Proceedings of Machine Learning Research 275:1–15, 2025 4th Conference on Causal Learning and Reasoning

Disparate Effect Of Missing Mediators On Transportability of Causal Effects

Vishwali Mhasawade VISHWALIM@NYU.EDU
New York University, Grossman School of Medicine

Rumi Chunara RUMI.CHUNARA@NYU.EDU
New York University, School of Global Public Health and Tandon School of Engineering

Editors: Biwei Huang and Mathias Drton

Abstract

The transport of mediation effects is an important question for upstream interventions, such as targeted development of parks to improve population health, as their effect on populations is mediated by factors like physical activity, which can vary from place to place. However, upstream treatment effect estimates could be biased when mediator variables are missing for the population where the effect is to be transported. We study this issue of the impact of missing mediators on transported effects, motivated by challenges in public health, wherein mediators are commonly missing but not at random. We propose a sensitivity analysis framework to quantify the impact of missing mediator data on transported mediation effects, identifying when the conditional transported mediation effect becomes insignificant for subgroups with missing data. Applied to longitudinal data from the Moving to Opportunity Study, a large-scale housing voucher experiment, this framework demonstrates the sensitivity of transported mediation effects to data missingness. In particular, we quantify the effect of missing mediators on transport effect estimates of voucher receipt in childhood, an upstream intervention on living location. We then assess the subsequent impact on the risk of mental health or substance use disorder mediated through parental health across sites. Our findings highlight that missing mediators can disparately impact effect estimates across population subgroups and provide a tangible understanding of how much missing data can be withstood for unbiased effect estimates in such mediated settings.

1. Introduction

There is a growing understanding of the importance of multi-level causes and determinants of health (Diez-Roux, 2000). For example, factors at the neighborhood level, such as the number of available healthy food resources, affect health outcomes like hypertension indirectly through individual-level behaviors like food consumption/diet (Motamedi et al., 2021; Stevenson et al., 2023). This perspective opens the possibility of intervening to improve health outcomes like hypertension at different levels. For example, at a neighborhood-level, increasing green space, or using an individual-level behavioral intervention to augment physical activity (Dalton and Jones, 2020; Stevenson et al., 2023). Notably, when estimating the causal effect of interventions at distal (e.g., neighborhood) levels, the mechanism through which such an intervention acts is often through individual-level factors like physical activity or diet behaviors (Mohnen et al., 2012). Therefore, when considering distal causes, the inclusion of mediators, factors that explain the mechanism through which a distal cause influences an outcome, becomes essential for assessing, comparing, and transporting treatment effects to other locations/sites. Accordingly, here we focus on questions in the context of the transportability of causal effects for interventions/exposures at a distal level for which there is a known mediator. Based on public health research, multiple contexts illustrate how distal factors influence health through mediators (Zhang et al., 2024; Frehlich et al., 2024).

We present a method for quantifying the impact of missing mediators on the estimated transported causal effect and study the relationship of missing mediators with causal effect. Specifically, we show how differences in transported causal effect can arise between minority and majority subgroup populations based on differing proportions of missing mediator data across the two subgroups.

We first highlight the challenges that missing mediators introduce in the setting of transporting causal effects. Transporting effects across populations where the source and target differ is a common task, especially when practitioners have access to observational data from multiple environments with potentially different populations. This is desirable when one wants to identify if the conclusions from a specific experiment apply to a different population, which is known as identifying the transportability of treatment effects or generalizing experimental results. Approaches aimed at identifying the transportability of treatment effects across two environments are based on two main assumptions: 1) knowledge about the differences across environments, which are usually represented using selection nodes in a selection diagram, an augmented causal graph, and 2) complete data is observed across both the environments, i.e., there are no missing attributes in the target environment (Cinelli and Pearl, 2021; Bareinboim and Pearl, 2012). However, it is difficult to obtain complete data from all the environments when assessing the transportability of causal effects due to different data collection policies, privacy concerns, etc. Accordingly, here, we focus on the second assumption regarding observing complete data and study under what conditions transportability is identifiable when the assumption is violated. Specifically, we focus on the setting where mediators are observed in the target environment but partially missing by focusing on transported mediation effects.

Missing mediator data poses an added challenge in transportability analyses, particularly when such data is absent in minority subgroups but available for majority groups, as is often the case in public health research (Simkhada et al., 2008). Indeed, it has been shown that respondents with missing information may not represent a random subset of population-based survey participants. In particular, disparities in missingness often correlate with relevant sociodemographic characteristics, such as neighborhood socioeconomic status (SES) (Kim et al., 2007; Krieger et al., 2008). For instance, missing data about individual-level behaviors, such as alcohol consumption, has been linked to neighborhood SES (Sania et al., 2021). This situation is problematic because SES, an upstream potential treatment or exposure, is a significant determinant of drinking behavior during pregnancy (Skagerstrom et al., 2011). Furthermore, pregnant women from low socioeconomic backgrounds are more likely to access antenatal care late in pregnancy, enroll late in research studies, and consequently, experience greater missing data in the early stages of pregnancy (Simkhada et al., 2008). In summary, it is essential to recognize that mediators for distal factors may be missing unevenly across population groups. Identifying exact patterns of missing data can be difficult, yet the scenario described here aligns with fundamental public health insights. Future work could extend this approach to account for other patterns of missingness as they are observed in real-world settings.

To study disparate missingness of mediators on transport effect estimation systematically, we first present a causal view of the setting using a directed acyclic graph (DAG) and highlight settings under which the causal effect can be transported under different missingness scenarios where the causal DAGs are based on real-world scenarios. We then adapt estimators for transporting causal effects under missing data and present a framework to evaluate the sensitivity of the estimated causal effect to missing mediator data. Our findings highlight that transported mediation effects are sensitive to missing mediator data. Specifically when the mediator data is missing the transported effect is different across subgroups such that it becomes non-significant for the subgroup with more missing data. We also demonstrate the effect of missing mediators on transported indirect treatment effect estimation from the Moving to Opportunity Study, which has been used to study the causal effect of moving to a new neighborhood (Sanbonmatsu et al., 2011).

