Transporting stochastic direct and indirect effects to new populations

Kara E. Rudolph*1, Jonathan Levy2, and Mark J. van der Laan2

1Department of Emergency Medicine, School of Medicine, University of California, Davis, Sacramento, California
2Division of Biostatistics, School of Public Health, University of California, Berkeley, California

Abstract

Transported mediation effects may contribute to understanding how and why interventions may work differently when applied to new populations. However, we are not aware of any estimators for such effects. Thus, we propose several different estimators of transported stochastic direct and indirect effects: an inverse-probability of treatment stabilized weighted estimator, a doubly robust estimator that solves the estimating equation, and a doubly robust substitution estimator in the targeted minimum loss-based framework. We demonstrate their finite sample properties in a simulation study.

Keywords: Causal inference; Mediation; Transportability; Generalizability; External validity; Instrumental variables; Targeted maximum likelihood estimation

1 Introduction

Often, an intervention, program, or policy that works in one place or population fails to replicate in another place or population [21] or can even have unintended harmful effects [6]. This is problematic from a public policy or public health perspective in that the goals of such interventions are to help—not harm, and problematic from a financial perspective in that limited resources may be not be spent optimally.

When such initiatives fail to replicate or have unintended effects in new populations, transportability theory and methods offer a chance to understand why. Transportability is the ability (based on identifying assumptions) to transport a causal effect from a source population to a new, target population, accounting for differences between the two populations.
(e.g., differences in compositional factors, treatment adherence, etc.) (16). Previous work developed estimators to transport total effects from a source to target population (20) or, similarly, to generalize effects from a sample to the population (4, 5, 13, 23).

In some cases, examining transportability of the total effect may shed light on reasons for lack of replication. However, in other cases, transporting the total effect may not identify the relevant differences and it may be beneficial to go further and examine transportability of the underlying mediation mechanisms. Although there has been work on the identification on transported indirect effects (2, 16), we are not aware of any previous work developing estimators for transporting mediation effects (direct and indirect effects) from a source to target population. Thus, we address this research gap by proposing several different estimators of stochastic direct and indirect effects: an inverse-probability of treatment stabilized weighted estimator, a doubly robust estimator that solves the estimating equation, and a doubly robust substitution estimator in the targeted minimum loss-based framework.

The paper is organized as follows. In Section 2, we introduce notation and the structural equations model generating our data. In Section 3, we define the parameters of interest, the transported stochastic direct and indirect effects, and give identification results. Sections 4, 5, and 6, detail an inverse-probability of treatment weighted estimator, an estimating equation estimator, and a targeted minimum loss-based estimator for estimating these transported stochastic direct and indirect effects. In Section 7, we present the results of simulation studies that demonstrates the relative performance of the aforementioned estimators in finite samples. Section 8 concludes.

2 Notation and structural equations model

The full data is generated by a structural equations model (SEM) (17, 26), which consists of the data generating process to which we would like to have access. The SEM first generates a random draw of a vector $U$ of unknown measurements (15), where $U = (U_S, U_W, U_A, U_Z, U_M, U_Y) \sim P_U$. Then our variables are generated in the following time ordering:

$$
S = f_S(U_S)
$$
$$
W = f_W(U_W, S)
$$
$$
A = f_A(U_A, W, S)
$$
$$
Z = f_Z(U_Z, A, W, S)
$$
$$
M = f_M(U_M, Z, W, S)
$$
$$
Y = f_Y(U_Y, Z, W, M),
$$

where $S$ is a binary indicator of site, $W$ is a vector of covariates, $A$ is a binary treatment, $Z$ is a binary intermediate variable, $M$ is a binary mediator, and $Y$ is a binary or continuous outcome. The SEM generates the full data as $(U, O) \sim P_{UO} \in \mathcal{M}^F$, our full-data statistical model. If we had access to the SEM, we could generate potential (i.e., counterfactual) outcomes (14, 18), which define our causal parameters of interest. We observe data $O = (S, W, A, Z, M, S \times Y)$ for $n$ participants, with the true distribution $O_1, ..., O_n \sim P_O \in \mathcal{M}$,
our observed data statistical model. Note, we will only observe the outcome, $Y$, for site $S = 1$.

We consider a structural causal model for these data aligned with our motivating example in which $A$ is an instrumental variable (IV) \( \hat{I} \), which puts several restrictions on the statistical model, $\mathcal{M}$, where $\mathcal{M}$ is the collection of probability distributions under which the causal quantities of interest are identified—we define and identify such quantities in the following section. The restrictions are: 1) $A$ is randomly assigned (possibly conditional on \((W, S)\)), and 2) there is no direct effect of $A$ on $M$ or of $A$ on $Y$—downstream effects of $A$ only operate through $Z$ (22). However, the methods we propose can be used in the more general case where $A$ directly affects $M$ and/or $Y$; we discuss such an extension in the appendix.

In our notation, we use lowercase letters to denote fixed, assignment values of variables and uppercase letters to denote observed values. We use subscripts for descriptive purposes—subscripts are not to be considered a variable. For instance, we use a capital letter in $p_Y$, the conditional density of $Y$, because it is a density of the random variable $Y$.

### 3 Parameters of interest

We consider two causal quantities of interest that we call transported stochastic direct and indirect effects. These causal quantities represent stochastic direct and indirect effects (22, 25) transported from a source population to a new, target population. Stochastic direct and indirect effects, also called randomized interventional direct and indirect effects (25), represent the 1) direct effect of $A$ on $Y$ not through $M$ and the 2) indirect effect of $A$ on $Y$ through $M$. As has been described previously (19), one can consider versions of these effects that condition on $Z$ and thus estimate the indirect pathway of $A$ to $M$ to $Y$, not through $Z$ (27), or versions that marginalize over $Z$ and thus estimate the combined indirect pathways of 1) $A$ to $Z$ to $M$ to $Y$ plus 2) $A$ to $M$ to $Y$ (22, 25). Adhering to the IV constraints on our statistical model, no effect operates through pathway $A$ to $M$ to $Y$, so we focus herein on the versions of these effects that marginalize over $Z$. Previously, such a stochastic intervention on $M$ has been defined (22, 25)

$$
\hat{g}_{M|a^*, W}(M \mid W) = \sum_z \Pr(M \mid Z = z, W)\Pr(Z = z \mid A = a^*, W).
$$

The subscript for $\hat{g}_{M|a^*, W}$ specifies that it is a conditional density of random variable $M$ given random variable $W$, and value $a^*$ for which a lower case letter indicates they are fixed and the same for all participants. Previously, the parameter of interest was defined as $\Psi^F(P_{UO}) = \mathbb{E}\left[Y_{a, g_{M|a^*, W}}\right]$ where the expectation is taken over the full data model and $Y_{a, g_{M|a^*, W}}$ is a potential outcome intervening on $A$ to set it to $a$, and then downstream, intervening on $M$ to set it to a random (i.e., stochastic) draw from the distribution of $M$ defined by $g_{M|a^*, W}(M \mid W)$ (22). We wish to transport this parameter to a new site where the outcome was not observed ($S = 0$), and thus make the following modification:

$$
\Psi^F(P_{UO}) = \mathbb{E}\left[Y_{a, g_{M|a^*, W, s}} \mid S = 0\right]
$$

3
where

\[ g_{M|a^*,W,s}^*(M \mid W) = \sum_z Pr(M \mid Z = z, W, S = s)Pr(Z = z \mid A = a^*, W, S = s), \]

and where we impose a certain \( a^* \) and a certain \( s \) in both the \( Z \) and \( M \) models. We note that \( g_{M|a^*,W,s}^*(M \mid W) \) represents any stochastic intervention, which can include stochastic draws from the true models, but can also include a data-dependent version, estimated from observed data distributions, which we denote \( \hat{g}_{M|a^*,W,s}^*(M \mid W) \).

The transported stochastic direct effect entails setting \( a^* \) to 0 and taking the difference in mean outcome between setting \( a \) to 1 and setting \( a \) to 0, denoted

\[ \mathbb{E} \left[ Y_{1,g_{M|0,W,s}^*} - Y_{0,g_{M|0,W,s}^*} \mid S = 0 \right] \]

and the transported stochastic indirect effect entails setting \( a = 1 \) and then taking the difference in mean outcome between setting \( a^* = 1 \) and \( a^* = 0 \), denoted

\[ \mathbb{E} \left[ Y_{1,g_{M|1,W,s}^*} - Y_{1,g_{M|0,W,s}^*} \mid S = 0 \right]. \]

### 4 Identifiability

To identify the stochastic direct effect and stochastic indirect effect we will need to impose additional assumptions on \( \mathcal{M}^F \) and \( \mathcal{M} \), listed below.

1. Positivity: For all \( S \) and \( W \) we need a positive probability of assigning any level of \( A \). For all combinations of \( S, W, \) and \( A = a \), we have a positive probability of any level of \( Z \). For \( S = 1 \) and all combinations of \( Z \) and \( W \) we need a positive probability of any level of the mediator, \( M \).

2. Common outcome model across sites: \( \mathbb{E} [Y \mid M, Z, W, S = 1] = \mathbb{E} [Y \mid M, Z, W, S = 0] \). The null hypothesis of a common outcome model may be tested nonparametrically \([12]\).

3. Sequential Randomization: \( Y_{am} \perp A \mid W, S \) and \( Y_{am} \perp M \mid W, Z, S \). This is akin to a two-time point longitudinal intervention where at the first time point, we statically intervene to set the treatment, \( A = a \), and at the second time point, we stochastically intervene on the mediator, \( M \).

**Theorem 4.1.** Given the above assumptions, we can establish the following identification:

\[
\Psi(P) = \Psi^F(P_{UX}) = \mathbb{E} \left[ \mathbb{E} \left[ \mathbb{E} g_{M|a^*,W,s} \mathbb{E} [Y \mid W, Z, M, S = 1] \mid W, Z \right] \mid W, a, S = 0 \right] \mid S = 0
\]

\[
= \mathbb{E} \left[ \mathbb{E} \left[ \sum_m \mathbb{E} Y \hat{g}_{M|a^*,W,s}(m \mid W) \mid M = m, W, Z, A = a, S = 1 \right] \mid A = a, W, S \right] \mid S = 0
\]

The proof is in the appendix.
5 Stabilized Inverse probability of treatment weighted estimator

First, we describe a stabilized inverse probability of treatment weighted (IPTW) estimator of $\Psi(P)$. The R code to implement this estimator is provided in the appendix. We describe model fitting using regression language for simplicity but note that machine learning can be used instead. We will use the knowledge of our smaller model, $M$, with restrictions detailed in Section 2, so do not include $A$ in the regression model for $M$.

We use the weights

$$H(M, Z, A, W, S) = \frac{\hat{g}_{M|a^*, W, s}(M \mid W)p_Z(Z \mid A = a, W, S = 0) p_W(S = 0 \mid W) I(S = 1, A = a)}{p_M(M \mid Z, W, S = 1) p_Z(Z \mid A = a, W, S = 1) p_A(a \mid W, S = 1) p_W(S = 1 \mid W) P_S(S = 0)}. \tag{2}$$

$\hat{g}_{M|a^*, W, s}(M|W)$ is a data-dependent stochastic intervention on $M$, which can be estimated

$$\hat{g}_{M|a^*, W, s}(M|W) = \sum_{z=0}^{1} P(M = 1 \mid Z = z, W, S = s) P(Z = z \mid A = a^*, W, S = s). P(M = 1 \mid Z = z, W, S = s)$$

can be estimated using a logistic regression estimating the probability of $M = 1$ given $Z$, $W$, and $S$ and thereby getting predicted probabilities for $M = 1$ setting $S = s$ and separately setting $Z = 1$ and $Z = 0$. $P(Z = z \mid A = a^*, W, S = s)$ can be estimated using a logistic regression estimating the probability of $Z = 1$ given $A, W$, and $S$ and thereby getting predicted probabilities for $Z = 1$ and for $Z = 0$, setting $A = a^*$ and $S = s$ and using observed values for $W$. We note that marginalizing out $Z$ introduces dependence on $A$.

The stabilized IPTW estimate of $\Psi(P)$ is the empirical mean of $Y$ weighted by $\hat{H}_n(M, Z, A, W, S)$, stabilized by the empirical mean of $H(M, Z, A, W, S)$:

$$\hat{\Psi}_{IPTW}^{n} = \frac{\hat{H}_n(M_i, Z_i, A_i, W_i, S_i)Y_i}{\frac{1}{n} \sum_i \hat{H}_n(M_i, Z_i, A_i, W_i, S_i)}, \tag{3}$$

where $\hat{H}_n$ is an approximation of $H(M, Z, A, W, S)$.

The standard error of the stabilized IPTW estimator is calculated as the sample standard deviation of the approximation of the IPTW influence curve of $\Psi(P)$, given by

$$IC(P) = \frac{\hat{H}_n(M_i, Z_i, A_i, W_i, S_i)}{\frac{1}{n} \sum_i \hat{H}_n(M_i, Z_i, A_i, W_i, S_i)} \left( Y_i - \frac{1}{n} \sum_i \hat{H}_n(M_i, Z_i, A_i, W_i, S_i)Y_i \right).$$

The stochastic direct effect (SDE) entails setting $a^*$ to 0 and taking the difference in estimates between setting $a$ to 1 and setting $a$ to 0. The corresponding influence curve approximation is a difference of the influence curve approximation for the parameter defined by setting $a^* = 0, a = 1$ and the influence curve approximation for the parameter defined by setting $a^* = 0, a = 0$. The stochastic indirect effect (SIE) entails setting $a = 1$ and then taking the difference in estimates between setting $a^* = 1$ and $a^* = 0$. The corresponding influence curve approximation is again a difference of the two influence curve approximations (one for parameter defined by setting $a^* = 1, a = 1$ and the other for $a^* = 0, a = 1$). For each estimator, we used the sample standard deviation of their respective influence curve approximations, divided by $\sqrt{n}$ for the standard error estimate.
6 Estimating equation estimator

Next, we describe two estimating equation (EE) estimators of \( \Psi(P) \): 1) one that incorporates the exclusion restrictions on our statistical model, \( M \), that there is no direct effect of \( A \) on \( M \) or of \( A \) on \( Y \), and 2) another that does not impose those restrictions. The R code to implement these estimators is provided in the appendix. We describe model fitting using regression language for simplicity but note that machine learning can be used instead.

6.1 Estimator incorporating exclusion restrictions

This EE estimator solves the efficient influence curve (EIC) equation of the parameter \( \Psi(P) \) for the restricted model where \( M \) and \( Y \) do not depend directly on \( A \). This EIC is given by

\[
D(P) = D_Y(P) + D_Z(P) + D_W(P), \quad \text{where}
\]

\[
D_Y(P) = (Y - \bar{Q}_Y(M, Z, W))
\]

\[
\times \hat{g}_{M|a, W, s}(M|W)p_Z(Z \mid A = a, W, S = 0)p_W(S = 0 \mid W)I(S = 1)
\]

\[
g_M(M \mid Z, W, S = 1)p_Z(Z \mid W, S = 1)p_W(S = 1 \mid W)P_S(S = 0),
\]

\[
D_Z(P) = (\bar{Q}_M(Z, W, S) - \bar{Q}_Z(a, W, S)) \frac{I(S = 0, A = a)}{p_A(a \mid W, S)p_S(S = 0)}, \quad \text{and}
\]

\[
D_W(P) = (\bar{Q}_Z(a, W, S) - \Psi(P)) \frac{I(S = 0)}{p_S(S = 0)} \quad \text{(notation is explained further below)}.
\]

We reiterate that this EIC is only efficient for model with the exclusion restrictions. We approximate \( D_Y(P) \) as follows. First, we estimate \( \mathbb{E}[Y \mid M, Z, W] \) as \( \bar{Q}_{Y,n}(M, Z, W) \), which can be calculated using predicted values from a regression of \( Y \) on \( M, Z, W \) among those with \( S = 1 \). \( p_Z(Z \mid W, S = 1) \) can be calculated \( p_Z(Z \mid W, S = 1) = \sum_{a=0}^{1} P(Z = 1 \mid A = a, W, S = 1)P(A = a \mid W, S = 1) \). Each of the remaining probabilities in \( D_Y(P) \) can be calculated using predicted probabilities from a logistic regression of the relevant conditional mean outcome models for \( M, Z \) and \( S \), as described in the stabilized IPTW estimator section.

Estimation of \( \hat{g}_{M|a, W, s}(M|W) \) is also described in the stabilized IPTW estimator section.

We next approximate \( D_Z(P) \). To do so, we apply the stochastic intervention on \( \bar{Q}_{Y,n}(M, Z, W) \) via the computation \( \bar{Q}_{M,n}(M, Z, W, S) = \mathbb{E}_{g_{M|a^*, W, s}}[\bar{Q}_{Y,n}(M, Z, W) \mid Z, W, S] \).

Specifically, we can generate predicted values of \( \bar{Q}_{Y,n}(1, Z, W) \) and \( \bar{Q}_{Y,n}(0, Z, W) \) and then marginalize over \( g_{M|a^*, W, s}(M|W): \sum_{m=0}^{1} \bar{Q}_{Y,n}(m, Z, W)g_{M|a^*, W, s}(m|W) \). Next, we estimate \( \bar{Q}_{Z,n}(a, W, S) \) by regressing \( \bar{Q}_{M,n}(Z, W, S) \) on \( A, W, \) and \( S \), and getting predicted values setting \( A = a \).

The estimate of \( \Psi(P) \) is obtained by solving \( \frac{1}{n} \sum_i \hat{D}_\Psi(O_i) = 0 \), where \( \hat{D}_{\Psi,n} \) is the EIC approximation as computed above. The variance of the EE estimate is the sample variance of the EIC approximation. The SDE, SIE, and their corresponding ICs and standard errors can be calculated as described in the stabilized IPTW estimator section. The conditions for consistency and asymptotic efficiency of this estimator are discussed in the appendix.
6.2 Estimator not imposing exclusion restrictions

This EE estimator solves the efficient influence curve (EIC) equation of the parameter \( \Psi(P) \) for the unrestricted model where \( M \) and \( Y \) may depend directly on \( A \). It will be inefficient under the restricted model where \( M \) and \( Y \) do not depend directly on \( A \). This EIC is given by

\[
D(P) = D_Y(P) + D_Z(P) + D_W(P), \quad \text{where} \\
D_Y(P) = (Y - Q_Y(M, Z, A, W)) \\
\times \frac{\hat{g}_{M|a^*, W, s}(M|W)p_Z(Z | A = a, W, S = 0)p_W(S = 0 | W)I(S = 1, A = a)}{g_M(M | Z, A, W, S = 1)p_Z(Z | A = a, W, S = 1)p_A(a | W, S = 1)p_W(S = 1 | W)P_S(S = 0)}, \\
D_Z(P) = (\hat{Q}_M(Z, A, W, S) - \hat{Q}_Z(a, W, S)) \frac{I(S = 0, A = a)}{p_A(a | W, S)p_S(S = 0)}, \quad \text{and} \\
D_W(P) = (\hat{Q}_Z(a, W, S) - \Psi(P)) \frac{I(S = 0)}{p_S(S = 0)} \quad (\text{notation is explained further below}).
\]

\[(5)\]

We approximate \( D_Y(P) \) as follows. First, we estimate \( \mathbb{E}[Y | M, Z, A, W] \) as \( \hat{Q}_{Y,n}(M, Z, A, W) \), which can be calculated using predicted values from a regression of \( Y \) on \( M, Z, A, W \) among those with \( S = 1 \). Each of the probabilities in \( D_Y(P) \) can be calculated using predicted probabilities from a logistic regression of the relevant conditional mean outcome models for \( M, Z, A, \) and \( S \). \( \hat{g}_{M|a^*, W, s}(M|W) \) can be estimated \( \hat{g}_{M|a^*, W, s}(M|W) = \sum_{z=0}^{1} P(M = 1|Z = z, A = a^*, W, S = s)P(Z = z|A = a^*, W, S = s) \).

We next approximate \( D_Z(P) \). To do so, we apply the stochastic intervention on \( \hat{Q}_{Y,n}(M, Z, A, W) \) via the computation \( \hat{Q}_{M,n}(Z, A, W, S) = \mathbb{E}_{\hat{g}_{M|a^*, W, s}}[\hat{Q}_{Y,n}(M, Z, A, W) | Z, A, W] \). Specifically, we can generate predicted values of \( \hat{Q}_{Y,n}(1, Z, A, W) \) and \( \hat{Q}_{Y,n}(0, Z, A, W) \) and then marginalize over \( \hat{g}_{M|a^*, W, s}(M|W) \): \( \sum_{m=0}^{1} \hat{Q}_{Y,n}(m, Z, A, W)\hat{g}_{M|a^*, W, s}(m|W) \). Next, we estimate \( \hat{Q}_{Z,n}(a, W, S) \) by regressing \( \hat{Q}_{M,n}(Z, A, W, S) \) on \( A, W, \) and \( S \), and getting predicted values setting \( A = a \).

The estimate of \( \Psi(P) \) is obtained by solving \( \frac{1}{n} \sum_i \hat{D}_n(O_i) = 0 \), where \( \hat{D}_n \) is the EIC approximation as computed above. The variance of the EE estimate is the sample variance of the EIC approximation. The SDE, SIE, and their corresponding ICs and standard errors can be calculated as described in the stabilized IPTW estimator section. The conditions for consistency and asymptotic efficiency of this estimator are discussed in the appendix.

7 Targeted minimum loss-based estimator

We now describe how to estimate \( \Psi(P) \) using targeted minimum loss-based estimation (TMLE). This estimation approach, which is just one of several TMLE approaches that could be used, uses sequential regression, updating the conditional outcome model at each stage to both solve the EIC equation while also lowering the empirical negative log-likelihood loss. The process is identical to a two time-point longitudinal intervention (10). Similar to
the EE estimator section, we describe two TMLE estimators of \( \Psi(P) \): 1) one that incorporates the exclusion restrictions on our statistical model, \( M \), that there is no direct effect of \( A \) on \( M \) or of \( A \) on \( Y \), and 2) another that does not impose those restrictions. We include R code for these estimators in the appendix. As in the previous sections, we describe model fitting using regression language for simplicity but note that machine learning can be used instead.

### 7.1 Estimator incorporating exclusion restrictions

This TMLE estimator solves the EIC equation of the parameter \( \Psi(P) \) for the restricted model where \( M \) and \( Y \) do not depend directly on \( A \).

Let \( \tilde{Q}_{Y,n}^{x}(M, Z, W) \) be an initial estimate of \( \tilde{E}[Y \mid M, Z, W] \). \( Q_{Y,n}^{0}(M, Z, W) \) can be estimated by predicted values from a regression of \( Y \) on \( M, Z, W \) among those with \( S = 1 \).

