Causal Fairness under distribution shifts

Alexander Kichutkin

1 Task Description

Solving (causal) fairness under distribution shifts involves understanding distribution shifts.

- Explore how causal fairness relations behave in presence of distribution shifts

In this write-up we will exclusively focus on the the outlined task. Using the notion of causal fairness at an x-specific level as introduced by Plečko and Bareinboim (2022)[4] and rules of causal inference we will investigate how causal fairness relations behave in presence of distribution shifts. We begin by briefly describing the elements of causal inference, causal fairness analysis and transportability theory needed. Afterwards, we will derive an expression for x-specific effects in the observational domain. Then we will investigate how these change under distribution shifts. Finally, we conclude by describing the consequences for transportability of causal fairness relations.

2 Preliminaries

2.1 Causal Inference

For our analysis we will need the rules of do-calculus and the rules and axioms of structural counterfactuals as laid out in Pearl (2009)[2].

2.1.1 do-calculus

The do-calculus follows the following rules:

do-calculus:
Consider a DAG G associated with a causal model, and P(.) being the induced probability distribution. For any disjoint subsets of variables X, Y, Z, and W, we have the following rules.

Rule 1 *(Insertion/deletion of observations)*:

$$P(y|do(x), z, w) = P(y|do(X), W) \ \textit{if} \ (Y \perp\!\!\!\perp Z|X, W)_{G_{\overline{X}}} \tag{1}$$

Rule 2 *(Action/observation exchange)*:

$$P(y|do(x),do(z),w) = P(y|do(X),z,W) \text{ } \textit{if} \text{ } (Y \perp\!\!\!\perp Z|X,W)_{G_{\overline{X},\underline{Z}}} \tag{2}$$

Rule 3 *(Insertion/deletion of actions)*:

$$P(y|do(x),do(z),w) = P(y|do(X),w) \text{ } \textit{if} \text{ } (Y \perp\!\!\!\perp Z|X,W)_{G_{\overline{X},\overline{Z(W)}}} \tag{3}$$

We will especially use *rule 2* in the final step to obtain an observable expression for the x-specific effects in section 3.

2.1.2 Axioms of Structural Counterfactuals and Restrictions Rules

Following Section 7.3.1 and 7.3.2 of Pearl (2009)[2] directly:

Consistency:
For any set of variables Y and X in a causal model, we have:

$$X(u) = x \Rightarrow Y(u) = Y_x(u) \tag{4}$$

Composition:
For any three sets of endogenous variables X, Y, and W in a causal model, we have:

$$W_x(u) = w \Rightarrow Y_{x,w} = Y_x(u) \tag{5}$$

Restrictions Rules:
Rule 1 *(exclusion restrictions)*:
For every variable Y having parents PA_Y and for every set of variables $Z \subset V$ disjoint of PA_Y, we have:

$$Y_{pa_Y}(u) = Y_{pa_Y,z}(u) \tag{6}$$

Rule 2 *(independence restrictions)*:
If $Z_1, ..., Z_k$ is any set of nodes in V not connected to Y via paths containing only U variables, we have:

$$Y_{pa_Y} \perp\!\!\!\perp Z_{1,pa_{z_1}}, ..., Z_{k,pa_{z_k}} \tag{7}$$

2.2 Transportability

Transportability theory addresses the problem of transferring information learned from experiments to a different environment, in which only passive observations can be collected. It received formal treatment by Bareinboim and Pearl (2011,2012)[3],[1], where we will mainly use results from the first paper. It introduces notations such as selection variables and selection diagrams to characterize different environments. The selection diagram allows to represent two different environments in one graph. Differences between the environments are indicated by the selection variables and their outgoing edges. Fig.1 shows our assumed source environment and fig2 shows how we can use selection variables and diagrams to represent different shifts compared to our source domain. We will mainly use the following theorem and definitions from [3] for our preliminary analysis.

Theorem 1: *Let D be the selection diagram characterizing Π and Π^*, and S a set of selection variables in D. The relation $R = P(y|do(x), z)$ is transportable from Π to Π^* if and only if the expression $P(y|do(x), z, s)$ is reducible, using the rules of do-calculus, to an expression in which S appears only as a conditioning variable in do-free terms.*

Theorem 1 is a declarative high-level statement indicating in which cases it is in principle possible to transport a relation R.

Definition: (Direct Transportability)
A causal relation R is said to be directly transportable from Π to Π^*, if $R(\Pi^*) = R(\Pi^*)$.

Definition: (Trivial Transportability)
A causal relation R is said to be trivially transportable from Π to Π^*, if $R(\Pi^*)$ is identifiable from (G^*, P^*).

The two definitions above describe, given that transportability with respect to *Theorem 1* is possible, the way we can transport the relation R. Direct transportability describes that R is invariant against distribution shifts as it stays the same. Trivial transportability on the other end can describe the opposite. If no part of R is invariant to the shift, but we have access to the observational distribution P^* and graph G^*, we can simply recompute R in the new environment. It could also be that only a few parts of R are not invariant. If we have access to the graph and observational distribution we can simply re-estimate those parts.

2.3 Causal Fairness Analysis

Plečko and Bareinboim (2022)[4] introduce a foundational framework for investigating Fair Machine Learning from a causal perspective. We will especially use the notion of the Standard Fairness Model (SFM), the Total Variation (TV) measure and its x-specific decomposition, which we will take directly from [4].

Definition: (Standard Fairness Model)
The standard fairness model (SFM) is the causal diagram G^{SFM} over endogenous variables X, Z, W, Y and given by fig1. The nodes represent:

- *the protected attribute, labeled X (e.g., gender, race, religion),*

- *the set of confounding variables Z, which are not causally influenced by the attribute X (e.g., demographic information, zip code),*
- *the set of mediator variables W that are possibly causally influenced by the attribute (e.g., educational level or other job-related information),*
- *the outcome variable Y (e.g., admissions, hiring, salary).*

Definition: *(Total Variation)*
The total variation (TV) is also known as the parity gap in the literature and is given by:

$$TV_{x_0,x_1}(y) = P(y|x_1) - P(y|x_1) \tag{8}$$

Here, x_0, x_1 are the two values of the protected attribute X, and Y is the outcome of interest.

It can be shown that the TV measure can be decomposed into the following more fine-grained measures.

Definition: *(Population level effects)*
On a population level. the TV can be decomposed into the NDE,NIE and Exp-SE as follows:

$$NDE_{x_0,x_1}(y) = P(y_{x_1,W_{x_0}}) - P(y_{x_0}) \tag{9}$$

$$NIE_{x_1,x_0}(y) = P(y_{x_1,W_{x_0}}) - P(y_{x_1}) \tag{10}$$

$$Exp - SE_x(y) = P(y|x) - P(y_x) \tag{11}$$

We can go even one step further and filter by specific $X = x$ to obtain the x-specific effects.

Definition: *(x-specific DE, IE, and SE)*
The x-total, direct, indirect, spurious effects are defined as follows:

$$x - DE_{x_0,x_1}(y|x_0) = P(y_{x_1,W_{x_0}}|x_0) - P(y_{x_0}|x_0) \tag{12}$$

$$x - IE_{x_1,x_0}(y|x_0) = P(y_{x_1,w_{x_0}}|x_0) - P(y_{x_1}|x_0) \tag{13}$$

$$x - SE_{x_1,x_0}(y|x_0) = P(y_{x_1}|x_0) - P(y_{x_1}|x_1) \tag{14}$$

Similarly, we can also filter for specific $Z = z$.

Definition: *(z-specific effects)*
The z-specific direct and indirect effects are defined as follows:

$$z - DE_{x_0,x_1}(y|z) = P(y_{x_1,W_{x_0}}|z) - P(y_{x_0}|z) \tag{15}$$

$$x - IE_{x_1,x_0}(y|z) = P(y_{x_1,w_{x_0}}|z) - P(y_{x_1}|z) \tag{16}$$

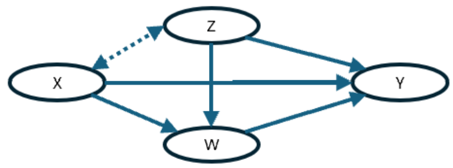


Figure 1: G^{SFM} representing the standard fairness model

The (x,z)-specific direct and indirect effects are defined as follows:

$$(x,z)-DE_{x_0,x_1}(y|x_0,z) = P(y_{x_1,W_{x_0}}|x_0,z) - P(y_{x_0}|x_0,z) \tag{17}$$

$$(x,z)-IE_{x_1,x_0}(y|x_0,z) = P(y_{x_1,w_{x_0}}|x_0,z) - P(y_{x_1}|x_0,z) \tag{18}$$

We can also

In the following we will derive observable expressions for the x-specific effects. Then, we will investigate how they behave under distribution shifts within the SFM structure.

3 Identifying the fairness map in the observational domain

In order to analyse the fairness measures of the TV family we need to derive an expression which is identifiable from the observational distribution for them. We begin at the population level and work our way through the fairness map.

3.1 Population level

At the population level we do not filter for specific covariates, but examine the full population.

3.1.1 $NDE_{x_0,x_1}(y)$

Using the consistency axiom and conditioning on W and Z we can rewrite the NDE as follows.

$$\begin{aligned} NDE_{x_0,x_1}(y) &= P(y_{x_1,W_{x_0}}) - P(y_{x_0}) \\ &= \Sigma_{z,w} P(y_{x_1,w}|w_{x_0},z)P(w_{x_0}|z)P(z) \\ &- P(y_{x_0,w}|w_{x_0},z)P(w_{x_0}|z)P(z) \end{aligned} \tag{19}$$

Using the consistency axiom and exclusions restrictions rule we have:

$$Z = Z_{x_0} = Z_{x_1} = z \Rightarrow W = W_z, Y = Y_z \tag{20}$$

This allows us to further rewrite the expression to:

$$= \Sigma_{z,w}[P(y_{x_1,w,z}|w_{x_0,z},z) - P(y_{x_0,w,z}|w_{x_0,z},z)]P(w_{x_0,z}|z)P(z) \tag{21}$$

We can apply the independence restrictions rule to obtain:

$$= \Sigma_{z,w}[P(y_{x_1,w,z}) - P(y_{x_0,w,z})]P(w_{x_0,z})P(z) \tag{22}$$

Finally, we can apply rule 2 of do-calculus to obtain an expression for the NDE in the observational domain.

$$= \Sigma_{z,w}[P(y|x_1,w,z) - P(y|x_0,w,z)]P(w|x_0,z)P(z) \tag{23}$$

3.1.2 $NIE_{x_1,x_0}(y)$

We will apply the same reasoning as for the NDE for the NIE and all following effect-levels. Thus, we obtain for the NIE:

$$\begin{aligned} NIE_{x_1,x_0}(y) &= P(y_{x_1,w_{x_0}}) - P(y_{x_1}) \\ &= \Sigma_{z,w}P(y|x_1,w,z)[P(w|x_0,z) - P(w|x_1,z)]P(z) \end{aligned} \tag{24}$$

3.1.3 $Exp - SE_x(y)$

Similarly, for the $Exp - SE$ we obtain:

$$\begin{aligned} Exp - SE_x(y) &= P(y|x) - P(y_x) \\ &= \Sigma_z P(y|x,z)[P(z|x) - P(z)] \end{aligned} \tag{25}$$

3.2 x-specific level

We now filter the population for specific $X = x$. We can re-use the results and reasoning from the population level as the only difference between both is mostly the conditioning on specific $X = x$.

