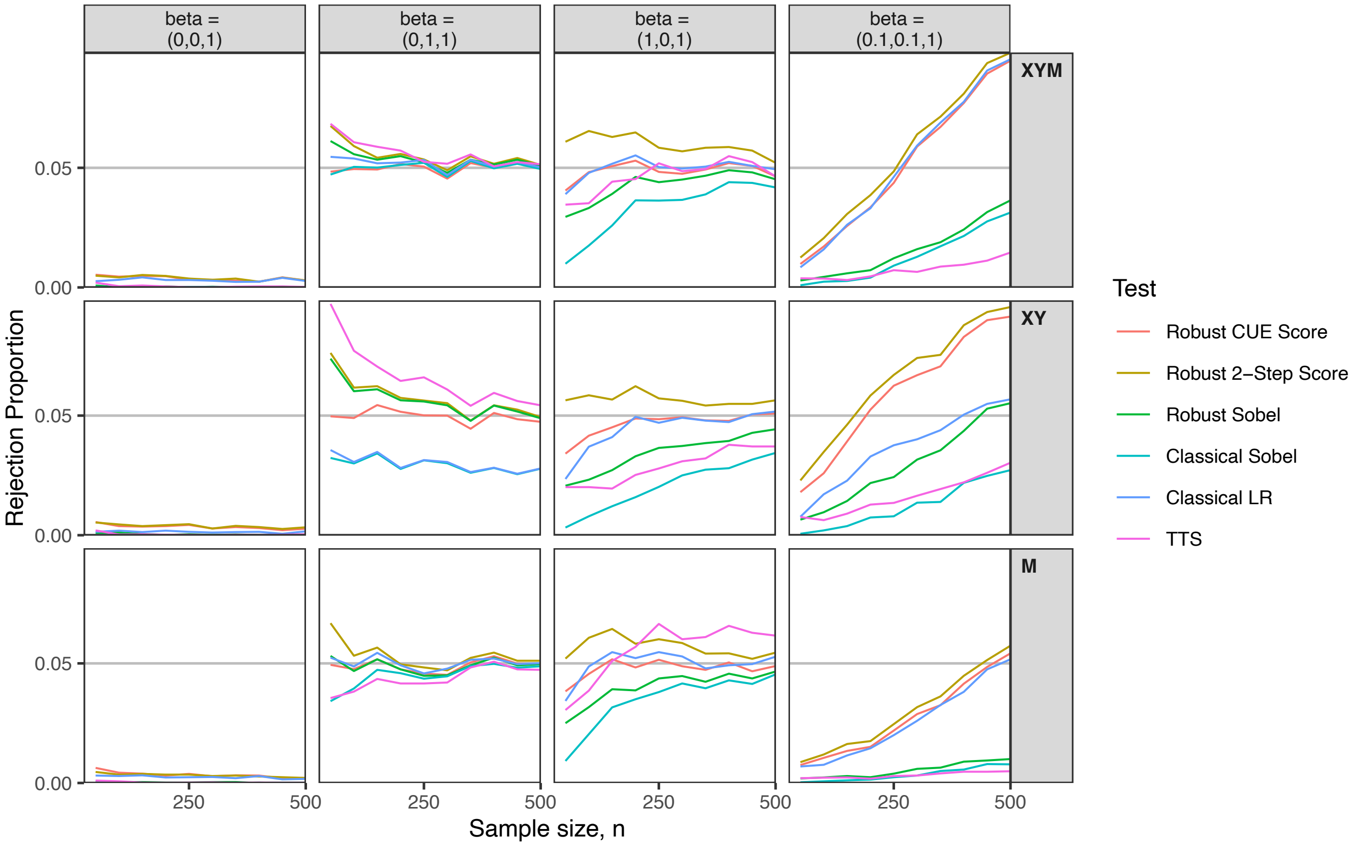

FIGURE 6.

Simulated results from data generating process B of the proportion of the 10^4 datasets for which the no-mediation hypothesis (H_0) is rejected at the 5% level testing using the CUE score, Two-step score, Robust Sobel, Classical Sobel, Classical LR, and TTS methods. Each column represents a different true β parameter, whilst each row gives the models which are correctly specified (those for which the misspecification indicator is equal to zero)

C. Simulation Data Tables

The Tables 4 to 9 present the results of the bias simulation where 10^3 dataset replicates were used to estimate the bias and variance of the G-estimator and the TTS based estimators. The variance of the G-estimators were calculated in 3 different ways, using the empirical variance over the 10^3 data set replicates, using theoretical results presented in this paper and using Bootstrap with 10^3 bootstrap iterations. For the TTS methods only the empirical variance was calculated. All variances estimates have been scaled by the sample size N. Theoretical and Bootstrap estimates are given as the mean over the 10^3 dataset replicates. The standard error of each quantity is given in brackets.