Next, we update that initial estimate using the weights

\[
H(M, Z, W, S) = \frac{\hat{g}_{M|a^*,W,s}(M|W) p_{Z}(Z \mid A = a, W, S = 0) p_{W}(S = 0 \mid W) I(S = 1)}{g_{M}(M \mid Z, W, S = 1) p_{Z}(Z \mid W, S = 1) p_{W}(S = 1 \mid W) P_{S}(S = 0)},
\]

which are approximated with \( \tilde{H}_{n}(M, Z, W, S) \). \( \tilde{Q}_{Y,n}^{0}(M, Z, W) \) is updated by performing a weighted parametric logistic regression of \( Y \) with \( \text{logit}(\tilde{Q}_{Y,n}^{0}(M, Z, W)) \) as an offset, intercept \( \epsilon_{Y,n} \), and weights \( \tilde{H}_{n}(M, Z, W, S) \). \( \epsilon_{Y,n} \) is the MLE fit of intercept \( \epsilon_{Y} \). The update is given by

\[
\tilde{Q}_{Y,n}^{x}(M, Z, W) = \tilde{Q}_{Y,n}^{0}(M, Z, W) \tilde{H}_{n}(M, Z, W, S).
\]

We then perform the stochastic intervention on \( \tilde{Q}_{Y,n}^{x}(M, Z, W) \) via the computation

\[
\tilde{Q}_{M,n}^{x}(Z, W, S) = \mathbb{E}_{g_{M|a^*,W,s}}[\tilde{Q}_{Y,n}^{x}(M, Z, W) \mid Z, W, S].
\]

This can be done by generating predicted values of \( \tilde{Q}_{Y,n}^{x}(1, Z, W) \) and \( \tilde{Q}_{Y,n}^{x}(0, Z, W) \) and then marginalizing over \( \hat{g}_{M|a^*,W,s}(M|W) = \sum_{m=0}^{1} \tilde{Q}_{Y,n}^{x}(m, Z, W) \tilde{H}_{n}(m, Z, W, S) \).

Next, we estimate \( \tilde{Q}_{Z,n}^{x}(a, W, S) \) by regressing \( \tilde{Q}_{M,n}^{x}(Z, W, S) \) on \( A, W, \) and \( S \), and getting predicted values setting \( A = a \). We then update this initial estimate using a second set of weights,

\[
H_{a}(a, W, S) = \frac{I(S = 0, A = a)}{P_{A}(A \mid W, S) P_{S}(S = 0)};
\]

in a weighted logistic regression of \( \text{logit}(\tilde{Q}_{M,n}^{x}(Z, W, S)) \) with \( \text{logit}(\tilde{Q}_{Z,n}^{x}(a, W, S)) \) as an offset, intercept \( \epsilon_{Z,n} \). \( \epsilon_{Z,n} \) is the MLE fit of intercept \( \epsilon_{Z} \). The updated estimate will be notated \( \tilde{Q}_{Z,n}^{x}(a, W, S) = \tilde{Q}_{Z,n}^{0}(a, W, S) \).

The empirical mean of \( \tilde{Q}_{Z,n}^{x}(a, W, S) \) among those for whom \( S = 0 \) is the TMLE estimate of \( \Psi(P) \). The influence curve approximation obtained by incorporating our updated model fits solves the empirical mean of the EIC equation (see Equation 6 in Section 6), replacing \( P \) with \( P_{n}^{*} \) defined by \( \tilde{Q}_{Y,n}^{x}(M, Z, W) \), \( \tilde{Q}_{M,n}(Z, W, S) \), \( \tilde{Q}_{Z,n}(a, W, S) \), our known \( P_{A}(A \mid W, S) \) and \( P_{S}(S = 0) \), our regression fit for \( P_{M}(M \mid Z, W, S) \), \( P_{Z}(Z \mid A, W, S) \) and \( P_{S|W}(S \mid W) \), and the empirical distribution as an estimate of \( P_{W}(W) \). We denote this EIC \( D_{n}^{*} \). The TMLE updating step also decreases the empirical loss of the model fits. The variance of the TMLE estimate is the sample variance of the EIC approximation. The SDE, SIE, and their corresponding ICs and standard errors can be calculated as described in the stabilized IPTW estimator section. The conditions for consistency and asymptotic efficiency of this estimator are discussed in the appendix.
7.2 Estimator not imposing exclusion restrictions

This TMLE estimator solves the EIC equation of the parameter $\Psi(P)$ for the unrestricted model where $M$ and $Y$ may depend directly on $A$. It will be inefficient under the restricted model where $M$ and $Y$ do not depend directly on $A$.

Let $\bar{Q}^0_{Y,n}(M, Z, A, W)$ be an initial estimate of $E[Y \mid M, Z, A, W]$. $\bar{Q}^0_{Y,n}(M, Z, A, W)$ can be estimated by predicted values from a regression of $Y$ on $M, Z, A, W$ among those with $S = 1$.

Next, we update that initial estimate using the weights, $H(M, Z, A, W, S)$, given in Equation 2, which are approximated with $\hat{H}_n(M, Z, A, W, S)$. $\bar{Q}^0_{Y,n}(M, Z, A, W)$ is updated by performing a weighted parametric logistic regression of $Y$ with $\logit(\bar{Q}^0_{Y,n}(M, Z, A, W))$ as an offset, intercept $\epsilon_Y$, and weights $\hat{H}_n(M, Z, A, W, S)$. $\epsilon_{Y,n}$ is the MLE fit of intercept $\epsilon_Y$.

The update is given by $\bar{Q}^*_{Y,n}(M, Z, A, W) = \bar{Q}^0_{Y,n}(\epsilon_{Y,n})(M, Z, A, W)$.

We then perform the stochastic intervention on $\bar{Q}^*_{Y,n}(M, Z, A, W)$ via the computation $\bar{Q}^*_{M,n}(Z, A, W, S) = \mathbb{E}_{\hat{g}_{M|a^*,W,S}}[\bar{Q}^*_{Y,n}(M, Z, A, W) \mid Z, A, W, S]$. This can be done by generating predicted values of $\hat{Q}^*_{Y,n}(1, Z, A, W)$ and $\hat{Q}^*_{Y,n}(0, Z, A, W)$ and then marginalizing over $\hat{g}_{m|a^*,W,S}(M|W)$: $\sum_{m=0}^1 \hat{Q}^*_{Y,n}(m, Z, A, W)\hat{g}_{m|a^*,W,S}(m|W)$.

Next, we estimate $\bar{Q}^*_{Z,n}(a, W, S)$ by regressing $\hat{Q}^*_{M,n}(Z, A, W, S)$ on $A, W, S$, and getting predicted values setting $A = a$. The remainder of the steps for this TMLE are identical to those for the restricted TMLE in the above subsection.

8 Simulation

8.1 Overview

We compare finite sample performance of our three types estimators in estimating the transported SDE and transported SIE using simulation. For both TMLE and EE, we include the 1) versions that are efficient under the restricted model in which $M$ and $Y$ do not depend directly on $A$ (henceforth TMLE efficient and EE efficient) and 2) versions that are more general in that they allow for the exclusion restriction not to hold but are inefficient under restricted model $\mathcal{M}$ (henceforth TMLE general and EE general). We show estimator performance in terms of absolute bias, efficiency, 95% confidence interval (CI) coverage, root mean squared error (RMSE), and percent of estimates lying outside the bounds of the parameter space across 1,000 simulations. For calculating the efficiency and the 95% CI coverage, we use both the IC and 500 bootstrap replicates.

We consider two data-generating mechanisms (DGMs) within the structural causal model described in Section 2. The DGMs are detailed in Table 1 using the same notation as in Section 2.

8.2 Results

First, in Table 2 we show results under correct specification of all models for sample sizes of $N=5,000$, $N=500$, and $N=100$ using DGM 1. Results using alternative DGMs were similar and are shown in the appendix.
Table 1: Simulation data-generating mechanisms.

| Data Generating Mechanism 1 | Data Generating Mechanism 2 |
|-----------------------------|-----------------------------|
| \( W_1 \sim \text{bernoulli} \) | \( W_1 \sim \text{bernoulli} \) |
| \( P(W_1 = 1) = 0.5 \) | \( P(W_1 = 1) = 0.5 \) |
| \( W_2 \sim \text{bernoulli} \) | \( W_2 \sim \text{bernoulli} \) |
| \( P(W_2 = 1) = \text{expit}(0.4 + 0.2W_1) \) | \( P(W_2 = 1) = \text{expit}(0.4 + 0.2W_1) \) |
| \( S \sim \text{bernoulli} \) | \( S \sim \text{bernoulli} \) |
| \( P(S = 1) = \text{expit}(3W_2 - 1) \) | \( P(S = 1) = \text{expit}(3W_2 - 1) \) |
| \( A \sim \text{bernoulli} \) | \( A \sim \text{bernoulli} \) |
| \( P(A = 1) = 0.5 \) | \( P(A = 1) = 0.5 \) |
| \( Z \sim \text{bernoulli} \) | \( Z \sim \text{bernoulli} \) |
| \( P(Z = 1) = \text{expit}(-0.1A - 0.2S + 0.2W_2 + 0.14AS + 0.2W_2S - 0.2AW_2S - 1) \) | \( P(Z = 1) = \text{expit}(-0.1A - 0.2S + 0.2W_2 + 0.14AS + 0.2W_2S - 0.2AW_2S - 1) \) |
| \( M \sim \text{bernoulli} \) | \( M \sim \text{bernoulli} \) |
| \( P(M = 1) = \text{expit}(1Z + 3ZW_2 + 0.2ZS - 0.2W_2S + 2W_2Z + 0.2S - 0.2ZW_2S - W_2 - 2) \) | \( P(M = 1) = \text{expit}(1Z + 3ZW_2 + 0.2ZS - 0.2W_2S + 2W_2Z + 0.2S - 0.2ZW_2S - W_2 - 2) \) |
| \( Y \sim \text{bernoulli} \) | \( Y \sim \text{bernoulli} \) |
| \( P(Y = 1) = \text{expit}(-6Z + 0.2ZW_2 + 2ZM + 2W_2M - 2W_2 + 4M + 1ZW_2M - 0.2) \) | \( P(Y = 1) = \text{expit}(\log(1.2) + \log(40)Z - \log(30)M - \log(1.2)W_2 - \log(40)W_2Z) \) |

Table 2 shows that under correct parametric model specifications, all estimators result in consistent estimates with 95% CI coverage close to 95% for sample sizes of \( N=5,000 \) and \( N=500 \). Influence curve-based efficiency is close to 100% of the efficiency bound for both the efficient TMLE and EE estimators for all sample sizes. For the transported direct effect, efficiency is about 3 times the efficiency bound for the general TMLE and EE estimators and about 4 times the efficiency bound for the stabilized IPTW estimator. For the transported indirect effect, IC- and bootstrapped-based efficiencies are similar and close to the efficiency bound for the TMLE and EE estimators. The stabilized IPTW estimator has efficiencies about 70% larger than the efficiency bound. In the appendix, we review the theory underlying the empirical finding that the stabilized IPTW estimator is least efficient.

Performance of all estimators degrades with a smaller sample size of \( N=100 \). For this sample size, bias slightly increases, the estimates stray from the efficiency bound, and IC-based coverage generally decreases (though bootstrapping can recover some of this lost coverage by approximating the true variance of the estimator).
Table 2: Simulation results comparing estimators of the transported stochastic direct effect and transported stochastic indirect effect under DGM 1 and correct model specification for various sample sizes. 1,000 simulations. Estimation methods compared include stabilized inverse probability weighting estimation (IPTW), solving the estimating equation (EE), and targeted minimum loss-based estimation (TMLE). For TMLE and EE, we compare versions of the estimators that incorporate the exclusion restrictions in our statistical model (TMLE efficient, EE efficient) and versions that do not (TMLE general, EE general). Efficiency and 95% CI coverage are calculated separately using 1) the influence curve (IC) and 2) bootstrapping (boot). Bias and RMSE values are averages across the simulations.

| Estimator          | Bias | Efficiency | 95% CI Coverage | RMSE | % Out of Bounds |
|--------------------|------|------------|-----------------|------|-----------------|
|                    |      | IC         | Boot            |      |                 |
|                    |      |            |                 |      |                 |
| **Transported stochastic direct effect** |      |            |                 |      |                 |
| N=5000             |      |            |                 |      |                 |
| TMLE efficient     | 0.000| 100.17     | 100.42          | 0.955| 0.957           | 0.008| 0               |
| TMLE general       | 0.001| 321.02     | 321.92          | 0.945| 0.942           | 0.027| 0               |
| EE efficient       | 0.000| 100.21     | 100.42          | 0.955| 0.957           | 0.008| 0               |
| EE general         | 0.001| 321.27     | 321.00          | 0.946| 0.942           | 0.027| 0               |
| IPTW               | 0.000| 412.67     | 388.11          | 0.954| 0.945           | 0.033| 0               |
| N=500              |      |            |                 |      |                 |
| TMLE efficient     | -0.004| 120.17    | 105.16          | 0.946| 0.957           | 0.028| 0               |
| TMLE general       | -0.004| 319.08    | 327.43          | 0.929| 0.936           | 0.089| 0               |
| EE efficient       | 0.000| 101.40     | 104.51          | 0.946| 0.957           | 0.028| 0               |
| EE general         | -0.004| 321.97    | 316.56          | 0.947| 0.935           | 0.088| 0               |
| IPTW               | -0.001| 428.43    | 419.16          | 0.951| 0.942           | 0.110| 0               |
| N=100              |      |            |                 |      |                 |
| TMLE efficient     | 0.003| 102.26     | 139.52          | 0.970| 0.992           | 0.064| 0               |
| TMLE general       | 0.007| 293.07     | 375.10          | 0.852| 0.945           | 0.218| 0               |
| EE efficient       | 0.004| 102.13     | 118.65          | 0.974| 0.989           | 0.060| 0               |
| EE general         | 0.005| 316.03     | 292.87          | 0.957| 0.930           | 0.186| 0               |
| IPTW               | 0.009| 442.80     | 456.03          | 0.909| 0.922           | 0.267| 0               |
| **Transported stochastic indirect effect** |      |            |                 |      |                 |
| N=5000             |      |            |                 |      |                 |
| TMLE efficient     | 0.000| 99.98      | 100.46          | 0.942| 0.942           | 0.004| 0               |
| TMLE general       | 0.000| 101.75     | 101.99          | 0.942| 0.942           | 0.004| 0               |
| EE efficient       | 0.000| 100.01     | 100.46          | 0.942| 0.942           | 0.004| 0               |
| EE general         | 0.000| 101.69     | 101.95          | 0.942| 0.943           | 0.004| 0               |
| IPTW               | 0.000| 166.56     | 141.45          | 0.980| 0.957           | 0.005| 0               |
| N=500              |      |            |                 |      |                 |
| TMLE efficient     | -0.001| 99.56     | 104.59          | 0.926| 0.929           | 0.012| 0               |
| TMLE general       | -0.001| 101.74    | 106.99          | 0.928| 0.935           | 0.012| 0               |
| EE efficient       | -0.001| 99.58     | 101.11          | 0.926| 0.929           | 0.012| 0               |
| EE general         | -0.001| 101.94    | 103.39          | 0.932| 0.930           | 0.012| 0               |
| IPTW               | -0.001| 177.68    | 156.98          | 0.948| 0.936           | 0.018| 0               |
| N=100              |      |            |                 |      |                 |
| TMLE efficient     | -0.003| 93.01     | 132.95          | 0.878| 0.944           | 0.033| 0               |
| TMLE general       | -0.001| 99.85     | 158.57          | 0.865| 0.952           | 0.041| 0               |
| EE efficient       | -0.003| 93.03     | 113.13          | 0.878| 0.924           | 0.033| 0               |
| EE general         | -0.002| 99.09     | 117.32          | 0.871| 0.915           | 0.036| 0               |
| IPTW               | 0.002| 167.51     | 206.17          | 0.812| 0.867           | 0.062| 0               |
Next, we show results under various model misspecifications in Table 3 for sample size N=5,000. We consider: 1) misspecification of the Y model, 2) misspecification of the Y and Z models, 3) misspecification of the Y and M models, 4) misspecification of the Y and S models, and 5) misspecification of the Z, M, and S models. We use DGM 2 for misspecification of the Y model and misspecification of the Y and M models. We use DGM 1 for misspecification of the Y and Z models, misspecification of the Y and S models, and misspecification of the Z, M, and S models. Full results for misspecified models under more DGMs and various sample sizes are shown in the appendix.

As derived from the robustness properties proven in the appendix, we expect the TMLE and EE estimators to be consistent if 1) the Y model is correctly specified or if 2) the Z, M, and S models are correctly specified. The stabilized IPTW estimator will be consistent only in the latter case: where the Z, M, and S models are correctly specified. Note that we assume A is randomly assigned, so we don’t examine misspecification of the A model. However, in cases where A is nonrandom, condition 2 above would change to require correct specification of the A, Z, M, and S models.

We see in Table 3 that, as expected, all estimators remain unbiased when the Y model is misspecified. Compared to the correctly specified case (Table 2, there is a slight reduction in efficiency, evidenced by efficiencies greater than 100% of the efficiency bound and slight overcoverage when using the IC as opposed to bootstrapping (Table 3).

When the Z, M, and S models are misspecified, the TMLE and EE estimators also remain unbiased, as expected, but the stabilized IPTW estimator is no longer consistent, resulting in 0% coverage of the transported SDE (Table 4). However, in this scenario, coverage of the TMLE and EE efficient estimators is anticonservative when using the IC, evidenced by efficiencies as low as 55% of the efficiency bound in for the transported SIE and coverage as low as 73%. This result is not unexpected—the EIC may no longer provide accurate inference under these misspecifications, because the misspecified models may bias the influence curve approximation. In this particular misspecified scenario, our influence curve approximation becomes a less variant mean-zero function of the observed data than our efficiency bound given by the true influence curve, causing poor coverage. In such instances of model misspecification, using the bootstrap can recover coverage (as long the fitting methods applied yield valid standard errors under the non-parametric bootstrap. For linear regressions, this is the case, but for data adaptive machine learning algorithms, such might not be the case (24)).

In the remaining model misspecifications shown in Table 3, none of the estimators are guaranteed to be unbiased. In the case where the Y and M models are misspecified, all three estimators perform poorly in estimating both the transported direct and indirect effects. Performances of the EE estimators are particularly poor in terms of bias and inefficiency. The marked inefficiency may be due, in part, to estimates lying outside the bounds of the parameter space. Indeed, we see that nearly 100% of the EE transported SDE estimates lie outside the parameter space. In contrast, even though the TMLE estimators solve the same influence curve, none of the TMLE estimates lie outside the parameter space, demonstrating its advantage as a substitution estimator.
Table 3: Simulation results comparing estimators of transported stochastic direct effect and transported stochastic indirect effect under various model misspecifications for sample size N=5,000. 1,000 simulations. Estimation methods compared include stabilized inverse probability weighting estimation (IPTW), solving the estimating equation (EE), and targeted minimum loss-based estimation (TMLE). For TMLE and EE, we compare versions of the estimators that incorporate the exclusion restrictions in our statistical model (TMLE efficient, EE efficient) and versions that do not (TMLE general, EE general). Bias and RMSE values are averages across the simulations. Efficiency and 95% CI coverage are calculated separately using 1) the influence curve (IC) and 2) bootstrapping (boot). The DGM used for each misspecification is noted below.

| Estimator                        | Bias     | Efficiency  | 95% CI Coverage | RMSE    | % Out of Bounds |
|----------------------------------|----------|-------------|------------------|---------|-----------------|
|                                  | IC       | Boot        | IC               | Boot    |                 |
| Y model misspecified, DGM 2      |          |             |                  |         |                 |
| Transported stochastic direct effect |          |             |                  |         |                 |
| TMLE efficient                   | 0.000    | 191.06      | 118.82           | 0.999   | 0.965           | 0.014 | 0               |
| TMLE general                     | 0.000    | 354.81      | 288.79           | 0.984   | 0.954           | 0.035 | 0               |
| EE efficient                     | 0.000    | 240.26      | 126.67           | 0.999   | 0.968           | 0.015 | 0               |
| EE general                       | 0.000    | 374.95      | 291.49           | 0.992   | 0.957           | 0.035 | 0               |
| IPTW                             | 0.000    | 300.52      | 276.72           | 0.973   | 0.957           | 0.034 | 0               |
| Transported stochastic indirect effect |          |             |                  |         |                 |
| TMLE efficient                   | 0.000    | 186.27      | 102.20           | 0.995   | 0.914           | 0.002 | 0               |
| TMLE general                     | 0.000    | 239.35      | 186.41           | 0.986   | 0.942           | 0.004 | 0               |
| EE efficient                     | 0.000    | 237.87      | 103.75           | 0.997   | 0.919           | 0.002 | 0               |
| EE general                       | 0.000    | 302.49      | 216.48           | 0.991   | 0.945           | 0.004 | 0               |
| IPTW                             | 0.000    | 223.45      | 174.48           | 0.988   | 0.941           | 0.003 | 0               |
| Y and Z models misspecified, DGM 1 |          |             |                  |         |                 |
| Transported stochastic direct effect |          |             |                  |         |                 |
| TMLE efficient                   | 0.032    | 75.49       | 77.35            | 0.000   | 0.000           | 0.032 | 0               |
| TMLE general                     | 0.188    | 973.02      | 463.03           | 0.223   | 0.003           | 0.192 | 0               |
| EE efficient                     | 0.033    | 80.22       | 81.92            | 0.000   | 0.000           | 0.034 | 0               |
| EE general                       | 0.545    | 1,112.09    | 1,008.79         | 0.000   | 0.000           | 0.552 | 0               |
| IPTW                             | 0.108    | 700.28      | 456.29           | 0.574   | 0.204           | 0.115 | 0               |
| Transported stochastic indirect effect |          |             |                  |         |                 |
| TMLE efficient                   | -0.014   | 192.60      | 92.53            | 0.380   | 0.004           | 0.014 | 0               |
| TMLE general                     | -0.040   | 492.45      | 143.03           | 0.185   | 0.000           | 0.040 | 0               |
| EE efficient                     | 0.018    | 307.60      | 121.67           | 0.767   | 0.029           | 0.019 | 0               |
| EE general                       | 0.066    | 911.65      | 686.60           | 0.447   | 0.231           | 0.072 | 0               |
| IPTW                             | -0.031   | 439.30      | 160.87           | 0.430   | 0.004           | 0.031 | 0               |
| Y and M models misspecified, DGM 2 |          |             |                  |         |                 |
| Transported stochastic direct effect |          |             |                  |         |                 |
| TMLE efficient                   | -0.272   | 280.08      | 149.17           | 0.000   | 0.000           | 0.273 | 0               |
| TMLE general                     | -0.272   | 698.69      | 321.77           | 0.147   | 0.001           | 0.276 | 0               |
| EE efficient                     | -1.110   | 852.67      | 735.38           | 0.000   | 0.000           | 1.118 | 99.2            |
| EE general                       | -1.111   | 1,093.33    | 877.83           | 0.000   | 0.000           | 1.121 | 98.4            |
| IPTW                             | -0.236   | 589.90      | 284.02           | 0.090   | 0.000           | 0.240 | 0               |
| Transported stochastic indirect effect |          |             |                  |         |                 |
| TMLE efficient                   | 0.104    | 249.26      | 173.89           | 0.000   | 0.000           | 0.106 | 0               |
| TMLE general                     | 0.103    | 370.51      | 279.70           | 0.036   | 0.016           | 0.106 | 0               |