3.2.1 $x - DE_{x_0,x_1}(y|x_0)$

Using the same logic as for the NDE, we obtain the following for the $x - DE$

$$\begin{aligned} x - DE_{x_0,x_1}(y|x_0) &= P(y_{x_1,W_{x_0}}|x_0) - P(y_{x_0}|x_0) \\ &= \Sigma_{z,w}P(y_{x_1,w}|x_0,w_{x_0},z)P(w_{x_0}|x_0,z)P(z|x_0) \\ &- P(y_{x_0,w}|x_0,w_{x_0},z)P(w_{x_0}|x_0,z)P(z|x_0) \\ &= \Sigma_{z,w}[P(y_{x_1,w,z}|x_0,w_{x_0,z},z) \\ &- P(y_{x_0,w,z}|x_0,w_{x_0,z},z)]P(w_{x_0,z}|x_0,z)P(z|x_0) \\ &= \Sigma_{z,w}[P(y_{x_1,w,z}) - P(y_{x_0,w,z})]P(w_{x_0,z})P(z|x_0) \\ &= \Sigma_{z,w}[P(y|x_1,w,z) - P(y|x_0,w,z)]P(w|x_0,z)P(z|x_0) \end{aligned} \tag{26}$$

3.2.2 $x - IE_{x_1,x_0}(y|x_0)$

Using the same logic as for the NIE, we obtain the following for the $x - IE$

$$
\begin{aligned}
x - IE_{x_1,x_0}(y|x_0) &= P(y_{x_1,w_{x_0}}|x_0) - P(y_{x_1}|x_0) \\
&= \Sigma_{z,w} P(y|x_1, w, z)P(w|x_0, z)P(z|x_0) \\
&- P(y_{x_1,w}|x_0, w_{x_1}, z)P(w_{x_1}|x_0, z)P(z|x_0) \\
&= \Sigma_{z,w} P(y|x_1, w, z)P(w|x_0, z)P(z|x_0) \\
&- P(y_{x_1,w,z}|x_0, w_{x_1,z}, z)P(w_{x_1,z}|x_0, z)P(z|x_0) \\
&= \Sigma_{z,w} P(y|x_1, w, z)P(w|x_0, z)P(z|x_0) \\
&- P(y_{x_1,w,z})P(w_{x_1,z})P(z|x_0) \\
&= \Sigma_{z,w} P(y|x_1, w, z)P(w|x_0, z)P(z|x_0) \\
&- P(y|x_1, w, z)P(w|x_1, z)P(z|x_0) \\
&= \Sigma_{z,w} P(y|x_1, w, z)[P(w|x_0, z) - P(w|x_1, z)]P(z|x_0)
\end{aligned} \tag{27}
$$

3.2.3 $x - SE_{x_1,x_0}(y|x_0)$

Using previous results and reasoning, we derive the following expression for the x-specific spurious effect in the observational domain.

$$
\begin{aligned}
x - SE_{x_1,x_0}(y|x_0, z) &= P(y_{x_1}|x_0) - P(y_{x_1}|x_1) \\
&= \Sigma_{z,w} P(y|x_1, w, z)P(w|x_1, z)P(z|x_0) \\
&- P(y_{x_1}, w|x_1, w_{x_1}, z)P(w_{x_1}|x_1, z)P(z|x_1) \\
&= \Sigma_{z,w} P(y|x_1, w, z)P(w|x_1, z)P(z|x_0) \\
&- P(y|x_1, w, z)P(w|x_1, z)P(z|x_1) \\
&= \Sigma_{z,w} P(y|x_1, w, z)P(w|x_1, z)[P(z|x_0) - P(z|x_1)] \\
&= \Sigma_{z} P(y|x_1, z)[P(z|x_0) - P(z|x_1)]
\end{aligned} \tag{28}
$$

3.3 (x,z)-specific level

At the $(x - z)$-specific level we now additionally filter the population on $Z = z$. Thus, we have an additional conditioning variable. Using the same reasoning and previous results, we obtain the following (x, z)-specific results:

3.3.1 $(x, z) - DE_{x_0,x_1}(y|x_0, z)$

$$
\begin{aligned}
(x, z) - DE_{x_0,x_1}(y|x_0, z) &= P(y_{x_1,W_{x_0}}|x_0, z) - P(y_{x_0}|x_0, z) \\
&= \Sigma_w P(y_{x_1,w}|x_0, w_{x_0}, z)P(w_{x_0}|x_0, z) \\
&- P(y_{x_0,w}|x_0, w_{x_0}, z)P(w_{x_0}|x_0, z) \\
&= \Sigma_w [P(y_{x_1,w,z}|x_0, w_{x_0,z}, z) - P(y_{x_0,w,z}|x_0, w_{x_0,z}, z)]P(w_{x_0,z}|x_0, z) \\
&= \Sigma_w [P(y_{x_1,w,z}) - P(y_{x_0,w,z})]P(w_{x_0,z}) \\
&= \Sigma_w [P(y|x_1, w, z) - P(y|x_0, w, z)]P(w|x_0, z)
\end{aligned} \tag{29}
$$

3.3.2 $(x,z) - IE_{x_1,x_0}(y|x_0,z)$

$$\begin{aligned}
(x,z) - IE_{x_1,x_0}(y|x_0,z) &= P(y_{x_1,w_{x_0}}|x_0,z) - P(y_{x_1}|x_0,z) \\
&= \Sigma_w P(y|x_1,w,z)P(w|x_0,z) \\
&- P(y_{x_1,w}|x_0,w_{x_1},z)P(w_{x_1}|x_0,z) \\
&= \Sigma_w P(y|x_1,w,z)P(w|x_0,z) \\
&- P(y_{x_1,w,z}|x_0,w_{x_1,z},z)P(w_{x_1,z}|x_0,z) \\
&= \Sigma_w P(y|x_1,w,z)P(w|x_0,z) \\
&- P(y_{x_1,w,z})P(w_{x_1,z}) \\
&= \Sigma_w P(y|x_1,w,z)P(w|x_0,z) \\
&- P(y|x_1,w,z)P(w|x_1,z) \\
&= \Sigma_w P(y|x_1,w,z)[P(w|x_0,z) - P(w|x_1,z)]
\end{aligned} \tag{30}$$

3.4 z-specific-level

3.4.1 $z - DE_{x_0,x_1}(y|z)$

$$\begin{aligned}
z - DE_{x_0,x_1}(y|z) &= P(y_{x_1,W_{x_0}}|z) - P(y_{x_0}|z) \\
&= \Sigma_w P(y_{x_1,w}|w_{x_0},z)P(w_{x_0}|z) \\
&- P(y_{x_0,w}|w_{x_0},z)P(w_{x_0}|z) \\
&= \Sigma_w [P(y_{x_1,w,z}|w_{x_0,z},z) - P(y_{x_0,w,z}|w_{x_0,z},z)]P(w_{x_0,z}|z) \\
&= \Sigma_w [P(y_{x_1,w,z}) - P(y_{x_0,w,z})]P(w_{x_0,z}) \\
&= \Sigma_w [P(y|x_1,w,z) - P(y|x_0,w,z)]P(w|x_0,z)
\end{aligned} \tag{31}$$

3.4.2 $z - IE_{x_1,x_0}(y|z)$

$$\begin{aligned}
z - IE_{x_1,x_0}(y|z) &= P(y_{x_1,w_{x_0}}|z) - P(y_{x_1}|z) \\
&= \Sigma_w P(y|x_1,w,z)P(w|x_0,z) \\
&- P(y_{x_1,w}|w_{x_1},z)P(w_{x_1}|z) \\
&= \Sigma_w P(y|x_1,w,z)P(w|x_0,z) \\
&- P(y_{x_1,w,z}|w_{x_1,z},z)P(w_{x_1,z}|z) \\
&= \Sigma_w P(y|x_1,w,z)P(w|x_0,z) \\
&- P(y_{x_1,w,z})P(w_{x_1,z}) \\
&= \Sigma_w P(y|x_1,w,z)P(w|x_0,z) \\
&- P(y|x_1,w,z)P(w|x_1,z) \\
&= \Sigma_w P(y|x_1,w,z)[P(w|x_0,z) - P(w|x_1,z)]
\end{aligned} \tag{32}$$

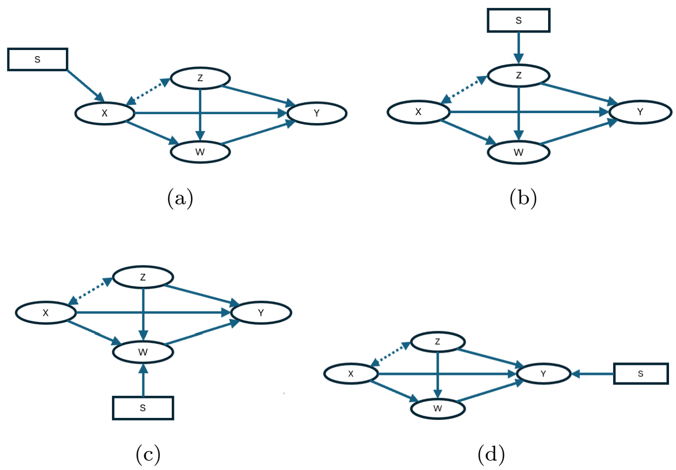


Figure 2: (a) G^{SFM} with a shift in X, indicated by selection variable S (b) G^{SFM} with a shift in Z, indicated by selection variable S (c) G^{SFM} with a shift in W, indicated by selection variable S (d) G^{SFM} with a shift in Y, indicated by selection variable S

4 Identifying the fairness map under distribution shifts

We are interested in exploring whether causal fairness relations are transportable over different domains. The relation of interest is the triplet $R = [x - DE_{x_0,x_1}(y|x_0), x - IE_{x_1,x_0}(y|x_0), x - SE_{x_1,x_0}(y)]$ and the question whether when computed in an source environment it is applicable to new, but seemingly similar ones. The structure of the SFM will represent our source domain.We will construct new target domains using selection variables and diagrams. For each possible shift within the SFM, we will analyse how the x-specific effects will change. For simplicity we also assume that all nodes of the SFM are single variables and not representing a set of variables.

4.1 Single Shifts

Single shifts describe the presence of an distribution shift in a single endogenous variable. We will represent it using a selection variable S.

4.1.1 Shift in X

A shift in X can be described using a selection variable S and augmenting G^{SFM} by adding S and an corresponding edge $S \to X$. The resulting selection diagram is depicted in Fig.2a.