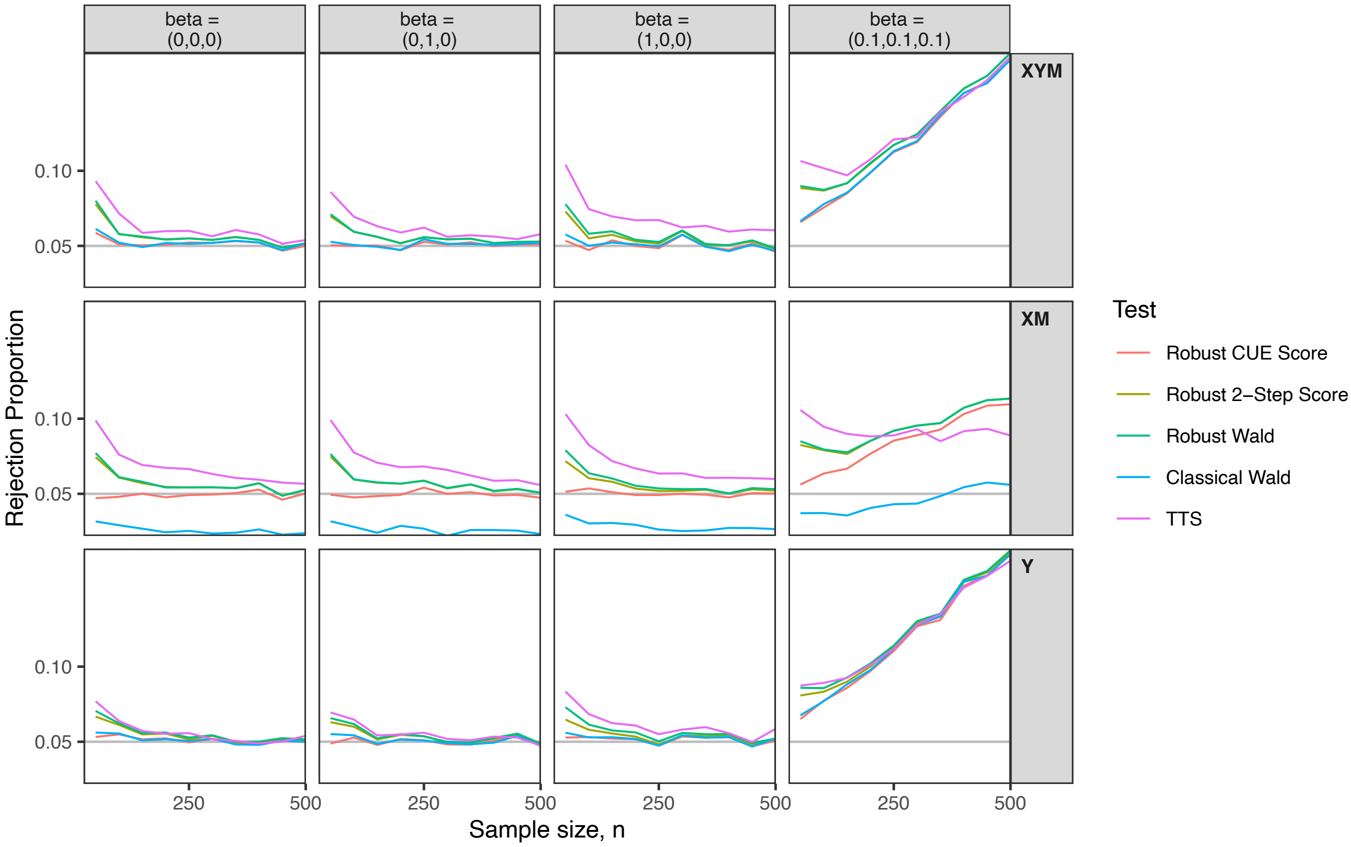

FIGURE 7.

Simulated results from data generating process B of the proportion of the 10^4 datasets for which the no-direct effect hypothesis (H_1) is rejected at the 5% level testing using the CUE score, Two-step score, Robust Wald, Classical Wald, and TTS methods. Each column represents a different true β parameter, whilst each row gives the models which are correctly specified (those for which the misspecification indicator is equal to zero)

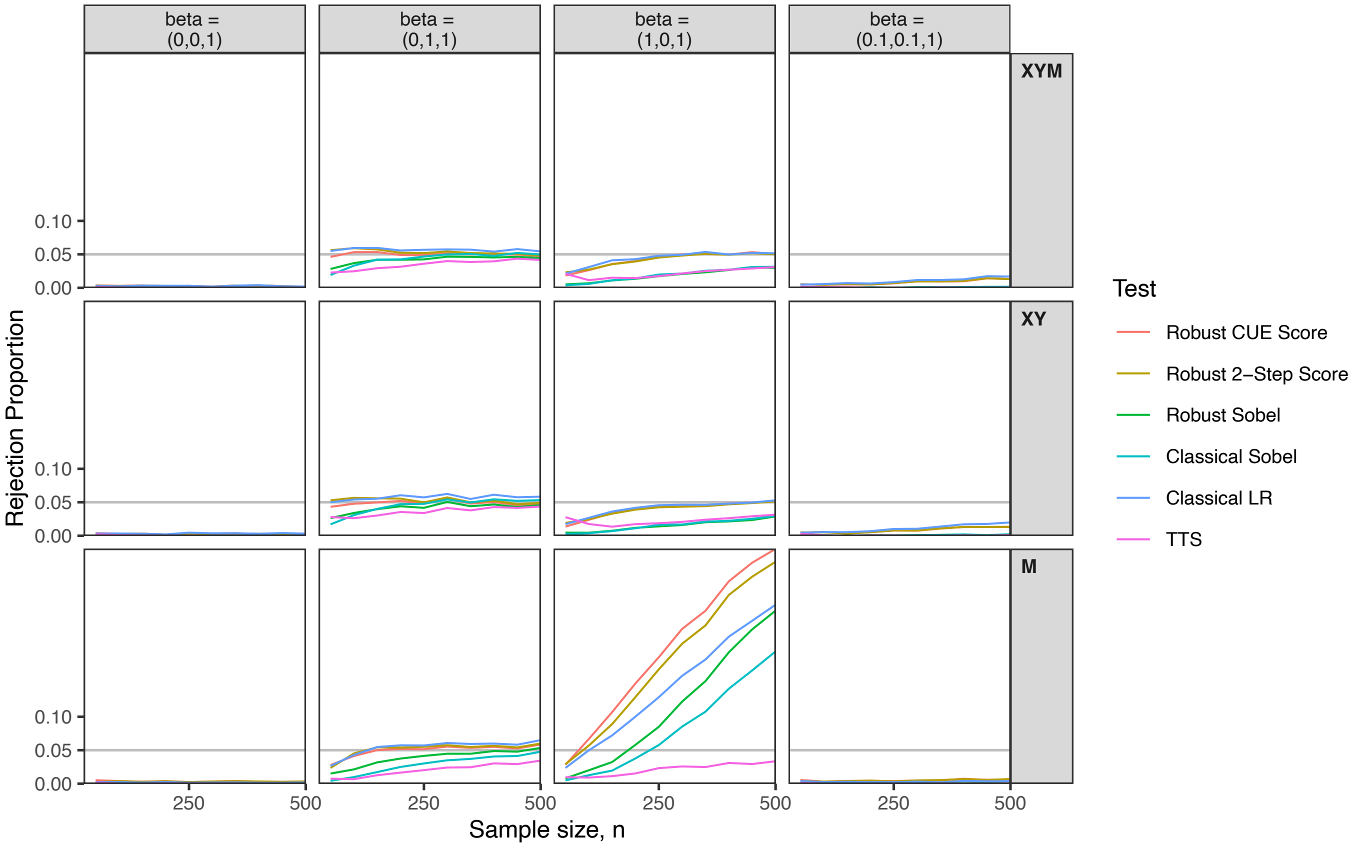

FIGURE 8.