2. Previous work

Transportability of causal effects. Identifying the transportability of causal effects across environments under selection bias of covariates has primarily focused on settings for which complete data in both the source and target environment are assumed (Pearl and Bareinboim, 2011; Bareinboim and Pearl, 2012, 2013, 2014). However, in applications such as epidemiology and public health, real-world challenges due to the high resources required for data collection can result in inability to consistently observe all covariates in each environment. In particular, social determinants of health, upstream treatments/exposures, represent one such class of attributes that can be difficult to obtain (Braveman and Gottlieb, 2014; Cantor and Thorpe, 2018).

A main focus of research on the identifiability of transported causal effects has been on identifying the various conditions under which an optimal adjustment set can be obtained, although restricted to the assumption that complete data is observed in the target environment ("Henckel et al., "2019"; Bareinboim and Pearl, 2012; Shpitser et al., 2012). In another line of work, selecting specific covariates, known as a separating set, to ensure that selection bias of the covariates does not affect the population average treatment effect (PATE) has also been discussed (Egami and Hartman, 2021). Huang (2024) propose a sensitivity analysis framework for assessing sensitivity to omitted covariates in the separating set in case of transported average treatment effects. However, to estimate causal effects in the target environment by correcting for selection bias, such a separating set would need to be modified when specific covariates are unobserved or have missing data. Thus, although sensitivity frameworks have been proposed for transported average treatment effects in the case of omitted covariates, the specific issue with transported mediation effects due to missing mediators in the target environment is largely unexplored.

Causal inference with missing covariates and selection bias. Transporting causal effects is challenging mainly due to potential covariate shifts across environments and the difficulty of collecting complete data across all environments. Quantifying and correcting the bias due to missingess requires prior information on what covariates, confounders, including colliders, moderators, treatments, or mediators, are observed in the target environment. For example, the setting in which pre-treatment covariates are only measured in the source environment but not in the target environment has been studied in previous work (Andrews and Oster, 2019). In this case, bias due to missingness and selection bias are both corrected using the target environment covariate means. Briefly, causal estimates in the source environment are adjusted based on the mean of the covariates from the target environment as selection bias introduces differences between the source and target mean values. The assumption in this related work is that even though the covariates are missing in the target population, their descriptive statistics are known a priori, or that the data is missing at random or missing completely at random. Depending on the missingness patterns in the target environment specifically, missing at random and missing completely at random patterns, different imputation strategies have also been studied (Mayer et al., 2021). However, assessing the impact of missingness on the bias between SATE and PATE for the missing not-at-random setting has not been addressed before. That is, how much missingness can be tolerated to ensure transportability when covariates are not missing randomly is an open question and specifically for missingness of mediators in case of transported mediation effects. Overall, quantifying the impact of missing data on bias between SATE and PATE in missing-not-at-random (MNAR) settings is a significant challenge. Further, given extensive public health research illustrates how distal factors influence health through mediators, how much missingness can be tolerated to ensure transportability of effects when mediators are missing is a critical gap. Addressing this question is essential to understand transported mediation effects accurately in the presence of MNAR covariates.

Thus, while previous work has focused on the identification of covariate adjustment sets under selection bias with inconsistent or missing covariates, the identification of transported causal effects with missing covariates, let alone missing meditators specifically, has not been explored to the best of our knowledge. Furthermore, quantifying the bias between SATE and PATE under selection bias is well established, however the effect of missing covariates and especially, missing mediators on the bias quantification is an open problem. In addition to addressing these research gaps, we present a framework to analyze the sensitivity to missing mediator data for estimators for the indirect effect mediated by the mediators.

3. Background

Generalizing or transporting causal effects from one observational setting consisting of a study sample in the source environment to the target environment of interest is a challenge when there are differences between the source and target environments, including those due to population distribution differences. Such population distribution differences across environments can arise when subgroups are underrepresented, or the treatment assignment in the target population is different from the source, in which case the estimated causal effect from the source environment cannot be directly transported to the target environment. Here, we focus on two main issues that may result in these differences: 1) selection bias; the observational data from the source environment is a non-randomized sample of the target population; and 2) missing mediators; there is missingess in the mediators in the target environment. We consider the setup where our variables are generated as follows:

$$\begin{aligned} S &= f_S(U_S) \\ W &= f_W(U_W, S) \\ A &= f_A(U_A, S) \\ R &= f_R(U_R, A, S) \\ C &= f_C(U_C, R, A, W, S) \\ Y &= f_Y(U_Y, A, C, W, S) \end{aligned}$$

where A represents the treatment, assumed to be binary in this case, which operates through the mediators, Y represents the outcome, R and C are mediators, where they lie on the path between the treatment, A and the outcome, Y, $A \to R \to C \to Y$. W represents subgroup membership and confounds C and Y. Here, W is assumed to be binary. We note that W may also be continuous, but for simplifications, we restrict to W being binary for the analysis. In order to represent the differences between environments that may arise due to differences in populations, data measurement procedures, or other reasons, we use selection diagrams (Pearl), which include additional nodes known as selection nodes, S. These selection nodes are commonly used to represent differences across environments which are known a priori (Bareinboim and Pearl, 2012). Thus, $S \to C$ represents that the only difference between the two environments is due to the distribution of the covariate C. [1] We first begin by discussing the transported average treatment effect before introducing the transported mediation effects, which are the main point of interest in this work. Considering $S = 1$ to be an indicator of the source environment and $S = 0$ to be an indicator of the target environment, the quantity of interest when transporting effects is the Population Average Treatment Effect (PATE) in the target environment,

$$\text{PATE} = E[Y^1 - Y^0 \mid S = 0]. \tag{1}$$

1. Note, an absence of an edge between S and C denotes that there are no differences with respect to C across environments, a property which is essential to identify if the causal effects can be transported from $S = 0$ to $S = 1$.

