SCM and meta-analysis

Combining evidence from heterogeneous observational and experimental environments is a hallmark of science — external validity. Drawing causal inference from such evidence is referred to as SCM meta-analysis or SCM data-fusion. In this note, we illustrate a formal approach to this exercise.[1]

Combining evidence from heterogeneous studies is nonparametrically achievable whenever permitted by the rules of do-calculus (Pearl, 1995) and where only passive observation (no *do*-operator) of the target domain is employed. The rules of do-calculus are below.

do-calculus

Let G be the DAG associated with a causal model and let $\Pr()$ be the probability distribution induced by the model. For any dis-joint set of variables X, Y, Z , and W the following rules apply.

Rule 1 (insertion/deletion of observations):
$\Pr(y \mid do(x), z, w) = \Pr(y \mid do(x), w)$ if $(Y \perp Z \mid X, W)_{G_{\overline{X}}}$ where $\perp$ refers to stochastic independence or d-separation in the graph.

Rule 2 (action/observation exchange):
$\Pr(y \mid do(x), z, w) = \Pr(y \mid do(x), do(z), w)$ if $(Y \perp Z \mid X, W)\ _{G_{\overline{X}\underline{Z}}}$.

Rule 3 (insertion/deletion of actions):
$\Pr(y \mid do(x), w) = \Pr(y \mid do(x), do(z), w)$ if $(Y \perp Z \mid X, W)_{G_{\overline{X}\,\overline{Z(W)}}}$
where $Z(W)$ is the set of Z -nodes that are not ancestors of any W -nodes in $G_{\overline{X}}$.

Direct transportability

Evidence can be transported from one domain to another provider there exists a common causal structure and the rules of do-calculus permit rescaling for differences in populations between the experimental and target domains. Graphically, this means the DAGs are similar except they contain S -nodes pointing towards variables whose mechanisms or distributions are suspected to differ between the domains. That is, for $S = 0$ no differences are suspected but for $S = 1$ differences between the domains are suspected regarding the variable(s) pointed to by the S -nodes.

Figure 1 illustrates two settings. The first in which direct transportability is satisfied and the second in which transportability is infeasible.

[1]This note draws on Bareinboim and Pearl, 2013, "Meta-transportability of causal effects: A formal approach," *Proceedings of the 16th International Conference on Artificial Intelligence and Statistics*, Bareinboim and Pearl, 2012, "Transportability of causal effects: Completeness results," *Proceedings of the 26th AAAI Conference on Artificial Intelligence, and* Bareinboim and Pearl, 2016, "Causal inference and the data-fusion problem," *Proceedings of the National Academy of Science.*

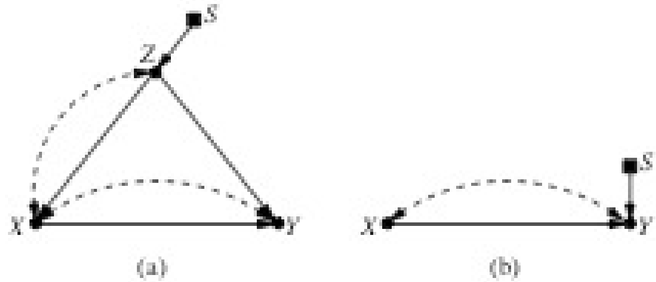

Figure 1: (a) Selection diagram illustrating when transportability among two domains is trivially solved through a simple recalibration. (b) Smallest selection diagram in which a causal relation is not transportable.

A graphical test for direct transportability is $(Y \perp S \mid X, Z)_{G_{\overline{X}}}$. That is, do-calculus transportability is satisfied if the transport formula can be written such that no *do*-operators are present in terms involving selection, $S = 1$. In figure 1(a), the subgraph $G_{\overline{X}}$ leaves the only path from S to Y the path through Z ; hence, direct transportability is satisfied. The transport formula is

$$P^* \left(y \mid do\left(x\right)\right) = \sum_z P^* \left(y \mid do\left(x\right), z\right) P^* \left(z \mid do\left(x\right)\right)$$

$$= \sum_z P^* \left(y \mid do\left(x\right), z\right) P^* \left(z\right)$$

$$= \sum_z P \left(y \mid do\left(x\right), z, S = 1\right) P \left(z \mid S = 1\right)$$

$$P^* \left(y \mid do\left(x\right)\right) = \sum_z P \left(y \mid do\left(x\right), z\right) P^* \left(z\right)$$

