Supplementary Materials

Journal Title
XX(X):0–6

DOI: 10.1177/ToBeAssigned
www.sagepub.com/
SAGE

This document provides the supplementary materials of the article "Causal Decomposition Analysis with Synergistic Interventions: A Triply-Robust Machine Learning Approach to Addressing Social Disparities".

Appendix A. Alternative Approaches for Synergistic Interventions

Equalizing the individual factor while setting the system-level factor to a pre-specified level. The expected potential outcome is defined as $E[Y(a, G_{M|R=0,C^M})|R = 1, C = c]$. This represents the average potential outcome for individuals with $R = 1$ and $C = c$, if the system-level factor is set to $A = a$ while the distribution of the individual factor was equal to that of $R = 0$ within the same level of $C^M = c^M$. This expected potential outcome can also be estimated using a triply-robust estimator. The estimation procedure is similar to that described in Section 4.2, except that factor A is not simulated from $P(A|R = 0, C^A)$ but is set to a pre-specified level a. Let $\nu^{A,\tilde{M}}_{y|R,X,A=a,C} \equiv E[\hat{\mu}_{R,X,A,Z,\tilde{M},C}|R, X, A = a, C]$. The triply-robust estimator for the potential outcome ($E[Y(a, G_{M|R=0,C^M})|R = 1, C = c]$) is as

follows:

$$
\begin{aligned}
=&\hat{E}[\hat{\nu}^{A,\tilde{M}}_{y|R,X,A=a,C} + \frac{I(A=a)}{\hat{\pi}_{A|R,X,C}}(\hat{\mu}_{y|R,X,A,Z,\tilde{M},C} - \hat{\nu}^{A,\tilde{M}}_{y|R,X,A=a,C}) \\
&+ \frac{I(A=a)\hat{\pi}_{M|R=0,C}}{\hat{\pi}_{A|R,X,C}\hat{\pi}_{M|R,X,A,Z,C}}(Y - \hat{\mu}_{y|R,X,A,Z,M,C})|R=r, C=c].
\end{aligned} \tag{1}
$$

Setting the individual factor to a pre-specified level while equalizing the system factor. The expected potential outcome is defined as $E[Y(G_{A|R=0,C^A}, M=m)|R=1, C=c]$. This represents the average potential outcome for individuals with $R=1$ and $C=c$, if the distribution of the system-level factor was equalized across groups while individual-level factor was set to $M=m$. The estimation procedure is similar, with factor M set to a pre-specified value m while factor A is simulated from $P(A|R=0,c)$. Here, $\hat{\nu}^{\tilde{A},m}_{y|R,X,\tilde{A},C} \equiv E[\hat{\mu}_{R,X,\tilde{A},Z,M=m,C}|R,X,\tilde{A},C]$. Then, the estimator is as follows:

$$
\begin{aligned}
=&\hat{E}[\hat{\nu}^{\tilde{A},m}_{y|R,X,\tilde{A},C} + \frac{\hat{\pi}_{A|R=0,C}}{\hat{\pi}_{A|R,X,C}}(\hat{\mu}_{y|R,X,\tilde{A},Z,M=m,C} - \hat{\nu}^{\tilde{A},m}_{y|R,X,\tilde{A},C}) \\
&+ \frac{\hat{\pi}_{A|R=0,C}I(M=m)}{\hat{\pi}_{A|R,X,C}\hat{\pi}_{M|R,X,A,Z,C}}(Y - \hat{\mu}_{y|R,X,A,Z,M=m,C})|R=r, C=c].
\end{aligned} \tag{2}
$$

Setting both individual and system factors to pre-specified values. The expected potential outcome is defined as $E[Y(A=a, M=m)|R=1, C=c]$. This represents the average potential outcome for individuals with $R=1$ and $C=c$, if the system-level factor was set to $A=a$ and the individual-level factor was set to $M=m$. We define $\hat{\nu}^{A,m}_{y|R,X,A=a,C} \equiv E[\hat{\mu}_{R,X,A,Z,M=m,C}|R,X,A=a,C]$. Then, the estimator is as follows:

$$
\begin{aligned}
=&\hat{E}[\hat{\nu}^{A,m}_{y|R,X,A=a,C} + \frac{I(A=a)}{\hat{\pi}_{A|R,X,C}}(\hat{\mu}_{y|R,X,A,Z,M=m,C} - \hat{\nu}^{A,m}_{y|R,X,A=a,C}) \\
&+ \frac{I(A=a)I(M=m)}{\hat{\pi}_{A|R,X,C}\hat{\pi}_{M|R,X,A,Z,C}}(Y - \hat{\mu}_{y|R,X,A,Z,M=m,C})|R=r, C=c].
\end{aligned} \tag{3}
$$