Simulated results from data generating process C of the proportion of the 10^4 datasets for which the no-mediation hypothesis (H_0) is rejected at the 5% level testing using the CUE score, Two-step score, Robust Sobel, Classical Sobel, Classical LR, and TTS methods. Each column represents a different true β parameter, whilst each row gives the models which are correctly specified (those for which the misspecification indicator is equal to zero)

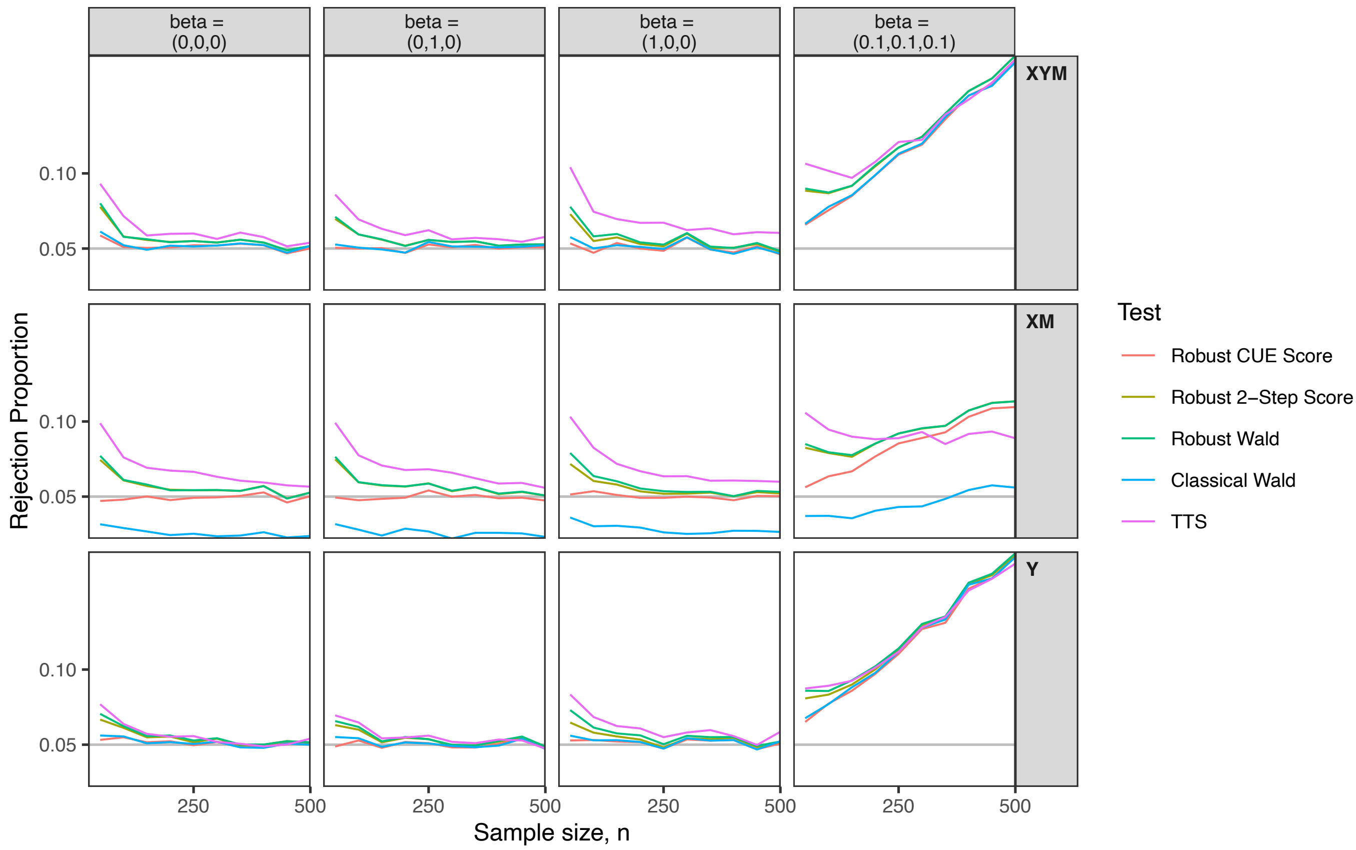

FIGURE 9.

Simulated results from data generating process C of the proportion of the 10^4 datasets for which the no-direct effect hypothesis (H_1) is rejected at the 5% level testing using the CUE score, Two-step score, Robust Wald, Classical Wald, and TTS methods. Each column represents a different true β parameter, whilst each row gives the models which are correctly specified (those for which the misspecification indicator is equal to zero)

Table 4: NDE estimand for Data Generating Process A

Beta	N	Spec	Bias_G	Bias_TTS	Var_G	Var_TTS	Theory_Var_G	Boot_Var_G
(0,0,0)	100	XYM	0.00205(0.00718)	-0.00188(0.0077)	5.15(0.23)	5.92(0.265)	4.76(0.0327)	4.97(0.035)
(0,0,0)	100	XY	-0.000128(0.00708)	-0.00336(0.00737)	5.01(0.224)	5.44(0.243)	4.75(0.0327)	4.94(0.0347)
(0,0,0)	100	XM	0.0171(0.00974)	0.0147(0.016)	9.49(0.425)	25.6(1.15)	9.37(0.115)	9.64(0.111)
(0,0,0)	100	M	0.843(0.0104)	0.932(0.0112)	10.9(0.488)	12.5(0.558)	10.2(0.108)	10.6(0.111)
(0,0,0)	100	Y	-0.00524(0.00728)	-0.0053(0.00738)	5.3(0.237)	5.45(0.244)	4.72(0.0299)	4.89(0.0319)
(0,0,0)	500	XYM	0.00409(0.00311)	0.00318(0.00327)	4.83(0.216)	5.35(0.239)	4.84(0.0142)	4.88(0.0156)
(0,0,0)	500	XY	0.00205(0.00308)	0.00241(0.0032)	4.73(0.212)	5.12(0.229)	4.82(0.0133)	4.84(0.0149)
(0,0,0)	500	XM	0.00245(0.00443)	0.00352(0.00791)	9.8(0.438)	31.3(1.4)	9.69(0.0554)	9.73(0.057)
(0,0,0)	500	M	0.861(0.00444)	0.941(0.00456)	9.88(0.442)	10.4(0.464)	10.4(0.0487)	10.5(0.0499)
(0,0,0)	500	Y	0.0026(0.00316)	0.00141(0.00321)	4.99(0.223)	5.14(0.23)	4.79(0.0125)	4.82(0.0143)
(0,0,0)	1000	XYM	-0.00136(0.00221)	-0.00103(0.00227)	4.87(0.218)	5.16(0.231)	4.85(0.00996)	4.86(0.0123)
(0,0,0)	1000	XY	0.000526(0.00218)	0.000912(0.00226)	4.76(0.213)	5.09(0.228)	4.83(0.00941)	4.83(0.0113)
(0,0,0)	1000	XM	-0.000147(0.00309)	0.00113(0.0051)	9.57(0.428)	26.1(1.17)	9.58(0.038)	9.59(0.0395)
(0,0,0)	1000	M	0.869(0.00331)	0.95(0.0034)	11(0.491)	11.6(0.518)	10.6(0.0383)	10.7(0.0406)
(0,0,0)	1000	Y	0.00184(0.00216)	0.00153(0.00221)	4.66(0.208)	4.88(0.218)	4.8(0.00878)	4.82(0.0113)
(1,1,1)	100	XYM	0.0107(0.00776)	0.00396(0.0103)	6.03(0.27)	10.5(0.471)	5.82(0.0437)	6.11(0.0459)
(1,1,1)	100	XY	-0.00412(0.00745)	-0.00188(0.00778)	5.55(0.249)	6.05(0.271)	5.13(0.036)	5.39(0.0388)
(1,1,1)	100	XM	-0.0089(0.0112)	-0.0113(0.0193)	12.4(0.557)	37.4(1.67)	12.2(0.152)	12.7(0.152)
(1,1,1)	100	M	0.838(0.0115)	0.914(0.0159)	13.3(0.596)	25.2(1.13)	12.8(0.145)	13.4(0.149)