Continued on next page
### Table 3 – continued from previous page

| Estimator       | Bias | Efficiency | 95% CI Coverage | RMSE | % Out of Bounds |
|-----------------|------|------------|-----------------|------|-----------------|
|                 |      | IC        | Boot            | IC   | Boot            |
| EE efficient    | 0.135| 336.39     | 227.03          | 0.000| 0.000           | 0.137 | 0              |
| EE general      | 0.137| 535.91     | 413.33          | 0.054| 0.016           | 0.141 | 0              |
| IPTW            | 0.061| 279.01     | 200.78          | 0.197| 0.090           | 0.063 | 0              |
| **Transported stochastic direct effect** |      |            |                 |      |                 |
| TMLE efficient  | -0.164| 285.91     | 168.14          | 0.000| 0.000           | 0.165 | 0              |
| TMLE general    | -0.164| 511.22     | 399.63          | 0.019| 0.010           | 0.167 | 0              |
| EE efficient    | -0.220| 415.64     | 191.91          | 0.000| 0.000           | 0.220 | 0              |
| EE general      | -0.218| 699.90     | 586.82          | 0.029| 0.023           | 0.224 | 0              |
| IPTW            | -0.067| 404.82     | 322.51          | 0.467| 0.319           | 0.073 | 0              |
| **Transported stochastic indirect effect** |      |            |                 |      |                 |
| TMLE efficient  | 0.222 | 507.95     | 186.55          | 0.000| 0.000           | 0.223 | 0              |
| TMLE general    | 0.222 | 600.12     | 343.91          | 0.000| 0.000           | 0.222 | 0              |
| EE efficient    | 0.278 | 1,033.84   | 216.14          | 0.000| 0.000           | 0.279 | 0              |
| EE general      | 0.278 | 1,201.77   | 675.26          | 0.000| 0.000           | 0.279 | 0              |
| IPTW            | 0.278 | 520.31     | 337.87          | 0.000| 0.000           | 0.278 | 0              |
| **Z, M, and S models misspecified, DGM 1** |      |            |                 |      |                 |
| Transformed stochastic direct effect |      |            |                 |      |                 |
| TMLE efficient  | 0.000 | 98.78      | 100.01          | 0.956| 0.960           | 0.008 | 0              |
| TMLE general    | 0.001 | 228.37     | 231.43          | 0.939| 0.941           | 0.020 | 0              |
| EE efficient    | 0.000 | 98.80      | 100.03          | 0.956| 0.960           | 0.008 | 0              |
| EE general      | 0.001 | 228.50     | 203.24          | 0.966| 0.942           | 0.018 | 0              |
| IPTW            | 0.254 | 396.11     | 406.66          | 0.000| 0.000           | 0.256 | 0              |
| Transformed stochastic indirect effect |      |            |                 |      |                 |
| TMLE efficient  | 0.001 | 54.95      | 100.51          | 0.727| 0.938           | 0.010 | 0              |
| TMLE general    | 0.001 | 129.68     | 136.04          | 0.879| 0.920           | 0.014 | 0              |
| EE efficient    | 0.000 | 54.95      | 100.46          | 0.732| 0.939           | 0.010 | 0              |
| EE general      | 0.001 | 130.08     | 137.03          | 0.891| 0.930           | 0.014 | 0              |
| IPTW            | 0.014 | 301.92     | 123.87          | 0.976| 0.716           | 0.018 | 0              |

### 9 Conclusion

In this paper, we defined and identified parameters that transport stochastic direct and indirect mediating effects from a source population (\( S = 1 \)) to a new, target population (\( S = 0 \)). Identification of such parameters rely on the typical sequential randomization and positivity assumptions of other stochastic mediation effects \([22, 25, 27]\) as well as a common outcome model assumption, described previously for transport estimators \([20]\), which can be tested nonparametrically \([12]\). Such parameters enable the prediction of mediating effects in new populations based on data about the mediation mechanism in a source population and the differing distributions of compositional characteristics between the two populations. Thus, transport SDE and SIE parameters contribute to understanding the how and why interventions may work differently and/or have differing effects when applied to new populations.

We proposed three estimation approaches for such effects: a stabilized IPTW estimator, an estimating equation approach, and a TMLE approach. For the EE and TMLE approaches, we describe a version of each estimator that is efficient under a statistical model...
with exclusion restrictions such that $A$ does not directly affect $M$ or $Y$, and another version that is efficient under a statistical model without those restrictions such that $A$ can directly affect $M$ and $Y$. The EE and TMLE estimators solve the EIC for a particular statistical model, which results in double robustness, meaning that they are unbiased if one either consistently estimates the $Y$ model or $Z, M, S$ models. The TMLE estimator has the additional advantage of staying within the bounds of the parameter space by virtue of being a substitution estimator. We demonstrated the finite sample advantage of staying within the parameter bounds in a simulation study; in a particularly challenging scenario, nearly all of the EE estimates were outside of the parameter space but the TMLE estimates stayed within bounds. The simulation study also demonstrated that, as expected, the stabilized IPTW estimator is less efficient than the EE and TMLE estimators under correct model specification and in many scenarios where models are misspecified.

We also saw empirical evidence that, even in cases where robustness properties guarantee estimator consistency, the EIC may no longer provide accurate inference when models are misspecified. If parametric models are used, as in our simulation, the bootstrap can be used to recover appropriate coverage in such scenarios. However, the bootstrap is not an appropriate strategy if data-adaptive methods are used in model fitting. An alternative approach is to target the nuisance parameters of the IC (e.g., $g_M, g_Z, g_S$, etc.) in addition to the parameter of interest in a TMLE. Applying such extra targeting to TMLE estimators we describe here is an area for future work.

To facilitate use of these methods, we include commented R code for implementing them in the appendix.

The estimators we propose are limited in that they consider a stochastic intervention on mediator, $M$, that is assumed known and estimated from observed data. However, we plan to extend them to a true, unknown stochastic intervention in the future. Another limitation is that the parameters are only identified if one assumes a common outcome model across the source and target populations. There will be some research questions for which it is not possible to establish evidence for or against this assumption, as in questions about predicting a long-term outcome in a new population. However, when the research question instead focuses on establishing the extent to which mechanisms are shared across populations, and the full set of data $O = (S, W, A, Z, M, Y)$ is observed for both populations, one can empirically test whether there is evidence against such a shared outcome model.

In the main text, we focused on transporting mediation estimates where an instrument, $A$, was statically intervened on and mediator $M$ was stochastically intervened on. Moreover, we were primarily concerned with a statistical model that imposed instrumental variable assumptions such as the exclusion restriction assumption. However, we describe how each estimator can be easily modified to accommodate statistical models that do not impose instrumental variable assumptions, allowing for a direct effect of $A$ on $M$ and of $A$ on $Y$. Extending our proposed estimators for data generating mechanisms that do not include an intermediate variable, $Z$ is straightforward. Thus, our transport mediation estimators can be applied to a wide-range of common data generating mechanisms.
References

[1] Angrist, J. D., Imbens, G. W. and Rubin, D. B. (1996) Identification of causal effects using instrumental variables. *Journal of the American statistical Association*, **91**, 444–455.

[2] Bareinboim, E. and Pearl, J. (2016) A general algorithm for deciding transportability of experimental results. *Journal of causal Inference*, **1**, 107–134.

[3] Benkeser, D., Carone, M., Laan, M. V. D. and Gilbert, P. (2017) Doubly robust non-parametric inference on the average treatment effect. *Biometrika*, **104**, 863–880.

[4] Cole, S. R. and Stuart, E. A. (2010) Generalizing evidence from randomized clinical trials to target populations: The actg 320 trial. *American journal of epidemiology*, **172**, 107–115.

[5] Frangakis, C. (2009) The calibration of treatment effects from clinical trials to target populations. *Clinical trials (London, England)*, **6**, 136.

[6] Kling, J. R., Liebman, J. B. and Katz, L. F. (2007) Experimental analysis of neighborhood effects. *Econometrica*, **75**, 83–119.

[7] van der Laan, M. (2016) A generally efficient targeted minimum loss based estimator. *U.C. Berkeley Division of Biostatistics Working Paper Series*, **343**. URL http://biostats.bepress.com/ucbbiostat/paper343.

[8] van der Laan, M. and Rose, S. (2011) *Targeted Learning*. New York: Springer.

[9] van der Laan, M. and Rubin, D. (2006) Targeted maximum likelihood learning. *U.C. Berkeley Division of Biostatistics Working Paper Series*. URL http://biostats.bepress.com/ucbbiostat/paper213.

[10] van der Laan, M. J. and Gruber, S. (2012) Targeted minimum loss based estimation of causal effects of multiple time point interventions. *The international journal of biostatistics*, **8**.

[11] Levy, J. (2019) Tutorial: Deriving the efficient influence curve for large models. *arXiv preprint arXiv:1903.01706*.

[12] Luedtke, A. R., Carone, M. and van der Laan, M. J. (2015) An omnibus nonparametric test of equality in distribution for unknown functions. *arXiv preprint arXiv:1510.04195*.

[13] Miettinen, O. S. (1972) Standardization of risk ratios. *American Journal of Epidemiology*, **96**, 383–388.

[14] Neyman, J. S. (1923) On the application of probability theory to agricultural experiments. essay on principles. section 9.(translated and edited by dm dabrowska and tp speed, statistical science (1990), 5, 465-480). *Annals of Agricultural Sciences*, **10**, 1–51.

[15] Pearl, J. (2009) *Causality*. Cambridge university press.
[16] Pearl, J. and Bareinboim, E. (2014) External validity: From do-calculus to transportability across populations. *Statistical Science*, 579–595.

[17] Pearl, J. et al. (2009) Causal inference in statistics: An overview. *Statistics surveys*, 3, 96–146.

[18] Rubin, D. B. (1974) Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology*, 66, 688.

[19] Rudolph, K. E., Goin, D. E., Paksarian, D., Crowder, R., Merikangas, K. R. and Stuart, E. A. (2018) Causal mediation analysis with observational data: Considerations and illustration examining mechanisms linking neighborhood poverty to adolescent substance use. *American journal of epidemiology*.

[20] Rudolph, K. E. and van der Laan, M. J. (2017) Robust estimation of encouragement-design intervention effects transported across sites. *Journal of the Royal Statistical Society Series B Statistical Methodology*, 79, 1509–1525.

[21] Rudolph, K. E., Schmidt, N. M., Crowder, R., Galin, J., Glymour, M. M., Ahern, J. and Osypuk, T. L. (2018) Composition or context: using transportability to understand drivers of site differences in a large-scale housing experiment. *Epidemiology*, 29, 199–206.

[22] Rudolph, K. E., Sofrygin, O., Zheng, W. and van der Laan, M. J. (2017) Robust and flexible estimation of data-dependent stochastic mediation effects: a proposed method and example in a randomized trial setting. *Epidemiologic Methods*, In Press.

[23] Stuart, E. A., Cole, S. R., Bradshaw, C. P. and Leaf, P. J. (2011) The use of propensity scores to assess the generalizability of results from randomized trials. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 174, 369–386.

[24] Van Der Vaart, A. W. and Wellner, J. A. (1996) Weak convergence. In *Weak convergence and empirical processes*, 16–28. Springer.

[25] VanderWeele, T. J. and Tchetgen Tchetgen, E. J. (2017) Mediation analysis with time varying exposures and mediators. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 79, 917–938.

[26] Wright, S. (1921) Correlation and causation. *Journal of agricultural research*, 20, 557–585.

[27] Zheng, W. and van der Laan, M. (2017) Longitudinal mediation analysis with time-varying mediators and exposures, with application to survival outcomes. *Journal of Causal Inference*. 17
A Identifiability

To identify the stochastic direct effect and stochastic indirect effect we will need to impose additional assumptions on $\mathcal{M}^F$ and $\mathcal{M}$, listed below.

1. Positivity: For all $S$ and $W$ we need a positive probability of assigning any level of $A$. For all combinations of $S, W,$ and $A=a$, we have a positive probability of any level of $Z$. For $S=1$ and all combinations of $Z$ and $W$ we need a positive probability of any level of the mediator, $M$.

2. Common outcome model across sites: $\mathbb{E}[Y | M, Z, W, S=1] = \mathbb{E}[Y | M, Z, W, S=0]$. The null hypothesis of a common outcome model may be tested nonparametrically \cite{12}.

3. Sequential Randomization: $Y_{am} \perp A | W, S$ and $Y_{am} \perp M | W, Z, S$. This is akin to a two-time point longitudinal intervention where at the first time point, we statically intervene to set the treatment, $A=a$, and at the second time point, we stochastically intervene on the mediator, $M$.

Theorem A.1.

\[
\Psi(P) = \Psi^F(P_{UX}) = \mathbb{E}\left[ \mathbb{E}\left[ \mathbb{E}[Y | M, Z, W, S=1] | W, Z \right] | W, a, S=0, S=0 \right]
\]

Proof:

\[
\Psi^F(P_{UX}) = \mathbb{E}\left[ Y_{a,\hat{g}_M[a^*,W,s]} | S=0 \right]
\]

tower law $\mathbb{E}\left[ Y_{a,\hat{g}_M[a^*,W,s]} | Z, A, W \right] | S=0$

assumption 2 $\mathbb{E}\left[ \sum_m \mathbb{E}[Y_{am}\hat{g}_M[a^*,W,s] | M=m, Z, A, W, S=1] | S=0 \right]
\]

tower law $\mathbb{E}\left[ \sum_m \mathbb{E}[Y_{am}\hat{g}_M[a^*,W,s] | M=m, Z, A, W, S=1] | A=a, W, S=0, S=0 \right]
\]

assumption 3 $\mathbb{E}\left[ \sum_m \mathbb{E}[Y_{am}\hat{g}_M[a^*,W,s] | M=m, Z, A=a, W, S=1] | A=a, W, S=0, S=0 \right]
\]

The conditional expectations are well-defined due to assumption 1 above.
B Estimator derivations

B.1 Stabilized inverse probability of treatment weighted estimator

This estimator solves the estimating equation \( P_n \hat{H}_n(Y - \Psi_n) = 0 \), where \( \hat{H}_n(M, Z, A, W, S) \) is an estimate of the true \( H \) given by

\[
H(M, Z, A, W, S) = \frac{\hat{g}_{M|a^*, W, a}(M \mid W) p_Z(Z \mid A, W, S = 0) p_{S|W}(S = 0 \mid W) I(S = 1, A = a)}{p_M(M \mid Z, W, S) p_Z(Z \mid A, W, S) p_A(A \mid W, S) p_{S|W}(S \mid W) P_S(S = 0)}.
\]

(8)

To solve the estimating equation, we find an initial estimate, \( \Psi^0_n = P_n \hat{H}_n Y \), i.e., the standard unstabilized IPTW estimator.

\[
\Psi^1_n - \Psi^0_n = - \left( \frac{d}{d\Psi} P_n \hat{H}_n(Y - \Psi^0_n) \right)^{-1} P_n \hat{H}_n(Y - \Psi^0_n),
\]

where our stabilized estimator is then \( \Psi^{IPTW}_n = \Psi^1_n = P_n \hat{H}_n Y / P_n \hat{H}_n \).

B.1.1 Efficiency

Theorem B.1. Assume

1. A1: \( \|H - \hat{H}_n\|_{L^2(P_0)} = o_P(n^{-0.5}) \)

2. A2: \( \hat{H}_n \) and \( \hat{D}_n \) are in a \( P_0 \) Donsker class.

Let \( IC_{IPTW} \) be the true influence curve or \( H(Y - \Psi(P_0))/P_0 H = H(Y - \Psi(P_0)) \), noting \( P_0 H = 1 \) for the true \( H \). Then \( \sqrt{n}(\Psi^{IPTW}_n - \Psi(P_0)) \) converges in distribution to \( \sqrt{n} \) times the empirical mean of the true influence curve. In this case, our standard error times \( \sqrt{n} \) will converge to the standard deviation of \( IC_{IPTW} \).

Proof:

\[
\Psi^{IPTW}_n - \Psi(P_0) = \Psi^0_n + P_n \hat{D}_n - \Psi(P_0)
= \Psi^0_n + (P_n - P_0) \hat{D}_n + P_0 \hat{D}_n - \Psi(P_0)
\]

By A2 \( \sqrt{n}(P_n - P_0) \hat{D}_n \) converges in distribution to a normal distribution of mean 0 and variance equal to the true variance of \( IC_{IPTW} \). Thus, we must show \( \Psi^0_n + P_0 \hat{D}_n - \Psi(P_0) = o_P(1/\sqrt{n}) \) under A1.

\[
\Psi^0_n + P_0 \hat{D}_n - \Psi(P_0) = P_n \hat{H}_n Y - P_0 H Y + \frac{P_0 \hat{H}_n Y}{P_n \hat{H}_n} - \frac{P_n \hat{H}_n P_0 \hat{H}_n Y}{P_n \hat{H}_n}
= (P_n - P_0) \frac{\Psi^0_n}{P_n \hat{H}_n} + P_0(\hat{H}_n - H) Y + \frac{P_0 \hat{H}_n Y}{P_n \hat{H}_n} \left( (P_0 - P_n) \hat{H}_n + P_0 (H - \hat{H}_n) \right)
= P_0(\hat{H}_n - H) Y + \frac{(P_0 - P_n) \hat{H}_n Y (P_0 - P_n) \hat{H}_n}{P_n \hat{H}_n} + \frac{P_0 \hat{H}_n Y}{P_n \hat{H}_n} P_0 (H - \hat{H}_n)
= P_0(\hat{H}_n - H) Y + \frac{(P_0 - P_n) \hat{H}_n Y ((P_0 - P_n) (\hat{H}_n - H) + H)}{P_n \hat{H}_n} + \frac{P_0 \hat{H}_n Y}{P_n \hat{H}_n} P_0 (H - \hat{H}_n)
\]

(9)
We can notice in Equation 9 that the first term is bounded above by $C\|\hat{H}_n - H\|_{L^2(P_0)}$ for some constant $C$. By A2, the second term is $o_P(1/\sqrt{n})$. By A1 the third term is also $o_P(1/\sqrt{n})$.

**Theorem B.2.** The IPTW estimator’s limiting variance under the assumptions of theorem B.1.1 is larger than that of the TMLE or EE under well-specified models, if the model is non-parametric. If we use the TMLE and EE incorporating the efficient influence curve for the restricted model, then these are even more efficient.

**Proof:**
Let’s assume the unrestricted model, $M^{np}$. We project the IPTW influence curve onto the tangent space subspace, $T_{Y\times S}$ consisting of $\{\gamma(O) | E[\gamma(O) | M, Z, A, W, S] = 0\}$ via the formula

$$\prod (IC_{\text{IPTW}}\|T_{Y\times S}) = IC_{\text{IPTW}}(O) - E[IC_{\text{IPTW}}(O) | M, Z, A, W, S] = D_Y^*(O)$$

Likewise we have $\prod (IC_{\text{IPTW}}(O)|T_Z) = E[IC_{\text{IPTW}}(O) | Z, A, W, S] - E[IC_{\text{IPTW}}(O) | A, W, S] = D_Z^*(O)$, where $T_Z$ is the tangent space consisting of $\{\gamma(O) | E[\gamma(O) | A, W, S] = 0\}$. The empirical mean of $\prod (IC_{\text{IPTW}}(O)|T_S) = E[IC_{\text{IPTW}}(O) | W] - E[IC_{\text{IPTW}}(O) | W]$ is the empirical mean of $D_Y^*(O)$, the $T_S$ component of the efficient influence curve for $M^{np}$. We have non-trivial projections of $IC_{\text{IPTW}}$ onto $T_M$ and $T_A$. Since the asymptotic variance of $IC_{\text{IPTW}}(O)$ is a sum of the variances of its orthogonal components, it must exceed that of the EE and TMLE estimators, and the proof is complete.

**B.1.2 Robustness**
The IPTW estimator requires all models for $M, Z, A$ and $S$ to be estimated at parametric rates in order to be consistent.

**B.2 Estimating equation estimator**

**B.2.1 Derivation of efficient influence curve for model without exclusion restrictions**
First we derive the efficient influence curve (EIC) for the less restricted case where $A$ may directly affect $M$ and $Y$. For more information on deriving an efficient influence curve in general, we refer the reader to a tutorial by Levy, 2019 (11). We will also allow our treatment mechanism not to be known, which will have no effect on the derivation but means that the EIC is applicable in situations when treatment is not randomly assigned.