4.1.1.1 Population Level

The NDE_{x_0,x_1} in presence of a shift in X, can be written as:

$$NDE_{x_0,x_1}(y|s) = P(y_{x_1,W_{x_0}}|s) - P(y_{x_0}|s) \tag{33}$$

Conditioning on W and Z again, we obtain:

$$\begin{aligned} &= \Sigma_{z,w} P(y_{x_1,w}|w_{x_0}, z, s) P(w_{x_0}|z, s) P(z|s) \\ &- P(y_{x_0,w}|w_{x_0}, z, s) P(w_{x_0}|z, s) P(z|s) \end{aligned} \tag{34}$$

Using Fig.2a) we can verify that $Z \perp\!\!\!\perp S$. Hence, we arrive at:

$$= \Sigma_{z,w}[P(y_{x_1,w}|w_{x_0}, z, s) - P(y_{x_0,w}|w_{x_0}, z, s)]P(w_{x_0}|z, s)P(z) \tag{35}$$

We can now proceed with the same steps as we have done in section 3.1.1 ending up with the following preliminary expression.

$$= \Sigma_{z,w}[P(y_{x_1,w,z}|w_{x_0,z}, z, s) - P(y_{x_0,w,z}|w_{x_0,z}, z, s)]P(w_{x_0,z}|z, s)P(z) \tag{36}$$

Using the independence restrictions rule and do-calculus, we end up with:

$$\begin{aligned} &= \Sigma_{z,w}[P(y_{x_1,w,z}) - P(y_{x_0,w,z})]P(w_{x_0,z})P(z) \\ &= \Sigma_{z,w}[P(y|x_1, w, z) - P(y|x_0, w, z)]P(w|x_0, z)P(z) \end{aligned} \tag{37}$$

Following the same logic we can derive the expressions for NIE_{x_1,x_0}, and $Exp-SE_x$. as:

$$\begin{aligned} NIE_{x_1,x_0}(y|s) &= P(y_{x_1,w_{x_0}}|s) - P(y_{x_1}|s) \\ &= \Sigma_{z,w} P(y|x_1, w, z)P(w|x_0, z)P(z) \\ &- P(y|x_1, w, z)P(w|x_1, z)P(z) \\ &= \Sigma_{z,w} P(y|x_1, w, z)[P(w|x_0, z) - P(w|x_1, z)]P(z) \end{aligned} \tag{38}$$

$$\begin{aligned} Exp - SE_x(y|s) &= P(y|x, s) - P(y_x|s) \\ &= \Sigma_{z,w} P(y|x, w, z)P(w|x, z)P(z|x, s) \\ &- P(y|x, w, z)P(w|x, z)P(z) \end{aligned} \tag{39}$$

As $Z \not\perp\!\!\!\perp S|X$, we write $P^*(z|x) = P(z|x, s)$ and obtain the final expression for $Exp - SE_x$.

$$= \Sigma_z P(y|x, z)(P^*(z|x) - P(z)) \tag{40}$$

Thus, in presence of a shift in X we would need to re-estimate $P^*(z|x) = P(z|x, s)$.

4.1.1.2 x-specific level

As in the previous section, we condition on Z and W.

$$\begin{aligned} x - DE_{x_0,x_1}(y|x_0, s) &= P(y_{x_1,W_{x_0}}|x_0, s) - P(y_{x_0}|x_0, s) \\ &= \Sigma_{z,w} P(y_{x_1,w}|z, w_{x_0}, x_0, s) P(w_{x_0}|z, x_0, s) P(z|x_0, s) \\ &- P(y_{x_0,w}|z, w_{x_0}, x_0, s) P(w_{x_0}|z, x_0, s) P(z|x_0, s) \end{aligned} \tag{41}$$

So far, the only difference to (8) is the additional conditioning variable S indicating the new environment. From the selection diagram one can see that $Y \perp\!\!\!\perp S|(X, Z)$ and $W \perp\!\!\!\perp S|(X, Z)$. Hence, we can write:

$$\begin{aligned} &= \Sigma_{z,w} P(y_{x_1,w}|z, w_{x_0}, x_0) P(w_{x_0}|z, x_0) P(z|x_0, s) \\ &- P(y_{x_0,w}|z, w_{x_0}, x_0) P(w_{x_0}|z, x_0) P(z|x_0, s) \end{aligned} \tag{42}$$

Consulting the selection diagram again, we see that X is a collider on the path from S to Z
($S \rightarrow X \leftarrow\text{-}\text{-}\text{-}\rightarrow Z$). Conditioning on X activates this path. Hence, $Z \not\perp\!\!\!\perp S|X$. Therefore, $P(Z)$ needs to be re-estimated in the new environment and we write $P^*(z|x_0) = P(z|x_0, s)$. The remainder follows the derivation of $x - DE_{x_0,x_1}(y|x_0)$. Thus, we obtain:

$$x - DE_{x_0,x_1}(y|x_0, s) = [P(y|x_1, w, z) - (y|x_0, w, z)] P(w|x_0, z) P^*(z|x_0) \tag{43}$$

For the indirect and spurious effect, we follow the same reasoning and obtain:

$$\begin{aligned} x - IE_{x_1,x_0}(y|x_0, s) &= P(y_{x_1,W_{x_0}}|x_0, s) - P(y_{x_1}|x_0, s) \\ &= \Sigma_{z,w} P(y|x_1, w, z) P(w|x_0, z) P^*(z|x_0) \\ &- P(y_{x_1,w}|x_0, w_{x_1}, z, s) P(w_{x_1}|x_0, z, s) P(z|x_0, s) \\ &= \Sigma_{z,w} P(y|x_1, w, z) P(w|x_0, z) P^*(z|x_0) \\ &- P(y_{x_1,w}|x_0, w_{x_1}, z) P(w_{x_1}|x_0, z) P(z|x_0, s) \\ &= \Sigma_{z,w} P(y|x_1, w, z) P(w|x_0, z) P^*(z|x_0) \\ &- P(y_{x_1,w}|x_0, w_{x_1}, z) P(w_{x_1}|x_0, z) P^*(z|x_0) \\ &= \Sigma_{z,w} P(y|x_1, w, z) [P(w|x_0, z) - P(w|x_1, z)] P^*(z|x_0) \end{aligned} \tag{44}$$

$$\begin{aligned} x - SE_{x_1,x_0}(y|x_0, s) &= P(y_{x_1}|x_0, s) - P(y_{x_1}|x_1, s) \\ &= \Sigma_{z,w} P(y|x_1, w, z) P(w|x_1, z) [P^*(z|x_0) - P^*(z|x_1)] \end{aligned} \tag{45}$$

The results indicate that transporting R is indeed possible when we have a shift in X. However, it requires to re-estimate $P^*(z|x)$ in the target environment.

4.1.1.3 (x,z)-specific and z-specific level

Note, that the affected conditional in presence of a shift in X is $P(Z|X)$. Therefore, we obtain invariant x and (x,z)-specific effects, given by:

$$\begin{aligned}(x,z)-DE_{x_0,x_1}(y|x_0,z) &= P(y_{x_1,W_{x_0}}|x_0,z)-P(y_{x_0}|x_0,z)\\ &= \Sigma_w[P(y|x_1,w,z)-P(y|x_0,w,z)]P(w|x_0,z)\end{aligned} \tag{46}$$

$$(x,z)-IE_{x_1,x_0}(y|x_0,z) = \Sigma_w P(y|x_1,w,z)[P(w|x_0,z)-P(w|x_1,z)] \tag{47}$$

$$z-DE_{x_0,x_1}(y|z) = \Sigma_w[P(y|x_1,w,z)-P(y|x_0,w,z)]P(w|x_0,z) \tag{48}$$

$$z-IE_{x_1,x_0}(y|z) = \Sigma_w P(y|x_1,w,z)[P(w|x_0,z)-P(w|x_1,z)] \tag{49}$$

4.1.2 Shift in Z

We can represent a shift in Z by using a selection variable S and augmenting G^{SFM} by adding S and an corresponding edge $S \to Z$. The resulting selection diagram is depicted in Fig.2b. Using the selection diagram we find that $Y \perp\!\!\!\perp S|(Z,X)$, $W \perp\!\!\!\perp S|(Z,X)$ and $Z \not\perp\!\!\!\perp S$.

4.1.2.1 Population level

$$\begin{aligned}NDE_{x_0,x_1}(y|s) &= P(y_{x_1,W_{x_0}}|s)-P(y_{x_0}|s)\\ &= \Sigma_{z,w}P(y_{x_1,w}|w_{x_0},z,s)P(w_{x_0}|z,s)P(z|s)\\ &- P(y_{x_0,w}|w_{x_0},z,s)P(w_{x_0}|z,s)P(z|s)\\ &= \Sigma_{z,w}[P(y_{x_1,w}|w_{x_0},z,s)-P(y_{x_0,w}|w_{x_0},z,s)]P(w_{x_0}|z,s)P^*(z)\\ &= \Sigma_{z,w}[P(y_{x_1,w,z}|w_{x_0,z},z,s)-P(y_{x_0,w,z}|w_{x_0,z},z,s)]P(w_{x_0,z}|z,s)P^*(z)\\ &= \Sigma_{z,w}[P(y_{x_1,w,z})-P(y_{x_0,w,z})]P(w_{x_0,z})P^*(z)\\ &= \Sigma_{z,w}[P(y|x_1,w,z)-P(y|x_0,w,z)]P(w|x_0,z)P^*(z)\end{aligned} \tag{50}$$

Following the same logic we can derive the expressions for NIE_{x_1,x_0}, and $Exp-SE_x$. as:

$$\begin{aligned}NIE_{x_1,x_0}(y|s) &= P(y_{x_1,w_{x_0}}|s)-P(y_{x_1}|s)\\ &= \Sigma_{z,w}P(y|x_1,w,z)P(w|x_0,z)P(z)\\ &- P(y|x_1,w,z)P(w|x_1,z)P(z)\\ &= \Sigma_{z,w}P(y|x_1,w,z)[P(w|x_0,z)-P(w|x_1,z)]P(z)\end{aligned} \tag{51}$$

$$\begin{aligned}Exp-SE_x(y|s) &= P(y|x,s)-P(y_x|s)\\ &= \Sigma_{z,w}P(y|x,w,z)P(w|x,z)P(z|x,s)\\ &- P(y|x,w,z)P(w|x,z)P(z)\end{aligned} \tag{52}$$

As $Z \not\perp\!\!\!\perp S|X$, we write $P^*(z|x) = P(z|x,s)$ and obtain the final expression for $Exp - SE_x$.

$$= \Sigma_z P(y|x,z)(P^*(z|x) - P(z)) \tag{53}$$

Thus, in presence of a shift in X we would need to re-estimate $P^*(z|x) = P(z|x,s)$.

4.1.2.2 x-specific level

$$\begin{aligned} x - DE_{x_0,x_1}(y|x_0,s) &= P(y_{x_1,w_{x_0}}|x_0,s) - P(y_{x_0}|x_0,s) \\ &= \Sigma_{z,w} P(y_{x_1,w}|x_0,w_{x_0},z,s)P(w_{x_0}|x_0,z,s)P(z|x_0,s) \\ &- P(y_{x_0,w}|x_0,w_{x_0},z,s)P(w_{x_0}|x_0,z,s)P(z|x_0,s) \\ &= \Sigma_{z,w}[P(y|x_1,w,z) - P(y|x_0,w,z)]P(w|x_0,z)P^*(z|x_0) \end{aligned} \tag{54}$$

$$\begin{aligned} x - IE_{x_1,x_0}(y|x_0,s) &= P(y_{x_1,w_{x_0}}|x_0,s) - P(y_{x_1}|x_0,s) \\ &= \Sigma_{z,w} P(y|x_1,w,z)P(w|x_0,z)P^*(z|x_0) \\ &- P(y|x_1,w,z)P(w|x_1,z)P^*(z|x_0) \\ &= \Sigma_{z,w} P(y|x_1,w,z)[P(w|x_0,z) - P(w|x_1,z)]P^*(z|x_0) \end{aligned} \tag{55}$$

$$\begin{aligned} x - SE_{x_1,x_0}(y|x_0,S) &= P(y_{x_1}|x_0,s) - P(y_{x_1}|x_1,s) \\ &= \Sigma_{z,w} P(y|x_1,w,z)P(w|x_1,z)P^*(z|x_0) \\ &- P(y|x_1,w,z)P(w|x_1,z)P^*(z|x_1)) \\ &= \Sigma_{z,w} P(y|x_1,w,z)P(w|x_1,z)[P^*(z|x_0) - P^*(z|x_1)] \end{aligned} \tag{56}$$

The result is similar to shifts in X. R is transportable in presence of a shift in Z. However, it requires to re-estimate $P^*(z|x)$ in the target environment.