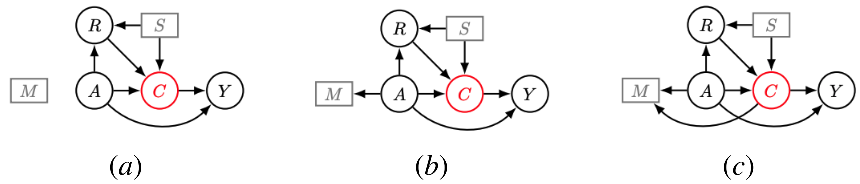

Figure 1: Different missingness patterns for the mediator C illustrated using causal graph for missing completely at random (MCAR) in (a) where $C \perp\!\!\!\perp M$, missing at random (MAR) in (b) where $C \perp\!\!\!\perp M \mid A, R$, and missing not at random (MNAR) in (c) where $C \not\perp\!\!\!\perp M \mid A, R$.

However, with the data in the source environment, we are only able to estimate the Sample Average Treatment Effect (SATE),

$$\text{SATE} = E[Y^1 - Y^0 \mid S = 1]. \tag{2}$$

where Y^a is the potential outcome had the treatment been set to $A = a$.

Mediated indirect effect estimation with missing mediator data in the target environment. The goal of this study is to understand if the indirect effect mediated by C, i.e., the effect along $A \to R \to C$, is affected by missing data in C. Our task is challenging because estimating the indirect effect of A on Y as mediated by C is difficult due to C having missing data in the target environment as well as distribution shift in W which affects C across the two environments, $S \to W \to C$. Specifically, we focus on the issue of where mediator data is missing in the target environment where our main goal is to assess the sensitivity of the transported mediation effect to the missingness in the mediator C. As illustrated in Figure 1, M denotes the set of missingness indicators of the variables in C.[2]

We illustrate the different missingness patterns with C as the specific mediator of interest with missing data, A as the treatment, Y as the outcome and R as another mediator along $A \to R \to C \to Y$ through causal graphs in Figure 1, respectively, before introducing the estimator for the transported indirect effect and the sensitivity framework.

These three settings, MCAR (missing completely at random), MAR (missing at random), and MNAR (missing not at random), represent missingness patterns commonly found in real-world datasets. For assessing the sensitivity of missing mediator data on transported effect estimates, we focus on the MNAR setup as illustrated next for two reasons. First, data about about individual behaviors, a common mediator of distal causes such as environment attributes, has been shown to commonly be missing not at random in public health research (Hallgren and Witkiewitz, 2013; Hallgren et al., 2016). Further, correcting for the bias in transported effects due to MNAR is challenging as opposed to MCAR and MAR as the missingness pattern cannot be estimated in MNAR but may be possible for the other two. In sum, the goal of this work is to show the sensitivity of the transported indirect effect to the missingness in the mediator data. We do not address how to handle the missing data, which is a suitable problem for future work.

4. Sensitivity of the causal estimator to missingness of mediators.

Next, we present a specific case of estimating the causal effect with missing mediators when the missingness is MNAR. In this case, MNAR leads to biased estimates of causal effects because the reason for missingness is related to the unobserved (missing) value itself, complicating any attempts to correct for the missing data in contrast with MAR, where missingness is related to the observed data but not the missing value, and MCAR, where missingness is unrelated to both the

2. The missing variables are represented in red in the causal graphs.

observed and missing data. Figure 2 represents one such setup. This is the causal graph we will be using for simulated analysis in this study. This approach is well-supported by robust public health frameworks that have demonstrated the causal relationship between neighborhood factors such as neighborhood socioeconomic status, which affect individual behaviors such as alcohol consumption and also affect individual health outcomes (Skagerstrom et al., 2011; Sania et al., 2021; Simkhada et al., 2008).

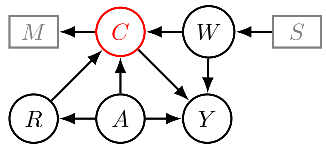

Figure 2: Figure representing the interaction between neighborhood factors and Cardiovascular disease risk (CVD); neighborhood SES (A) affects individual behaviors related to CVD, such as alcohol consumption, which are mediators (C), while C also affects CVD risk (Y). A also affects the alcohol resources in the neighborhood, R which in turn affects C. Alcohol consumption and CVD risk can be confounded by 'race', W, which also defines minority and majority groups based on the value of 'race'. Behavioral data for individuals related to alcohol consumption, the mediators can be missing for specific subgroups (Sania et al., 2021), which is denoted by the missingness indicator (M). We also account for a distribution shift across source and target environments, which is represented by the selection node (S).

4.1. Estimator in the target environment with missing data and selection bias

We assume that in the target domain, the mediators are missing, and there is a distribution shift across the source and target environments. Thus, we have the following edges in the DAG, $S \to W \to C$.

We represent the potential outcome in the source environment as follows:

$$\Psi\left(P_{S=1}\right) = \mathbb{E}\left[Y_{a, g_{C|a^*,w,S=1}}\right], \tag{3}$$

where the expectation is taken over the full data model and $Y_{a,g_{C|a^*,w,s}}$ is a potential outcome intervening on A to set it to a specific treatment a, and then intervening on C to set it to a random draw from the distribution of C, i.e., a stochastic intervention and $g_{C|a^*,w,s}$ captures this intervention effect on C, where the distribution is defined as follows:

$$g_{C|a^*,w,S=1} = \sum_r P\left(C = c \mid A = a^*, R = r, S = 1, W = w\right) \times P\left(R = r \mid A = a^*\right). \tag{4}$$

When the effect is transported to the target environment ($S = 0$), we make the following modification to the potential outcome:

$$\Psi(P_{S=0}) = \mathbb{E}\left[Y_{a, g_{C|a^*,w,S=0}}\right], \tag{5}$$

where we have to adjust for the mediators in the target environment,

$$g_{C|a^*,w,S=0} = \sum_r P\left(C = c \mid A = a^*, R = r, S = 0, W\right) \times P\left(R = r \mid A = a^*, S = 0\right).$$