(1,1,1)	100	Y	-0.0135(0.0073)	-0.00428(0.00813)	5.33(0.238)	6.61(0.296)	5.7(0.0387)	5.95(0.0407)
(1,1,1)	500	XYM	0.00438(0.00351)	0.00847(0.00445)	6.14(0.275)	9.92(0.444)	5.85(0.0192)	5.9(0.0211)
(1,1,1)	500	XY	0.00309(0.00316)	0.00385(0.0033)	4.99(0.223)	5.44(0.243)	5.16(0.0151)	5.21(0.0173)
(1,1,1)	500	XM	-0.00834(0.00511)	-0.00323(0.0115)	13(0.583)	66.5(2.98)	12.6(0.0688)	12.7(0.0705)
(1,1,1)	500	M	0.872(0.00511)	0.954(0.00691)	13(0.584)	23.9(1.07)	13.3(0.0709)	13.4(0.0732)
(1,1,1)	500	Y	0.00428(0.00336)	0.0034(0.00358)	5.64(0.252)	6.42(0.287)	5.7(0.0161)	5.75(0.0182)
(1,1,1)	1000	XYM	4.3e-05(0.00237)	-0.000762(0.00309)	5.6(0.251)	9.53(0.426)	5.81(0.0138)	5.85(0.0163)
(1,1,1)	1000	XY	-8.16e-05(0.00223)	0.000244(0.00232)	4.98(0.223)	5.37(0.24)	5.16(0.0109)	5.2(0.0131)
(1,1,1)	1000	XM	0.00194(0.00361)	0.0103(0.00841)	13(0.583)	70.7(3.16)	12.7(0.0531)	12.7(0.0564)
(1,1,1)	1000	M	0.863(0.0037)	0.943(0.00521)	13.7(0.611)	27.2(1.22)	13.5(0.0581)	13.6(0.061)
(1,1,1)	1000	Y	0.00272(0.00243)	0.0014(0.00259)	5.92(0.265)	6.7(0.3)	5.74(0.0115)	5.76(0.014)

Table 5: NIDE estimand for Data Generating Process A

Beta	N	Spec	Bias_G	Bias_TTS	Var_G	Var_TTS	Theory_Var_G	Boot_Var_G
(0,0,0)	100	XYM	-0.000762(0.00069)	-0.000561(0.00123)	0.0476(0.00213)	0.151(0.00675)	0.0974(0.00342)	0.154(0.154)
(0,0,0)	100	XY	0.00211(0.000644)	0.00228(0.00151)	0.0414(0.00185)	0.229(0.0103)	0.0731(0.00303)	0.112(0.112)
(0,0,0)	100	XM	0.00241(0.00128)	0.0055(0.0022)	0.163(0.0073)	0.486(0.0217)	0.29(0.0112)	0.445(0.445)
(0,0,0)	100	M	0.000992(0.00121)	0.00284(0.00223)	0.146(0.00651)	0.497(0.0223)	0.271(0.0101)	0.413(0.413)
(0,0,0)	100	Y	0.00209(0.00181)	0.00264(0.00228)	0.328(0.0147)	0.518(0.0232)	0.311(0.0075)	0.36(0.36)
(0,0,0)	500	XYM	-2.61e-05(0.000136)	-0.000147(0.000248)	0.0093(0.000416)	0.0308(0.00138)	0.019(0.000606)	0.029(0.029)
(0,0,0)	500	XY	9.12e-06(0.000105)	-6.06e-06(0.000298)	0.00556(0.000249)	0.0445(0.00199)	0.0125(0.000427)	0.0192(0.0192)
(0,0,0)	500	XM	-0.000496(0.000261)	-0.000989(0.000548)	0.0341(0.00153)	0.15(0.00673)	0.0621(0.00217)	0.0921(0.0921)
(0,0,0)	500	M	0.000276(0.000243)	0.000609(0.000412)	0.0295(0.00132)	0.0847(0.00379)	0.0536(0.00184)	0.0796(0.0796)
(0,0,0)	500	Y	-0.00153(0.000734)	-0.00077(0.000933)	0.27(0.0121)	0.436(0.0195)	0.27(0.00291)	0.279(0.279)
(0,0,0)	1000	XYM	2.98e-05(6.83e-05)	-0.000152(0.000112)	0.00466(0.000209)	0.0125(0.000561)	0.00954(0.000296)	0.0145(0.0145)
(0,0,0)	1000	XY	-3.7e-05(5.58e-05)	-1.95e-05(0.000161)	0.00312(0.00014)	0.0259(0.00116)	0.00646(0.000211)	0.00974(0.00974)
(0,0,0)	1000	XM	-2.77e-05(0.000126)	5.37e-05(0.000268)	0.0158(0.000706)	0.0716(0.0032)	0.0292(0.000954)	0.0437(0.0437)
(0,0,0)	1000	M	4.46e-05(0.000115)	4.94e-05(0.000194)	0.0133(0.000594)	0.0375(0.00168)	0.0267(0.000878)	0.0399(0.0399)
(0,0,0)	1000	Y	-0.00055(0.000508)	-0.000439(0.000653)	0.258(0.0115)	0.426(0.0191)	0.266(0.00217)	0.269(0.269)
(1,1,1)	100	XYM	0.003(0.00743)	0.011(0.0103)	5.52(0.247)	10.7(0.478)	5.88(0.0507)	6.16(6.16)
(1,1,1)	100	XY	-0.000413(0.0103)	-0.00162(0.0178)	10.5(0.472)	31.7(1.42)	9.58(0.109)	9.8(9.8)
(1,1,1)	100	XM	0.00459(0.0093)	0.0116(0.0145)	8.65(0.387)	21(0.942)	8.03(0.0955)	8.51(8.51)
(1,1,1)	100	M	-0.00994(0.00879)	0.0058(0.0134)	7.73(0.346)	17.9(0.802)	7.39(0.0889)	7.91(7.91)