Appendix B. Identification Results

The expected potential outcome, $E[Y(G_{a|0,C^A}, G_{m|R=0,c})|R=r, C=c]$, can be expressed as:

$$
\begin{aligned}
&= \sum_{a,m} E[Y(a,m)|R=r, G_{a|0,C^A}=a, G_{m|0,C^M}=m, C=c] P(G_{a|0,C^A}=a|R=r, C=c) \\
&\quad P(G_{m|0,C^M}=m|R=r, G_{a|0,C^A}=a, C=c) \\
&= \sum_{a,m} E[Y(a,m)|R=r, C=c] P(G_{a|0,C^A}=a) P(G_{m|0,C^M}=m) \\
&= \sum_{x,a,m} E[Y(a,m)|R=r, X=x, C=c] P(X=x|R=r, C=c) P(G_{a|0,C^A}=a) P(G_{m|0,C^M}=m) \\
&= \sum_{x,a,m} E[Y(a,m)|R=r, X=x, A=a, C=c] P(X=x|R=r, C=c) P(G_{a|0,C^A}=a) P(G_{m|0,C^M}=m) \\
&= \sum_{x,a,z,m} E[Y(a,m)|R=r, X=x, A=a, Z=z, C=c] P(X=x|R=r, C=c) P(G_{a|0,C^A}=a) \\
&\quad P(Z=z|R=r, X=x, A=a, C=c) P(G_{m|0,C^M}=m) \\
&= \sum_{x,a,z,m} E[Y(a,m)|R=r, X=x, A=a, Z=z, M=m, C=c] P(X=x|R=r, C=c) P(G_{a|0,C^A}=a) \\
&\quad P(Z=z|R=r, X=x, A=a, C=c) P(G_{m|0,C^M}=m) \\
&= \sum_{x,a,z,m} E[Y|R=r, X=x, A=a, Z=z, M=m, C=c] P(X=x|R=r, C=c) P(A=a|R=0, C^A=c^A) \\
&\quad P(Z=z|R=r, X=x, A=a, C=c) P(M=m|R=0, C^M=c^M)
\end{aligned}
\tag{4}
$$

The first, third, and fifth equalities are due to the law of total probability. The second equality is because $G_{a|0,C^A}$ and $G_{m|0,C^M}$ are random and thus are independent of any observed variables. The fourth equality is due to B1 ($Y(a,m) \perp A=a|R=r, X=x, C=c$). The sixth equality is due to A1 ($Y(a,m) \perp M=m|R=r, X=x, A=a, Z=z, C=c$). The seventh equality is due to B3 (consistency) and the definition of $G_{m|0,C^M}$ and $G_{m|0,C^A}$. This completes the proof.

Appendix C. Different Estimators

The last expression of equation (4) can be rewritten as:

$$
\begin{aligned}
&= E_X\Big[E\big[E[Y|R,X,A,Z,M,C]P(M|R=0,C^M)\big|R,X,A,C\big]P(A|R=0,C^A)|R=r,C=c\Big] \\
&= E_{X,A}\Big[E\big[\frac{E[Y|R,X,A,Z,M,C]}{P(A|R,X,C)}P(M|R=0,C^M)|R,X,A,C)\big]P(A|R=0,C^A) \\
&\quad \times P(A|R,X,C)|R=r,C=c\Big] \\
&= E_X\Big[E\big[\frac{E[Y|R,X,A,Z,M,C]}{P(A|R,X,C)}P(M|R=0,C^M)|R,X,C\big]P(A|R=0,C^A)|R=r,C=c\Big] \\
&= E\big[\frac{E[Y|R,X,A,Z,M,C]}{P(A|R,X,C)}P(A|R=0,C^A)P(M|R=0,C^M)|R=r,C=c\big] \\
&= E\big[\frac{P(A|R=0,C^A)P(M|R=0,C^M)}{P(A|R,X,C)}E[\frac{Y}{P(M|R,X,A,Z,C)}|R,X,A,Z,M,C] \\
&\quad \times P(M|R,X,A,Z,C)|R=r,C=c\big] \\
&= E\big[\frac{P(A|R=0,C^A)P(M|R=0,C^M)}{P(A|R,X,C)}E[\frac{Y}{P(M|R,X,A,Z,C)}|R,X,A,Z,C]|R=r,C=c\big] \\
&= E\big[\frac{P(A|R=0,C^A)P(M|R=0,C^M)}{P(A|R,X,C)P(M|R,X,A,Z,C)}Y|R=r,C=c\big]
\end{aligned}
\tag{5}
$$

The first equality is due to the law of iterated expectations. The second equality holds because $P(A|R,X,C)$ is a function of $\{R,X,A,C\}$, so it can be factored into the expectation. The third equality is due to the law of total probability over A, where $P(A|R,X,C)$ integrates out A. The fourth equality is due to the law of iterated expectations to marginalize over X. The fifth equality holds because $P(M|R,X,A,Z,C)$ is a function of $\{R,X,A,Z,C\}$, so it can be factored into the expectation. The sixth equality holds by marginalizing over M. The first equality is the imputation estimator; the fourth equality is the imputation-then-weighting estimator; the seventh equality is the weighting estimator. This completes the proof.

