Causality in Econometrics and Statistics: Structural Models are Causal Models (Do-Calculus Extract)

by Rodrigo Pinto & James J. Heckman

James J. Heckman

Econ 312, Spring 2023

THE UNIVERSITY OF CHICAGO

Do Operators

The Do-Calculus

- **Attempt:** Counterfactual manipulations using the empirical model.
- **Intent:** Expressions obtained from a hypothetical model.
- **Tools:** Uses causal/graphical/statistical rules outside statistics.
- **Fixing:** Uses $do(X) = x$ for fixing X at x in the DAG for all X-inputs (does not allow to target causal links separately).
- **Flexibility:** Does not easily define complex treatments, such as treatment on the treated, i.e., $E_{\mathrm{E}}(Y|X=1,\tilde{X}=1) - E_{\mathrm{E}}(Y|X=1,\tilde{X}=0)$.

In Contrast: Identification using the hypothetical model is transparent and does not require additional causal rules, only standard statistical tools.

Definition of the Do-Operator (which is Fixing)

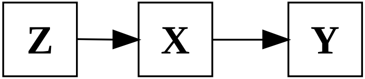
The **Do-operator** is based on the **Truncated Factorization** of the probability factor of the fixed variable is deleted:

Let $X \subset V$: Then

$\Pr(V(x) = v) = \Pr(V_1 = v_1, \ldots, V_{m+n} = v_{m+n}, | do(X) = x)$ and:

$$\Pr(V(x) = v) = \begin{cases} \prod_{V_i \in V \setminus X} P(V_i = v_i | pa(V_i)) & \text{if } v \text{ is consistent with } x; \\ 0 & \text{if } v \text{ is inconsistent with } x. \end{cases}$$

Example of the Do-Operator



- **Variables:** Y, X, Z
- **Factorization:**

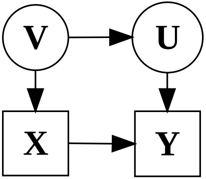
$$\begin{aligned}\Pr(Y, X, Z) &= \Pr(Y|Z, X)\Pr(X|Z)\Pr(Z)\\ &= \Pr(Y|X)\Pr(X|Z)\Pr(Z)\end{aligned}$$

- **Do-operator:** $\Pr(Z, Y|do(X) = x) = \Pr(Y|X = x)\Pr(Z)$
- **Conditional operator:**

$$\begin{aligned}\Pr(Y, Z|X = x) &= \Pr(Y|Z, X = x)\Pr(X|Z, X = x)\Pr(Z|X = x)\\ &= \Pr(Y|X = x)\Pr(Z|X = x)\end{aligned}$$

Do-operator targets variables, not causal links.

Example of the Do-Operator



- **Variables:** Y, X, U, V
- **Factorization:** $\Pr(V, U, X, Y) = \Pr(Y|U, X)\Pr(X|V)\Pr(U|V)\Pr(V)$
- **Do-operator:** $\Pr(V, U, Y|do(X) = x) = \Pr(Y|U, X = x)\Pr(U|V)\Pr(V)$
- **Conditional operator:**

$$\begin{aligned}\Pr(V, U, Y|X = x) &= \Pr(Y|U, V, X = x)\Pr(U|V, X = x)\Pr(V|X = x)\\ &= \Pr(Y|U, X = x)\Pr(U|V)\Pr(V|X = x)\end{aligned}$$

Do-operator targets variables, not causal links.

Comparison: Hypothetical Model and Do-Operator

Fixing within Standard Probability Theory

Fixing in the empirical model is translated to statistical conditioning in the hypothetical model:

$$\underbrace{E_{\mathrm{E}}(Y(x))}_{\text{Causal Operation Empirical Model}} = \underbrace{E_{\mathrm{H}}(Y|\tilde{X} = x)}_{\text{Statistical Operation Hypothetical Model}}$$

do-Operator and Statistical Conditioning

Let $\tilde{X}$ be the hypothetical variable in G_{H} associated with variable X in the empirical model G_{E}, such that $Ch_{\mathrm{H}}(\tilde{X}) = Ch_{\mathrm{E}}(X)$, then:

$$\mathbf{P}_{\mathrm{H}}(\mathcal{T}_{\mathrm{E}} \setminus \{X\}|\tilde{X} = x) = \mathbf{P}_{\mathrm{E}}(\mathcal{T}_{\mathrm{E}} \setminus \{X\}|do(X) = x).$$

THE UNIVERSITY OF CHICAGO

Defining the Do-Calculus

What is the do-calculus?