(1,1,1)	100	Y	0.831(0.0111)	0.921(0.0124)	12.4(0.553)	15.5(0.693)	11.3(0.111)	11.7(11.7)
(1,1,1)	500	XYM	0.000102(0.00352)	-0.00223(0.0044)	6.19(0.277)	9.67(0.433)	5.84(0.0236)	5.9(5.9)
(1,1,1)	500	XY	-0.00721(0.00441)	-0.0119(0.00778)	9.73(0.435)	30.3(1.35)	9.88(0.0587)	9.9(9.9)
(1,1,1)	500	XM	0.00657(0.00394)	0.00122(0.00917)	7.77(0.348)	42.1(1.88)	7.92(0.0416)	8.03(8.03)
(1,1,1)	500	M	-0.00231(0.00376)	-0.00462(0.00599)	7.06(0.316)	17.9(0.803)	7.42(0.0403)	7.58(7.58)
(1,1,1)	500	Y	0.855(0.00516)	0.937(0.0055)	13.3(0.596)	15.1(0.677)	11.8(0.0618)	11.8(11.8)
(1,1,1)	1000	XYM	-0.00296(0.00241)	-0.00196(0.00314)	5.83(0.261)	9.83(0.44)	5.84(0.0181)	5.87(5.87)
(1,1,1)	1000	XY	0.00151(0.00319)	-0.00187(0.00545)	10.2(0.454)	29.7(1.33)	9.99(0.0444)	10(10)
(1,1,1)	1000	XM	0.00169(0.00286)	-0.00207(0.00693)	8.18(0.366)	48.1(2.15)	7.89(0.0308)	7.93(7.93)
(1,1,1)	1000	M	-0.000208(0.0028)	0.00131(0.00458)	7.81(0.35)	21(0.938)	7.45(0.0309)	7.6(7.6)
(1,1,1)	1000	Y	0.865(0.00339)	0.946(0.00361)	11.5(0.514)	13.1(0.584)	11.8(0.0447)	11.8(11.8)

Table 6: NDE estimand for Data Generating Process B

Beta	N	Spec	Bias_G	Bias_TTS	Var_G	Var_TTS	Theory_Var_G	Boot_Var_G
(0,0,0)	100	XYM	-0.00368(0.00703)	-0.00329(0.00757)	4.95(0.221)	5.73(0.256)	4.8(0.0329)	5.02(0.0354)
(0,0,0)	100	XY	-0.0067(0.00723)	-0.00355(0.00782)	5.22(0.234)	6.11(0.273)	4.8(0.0342)	5(0.036)
(0,0,0)	100	XM	0.00277(0.00967)	0.0136(0.0149)	9.35(0.419)	22.2(0.993)	9.1(0.0968)	9.44(0.097)
(0,0,0)	100	M	0.859(0.0106)	0.946(0.0112)	11.2(0.499)	12.5(0.558)	10.1(0.0966)	10.5(0.0994)
(0,0,0)	100	Y	0.014(0.00713)	0.0153(0.00731)	5.09(0.228)	5.34(0.239)	4.7(0.029)	4.87(0.0308)
(0,0,0)	500	XYM	0.000774(0.00305)	-0.000145(0.00323)	4.65(0.208)	5.23(0.234)	4.81(0.0137)	4.85(0.0154)
(0,0,0)	500	XY	0.00594(0.00324)	0.00598(0.00343)	5.24(0.234)	5.87(0.263)	4.83(0.0136)	4.85(0.0155)
(0,0,0)	500	XM	-0.00153(0.00435)	-0.00482(0.00808)	9.45(0.423)	32.7(1.46)	9.63(0.0557)	9.69(0.0565)
(0,0,0)	500	M	0.864(0.00459)	0.947(0.00475)	10.5(0.472)	11.3(0.504)	10.6(0.0526)	10.7(0.0541)
(0,0,0)	500	Y	-0.00428(0.00308)	-0.00494(0.00313)	4.73(0.212)	4.89(0.219)	4.75(0.012)	4.78(0.0138)
(0,0,0)	1000	XYM	0.000397(0.00226)	0.00151(0.00238)	5.12(0.229)	5.65(0.253)	4.83(0.00957)	4.86(0.0116)
(0,0,0)	1000	XY	-0.000543(0.00224)	-0.00172(0.00232)	5(0.224)	5.38(0.241)	4.82(0.00953)	4.83(0.012)
(0,0,0)	1000	XM	0.00162(0.00325)	0.00184(0.00547)	10.6(0.473)	30(1.34)	9.67(0.0395)	9.7(0.0425)
(0,0,0)	1000	M	0.868(0.00323)	0.948(0.00333)	10.5(0.468)	11.1(0.498)	10.6(0.0381)	10.7(0.0415)
(0,0,0)	1000	Y	0.00221(0.00214)	0.00212(0.00218)	4.58(0.205)	4.73(0.212)	4.76(0.0087)	4.78(0.011)
(1,1,1)	100	XYM	0.012(0.00748)	0.0165(0.0104)	5.59(0.25)	10.7(0.481)	5.48(0.0399)	5.76(0.0431)
(1,1,1)	100	XY	0.0157(0.00721)	0.0209(0.00754)	5.2(0.233)	5.68(0.254)	5.07(0.0358)	5.29(0.0377)
(1,1,1)	100	XM	0.000393(0.0108)	-0.0156(0.0217)	11.7(0.525)	47.3(2.12)	11(0.127)	11.5(0.128)
(1,1,1)	100	M	0.849(0.0109)	0.917(0.0166)	11.9(0.532)	27.6(1.24)	11.7(0.133)	12.2(0.136)