Appendix D. Simulation Results

Table S1. Bias and RMSE by Sample Size, Simulation, and Models

			GLM				XGBoost				XGBoost with Crossfitting			
N	**Sim**	**Property**	δ_{imp}	δ_{wgt}	δ_{iw}	δ_{tr}	δ_{imp}	δ_{wgt}	δ_{iw}	δ_{tr}	δ_{imp}	δ_{wgt}	δ_{iw}	δ_{tr}
500	1	Bias	0.099	-0.288	0.032	0.254	-0.259	-0.326	-0.238	-0.104	-0.131	-0.292	-0.146	0.0118
		RMSE	0.259	0.371	0.327	0.263	0.259	0.371	0.327	0.263	0.259	0.371	0.327	0.263
	2	Bias	0.086	-0.282	0.026	0.217	-0.273	-0.333	-0.265	-0.118	-0.177	-0.256	-0.172	-0.0265
		RMSE	0.273	0.382	0.334	0.316	0.273	0.382	0.334	0.316	0.273	0.382	0.334	0.316
	3	Bias	0.107	-0.282	0.047	0.274	-0.252	-0.311	-0.234	-0.086	-0.144	-0.291	-0.152	0.0172
		RMSE	0.252	0.365	0.311	0.270	0.252	0.365	0.311	0.270	0.252	0.365	0.311	0.270
	4	Bias	0.088	-0.285	0.038	0.238	-0.270	-0.320	-0.245	-0.121	-0.183	-0.283	-0.174	-0.0193
		RMSE	0.270	0.351	0.321	0.265	0.270	0.351	0.321	0.265	0.270	0.351	0.321	0.265
1000	1	Bias	0.083	-0.258	0.025	0.218	-0.110	-0.292	-0.127	-0.078	-0.083	-0.292	-0.146	0.0118
		RMSE	0.205	0.371	0.327	0.207	0.205	0.371	0.327	0.207	0.205	0.371	0.327	0.207
	2	Bias	0.072	-0.243	0.018	0.199	-0.153	-0.292	-0.234	-0.059	-0.183	-0.343	-0.165	-0.0265
		RMSE	0.228	0.313	0.314	0.218	0.228	0.313	0.314	0.218	0.228	0.313	0.314	0.218
	3	Bias	0.093	-0.265	0.033	0.265	-0.144	-0.271	-0.167	-0.087	-0.144	-0.291	-0.152	0.0172
		RMSE	0.196	0.317	0.312	0.186	0.196	0.317	0.312	0.186	0.196	0.317	0.312	0.186
	4	Bias	0.088	-0.255	0.028	0.221	-0.166	-0.283	-0.177	-0.095	-0.183	-0.283	-0.174	-0.0193
		RMSE	0.263	0.325	0.320	0.217	0.263	0.325	0.320	0.217	0.263	0.325	0.320	0.217
2000	1	Bias	-0.00636	-0.114	-0.00996	-0.0215	-0.0831	-0.223	-0.167	-0.0789	-0.131	-0.292	-0.146	0.0118
		RMSE	0.0304	0.153	0.0701	0.0755	0.0838	0.223	0.167	0.0803	0.131	0.325	0.149	0.170
	2	Bias	-0.185	-0.0429	-0.175	-0.0271	-0.131	-0.227	-0.201	-0.118	-0.177	-0.256	-0.172	-0.0265
		RMSE	0.185	0.115	0.175	0.0875	0.131	0.227	0.201	0.118	0.177	0.313	0.172	0.176
	3	Bias	-0.0925	-0.183	-0.00639	-0.00789	-0.0971	-0.220	-0.165	-0.0863	-0.144	-0.291	-0.152	0.0172
		RMSE	0.0925	0.185	0.0456	0.0615	0.0972	0.220	0.165	0.0873	0.144	0.323	0.157	0.164
	4	Bias	-0.184	-0.114	-0.0827	-0.260	-0.133	-0.222	-0.149	-0.0193	-0.183	-0.283	-0.174	-0.0193
		RMSE	0.185	0.125	0.112	0.260	0.153	0.222	0.172	0.190	0.183	0.328	0.174	0.172

Note. True Value: 0.385, N:Sample Size, Sim 1: the Imputation estimator is correctly specified, Sim 2: the weighting estimator is correctly specified, Sim 3: the imputation-then-weighting estimator is correctly specified, Sim 4: all are incorrectly specified.