The first line indicates the target domain causal effect can be expanded to include Z by Bayes chain rule. The second line employs do-calculus rule 3 to delete $do\left(x\right)$ from the second term. The third line expresses the same quantity for the experimental domain where $S = 1$ indicates selection variables identifying differences between the domains. The last line is the transport formula. Since S is independent of Y conditional on $do\left(x\right)$ and z , the first term is transported from the experimental domain and the second term (drawn from the target domain) rescales or recalibrates. For instance, suppose we have two settings which differ only by firm size, say Z , the transport formula indicates manipulation to recover $P\left(y \mid do\left(x\right), z\right)$ in one experiment can be used infer the causal effect in the target setting by simply reweighting by $P^*\left(z\right)$, the target probability distribution for size.

Figure 1(b) illustrates a setting in which direct transportability is infeasible, in fact, the smallest selection diagram exhibiting non-transportability. The example below is based on BP12's proof.

Suppose X, U and Y in figure 1(b) are binary variables and there are two models representing the two domains, M_1 and M_2 . M_1 is defined by $X_1 = U + U_{x_1}$ and $P_1(U) = 1/2$ while M_2 is defined similarly, $X_2 = U + U_{x_2}$ and $P_2(U) = 1/2$ except outcome, Y_1 and Y_2 , is generated by different mechanisms in the two models (as indicated below). The selection DAG implies

$$P_1(X \mid S) = P_2(X \mid S), S = \{0, 1\}$$

$$P_1(Y \mid X, S) = P_2(Y \mid X, S), S = \{0, 1\}$$

$$P_1(Y \mid do(X), S = 0) = P_2(Y \mid do(X), S = 0)$$

Outcome is generated as follows.

X	S	U	Y_1	Y_2	$P_1(Y, X, S, U)$	$P_2(Y, X, S, U)$
0	0	0	0	0	1/8	1/8
0	0	1	1	1	1/8	1/8
0	1	0	1	1	1/4	1/8
0	1	1	0	1	0	1/8
1	0	0	1	1	1/8	1/8
1	0	1	0	0	1/8	1/8
1	1	0	0	1	0	1/8
1	1	1	1	1	1/4	1/8

This data generating process (DGP) is consistent with the DAG (satisfies the conditions above including if there are no differences in the two domains, $S = 0$, the causal effects are the same) and demonstrates there exist values of X and Y such that

$$P_1(Y \mid do(X), S = 1) \neq P_2(Y \mid do(X), S = 1)$$

As U is a back-door into X

$$P_i(Y \mid do(X), S = 1) = \sum_u P_i(Y \mid X = x, U = u) P_i(u)$$

In particular,

$$P_1(Y = 1 \mid do(X = 0), S = 1) = 1/2$$

while

$$P_2(Y = 1 \mid do(X = 0), S = 1) = 1$$

and

$$P_1(Y = 1 \mid do(X = 1), S = 1) = 1/2$$

while

$$P_2(Y = 1 \mid do(X = 1), S = 1) = 1$$

Hence, figure 1(b) is not transportable.[2]

Meta-transportability

Meta-transportability from multiple heterogeneous experiments extends transportability of causal effects from a single experiment to the target population. A formal, expanded foray into external validity. Figure 2 illustrates meta-transportability or μ -transportability.

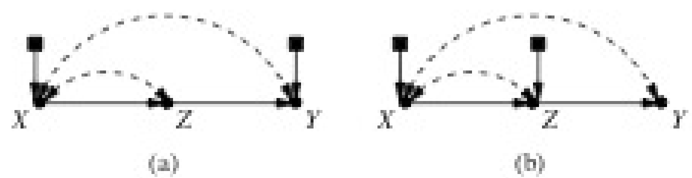

Figure 2: Selection diagrams illustrating impossibility of obtaining $P^*(y|do(x))$ through individual transportability from π_a and π_b to π^*, yet a more elaborated analysis yield the desired result combining different pieces from both domains.