(1,1,1)	100	Y	0.00425(0.00763)	0.0164(0.00931)	5.82(0.261)	8.66(0.388)	5.47(0.0368)	5.71(0.0393)
(1,1,1)	500	XYM	0.000856(0.00326)	-0.00715(0.0112)	5.33(0.238)	63.1(2.82)	5.4(0.0175)	5.45(0.0193)
(1,1,1)	500	XY	-0.00366(0.00313)	-0.00349(0.00337)	4.91(0.22)	5.67(0.254)	5.12(0.0148)	5.16(0.0159)
(1,1,1)	500	XM	0.00671(0.0049)	-0.0199(0.0274)	12(0.537)	374(16.7)	11.3(0.0617)	11.3(0.0635)
(1,1,1)	500	M	0.862(0.00503)	1.19(0.24)	12.7(0.567)	28800(1290)	12.4(0.0697)	12.5(0.0707)
(1,1,1)	500	Y	-0.000415(0.00329)	-0.00134(0.00975)	5.41(0.242)	47.5(2.13)	5.5(0.0164)	5.53(0.018)
(1,1,1)	1000	XYM	0.00236(0.00228)	0.0089(0.0455)	5.21(0.233)	2070(92.6)	5.43(0.0119)	5.46(0.0144)
(1,1,1)	1000	XY	-0.000313(0.00221)	-0.00102(0.00234)	4.87(0.218)	5.48(0.245)	5.1(0.0106)	5.12(0.0129)
(1,1,1)	1000	XM	0.00758(0.00329)	0.113(0.0773)	10.8(0.483)	5980(268)	11.4(0.0459)	11.5(0.048)
(1,1,1)	1000	M	0.864(0.00339)	0.87(0.0548)	11.5(0.513)	3010(135)	12.3(0.0469)	12.3(0.0508)
(1,1,1)	1000	Y	-0.000468(0.00235)	-0.199(0.221)	5.51(0.247)	49000(2190)	5.5(0.0108)	5.52(0.0133)

Table 7: NIDE estimand for Data Generating Process B

Beta	N	Spec	Bias_G	Bias_TTS	Var_G	Var_TTS	Theory_Var_G	Boot_Var_G
(0,0,0)	100	XYM	0.000217(0.000709)	0.00168(0.00121)	0.0503(0.00225)	0.147(0.00659)	0.0984(0.00356)	0.155(0.155)
(0,0,0)	100	XY	0.000536(0.000679)	0.00193(0.00154)	0.0461(0.00206)	0.237(0.0106)	0.0742(0.00289)	0.117(0.117)
(0,0,0)	100	XM	-0.000557(0.00124)	-0.00262(0.00261)	0.154(0.00688)	0.679(0.0304)	0.285(0.0112)	0.441(0.441)
(0,0,0)	100	M	0.000394(0.00114)	0.0045(0.00231)	0.13(0.00582)	0.535(0.0239)	0.242(0.0106)	0.381(0.381)
(0,0,0)	100	Y	-0.000822(0.0016)	-0.0018(0.00223)	0.258(0.0115)	0.498(0.0223)	0.28(0.00735)	0.332(0.332)
(0,0,0)	500	XYM	8.56e-05(0.000149)	0.000237(0.000246)	0.0111(0.000495)	0.0302(0.00135)	0.0196(0.000666)	0.0296(0.0296)
(0,0,0)	500	XY	1.89e-05(0.000119)	0.000217(0.000285)	0.00702(0.000314)	0.0407(0.00182)	0.0141(0.000477)	0.0213(0.0213)
(0,0,0)	500	XM	3.9e-05(0.000242)	-0.000221(0.000528)	0.0293(0.00131)	0.139(0.00624)	0.0571(0.00189)	0.0856(0.0856)
(0,0,0)	500	M	1.1e-05(0.000228)	0.000383(0.000398)	0.026(0.00116)	0.0794(0.00355)	0.0517(0.00181)	0.0779(0.0779)
(0,0,0)	500	Y	0.000357(0.000675)	0.000558(0.000943)	0.228(0.0102)	0.444(0.0199)	0.222(0.00265)	0.23(0.23)
(0,0,0)	1000	XYM	1.94e-05(7.09e-05)	-0.000156(0.000121)	0.00502(0.000225)	0.0146(0.000652)	0.00985(0.000324)	0.0148(0.0148)
(0,0,0)	1000	XY	4.19e-05(6.36e-05)	7.49e-05(0.000152)	0.00405(0.000181)	0.023(0.00103)	0.00729(0.000251)	0.0108(0.0108)
(0,0,0)	1000	XM	-0.000164(0.000122)	-4.87e-05(0.000261)	0.0149(0.000665)	0.0681(0.00305)	0.0293(0.000969)	0.044(0.044)
(0,0,0)	1000	M	-0.000187(0.000119)	-0.000146(0.000199)	0.0142(0.000636)	0.0396(0.00177)	0.0258(0.000902)	0.0388(0.0388)
(0,0,0)	1000	Y	-0.000691(0.000461)	-0.000609(0.000637)	0.213(0.00953)	0.406(0.0182)	0.219(0.00194)	0.223(0.223)
(1,1,1)	100	XYM	-0.0104(0.00965)	-0.017(0.0121)	9.3(0.416)	14.6(0.652)	8.7(0.11)	9.09(9.09)
(1,1,1)	100	XY	-0.0215(0.0116)	-0.0301(0.0175)	13.4(0.6)	30.5(1.36)	12.6(0.145)	12.9(12.9)
(1,1,1)	100	XM	0.00847(0.00995)	0.00372(0.0188)	9.89(0.443)	35.2(1.57)	9.97(0.128)	10.6(10.6)
(1,1,1)	100	M	-0.0017(0.00999)	0.0244(0.0162)	9.97(0.446)	26.2(1.17)	9.42(0.129)	10(10)