One way to decompose the total effect of exposure $A = a$ don't he outcome Y is through natural direct effect (NDE) and natural indirect effect (NIE). NDE is the effect of A on Y that is not mediated by the mediator C while NIE is the effect of A on Y through C, $A \to C \to Y$. However, NDE and NIE require the assumption that there is no measured or unmeasured post-treatment confounder of the mediator and the outcome. This assumption is challenging especially in public health datasets with multi-level data. Accordingly, we focus on the stochastic direct and indirect effects, which are similar to natural direct and indirect effects but do not require assumptions of no measured or unmeasured post-treatment confounder of the mediator-outcome relationship (VanderWeele and Tchetgen Tchetgen, 2017). The transported stochastic direct effect (SDE) can be evaluated by setting a^* to 0 and taking the mean difference in the outcome between setting a to 1 and setting a to 0,

$$\text{SDE} = \mathbb{E}\left[Y_{1,g_{C|0,w,s}} - Y_{0,g_{C|0,w,s}} \mid S=0\right]. \tag{6}$$

and the transported stochastic indirect effect (SIE) can be evaluated by setting $a = 1$ and then taking the difference in the mean outcome between setting $a^* = 1$ and $a^* = 0$, denoted as:

$$\text{SIE} = \mathbb{E}\left[Y_{1,g_{C|1,w,s}} - Y_{1,g_{C|0,w,s}} \mid S=0\right]. \tag{7}$$

Since only SIE depends on the effect of the mediator, represented by $g_{C|0,S=s,W}$, we focus on the SIE for sensitivity to the missingness of the mediator.

Moreover, since we are interested in assessing how SIE would vary across the missingness of the mediator for majority and minority subgroups of the data, we focus on $\text{SIE}_{W=0}$ to represent the SIE in the target environment for the minority group and compare this against $\text{SIE}_{W=1}$, the SIE in the target environment for the majority group. We define the conditional stochastic effect for a specific value of $W = w*$ as follows:

$$\text{SIE}_{w^*} = \mathbb{E}\left[Y_{1,g_{C|1,s,w=w*}} - Y_{1,g_{C|0,s,w=w*}} \mid S=0, W=w*\right]. \tag{8}$$

Since our goal is assessing the sensitivity of SIE_w, our estimand of interest, to missingness in C, we first introduce the estimator for SIE_w in next section and then follow with the sensitivity framework.

4.2. Targeted minimum loss-based estimator (TMLE) for the target environment

For our estimation of the transported causal effect under missing mediator data, we adapt the TMLE estimator (Rudolph et al., 2021). We start by estimating, $\Psi\left(P_{S=s}\right)$ helpful for obtaining SIE. To do this, we write the efficient influence curve equation of the parameter (EIC). An EIC is defined in terms of the statistical model, M defined by the causal graph in Figure 2 and the target parameter, $\Psi(P_{S=s})$, as the canonical gradient of the pathwise derivative of the $\Psi(P_{S=s})$ along each possible submodel of M. Levy et al. (2021) provides a comprehensive background of the efficient influence function of the estimand, $D(P)$. If a regular asymptotically linear (RAL) estimator has an influence curve equal to EIC, then the estimator is asymptotically efficient, meaning it is of minimum variance for an RAL estimator. Accordingly, the empirical mean of the influence curve provides a linear approximation of the estimator. Thus, the empirical mean of the influence curve is the parameter of interest here.

The EIC of this parameter in the nonparametric model is given by

$$\begin{aligned}
&D(P) = D_Y(P) + D_R(P), \text{ where}\\
&D_Y(P) = \left(Y - \bar{Q}_Y\left(A,C,R\right)\right) \times \frac{g_{C,n|a^*,s}(C) P_R\left(R \mid A=a, S=0\right) I\left(S=1\right)}{g_C(C \mid R, S=1) P_R\left(R \mid S=1\right) P_S(S=0)}\\
&D_R(P) = \left(\bar{Q}_C\left(a,R,S\right) - \Psi(P)\right) \frac{I(S=0)}{P_S(S=0)}
\end{aligned} \tag{9}$$

Let $\bar{Q}_{Y,n}(R,C,A)$ be an initial estimate of $\mathbb{E}[Y \mid C,R,A]$ where n denotes the number of samples. $\bar{Q}_{Y,n}(R,C,A)$ can be estimated by predicted values from a regression of Y on C,R,A among those with $S=1$ in the source environment. We update the initial estimate of $\bar{Q}_{Y,n}(R,C,A)$ using the following weights,

$$H(C,R,A,S) = \frac{g_{C,n|a^*,s}(C) P_R(R \mid A=a, S=0)\, I(S=1, A=a)}{P_C(C \mid R, S=1)\, P_R(C \mid A=a, S=1)\, P_A(A=a \mid S=1) P_S(S=0)} \tag{10}$$

$\bar{Q}_{Y,n}(R,C,A)$ is then updated by performing a weighted logistic regression of Y with $\bar{Q}_{Y,n}(R,C,A)$ as an offset and intercept ϵ_Y and weights $H_n(C,R,A,S)$. The update is given by $\bar{Q}_{Y,n}(R,C,A)$ = $\bar{Q}_{Y,n}(\epsilon_{Y,n})(R,C,A)$.

Following standard procedure for doubly robust estimators (Van der Laan et al., 2011), we can then perform the stochastic intervention on $\bar{Q}_{Y,n}(R,C,A)$ via the computation

$$\bar{Q}_{C,n}(R,A,C) = \mathbb{E}_{g_{C,n|a^*,s}}\left[\bar{Q}_{Y,n}(R,C,A) \mid A,S\right].$$

This can be done by generating predicted values of $\bar{Q}_{Y,n}(R,1,A)$ and $\bar{Q}_{Y,n}(R,0,A)$ and then marginalizing over $g_{C,n|a^*,s}(C) : \sum_{c=0}^{1} \bar{Q}_{Y,n}(R,c,A) g_{C,n|a^*,s}(c)$.

The empirical mean of $\bar{Q}_{C,n}(R,A,C)$ among those for whom $S=0$ is the TMLE estimate of $\psi(P)$. It solves $\frac{1}{n}\sum_{i=0}^{n} D_n^*(P) = 0$. The transported stochastic direct effect (SDE) can be evaluated by setting a^* to 0 and taking the difference in $\psi(P)$ setting a to 1 versus setting a to 0. The transported stochastic indirect effect (SIE) entails setting $a=1$ and then taking the difference in $\psi(P)$ setting $a^*=1$ versus $a^*=0$. The corresponding EIC is the difference in EIC for the parameter defined by setting $a^*=1, a=1$ and the EIC for the parameter defined by setting $a^*=0, a=1$.