A set of three graphical/statistical **rules** that **convert** expressions of causal inference into probability equations.

1. **Goal:** Identify causal effects from non-experimental data.
2. **Application:** Bayesian network structure, i.e., Directed Acyclic Graph (DAG) that represents causal relationships.
3. **Identification method:** Iteration of do-calculus rules to generate a function that describes treatment effects statistics as a function of the observed variables only (Tian and Pearl 2002, Tian and Pearl 2003).

Characteristics of Pearl's Do-Calculus

1. **Information:** DAG only provides information on the causal relation among variables.
2. **Not Suited** for examining assumptions on functional forms.
3. **Identification:** If this information is sufficient to identify causal effects, then:
4. **Completeness**:
   1. There exists a **sequence** of application of the Do-Calculus that
   2. **generates** a formula for causal effects based on observational quantities (Huang and Valtorta 2006, Shpitser and Pearl 2006)
5. **Limitation:** Does not allow for additional information outside the DAG framework.
   1. **Only** applies to the information content of a DAG.
   2. **IV** is not identified through Do-calculus
   3. **Why?** requires assumptions outside DAG: linearity, monotonicity, separability.

Notation for the Do-Calculus

More notation is needed to define these rules:

DAG Notation

Let X, Y, Z be arbitrary disjoint sets of variables (nodes) in a causal graph G.

- $G_{\overline{X}}$: DAG that modifies G by deleting the arrows pointing *to* X.
- $G_{\underline{X}}$: DAG that modifies G by deleting arrows emerging *from* X.
- $G_{\overline{X},\underline{Z}}$: DAG that modifies G by deleting arrows pointing *to* X and *emerging* from Z.

THE UNIVERSITY OF CHICAGO

Examples of DAG Notation

G

$$\begin{tikzcd} \boldsymbol{V} \arrow[r] \arrow[d] & \boldsymbol{U} \arrow[d] \\ \boldsymbol{X} \arrow[r] & \boldsymbol{Y} \end{tikzcd}$$

$G_{\underline{X}}$

$$\begin{tikzcd} \boldsymbol{V} \arrow[r] \arrow[d] & \boldsymbol{U} \arrow[d] \\ \boldsymbol{X} & \boldsymbol{Y} \end{tikzcd}$$

$G_{\overline{X}}$

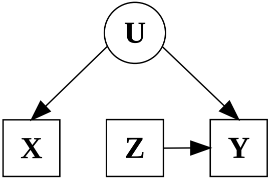
$$\begin{tikzcd} \boldsymbol{V} \arrow[r] & \boldsymbol{U} \arrow[d] \\ \boldsymbol{X} \arrow[r] & \boldsymbol{Y} \end{tikzcd}$$

$G_{\overline{X},\underline{U}}$

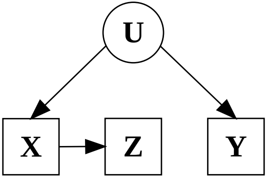
$$\begin{tikzcd} \boldsymbol{V} \arrow[r] & \boldsymbol{U} \\ \boldsymbol{X} \arrow[r] & \boldsymbol{Y} \end{tikzcd}$$