(1,1,1)	100	Y	0.834(0.0119)	0.906(0.0135)	14.1(0.632)	18.1(0.811)	13.8(0.141)	14.3(14.3)
(1,1,1)	500	XYM	-0.0118(0.00416)	-0.00184(0.0116)	8.65(0.387)	67.8(3.04)	8.61(0.0449)	8.68(8.68)
(1,1,1)	500	XY	0.0018(0.00523)	-0.00395(0.00823)	13.7(0.611)	33.9(1.52)	13.1(0.0755)	13.2(13.2)
(1,1,1)	500	XM	0.000698(0.00445)	0.0333(0.0266)	9.89(0.443)	353(15.8)	9.88(0.0565)	9.99(9.99)
(1,1,1)	500	M	0.00828(0.00433)	-0.238(0.24)	9.37(0.419)	28800(1290)	9.39(0.0647)	9.65(9.65)
(1,1,1)	500	Y	0.855(0.00529)	0.938(0.0106)	14(0.626)	55.8(2.5)	14.5(0.0749)	14.6(14.6)
(1,1,1)	1000	XYM	0.00217(0.00298)	-0.00639(0.0457)	8.9(0.398)	2090(93.3)	8.71(0.0361)	8.75(8.75)
(1,1,1)	1000	XY	-0.00308(0.00367)	-0.0062(0.00601)	13.5(0.603)	36.1(1.62)	13.2(0.0548)	13.2(13.2)
(1,1,1)	1000	XM	0.00129(0.00319)	-0.0985(0.0772)	10.2(0.456)	5960(267)	9.84(0.0422)	9.92(9.92)
(1,1,1)	1000	M	-0.00305(0.00302)	0.0711(0.055)	9.12(0.408)	3020(135)	9.39(0.0538)	9.55(9.55)
(1,1,1)	1000	Y	0.862(0.00374)	1.14(0.221)	14(0.624)	49000(2190)	14.7(0.0584)	14.7(14.7)

Table 8: NDE estimand for Data Generating Process C

Beta	N	Spec	Bias_G	Bias_TTS	Var_G	Var_TTS	Theory_Var_G	Boot_Var_G
(0,0,0)	100	XYM	-0.00301(0.00656)	0.00107(0.00705)	4.31(0.193)	4.97(0.222)	4.77(0.0313)	6.54(0.248)
(0,0,0)	100	XY	0.00514(0.00733)	0.00348(0.00803)	5.37(0.24)	6.45(0.288)	4.79(0.0321)	7.62(0.303)
(0,0,0)	100	XM	0.00579(0.0101)	0.0267(0.0184)	10.2(0.455)	34(1.52)	9.3(0.122)	12.3(0.503)
(0,0,0)	100	M	0.836(0.0105)	0.92(0.0111)	11(0.494)	12.3(0.552)	9.96(0.0995)	80.9(1.9)
(0,0,0)	100	Y	0.00444(0.00685)	0.00214(0.00691)	4.69(0.21)	4.78(0.214)	4.59(0.0306)	7.85(0.311)
(0,0,0)	500	XYM	0.000106(0.00301)	0.00157(0.00314)	4.52(0.202)	4.94(0.221)	4.83(0.0136)	6.86(0.264)
(0,0,0)	500	XY	0.000599(0.00315)	0.0014(0.00335)	4.98(0.223)	5.62(0.252)	4.83(0.0139)	6.85(0.267)
(0,0,0)	500	XM	-0.00674(0.00455)	-0.00605(0.00887)	10.4(0.464)	39.3(1.76)	9.64(0.0593)	12.3(0.488)
(0,0,0)	500	M	0.862(0.00472)	0.944(0.0048)	11.1(0.499)	11.5(0.515)	10.5(0.0534)	378(4.19)
(0,0,0)	500	Y	0.000134(0.00313)	0.000521(0.00314)	4.9(0.219)	4.93(0.221)	4.61(0.0121)	14.2(0.497)
(0,0,0)	1000	XYM	-0.00227(0.00225)	-0.00227(0.00238)	5.07(0.227)	5.67(0.254)	4.83(0.00983)	7.16(0.288)
(0,0,0)	1000	XY	-0.00111(0.00221)	-0.000989(0.0023)	4.88(0.218)	5.28(0.236)	4.84(0.0101)	7.07(0.258)
(0,0,0)	1000	XM	-0.00127(0.00302)	0.000813(0.00527)	9.12(0.408)	27.8(1.24)	9.64(0.0377)	11.7(0.472)
(0,0,0)	1000	M	0.863(0.00337)	0.942(0.0034)	11.4(0.509)	11.6(0.518)	10.6(0.0371)	742(5.98)
(0,0,0)	1000	Y	-0.000594(0.00219)	-0.00018(0.00222)	4.8(0.215)	4.91(0.22)	4.64(0.00849)	21.4(0.633)
(1,1,1)	100	XYM	0.00449(0.00736)	0.00319(0.00818)	5.42(0.242)	6.69(0.299)	5.11(0.0375)	70.7(1.34)
(1,1,1)	100	XY	-0.00356(0.00714)	0.00596(0.00833)	5.1(0.228)	6.95(0.311)	4.97(0.0351)	73.6(1.33)
(1,1,1)	100	XM	0.0209(0.0103)	-0.00588(0.0212)	10.6(0.474)	45.1(2.02)	9.77(0.121)	79.3(1.83)
(1,1,1)	100	M	0.888(0.0109)	0.924(0.0115)	11.9(0.534)	13.3(0.595)	10.9(0.108)	298(3.81)

(1,1,1)	100	Y	-0.00842(0.00693)	-0.00814(0.00754)	4.81(0.215)	5.69(0.255)	5.03(0.0372)	57.5(1.1)
(1,1,1)	500	XYM	-0.000439(0.00312)	-0.000264(0.00336)	4.87(0.218)	5.64(0.252)	5.06(0.0155)	323(2.79)
(1,1,1)	500	XY	-0.00397(0.0032)	-0.00326(0.00347)	5.11(0.229)	6.03(0.27)	5.04(0.0154)	336(2.86)
(1,1,1)	500	XM	0.0303(0.00452)	-0.00829(0.0084)	10.2(0.457)	35.3(1.58)	10.2(0.0614)	352(3.92)
(1,1,1)	500	M	0.901(0.00479)	0.935(0.00495)	11.5(0.513)	12.3(0.548)	11.2(0.0545)	1460(8.46)
(1,1,1)	500	Y	0.00265(0.00313)	0.00292(0.00328)	4.91(0.219)	5.38(0.241)	5.02(0.0151)	267(2.49)
(1,1,1)	1000	XYM	-0.00664(0.0022)	-0.00507(0.00238)	4.83(0.216)	5.65(0.253)	5.07(0.0111)	629(3.77)
(1,1,1)	1000	XY	-0.00122(0.00242)	-0.000419(0.00257)	5.86(0.262)	6.6(0.296)	5.04(0.0108)	675(4.3)
(1,1,1)	1000	XM	0.029(0.00318)	-0.0139(0.00569)	10.1(0.453)	32.4(1.45)	10.4(0.0411)	691(5.43)
(1,1,1)	1000	M	0.918(0.00339)	0.951(0.00349)	11.5(0.515)	12.2(0.546)	11.5(0.0405)	2970(12.1)
(1,1,1)	1000	Y	0.000258(0.00229)	-0.000654(0.00239)	5.24(0.234)	5.73(0.257)	5.05(0.0111)	523(3.62)