Since the empirical mean of $\bar{Q}_{C,n}(R,A,C)$ among those for whom $S=0$ is the TMLE estimate, and the expectation is taken over $g_{C,n|a^*,s}(c)$, having missing values in C will influence the effect.

We now have the estimator for SIE, which depends on C. Accordingly, missingness in C can bias our estimate of SIE. We next describe the specific weights that are affected by missingness of mediators in the target distribution as $\mathbf{w} = g_{C,n|a^*,s}(C)$. Importantly, if there is missingness in C, the estimate of $\mathbf{w}$ will be biased. We denote the biased estimate by $\mathbf{w}^* = g_{C^*,n|a^*,s}(C^*)$, where C^* denotes C without the missing data in C. That is, we only select the samples for which C is not missing or ignore the individuals for which C is missing.

4.3. Sensitivity framework

In order to assess the sensitivity of the transported effects to missing mediators, we focus on the approach proposed by Tan (2006), specific to the weights for individuals which accounts for the bias introduced by individual-level error due to missingness where $i = 1, \ldots n$, as follows:

$$\lambda^{-1} \leq \frac{\mathbf{w}_i *}{\mathbf{w}_i} \leq \lambda, \tag{11}$$

where $\lambda \geq 1$. Here, $\mathbf{w}_i *$ represent the weights in the EIC when ith instance has missing data and is excluded in estimating the weights and $\mathbf{w}_i$ represents the setting where ith instance does not have missingness and is accounted for in the estimation of SIE.

For the indirect effect in the target environment, we consider the effect of missing data in C on $g_{C*|a^*,S=0}(C* \mid A, S=0)$ with k missing samples for C as follows:

Definition 1 *Let R^2 be the residual variation in the true weights $\mathbf{w}$, not explained by the estimated weights $\mathbf{w}^*$ with missingness in samples denoted by i:*

$$R^2 := 1 - \frac{var(\mathbf{w}^*_{-(i)} \mid i \in k)}{var(\mathbf{w}_i)} \tag{12}$$

When there is missing data in C, the weight w_i as demonstrated in Equation 4.2 is biased. As the bias due to missingess in C increases, so does the bias in the estimate of SIE*.

Accordingly, the estimate of SIE* will be a biased function of R^2. We represent SIE* as follows:

$$\begin{aligned}\text{SIE}^* &= \mathbb{E}_{g^*_{C,n|a^*=1,s=0}}\left[\bar{Q}^*_{Y,n}(R,C,A=1) \mid A=1,S=0\right]\\ &\quad - \mathbb{E}_{g^*_{C,n|a^*=0,s=0}}\left[\bar{Q}^*_{Y,n}(R,C,A=1) \mid A=1,S=0\right]\\ &\quad := \sigma(R^2)\text{SIE},\end{aligned} \tag{13}$$

where $\sigma(R^2)$ is a function of the residual variance that is not explained by the missing weights, $\mathbf{w}_{-i}$. We derive the confidence intervals for the SIE* following the approach developed by Huang and Pimentel (2022) as follows:

$$\text{CI}(\alpha) = \left[D_{\alpha/2}\left(\inf_{\bar{w}\in R^2}\text{SIE}^*\left(\sigma\left(R^2\right)\right)\right), D_{1-\alpha/2}\left(\sup_{\bar{w}\in R^2}\text{SIE}^*\left(\sigma\left(R^2\right)\right)\right)\right] \tag{14}$$

5. Simulation Study

5.1. Data and Variables

Based on the causal graph presented in Figure 2, we sample data from a distribution as follows:

$$\begin{aligned}&A \sim \text{Bernoulli}\,(0.5)\\ &W \sim \text{Bernoulli}\,(0.5 * S + \mathcal{N}\,(0, 0.1))\\ &R := 0.7 * A + \mathcal{N}\,(0, 0.5)\\ &C := 1.5 * R + 0.2 * W + 0.8 * (1 - W) + \mathcal{N}\,(0, 0.5)\\ &Y = \text{Bernoulli}\,(\sigma\,(0.2 * A + 2.5 * C + -0.7 * W))\\ &M := \text{Bernoulli}\,(\sigma\,(0 * C * W + \lambda * C * (1 - W)))\\ &\sigma(x) = 1/(1 + \exp(-x))\end{aligned}$$

Here, $S = 1$ denotes the target environment, $W = 1$ denotes the majority group with complete mediator data without missingness, $W = 0$ denotes the minority group with missing mediator data, and λ represents the proportion of missingness. The specific parameters for the data-generating process were chosen to allow for the indirect effect of A on Y to be conditionally the same given W for the majority ($W = 1$) and minority groups ($W = 0$). Following the data-generating process in accordance with the causal graph, we varied the proportion of missingness of the mediator C from 0.1 to 0.9 for the minority group, thus affecting the bias in R^2 for the minority group while not affecting the majority group.

5.2. Results

Figure 3 shows that the stochastic indirect effect (SIE) is homogenous across majority and minority groups without any missingness in C; however, increasing the bias in R^2 makes the SIE heterogeneous across the groups. Moreover, the indirect effect, mediated by C, becomes insignificant for bias in $R^2 > 0.29$ for the minority group, which has missing data for C. On the contrary, the indirect effect remains significant for the majority group, which does not have missing data for C. For the minority group, as missingness is increased in C, the indirect effect becomes insignificant even though it is significant with no missing. For a bias in R^2 of 0.9, SIE is approximately 0.31 (-5.2, 6.51), where the numbers in the parentheses represent the 95% confidence interval. With bias in R^2 of 0.1, SIE is equal to 2.72 (2.62, 2.81). In the case of the majority group, SIE remains significant. For a bias in $R2$ of 0.9, it is approximately equal to 2.79 (0.41, 4.41) and thus remains significant with an increasing bias in $R2$.