Examples of DAG Notation

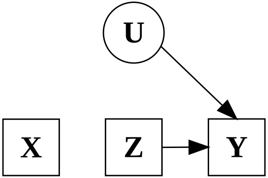
$G_{\underline{X}} = G_{\overline{Z}}$



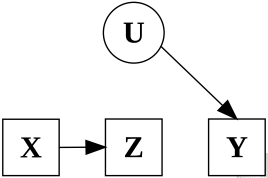
$G_{\underline{Z}}$



$G_{\overline{X},\overline{Z}}$



$G_{\overline{X},\underline{Z}}$



G

$$\begin{tikzcd}
 & \boldsymbol{V} \arrow[r] \arrow[d] & \boldsymbol{U} \arrow[d] \\
\boldsymbol{Z} \arrow[r] & \boldsymbol{X} \arrow[r] & \boldsymbol{Y} \\
 & \boldsymbol{W} \arrow[u] \arrow[ur] &
\end{tikzcd}$$

$G_{\overline{X}}$

$$\begin{tikzcd}
 & \boldsymbol{V} \arrow[r] & \boldsymbol{U} \arrow[d] \\
\boldsymbol{Z} & \boldsymbol{X} \arrow[r] & \boldsymbol{Y} \\
 & \boldsymbol{W} \arrow[ur] &
\end{tikzcd}$$

$G_{\overline{X},\overline{Y}}$

$$\begin{tikzcd}
 & \boldsymbol{V} \arrow[r] & \boldsymbol{U} \arrow[d] \\
\boldsymbol{Z} & \boldsymbol{X} \arrow[r] & \boldsymbol{Y} \\
 & \boldsymbol{W} \arrow[ur] &
\end{tikzcd}$$

$G_{\overline{X},\overline{Z(W)}}$

$$\begin{tikzcd}
 & \boldsymbol{V} \arrow[r] & \boldsymbol{U} \arrow[d] \\
\boldsymbol{Z} & \boldsymbol{X} \arrow[r] & \boldsymbol{Y} \\
 & \boldsymbol{W} \arrow[ur] &
\end{tikzcd}$$

Do-Calculus Rules

- Assumes the Local Markov Condition and independence of ϵ.

Let G be a DAG and let X, Y, Z, W be any disjoint sets of variables. The do-calculus rules are:

- **Rule 1:** *Insertion/deletion of observations:*
  $Y \perp\!\!\!\perp Z|(X, W)$ under $G_{\overline{X}}$
  $$\Rightarrow \mathbf{P}(Y|do(X), Z, W) = \mathbf{P}(Y|do(X), W).$$
- **Rule 2:** *Action/observation exchange:*
  $Y \perp\!\!\!\perp Z|(X, W)$ under $G_{\overline{X}, \underline{Z}}$
  $$\Rightarrow \mathbf{P}(Y|do(X), do(Z), W) = \mathbf{P}(Y|do(X), Z, W).$$
- **Rule 3:** *Insertion/deletion of actions:*
  $Y \perp\!\!\!\perp Z|(X, W)$ under $G_{\overline{X}, \overline{Z(W)}}$
  $$\Rightarrow \mathbf{P}(Y|do(X), do(Z), W) = \mathbf{P}(Y|do(X), W),$$
  where $Z(W)$ is the set of Z-nodes that are not ancestors of any W-node in $G_{\overline{X}}$.

Understanding the Rules of Do-Calculus

Let G be a DAG then for any disjoint sets of variables X, Y, Z, W :
Rule 1: Insertion/deletion of observations

$$\text{If } \underbrace{Y \perp\!\!\!\perp Z|(X, W)}_{\text{Statistical Relation}} \text{ under } \underbrace{G_{\overline{X}}}_{\text{Graphic Criterion}} \text{ then}$$

$$\underbrace{\Pr(Y|do(X), Z, W) = \Pr(Y|do(X), W)}_{\text{Equivalent Probability Expression}}$$

Do-Calculus Exercise

G

$$\begin{tikzcd} \boldsymbol{V} \arrow[r] \arrow[d] & \boldsymbol{U} \arrow[d] \\ \boldsymbol{X} \arrow[r] & \boldsymbol{Y} \end{tikzcd}$$

$G_{\underline{X}}$

$$\begin{tikzcd} \boldsymbol{V} \arrow[r] \arrow[d] & \boldsymbol{U} \arrow[d] \\ \boldsymbol{X} & \boldsymbol{Y} \end{tikzcd}$$