Appendix E. Details of Case Study

We provide detailed steps for estimating disparity reduction for Black students compared to White students using the four estimators introduced in Section 4.2. The estimates were obtained from the original data. We begin by outlining the procedure using a GLM.

1. Estimate the average math score among Black students within the same levels of gender and native language ($E[Y|R=1,C=c]$, where $R=1$ indicates Black students). To do this, we fit a linear model regressing math score on gender and native language among Black students. The average math score within this subgroup can be obtained from the intercept of the fitted linear model. In our case study, $\hat{E}[Y|R=1,C=c]=0.190$.
2. Estimate the potential math score among Blacks if their access to high-performing schools and Algebra I instruction by 9th grade were equalized, using the Pure Imputation Estimator. Following the procedure described in Section 4.2, we obtain $\hat{E}[\hat{\nu}_{y|R,X,\tilde{A},C}|R=1,C=c]=0.278$.
3. Estimate the potential math score among Blacks if their access to high-performing schools and Algebra I instruction by 9th grade were equalized, using the Weighting Estimator. Following the procedure described in Section 4.2, we obtain $\hat{E}[\frac{\hat{\pi}_{A|R=0,C^A}\hat{\pi}_{M|R=0,C^M}}{\hat{\pi}_{A|R,X,C}\hat{\pi}_{M|R,X,A,Z,C}}Y|R=1,C=c]=0.281$.
4. Calculate the potential math score among Blacks if their access to high-performing schools and Algebra I instruction by 9th grade were equalized, using the Imputation-then-Weighting estimator. Following the procedure described in Section 4.2, we obtain $\hat{E}[\frac{\hat{\pi}_{A|R=0,C^A}}{\hat{\pi}_{A|R,X,C}}\hat{\mu}_{y|R,X,A,Z,\tilde{M},C}|R=1,C=c]=0.271$.
5. Combining three estimators, we calculate the potential math score using a triply-robust estimator as shown in equation (6). We obtain $\hat{E}[\hat{\nu}_{y|R,X,\tilde{A},C}+\frac{\hat{\pi}_{A|R=0,C^A}}{\hat{\pi}_{A|R,X,C}}(\hat{\mu}_{y|R,X,\tilde{A},Z,\tilde{M},C}-\hat{\nu}_{y|R,X,\tilde{A},C})+\frac{\hat{\pi}_{A|R=0,C^A}\hat{\pi}_{M|R=0,C^M}}{\hat{\pi}_{A|R,X,C}\hat{\pi}_{M|R,X,A,Z,C}}(Y-\hat{\mu}_{y|R,X,A,Z,M,C})|R=1,C=c]=0.286$.
6. Calculate disparity reduction as the difference between the average math score among Black students ($\hat{E}[Y|R=r,C=c]$) and the potential math score among Black students ($\hat{E}[Y(G_{a|0,C^A},G_{m|0,C^M})|R=r,C=c]$ using aforementioned estimators. We obtained $\hat{\delta}_c^{imp}(1)=-0.134$ using the imputation estimator; $\hat{\delta}_c^{wgt}(1)=0.026$ using the weighting estimator;

$\hat{\delta}_c^{iw}(1) = -0.263$ using the imputation-then-weighting estimator; and $\hat{\delta}_c^{tr}(1) = -0.096$ using the triply-robust estimator.

7. Standard errors were computed based on 500 clustered bootstraps based on school ID.

The same steps apply when using XGBoost without cross-fitting. The only difference is that the models must be trained using XGBoost. For model training, we used default settings for `max depth` (6) and `nrounds` (100). To prevent overfitting, we set `early stopping rounds=10`. The resulting estimates are: $\hat{\delta}_c^{imp}(1) = -0.089$ using the imputation estimator; $\hat{\delta}_c^{wgt}(1) = -0.160$ using the weighting estimator; $\hat{\delta}_c^{iw}(1) = -0.105$ using the imputation-then-weighting estimator; and $\hat{\delta}_c^{tr}(1) = -0.185$ using the triply-robust estimator.

For XGBoost with cross-fitting, we trained models using different data than those used for prediction to avoid overfitting. Specifically, the data were split into $K = 5$ folds (I_1, I_2, I_3, I_4, and I_5). For each fold $k \in K$, models were trained using the data excluding that fold (I_{-k}) and predictions were obtained from the withheld data (I_k). The procedure were repeated for K=5 folds, and disparity reduction was estimated based on the full sample. The resulting estimates are: $\hat{\delta}_c^{imp}(1) = -0.134$ using the imputation estimator; $\hat{\delta}_c^{wgt}(1) = 0.026$ using the weighting estimator; $\hat{\delta}_c^{iw}(1) = -0.264$ using the imputation-then-weighting estimator; and $\hat{\delta}_c^{tr}(1) = -0.191$ using the triply-robust estimator.