4.1.2.3 (x,z)-specific and z-specific level

Similar reasoning applies here as well. Thus, we will have invariant z- and (x,z)-specific effects given by:

$$\begin{aligned} (x,z) - DE_{x_0,x_1}(y|x_0,z) &= P(y_{x_1,W_{x_0}}|x_0,z) - P(y_{x_0}|x_0,z) \\ &= \Sigma_w[P(y|x_1,w,z) - P(y|x_0,w,z)]P(w|x_0,z) \end{aligned} \tag{57}$$

$$(x,z) - IE_{x_1,x_0}(y|x_0,z) = \Sigma_w P(y|x_1,w,z)[P(w|x_0,z) - P(w|x_1,z)] \tag{58}$$

$$z - DE_{x_0,x_1}(y|z) = \Sigma_w[P(y|x_1,w,z) - P(y|x_0,w,z)]P(w|x_0,z) \tag{59}$$

$$z - IE_{x_1,x_0}(y|z) = \Sigma_w P(y|x_1,w,z)[P(w|x_0,z) - P(w|x_1,z)] \tag{60}$$

4.1.3 Shift in W

4.1.3.1 population level

The shifts at the population level are trivially transportable and given by:

$$NDE_{x_0,x_1}(y|s) = \Sigma_{z,w}[P(y|x_1,w,z) - P(y|x_0,w,z)]P^*(w|x_0,z)P(z) \quad (61)$$

$$NIE_{x_1,x_0}(y|s) = \Sigma_{z,w}P(y|x_1,w,z)[P^*(w|x_0,z) - P^*(w|x_1,z)]P(z) \quad (62)$$

$$Exp - SE_x(y|s) = P(y|x,w,z)P^*(w|x,z)(P(z|x) - P(z)) \quad (63)$$

4.1.3.2 x-specific level

We can represent a shift in W by using a selection variable S and augmenting G^{SFM} by adding S and an corresponding edge $S \rightarrow W$. The resulting selection diagram is depicted in Fig.2c. Using the selection diagram we find that $Y \perp\!\!\!\perp S|(Z,W,X)$, $Z \perp\!\!\!\perp S$ and $W \not\perp\!\!\!\perp S$. Using these (in)dependence relations we obtain the following for the x-specific effects.

$$\begin{aligned} x - DE_{x_0,x_1}(y|x_0,s) &= P(y_{x_1,w_{x_0}}|x_0,s) - P(y_{x_0}|x_0,s) \\ &= \Sigma_{z,w}P(y_{x_1,w}|x_0,w_{x_0},z)P(w_{x_0}|x_0,z,s)P(z|x_0) \\ &- P(y_{x_0,w}|x_0,w_{x_0},z,)P(w_{x_0}|x_0,z,s)P(z|x_0) \\ &= \Sigma_{z,w}[P(y|x_1,w,z) - P(y|x_0,w,z)]P^*(w|x_0,z)P(z|x_0) \end{aligned} \quad (64)$$

$$\begin{aligned} x - IE_{x_1,x_0}(y|x_0,s) &= P(y_{x_1,w_{x_0}}|x_0,s) - P(y_{x_1}|x_0,s) \\ &= \Sigma_{z,w}P(y|x_1,w,z)P^*(w|x_0,z)P(z|x_0) \\ &- P(y_{x_1,w}|x_0,w_{x_1},z)P(w_{x_1}|x_0,z,s)P(z|x_0) \\ &= \Sigma_{z,w}P(y|x_1,w,z)[P(w^*|x_0,z) - P(w^*|x_1,z)]P(z|x_0) \end{aligned} \quad (65)$$

$$\begin{aligned} x - SE_{x_1,x_0}(y|x_0,s) &= P(y_{x_1}|x_0,s) - P(y_{x_1}|x_1,s) \\ &= \Sigma_{z,w}P(y|x_1,w,z)P^*(w|x_1,z)P(z|x_0) \\ &- P(y_{x_1,w}|x_1,w_{x_1},z)P(w_{x_1}|x_1,z,s)P(z|x_1) \\ &= \Sigma_{z,w}P(y|x_1,w,z)P^*(w|x_1,z)[P(z|x_0) - P(z|x_1)] \end{aligned} \quad (66)$$

R is transportable when there is a shift in W.However, it requires to re-estimate $P^*(w|x,z)$ in the target environment.

4.1.3.3 (x,z)-specific level

Note, that in this case, the (x,z)- and z-specific level does not remain invariant anymore. Specifically, we will have the following trivially transportable expressions:

$$\begin{aligned} (x,z) - DE_{x_0,x_1}(y|x_0,z) &= P(y_{x_1,W_{x_0}}|x_0,z) - P(y_{x_0}|x_0,z) \\ &= \Sigma_w[P(y|x_1,w,z) - P(y|x_0,w,z)]P^*(w|x_0,z) \end{aligned} \quad (67)$$

$$(x,z)-IE_{x_1,x_0}(y|x_0,z)=\Sigma_w P(y|x_1,w,z)[P^*(w|x_0,z)-P^*(w|x_1,z)] \tag{68}$$

$$z-DE_{x_0,x_1}(y|z)=\Sigma_w[P(y|x_1,w,z)-P(y|x_0,w,z)]P^*(w|x_0,z) \tag{69}$$

$$z-IE_{x_1,x_0}(y|z)=\Sigma_w P(y|x_1,w,z)[P^*(w|x_0,z)-P^*(w|x_1,z)] \tag{70}$$

4.1.4 Shift in Y

4.1.4.1 population level

The population level effects are trivially transportable and are given by:

$$NDE_{x_0,x_1}(y|s)=\Sigma_{z,w}[P^*(y|x_1,w,z)-P^*(y|x_0,w,z)]P(w|x_0,z)P(z) \tag{71}$$

$$NIE_{x_1,x_0}(y|s)=\Sigma_{z,w}P^*(y|x_1,w,z)[P(w|x_0,z)-P(w|x_1,z)]P(z) \tag{72}$$

$$Exp-SE_x(y|s)=P^*(y|x,w,z)P(w|x,z)(P(z|x)-P(z)) \tag{73}$$

4.1.4.2 x-specific level

We can represent a shift in Y by using a selection variable S and augmenting G^{SFM} by adding S and an corresponding edge $S\to Y$. The resulting selection diagram is depicted in Fig.2e. Using the selection diagram we find that $Y \not\perp\!\!\!\perp S|$ and $(X,Z,W)\perp\!\!\!\perp S$. Using these (in)dependence relations we obtain the following for the x-specific effects.

$$\begin{aligned} x-DE_{x_0,x_1}(y|x_0,s) &= P(y_{x_1,w_{x_0}}|x_0,s)-P(y_{x_0}|x_0,s)\\ &= \Sigma_{z,w}P(y_{x_1,w}|x_0,w_{x_0},z,s)P(w_{x_0}|x_0,z)P(z|x_0)\\ &- P(y_{x_0,w}|x_0,w_{x_0},z,s)P(w_{x_0}|x_0,z)P(z|x_0)\\ &= \Sigma_{z,w}[P^*(y|x_1,w,z)-P^*(y|x_0,w,z)]P(w|x_0,z)P(z|x_0) \end{aligned} \tag{74}$$

$$\begin{aligned} x-IE_{x_1,x_0}(y|x_0,s) &= P(y_{x_1,w_{x_0}}|x_0,s)-P(y_{x_1}|x_0,s)\\ &= \Sigma_{z,w}P(y|x_1,w,z,s)P(w|x_0,z)P(z|x_0)\\ &- P(y_{x_1,w}|x_0,w_{x_1},z,s)P(w_{x_1}|x_0,z)P(z|x_0)\\ &= \Sigma_{z,w}P^*(y|x_1,w,z)[P(w|x_0,z)-P(w|x_1,z)]P(z|x_0) \end{aligned} \tag{75}$$

$$\begin{aligned} x-SE_{x_1,x_0}(y|x_0,s) &= P(y_{x_1}|x_0,s)-P(y_{x_1}|x_1,s)\\ &= \Sigma_{z,w}P(y|x_1,w,z,s)P(w|x_1,z)P(z|x_0)\\ &- P(y_{x_1,w}|x_1,w_{x_1},z,s)P(w_{x_1}|x_1,z)P(z|x_1)\\ &= \Sigma_{z,w}P^*(y|x_1,w,z)P(w|x_1,z)[P(z|x_0)-P(z|x_1)] \end{aligned} \tag{76}$$

R is transportable when there is a shift in Y.However, it requires to re-estimate $P^*(y|x,w,z)$ in the target environment.

4.1.4.3 (x,z)-specific level

Similarly as before, the effects become trivially transportable and given by:

$$\begin{aligned}(x,z) - DE_{x_0,x_1}(y|x_0,z) &= P^*(y_{x_1,W_{x_0}}|x_0,z) - P^*(y_{x_0}|x_0,z) \\ &= \Sigma_w[P^*(y|x_1,w,z) - P^*(y|x_0,w,z)]P(w|x_0,z)\end{aligned} \tag{77}$$

$$(x,z) - IE_{x_1,x_0}(y|x_0,z) = \Sigma_w P^*(y|x_1,w,z)[P(w|x_0,z) - P(w|x_1,z)] \tag{78}$$

$$z - DE_{x_0,x_1}(y|z) = \Sigma_w[P^*(y|x_1,w,z) - P^*(y|x_0,w,z)]P(w|x_0,z) \tag{79}$$

$$z - IE_{x_1,x_0}(y|z) = \Sigma_w P^*(y|x_1,w,z)[P(w|x_0,z) - P(w|x_1,z)] \tag{80}$$

4.2 Compound Shifts

Compound shifts describe the presence of distribution shifts in multiple variables. S now becomes a set of selection variables. The representation remains the same. We augment G^{SFM} by adding each member and their corresponding edge to obtain the selection diagram.

4.2.1 Shift in X,Z

4.2.1.1 population level

The effects are trivially transportable and are given by:

$$NDE_{x_0,x_1}(y|s) = \Sigma_{z,w}[P(y|x_1,w,z) - P(y|x_0,w,z)]P(w|x_0,z)P^*(z) \tag{81}$$

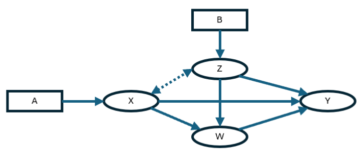
$$NIE_{x_1,x_0}(y|s) = \Sigma_{z,w}P(y|x_1,w,z)[P(w|x_0,z) - P(w|x_1,z)]P^*(z) \tag{82}$$

$$Exp - SE_x(y|s) = P(y|x,w,z)P(w|x,z)(P^*(z|x) - P^*(z)) \tag{83}$$

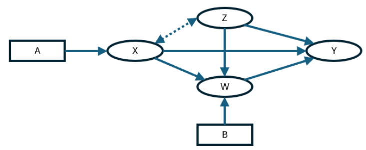
4.2.1.2 x-specific level

We can represent a compound shift in X and Z by the set of selection variables $S = \{A, B\}$. Then we can augment G^{SFM} including nodes A and B as well as the edges $A \to X$ and $B \to Z$. The resulting selection diagram is depicted in Fig.3a. From there we can obtain that $Y \perp\!\!\!\perp (A,B)|(Z,X)$, $W \perp\!\!\!\perp (A,B)|(X,Z)$ and $Z \not\perp\!\!\!\perp (A,B)|X$. Using these (in)dependence relations and previous results we find for the x-specific effects:

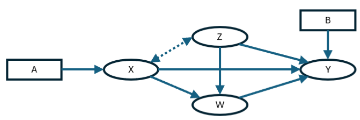
$$\begin{aligned}x - DE_{x_0,x_1}(y|x_0,a,b) &= P(y_{x_1,w_{x_0}}|x_0,a,b) - P(y_{x_0}|x_0,a,b) \\ &= \Sigma_{z,w}P(y_{x_1,w}|x_0,w_{x_0},z)P(w_{x_0}|x_0,z)P(z|x_0,a,b) \\ &- P(y_{x_0,w}|x_0,w_{x_0},z)P(w_{x_0}|x_0,z)P(z|x_0,a,b) \\ &= \Sigma_{z,w}[P(y|x_1,w,z) - P(y|x_0,w,z)]P(w|x_0,z)P^*(z|x_0)\end{aligned} \tag{84}$$



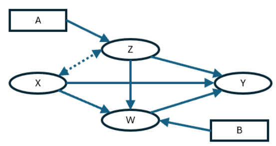
(a)



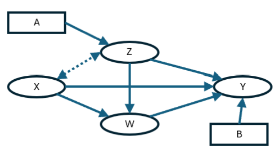
(b)



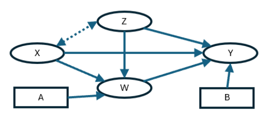
(c)



(d)



(e)



(f)

Figure 3: (a) G^{SFM} with a shift in X and Z, indicated by $S = \{A, B\}$ (b) G^{SFM} with a shift in X and W, indicated by $S = \{A, B\}$ (c) G^{SFM} with a shift in X and Y, indicated by $S = \{A, B\}$ (d) G^{SFM} with a shift in Z and W, indicated by $S = \{A, B\}$ (e) G^{SFM} with a shift in Z and Y, indicated by $S = \{A, B\}$ (f) G^{SFM} with a shift in W and Y, indicated by $S = \{A, B\}$

$$
\begin{aligned}
x-IE_{x_1,x_0}(y|x_0,a,b) &= P(y_{x_1,w_{x_0}}|x_0,a,b) - P(y_{x_1}|x_0,a,b) \\
&= \Sigma_{z,w} P(y|x_1,w,z)P(w|x_0,z)P^*(z|x_0) - P(y_{x_1}|x_0,a,b) \\
&= \Sigma_{z,w} P(y|x_1,w,z)[P(w|x_0,z) - P(w|x_1,z)]P^*(z|x_0)
\end{aligned}
\tag{85}
$$

$$
\begin{aligned}
x-SE_{x_1,x_0}(y|x_0,a,b) &= P(y_{x_1}|x_0,a,b) - P(y_{x_1}|x_1,a,b) \\
&= \Sigma_{z,w} P(y|x_1,w,z)P(w|x_1,z)P^*(z|x_0) \\
&- P(y_{x_1,w}|x_1,w_{x_1},z)P(w_{x_1}|x_1,z)P^*(z|x_1) \\
&= \Sigma_{z,w} P(y|x_1,w,z)P(w|x_1,z)[P^*(z|x_0) - P^*(z|x_1)]
\end{aligned}
\tag{86}
$$

R is transportable when there is a compound shift in X and Z.However, it requires to re-estimate $P^*(z|x)$ in the target environment.

4.2.1.3 (x,z)-specific level

As for the individual single shifts, the effects remain invariant at this level and given by:

$$
\begin{aligned}
(x,z)-DE_{x_0,x_1}(y|x_0,z) &= P(y_{x_1,W_{x_0}}|x_0,z) - P(y_{x_0}|x_0,z) \\
&= \Sigma_w [P(y|x_1,w,z) - P(y|x_0,w,z)]P(w|x_0,z)
\end{aligned}
\tag{87}
$$

$$
(x,z)-IE_{x_1,x_0}(y|x_0,z) = \Sigma_w P(y|x_1,w,z)[P(w|x_0,z) - P(w|x_1,z)] \tag{88}
$$

$$
z-DE_{x_0,x_1}(y|z) = \Sigma_w [P(y|x_1,w,z) - P(y|x_0,w,z)]P(w|x_0,z) \tag{89}
$$

$$
z-IE_{x_1,x_0}(y|z) = \Sigma_w P(y|x_1,w,z)[P(w|x_0,z) - P(w|x_1,z)] \tag{90}
$$

4.2.2 Shift in X,W

4.2.2.1 population level

The effects are trivially transportable and are given by:

$$
NDE_{x_0,x_1}(y|s) = \Sigma_{z,w} [P(y|x_1,w,z) - P(y|x_0,w,z)]P^*(w|x_0,z)P(z) \tag{91}
$$

$$
NIE_{x_1,x_0}(y|s) = \Sigma_{z,w} P(y|x_1,w,z)[P^*(w|x_0,z) - P^*(w|x_1,z)]P(z) \tag{92}
$$

$$
Exp-SE_x(y|s) = P(y|x,w,z)P^*(w|x,z)(P^*(z|x) - P(z)) \tag{93}
$$

4.2.2.2 x-specific level

We can represent a compound shift in X and W by the set of selection variables $S = \{A, B\}$. Then we can augment G^{SFM} including nodes A and B as well as the edges $A \to X$ and $B \to W$. The resulting selection diagram is depicted in Fig.3b. From there we can obtain that $Y \perp\!\!\!\perp (A,B)|(Z,X,W)$, $W \perp\!\!\!\perp A|(X,Z)$,

$W \not\perp\!\!\!\perp (B)$, $Z \not\perp\!\!\!\perp A|X$ and $Z \perp\!\!\!\perp B$. Using these (in)dependence relations and previous results we find for the x-specific effects:

$$\begin{aligned} x - DE_{x_0,x_1}(y|x_0,a,b) &= P(y_{x_1,w_{x_0}}|x_0,a,b) - P(y_{x_0}|x_0,a,b) \\ &= \Sigma_{z,w} P(y_{x_1,w}|x_0,w_{x_0},z)P(w_{x_0}|x_0,z,b)P(z|x_0,a) \\ &- P(y_{x_0,w}|x_0,w_{x_0},z)P(w_{x_0}|x_0,z,b)P(z|x_0,a) \\ &= \Sigma_{z,w}[P(y|x_1,z) - P(y|x_0,z)]P^*(w|x_0,z)P^*(z|x_0) \end{aligned} \tag{94}$$

$$\begin{aligned} x - IE_{x_1,x_0}(y|x_0,a,b) &= P(y_{x_1,w_{x_0}}|x_0,a,b) - P(y_{x_1}|x_0,a,b) \\ &= \Sigma_{z,w} P(y|x_1,w,z)P^*(w|x_0,z)P^*(z|x_0) \\ &- P(y|x_1,w,z)P^*(w|x_1,z)P^*(z|x_0) \\ &= \Sigma_{z,w} P(y|x_1,w,z)[P^*(w|x_0,z) - P^*(w|x_1,z)]P^*(z|x_0) \end{aligned} \tag{95}$$

$$\begin{aligned} x - SE_{x_1,x_0}(y|x_0,a,b) &= P(y_{x_1}|x_0,a,b) - P(y_{x_1}|x_1,a,b) \\ &= \Sigma_{z,w} P(y|x_1,w,z)P^*(w|x_1,z)P^*(z|x_0) \\ &- P(y|x_1,w,z)P^*(w|x_1,z)P^*(z|x_1) \\ &= \Sigma_{z,w} P(y|x_1,w,z)P^*(w|x_1,z)[P^*(z|x_0) - P^*(z|x_1)] \end{aligned} \tag{96}$$

R is transportable when there is a compound shift in X and W.However, it requires to re-estimate $P^*(z|x)$ and $P^*(w|x,z)$ in the target environment.

4.2.2.3 (x,z)-specific level

The effects become trivially transportable and are given by:

$$\begin{aligned} (x,z) - DE_{x_0,x_1}(y|x_0,z) &= P(y_{x_1,W_{x_0}}|x_0,z) - P(y_{x_0}|x_0,z) \\ &= \Sigma_w [P(y|x_1,w,z) - P(y|x_0,w,z)]P^*(w|x_0,z) \end{aligned} \tag{97}$$

$$(x,z) - IE_{x_1,x_0}(y|x_0,z) = \Sigma_w P(y|x_1,w,z)[P^*(w|x_0,z) - P^*(w|x_1,z)] \tag{98}$$

$$z - DE_{x_0,x_1}(y|z) = \Sigma_w [P(y|x_1,w,z) - P(y|x_0,w,z)]P^*(w|x_0,z) \tag{99}$$

$$z - IE_{x_1,x_0}(y|z) = \Sigma_w P(y|x_1,w,z)[P^*(w|x_0,z) - P^*(w|x_1,z)] \tag{100}$$

4.2.3 Shift in X,Y

4.2.3.1 population level

The effects are trivially transportable and are given by:

$$NDE_{x_0,x_1}(y|s) = \Sigma_{z,w}[P^*(y|x_1,w,z) - P^*(y|x_0,w,z)]P(w|x_0,z)P(z) \tag{101}$$

$$NIE_{x_1,x_0}(y|s) = \Sigma_{z,w} P^*(y|x_1,w,z)[P(w|x_0,z) - P(w|x_1,z)]P(z) \tag{102}$$

$$Exp - SE_x(y|s) = P^*(y|x,w,z)P(w|x,z)(P^*(z|x) - P(z)) \tag{103}$$

4.2.3.2 x-specific level

We can represent a compound shift in X and Y by the set of selection variables $S = \{A, B\}$. Then we can augment G^{SFM} including nodes A and B as well as the edges $A \rightarrow X$ and $B \rightarrow Y$. The resulting selection diagram is depicted in Fig.3c. From there we can obtain that $Y \perp\!\!\!\perp A|(Z, X)$, $Y \not\perp\!\!\!\perp B$, $W \perp\!\!\!\perp A|(X, Z)$, $W \perp\!\!\!\perp B$, $Z \not\perp\!\!\!\perp A|X$ and $Z \perp\!\!\!\perp B$. Using these (in)dependence relations and previous results we find for the x-specific effects:

$$\begin{aligned} x - DE_{x_0,x_1}(y|x_0, a, b) &= P(y_{x_1,w_{x_0}}|x_0, a, b) - P(y_{x_0}|x_0, a, b) \\ &= \Sigma_{z,w}[P^*(y|x_1, w, z) - P^*(y|x_0, w, z)]P(w|x_0, z)P^*(z|x_0) \end{aligned} \tag{104}$$

$$\begin{aligned} x - IE_{x_1,x_0}(y|x_0, a, b) &= P(y_{x_1,w_{x_0}}|x_0, a, b) - P(y_{x_1}|x_0, a, b) \\ &= \Sigma_{z,w}P^*(y|x_1, w, z)[P(w|x_0, z) - P(w|x_1, z)]P^*(z|x_0) \end{aligned} \tag{105}$$

$$\begin{aligned} x - SE_{x_1,x_0}(y|x_0, a, b) &= P(y_{x_1}|x_0, a, b) - P(y_{x_1}|x_1, a, b) \\ &= \Sigma_{z,w}P^*(y|x_1, w, z)P(w|x_1, z)[P^*(z|x_0) - P^*(z|x_1)] \end{aligned} \tag{106}$$

R is transportable when there is a compound shift in X and W.However, it requires to re-estimate $P^*(z|x)$ and $P^*(y|x, w, z)$ in the target environment.