**Theorem B.3.** Consider a non-parametric model or semiparametric model with one or both the treatment and mediator mechanisms known (mechanisms for $A$ and $M$). Consider the parameter defined by

$$\Psi(P) = E\left[\sum_m E[Y_{\hat{g}_M(a, W, S)}(m) | W] | M = m, W, Z, A = a, S = 1 \right] | A = a, W, S] | S = 0$$
where the expectations are taken with respect to \( P \). Then the efficient influence curve is given by
\[
D^*(P)(O) = D_Y^*(P)(O) + D_Z^*(P)(O) + D_W^*(P)(O)
\]

where
\[
D_Y^*(P)(O) = (Y - \mathbb{E}[Y \mid M, Z, A, W]) \times \\
g_M^{a,W}(M \mid W)p_Z(Z \mid A, W, S = 0) p_{S\mid W}(S = 0 \mid W) I(S = 1, A = a) / g_M(M \mid Z, A, W, S) p_Z(Z \mid A, W, S) p_{A \mid W, S}(A \mid W, S) p_{S\mid W}(S \mid W) p_S(S = 0)
\]
\[
D_Z^*(P)(O) = (\bar{Q}_M(Z, A, W, S) - \bar{Q}_Z(a, W, S)) I(S = 0, A = a) / p_{A \mid W, S} p_S(S = 0)
\]
\[
D_W^*(P)(O) = (\bar{Q}_Z(a, W, S) - \Psi(P)) I(S = 0) / p_S(S = 0)
\]

Proof: The density factorizes as follows: \( p(O) = p_{Y \times S}(Y \times \gamma \mid M, Z, A, W, S)p_M(M \mid Z, A, W, S)p_Z(Z \mid A, Z, W, S)p_A(A \mid W, S)p_W(S \mid W) p_S(S) \)
This means \( p_{\gamma}(O) = p_{Y \times S}(Y \times \gamma \mid M, Z, A, W, S)p_{M,\gamma}(M \mid Z, A, W, S)p_{Z,\gamma}(Z \mid A, W, S)p_{A,\gamma}(A \mid W, S)p_{W,S}(W \mid S)p_S(S) \)

\[ \frac{d}{de}(p_{\gamma}(Y \times S \mid M, Z, A, W, S)) \bigg|_{e=0} = (\gamma(0) - \mathbb{E}[\gamma(O) \mid M, Z, A, W, S]) p_{Y \times S}(Y \times S \mid M, Z, A, W, S) \quad (10) \]
\[ \frac{d}{de}(p_{Z,\gamma}(Z \mid A, W, S)) \bigg|_{e=0} = (\mathbb{E}[\gamma(O) \mid Z, A, W, S] - \mathbb{E}[\gamma(O) \mid A, W, S]) p_{Z}(Z \mid A, W, S) \quad (11) \]
\[ \frac{d}{de}(p_{W,S}(W \mid S)) \bigg|_{e=0} = (\mathbb{E}[\gamma(O) \mid W, S] - \mathbb{E}[\gamma(O) \mid S]) p_{W,S}(W \mid S) \quad (12) \]

Our estimand, which identifies our parameter of interest (see Section 3 of the main text) is given by
\[
\Psi(P) = \int y p_Y(y \mid m, z, a, w, s = 1) \bar{g}_{M,a,W,s}(m \mid a, w, s = 0) p_{W,S}(w \mid s = 0) dv(y, m, z, w)
\]

We then take the pathwise derivative for a path along score, \( \gamma \). We note that this derivative is unaffected by knowledge of the treatment mechanism, \( E[A \mid S, W] \), or the mediator mechanism, \( E[M \mid Z, A, W, S] \), due to the estimand not depending on these models as well as the fact that scores, \( \gamma_A \) and \( \gamma_M \) are orthogonal (have 0 covariance) to \( \gamma_Y, \gamma_Z, \gamma_W \) in the Hilbert Space \( L^2(P) \). This is why for a semi-parametric model where the treatment mechanism is known, the efficient influence curve will be the same as that for \( \mathcal{M}^{op} \).

\[ \frac{d}{de}\Psi(P) \bigg|_{e=0} = \left[ \frac{d}{de} \int y p_Y(y \mid m, z, a, w, s = 1) \bar{g}_{M,a,W,s}(m \mid a, w, s = 0) p_{W,S}(w \mid s = 0) dv(y, m, z, w) \bigg|_{e=0} \right] \]
\[ + \frac{d}{de} \int y p_Y(y \mid m, z, a, w, s = 1) \bar{g}_{M,a,W,s}(m \mid a, w, s = 0) p_{W,S}(w \mid s = 0) dv(y, m, a, z, w) \bigg|_{e=0} \]
\[ + \frac{d}{de} \int y p_Y(y \mid m, z, a, w, s = 1) \bar{g}_{M,a,W,s}(m \mid a, w, s = 0) p_{W,S}(w \mid s = 0) dv(y, m, a, z, w) \bigg|_{e=0} \]

The first term in (13)
\[
\frac{d}{de} \int \left( \left. g_{M|a^*, W,s} (m \mid w) p_Z (z \mid x = a, w, s = 0) p_{W \mid S} (w \mid s = 0) \right|_{e=0} \right) dv(y, m, z, w)
\]
\[
= \int \frac{d}{de} \left( \left. g_{M|a^*, W,s} (m \mid w) \right|_{e=0} \right) dv(y, m, z, w)
\]
\[
= \int \left( \left. g_{M|a^*, W,s} (m \mid w) \right|_{e=0} \right) dv(y, m, z, w)
\]
\[
\frac{I(s = 1, x = a)}{p_A (x \mid w, s)} \int_{e=0} p_{W \mid S} (w \mid s) p_S (s = 1) \int \left( \left. g_{M|a^*, W,s} (m \mid w) \right|_{e=0} \right) dv(y, m, z, w, s)
\]

where

\[
D_Y^*(P)(O) = (Y - E[Y \mid M, Z, A, W]) \frac{\hat{g}_{M|a^*, W,s} (M \mid W)}{g_M (M \mid Z, A, W, S)} p_Z (Z \mid A, W, S = 0) p_{S \mid W} (S = 0 \mid W) I(S = 1, A = a)
\]

The reader may notice \(D_Y^*(P)(O)\) is not a mean 0 function of \(Y \mid M, Z, W\) because it also depends on the variable, \(A\). Hence, it is not an element of the tangent space under the restricted model, \(\mathcal{M}\), that incorporates the instrumental variable exclusion restrictions as described in Section 2 of the main text, which is necessarily a proper subspace of \(L^2(P)\) for \(P \in \mathcal{M}^{\text{up}}\). Therefore, \(D_Y^*(P)(O)\) has an extra orthogonal component in addition to the efficient influence curve under \(\mathcal{M}\), making its variance necessarily bigger than the lower bound under model \(\mathcal{M}\). This is why the estimators using \(D_Y^*(P)(O)\) (the EE and TMLE estimators) are not asymptotically efficient for model, \(\mathcal{M}\) but are so for \(\mathcal{M}^{\text{up}}\).

The second term in [13]:

\[
\frac{d}{de} \int \left( \left. g_{M|a^*, W,s} (m \mid w) p_Z (z \mid x = a, w, s = 0) p_{W \mid S} (w \mid s = 0) \right|_{e=0} \right) dv(y, m, z, w)
\]
\[
= \int \frac{d}{de} \left( \left. g_{M|a^*, W,s} (m \mid w) \right|_{e=0} \right) dv(y, m, z, w)
\]
\[
= \int \left( \left. g_{M|a^*, W,s} (m \mid w) \right|_{e=0} \right) dv(y, m, z, w)
\]
\[
\frac{I(s = 1, x = a)}{p_A (x \mid w, s)} \int_{e=0} p_{W \mid S} (w \mid s) p_S (s = 1) \int \left( \left. g_{M|a^*, W,s} (m \mid w) \right|_{e=0} \right) dv(y, m, z, w, s)
\]

where

\[
D_Y^*(P)(O) = (Y - E[Y \mid M, Z, A, W]) \frac{\hat{g}_{M|a^*, W,s} (M \mid W)}{g_M (M \mid Z, A, W, S)} p_Z (Z \mid A, W, S = 0) p_{A \mid W, S} (A \mid W, S) p_{S \mid W} (S = 1 \mid W) p_S (s = 0) \int dP(o)
\]
We substitute
\[ \tilde{Q}_M(z, x, w) = \mathbb{E}_{\hat{g}_{M|a^*, W,s}}(\mathbb{E}[Y \mid M, A, Z, W] \mid z, x, w, s = 1) \]
\[ \tilde{Q}_Z(x, w, s) = \mathbb{E}_{P_{Z|A,W,S}}[\mathbb{E}_{\hat{g}_{M|a^*, W,s}}(\tilde{Q}_M(Z, A, W) \mid x, w, s)] \]
and since \( x \) represents the treatment, \( A \), in the integrals above, we get
\[
D^*_Z(P)(O) = (\tilde{Q}_M(Z, A, W) - \tilde{Q}_Z(A, W, S)) \frac{I(S=0, A=a)}{\rho_A(A|W,S)p_Z(S=0)}. 
\]
The third term in [12] is the conditional expectation of the treatment mechanism given \( z, w, s \), and \( S, \delta \),
\[
\frac{d}{d\epsilon} \int yp y(y \mid m, z, a, w, s = 1) \hat{g}_{M|a^*, W,s}(m \mid w)p_Z(z \mid a, w, s = 0) p_{W|S,s}(w \mid s = 0) dw(y, m, z, w) \bigg| \epsilon = 0 
\]
\[ = \int yp y(y \mid m, z, w) \hat{g}_{M|a^*, W,s}(m \mid w)p_Z(z \mid a, w, s) \frac{d}{d\epsilon} p_{W|S,s}(w \mid s) \bigg| \epsilon = 0 \rho_Z(s) dv(y, m, z, w, s) 
\]
\[ = \int S(o) (\tilde{Q}_Z(x = a, w, s) - \Psi(P)) \frac{I(s = 0)}{\rho_Z(s = 0)} dP(o) 
\]
\[ = (\gamma, D^*_W)_{I^*_Z(P)} \]
where \( D^*_W(P)(O) = (\tilde{Q}_Z(A = a, W, S) - \Psi(P)) \frac{I(S=0)}{\rho_Z(S=0)} \)
Thus the efficient influence curve is the sum of its orthogonal components:
\[
D^*(P)(O) = D^*_Y(P)(O) + D^*_Z(P)(O) + D^*_W(P)(O) 
\]

### B.2.2 Derivation of efficient influence curve for model with exclusion restrictions

Next, we derive the EIC for the model that restricts \( A \) to affect \( M \) and \( Y \) only through \( Z \). Similar to the above subsection, we allow our treatment mechanism not to be known.

**Theorem B.4.** The efficient influence curve for our restricted model, where \( M \) and \( Y \) do not depend directly on \( A \), is given by
\[
D^*(P)(O) = D^*_{Y,r}(P)(O) + D^*_Z(P)(O) + D^*_W(P)(O) 
\]
where
\[
D^*_{Y,r}(P)(O) = \left( y - \mathbb{E}[y \mid m, z, w] \right) \frac{\hat{g}_{M|a^*, W,s}(m \mid w)p_Z(z \mid a_0, w, s = 0)p_{S|W}(s = 0 \mid w)I(s = 1)}{g_{M,r}(m \mid z, w, s)p_{Z|W,S}(z \mid w, s)p_{S|W}(s \mid w)\rho_S(s = 0)} 
\]

**Proof:**
We can note that our only task here is to project \( D^*_r(P) \), our component of the influence curve in \( T_Y \), onto the subspace of \( T_Y \) given by
\[
T_Y = \{ \gamma : \mathbb{E}(\gamma(O) \mid Y, S, M, Z, W, S) = 0, \mathbb{E}(\gamma(O))^2 < \infty \}. 
\]
\( \rho_Y \) is the conditional density of \( y \) given \( m, z, w \) and \( p_{M,r} \) is the conditional density of \( m \) given \( z, w, s \) in the restricted model, i.e. we don’t put the instrument, \( a \), in
those conditional statements as that is the model assumption. We remind the reader that a "bar" signifies the variable and all past variables as in, $\bar{M} = m, z, x, w, s$.

Notice the following:

$$
\begin{align*}
p_{A \mid Y, S, r}(x \mid ys, m, z, w, s = 1) &= \frac{p_{A,r}(x, ys, m, z, w, s = 1)}{p_{O/A}(ys, m, z, w, s = 1)} \\
&= \frac{p_{Y,r}(y \mid m, z, w)p_{M,r}(m \mid z, w, s = 1)p_{\bar{Z}}(z, x, w, s)}{p_{Z/A}(z, w, s = 1)} \\
&= \frac{p_{\bar{Z}}(z, x, w, s = 1)}{p_{Z/A}(z, w, s = 1)}
\end{align*}
$$

(14)

$$
\begin{align*}
p_{A,Y, S, r}(x, ys \mid m, z, w, s = 1) &= \frac{p_{Y,r}(ys, x, m, z, w, s = 1)}{p_{M,r}(m, z, w, s = 1)} \\
&= \frac{p_{Y,r}(y \mid m, z, w)p_{\bar{Z}}(z, x, w, s = 1)}{p_{Z/A}(z, w, s = 1)}
\end{align*}
$$

(15)
Thus from (14) and (15)

$$\Pi(D^r_Y || Y_{r,r})$$

$$= \mathbb{E}(D^r_Y(O) | Y_S, M, Z, W, S) - \mathbb{E}(D^r_Y(O) | M, Z, W, S)$$

$$= \int (y - \mathbb{E}[y | m, z, w]) \times$$

$$\frac{\hat{g}_{M,a^*,W,a}(m | w)p_Z(z | a, w, s = 0)p_{S|W}(s = 0 | w)I(s = 1, x = a)}{g_{M,r}(m | z, w, s = 1)p_Z(z | a, w, s = 1)g_A(a | w, 1)p_{S|W}(1 | w)p_S(0)} p_{A|Y|S,r}(x | ys, m, z, w) d\nu(x)$$

$$- \int (y - \mathbb{E}[y | m, z, w]) \times$$

$$\frac{\hat{g}_{M,a^*,W,a}(m | w)p_Z(z | a, w, s = 0)p_{S|W}(s = 0 | w)I(s = 1, x = a)}{g_{M,r}(m | z, w, s = 1)p_Z(z | a, w, s = 1)g_A(a | w, 1)p_{S|W}(1 | w)p_S(0)} p_{A|Y|S,r}(x | ys, m, z, w) d\nu(x, ys)$$

remembering we are integrating with respect to x and all else is fixed in the first integral

All is fixed but x and ys in the second integral. Since I(s=1), ys = 1 and s = 1

$$= \int (y - \mathbb{E}[y | m, z, w]) \times$$

$$\frac{\hat{g}_{M,a^*,W,a}(m | w)p_Z(z | a, w, s = 0)p_{S|W}(s = 0 | w)I(s = 1, x = a)}{g_{M,r}(m | z, w, s = 1)p_Z(z | a, w, s = 1)g_A(a | w, 1)p_{S|W}(1 | w)p_S(0)} p_{A|Y|S,r}(x | ys, m, z, w, s = 1) d\nu(x)$$

$$- \int (y - \mathbb{E}[y | m, z, w]) \times$$

$$\frac{\hat{g}_{M,a^*,W,a}(m | w)p_Z(z | a, w, s = 0)p_{S|W}(s = 0 | w)I(s = 1, x = a)}{g_{M,r}(m | z, w, s = 1)p_Z(z | a, w, s = 1)g_A(a | w, 1)p_{S|W}(1 | w)p_S(0)} p_{A|Y|S,r}(x, ys | m, z, w, s = 1) d\nu(x, ys)$$

use (14) and (15) for the 1st and 2nd integrals respectively:

$$= \int (y - \mathbb{E}[y | m, z, w]) \times$$

$$\frac{\hat{g}_{M,a^*,W,a}(m | w)p_Z(z | a, w, s = 0)p_{S|W}(s = 0 | w)I(s = 1, x = a)}{g_{M,r}(m | z, w, s = 1)p_Z(z | a, w, s = 1)g_A(a | w, 1)p_{S|W}(1 | w)p_S(0)} p_{Z}(z, x, w, s = 1) d\nu(x)$$

$$- \int (y - \mathbb{E}[y | m, z, w]) p_{Y,r}(y | m, z, w) d\nu(y) \times$$

$$\frac{\hat{g}_{M,a^*,W,a}(m | w)p_Z(z | a, w, s = 0)p_{S|W}(s = 0 | w)I(s = 1, x = a)}{g_{M,r}(m | z, w, s = 1)p_Z(z | a, w, s = 1)g_A(a | w, 1)p_{S|W}(1 | w)p_S(0)} p_{Z}(z, w, s = 1) d\nu(x)$$

$$= (y - \mathbb{E}[y | m, z, w]) \times$$

$$\frac{\hat{g}_{M,a^*,W,a}(m | w)p_Z(z | a, w, s = 0)p_{S|W}(s = 0 | w)I(s = 1, x = a)}{g_{M,r}(m | z, w, s = 1)p_Z(z | a, w, s = 1)g_A(a | w, 1)p_{S|W}(1 | w)p_S(0)} p_{Z}(z, w, s = 1) d\nu(x)$$

And the proof is complete since the other components of the unrestricted model’s influence curve will remain the same. The reader may note that $p_{Z|W,S}(z | w, s) = p_{Z}(z | 1, w, s)g_A(1 | w, s) + p_{Z}(z | 0, w, s)g_A(0 | w, s)$, so we need not perform any additional regressions for this restricted model.

### B.2.3 Robustness

Here we derive the remainder term for the EE estimator. We first form an initial estimate

$$\hat{\Psi}^0_n = \sum_{i=1}^{n} \frac{\mathbb{I} (S_i = 0)}{\sum_{i=1}^{n} \mathbb{I} (S = 0)} Q^0_{L_0,N}(A_i = a, W_i, S_i)$$

25
Then we update this estimate by adding the empirical mean of the approximated influence curve. If we call our approximated influence curve, $\hat{D}_{EE}^n$ our estimating equation (EE) estimate is given by

$$\hat{\Psi}^1_n = \hat{\Psi}^0_n + \sum_{i=1}^n \hat{D}_{EE}^n(O_i)$$

This then leads to a second order expansion

$$\hat{\Psi}^1_n - \Psi(P_0) = (P_n - P_0)\hat{D}_{EE}^n(O) + R_2(P_n, P_0)$$

where $R_2(P_n, P_0) = \hat{\Psi}^0_n - \Psi(P_0) + P_0\hat{D}_{EE}^n(O)$.

The behavior of $R_2$ determines the robustness of our estimator and conditions under which we can guarantee consistency and efficiency [8, 9].

**Theorem B.5.**

$$R_2(P_n, P_0) = \int \frac{(\bar{Q}_Y(m, z, w) - \bar{Q}_Y(m, z, w))}{g(o)} f_{1,0} f_{2,0} f_{3,0} f_{4,0} f_{5,0} f_{6,0} f_{7,0}(o) dv(o)$$

$$= \int (\bar{Q}_Y(m, z, w) - \bar{Q}_Y,0(m, z, w)) \sum_{i=1}^6 (f_{i,0} - f_{i})(o) h_i(o)$$

Cauchy–Schwarz \leq k \sum_{i=1}^6 \|\bar{Q}_Y - \bar{Q}_Y,0\|_{L^2(P_0)} \|f_{i,0} - f_{i}\|_{L^2(P_0)}$$

Where we substituted the following: $f_{1,0}(o) = g_{M,0}(m \mid m, z, w)$, $f_{2,0} = p_{Z,0}(m \mid a, w, 1)$, $f_{3,0} = p_{Z,0}(m \mid a, w, 0)$, $f_{4,0} = p_{A,0}(x = a \mid w, 1)$, $f_{5,0} = p_{A,0}(x = a \mid w, 0)$, $f_{6,0} = p_{S,0}(s = 1 \mid w)$ and $f_{7,0} = p_{S,0}(s = 0 \mid w)$ and dropping the subscript, 0, indicates the estimated counterpart. $h_i$ is a bounded function by the positivity assumption (see section 4) and thus the last inequality holds with a sufficiently large $k$. 

26
\[
R_2 = \Psi_n - \Psi(P_0) + P_0 \left\{ \left( Y - \bar{Q}_Y(M, Z, W) \right) \frac{\hat{g}_{a*, W,S}(M \mid W)p_Z(Z \mid A, W, S = 0)p_{S|W}(S = 0 \mid W)I(S = 1, A = a)}{g_{M}(M \mid Z, W, S)p_Z(Z \mid A, W, S)p_{A}(A \mid W, S)p_{S|W}(S \mid W)P(S = 0)} \right\}
\]

\[
+ P_0 \left\{ \left( Q_{L_1}(Z, A, W) - Q_{L_0}(A, W, S) \right) \frac{I(S = 0, A = a)}{p_A(A \mid W, S)P(S = 0)} \right\}
\]

\[
\approx P_0 \left\{ \left( Y - \bar{Q}(M, Z, W) \right) \frac{\hat{g}_{a*, W,S}(M \mid W,S = 0)p_Z(Z \mid A, W, S = 0)p_{S|W}(S = 0 \mid W)I(S = 1, A = a)}{g_{M}(M \mid Z, W, S)p_Z(Z \mid A, W, S)p_{A}(A \mid W, S)p_{S|W}(S \mid W)P(S = 0)} \right\}
\]

\[
+ P_0 \left\{ \left( \bar{Q}_{L_1}(Z, A, W) - \bar{Q}_{L_0}(A, W, S) \right) \frac{I(S = 0, A = a)}{p_A(A \mid W, S)P(S = 0)} \right\}
\]

\[
= P_0 \left\{ \left( \bar{Q}_{Y,0}(M, Z, W) - \bar{Q}_Y(M, Z, W) \right) \frac{\hat{g}_{a*, W,S}(M \mid W,S = 0)p_Z(Z \mid A, W, S = 0)p_{S|W}(S = 0 \mid W)I(S = 1, A = a)}{g_{M}(M \mid Z, W, S)p_Z(Z \mid A, W, S)p_{A}(A \mid W, S)p_{S|W}(S \mid W)P(S = 0)} \right\}
\]

\[
+ P_0 \left\{ \left( Q_{L_1,0}(Z, A, W) - Q_{L_0,0}(A, W, S) \right) \frac{I(S = 0, A = a)}{p_A(A \mid W, S)P(S = 0)} \right\}
\]

The treatment mechanism (model for A) being well-specified makes term 4 disappear above. Clearly terms 3 and 5 only cancel if the models for M, Z and S also are well-specified.

Integrating terms 3 and 5, we get

\[
\int (\bar{Q}_{Y,0}(m, Z, W) - \bar{Q}_Y(m, Z, W)) \hat{g}_{a*, W}(m \mid Z) \frac{g_{M,0}(m \mid z, w, 1) p_{Z,0}(Z \mid a, w, 1) p_{Z}(Z \mid a, w, 0) p_{A}(A \mid w, 1) p_{S|W}(1 \mid w)P(s = 0)}{g_{M}(m \mid z, w, 1) p_{Z}(Z \mid a, w, 1) p_{Z}(Z \mid a, w, 0) p_{A}(A \mid w, 1) p_{S|W}(1 \mid w)P(s = 0)} \ dv(o)
\]

If we add fractions and put everything in one integral we have:

\[
\int (\bar{Q}_Y(m, z, w) - \bar{Q}_{Y,0}(m, z, w)) \frac{f_{1,0}f_{2,0}f_{3,0}f_{4,0}f_{5,0}f_{6,0}f_{7}(o) - f_{1}f_{2}f_{3}f_{4}f_{5}f_{6}f_{7}(o)}{g(o)} \ dv(o)
\]

\[
= \int (\bar{Q}_Y(m, z, w) - \bar{Q}_{Y,0}(m, z, w)) \sum_{i=1}^{6} (f_{i,0} - f_{i}) (o) h_i(o)
\]

By Cauchy-Schwarz

\[
\| \bar{Q}_Y - \bar{Q}_{Y,0} \|_{L^2(P_0)} \leq \sqrt{k} \sum_{i=1}^{6} \| f_{i,0} - f_{i} \|_{L^2(P_0)}
\]

h_i is a bounded function by the positivity assumption (see section 4) and thus the last inequality holds with a sufficiently large k, completing the proof.

We conclude that a product of $L^2(P_0)$ norms between the bias in estimating the outcome model and the bias of the regression models for $M \mid Z, W, S, Z \mid A, W, S, A \mid W, S$
and \( S \mid W \) must be such that the product converges to 0 in probability when multiplied by \( \sqrt{n} \). Also, \( \sqrt{n} \| \tilde{Q}_{L_1} - \tilde{Q}_{L_1,0} \|_{L^2 P_0} \| p_A - p_{A,0} \|_{L^2 P_0} \) must also converge to 0 in probability for the TMLE to be consistent and asymptotically linear with limiting variance that of the TMLE influence curve (efficient for the broader model including dependence of \( Y \) and \( M \) on \( A \)). Such conditions are guaranteed asymptotically when using the highly adaptive lasso to fit the regressions if the true regressions are of finite sectional variation norm and are left-hand continuous with right-hand limits \( \frown \).

To summarize, if the outcome model is misspecified then we need to correctly specify all the models \( M \mid Z, W, S, Z \mid A, W, S, A \mid W, S \) and \( S \mid W \) for the EE estimator to be consistent and asymptotically efficient. If we specify the outcome model correctly so that \( \sqrt{n} \| \tilde{Q}_Y - \tilde{Q}_{Y,0} \|_{L^2 (P_0)} \) converges to 0 in probability, then clearly term 1 disappears by the Cauchy-Schwarz inequality, and we have to consider term 2, which disappears if \( \sqrt{n} \| \tilde{Q}_{L_1} - \tilde{Q}_{L_1,0} \|_{L^2 P_0} \| p_A - p_{A,0} \|_{L^2 P_0} \) converges to 0 in probability. Thus if we correctly specify the regression, \( \tilde{Q}_{L_1} (Z, W) \mid A, W, S \), along with the regression for the outcome, then we need not specify any other model correctly to obtain a consistent estimator. If we correctly specify the outcome and misspecify the second regression, \( \tilde{Q}_{L_1} \mid A, W, S \), then we need to correctly specify the treatment mechanism in order to be consistent.