Table 9: NIDE estimand for Data Generating Process C

Beta	N	Spec	Bias_G	Bias_TTS	Var_G	Var_TTS	Theory_Var_G	Boot_Var_G
(0,0,0)	100	XYM	-0.000386(0.000758)	1.94e-05(0.0011)	0.0574(0.00257)	0.12(0.00537)	0.0994(0.273)	9.98(9.98)
(0,0,0)	100	XY	0.000943(0.000678)	0.00172(0.00138)	0.0459(0.00206)	0.19(0.00849)	0.0945(0.21)	9.26(9.26)
(0,0,0)	100	XM	0.00118(0.00118)	0.00274(0.00183)	0.14(0.00625)	0.336(0.015)	0.241(0.596)	22.5(22.5)
(0,0,0)	100	M	-0.00348(0.00116)	-0.00133(0.00158)	0.135(0.00602)	0.251(0.0112)	0.224(0.58)	21.9(21.9)
(0,0,0)	100	Y	-0.00104(0.00105)	-0.000516(0.00147)	0.111(0.00495)	0.216(0.00967)	0.159(0.179)	8.59(8.59)
(0,0,0)	500	XYM	0.000136(0.000144)	-0.000222(0.000225)	0.0103(0.000461)	0.0252(0.00113)	0.0198(0.223)	9.59(9.59)
(0,0,0)	500	XY	-0.000105(0.000137)	-7.15e-05(0.00026)	0.00934(0.000418)	0.0339(0.00152)	0.0192(0.221)	9.3(9.3)
(0,0,0)	500	XM	0.000125(0.000217)	0.00019(0.00033)	0.0234(0.00105)	0.0543(0.00243)	0.0461(0.506)	21.8(21.8)
(0,0,0)	500	M	-0.000259(0.000191)	0.000187(0.000271)	0.0183(0.000817)	0.0369(0.00165)	0.0392(0.467)	20.2(20.2)
(0,0,0)	500	Y	0.000414(0.000398)	0.000288(0.000525)	0.079(0.00354)	0.138(0.00616)	0.0888(0.16)	8.2(8.2)
(0,0,0)	1000	XYM	-6.41e-05(7.27e-05)	5.69e-05(0.000113)	0.00529(0.000237)	0.0128(0.000571)	0.0101(0.202)	9.45(9.45)
(0,0,0)	1000	XY	3.72e-05(6.75e-05)	2.97e-05(0.000135)	0.00456(0.000204)	0.0182(0.000813)	0.0099(0.193)	9.01(9.01)
(0,0,0)	1000	XM	-1.45e-05(0.000103)	-9.14e-05(0.000177)	0.0107(0.000477)	0.0315(0.00141)	0.0227(0.419)	20.6(20.6)
(0,0,0)	1000	M	-0.000306(0.000115)	9.28e-05(0.000152)	0.0133(0.000595)	0.0231(0.00103)	0.0224(0.506)	21.3(21.3)
(0,0,0)	1000	Y	6.49e-05(0.00027)	-0.00032(0.000358)	0.0727(0.00325)	0.128(0.00572)	0.0769(0.162)	8.28(8.28)
(1,1,1)	100	XYM	0.0155(0.00379)	0.00598(0.0045)	1.44(0.0643)	2.02(0.0905)	1.39(1.13)	73.4(73.4)
(1,1,1)	100	XY	0.0132(0.00346)	-0.00869(0.00528)	1.2(0.0535)	2.78(0.125)	1.23(1.21)	78.1(78.1)
(1,1,1)	100	XM	-0.033(0.00366)	-0.00466(0.00569)	1.34(0.0599)	3.24(0.145)	1.43(1.29)	62.1(62.1)
(1,1,1)	100	M	-0.0495(0.0036)	-0.00195(0.00502)	1.29(0.0579)	2.52(0.113)	1.35(1.19)	56.9(56.9)

(1,1,1)	100	Y	0.112(0.00395)	0.11(0.00469)	1.56(0.0699)	2.2(0.0985)	1.57(1)	63.5(63.5)
(1,1,1)	500	XYM	0.00586(0.00163)	-0.00301(0.00181)	1.32(0.0591)	1.65(0.0736)	1.26(2.43)	322(322)
(1,1,1)	500	XY	0.0154(0.00148)	0.00137(0.00175)	1.1(0.0492)	1.53(0.0683)	1.14(2.49)	343(343)
(1,1,1)	500	XM	-0.0269(0.00157)	0.000325(0.00247)	1.24(0.0553)	3.05(0.137)	1.26(2.57)	234(234)
(1,1,1)	500	M	-0.0453(0.00149)	-0.00169(0.00193)	1.12(0.0499)	1.86(0.0833)	1.13(2.63)	206(206)
(1,1,1)	500	Y	0.111(0.00167)	0.107(0.00192)	1.4(0.0626)	1.84(0.0823)	1.44(2.08)	268(268)
(1,1,1)	1000	XYM	0.00812(0.00112)	-0.00125(0.00122)	1.25(0.0558)	1.49(0.0669)	1.26(3.39)	644(644)
(1,1,1)	1000	XY	0.0115(0.00106)	-0.00175(0.00123)	1.13(0.0505)	1.51(0.0677)	1.12(3.57)	675(675)
(1,1,1)	1000	XM	-0.0276(0.00111)	-0.000538(0.00162)	1.23(0.0549)	2.64(0.118)	1.24(3.78)	451(451)
(1,1,1)	1000	M	-0.0464(0.00102)	-0.0035(0.00134)	1.05(0.0469)	1.81(0.0809)	1.09(3.57)	385(385)
(1,1,1)	1000	Y	0.113(0.0012)	0.109(0.00141)	1.45(0.0647)	1.98(0.0888)	1.44(2.92)	529(529)