4.2.3.3 (x, z)-specific level

The effects become trivially transportable and given by:

$$\begin{aligned} (x, z) - DE_{x_0,x_1}(y|x_0, z) &= P(y_{x_1,W_{x_0}}|x_0, z) - P(y_{x_0}|x_0, z) \\ &= \Sigma_w[P^*(y|x_1, w, z) - P^*(y|x_0, w, z)]P(w|x_0, z) \end{aligned} \tag{107}$$

$$(x, z) - IE_{x_1,x_0}(y|x_0, z) = \Sigma_w P^*(y|x_1, w, z)[P(w|x_0, z) - P(w|x_1, z)] \tag{108}$$

$$z - DE_{x_0,x_1}(y|z) = \Sigma_w[P^*(y|x_1, w, z) - P^*(y|x_0, w, z)]P(w|x_0, z) \tag{109}$$

$$z - IE_{x_1,x_0}(y|z) = \Sigma_w P^*(y|x_1, w, z)[P(w|x_0, z) - P(w|x_1, z)] \tag{110}$$

4.2.4 Shift in Z,W

4.2.4.1 population level

The effects are trivially transportable and are given by:

$$NDE_{x_0,x_1}(y|s) = \Sigma_{z,w}[P(y|x_1, w, z) - P(y|x_0, w, z)]P^*(w|x_0, z)P^*(z) \tag{111}$$

$$NIE_{x_1,x_0}(y|s) = \Sigma_{z,w}P(y|x_1, w, z)[P^*(w|x_0, z) - P^*(w|x_1, z)]P^*(z) \tag{112}$$

$$Exp - SE_x(y|s) = P(y|x, w, z)P^*(w|x, z)(P^*(z|x) - P^*(z)) \tag{113}$$

4.2.4.2 x-specific level

We can represent a compound shift in Z and W by the set of selection variables $S = \{A, B\}$. Then we can augment G^{SFM} including nodes A and B as well as the edges $A \to Z$ and $B \to W$. The resulting selection diagram is depicted in Fig.3d. From there we can obtain that $Y \perp\!\!\!\perp (A,B)|(Z,X,W)$, $W \perp\!\!\!\perp A|(X,Z)$, $W \not\perp\!\!\!\perp B$, $Z \not\perp\!\!\!\perp A|X$ and $Z \perp\!\!\!\perp B$. Using these (in)dependence relations and previous results we find for the x-specific effects:

$$\begin{aligned} x - DE_{x_0,x_1}(y|x_0,a,b) &= P(y_{x_1,w_{x_0}}|x_0,a,b) - P(y_{x_0}|x_0,a,b) \\ &= \Sigma_{z,w}[P(y|x_1,w,z) - P(y|x_0,w,z)]P^*(w|x_0,z)P^*(z|x_0) \end{aligned} \tag{114}$$

$$\begin{aligned} x - IE_{x_1,x_0}(y|x_0,a,b) &= P(y_{x_1,w_{x_0}}|x_0,a,b) - P(y_{x_1}|x_0,a,b) \\ &= \Sigma_{z,w}P(y|x_1,w,z)[P^*(w|x_0,z) - P^*(w|x_1,z)]P^*(z|x_0) \end{aligned} \tag{115}$$

$$\begin{aligned} x - SE_{x_1,x_0}(y|x_0,a,b) &= P(y_{x_1}|x_0,a,b) - P(y_{x_1}|x_1,a,b) \\ &= \Sigma_{z,w}P(y|x_1,w,z)P^*(w|x_1,z)[P^*(z|x_0) - P^*(z|x_1)] \end{aligned} \tag{116}$$

R is transportable when there is a compound shift in Z and W.However, it requires to re-estimate $P^*(z|x)$ and $P^*(w|x,z)$ in the target environment.

4.2.4.3 (x,z)-specific level

The effects are trivially transportable and given by:

$$\begin{aligned} (x,z) - DE_{x_0,x_1}(y|x_0,z) &= P(y_{x_1,W_{x_0}}|x_0,z) - P(y_{x_0}|x_0,z) \\ &= \Sigma_w[P(y|x_1,w,z) - P(y|x_0,w,z)]P^*(w|x_0,z) \end{aligned} \tag{117}$$

$$(x,z) - IE_{x_1,x_0}(y|x_0,z) = \Sigma_w P(y|x_1,w,z)[P^*(w|x_0,z) - P^*(w|x_1,z)] \tag{118}$$

$$z - DE_{x_0,x_1}(y|z) = \Sigma_w[P(y|x_1,w,z) - P(y|x_0,w,z)]P^*(w|x_0,z) \tag{119}$$

$$z - IE_{x_1,x_0}(y|z) = \Sigma_w P(y|x_1,w,z)[P^*(w|x_0,z) - P^*(w|x_1,z)] \tag{120}$$

4.2.5 Shift in Z,Y

4.2.5.1 population level

The effects are trivially transportable and are given by:

$$NDE_{x_0,x_1}(y|s) = \Sigma_{z,w}[P^*(y|x_1,w,z) - P^*(y|x_0,w,z)]P(w|x_0,z)P^*(z) \tag{121}$$

$$NIE_{x_1,x_0}(y|s) = \Sigma_{z,w}P^*(y|x_1,w,z)[P(w|x_0,z) - P(w|x_1,z)]P^*(z) \tag{122}$$

$$Exp - SE_x(y|s) = P^*(y|x,w,z)P(w|x,z)(P^*(z|x) - P^*(z)) \tag{123}$$

4.2.5.2 x-specific level

We can represent a compound shift in Z and Y by the set of selection variables $S = \{A, B\}$. Then we can augment G^{SFM} including nodes A and B as well as the edges $A \to Z$ and $B \to Y$. The resulting selection diagram is depicted in Fig.3e. From there we can obtain that $Y \perp\!\!\!\perp A|(Z, X)$, $Y \not\perp\!\!\!\perp B$, $W \perp\!\!\!\perp A|(X, Z)$, $W \perp\!\!\!\perp B$, $Z \not\perp\!\!\!\perp A$ and $Z \perp\!\!\!\perp B$. Using these (in)dependence relations and previous results we find for the x-specific effects:

$$\begin{aligned} x - DE_{x_0,x_1}(y|x_0, a, b) &= P(y_{x_1,w_{x_0}}|x_0, a, b) - P(y_{x_0}|x_0, a, b) \\ &= \Sigma_{z,w}[P^*(y|x_1, w, z) - P^*(y|x_0, w, z)]P(w|x_0, z)P^*(z|x_0) \end{aligned} \tag{124}$$

$$\begin{aligned} x - IE_{x_1,x_0}(y|x_0, a, b) &= P(y_{x_1,w_{x_0}}|x_0, a, b) - P(y_{x_1}|x_0, a, b) \\ &= \Sigma_{z,w}P^*(y|x_1, w, z)[P(w|x_0.z) - P(w|x_1, z)]P^*(z|x_0) \end{aligned} \tag{125}$$

$$\begin{aligned} x - SE_{x_1,x_0}(y|x_0, a, b) &= P(y_{x_1}|x_0, a, b) - P(y_{x_1}|x_1, a, b) \\ &= \Sigma_{z,w}P^*(y|x_1, w, z)P(w|x_1, z)[P^*(z|x_0) - P^*(z|x_1)] \end{aligned} \tag{126}$$

R is transportable when there is a compound shift in Z and W.However, it requires to re-estimate $P^*(z|x)$ and $P^*(y|x, w, z)$ in the target environment.

4.2.5.3 (x,z)-specific level

The shifts become trivially transportable and are given by:

$$\begin{aligned} (x, z) - DE_{x_0,x_1}(y|x_0, z) &= P(y_{x_1,W_{x_0}}|x_0, z) - P(y_{x_0}|x_0, z) \\ &= \Sigma_w[P^*(y|x_1, w, z) - P^*(y|x_0, w, z)]P(w|x_0, z) \end{aligned} \tag{127}$$

$$(x, z) - IE_{x_1,x_0}(y|x_0, z) = \Sigma_w P^*(y|x_1, w, z)[P(w|x_0, z) - P(w|x_1, z)] \tag{128}$$

$$z - DE_{x_0,x_1}(y|z) = \Sigma_w[P^*(y|x_1, w, z) - P^*(y|x_0, w, z)]P(w|x_0, z) \tag{129}$$

$$z - IE_{x_1,x_0}(y|z) = \Sigma_w P^*(y|x_1, w, z)[P(w|x_0, z) - P(w|x_1, z)] \tag{130}$$

4.2.6 Shift in W,Y

4.2.6.1 population level

The effects are trivially transportable and are given by:

$$NDE_{x_0,x_1}(y|s) = \Sigma_{z,w}[P^*(y|x_1, w, z) - P^*(y|x_0, w, z)]P^*(w|x_0, z)P(z) \tag{131}$$

$$NIE_{x_1,x_0}(y|s) = \Sigma_{z,w}P^*(y|x_1, w, z)[P^*(w|x_0, z) - P^*(w|x_1, z)]P(z) \tag{132}$$

$$Exp - SE_x(y|s) = P^*(y|x, w, z)P^*(w|x, z)(P(z|x) - P(z)) \tag{133}$$

4.2.6.2 x-specific level

We can represent a compound shift in W and Y by the set of selection variables $S = \{A, B\}$. Then we can augment G^{SFM} including nodes A and B as well as the edges $A \rightarrow W$ and $B \rightarrow Y$. The resulting selection diagram is depicted in Fig.3f. From there we can obtain that $Y \perp\!\!\!\perp A|(X, Z, W)$, $Y \not\perp\!\!\!\perp B$, $W \not\perp\!\!\!\perp A$, $W \perp\!\!\!\perp B$, and $Z \perp\!\!\!\perp (A, B)$. Using these (in)dependence relations and previous results we find for the x-specific effects:

$$\begin{aligned} x - DE_{x_0,x_1}(y|x_0, a, b) &= P(y_{x_1,w_{x_0}}|x_0, a, b) - P(y_{x_0}|x_0, a, b) \\ &= \Sigma_{z,w}[P^*(y|x_1, w, z) - P^*(y|x_0, w, z)]P^*(w|x_0, z)P(z|x_0) \end{aligned} \tag{134}$$

$$\begin{aligned} x - IE_{x_1,x_0}(y|x_0, a, b) &= P(y_{x_1,w_{x_0}}|x_0, a, b) - P(y_{x_1}|x_0, a, b) \\ &= \Sigma_{z,w}P^*(y|x_1, w, z)[P^*(w|x_0, z) - P^*(w|x_1, z)]P(z|x_0) \end{aligned} \tag{135}$$

$$\begin{aligned} x - SE_{x_1,x_0}(y|x_0, a, b) &= P(y_{x_1}|x_0, a, b) - P(y_{x_1}|x_1, a, b) \\ &= \Sigma_{z,w}P^*(y|x_1, w, z)P^*(w|x_1, z)[P(z|x_0) - P(z|x_1)] \end{aligned} \tag{136}$$

R is transportable when there is a compound shift in W and Y.However, it requires to re-estimate $P^*(w|x, z)$ and $P^*(y|x, w, z)$ in the target environment.