### B.3 Targeted minimum loss-based estimator

#### B.3.1 Derivation of efficient influence curve

The derivations of the EICs for the less restricted and more restricted models are given in subsections [B.2.1](#) and [B.2.2](#).

#### B.3.2 Robustness

For our TMLE estimator: \( \Psi(P^*_n) - \Psi(P_0) = (P_n - P_0)D^*(P^*_n) + R_2(P^*_n, P_0) \) where \( P_0 \) is the true observed data generating distribution and \( P^*_n \) is the TMLE updated estimation of \( P_0 \). Since the empirical mean of \( D^*(P^*_n) \) = 0 by virtue of the TMLE algorithm, we have \( R_2(P^*_n, P_0) = \Psi(P^*_n) - \Psi(P_0) + P_0D^*(P^*_n) \). The robustness properties of the TMLE estimator are identical to the EE estimator, see Subsection [B.2.3](#).
C Simulation

Table 4: Simulation data-generating mechanisms.

| W_1 ~ bernoulli | P(W_1 = 1) = 0.5 |
| W_2 ~ bernoulli | P(W_2 = 1) = \text{expit}(0.4 + 0.2W_1) |
| S ~ bernoulli   | P(S = 1) = \text{expit}(3W_2 - 1) |
| A ~ bernoulli   | P(A = 1) = 0.5 |
| Z ~ bernoulli   | P(Z = 1) = \text{expit}(-3A + -0.2S + 2W_2 + 0.2AW_2 - 0.2AS + 0.2W_2S + 2AW_2S - 0.2) |
| M ~ bernoulli   | P(M = 1) = \text{expit}(1Z + 6W_2Z - 2W_2 - 2) |
| Y ~ bernoulli   | P(Y = 1) = \text{expit}(\log(1.2) + \log(40)Z - \log(30)M - \log(1.2)W_2 - \log(40)W_2Z) |

Data Generating Mechanism 2

| W_1 ~ bernoulli | P(W_1 = 1) = 0.5 |
| W_2 ~ bernoulli | P(W_2 = 1) = \text{expit}(0.4 + 0.2W_1) |
| S ~ bernoulli   | P(S = 1) = \text{expit}(3W_2 - 1) |
| A ~ bernoulli   | P(A = 1) = 0.5 |
| Z ~ bernoulli   | P(Z = 1) = \text{expit}(-0.1A + -0.2S + 0.2W_2 + 5AW_2 + 0.14AS + 0.2W_2S - 0.2AW_2S - 1) |
| M ~ bernoulli   | P(M = 1) = \text{expit}(1Z + 3ZW_2 + 0.2ZS - 0.2W_2S + 2W_2Z + 0.2S - 0.2ZW_2S - W_2 - 2) |
| Y ~ bernoulli   | P(Y = 1) = \text{expit}(-6Z + 0.2ZW_2 + 2ZM + 2W_2M - 2W_2 + 4M + 1ZW_2M - 0.2) |

Data Generating Mechanism 3

| W_1 ~ bernoulli | P(W_1 = 1) = 0.5 |
| W_2 ~ bernoulli | P(W_2 = 1) = \text{expit}(0.4 + 0.2W_1) |
| S ~ bernoulli   | P(S = 1) = \text{expit}(3W_2 - 1) |
| A ~ bernoulli   | P(A = 1) = 0.5 |
| Z ~ bernoulli   | P(Z = 1) = \text{expit}(-3A + 2S + 2W_2 + 0.2AW_2 - 0.2AS + 0.2W_2S + 2AW_2S - 0.2) |
| M ~ bernoulli   | P(M = 1) = \text{expit}(3Z - 0.2ZW_2 + 0.2ZS - 0.2W_2S + 2W_2Z + 0.2S - 0.2ZW_2S - W_2 - 2) |
| Y ~ bernoulli   | P(Y = 1) = \text{expit}(-6Z + 0.2ZW_2 + 2ZM + 2W_2M - 0.2W_2 + 4M + 1ZW_2M - 0.2) |

DGM 1 is intended to break when Y and M models are misspecified, especially. DGM 2 is intended to break when Y and Z models are misspecified, especially. DGM 3 is intended to break when Y and S models are misspecified, especially.

Table 5: Simulation results comparing estimators of the transported direct and indirect effects for DGP 1 under well-specified models for sample sizes 100, 500 and 5000

| Estimator     | Bias  | Efficiency | 95%CI Cov | RMSE | % Out of Bds |
|---------------|-------|------------|-----------|------|--------------|
|               | IC Bootstrapping | IC Bootstrapping |         |      |              |
| SDE_tmle_eff  | 0.015 | 86.77      | 161.49    | 0.807| 0.959        | 0.114    | 0       |
| SDE_EE_eff    | 0.016 | 86.77      | 129.06    | 0.812| 0.919        | 0.112    | 0       |
| SDE_tmle      | 0.028 | 201.37     | 292.39    | 0.82 | 0.941        | 0.252    | 0       |
| SDE_EE        | 0.024 | 216.18     | 4222.44   | 0.923| 0.924        | 0.210    | 0       |
| SDE_iptw      | 0.039 | 301.76     | 305.12    | 0.906| 0.906        | 0.273    | 0       |

Transport stochastic indirect effect

| Estimator     | Bias  | Efficiency | 95%CI Cov | RMSE | % Out of Bds |
|---------------|-------|------------|-----------|------|--------------|
| SIE_tmle_eff  | -0.005| 103.450    | 52.71     | 0.833| 0.961        | 0.044    | 0       |
| SIE_EE_eff    | -0.004| 102.77     | 144.73    | 0.843| 0.907        | 0.038    | 0       |

Continued on next page
Table 5 – continued from previous page

| Estimator   | Bias | Efficiency | 95%CI | Cov | RMSE | % Out of Bds |
|-------------|------|------------|-------|-----|------|--------------|
|             | IC   | Bootstrapping | IC   | Bootstrapping |       |              |
| SIE_tmle    | -0.007 | 126.57 | 545.44 | 0.791 | 0.948 | 0.056 | 0 |
| SIE_EE      | -0.004 | 118.11 | 158.02 | 0.829 | 0.910 | 0.039 | 0 |
| SIE_iptw    | -0.020 | 197.71 | 193.86 | 0.745 | 0.741 | 0.052 | 0 |
| **DGM 1, N=500** |      |            |       |      |      |      |
| Transport stochastic direct effect |       |            |       |      |      |      |
| SDE_tmle_eff | 0.0005 | 96.77 | 101.58 | 0.945 | 0.950 | 0.039 | 0 |
| SDE_EE_eff  | 0.0005 | 96.81 | 101.20 | 0.946 | 0.948 | 0.039 | 0 |
| SDE_tmle    | 0.0002 | 225.77 | 236.31 | 0.933 | 0.941 | 0.091 | 0 |
| SDE_EE      | 0.0003 | 227.70 | 226.59 | 0.948 | 0.936 | 0.090 | 0 |
| SDE_iptw    | 0.005  | 300.50 | 283.20 | 0.948 | 0.934 | 0.111 | 0 |
| Transport stochastic indirect effect |       |            |       |      |      |      |
| SIE_tmle_eff | 0     | 101.96 | 110.93 | 0.938 | 0.940 | 0.008 | 0 |
| SIE_EE_eff  | 0     | 102.35 | 107.99 | 0.939 | 0.939 | 0.008 | 0 |
| SIE_tmle    | 0.0001 | 125.37 | 166.92 | 0.920 | 0.949 | 0.011 | 0 |
| SIE_EE      | 0.0001 | 126.16 | 131.84 | 0.925 | 0.940 | 0.011 | 0 |
| SIE_iptw    | -0.003 | 243.47 | 211.11 | 0.894 | 0.874 | 0.015 | 0 |
| **DGM 1, N=5000** |      |            |       |      |      |      |
| Transport stochastic direct effect |       |            |       |      |      |      |
| SDE_tmle_eff | -0.0001 | 100.39 | 100.59 | 0.949 | 0.949 | 0.013 | 0 |
| SDE_EE_eff  | -0.0001 | 100.39 | 100.59 | 0.948 | 0.949 | 0.013 | 0 |
| SDE_tmle    | 0.001  | 227.08 | 227.72 | 0.96  | 0.959 | 0.028 | 0 |
| SDE_EE      | 0.001  | 227.21 | 227.14 | 0.96  | 0.958 | 0.028 | 0 |
| SDE_iptw    | 0.002  | 301.43 | 277.55 | 0.965 | 0.948 | 0.033 | 0 |
| Transport stochastic indirect effect |       |            |       |      |      |      |
| SIE_tmle_eff | -0.0001 | 101.12 | 101.44 | 0.932 | 0.924 | 0.002 | 0 |
| SIE_EE_eff  | -0.0001 | 101.23 | 101.43 | 0.932 | 0.925 | 0.002 | 0 |
| SIE_tmle    | -0.0001 | 130.34 | 131    | 0.939 | 0.940 | 0.003 | 0 |
| SIE_EE      | -0.0001 | 130.50 | 130.94 | 0.940 | 0.941 | 0.003 | 0 |
| SIE_iptw    | -0.0003 | 222.77 | 175.23 | 0.984 | 0.955 | 0.003 | 0 |

Table 6: Simulation results comparing estimators of the transported direct and indirect effects for DGP 1 under \(Z\), \(M\), and \(S\) models misspecified. For sample sizes 100, 500 and 5000.

| Estimator   | Bias | Efficiency | 95%CI | Cov | RMSE | % Out of Bds |
|-------------|------|------------|-------|-----|------|--------------|
|             | IC   | Bootstrapping | IC   | Bootstrapping |       |              |
| **DGM 1, N=100** |      |            |       |      |      |      |
| Transport stochastic direct effect |       |            |       |      |      |      |
| SDE_tmle_eff | 0.029 | 60.77 | 125.16 | 0.529 | 0.879 | 0.139 | 0 |
| SDE_EE_eff  | 0.029 | 60.38 | 121.08 | 0.539 | 0.873 | 0.134 | 0 |
| SDE_tmle    | 0.026 | 87.39 | 167.61 | 0.609 | 0.922 | 0.177 | 0 |
| SDE_EE      | 0.029 | 93.83 | 134.66 | 0.693 | 0.878 | 0.159 | 0 |
| SDE_iptw    | 0.027 | 182.41 | 172.73 | 0.867 | 0.836 | 0.209 | 0 |
| Transport stochastic indirect effect |       |            |       |      |      |      |
| SIE_tmle_eff | -0.009 | 54.10 | 200.16 | 0.616 | 0.932 | 0.076 | 0 |
| SIE_EE_eff  | -0.010 | 56.47 | 166.06 | 0.626 | 0.919 | 0.075 | 0 |
| SIE_tmle    | -0.012 | 69.12 | 244.45 | 0.572 | 0.923 | 0.089 | 0 |

Continued on next page
Table 6 – continued from previous page

| Estimator  | Bias  | Efficiency | 95% CI Cov | RMSE | % Out of Bds |
|------------|-------|------------|------------|------|--------------|
|            | IC    | Bootstrapping | IC       | Bootstrapping |       |              |
| SIE_EE     | -0.010 | 77.87 | 172.43 | 0.669 | 0.912 | 0.079 | 0 |
| SIE_iptw   | -0.007 | 120.77 | 112.76 | 0.833 | 0.856 | 0.055 | 0 |

DGM 1, N=500

Transport stochastic direct effect

| Estimator  | Bias  | Efficiency | 95% CI Cov | RMSE | % Out of Bds |
|------------|-------|------------|------------|------|--------------|
|            | IC    | Bootstrapping | IC       | Bootstrapping |       |              |
| SDE_tmle_eff | -0.005 | 53.490 | 100.99 | 0.643 | 0.879 | 0.064 | 0 |
| SDE_EE_eff | -0.005 | 53.66 | 98.63 | 0.643 | 0.879 | 0.063 | 0 |
| SDE_tmle   | -0.007 | 91.11 | 141.58 | 0.747 | 0.895 | 0.092 | 0 |
| SDE_EE     | -0.006 | 97.42 | 125.71 | 0.843 | 0.920 | 0.078 | 0 |
| SDE_iptw   | 0.012  | 149.01 | 144.41 | 0.957 | 0.935 | 0.087 | 0 |

Transport stochastic indirect effect

| Estimator  | Bias  | Efficiency | 95% CI Cov | RMSE | % Out of Bds |
|------------|-------|------------|------------|------|--------------|
|            | IC    | Bootstrapping | IC       | Bootstrapping |       |              |
| SIE_tmle_eff | 0.001  | 64.72 | 97.48 | 0.826 | 0.948 | 0.019 | 0 |
| SIE_EE_eff | 0.0002 | 65.52 | 95.96 | 0.839 | 0.952 | 0.018 | 0 |
| SIE_tmle   | -0.0004 | 96.56 | 142.93 | 0.797 | 0.932 | 0.027 | 0 |
| SIE_EE     | -0.001 | 101.28 | 118.11 | 0.889 | 0.951 | 0.023 | 0 |
| SIE_iptw   | -0.0002 | 142.27 | 134.94 | 0.910 | 0.895 | 0.028 | 0 |

DGM 1, N=5000

Transport stochastic direct effect

| Estimator  | Bias  | Efficiency | 95% CI Cov | RMSE | % Out of Bds |
|------------|-------|------------|------------|------|--------------|
|            | IC    | Bootstrapping | IC       | Bootstrapping |       |              |
| SDE_tmle_eff | -0.0001 | 48.68 | 93.50 | 0.677 | 0.942 | 0.019 | 0 |
| SDE_EE_eff | -0.0002 | 48.71 | 93.31 | 0.676 | 0.938 | 0.019 | 0 |
| SDE_tmle   | -0.0001 | 94.62 | 150.35 | 0.768 | 0.947 | 0.030 | 0 |
| SDE_EE     | -0.0002 | 94.81 | 124.45 | 0.854 | 0.948 | 0.025 | 0 |
| SDE_iptw   | 0.014  | 138.90 | 132.79 | 0.923 | 0.912 | 0.031 | 0 |

Transport stochastic indirect effect

| Estimator  | Bias  | Efficiency | 95% CI Cov | RMSE | % Out of Bds |
|------------|-------|------------|------------|------|--------------|
|            | IC    | Bootstrapping | IC       | Bootstrapping |       |              |
| SIE_tmle_eff | 0.0001 | 61.13 | 77.84 | 0.875 | 0.947 | 0.006 | 0 |
| SIE_EE_eff | 0.0001 | 61.18 | 75.11 | 0.888 | 0.945 | 0.005 | 0 |
| SIE_tmle   | 0.0002 | 94.19 | 117.94 | 0.885 | 0.955 | 0.008 | 0 |
| SIE_EE     | 0.0001 | 94.53 | 10 | 0.941 | 0.957 | 0.007 | 0 |
| SIE_iptw   | 0.002  | 128.26 | 123.34 | 0.949 | 0.937 | 0.009 | 0 |

Table 7: Simulation results comparing estimators of the transported direct and indirect effects for DGP 1 under Y and S models misspecified. For sample sizes 100, 500 and 5000.
Table 7 – continued from previous page

| Estimator          | Bias  | Efficiency | 95% CI | Cov | RMSE | % Out of Bds |
|--------------------|-------|------------|--------|-----|------|--------------|
|                    |       | IC Bootstrapping | IC Bootstrapping |     |      |              |
| **DGM 1, N=500**   |       |             |        |     |      |              |
| Transport stochastic direct effect |       |             |        |     |      |              |
| SDE_tmle_eff       | 0.11  | 140.17     | 175    | 0.528 | 0.577 | 0.149        | 0             |
| SDE_EE_eff         | 0.141 | 294.11     | 171.95 | 0.577 | 0.546 | 0.172        | 0             |
| SDE_tmle           | 0.120 | 254.54     | 308.58 | 0.524 | 0.571 | 0.212        | 0             |
| SDE_EE             | 0.141 | 454.66     | 307.74 | 0.514 | 0.475 | 0.235        | 0             |
| SDE_iptw           | 0.173 | 236.52     | 255.31 | 0.477 | 0.521 | 0.213        | 0             |
| Transport stochastic indirect effect |       |             |        |     |      |              |
| SIE_tmle_eff       | 0.026 | 290.33     | 468.23 | 0.429 | 0.673 | 0.040        | 0             |
| SIE_EE_eff         | 0.013 | 638.03     | 393.79 | 0.826 | 0.664 | 0.034        | 0             |
| SIE_tmle           | 0.022 | 384.49     | 542.07 | 0.606 | 0.803 | 0.039        | 0             |
| SIE_EE             | 0.013 | 739.29     | 525.56 | 0.810 | 0.728 | 0.041        | 0             |
| SIE_iptw           | 0.024 | 409.12     | 449.20 | 0.666 | 0.693 | 0.036        | 0             |
| **DGM 1, N=5000**  |       |             |        |     |      |              |
| Transport stochastic direct effect |       |             |        |     |      |              |
| SDE_tmle_eff       | 0.045 | 226.69     | 190.97 | 0.717 | 0.512 | 0.054        | 0             |
| SDE_EE_eff         | 0.078 | 519.45     | 183.26 | 0.910 | 0.064 | 0.083        | 0             |
| SDE_tmle           | 0.051 | 490.16     | 587.79 | 0.854 | 0.924 | 0.093        | 0             |
| SDE_EE             | 0.082 | 803.98     | 652.80 | 0.860 | 0.834 | 0.118        | 0             |
| SDE_iptw           | 0.124 | 390.64     | 378.85 | 0.244 | 0.164 | 0.135        | 0             |
| Transport stochastic indirect effect |       |             |        |     |      |              |
| SIE_tmle_eff       | 0.036 | 332.17     | 319.04 | 0.074 | 0.106 | 0.037        | 0             |
| SIE_EE_eff         | 0.028 | 883.10     | 278.05 | 0.554 | 0.129 | 0.029        | 0             |
| SIE_tmle           | 0.035 | 410.83     | 458.58 | 0.136 | 0.157 | 0.037        | 0             |
| SIE_EE             | 0.027 | 1059.64    | 688.37 | 0.740 | 0.391 | 0.031        | 0             |
| SIE_iptw           | 0.028 | 380.58     | 339.93 | 0.173 | 0.171 | 0.030        | 0             |

Table 8: Simulation results comparing estimators of the transported direct and indirect effects for DGP 1 under well-specified models under the Y and Z models misspecified. For sample sizes 100, 500 and 5000.

| Estimator          | Bias  | Efficiency | 95% CI | Cov | RMSE | % Out of Bds |
|--------------------|-------|------------|--------|-----|------|--------------|
|                    |       | IC Bootstrapping | IC Bootstrapping |     |      |              |
| **DGM 1, N=100**   |       |             |        |     |      |              |
| Transport stochastic direct effect |       |             |        |     |      |              |
| SDE_tmle_eff       | 0.195 | 92.71      | 116.32 | 0.395 | 0.524 | 0.222        | 0             |
| SDE_EE_eff         | 0.216 | 113.89     | 117.45 | 0.453 | 0.469 | 0.241        | 0             |
| SDE_tmle           | 0.076 | 261.30     | 292.91 | 0.824 | 0.865 | 0.323        | 0             |
| SDE_EE             | 0.082 | 326.80     | 293.93 | 0.921 | 0.874 | 0.321        | 0.100         |
| SDE_iptw           | 0.040 | 271.53     | 271.20 | 0.892 | 0.889 | 0.281        | 0             |
| Transport stochastic indirect effect |       |             |        |     |      |              |
| SIE_tmle_eff       | -0.025| 155.53     | 155.39 | 0.740 | 0.688 | 0.055        | 0             |
| SIE_EE_eff         | -0.018| 181.09     | 119.44 | 0.794 | 0.665 | 0.046        | 0             |
| SIE_tmle           | -0.023| 167.73     | 219.28 | 0.664 | 0.756 | 0.069        | 0             |
| SIE_EE             | -0.016| 219.30     | 160.33 | 0.733 | 0.671 | 0.060        | 0             |
| SIE_iptw           | -0.017| 180.22     | 176.80 | 0.818 | 0.803 | 0.048        | 0             |
| **DGM 1, N=5000**  |       |             |        |     |      |              |
| Transport stochastic direct effect |       |             |        |     |      |              |
| SDE_tmle_eff       |       |            |       |     |      |              |
| SDE_EE_eff         |       |            |       |     |      |              |
| SDE_tmle           |       |            |       |     |      |              |
| SDE_EE             |       |            |       |     |      |              |
| SDE_iptw           |       |            |       |     |      |              |

Continued on next page
Table 8 – continued from previous page

| Estimator       | Bias   | Efficiency | 95%CI Cov | RMSE | % Out of Bds |
|-----------------|--------|------------|-----------|------|--------------|
|                 | IC     | Bootstrapping | IC      | Bootstrapping |         |
| **Transport stochastic direct effect** |         |            |          |      |              |
| SDE_tmle_eff    | 0.182  | 102.42     | 114.90   | 0.016 | 0.032        | 0.188   | 0            |
| SDE_EE_eff      | 0.203  | 112.52     | 107.23   | 0.009 | 0.006        | 0.208   | 0            |
| SDE_tmle        | 0.041  | 322.39     | 327.28   | 0.930 | 0.924        | 0.144   | 0            |
| SDE_EE          | 0.039  | 334.96     | 319.05   | 0.942 | 0.929        | 0.138   | 0            |
| SDE_iptw        | 0.006  | 278.81     | 276.85   | 0.936 | 0.929        | 0.115   | 0            |
| **Transport stochastic indirect effect** |         |            |          |      |              |
| SIE_tmle_eff    | -0.006 | 200.08     | 124.34   | 0.903 | 0.808        | 0.011   | 0            |
| SIE_EE_eff      | -0.003 | 212.20     | 126.88   | 0.926 | 0.843        | 0.010   | 0            |
| SIE_tmle        | -0.002 | 241.17     | 229.07   | 0.883 | 0.872        | 0.016   | 0            |
| SIE_EE          | -0.003 | 283.03     | 231.65   | 0.892 | 0.862        | 0.017   | 0            |
| SIE_iptw        | -0.003 | 216.89     | 203.55   | 0.921 | 0.899        | 0.014   | 0            |