1. LMC to X under $G_{\underline{X}}$ generates $X \perp\!\!\!\perp (U, Y)|V \Rightarrow X \perp\!\!\!\perp (U, Y)|V$.
2. Now if $X \perp\!\!\!\perp (U, Y)|V$ holds under $G_{\underline{X}}$, then, by **Rule 2**,

$$\mathbf{P}(Y|do(X), V) = \mathbf{P}(Y|X, V). \tag{1}$$

$$\therefore E(Y|do(X) = x) = \underbrace{\int E(Y|V = v, do(X) = x)dF_V(v)}_{\text{Using do(X),i.e. Fixing } X}$$

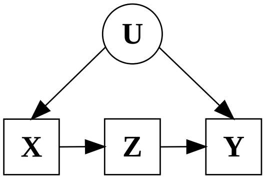
$$= \underbrace{\int E(Y|V = v, X = x)dF_V(v)}_{\text{Replace "do" with Standard Statistical Conditioning}} \text{ by Equation(1)}$$

Do-Calculus Exercise : The Front-Door Model

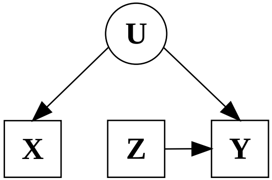
Using the Do-Calculus : Task 1 – Compute $\Pr(Z|do(X))$

$X \perp\!\!\!\perp Z$ in $G_{\underline{X}}$, by **Rule 2**, $\Pr(Z|do(X)) = \Pr(Z|X)$.

G



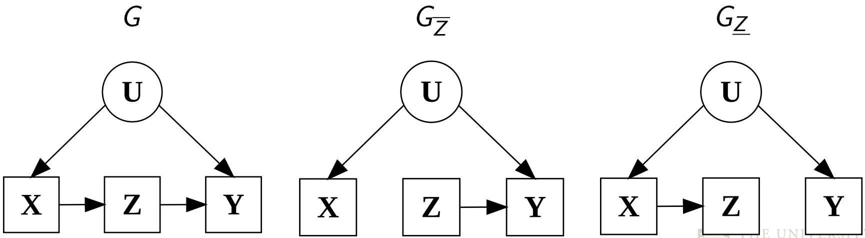
$G_{\underline{X}}$



Using the Do-Calculus : Task 2 – Compute $\Pr(Y|do(Z))$

$Z \perp\!\!\!\perp X$ in $G_{\overline{Z}}$, by **Rule 3**, $\Pr(X|do(Z)) = \Pr(X)$
$Z \perp\!\!\!\perp Y|X$ in $G_{\underline{Z}}$, by **Rule 2**, $\Pr(Y|X, do(Z)) = \Pr(Y|X, Z)$

$$\begin{aligned}\therefore \Pr(Y|do(Z)) &= \sum_X \Pr(Y|X, do(Z)) \Pr(X|do(Z)) \\ &= \sum_X \Pr(Y|X, Z) \Pr(X)\end{aligned}$$



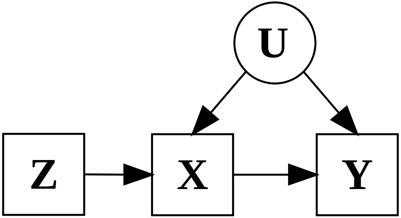
Link to Roy Model Material

Limitations of Do-Calculus for Econometric Identification

Failure of Do-Calculus
Does Not Generates Standard IV Results

The simplest instrumental variable model consists of four variables:

1. A confounding variable U that is external and unobserved.
2. An external instrumental variable Z.
3. An observed variable X caused by U and Z.
4. An outcome Y caused by U and X.



Do-Calculus Non-identification of the IV Model

- **Limitation:** IV model is not identified by literature that relies exclusively on DAGs.
- **Why?:** IV identification relies on assumptions outside the scope of DAG literature.
- **LMC:** generates the conditional independence relationships: $Y \perp\!\!\!\perp Z|(U, X)$ and $U \perp\!\!\!\perp Z$.
- **TSLS:** $X \not\perp\!\!\!\perp Z$ holds, thus, the IV model satisfy the necessary criteria to apply the method of Two Stage Least Squares (TSLS).
- **Assumption Outside of DAGs:** TSLS identifies the IV model under linearity.