Neither π_a or π_b are individually transportable to π^* , however, collectively they can be transported to the target domain.

$$P^*(y \mid do(x)) = \sum_z P^*(y \mid do(x), z)\, P^*(z \mid do(x))$$

$$= \sum_z P^*(y \mid do(x), do(z))\, P^*(z \mid do(x))$$

$$= \sum_z P^*(y \mid do(z))\, P^*(z \mid do(x))$$

The first line is Bayes chain rule to insert observation of Z . The second line utilizes do-calulus rule 2 interchanging observation with action on Z . The third line employs rule 3 to delete $do(x)$. Since $(S \perp Y \mid Z)_{G^{(b)}_{\overline{Z}}}$, the first term in the last expression can be transported from π_b . Also, since $(S \perp Z \mid X)_{G^{(a)}_{\overline{X}}}$,

[2] Suppose there is no bow in 1(b). Then, $P^*(y \mid do(x)) = P^*(y \mid x)$ so the distribution with $S = 1$ is free of *do*-operators. To see that this allows transportability suppose we only observe $P^*(y \mid x, z)$ (where there may or may not exist a path $Z \to Y$) and $(X, Z \perp S)$, plus $(X \perp Z)$. Then, $P(z \mid x)$ or $P(z)$ are transportable from the source domain yielding

$$P^*(y \mid do(x)) = \sum_z P^*(y \mid x, z)\, P(z)$$

the second term can be transported from π_a . Hence, the causal effect in the target domain can be μ -transported.

$$P^* (y \mid do\,(x)) = \sum_z P^{(b)} (y \mid do\,(z))\, P^{(a)} (z \mid do\,(x))$$

Figure 3 represents are more ambitious setting for μ -transportability.

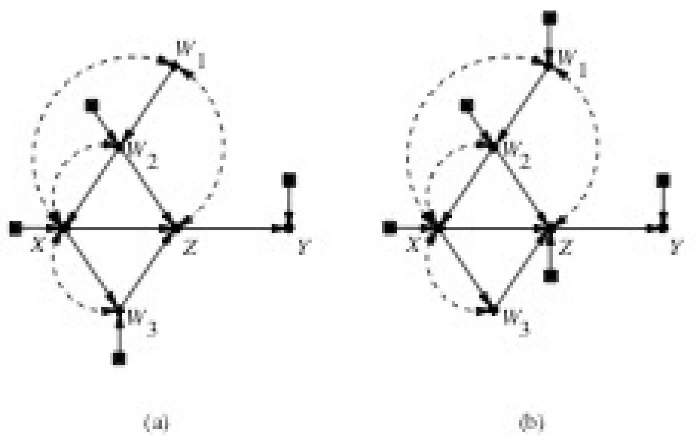

Figure 3: Selection diagrams illustrating a more involved analysis that yields an estimand (Eq. (1)) for the target quantity which combines information from three domains, the two sources π_a and π_b together with the target π^*.

A combination of passive observation from the target domain and active manipulation from π_a and π_b produces the μ - transport formula (1).

$$\begin{aligned} P^* (y \mid do\,(x)) = \sum_{w_1, w_2, w_3, z} & P^* (y \mid z)\, P^{(a)} (w_1, z \mid do\,(x)\,, do\,(w_2)\,, do\,(w_3)) \times \\ & P^* (w_2 \mid w_1)\, P^{(b)} (w_3 \mid do\,(x)\,, do\,(w_1)\,, do\,(w_2)) \end{aligned} \tag{1}$$

Equation (1) follows from do-calculus. First, apply Bayes chain rule.

$$\begin{aligned} P^* (y \mid do\,(x)) = \sum_{w_1, w_2, w_3, z} & P^* (y \mid do\,(x)\,, w_1, w_2, w_3, z)\, P^* (z \mid do\,(x)\,, w_1, w_2, w_3) \times \\ & P^* (w_3 \mid do\,(x)\,, w_1, w_2)\, P^* (w_2 \mid do\,(x)\,, w_1)\, P^* (w_1 \mid do\,(x)) \end{aligned} \tag{2}$$

Since Z d-separates the other variables from Y , the first term simplifies immediately to the first term in (1).

$$P^{*}\left(y \mid do\left(x\right), w_{1}, w_{2}, w_{3}, z\right) = P^{*}\left(y \mid z\right) \tag{3}$$