**DGM 1, N=5000**

| Estimator       | Bias   | Efficiency | 95%CI Cov | RMSE | % Out of Bds |
|-----------------|--------|------------|-----------|------|--------------|
|                 | IC     | Bootstrapping | IC      | Bootstrapping |         |
| **Transport stochastic direct effect** |         |            |          |      |              |
| SDE_tmle_eff    | 0.181  | 104.08     | 117.69   | 0   | 0            | 0.181   | 0            |
| SDE_EE_eff      | 0.202  | 111.38     | 105.15   | 0   | 0            | 0.202   | 0            |
| SDE_tmle        | 0.050  | 327.02     | 332.07   | 0.782 | 0.791        | 0.064   | 0            |
| SDE_EE          | 0.047  | 328.78     | 317.31   | 0.805 | 0.793        | 0.061   | 0            |
| SDE_iptw        | 0.014  | 275.20     | 271.82   | 0.948 | 0.941        | 0.036   | 0            |
| **Transport stochastic indirect effect** |         |            |          |      |              |
| SIE_tmle_eff    | -0.003 | 179.01     | 95.78    | 0.965 | 0.718        | 0.004   | 0            |
| SIE_EE_eff      | -0.0004| 181.99     | 101.47   | 0.994 | 0.923        | 0.002   | 0            |
| SIE_tmle        | 0.001  | 222.07     | 192.19   | 0.975 | 0.948        | 0.004   | 0            |
| SIE_EE          | 0      | 247.54     | 198.36   | 0.987 | 0.946        | 0.004   | 0            |
| SIE_iptw        | 0      | 195.39     | 170.37   | 0.975 | 0.954        | 0.004   | 0            |

Table 9: Simulation results comparing estimators of the transported direct and indirect effects for DGP 1 under Y and M models misspecified. For sample sizes 100, 500 and 5000.

| Estimator       | Bias   | Efficiency | 95%CI Cov | RMSE | % Out of Bds |
|-----------------|--------|------------|-----------|------|--------------|
|                 | IC     | Bootstrapping | IC      | Bootstrapping |         |
| **DGM 1, N=100** |         |            |          |      |              |
| SDE_tmle_eff    | -0.133 | 245.31     | 239.69   | 0.782 | 0.837        | 0.283   | 0            |
| SDE_EE_eff      | -1.216 | 1,423.80   | 1,532.61 | 0.895 | 0.890        | 1.984   | 51.100       |
| SDE_tmle        | -0.213 | 429.80     | 338.75   | 0.718 | 0.790        | 0.418   | 0            |
| SDE_EE          | -1.259 | 1,611.01   | 1,632.61 | 0.953 | 0.949        | 2.078   | 56.300       |
| SDE_iptw        | -0.187 | 1,098.72   | 320.43   | 0.893 | 0.818        | 0.382   | 0            |
| **DGM 1, N=500** |         |            |          |      |              |
| SIE_tmle_eff    | -0.001 | 233.32     | 208.17   | 0.735 | 0.712        | 0.141   | 0            |
| SIE_EE_eff      | 0.075  | 394.06     | 367.59   | 0.834 | 0.830        | 0.215   | 0.700        |
| SIE_tmle        | -0.044 | 166.40     | 205.14   | 0.539 | 0.575        | 0.139   | 0            |
| SIE_EE          | 0.018  | 382.64     | 295.24   | 0.664 | 0.627        | 0.227   | 1.300        |
| SIE_iptw        | -0.038 | 199.78     | 157.01   | 0.653 | 0.563        | 0.134   | 0            |
| **DGM 1, N=5000** |         |            |          |      |              |
| SDE_tmle_eff    | -0.260 | 303.14     | 180.27   | 0.482 | 0.163        | 0.272   | 0            |
| Continued on next page |         |            |          |      |              |
Table 9 – continued from previous page

| Estimator      | Bias  | Efficiency | 95%CI Cov | RMSE | % Out of Bds |
|----------------|-------|------------|-----------|------|--------------|
|                |       | IC         | Bootstrapping | IC | Bootstrapping |
| SDE_EE_eff     | -1.181 | 1,040.61  | 1,021.97 | 0.080 | 0.066 | 1.272 | 76.400 |
| SDE_tmle       | -0.309 | 670.04     | 318.17 | 0.823 | 0.439 | 0.349 | 0     |
| SDE_EE         | -1.238 | 1,290.55   | 1,219.61 | 0.360 | 0.305 | 1.354 | 77.100 |
| SDE_iptw       | -0.267 | 732.08     | 298.20 | 0.916 | 0.505 | 0.304 | 0     |

Transport stochastic indirect effect

| Estimator      | Bias  | Efficiency | 95%CI Cov | RMSE | % Out of Bds |
|----------------|-------|------------|-----------|------|--------------|
|                |       | IC         | Bootstrapping | IC | Bootstrapping |
| SIE_tmle_eff   | 0.095 | 305.14     | 220.10 | 0.712 | 0.467 | 0.108 | 0     |
| SIE_EE_eff     | 0.137 | 442.19     | 306.25 | 0.806 | 0.424 | 0.156 | 0     |
| SIE_tmle       | 0.060 | 334.84     | 297.52 | 0.892 | 0.808 | 0.094 | 0     |
| SIE_EE         | 0.112 | 591.83     | 473.35 | 0.965 | 0.828 | 0.163 | 0     |
| SIE_iptw       | 0.045 | 316.22     | 259.63 | 0.926 | 0.814 | 0.075 | 0     |

DGM 1, N=5000

Transport stochastic direct effect

| Estimator      | Bias  | Efficiency | 95%CI Cov | RMSE | % Out of Bds |
|----------------|-------|------------|-----------|------|--------------|
|                |       | IC         | Bootstrapping | IC | Bootstrapping |
| SDE_tmle_eff   | -0.272 | 280.08    | 149.17 | 0     | 0     | 0.273 | 0     |
| SDE_EE_eff     | -1.110 | 852.67     | 735.38 | 0     | 0     | 1.118 | 99.210 |
| SDE_tmle       | -0.272 | 698.69     | 321.77 | 0.147 | 0.001 | 0.276 | 0     |
| SDE_EE         | -1.111 | 1,093.33   | 877.83 | 0     | 0     | 1.121 | 98.410 |
| SDE_iptw       | -0.236 | 589.90     | 284.02 | 0.090 | 0     | 0.240 | 0     |

Transport stochastic indirect effect

| Estimator      | Bias  | Efficiency | 95%CI Cov | RMSE | % Out of Bds |
|----------------|-------|------------|-----------|------|--------------|
|                |       | IC         | Bootstrapping | IC | Bootstrapping |
| SIE_tmle_eff   | 0.104 | 249.26     | 173.89 | 0     | 0     | 0.106 | 0     |
| SIE_EE_eff     | 0.135 | 336.39     | 227.03 | 0     | 0     | 0.137 | 0     |
| SIE_tmle       | 0.103 | 370.51     | 279.70 | 0.036 | 0.016 | 0.106 | 0     |
| SIE_EE         | 0.137 | 535.91     | 413.33 | 0.054 | 0.016 | 0.141 | 0     |
| SIE_iptw       | 0.061 | 279.01     | 200.78 | 0.197 | 0.090 | 0.063 | 0     |

Table 10: Simulation results comparing estimators of the transported direct and indirect effects for DGP 1 under the Y model misspecified. For sample sizes 100, 500 and 5000.

| Estimator      | Bias  | Efficiency | 95%CI Cov | RMSE | % Out of Bds |
|----------------|-------|------------|-----------|------|--------------|
|                |       | IC         | Bootstrapping | IC | Bootstrapping |
| DGM 1, N=100   |       |            |            |      |              |

Transport stochastic direct effect

| Estimator      | Bias  | Efficiency | 95%CI Cov | RMSE | % Out of Bds |
|----------------|-------|------------|-----------|------|--------------|
|                |       | IC         | Bootstrapping | IC | Bootstrapping |
| SDE_tmle_eff   | 0.045 | 157.49     | 170.41 | 0.872 | 0.893 | 0.143 | 0     |
| SDE_EE_eff     | 0.050 | 232.07     | 170.51 | 0.894 | 0.872 | 0.149 | 0     |
| SDE_tmle       | 0.010 | 282.59     | 314.26 | 0.876 | 0.912 | 0.273 | 0     |
| SDE_EE         | 0.022 | 355.18     | 289.81 | 0.950 | 0.902 | 0.258 | 0     |
| SDE_iptw       | 0.014 | 297.66     | 306.32 | 0.897 | 0.909 | 0.267 | 0     |

Transport stochastic indirect effect

| Estimator      | Bias  | Efficiency | 95%CI Cov | RMSE | % Out of Bds |
|----------------|-------|------------|-----------|------|--------------|
|                |       | IC         | Bootstrapping | IC | Bootstrapping |
| SIE_tmle_eff   | -0.019 | 161.13   | 174.95 | 0.722 | 0.777 | 0.046 | 0     |
| SIE_EE_eff     | -0.019 | 177.84   | 127.50 | 0.751 | 0.710 | 0.044 | 0     |
| SIE_tmle       | -0.020 | 175.47   | 241.36 | 0.638 | 0.742 | 0.055 | 0     |
| SIE_EE         | -0.017 | 211.04   | 162.19 | 0.705 | 0.668 | 0.051 | 0     |
| SIE_iptw       | -0.018 | 202.31   | 200.29 | 0.764 | 0.749 | 0.051 | 0     |

DGM 1, N=500

Transport stochastic direct effect

| Estimator      | Bias  | Efficiency | 95%CI Cov | RMSE | % Out of Bds |
|----------------|-------|------------|-----------|------|--------------|
|                |       | IC         | Bootstrapping | IC | Bootstrapping |
| SDE_tmle_eff   | 0.013 | 185.58   | 128.33 | 0.983 | 0.918 | 0.053 | 0     |
| SDE_EE_eff     | 0.014 | 240.63   | 137.63 | 0.990 | 0.915 | 0.056 | 0     |
| SDE_tmle       | -0.0001 | 339.55 | 291.55 | 0.975 | 0.939 | 0.112 | 0     |

Continued on next page
Table 10 – continued from previous page

| Estimator          | Bias  | Efficiency | 95%CI Cov Bootstrapping | RMSE | % Out of Bds |
|--------------------|-------|------------|--------------------------|------|--------------|
|                    |       | IC         | IC Bootstrapping         |      |              |
| SDE_EE             | 0.002 | 370.42     | 288.90                   | 0.988| 0.939        | 0.112 | 0 |
| SDE_iptw           | 0.005 | 299.38     | 283.42                   | 0.957| 0.943        | 0.108 | 0 |

Transport stochastic indirect effect

| Estimator          | Bias  | Efficiency | 95%CI Cov Bootstrapping | RMSE | % Out of Bds |
|--------------------|-------|------------|--------------------------|------|--------------|
|                    |       | IC         | IC Bootstrapping         |      |              |
| SIE_tmle_eff       | -0.002| 217.44     | 131.94                   | 0.930| 0.880        | 0.010 | 0 |
| SIE_EE_eff         | -0.003| 251        | 128.78                   | 0.926| 0.842        | 0.010 | 0 |
| SIE_tmle           | -0.003| 255.86     | 222.71                   | 0.883| 0.879        | 0.015 | 0 |
| SIE_EE             | -0.004| 308.32     | 222.67                   | 0.885| 0.833        | 0.016 | 0 |
| SIE_iptw           | -0.002| 249.26     | 214.45                   | 0.905| 0.889        | 0.015 | 0 |

DGM 1, N=5000

Transport stochastic direct effect

| Estimator          | Bias  | Efficiency | 95%CI Cov Bootstrapping | RMSE | % Out of Bds |
|--------------------|-------|------------|--------------------------|------|--------------|
|                    |       | IC         | IC Bootstrapping         |      |              |
| SDE_tmle_eff       | -0.0002| 191.06     | 118.82                   | 0.999| 0.965        | 0.014 | 0 |
| SDE_EE_eff         | -0.0004| 240.26     | 126.67                   | 0.999| 0.968        | 0.015 | 0 |
| SDE_tmle           | 0      | 354.81     | 288.79                   | 0.984| 0.954        | 0.035 | 0 |
| SDE_EE             | -0.0002| 374.95     | 291.49                   | 0.992| 0.957        | 0.035 | 0 |
| SDE_iptw           | -0.0001| 300.52     | 276.72                   | 0.973| 0.957        | 0.034 | 0 |

Table 11: Simulation results comparing estimators of the transported direct and indirect effects for DGP 2 under well-specified models for sample sizes 100, 500 and 5000.

| Estimator          | Bias  | Efficiency | 95%CI Cov Bootstrapping | RMSE | % Out of Bds |
|--------------------|-------|------------|--------------------------|------|--------------|
|                    |       | IC         | IC Bootstrapping         |      |              |
| DGM 2, N=100       |       |            |                          |      |              |

Transport stochastic direct effect

| Estimator          | Bias  | Efficiency | 95%CI Cov Bootstrapping | RMSE | % Out of Bds |
|--------------------|-------|------------|--------------------------|------|--------------|
|                    |       | IC         | IC Bootstrapping         |      |              |
| SDE_tmle_eff       | 0.003 | 102.26     | 139.52                   | 0.970| 0.992        | 0.064 | 0 |
| SDE_EE_eff         | 0.004 | 102.13     | 118.65                   | 0.974| 0.989        | 0.060 | 0 |
| SDE_tmle           | 0.007 | 293.07     | 375.10                   | 0.852| 0.945        | 0.218 | 0 |
| SDE_EE             | 0.005 | 316.03     | 292.87                   | 0.957| 0.930        | 0.186 | 0 |
| SDE_iptw           | 0.009 | 442.80     | 456.03                   | 0.909| 0.922        | 0.267 | 0 |

Transport stochastic indirect effect

| Estimator          | Bias  | Efficiency | 95%CI Cov Bootstrapping | RMSE | % Out of Bds |
|--------------------|-------|------------|--------------------------|------|--------------|
|                    |       | IC         | IC Bootstrapping         |      |              |
| SIE_tmle_eff       | -0.003| 93.01      | 132.95                   | 0.878| 0.944        | 0.033 | 0 |
| SIE_EE_eff         | -0.003| 93.03      | 113.13                   | 0.878| 0.924        | 0.033 | 0 |
| SIE_tmle           | -0.001| 99.85      | 158.57                   | 0.865| 0.952        | 0.041 | 0 |
| SIE_EE             | -0.002| 99.09      | 117.32                   | 0.871| 0.915        | 0.036 | 0 |
| SIE_iptw           | 0.002 | 167.51     | 206.17                   | 0.812| 0.867        | 0.062 | 0 |

DGM 2, N=500

Transport stochastic direct effect

| Estimator          | Bias  | Efficiency | 95%CI Cov Bootstrapping | RMSE | % Out of Bds |
|--------------------|-------|------------|--------------------------|------|--------------|
|                    |       | IC         | IC Bootstrapping         |      |              |
| SDE_tmle_eff       | -0.0003| 101.37     | 105.16                   | 0.946| 0.957        | 0.028 | 0 |
| SDE_EE_eff         | -0.0003| 101.40     | 104.51                   | 0.946| 0.957        | 0.028 | 0 |
| SDE_tmle           | -0.004| 319.08     | 327.43                   | 0.929| 0.936        | 0.089 | 0 |
| SDE_EE             | -0.004| 321.97     | 316.56                   | 0.947| 0.935        | 0.088 | 0 |
| SDE_iptw           | -0.001| 428.43     | 419.16                   | 0.951| 0.942        | 0.110 | 0 |

Continued on next page
Table 11 – continued from previous page

| Estimator            | Bias      | Efficiency | 95% CI | Cov       | RMSE     | % Out of Bds |
|----------------------|-----------|------------|--------|-----------|----------|--------------|
|                      | IC Bootstrapping | IC Bootstrapping |       |           |          |              |
| Transport stochastic indirect effect |
| SIE_tmle_eff          | -0.001    | 99.56      | 104.59 | 0.926     | 0.929    | 0.012        | 0            |
| SIE_EE_eff            | -0.001    | 99.58      | 101.11 | 0.926     | 0.929    | 0.012        | 0            |
| SIE_tmle              | -0.0005   | 101.74     | 106.99 | 0.928     | 0.935    | 0.012        | 0            |
| SIE_EE                | -0.0005   | 101.94     | 103.39 | 0.932     | 0.930    | 0.012        | 0            |
| SIE_iptw              | -0.0005   | 177.68     | 156.98 | 0.948     | 0.936    | 0.018        | 0            |
| DGM 2, N=5000         |
| Transport stochastic direct effect |
| SDE_tmle_eff          | -0.0003   | 100.17     | 100.42 | 0.955     | 0.957    | 0.008        | 0            |
| SDE_EE_eff            | -0.0003   | 100.21     | 100.42 | 0.955     | 0.957    | 0.008        | 0            |
| SDE_tmle              | 0.001     | 321.02     | 321.92 | 0.945     | 0.942    | 0.027        | 0            |
| SDE_EE                | 0.001     | 321.27     | 321    | 0.946     | 0.942    | 0.027        | 0            |
| SDE_iptw              | 0.0002    | 412.67     | 388.11 | 0.954     | 0.945    | 0.033        | 0            |
| Transport stochastic indirect effect |
| SIE_tmle_eff          | -0.0001   | 99.98      | 100.46 | 0.942     | 0.942    | 0.004        | 0            |
| SIE_EE_eff            | -0.0001   | 100.01     | 100.46 | 0.942     | 0.942    | 0.004        | 0            |
| SIE_tmle              | -0.0001   | 101.75     | 101.99 | 0.942     | 0.942    | 0.004        | 0            |
| SIE_EE                | -0.0001   | 101.69     | 101.95 | 0.942     | 0.943    | 0.004        | 0            |
| SIE_iptw              | 0         | 166.56     | 141.45 | 0.980     | 0.957    | 0.005        | 0            |

Table 12: Simulation results comparing estimators of the transported direct and indirect effects for DGP 2 under Z, M, and S models misspecified. For sample sizes 100, 500 and 5000.

| Estimator            | Bias      | Efficiency | 95% CI | Cov       | RMSE     | % Out of Bds |
|----------------------|-----------|------------|--------|-----------|----------|--------------|
|                      | IC Bootstrapping | IC Bootstrapping |       |           |          |              |
| DGM 2, N=100         |
| Transport stochastic direct effect |
| SDE_tmle_eff          | 0.001     | 90.70      | 108.20 | 0.918     | 0.988    | 0.062        | 0            |
| SDE_EE_eff            | 0.001     | 90.72      | 108.25 | 0.918     | 0.988    | 0.062        | 0            |
| SDE_tmle              | 0.007     | 191.08     | 235.80 | 0.861     | 0.956    | 0.141        | 0            |
| SDE_EE                | 0.004     | 207.43     | 191.55 | 0.957     | 0.951    | 0.121        | 0            |
| SDE_iptw              | 0.206     | 367.51     | 375.66 | 0.729     | 0.756    | 0.345        | 0            |
| Transport stochastic indirect effect |
| SIE_tmle_eff          | 0.002     | 51.42      | 132.19 | 0.575     | 0.927    | 0.082        | 0            |
| SIE_EE_eff            | 0.002     | 51.31      | 128.88 | 0.577     | 0.920    | 0.080        | 0            |
| SIE_tmle              | 0.002     | 96.11      | 158.41 | 0.721     | 0.892    | 0.099        | 0            |
| SIE_EE                | 0.002     | 109.50     | 151.53 | 0.779     | 0.897    | 0.098        | 0            |
| SIE_iptw              | -0.054    | 537.90     | 138.25 | 0.886     | 0.846    | 0.110        | 0            |
| DGM 2, N=500          |
| Transport stochastic direct effect |
| SDE_tmle_eff          | 0.0002    | 99.14      | 105.22 | 0.946     | 0.969    | 0.028        | 0            |
| SDE_EE_eff            | 0.0002    | 99.14      | 105.22 | 0.946     | 0.969    | 0.028        | 0            |
| SDE_tmle              | -0.002    | 222.27     | 220.01 | 0.941     | 0.947    | 0.061        | 0            |
| SDE_EE                | -0.002    | 224.49     | 199.30 | 0.971     | 0.952    | 0.055        | 0            |
| SDE_iptw              | 0.247     | 396.13     | 428.63 | 0.385     | 0.423    | 0.273        | 0            |
| Transport stochastic indirect effect |

Continued on next page
### Table 12 – continued from previous page

| Estimator | Bias  | Efficiency | 95% CI Cov | RMSE | % Out of Bds |
|-----------|-------|------------|------------|------|--------------|
|           |       | IC Bootstrapping | IC Bootstrapping |     |              |
| SIE_tmle_eff | -0.0004 | 51.46 | 110.33 | 0.622 | 0.934 | 0.032 | 0 |
| SIE_EE_eff  | -0.0005 | 51.46 | 109.03 | 0.625 | 0.933 | 0.032 | 0 |
| SIE_tmle    | 0.001  | 110.16 | 133.88 | 0.850 | 0.921 | 0.044 | 0 |
| SIE_EE      | 0.0003 | 114.19 | 136.01 | 0.871 | 0.926 | 0.047 | 0 |
| SIE_iptw    | 0.005  | 389.26 | 142.21 | 0.972 | 0.943 | 0.044 | 0 |

| DGM 2, N=5000 |

#### Transport stochastic direct effect

| Estimator | Bias  | Efficiency | 95% CI Cov | RMSE | % Out of Bds |
|-----------|-------|------------|------------|------|--------------|
|           |       | IC Bootstrapping | IC Bootstrapping |     |              |
| SDE_tmle_eff | 0      | 98.78 | 100.01 | 0.956 | 0.960 | 0.008 | 0 |
| SDE_EE_eff  | 0      | 98.80 | 100.03 | 0.956 | 0.960 | 0.008 | 0 |
| SDE_tmle    | 0.001  | 228.37 | 231.43 | 0.939 | 0.941 | 0.020 | 0 |
| SDE_EE      | 0.001  | 228.50 | 203.24 | 0.966 | 0.942 | 0.018 | 0 |
| SDE_iptw    | 0.254  | 396.11 | 406.66 | 0      | 0      | 0.256 | 0 |

#### Transport stochastic indirect effect

| Estimator | Bias  | Efficiency | 95% CI Cov | RMSE | % Out of Bds |
|-----------|-------|------------|------------|------|--------------|
|           |       | IC Bootstrapping | IC Bootstrapping |     |              |
| SIE_tmle_eff | 0.0005 | 54.95 | 100.51 | 0.727 | 0.938 | 0.010 | 0 |
| SIE_EE_eff  | 0.0004 | 54.95 | 100.46 | 0.732 | 0.939 | 0.010 | 0 |
| SIE_tmle    | 0.001  | 129.68 | 136.04 | 0.879 | 0.920 | 0.014 | 0 |
| SIE_EE      | 0.001  | 130.08 | 137.03 | 0.891 | 0.930 | 0.014 | 0 |
| SIE_iptw    | 0.014  | 301.92 | 123.87 | 0.976 | 0.716 | 0.018 | 0 |

### Table 13: Simulation results comparing estimators of the transported direct and indirect effects for DGP 2 under \( Y \) and \( S \) models misspecified. For sample sizes 100, 500 and 5000.

| Estimator | Bias  | Efficiency | 95% CI Cov | RMSE | % Out of Bds |
|-----------|-------|------------|------------|------|--------------|
|           |       | IC Bootstrapping | IC Bootstrapping |     |              |
| DGM 2, N=100 |

#### Transport stochastic direct effect

| Estimator | Bias  | Efficiency | 95% CI Cov | RMSE | % Out of Bds |
|-----------|-------|------------|------------|------|--------------|
|           |       | IC Bootstrapping | IC Bootstrapping |     |              |
| SDE_tmle_eff | -0.048 | 198.25 | 193.84 | 0.963 | 0.921 | 0.146 | 0 |
| SDE_EE_eff  | -0.102 | 315.28 | 200.46 | 0.946 | 0.875 | 0.181 | 0 |
| SDE_tmle    | -0.057 | 317.69 | 365.58 | 0.819 | 0.880 | 0.245 | 0 |
| SDE_EE      | -0.102 | 554.59 | 425.87 | 0.960 | 0.811 | 0.340 | 1.100 |
| SDE_iptw    | 0.054  | 332.61 | 378.44 | 0.844 | 0.870 | 0.249 | 0 |