4.2.6.3 (x,z)-specific level

The effects are trivially transportable and are given by:

$$\begin{aligned} (x, z) - DE_{x_0,x_1}(y|x_0, z) &= P(y_{x_1,W_{x_0}}|x_0, z) - P(y_{x_0}|x_0, z) \\ &= \Sigma_w[P^*(y|x_1, w, z) - P^*(y|x_0, w, z)]P^*(w|x_0, z) \end{aligned} \tag{137}$$

$$(x, z) - IE_{x_1,x_0}(y|x_0, z) = \Sigma_w P^*(y|x_1, w, z)[P^*(w|x_0, z) - P^*(w|x_1, z)] \tag{138}$$

$$z - DE_{x_0,x_1}(y|z) = \Sigma_w[P^*(y|x_1, w, z) - P^*(y|x_0, w, z)]P^*(w|x_0, z) \tag{139}$$

$$z - IE_{x_1,x_0}(y|z) = \Sigma_w P^*(y|x_1, w, z)[P^*(w|x_0, z) - P^*(w|x_1, z)] \tag{140}$$

4.2.7 Shift in X,Z,W

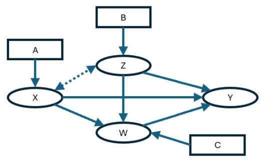
4.2.7.1 population level

The effects are trivially transportable and are given by:

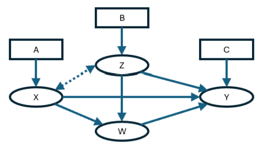
$$NDE_{x_0,x_1}(y|s) = \Sigma_{z,w}[P(y|x_1, w, z) - P(y|x_0, w, z)]P^*(w|x_0, z)P^*(z) \tag{141}$$

$$NIE_{x_1,x_0}(y|s) = \Sigma_{z,w}P(y|x_1, w, z)[P^*(w|x_0, z) - P^*(w|x_1, z)]P^*(z) \tag{142}$$

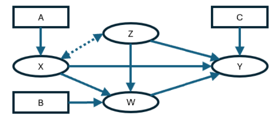
$$Exp - SE_x(y|s) = P(y|x, w, z)P^*(w|x, z)(P^*(z|x) - P^*(z)) \tag{143}$$



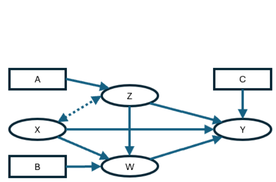
(a)



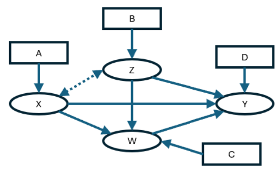
(b)



(c)



(d)



(e)

Figure 4: (a) G^{SFM} with a shift in X, Z and W indicated by $S = \{A, B, C\}$ (b) G^{SFM} with a shift in X, Z and Y indicated by $S = \{A, B, C\}$ (c) G^{SFM} with a shift in X, W and W indicated by $S = \{A, B, C\}$ (d) G^{SFM} with a shift in X, W and Y indicated by $S = \{A, B, C\}$ (e) G^{SFM} with a shift in X, Z, W and Y indicated by $S = \{A, B, C, D\}$

4.2.7.2 x-specific level

We can represent a compound shift in X, Z and W by the set of selection variables $S = \{A, B, C\}$. Then we can augment G^{SFM} including nodes A, B and C as well as the edges $A \to X$, $B \to Z$ and $C \to W$. The resulting selection diagram is depicted in Fig.4a. From there we can obtain that
$Y \perp\!\!\!\perp (A, B, C)|(X, Z, W)$, $W \not\!\perp\!\!\!\perp C$, $W \perp\!\!\!\perp (A, B)|(X, Z)$, $Z \perp\!\!\!\perp C$ and $Z \not\!\perp\!\!\!\perp (A, B)|X$. Using these (in)dependence relations and previous results we find for the x-specific effects:

$$\begin{aligned} x - DE_{x_0,x_1}(y|x_0, a, b, c) &= P(y_{x_1,w_{x_0}}|x_0, a, b, c) - P(y_{x_0}|x_0, a, b, c) \\ &= \Sigma_{z,w}[P(y|x_1, w, z) - P(y|x_0, w, z)]P^*(w|x_0, z)P^*(z|x_0) \end{aligned} \tag{144}$$

$$\begin{aligned} x - IE_{x_1,x_0}(y|x_0, a, b, c) &= P(y_{x_1,w_{x_0}}|x_0, a, b, c) - P(y_{x_1}|x_0, a, b, c) \\ &= \Sigma_{z,w}P(y|x_1, w, z)[P^*(w|x_0, z) - P^*(w|x_1, z)]P^*(z|x_0) \end{aligned} \tag{145}$$

$$\begin{aligned} x - SE_{x_1,x_0}(y|x_0, a, b, c) &= P(y_{x_1}|x_0, a, b, c) - P(y_{x_1}|x_1, a, b, c) \\ &= \Sigma_{z,w}P(y|x_1, w, z)P^*(w|x_1, z)[P^*(z|x_0) - P^*(z|x_1)] \end{aligned} \tag{146}$$

R is transportable when there is a compound shift in X, Z and W. However, it requires to re-estimate $P^*(z|x)$ and $P^*(w|x, z)$ in the target environment.

4.2.7.3 (x,z)-specific level

The effects are trivially transportable and given by:

$$\begin{aligned}(x,z) - DE_{x_0,x_1}(y|x_0,z) &= P(y_{x_1,W_{x_0}}|x_0,z) - P(y_{x_0}|x_0,z)\\ &= \Sigma_w[P^*(y|x_1,w,z) - P^*(y|x_0,w,z)]P^*(w|x_0,z)\end{aligned} \tag{147}$$

$$(x,z) - IE_{x_1,x_0}(y|x_0,z) = \Sigma_w P^*(y|x_1,w,z)[P^*(w|x_0,z) - P^*(w|x_1,z)] \tag{148}$$

$$z - DE_{x_0,x_1}(y|z) = \Sigma_w[P^*(y|x_1,w,z) - P^*(y|x_0,w,z)]P^*(w|x_0,z) \tag{149}$$

$$z - IE_{x_1,x_0}(y|z) = \Sigma_w P^*(y|x_1,w,z)[P^*(w|x_0,z) - P^*(w|x_1,z)] \tag{150}$$

4.2.8 Shift in X,Z,Y

4.2.8.1 population level

$$NDE_{x_0,x_1}(y|s) = \Sigma_{z,w}[P^*(y|x_1,w,z) - P^*(y|x_0,w,z)]P^*(w|x_0,z)P^*(z) \tag{151}$$

$$NIE_{x_1,x_0}(y|s) = \Sigma_{z,w}P^*(y|x_1,w,z)[P^*(w|x_0,z) - P^*(w|x_1,z)]P^*(z) \tag{152}$$

$$Exp - SE_x(y|s) = P^*(y|x,w,z)P^*(w|x,z)(P^*(z|x) - P^*(z)) \tag{153}$$

4.2.8.2 x-specific level

We can represent a compound shift in X, Z and Y by the set of selection variables $S = \{A,B,C\}$. Then we can augment G^{SFM} including nodes A, B and C as well as the edges $A \to X$, $B \to Z$ and $C \to Y$. The resulting selection diagram is depicted in Fig.4b. From there we can obtain that
$Y \perp\!\!\!\perp (A,B)|(X,Z)$, $Y \not\perp\!\!\!\perp C$, $W \perp\!\!\!\perp (A,B,C)|(X,Z)$, $Z \perp\!\!\!\perp C$ and $Z \not\perp\!\!\!\perp (A,B)|X$. Using these (in)dependence relations and previous results we find for the x-specific effects:

$$\begin{aligned}x - DE_{x_0,x_1}(y|x_0,a,b,c) &= P(y_{x_1,w_{x_0}}|x_0,a,b,c) - P(y_{x_0}|x_0,a,b,c)\\ &= \Sigma_{z,w}[P^*(y|x_1,w,z) - P^*(y|x_0,w,z)]P(w|x_0,z)P^*(z|x_0)\end{aligned} \tag{154}$$

$$\begin{aligned}x - IE_{x_1,x_0}(y|x_0,a,b,c) &= P(y_{x_1,w_{x_0}}|x_0,a,b,c) - P(y_{x_1}|x_0,a,b,c)\\ &= \Sigma_{z,w}P^*(y|x_1,w,z)[P(w|x_0,z) - P(w|x_1,z)]P^*(z|x_0)\end{aligned} \tag{155}$$

$$\begin{aligned}x - SE_{x_1,x_0}(y|x_0,a,b,c) &= P(y_{x_1}|x_0,a,b,c) - P(y_{x_1}|x_1,a,b,c)\\ &= \Sigma_{z,w}P^*(y|x_1,w,z)P(w|x_1,z)[P^*(z|x_0) - P^*(z|x_1)]\end{aligned} \tag{156}$$

R is transportable when there is a compound shift in W and Y.However, it requires to re-estimate $P^*(z|x)$ and $P^*(y|x,w,z)$ in the target environment.

4.2.8.3 (x,z)-specific level

The shifts become trivially transportable and are given by:

$$\begin{aligned}(x,z)-DE_{x_0,x_1}(y|x_0,z) &= P(y_{x_1,W_{x_0}}|x_0,z)-P(y_{x_0}|x_0,z)\\ &= \Sigma_w[P^*(y|x_1,w,z)-P^*(y|x_0,w,z)]P(w|x_0,z)\end{aligned} \tag{157}$$

$$(x,z)-IE_{x_1,x_0}(y|x_0,z)=\Sigma_w P^*(y|x_1,w,z)[P(w|x_0,z)-P(w|x_1,z)] \tag{158}$$

$$z-DE_{x_0,x_1}(y|z)=\Sigma_w[P^*(y|x_1,w,z)-P^*(y|x_0,w,z)]P(w|x_0,z) \tag{159}$$

$$z-IE_{x_1,x_0}(y|z)=\Sigma_w P^*(y|x_1,w,z)[P(w|x_0,z)-P(w|x_1,z)] \tag{160}$$

4.2.9 Shift in X,W,Y

4.2.9.1 population level

$$NDE_{x_0,x_1}(y|s)=\Sigma_{z,w}[P^*(y|x_1,w,z)-P^*(y|x_0,w,z)]P^*(w|x_0,z)P(z) \tag{161}$$

$$NIE_{x_1,x_0}(y|s)=\Sigma_{z,w}P^*(y|x_1,w,z)[P^*(w|x_0,z)-P^*(w|x_1,z)]P(z) \tag{162}$$

$$Exp-SE_x(y|s)=P^*(y|x,w,z)P^*(w|x,z)(P^*(z|x)-P(z)) \tag{163}$$

4.2.9.2 x-specific level

We can represent a compound shift in X, W and Y by the set of selection variables $S=\{A,B,C\}$. Then we can augment G^{SFM} including nodes A, B and C as well as the edges $A\to X$, $B\to W$ and $C\to Y$. The resulting selection diagram is depicted in Fig.4c. From there we can obtain that
$Y \perp\!\!\!\perp (A,B)|(X,Z,W)$, $Y \not\perp\!\!\!\perp C$, $W \not\perp\!\!\!\perp B$, $W \perp\!\!\!\perp (A,C)|(X,Z)$, $Z \perp\!\!\!\perp (B,C)$ and $Z \not\perp\!\!\!\perp A|X$. Using these (in)dependence relations and previous results we find for the x-specific effects:

$$\begin{aligned}x-DE_{x_0,x_1}(y|x_0,a,b,c) &= P(y_{x_1,w_{x_0}}|x_0,a,b,c)-P(y_{x_0}|x_0,a,b,c)\\ &= \Sigma_{z,w}[P^*(y|x_1,w,z)-P^*(y|x_0,w,z)]P^*(w|x_0,z)P^*(z|x_0)\end{aligned} \tag{164}$$

$$\begin{aligned}x-IE_{x_1,x_0}(y|x_0,a,b,c) &= P(y_{x_1,w_{x_0}}|x_0,a,b,c)-P(y_{x_1}|x_0,a,b,c)\\ &= \Sigma_{z,w}P^*(y|x_1,w,z)[P^*(w|x_0,z)-P^*(w|x_1,z)]P^*(z|x_0)\end{aligned} \tag{165}$$

$$\begin{aligned}x-SE_{x_1,x_0}(y|x_0,a,b,c) &= P^*(y_{x_1}|x_0,a,b,c)-P^*(y_{x_1}|x_1,a,b,c)\\ &= \Sigma_{z,w}P^*(y|x_1,w,z)P^*(w|x_1,z)[P^*(z|x_0)-P^*(z|x_1)]\end{aligned} \tag{166}$$

R is transportable when there is a compound shift in W and Y.However, it requires to re-estimate $P^*(z|x)$, $P^*(w|x,z)$ and $P^*(y|x,w,z)$ in the target environment.