THE UNIVERSITY OF CHICAGO

Do-Calculus and IV

The Do-Calculus applied to the IV Model generates:

1. $\Pr(Y|do(X), do(Z)) = \Pr(Y|do(X), Z) = \Pr(Y|do(X))$,
2. $\Pr(Y|do(Z)) = \Pr(Y|Z)$

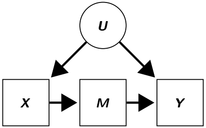
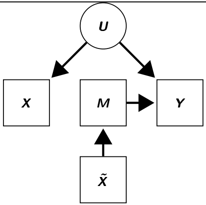
Only establishes the exogeneity of the instrumental variable Z.
Insufficient to identify $\Pr(Y|do(X))$.

- The instrumental variable model is not identified applying the rules of the do-calculus.
- Indeed, in this framework it is impossible to identify the causal effect of X on Y without additional information.
- The instrumental variable model is identified under further assumptions such as linearity, separability, monotonicity.
- However, these assumptions are outside the scope of the do-calculus.

Comparing Analyses Based on the Do-Calculus with Those from the Hypothetical Model

- We illustrate the use of the do-calculus and the hypothetical model approaches by identifying the causal effects of a well-known model that Pearl (2009) calls the "Front-Door model."
- It consists of four variables: (1) an external unobserved variable U; (2) an observed variable X caused by U; (3) an observed variable M caused by X; and (4) an outcome Y caused by U and M.

"Front-Door" Empirical and Hypothetical Models

1. Pearl's "Front-Door" Empirical Model	2. Our Version of the "Front-Door" Hypothetical Model
$\mathcal{T} = \{U, X, M, Y\}$ $\boldsymbol{\epsilon} = \{\epsilon_U, \epsilon_X, \epsilon_M, \epsilon_Y\}$ $Y = f_Y(M, U, \epsilon_Y)$ $X = f_X(U, \epsilon_X)$ $M = f_M(X, \epsilon_M)$ $U = f_U(\epsilon_U)$	$\mathcal{T} = \{U, X, M, Y, \tilde{X}\}$ $\boldsymbol{\epsilon} = \{\epsilon_U, \epsilon_X, \epsilon_M, \epsilon_Y\}$ $Y = f_Y(M, U, \epsilon_Y)$ $X = f_X(U, \epsilon_X)$ $M = f_M(\tilde{X}, \epsilon_M)$ $U = f_U(\epsilon_U)$
	
$Parent(U) = \varnothing,$ $Parent(X) = \{U\}$ $Parent(M) = \{X\}$ $Parent(Y) = \{M, U\}$	$Parent(U) = Parent(\tilde{X}) = \varnothing,$ $Parent(X) = \{U\}$ $Parent(M) = \{\tilde{X}\}$ $Parent(Y) = \{M, U\}$
$Y \perp\!\!\!\perp X \vert (M, U)$ $M \perp\!\!\!\perp U \vert X$	$Y \perp\!\!\!\perp (\tilde{X}, X) \vert (M, U)$ $M \perp\!\!\!\perp (U, X) \vert \tilde{X}$ $X \perp\!\!\!\perp (M, \tilde{X}, Y) \vert U$ $U \perp\!\!\!\perp (M, \tilde{X})$ $\tilde{X} \perp\!\!\!\perp (X, U)$
$\mathbf{P}_{\mathrm{E}}(Y, M, X, U) =$ $\mathbf{P}_{\mathrm{E}}(Y \vert M, U)\, \mathbf{P}_{\mathrm{E}}(X \vert U)\, \mathbf{P}_{\mathrm{E}}(M \vert X)\, \mathbf{P}_{\mathrm{E}}(U)$	$\mathbf{P}_{\mathrm{H}}(Y, M, X, U, \tilde{X}) =$ $\mathbf{P}_{\mathrm{H}}(Y \vert M, U)\, \mathbf{P}(X \vert U)\, \mathbf{P}_{\mathrm{H}}(M \vert \tilde{X})\, \mathbf{P}_{\mathrm{H}}(U)\, \mathbf{P}_{\mathrm{H}}(\tilde{X})$
$\mathbf{P}_{\mathrm{E}}(Y, M, U \vert do(X) = x) =$ $\mathbf{P}_{\mathrm{E}}(Y \vert M, U)\, \mathbf{P}_{\mathrm{E}}(M \vert X = x)\, \mathbf{P}_{\mathrm{E}}(U)$	$\mathbf{P}_{\mathrm{H}}(Y, M, U, X \vert \tilde{X} = x) =$ $\mathbf{P}_{\mathrm{H}}(Y \vert M, U)\, \mathbf{P}(X \vert U)\, \mathbf{P}_{\mathrm{H}}(M \vert \tilde{X} = x)\, \mathbf{P}_{\mathrm{H}}(U)$