#### Transport stochastic indirect effect

| Estimator | Bias  | Efficiency | 95% CI Cov | RMSE | % Out of Bds |
|-----------|-------|------------|------------|------|--------------|
|           |       | IC Bootstrapping | IC Bootstrapping |     |              |
| SIE_tmle_eff | 0.142  | 362.76 | 392.68 | 0.589 | 0.652 | 0.200 | 0 |
| SIE_EE_eff  | 0.189  | 708.70 | 403.57 | 0.940 | 0.517 | 0.246 | 0 |
| SIE_tmle    | 0.118  | 368.02 | 420.49 | 0.606 | 0.748 | 0.201 | 0 |
| SIE_EE      | 0.184  | 749.61 | 475.97 | 0.944 | 0.682 | 0.267 | 0 |
| SIE_iptw    | 0.182  | 384.16 | 427.11 | 0.513 | 0.627 | 0.244 | 0 |

### DGM 2, N=5000

#### Transport stochastic direct effect

| Estimator | Bias  | Efficiency | 95% CI Cov | RMSE | % Out of Bds |
|-----------|-------|------------|------------|------|--------------|
|           |       | IC Bootstrapping | IC Bootstrapping |     |              |
| SDE_tmle_eff | -0.162 | 309.07 | 214.95 | 0.395 | 0.121 | 0.170 | 0 |
| SDE_EE_eff  | -0.219 | 453.68 | 234.34 | 0.553 | 0.029 | 0.227 | 0 |
| SDE_tmle    | -0.173 | 519.71 | 423.35 | 0.712 | 0.599 | 0.206 | 0 |
| SDE_EE      | -0.226 | 758.85 | 619.39 | 0.817 | 0.613 | 0.281 | 0 |
| SDE_iptw    | -0.069 | 429.11 | 376.48 | 0.942 | 0.794 | 0.118 | 0 |

#### Transport stochastic indirect effect

| Estimator | Bias  | Efficiency | 95% CI Cov | RMSE | % Out of Bds |
|-----------|-------|------------|------------|------|--------------|
|           |       | IC Bootstrapping | IC Bootstrapping |     |              |
| SIE_tmle_eff | 0.220  | 529.14 | 274.22 | 0.016 | 0.012 | 0.222 | 0 |
| SIE_EE_eff  | 0.277  | 1,080.48 | 309.93 | 0.203 | 0.006 | 0.280 | 0 |

Continued on next page
Table 13 – continued from previous page

| Estimator  | Bias      | Efficiency | 95% CI Cov | RMSE | % Out of Bds |
|------------|-----------|------------|------------|------|--------------|
|            | IC        | Bootstrapping | IC        | Bootstrapping |     |
| SIE_tmle   | 0.213     | 600.82     | 434.83     | 0.060 | 0.052 | 0.219 | 0  |
| SIE_EE     | 0.278     | 1,253.70   | 726.29     | 0.442 | 0.132 | 0.292 | 0  |
| SIE_iptw   | 0.273     | 544.21     | 457.14     | 0.022 | 0.031 | 0.278 | 0  |

DGM 2, N=5000

Transport stochastic direct effect

| Estimator  | Bias      | Efficiency | 95% CI Cov | RMSE | % Out of Bds |
|------------|-----------|------------|------------|------|--------------|
|            | IC        | Bootstrapping | IC        | Bootstrapping |     |
| SDE_tmle_eff | -0.164    | 285.91     | 168.14     | 0    | 0    | 0.165 | 0  |
| SDE_EE_eff  | -0.220    | 415.64     | 191.91     | 0    | 0    | 0.22  | 0  |
| SDE_tmle    | -0.164    | 511.22     | 399.63     | 0.019 | 0.010 | 0.167 | 0  |
| SDE_EE      | -0.218    | 699.90     | 586.82     | 0.029 | 0.023 | 0.224 | 0  |
| SDE_iptw    | -0.067    | 404.82     | 322.51     | 0.467 | 0.319 | 0.073 | 0  |

Transport stochastic indirect effect

| Estimator  | Bias      | Efficiency | 95% CI Cov | RMSE | % Out of Bds |
|------------|-----------|------------|------------|------|--------------|
|            | IC        | Bootstrapping | IC        | Bootstrapping |     |
| SIE_tmle_eff | 0.222     | 507.95     | 180.55     | 0    | 0    | 0.223 | 0  |
| SIE_EE_eff  | 0.278     | 1,033.84   | 216.14     | 0    | 0    | 0.279 | 0  |
| SIE_tmle    | 0.222     | 600.12     | 343.91     | 0    | 0    | 0.222 | 0  |
| SIE_EE      | 0.278     | 1,201.77   | 675.26     | 0    | 0    | 0.279 | 0  |
| SIE_iptw    | 0.278     | 520.31     | 337.87     | 0    | 0    | 0.278 | 0  |

Table 14: Simulation results comparing estimators of the transported direct and indirect effects for DGP 2 under Y and Z models misspecified. For sample sizes 100, 500 and 5000.

| Estimator  | Bias      | Efficiency | 95% CI Cov | RMSE | % Out of Bds |
|------------|-----------|------------|------------|------|--------------|
|            | IC        | Bootstrapping | IC        | Bootstrapping |     |
| DGM 2, N=100

Transport stochastic direct effect

| Estimator  | Bias      | Efficiency | 95% CI Cov | RMSE | % Out of Bds |
|------------|-----------|------------|------------|------|--------------|
|            | IC        | Bootstrapping | IC        | Bootstrapping |     |
| SDE_tmle_eff | 0.028     | 73.84      | 84.89      | 0.881 | 0.976 | 0.053 | 0  |
| SDE_EE_eff  | 0.031     | 81.31      | 95.66      | 0.884 | 0.969 | 0.058 | 0  |
| SDE_tmle    | 0.146     | 892.01     | 513.89     | 0.914 | 0.847 | 0.360 | 0  |
| SDE_EE      | 0.576     | 1,247.94   | 1,341.48   | 0.935 | 0.923 | 0.966 | 24.300 |
| SDE_iptw    | 0.105     | 878.55     | 480.40     | 0.945 | 0.856 | 0.326 | 0  |

Transport stochastic indirect effect

| Estimator  | Bias      | Efficiency | 95% CI Cov | RMSE | % Out of Bds |
|------------|-----------|------------|------------|------|--------------|
|            | IC        | Bootstrapping | IC        | Bootstrapping |     |
| SIE_tmle_eff | -0.046    | 165.05     | 120.15     | 0.815 | 0.728 | 0.086 | 0  |
| SIE_EE_eff  | -0.029    | 249.13     | 155.58     | 0.819 | 0.747 | 0.105 | 0  |
| SIE_tmle    | -0.068    | 413.29     | 276.43     | 0.799 | 0.695 | 0.132 | 0  |
| SIE_EE      | -0.003    | 724.66     | 686.48     | 0.825 | 0.837 | 0.467 | 5  |
| SIE_iptw    | -0.069    | 644.23     | 139.24     | 0.770 | 0.529 | 0.104 | 0  |

DGM 2, N=500

Transport stochastic direct effect

| Estimator  | Bias      | Efficiency | 95% CI Cov | RMSE | % Out of Bds |
|------------|-----------|------------|------------|------|--------------|
|            | IC        | Bootstrapping | IC        | Bootstrapping |     |
| SDE_tmle_eff | 0.031     | 74.57      | 78.71      | 0.647 | 0.693 | 0.038 | 0  |
| SDE_EE_eff  | 0.033     | 79.85      | 84.01      | 0.658 | 0.697 | 0.040 | 0  |
| SDE_tmle    | 0.179     | 1,002.95   | 559.87     | 0.991 | 0.735 | 0.224 | 0  |
| SDE_EE      | 0.555     | 1,171.51   | 1,116.10   | 0.584 | 0.545 | 0.626 | 6.800 |
| SDE_iptw    | 0.103     | 740.10     | 476.14     | 0.995 | 0.851 | 0.164 | 0  |

Transport stochastic indirect effect

| Estimator  | Bias      | Efficiency | 95% CI Cov | RMSE | % Out of Bds |
|------------|-----------|------------|------------|------|--------------|
|            | IC        | Bootstrapping | IC        | Bootstrapping |     |
| SIE_tmle_eff | -0.015    | 200.43     | 101.69     | 0.960 | 0.728 | 0.019 | 0  |
| SIE_EE_eff  | 0.017     | 329.17     | 147.24     | 0.993 | 0.808 | 0.026 | 0  |
| SIE_tmle    | -0.040    | 469.27     | 531        | 0.853 | 0.667 | 0.064 | 0  |
| SIE_EE      | 0.075     | 931.81     | 763.34     | 0.909 | 0.817 | 0.156 | 0.200 |

Continued on next page
Table 14 – continued from previous page

| Estimator | Bias  | Efficiency | 95%CI Cov | RMSE   | % Out of Bds |
|-----------|-------|------------|-----------|--------|-------------|
|           |       | IC         | Bootstrapping | IC     | Bootstrapping |
| SIE_iptw  | -0.033| 518.39     | 166.53     | 0.729  | 0.519        | 0.041 | 0           |
| DGM 2, N=5000 |

Transport stochastic direct effect

| Estimator          | Bias  | Efficiency | 95%CI Cov | RMSE   | % Out of Bds |
|--------------------|-------|------------|-----------|--------|-------------|
| SDE_tmle_eff       | 0.032 | 75.49      | 77.35     | 0      | 0           |
| SDE_EE_eff         | 0.033 | 80.22      | 81.92     | 0      | 0           |
| SDE_tmle           | 0.188 | 973.02     | 463.63    | 0.223  | 0.003       |
| SDE_EE             | 0.545 | 1,112.09   | 1,008.79  | 0      | 0           |
| SDE_iptw           | 0.108 | 700.28     | 456.29    | 0.574  | 0.204       |

Transport stochastic indirect effect

| Estimator          | Bias  | Efficiency | 95%CI Cov | RMSE   | % Out of Bds |
|--------------------|-------|------------|-----------|--------|-------------|
| SIE_tmle_eff       | -0.014| 192.60     | 92.53     | 0.380  | 0.004       |
| SIE_EE_eff         | 0.018 | 307.60     | 121.67    | 0.767  | 0.029       |
| SIE_tmle           | -0.040| 492.45     | 143.03    | 0.185  | 0           |
| SIE_EE             | 0.066 | 911.65     | 686.60    | 0.447  | 0.231       |
| SIE_iptw           | -0.031| 439.30     | 160.87    | 0.430  | 0.004       |

Table 15: Simulation results comparing estimators of the transported direct and indirect effects for DGP 2 under Y and M models misspecified. For sample sizes 100, 500 and 5000.

| Estimator          | Bias  | Efficiency | 95%CI Cov | RMSE   | % Out of Bds |
|--------------------|-------|------------|-----------|--------|-------------|
|                   |       | IC         | Bootstrapping | IC     | Bootstrapping |
| DGM 2, N=100 |

Transport stochastic direct effect

| Estimator          | Bias  | Efficiency | 95%CI Cov | RMSE   | % Out of Bds |
|--------------------|-------|------------|-----------|--------|-------------|
| SDE_tmle_eff       | 0.028 | 122.23     | 210.47    | 0.825  | 0.985       |
| SDE_EE_eff         | 0.055 | 561.86     | 1131.74   | 0.905  | 0.998       |
| SDE_tmle           | 0.034 | 797.09     | 552.54    | 0.834  | 0.825       |
| SDE_EE             | 0.026 | 1,816.94   | 1828.34   | 0.974  | 0.972       |
| SDE_iptw           | 0.028 | 639.31     | 509.92    | 0.872  | 0.851       |

Transport stochastic indirect effect

| Estimator          | Bias  | Efficiency | 95%CI Cov | RMSE   | % Out of Bds |
|--------------------|-------|------------|-----------|--------|-------------|
| SIE_tmle_eff       | -0.030| 283.99     | 172.61    | 0.842  | 0.780       |
| SIE_EE_eff         | -0.030| 534.83     | 549.34    | 0.885  | 0.908       |
| SIE_tmle           | -0.030| 252.15     | 207.16    | 0.686  | 0.734       |
| SIE_EE             | 0.034 | 517.82     | 503.88    | 0.839  | 0.830       |
| SIE_iptw           | -0.021| 268.93     | 186.20    | 0.692  | 0.769       |

DGM 2, N=5000 |

Transport stochastic direct effect

| Estimator          | Bias  | Efficiency | 95%CI Cov | RMSE   | % Out of Bds |
|--------------------|-------|------------|-----------|--------|-------------|
| SDE_tmle_eff       | 0.038 | 79.10      | 179.89    | 0.506  | 0.918       |
| SDE_EE_eff         | 0.077 | 208.62     | 728.54    | 0.299  | 0.970       |
| SDE_tmle           | 0.033 | 1,209.81   | 582.41    | 0.991  | 0.968       |
| SDE_EE             | 0.077 | 1,870.98   | 1,799.55  | 0.987  | 0.990       |
| SDE_iptw           | 0.022 | 737.59     | 523.97    | 0.989  | 0.963       |

Transport stochastic indirect effect

| Estimator          | Bias  | Efficiency | 95%CI Cov | RMSE   | % Out of Bds |
|--------------------|-------|------------|-----------|--------|-------------|
| SIE_tmle_eff       | -0.036| 213.86     | 103.92    | 0.525  | 0.230       |
| SIE_EE_eff         | -0.018| 252.81     | 228.74    | 0.395  | 0.324       |
| SIE_tmle           | -0.035| 290.63     | 159.72    | 0.681  | 0.399       |
| SIE_EE             | -0.017| 354.25     | 346.99    | 0.588  | 0.565       |
| SIE_iptw           | -0.029| 157.96     | 135.95    | 0.537  | 0.425       |

DGM 2, N=5000 |

Continued on next page
Table 15 – continued from previous page

| Estimator       | Bias  | Efficiency | 95%CI | Cov RMSE | % Out of Bds |
|-----------------|-------|------------|-------|----------|--------------|
|                 | IC Bootstrapping | IC Bootstrapping |       |          |              |
| **Transport stochastic direct effect** |       |            |       |          |              |
| SDE_tmle_eff    | 0.039 | 72.52      | 169.93| 0.039    | 0.211        |
| SDE_EE_eff      | 0.074 | 116.09     | 615.71| 0.073    | 0.715        |
| SDE_tmle        | 0.040 | 1,191.15   | 425.74| 1        | 0.780        |
| SDE_EE          | 0.077 | 1,623.22   | 1278.49| 0.955    | 0.889        |
| SDE_iptw        | 0.026 | 710.95     | 429.68| 0.992    | 0.874        |
| **Transport stochastic indirect effect** |       |            |       |          |              |
| SIE_tmle_eff    | -0.037| 130.59     | 63.97 | 0.002    | 0.037        |
| SIE_EE_eff      | -0.020| 126.70     | 105.49| 0.178    | 0.129        |
| SIE_tmle        | -0.036| 176.50     | 85.45 | 0.005    | 0.037        |
| SIE_EE          | -0.020| 164.23     | 147.83| 0.232    | 0.206        |
| SIE_iptw        | -0.030| 134.37     | 83.04 | 0        | 0.031        |

Table 16: Simulation results comparing estimators of the transported direct and indirect effects for DGP 2 under $Y$ model misspecified. For sample sizes 100, 500 and 5000.

| Estimator       | Bias  | Efficiency | 95%CI | Cov RMSE | % Out of Bds |
|-----------------|-------|------------|-------|----------|--------------|
|                 | IC Bootstrapping | IC Bootstrapping |       |          |              |
| **DGM 2, N=100** |       |            |       |          |              |
| SDE_tmle_eff    | 0.016 | 117.52     | 135.93| 0.928    | 0.991        |
| SDE_EE_eff      | 0.017 | 134.03     | 127.09| 0.971    | 0.994        |
| SDE_tmle        | -0.005| 437.29     | 464.11| 0.856    | 0.888        |
| SDE_EE          | 0.003 | 510.08     | 415.06| 0.971    | 0.887        |
| SDE_iptw        | -0.004| 437.59     | 450.42| 0.892    | 0.911        |
| **DGM 2, N=500** |       |            |       |          |              |
| SIE_tmle_eff    | -0.015| 154.70     | 134.11| 0.741    | 0.759        |
| SIE_EE_eff      | -0.016| 158.75     | 111.49| 0.752    | 0.711        |
| SIE_tmle        | -0.013| 170.37     | 205.62| 0.649    | 0.798        |
| SIE_EE          | -0.015| 185.12     | 139.18| 0.700    | 0.667        |
| SIE_iptw        | -0.004| 166.40     | 199.63| 0.795    | 0.857        |
| **DGM 2, N=5000** |       |            |       |          |              |
| SIE_tmle_eff    | -0.001| 119.51     | 107.23| 0.981    | 0.978        |
| SIE_EE_eff      | -0.001| 122.98     | 106.84| 0.990    | 0.978        |
| SIE_tmle        | -0.002| 527.31     | 463.49| 0.974    | 0.950        |
| SIE_EE          | -0.0005| 539.36    | 468.44| 0.981    | 0.957        |
| SIE_iptw        | -0.001| 429.84     | 422.32| 0.959    | 0.951        |
| **DGM 2, N=5000** |       |            |       |          |              |
| SIE_tmle_eff    | 0.0002| 207.64     | 113.29| 0.993    | 0.962        |
| SIE_EE_eff      | 0.0003| 204.81     | 112.47| 0.993    | 0.962        |
| SIE_tmle        | -0.0001| 243.79   | 185.49| 0.979    | 0.951        |
| SIE_EE          | -0.0001| 240.52   | 177.94| 0.978    | 0.944        |
| SIE_iptw        | 0.0003| 177.08     | 158.06| 0.969    | 0.953        |
| **DGM 2, N=5000** |       |            |       |          |              |
| SIE_tmle_eff    | -0.0003| 111.40   | 100.18| 0.963    | 0.936        |
| **Continued on next page** |       |            |       |          |              |
Table 16 - continued from previous page

| Estimator          | Bias  | Efficiency | 95%CI | Cov | RMSE | % Out of Bds |
|--------------------|-------|------------|-------|-----|------|--------------|
|                    |       | IC         | Bootstrapping | IC | Bootstrapping |          |
| SDE_EE_eff         | -0.0003 | 111.55 | 100.18 | 0.968 | 0.937 | 0.009 | 0 |
| SDE_tmle           | 0.002  | 508.10 | 414.57 | 0.984 | 0.947 | 0.034 | 0 |
| SDE_EE             | 0.002  | 508.50 | 414.73 | 0.984 | 0.947 | 0.034 | 0 |
| SDE_iptw           | 0.002  | 413.87 | 389.34 | 0.958 | 0.946 | 0.032 | 0 |

Transport stochastic indirect effect

| Estimator          | Bias  | Efficiency | 95%CI | Cov | RMSE | % Out of Bds |
|--------------------|-------|------------|-------|-----|------|--------------|
|                    |       | IC         | Bootstrapping | IC | Bootstrapping |          |
| SIE_tmle_eff       | -0.0001 | 195.72 | 101.34 | 0.999 | 0.941 | 0.004 | 0 |
| SIE_EE_eff         | -0.0001 | 189.44 | 101.82 | 0.999 | 0.942 | 0.004 | 0 |
| SIE_tmle           | -0.0002 | 231.77 | 163.22 | 0.992 | 0.954 | 0.006 | 0 |
| SIE_EE             | -0.0002 | 220.13 | 155.51 | 0.993 | 0.958 | 0.005 | 0 |
| SIE_iptw           | -0.0002 | 166.19 | 141.12 | 0.981 | 0.951 | 0.005 | 0 |

Table 17: Simulation results comparing estimators of the transported direct and indirect effects for DGP 3 under well-specified models for sample sizes 100, 500 and 5000.

| Estimator          | Bias  | Efficiency | 95%CI | Cov | RMSE | % Out of Bds |
|--------------------|-------|------------|-------|-----|------|--------------|
|                    |       | IC         | Bootstrapping | IC | Bootstrapping |          |
| DGM 3, N=100       |       |            |            |     |      |             |
| Transport stochastic direct effect
| SDE_tmle_eff       | -0.020 | 93.10 | 158.60 | 0.823 | 0.955 | 0.115 | 0 |
| SDE_EE_eff         | -0.017 | 93.22 | 113.95 | 0.828 | 0.986 | 0.099 | 0 |
| SDE_tmle           | -0.030 | 126.80 | 239.56 | 0.809 | 0.951 | 0.199 | 0 |
| SDE_EE             | -0.020 | 139.54 | 141.15 | 0.890 | 0.875 | 0.140 | 0 |
| SDE_iptw           | -0.043 | 346.24 | 324.14 | 0.840 | 0.871 | 0.327 | 0 |
| Transport stochastic indirect effect
| SIE_tmle_eff       | 0.011  | 90.59 | 159.14 | 0.803 | 0.929 | 0.114 | 0 |
| SIE_EE_eff         | 0.008  | 90.40 | 159.14 | 0.809 | 0.905 | 0.104 | 0 |
| SIE_tmle           | 0.018  | 90.81 | 173.84 | 0.742 | 0.907 | 0.152 | 0 |
| SIE_EE             | 0.009  | 95.89 | 173.84 | 0.782 | 0.886 | 0.107 | 0 |
| SIE_iptw           | 0.085  | 154.04 | 133.33 | 0.634 | 0.603 | 0.192 | 0 |
| DGM 3, N=500       |       |            |            |     |      |             |
| Transport stochastic direct effect
| SDE_tmle_eff       | -0.002 | 99.42 | 106.45 | 0.931 | 0.948 | 0.043 | 0 |
| SDE_EE_eff         | -0.002 | 99.45 | 106.04 | 0.931 | 0.947 | 0.043 | 0 |
| SDE_tmle           | -0.003 | 168.58 | 169.76 | 0.933 | 0.939 | 0.080 | 0 |
| SDE_EE             | -0.003 | 177.72 | 170.88 | 0.960 | 0.941 | 0.085 | 0 |
| SDE_iptw           | 0.042  | 466.12 | 416.46 | 0.773 | 0.780 | 0.194 | 0 |
| Transport stochastic indirect effect
| SIE_tmle_eff       | -0.0005 | 100.48 | 107.39 | 0.928 | 0.937 | 0.045 | 0 |
| SIE_EE_eff         | -0.0005 | 100.48 | 106.91 | 0.927 | 0.937 | 0.045 | 0 |
| SIE_tmle           | -0.001 | 107.84 | 116.48 | 0.926 | 0.937 | 0.050 | 0 |
| SIE_EE             | -0.001 | 108.51 | 114.77 | 0.930 | 0.938 | 0.048 | 0 |
| SIE_iptw           | -0.007 | 174.62 | 163.60 | 0.941 | 0.944 | 0.062 | 0 |
| DGM 3, N=5000      |       |            |            |     |      |             |
| Transport stochastic direct effect
| SDE_tmle_eff       | -0.001 | 100.04 | 126.93 | 0.941 | 0.962 | 0.015 | 0 |
| SDE_EE_eff         | -0.0005 | 100.02 | 113.84 | 0.943 | 0.955 | 0.014 | 0 |
| SDE_tmle           | 0.001  | 214.33 | 238.29 | 0.961 | 0.965 | 0.032 | 0 |