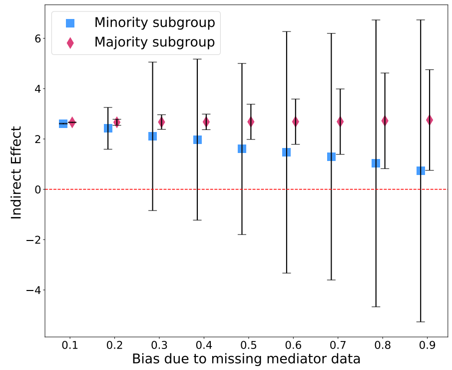

Figure 3: Results from the sensitivity analysis under the sensitivity framework. We vary the bias in R^2 measure due to missing mediator data across the x-axis for the minority and majority groups independently and plot the range of estimated stochastic indirect effect (SIE) values on the y-axis. The solid bar denotes the point estimate bounds for a specified bias in R^2 value, estimated as the point estimate plus 95% confidence intervals. We estimate a bias in R^2 = 0.29, such that if $R^2 \geq 0.29$, the intervals contain the null estimate for the SIE.

6. Application of Sensitivity Framework to Moving to Opportunity Study

6.1. Background

Moving to Opportunity (MTO) was a longitudinal randomized trial conducted by the U.S. Department of Housing and Urban Development from 1994 to 2007 in five cities in the United States: Baltimore, Boston, Chicago, Los Angeles, and New York (Sanbonmatsu et al., 2011). The MTO study provides an opportunity to uncover how improving the neighborhood improves individual-level health outcomes. Moreover, since this is a multicity dataset, we can assess the transportability of interventions across environments, specifically cities. In this study, families living in high-rise public housing in these cities could sign up to be randomized at baseline, the starting point of the experiment, to receive a Section 8 housing voucher that they could use to move out of public housing and into rental housing on the private market. The adult participants and their children were then surveyed at two follow-up time points. The goal was to estimate the effects of the housing intervention on economic, educational, and health outcomes of adults and children over time. Previous work has also used this dataset to assess the transportability of interventions (although not focusing on issues related to missing data) (Rudolph et al., 2021). The fact that this dataset has already been previously used to study transportability further motivates its use in assessing transportability challenges under missing data.

6.2. Data and Variables

The treatment considered here is receipt of a Section 8 housing voucher, (A). The outcome is child mental health or substance abuse during the follow-up years (Y). The mediator we focus on is overall parental health, (C), which is an established mediator of the effect of neighborhood factors on youth substance abuse (Buu et al., 2009). We control for an extensive set of covariates, W, at the individual and family levels, including sociodemographics, neighborhood characteristics.

Based on prior research with MTO (Rudolph et al., 2018), we consider data from cities of: New York, Los Angeles, Boston, and Chicago. Data from Baltimore was not considered in the analysis since voucher receipt was not associated with moving to a lower-poverty neighborhood there, unlike

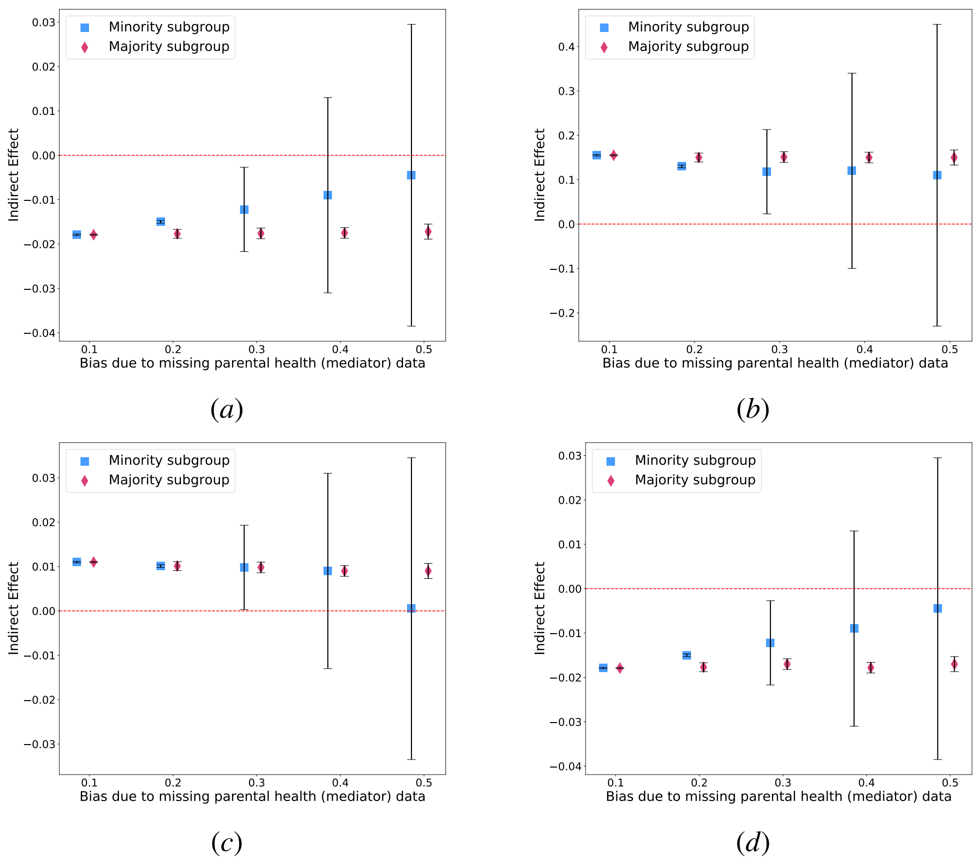

Figure 4: Results from the sensitivity analysis under the sensitivity framework for MTO data with Los Angeles as the target environment in (a), Boston as the target environment in (b), New York City as the target environment in (c), and Chicago as the target environment in (d). We vary the bias in R^2 due to missing mediator (parental health) across the x-axis for the minority and majority groups independently and plot the range of estimated stochastic indirect effect (SIE) values on the y-axis. The solid bar denotes the point estimate bounds for a specified proportion of missing mediator data, estimated as the point estimate plus 95% confidence intervals.

in other cities (Rudolph et al., 2018). When data from one city was used as the target environment, data from the other three cities was combined as the source environment. For example, with New York as the target environment, data from Los Angeles, Boston, and Chicago was combined as the source environment.