4.2.9.3 (x,z)-specific level

The effects are trivially transportable and are given by:

$$\begin{aligned}(x,z)-DE_{x_0,x_1}(y|x_0,z) &= P(y_{x_1,W_{x_0}}|x_0,z)-P(y_{x_0}|x_0,z)\\ &= \Sigma_w[P^*(y|x_1,w,z)-P^*(y|x_0,w,z)]P^*(w|x_0,z)\end{aligned} \tag{167}$$

$$(x,z)-IE_{x_1,x_0}(y|x_0,z)=\Sigma_w P^*(y|x_1,w,z)[P^*(w|x_0,z)-P^*(w|x_1,z)] \tag{168}$$

$$z-DE_{x_0,x_1}(y|z)=\Sigma_w[P^*(y|x_1,w,z)-P^*(y|x_0,w,z)]P^*(w|x_0,z) \tag{169}$$

$$z-IE_{x_1,x_0}(y|z)=\Sigma_w P^*(y|x_1,w,z)[P^*(w|x_0,z)-P^*(w|x_1,z)] \tag{170}$$

4.2.10 Shift in Z,W,Y

4.2.10.1 population level

The effects are trivially transportable and are given by:

$$NDE_{x_0,x_1}(y|s)=\Sigma_{z,w}[P^*(y|x_1,w,z)-P^*(y|x_0,w,z)]P^*(w|x_0,z)P^*(z) \tag{171}$$

$$NIE_{x_1,x_0}(y|s)=\Sigma_{z,w}P^*(y|x_1,w,z)[P^*(w|x_0,z)-P^*(w|x_1,z)]P^*(z) \tag{172}$$

$$Exp-SE_x(y|s)=P^*(y|x,w,z)P^*(w|x,z)(P^*(z|x)-P^*(z)) \tag{173}$$

4.2.10.2 x-specific level

We can represent a compound shift in Z, W and Y by the set of selection variables $S=\{A,B,C\}$. Then we can augment G^{SFM} including nodes A, B and C as well as the edges $A\to Z$, $B\to W$ and $C\to Y$. The resulting selection diagram is depicted in Fig.4d. From there we can obtain that $Y \perp\!\!\!\perp (A,B)|(X,Z,W)$, $Y \not\perp\!\!\!\perp C$, $W \not\perp\!\!\!\perp B$, $W \perp\!\!\!\perp (A,C)|(X,Z)$, $Z \perp\!\!\!\perp (B,C)$ and $Z \not\perp\!\!\!\perp A$. Using these (in)dependence relations and previous results we find for the x-specific effects:

$$\begin{aligned}x-DE_{x_0,x_1}(y|x_0,a,b,c) &= P(y_{x_1,w_{x_0}}|x_0,a,b,c)-P(y_{x_0}|x_0,a,b,c)\\ &= \Sigma_{z,w}[P^*(y|x_1,w,z)-P^*(y|x_0,w,z)]P^*(w|x_0,z)P^*(z|x_0)\end{aligned} \tag{174}$$

$$\begin{aligned}x-IE_{x_1,x_0}(y|x_0,a,b,c) &= P(y_{x_1,w_{x_0}}|x_0,a,b,c)-P(y_{x_1}|x_0,a,b,c)\\ &= \Sigma_{z,w}P^*(y|x_1,w,z)[P^*(w|x_0,z)-P^*(w|x_1,z)]P^*(z|x_0)\end{aligned} \tag{175}$$

$$\begin{aligned}x-SE_{x_1,x_0}(y|x_0,a,b,c) &= P^*(y_{x_1}|x_0,a,b,c)-P^*(y_{x_1}|x_1,a,b,c)\\ &= \Sigma_{z,w}P^*(y|x_1,w,z)P^*(w|x_1,z)[P^*(z|x_0)-P^*(z|x_1)]\end{aligned} \tag{176}$$

4.2.10.3 (x,z)-specific level

The effects are trivially transportable and given by:

$$\begin{aligned}(x,z)-DE_{x_0,x_1}(y|x_0,z) &= P(y_{x_1,W_{x_0}}|x_0,z) - P(y_{x_0}|x_0,z) \\ &= \Sigma_w[P^*(y|x_1,w,z) - P^*(y|x_0,w,z)]P^*(w|x_0,z)\end{aligned} \tag{177}$$

$$(x,z)-IE_{x_1,x_0}(y|x_0,z) = \Sigma_w P^*(y|x_1,w,z)[P^*(w|x_0,z) - P^*(w|x_1,z)] \tag{178}$$

$$z-DE_{x_0,x_1}(y|z) = \Sigma_w[P^*(y|x_1,w,z) - P^*(y|x_0,w,z)]P^*(w|x_0,z) \tag{179}$$

$$z-IE_{x_1,x_0}(y|z) = \Sigma_w P^*(y|x_1,w,z)[P^*(w|x_0,z) - P^*(w|x_1,z)] \tag{180}$$

4.2.11 Shift in X,Z,W,Y

4.2.11.1 population level

The effects are trivially transportable and are given by:

$$NDE_{x_0,x_1}(y|s) = \Sigma_{z,w}[P^*(y|x_1,w,z) - P^*(y|x_0,w,z)]P^*(w|x_0,z)P^*(z) \tag{181}$$

$$NIE_{x_1,x_0}(y|s) = \Sigma_{z,w}P^*(y|x_1,w,z)[P^*(w|x_0,z) - P^*(w|x_1,z)]P^*(z) \tag{182}$$

$$Exp-SE_x(y|s) = P^*(y|x,w,z)P^*(w|x,z)(P^*(z|x) - P^*(z)) \tag{183}$$

4.2.11.2 x-specific level

A special case of a compound shift occurs when there is a distribution shift in all endogenous variables. We can represent it with the set of selection variables $S = \{A,B,C,D\}$. Then we can augment G^{SFM} including nodes and corresponding edges of the selection variables. The resulting selection diagram is depicted in Fig.4e. It is trivial to see that with a shift in all endogenous variables we must have the following for the x-specific effects:

$$x-DE_{x_0,x_1}(y|x_0,a,b,c) = \Sigma_{z,w}[P^*(y|x_1,w,z) - P^*(y|x_0,w,z)]P^*(w|x_0,z)P^*(z|x_0)$$

$$x-IE_{x_1,x_0}(y|x_0,a,b,c) = \Sigma_{z,w}P^*(y|x_1,w,z)[P^*(w|x_0,z) - P^*(w|x_1,z)]P^*(z|x_0)$$

$$x-SE_{x_1,x_0}(y|x_0,a,b,c) = \Sigma_{z,w}P^*(y|x_1,w,z)P^*(w|x_1,z)[P^*(z|x_0) - P^*(z|x_1)]$$

R is transportable when there is a compound shift in all endogenous variables. However, it requires to re-estimate $P^*(z|x)$, $P^*(w|x,z)$ and $P^*(y|x,w,z)$ in the target environment.

4.2.11.3 (x,z)-specific level

The effects are trivially transportable and given by:

$$(x,z) - DE_{x_0,x_1}(y|x_0,z) = P(y_{x_1,W_{x_0}}|x_0,z) - P(y_{x_0}|x_0,z) \\ = \Sigma_w[P^*(y|x_1,w,z) - P^*(y|x_0,w,z)]P^*(w|x_0,z) \tag{184}$$

$$(x,z) - IE_{x_1,x_0}(y|x_0,z) = \Sigma_w P^*(y|x_1,w,z)[P^*(w|x_0,z) - P^*(w|x_1,z)] \tag{185}$$

$$z - DE_{x_0,x_1}(y|z) = \Sigma_w[P^*(y|x_1,w,z) - P^*(y|x_0,w,z)]P^*(w|x_0,z) \tag{186}$$

$$z - IE_{x_1,x_0}(y|z) = \Sigma_w P^*(y|x_1,w,z)[P^*(w|x_0,z) - P^*(w|x_1,z)] \tag{187}$$

5 Transportability of x-specific effects

The previous two sections derived an observable expression for the x-specific effects and using the triplet $R = [x - DE_{x_0,x_1}(y|x_0), x - IE_{x_1,x_0}(y|x_0), x - SE_{x_1,x_0}(y)]$ as the relation of interest, investigated how it changes for all possible shifts within the SFM structure. Now, we will summarize what it implies for transportability between environments. For this we will first make a few assumptions.

- We have environments Π and Π^* which both have the same endogenous variables V and obey to the structure of the SFM.
- Π and Π^* are characterized by their observational distributions $P(V)$ and $P^*(V)$ and G^{SFM}.
- Access to $P(V)$ is necessary as Π is the source domain where we initially compute R.

Now, for different access to $P^*(V)$ we will analyse the requirements for transportability of R.

Full access to $P^*(V)$:
In section 3 we showed that the x-specific effects can be fully computed from the observational distribution. As we have access to $P^*(V)$ and know Π^* adheres to the SFM, we can simply compute the x-specific effect in the new environment. We say R is trivially transportable for all possible shifts.

Access to only any single marginal observational distribution:
Within the SFM and for the x-specific effects we work with conditional probabilities. As they involve at least 2 variables, there is no shift within the SFM which we could re-estimate with only the observational distribution of a single variable.

Access to $P^*(X,Z)$ or $P^*(X,Z,Y)$:

In this case we can model single shifts in X and Z, as well as a compound shift in X and Z. For these $P^*(z|x)$ needs to be re-estimated in the target environment

Access to $P^*(X,Y)$, $P^*(X,W)$, $P^*(W,Z)$, $P^*(W,Y)$,$P^*(Z,Y)$, $P^*(W,Z,Y)$ or $P^*(X,W,Y)$:
There is no shift within the SFM which we could model having only access to one of the above observational distributions.

Access to $P^*(X,Z,W)$:
In this case we can model single shifts in either X,Z and W, as well as compound shifts in all of them and any of their combinations of length 2.

5.1 Implications for (x,z)-specific effects

Most results for the x-specific effect will also hold with respect to the (x,z)-specific effects there we further filter the population for a specific z. However, shifts in the protected attribute X and Z should become directly transportable. If these individual shifts are directly transportable, so should be their compound shift (X,Z). The reasoning behind this can be explained in graphical terms. Looking at fig 3a, we see that there is a bi-directed edge between X and Z. If we were now to condition on $X=x$, we would activate a backdoor-path from A into Z. For the (x,z) specific effects we additionally condition on $Z=z$. This blocks the back-door path making Z invariant to shifts in X.

References

[1] Elias Bareinboim and Judea Pearl. "Transportability of causal effects: Completeness results". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 26. 1. 2012, pp. 698–704.

[2] Judea Pearl. *Causality*. Cambridge university press, 2009.

[3] Judea Pearl and Elias Bareinboim. "Transportability of causal and statistical relations: A formal approach". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 25. 1. 2011, pp. 247–254.

[4] Drago Plecko and Elias Bareinboim. "Causal fairness analysis". In: *arXiv preprint arXiv:2207.11385* (2022).