Rule 2 exchanges action with observation of W_2 and W_3 in the next term focused on Z . Further, π^a transports to replace π^* since all the S -nodes are detached in the subgraph $G_{\overline{XW_2W_3}}$ except for the one into Y, a collider, and $(S \perp Z \mid X, W_1, W_2, W_3)_{G^{(a)}_{\overline{XW_2W_3}}}$.

$$P^{*}\left(z \mid do\left(x\right), w_{1}, w_{2}, w_{3}\right) = P^{(a)}\left(z \mid do\left(x\right), w_{1}, do\left(w_{2}\right), do\left(w_{3}\right)\right)$$

In addition, rule 3 inserts action for W_2 and W_3 on W_1 in either the target or source (a) domain.

$$P^{*}\left(w_{1} \mid do\left(x\right)\right) = P^{(a)}\left(w_{1} \mid do\left(x\right), do\left(w_{2}\right), do\left(w_{3}\right)\right)$$

Collectively, this yields the second term in (1)

$$P^{*}\left(z \mid do\left(x\right), w_{1}, w_{2}, w_{3}\right) P^{*}\left(w_{1} \mid do\left(x\right)\right) = P^{(a)}\left(w_{1}, z \mid do\left(x\right), do\left(w_{2}\right), do\left(w_{3}\right)\right) \tag{4}$$

Rule 3 deletes $do\left(x\right)$ from W_2 leading to the third term in (1).

$$P^{*}\left(w_{2} \mid do\left(x\right), w_{1}\right) = P^{*}\left(w_{2} \mid w_{1}\right) \tag{5}$$

Rule 2 replaces observation with action on W_1 and W_2 in W_3. Further, as $(S \perp W_3 \mid X, W_1, W_2)_{G^{(b)}_{\overline{XW_1W_2}}}$ the last term of the μ-transport formula is derived.

$$P^{*}\left(w_{3} \mid do\left(x\right), w_{1}, w_{2}\right) = P^{(b)}\left(w_{3} \mid do\left(x\right), do\left(w_{1}\right), do\left(w_{2}\right)\right) \tag{6}$$

The demonstration is completed by substituting the equations (3), (4), (5), and (6) into equation (2) to produce equation (1).

μ-transport infeasibility

Lest we believe μ-transportability is always feasible we explore one more example to illustrate μ-transportability failure in figure 4.

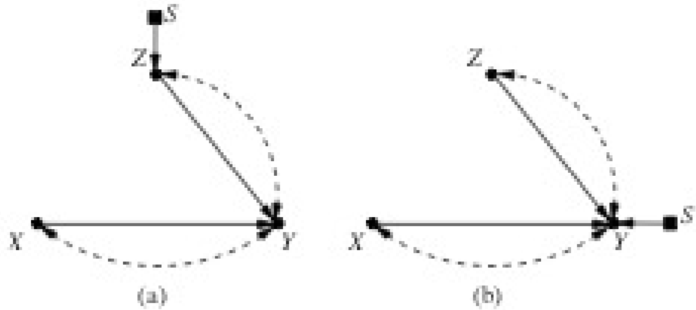

Figure 4: Collection of heterogeneous selection diagrams in which the target relation $P^*(y|do(x))$ is not μ-transportable from both domains (see Theorem 2).

The rules of do-calculus do not allow expressing the causal effect, $P^*\left(y \mid do\left(x\right)\right)$ without manipulation (do-operators) in the target domain. do-calculus rules 1 and 3 do not allow insertion of observation or action on Z. However, a common identification strategies employs Bayes chain rule to insert Z.

$$P^*\left(y \mid do\left(x\right)\right) = \sum_z P^*\left(y \mid do\left(x\right), z\right) P^*\left(z \mid do\left(x\right)\right)$$

First, the second term can be drawn from source (b) as Y is a collider and $\left(S \perp Z \mid X\right)_{G^{(b)}_{\overline{X}}}$.

$$P^*\left(z \mid do\left(x\right)\right) = P^{(b)}\left(z \mid do\left(x\right)\right)$$

Second, although $\left(S \perp Y \mid X, Z\right)_{G^{(a)}_{\overline{Z}}}$ this doesn't apply since do-calculus doesn't permit insertion or exchange of $do\left(Z\right)$ in the first term. Consequently, we cannot eliminate action, $do\left(x\right)$, in the target domain and μ-transportation fails.