To assess the sensitivity to missingness, we induce missingness in the mediator, 'parental health' for the minority group. This systematic induction of missingness allows us to assess at what proportions of missingness the causal effect of receiving a Section 8 voucher becomes insignificant. Missingness is introduced randomly in the mediator (parental health), C, for 0.1, 0.2, 0.3, 0.4, and 0.5 proportion of the total sample for the minority group. Since African Americans and Latinos comprised the majority of the racial and ethnic groups, we considered African Americans as the minority group ($W = 0$) and Latinos as the majority group ($W = 1$). This group assignment follows from previous studies that have focused on African American and Latino subgroups for analysis in MTO (Rudolph et al., 2018).

6.3. Results

We assess the point estimate of the stochastic indirect effect along with the confidence interval as a function of the variance in the weights not explained by the missing mediator data, R^2. With a missingness proportion close to 0.1, the point estimate is approximately -0.018 for Los Angeles, 0.161 for Boston, 0.012 for New York, and -0.018 for Chicago for the majority and minority groups as shown in Figure 4. This effect direction aligns with previous research, which found that moving to lower-poverty neighborhoods reduces overall substance abuse and improves mental health for youth in certain cities such as Los Angeles and Chicago but not in the case of New York and Boston (Sanbonmatsu et al., 2011).

However, as missingness increases, the indirect transported causal effect estimate is close to -0.009 for the minority group with confidence intervals (-0.03, 0.015) for a bias of approximately 0.37, suggesting that the effect becomes insignificant for the minority group for Los Angeles. On the other hand, for the majority group, the stochastic indirect effect is close to -0.192 with a confidence interval (-0.182, -0.206) and is thus significant for the majority group. As bias due to missingness increases to 0.5, the indirect effect in Los Angeles for the minority group is close to -0.002 with a confidence interval (-0.044, 0.031), and for the majority group is approximately -0.191 with a confidence interval (-0.182, -0.206). We also observe similar effects in the case of Chicago. However, in the case of Boston, with increasing missingness, the stochastic indirect effect for the minority group is 0.101 with a confidence interval (-0.210,0.422) while that for the majority group is 0.160 with a confidence interval (0.159,0.161) for a bias of 0.5 due to missing parental data (mediator). In the case of New York, for a 0.5 bias due to missingness, the indirect effect for the minority group is 0.001 with a confidence interval (-0.032,0.033) for the minority group. As the bias due to missingness increases, we find that the transported indirect effect for Los Angles, Boston, New York, and Chicago becomes insignificant for the minority group but remains significant for the majority group.

7. Discussion

In this work, we focus on the issue of transporting mediation effects of interventions from a source environment to a target environment when mediator data is missing in the target environment. This is an important problem because of many situations in which it has been demonstrated that mediators, such as individual behaviors like smoking status or alcohol consumption disproportionately affect minority subgroups due to upstream exposures, such as neighborhood SES, which influence them. Moreover, data about these individual behaviors are commonly obtained inconsistently across minority groups, resulting in missing data. Since these individual behaviors mediate the effect of population-level factors such as neighborhood socioeconomic status on health outcomes, our study examines to what degree their missingness can bias the estimated indirect causal effect for the minority subgroup compared to the majority subgroup. Our contributions are: 1) introducing a sensitivity framework to assess the impact of missing mediator data on transported indirect effects, and 2) demonstrating that transported indirect effect is impacted by missingess in the mediator data. In simulated data, analyzing the residual bias in the estimated transported indirect effect by varying the magnitude of the missingness shows that the transported indirect effect becomes insignificant for the minority group as the missingness increases beyond a threshold (of 0.29 for our causal graph). We observe similar characteristics (with a threshold of 0.26), when evaluating the effect of moving to a better neighborhood on child mental health and substance abuse where there is also missingness in the mediator, overall parental health for LA, Boston, New York and Chicago. Missingness in parental health renders the effect of moving to a better neighborhood on child mental health and substance abuse insignificant for the minority group.

Certain limitations to this framework should be noted. Our approach relies on the assumption that the mediator's missingness pattern in the target environment is known (Simkhada et al., 2008); adopting the framework where this assumption fails is a potential future direction. While this assumption is supported by knowledge from public health, it should be noted that the specific missingness pattern is broad enough to allow multiple factors to affect missingness in the mediator. Moreover, while the analysis in this study was performed by only including the samples for which data is not missing, adapting different missing data imputation strategies to assess the bias in the transported causal effect is left for the future. This is challenging, considering that the missingness pattern we are concerned with is specific to 'missing not at random,' which can involve potential bias depending on the imputation strategy. Moreover, considering settings where there could be unmeasured confounding between the treatment and the outcome, the treatment and the mediator is another interesting future direction.

References

Isaiah Andrews and Emily Oster. A simple approximation for evaluating external validity bias. *Economics Letters*, 178:58–62, 2019.

Elias Bareinboim and Judea Pearl. Transportability of causal effects: Completeness results. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 26, 2012.

Elias Bareinboim and Judea Pearl. Meta-transportability of causal effects: A formal approach. In *Artificial Intelligence and Statistics*, pages 135–143. PMLR, 2013.

Elias Bareinboim and Judea Pearl. Transportability from multiple environments with limited experiments: Completeness results. *Advances in neural information processing systems*, 27:280–288, 2014.

Paula Braveman and Laura Gottlieb. The social determinants of health: it's time to consider the causes of the causes. *Public health reports*, 129(1_suppl2):19–31, 2014.

Anne Buu, Cydney Dipiazza, Jing Wang, Leon I Puttler, Hiram E Fitzgerald, and Robert A Zucker. Parent, family, and neighborhood effects on the development of child substance use and other psychopathology from preschool to the start of adulthood. *Journal of studies on alcohol and drugs*, 70(4):489–498, 2009.

Michael N Cantor and Lorna Thorpe. Integrating data on social determinants of health into electronic health records. *Health Affairs*, 37(4):585–590, 2018.

Carlos Cinelli and Judea Pearl. Generalizing experimental results by leveraging knowledge of mechanisms. *European Journal of Epidemiology*, 36(2):149–164, 2021.

Alice M Dalton and Andrew P Jones. Residential neighbourhood greenspace is associated with reduced risk of cardiovascular disease: A prospective cohort study. *PLoS One*, 15(1):e0226524, 2020.