Continued on next page
Table 17 – continued from previous page

| Estimator  | Bias   | Efficiency | 95%CI | Cov | RMSE | % Out of Bds |
|------------|--------|------------|-------|-----|------|--------------|
|            | IC     | Bootstrapping IC | Bootstrapping |     |      |              |
| SDE_EE     | 0.003  | 215.45     | 222.18 | 0.964 | 0.964 | 0.031 0      |
| SDE_iptw   | 0.006  | 475.36     | 466.38 | 0.934 | 0.937 | 0.059 0      |

Transport stochastic indirect effect

| Estimator     | Bias   | Efficiency | 95%CI | Cov | RMSE | % Out of Bds |
|---------------|--------|------------|-------|-----|------|--------------|
| SIE_tmle_eff  | 0      | 100.07     | 118.92 | 0.939 | 0.938 | 0.015 0      |
| SIE_EE_eff    | -0.002 | 100.07     | 100.78 | 0.940 | 0.936 | 0.014 0      |
| SIE_tmle      | -0.004 | 108.04     | 125.46 | 0.938 | 0.946 | 0.016 0      |
| SIE_EE        | -0.005 | 108.09     | 108.55 | 0.944 | 0.944 | 0.014 0      |
| SIE_iptw      | -0.001 | 147.09     | 123.72 | 0.967 | 0.942 | 0.016 0      |

Table 18: Simulation results comparing estimators of the transported direct and indirect effects for DGP 3 under $Z$, $M$, and $S$ models misspecified. For sample sizes 100, 500 and 5000.

| Estimator  | Bias   | Efficiency | 95%CI | Cov | RMSE | % Out of Bds |
|------------|--------|------------|-------|-----|------|--------------|
|            | IC     | Bootstrapping IC | Bootstrapping |     |      |              |

DGM 3, $N=100$

Transport stochastic direct effect

| Estimator     | Bias   | Efficiency | 95%CI | Cov | RMSE | % Out of Bds |
|---------------|--------|------------|-------|-----|------|--------------|
| SDE_tmle_eff  | -0.013 | 71.12      | 119.65 | 0.744 | 0.896 | 0.097 0      |
| SDE_EE_eff    | -0.013 | 71.11      | 119.59 | 0.745 | 0.896 | 0.097 0      |
| SDE_tmle      | -0.011 | 95.20      | 159.52 | 0.781 | 0.942 | 0.124 0      |
| SDE_EE        | -0.012 | 95.46      | 131.92 | 0.844 | 0.920 | 0.106 0      |
| SDE_iptw      | -0.110 | 196.15     | 206.68 | 0.877 | 0.889 | 0.200 0      |

Transport stochastic indirect effect

| Estimator     | Bias   | Efficiency | 95%CI | Cov | RMSE | % Out of Bds |
|---------------|--------|------------|-------|-----|------|--------------|
| SIE_tmle_eff  | 0.005  | 29.99      | 132.19 | 0.390 | 0.899 | 0.050 0      |
| SIE_EE_eff    | 0.005  | 30.01      | 131.03 | 0.389 | 0.898 | 0.050 0      |
| SIE_tmle      | 0.003  | 38.27      | 150.41 | 0.438 | 0.915 | 0.054 0      |
| SIE_EE        | 0.005  | 38.70      | 132.89 | 0.468 | 0.901 | 0.051 0      |
| SIE_iptw      | -0.045 | 85.72      | 97.43  | 0.533 | 0.574 | 0.060 0      |

DGM 3, $N=500$

Transport stochastic direct effect

| Estimator     | Bias   | Efficiency | 95%CI | Cov | RMSE | % Out of Bds |
|---------------|--------|------------|-------|-----|------|--------------|
| SDE_tmle_eff  | -0.002 | 69.65      | 106.13 | 0.796 | 0.955 | 0.042 0      |
| SDE_EE_eff    | -0.003 | 69.61      | 106.09 | 0.797 | 0.954 | 0.042 0      |
| SDE_tmle      | -0.002 | 95.76      | 142.10 | 0.792 | 0.934 | 0.057 0      |
| SDE_EE        | -0.002 | 95.89      | 121.23 | 0.868 | 0.946 | 0.048 0      |
| SDE_iptw      | -0.096 | 198.50     | 194.22 | 0.761 | 0.755 | 0.120 0      |

Transport stochastic indirect effect

| Estimator     | Bias   | Efficiency | 95%CI | Cov | RMSE | % Out of Bds |
|---------------|--------|------------|-------|-----|------|--------------|
| SIE_tmle_eff  | -0.001 | 27.56      | 106.69 | 0.371 | 0.930 | 0.023 0      |
| SIE_EE_eff    | -0.0004| 27.57      | 106.65 | 0.368 | 0.930 | 0.023 0      |
| SIE_tmle      | -0.001 | 37.31      | 113.59 | 0.464 | 0.930 | 0.024 0      |
| SIE_EE        | -0.0004| 37.46      | 108.94 | 0.489 | 0.933 | 0.023 0      |
| SIE_iptw      | -0.055 | 71.87      | 65.11  | 0.036 | 0.042 | 0.057 0      |

DGM 3, $N=5000$

Transport stochastic direct effect

| Estimator     | Bias   | Efficiency | 95%CI | Cov | RMSE | % Out of Bds |
|---------------|--------|------------|-------|-----|------|--------------|
| SDE_tmle_eff  | -0.0004| 66.47      | 110.54 | 0.804 | 0.962 | 0.013 0      |
| SDE_EE_eff    | -0.0005| 66.45      | 110.18 | 0.805 | 0.961 | 0.013 0      |
| SDE_tmle      | -0.0004| 93.90      | 152.31 | 0.802 | 0.966 | 0.018 0      |
| SDE_EE        | -0.0005| 93.90      | 125.67 | 0.890 | 0.963 | 0.015 0      |

Continued on next page
### Table 18 – continued from previous page

| Estimator      | Bias  | Efficiency | 95% CI | Cov    | RMSE  | % Out of Bds |
|----------------|-------|------------|--------|--------|-------|--------------|
|                |       | IC         | Bootstrapping | IC     | Bootstrapping |
| SDE_iptw       | -0.101 | 187.34     | 181.99 | 0.002  | 0.001 | 0.103        | 0             |
| SIE_tmle_eff   | -0.0002 | 26.08      | 109.30 | 0.382  | 0.937 | 0.007        | 0             |
| SIE_EE_eff     | -0.0002 | 26.090     | 106.07 | 0.385  | 0.940 | 0.007        | 0             |
| SIE_tmle       | -0.0002 | 35.98      | 113.49 | 0.497  | 0.929 | 0.007        | 0             |
| SIE_EE         | 0.0001  | 36         | 107.81 | 0.491  | 0.949 | 0.007        | 0             |
| SIE_iptw       | -0.057  | 61.10      | 53.42  | 0      | 0     | 0.058        | 0             |

Transport stochastic indirect effect

### Table 19: Simulation results comparing estimators of the transported direct and indirect effects for DGP 3 under Y and S models misspecified. For sample sizes 100, 500 and 5000.

| Estimator      | Bias  | Efficiency | 95% CI | Cov    | RMSE  | % Out of Bds |
|----------------|-------|------------|--------|--------|-------|--------------|
|                |       | IC         | Bootstrapping | IC     | Bootstrapping |
| DGM 3, N=100   |       |            |         |        |       |              |
| SDE_tmle_eff   | -0.213 | 74.55      | 170.93 | 0.230  | 0.654 | 0.266        | 0             |
| SDE_EE_eff     | -0.254 | 88.47      | 125.74 | 0.115  | 0.274 | 0.279        | 0             |
| SDE_tmle       | -0.180 | 119.92     | 267.80 | 0.519  | 0.819 | 0.297        | 0             |
| SDE_EE         | -0.238 | 150.22     | 177.67 | 0.453  | 0.574 | 0.284        | 0             |
| SDE_iptw       | -0.210 | 308.35     | 290.33 | 0.788  | 0.862 | 0.328        | 0             |
| SIE_tmle_eff   | 0.124  | 43.16      | 103.06 | 0.274  | 0.492 | 0.185        | 0             |
| SIE_EE_eff     | 0.164  | 53.82      | 41.56  | 0.194  | 0.120 | 0.193        | 0             |
| SIE_tmle       | 0.120  | 44.87      | 109.68 | 0.273  | 0.478 | 0.186        | 0             |
| SIE_EE         | 0.162  | 58.43      | 54.32  | 0.226  | 0.178 | 0.192        | 0             |
| SIE_iptw       | 0.113  | 146.94     | 107.72 | 0.513  | 0.500 | 0.173        | 0             |
| DGM 3, N=500   |       |            |         |        |       |              |
| SDE_tmle_eff   | -0.142 | 93.50      | 153.26 | 0.236  | 0.334 | 0.160        | 0             |
| SDE_EE_eff     | -0.197 | 202.02     | 148.83 | 0.287  | 0.284 | 0.219        | 0             |
| SDE_tmle       | -0.136 | 204.33     | 296.39 | 0.477  | 0.717 | 0.195        | 0             |
| SDE_EE         | -0.191 | 371.54     | 301.70 | 0.301  | 0.280 | 0.260        | 0             |
| SDE_iptw       | -0.164 | 329.74     | 244.23 | 0.727  | 0.501 | 0.197        | 0             |
| SIE_tmle_eff   | 0.124  | 43.16      | 103.06 | 0.274  | 0.492 | 0.185        | 0             |
| SIE_EE_eff     | 0.164  | 53.82      | 41.56  | 0.194  | 0.120 | 0.193        | 0             |
| SIE_tmle       | 0.120  | 44.87      | 109.68 | 0.273  | 0.478 | 0.186        | 0             |
| SIE_EE         | 0.162  | 58.43      | 54.32  | 0.226  | 0.178 | 0.192        | 0             |
| SIE_iptw       | 0.113  | 146.94     | 107.72 | 0.513  | 0.500 | 0.173        | 0             |
| DGM 3, N=5000  |       |            |         |        |       |              |
| SDE_tmle_eff   | -0.056 | 203.22     | 244.19 | 0.477  | 0.548 | 0.065        | 0             |
| SDE_EE_eff     | -0.070 | 875.83     | 342.74 | 0.962  | 0.596 | 0.082        | 0             |
| SDE_tmle       | -0.062 | 518.76     | 616    | 0.813  | 0.889 | 0.111        | 0             |
| SDE_EE         | -0.074 | 1,340.79   | 935.49 | 0.862  | 0.865 | 0.162        | 0             |
| SDE_iptw       | -0.099 | 417.60     | 429.88 | 0.620  | 0.669 | 0.118        | 0             |

Transport stochastic indirect effect

Continued on next page
Table 19 – continued from previous page

| Estimator   | Bias  | Efficiency | 95%CI | Cov | RMSE | % Out of Bds |
|-------------|-------|------------|-------|-----|------|--------------|
|             | IC    | Bootstrapping | IC    | Bootstrapping |       |              |
| SIE_tmle_eff| 0.125 | 147.45     | 105.41| 0   | 0    | 0.126        | 0  |
| SIE_EE_eff  | 0.133 | 67.89      | 35    | 0   | 0    | 0.134        | 0  |
| SIE_tmle    | 0.122 | 158.62     | 145.54| 0   | 0    | 0.124        | 0  |
| SIE_EE      | 0.133 | 74.73      | 47.12 | 0   | 0    | 0.134        | 0  |
| SIE_iptw    | 0.129 | 157.15     | 126.67| 0   | 0    | 0.130        | 0  |

Table 20: Simulation results comparing estimators of the transported direct and indirect effects for DGP 3 under Y and Z models misspecified. For sample sizes 100, 500 and 5000.

| Estimator   | Bias  | Efficiency | 95%CI | Cov | RMSE | % Out of Bds |
|-------------|-------|------------|-------|-----|------|--------------|
|             | IC    | Bootstrapping | IC    | Bootstrapping |       |              |
| Transport stochastic direct effect |
| DGM 3, N=100 |
| SDE_tmle_eff | -0.155 | 79.58      | 135.64| 0.422 | 0.658 | 0.195        | 0  |
| SDE_EE_eff  | -0.150 | 117.23     | 136.56| 0.580 | 0.714 | 0.188        | 0  |
| SDE_tmle    | 0.005 | 264.09     | 290.02| 0.806 | 0.874 | 0.287        | 0  |
| SDE_EE      | 0.028 | 380.69     | 299.92| 0.939 | 0.880 | 0.292        | 0.300 |
| SDE_iptw    | 0.002 | 280.33     | 287.94| 0.867 | 0.883 | 0.272        | 0  |
| Transport stochastic indirect effect |
| DGM 3, N=500 |
| SIE_tmle_eff | 0.047 | 111.60     | 126.11| 0.640 | 0.639 | 0.086        | 0  |
| SIE_EE_eff  | 0.038 | 170.70     | 110.67| 0.863 | 0.755 | 0.070        | 0  |
| SIE_tmle    | 0.046 | 120.45     | 151.05| 0.626 | 0.640 | 0.094        | 0  |
| SIE_EE      | 0.032 | 195.13     | 146.58| 0.691 | 0.639 | 0.088        | 0  |
| SIE_iptw    | 0.030 | 151.98     | 140.84| 0.760 | 0.809 | 0.076        | 0  |
| Transport stochastic direct effect |
| DGM 3, N=5000 |
| SDE_tmle_eff | -0.103 | 111.77    | 124.79| 0.307 | 0.383 | 0.112        | 0  |
| SDE_EE_eff  | -0.106 | 158.80    | 114.52| 0.526 | 0.266 | 0.114        | 0  |
| SDE_tmle    | 0.116 | 397.58     | 376.82| 0.799 | 0.735 | 0.192        | 0  |
| SDE_EE      | 0.120 | 481.43     | 371.48| 0.942 | 0.746 | 0.191        | 0  |
| SDE_iptw    | 0.090 | 339.25     | 328.30| 0.812 | 0.776 | 0.161        | 0  |
| Transport stochastic indirect effect |
| DGM 3, N=5000 |
| SIE_tmle_eff | -0.002 | 141.83   | 125.67| 0.977 | 0.954 | 0.022        | 0  |
| SIE_EE_eff  | 0.0002 | 214.39   | 101.32| 0.993 | 0.953 | 0.018        | 0  |
| SIE_tmle    | -0.004 | 151.37   | 186.26| 0.892 | 0.923 | 0.035        | 0  |
| SIE_EE      | -0.004 | 269.77   | 193.99| 0.973 | 0.938 | 0.036        | 0  |
| SIE_iptw    | -0.002 | 148.16   | 139.08| 0.935 | 0.921 | 0.026        | 0  |
| Transport stochastic direct effect |
| DGM 3, N=5000 |
| SDE_tmle_eff | -0.098 | 105.11  | 103.48| 0   | 0   | 0.099        | 0  |
| SDE_EE_eff  | -0.104 | 144.61  | 97.28 | 0   | 0   | 0.104        | 0  |
| SDE_tmle    | 0.128 | 387.96   | 351.80| 0.210 | 0.178 | 0.135        | 0  |
| SDE_EE      | 0.127 | 440.62   | 351.63| 0.282 | 0.183 | 0.134        | 0  |
| SDE_iptw    | 0.095 | 334.43   | 309.47| 0.343 | 0.296 | 0.102        | 0  |
| Transport stochastic indirect effect |
| SIE_tmle_eff | -0.005 | 122.10 | 87.11 | 0.939 | 0.826 | 0.007        | 0  |
| SIE_EE_eff  | -0.001 | 180.44 | 76.76 | 1   | 0.932 | 0.005        | 0  |

Continued on next page
| Estimator    | Bias | Efficiency | 95% CI | Cov | RMSE | % Out of Bds |
|-------------|------|------------|--------|-----|------|--------------|
|             | IC   | Bootstrapping | IC     | Bootstrapping |     |              |
| SIE_tmle    | -0.014 | 143.41 | 146.80 | 0.602 | 0.626 | 0.017 | 0 |
| SIE_EE      | -0.005 | 230.11 | 164.35 | 0.992 | 0.931 | 0.011 | 0 |
| SIE_iptw    | -0.005 | 121.01 | 114.92 | 0.881 | 0.859 | 0.009 | 0 |

Table 20: Simulation results comparing estimators of the transported direct and indirect effects for DGP 3 under Y and M models misspecified. For sample sizes 100, 500 and 5000.

| Estimator    | Bias | Efficiency | 95% CI | Cov | RMSE | % Out of Bds |
|-------------|------|------------|--------|-----|------|--------------|
|             | IC   | Bootstrapping | IC     | Bootstrapping |     |              |
| SIE_tmle    | -0.014 | 143.41 | 146.80 | 0.602 | 0.626 | 0.017 | 0 |
| SIE_EE      | -0.005 | 230.11 | 164.35 | 0.992 | 0.931 | 0.011 | 0 |
| SIE_iptw    | -0.005 | 121.01 | 114.92 | 0.881 | 0.859 | 0.009 | 0 |

| Estimator    | Bias | Efficiency | 95% CI | Cov | RMSE | % Out of Bds |
|-------------|------|------------|--------|-----|------|--------------|
|             | IC   | Bootstrapping | IC     | Bootstrapping |     |              |
| SIE_tmle    | -0.001 | 108.17 | 131.11 | 0.573 | 0.583 | 0.196 | 0 |
| SIE_EE      | 0.0071 | 203.97 | 155.45 | 0.656 | 0.628 | 0.197 | 0.100 |
| SIE_tmle    | 0.109 | 109.89 | 121.39 | 0.533 | 0.512 | 0.204 | 0 |
| SIE_EE      | 0.082 | 207.74 | 164.23 | 0.622 | 0.575 | 0.211 | 0.100 |
| SIE_iptw    | 0.094 | 192.23 | 119.34 | 0.644 | 0.562 | 0.183 | 0 |

Table 21: Simulation results comparing estimators of the transported direct and indirect effects for DGP 3 under Y and M models misspecified. For sample sizes 100, 500 and 5000.
Table 21 – continued from previous page

| Estimator  | Bias   | Efficiency IC | 95% CI Cov IC | RMSE Bootstrapping | % Out of Bds | Cov Bootstrapping |
|------------|--------|---------------|--------------|--------------------|-------------|------------------|
| SIE_iptw   | -0.015 | 129.58        | 120.45       | 0.862              | 0.825       | 0.023            |

Table 22: Simulation results comparing estimators of the transported direct and indirect effects for DGP 3 under $Y$ model misspecified. For sample sizes 100, 500 and 5000.

| Estimator         | Bias   | Efficiency IC | 95% CI Cov IC | RMSE Bootstrapping | % Out of Bds | Cov Bootstrapping |
|-------------------|--------|---------------|--------------|--------------------|-------------|------------------|
| **DGM 3, N=100**  |        |               |              |                    |             |                  |
| Transport stochastic direct effect |        |               |              |                    |             |                  |
| SDE_tmle_eff      | -0.067 | 114.03        | 177.02       | 0.731              | 0.893       | 0.158            |
| SDE_EE_eff        | -0.095 | 153.08        | 160.44       | 0.766              | 0.817       | 0.161            |
| SDE_tmle          | 0.019  | 189.07        | 312.17       | 0.598              | 0.876       | 0.314            |
| SDE_EE            | -0.063 | 335.43        | 258.40       | 0.857              | 0.854       | 0.335            |
| SDE_iptw          | -0.015 | 349.83        | 326.79       | 0.868              | 0.883       | 0.306            |
| Transport stochastic indirect effect |        |               |              |                    |             |                  |
| SIE_tmle_eff      | 0.093  | 107.56        | 142.73       | 0.533              | 0.582       | 0.200            |
| SIE_EE_eff        | 0.094  | 172.64        | 120.52       | 0.612              | 0.562       | 0.185            |
| SIE_tmle          | 0.101  | 109.75        | 138.89       | 0.516              | 0.532       | 0.203            |
| SIE_EE            | 0.096  | 180.93        | 132.24       | 0.604              | 0.545       | 0.193            |
| SIE_iptw          | 0.083  | 159.86        | 136.27       | 0.648              | 0.589       | 0.183            |
| **DGM 3, N=500**  |        |               |              |                    |             |                  |
| Transport stochastic direct effect |        |               |              |                    |             |                  |
| SDE_tmle_eff      | -0.007 | 147.33        | 140.13       | 0.954              | 0.949       | 0.054            |
| SDE_EE_eff        | -0.025 | 215.05        | 133.32       | 0.971              | 0.923       | 0.056            |
| SDE_tmle          | 0.066  | 411.77        | 462.31       | 0.607              | 0.674       | 0.238            |
| SDE_EE            | -0.025 | 729.40        | 644.29       | 0.992              | 0.752       | 0.314            |
| SDE_iptw          | 0.042  | 463.89        | 418.71       | 0.768              | 0.756       | 0.200            |
| Transport stochastic indirect effect |        |               |              |                    |             |                  |
| SIE_tmle_eff      | -0.009 | 113.07        | 157.65       | 0.898              | 0.955       | 0.060            |
| SIE_EE_eff        | 0.004  | 270.99        | 147.05       | 0.985              | 0.959       | 0.058            |
| SIE_tmle          | -0.005 | 126.18        | 184.44       | 0.896              | 0.958       | 0.068            |
| SIE_EE            | 0.003  | 293.04        | 190.52       | 0.970              | 0.950       | 0.074            |
| SIE_iptw          | -0.005 | 176.15        | 164.44       | 0.937              | 0.944       | 0.066            |
| **DGM 3, N=5000** |        |               |              |                    |             |                  |
| Transport stochastic direct effect |        |               |              |                    |             |                  |
| SDE_tmle_eff      | -0.0002| 147.13        | 118.60       | 0.984              | 0.947       | 0.015            |
| SDE_EE_eff        | -0.001 | 250.94        | 124.08       | 0.992              | 0.950       | 0.014            |
| SDE_tmle          | 0.006  | 574.32        | 554.59       | 0.925              | 0.910       | 0.076            |
| SDE_EE            | -0.002 | 716.64        | 643.04       | 0.947              | 0.910       | 0.087            |
| SDE_iptw          | 0.004  | 470.22        | 463.33       | 0.903              | 0.915       | 0.063            |
| Transport stochastic indirect effect |        |               |              |                    |             |                  |
| SIE_tmle_eff      | 0.0001 | 130.82        | 108.84       | 0.984              | 0.952       | 0.014            |
| SIE_EE_eff        | 0.001  | 220.90        | 108.07       | 1                 | 0.939       | 0.015            |
| SIE_tmle          | -0.0001| 146.59        | 126.90       | 0.955              | 0.939       | 0.017            |
| SIE_EE            | 0.001  | 240.50        | 147.62       | 0.999              | 0.938       | 0.020            |
| SIE_iptw          | -0.0001| 149.24        | 124.84       | 0.964              | 0.940       | 0.017            |