Ana V Diez-Roux. Multilevel analysis in public health research. *Annual review of public health*, 21(1): 171–192, 2000.

Naoki Egami and Erin Hartman. Covariate selection for generalizing experimental results: Application to a large-scale development program in uganda. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 2021.

Levi Frehlich, Tanvir C Turin, Patricia K Doyle-Baker, Justin J Lang, and Gavin R McCormack. Mediation analysis of the associations between neighbourhood walkability and greenness, accelerometer-measured physical activity, and health-related fitness in urban dwelling canadians. *Preventive Medicine*, 178:107792, 2024.

Kevin A Hallgren and Katie Witkiewitz. Missing data in alcohol clinical trials: a comparison of methods. *Alcoholism: Clinical and Experimental Research*, 37(12):2152–2160, 2013.

Kevin A Hallgren, Katie Witkiewitz, Henry R Kranzler, Daniel E Falk, Raye Z Litten, Stephanie S O'Malley, Raymond F Anton, and In conjunction with the Alcohol Clinical Trials Initiative (ACTIVE) Workgroup. Missing data in alcohol clinical trials with binary outcomes. *Alcoholism: Clinical and Experimental Research*, 40(7):1548–1557, 2016.

Leonard "Henckel, Emilija Perković, and Marloes H" Maathuis. "graphical criteria for efficient total effect estimation via adjustment in causal linear models". *"arXiv preprint arXiv:1907.02435"*, "2019".

Melody Huang and Samuel D Pimentel. Variance-based sensitivity analysis for weighting estimators result in more informative bounds. *arXiv preprint arXiv:2208.01691*, 2022.

Melody Y Huang. Sensitivity analysis for the generalization of experimental results. *Journal of the Royal Statistical Society Series A: Statistics in Society*, page qnae012, 2024.

Soowon Kim, Susan Egerter, Catherine Cubbin, Eugene R Takahashi, and Paula Braveman. Potential implications of missing income data in population-based surveys: an example from a postpartum survey in california. *Public health reports*, 122(6):753–763, 2007.

Nancy Krieger, Jarvis T Chen, James H Ware, and Afamia Kaddour. Race/ethnicity and breast cancer estrogen receptor status: impact of class, missing data, and modeling assumptions. *Cancer Causes & Control*, 19 (10):1305–1318, 2008.

Jonathan Levy, Mark van der Laan, Alan Hubbard, and Romain Pirracchio. A fundamental measure of treatment effect heterogeneity. *Journal of Causal Inference*, 9(1):83–108, 2021.

Imke Mayer, Julie Josse, and Traumabase Group. Generalizing treatment effects with incomplete covariates. *arXiv preprint arXiv:2104.12639*, 2021.

Sigrid M Mohnen, Beate Völker, Henk Flap, and Peter P Groenewegen. Health-related behavior as a mechanism behind the relationship between neighborhood social capital and individual health-a multilevel analysis. *BMC public health*, 12:1–12, 2012.

Amir Motamedi, Maryam Ekramzadeh, Ehsan Bahramali, Mojtaba Farjam, and Reza Homayounfar. Diet quality in relation to the risk of hypertension among iranian adults: cross-sectional analysis of fasa persian cohort study. *Nutrition Journal*, 20(1):1–10, 2021.

J Pearl. Causality: Models, reasoning, and inference 2000 cambridge.

Judea Pearl and Elias Bareinboim. Transportability of causal and statistical relations: A formal approach. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 25, 2011.

Kara E Rudolph, Nicole M Schmidt, M Maria Glymour, Rebecca Crowder, Jessica Galin, Jennifer Ahern, and Theresa L Osypuk. Composition or context: using transportability to understand drivers of site differences in a large-scale housing experiment. *Epidemiology (Cambridge, Mass.)*, 29(2):199, 2018.

Kara E Rudolph, Jonathan Levy, and Mark J van der Laan. Transporting stochastic direct and indirect effects to new populations. *Biometrics*, 77(1):197–211, 2021.

Lisa Sanbonmatsu, Lawrence F Katz, Jens Ludwig, Lisa A Gennetian, Greg J Duncan, Ronald C Kessler, Emma K Adam, Thomas McDade, and Stacy T Lindau. Moving to opportunity for fair housing demonstration program: Final impacts evaluation. 2011.

Ayesha Sania, Nicolò Pini, Morgan Nelson, Michael M Myers, Lauren C Shuffrey, Maristella Lucchini, Amy J Elliott, Hein J Odendaal, and William Fifer. The k nearest neighbor algorithm for imputation of missing longitudinal prenatal alcohol data. *Available at SSRN 4065215*, 2021.

Ilya Shpitser, Tyler VanderWeele, and James M Robins. On the validity of covariate adjustment for estimating causal effects. *arXiv preprint arXiv:1203.3515*, 2012.

Bibha Simkhada, Edwin R van Teijlingen, Maureen Porter, and Padam Simkhada. Factors affecting the utilization of antenatal care in developing countries: systematic review of the literature. *Journal of advanced nursing*, 61(3):244–260, 2008.

Janna Skagerstrom, Grace Chang, and Per Nilsen. Predictors of drinking during pregnancy: a systematic review. *Journal of women's health*, 20(6):901–913, 2011.

Andrew C Stevenson, Rachel C Colley, Kaberi Dasgupta, Leia M Minaker, Mylene Riva, Michael J Widener, and Nancy A Ross. Neighborhood fast-food environments and hypertension in canadian adults. *American Journal of Preventive Medicine*, 2023.

Zhiqiang Tan. A distributional approach for causal inference using propensity scores. *Journal of the American Statistical Association*, 101(476):1619–1637, 2006.

Mark J Van der Laan, Sherri Rose, et al. *Targeted learning: causal inference for observational and experimental data*, volume 4. Springer, 2011.

Tyler J VanderWeele and Eric J Tchetgen Tchetgen. Mediation analysis with time varying exposures and mediators. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 79(3):917–938, 2017.

Miao Zhang, Salman Rahman, Vishwali Mhasawade, and Rumi Chunara. Utilizing big data without domain knowledge impacts public health decision-making. *Proceedings of the National Academy of Sciences*, 121 (39):e2402387121, 2024.