- The do-calculus identifies $\mathbf{P}(Y|do(X))$ through four steps which we now perform.
- Steps 1, 2 and 3 identify $\mathbf{P}(M|do(X))$, $\mathbf{P}(Y|do(M))$ and $\mathbf{P}(Y|M, do(X))$ respectively.

1. Invoking LMC for variable M of DAG $G_{\underline{X}}$ generates $X \perp\!\!\!\perp M$. Thus, by Rule 2 of the do-calculus, we obtain $\mathbf{P}(M|do(X)) = \mathbf{P}(M|X)$.

2. Invoking LMC for variable M of DAG $G_{\overline{M}}$ generates $X \perp\!\!\!\perp M$. Thus, by Rule 3 of the do-calculus, $\mathbf{P}(X|do(M)) = \mathbf{P}(X)$. In addition, applying LMC for variable M of DAG $G_{\underline{M}}$ generates $M \perp\!\!\!\perp Y|X$. Thus, by Rule 2 of do-calculus, $\mathbf{P}(Y|X, do(M)) = \mathbf{P}(Y|X, M)$.

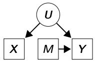
   $$\begin{aligned}\text{Therefore } \mathbf{P}(Y|do(M)) &= \sum_{x' \in \text{supp}(X)} \mathbf{P}(Y|X = x', do(M))\,\mathbf{P}(X = x'|do(M)) \\ &= \sum_{x' \in \text{supp}(X)} \mathbf{P}(Y|X = x', M)\,\mathbf{P}(X = x'),\end{aligned}$$

   where "supp" means support.

3. Invoking LMC for variable M of DAG $G_{\overline{X},\underline{M}}$ generates $Y \perp\!\!\!\perp M|X$.

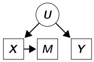
Do-Calculus and the Front-Door Model

1. Modified Front-Door Model $G_{\underline{X}} = G_{\overline{M}}$



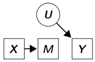
$$(Y, M) \perp\!\!\!\perp X | U$$
$$(X, U) \perp\!\!\!\perp M$$

2. Modified Front-Door Model $G_{\underline{M}}$



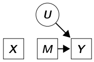
$$(X, M) \perp\!\!\!\perp Y | U$$
$$(Y, U) \perp\!\!\!\perp M | X$$

3. Modified Front-Door Model $G_{\overline{X}, \underline{M}}$



$$(X, M) \perp\!\!\!\perp (Y, U)$$

4. Modified Front-Door Model $G_{\overline{X}, \overline{M}}$



$$(Y, M, U) \perp\!\!\!\perp X$$
$$U \perp\!\!\!\perp M$$

These rules are intended to supplement standard statistical tools with a new set of "do" operations.

1. Thus, by Rule 2 of the do-calculus, $\mathbf{P}(Y|M, do(X)) = \mathbf{P}(Y|do(M), do(X))$. In addition, applying LMC for variable X of DAG $G_{\overline{X},\overline{M}}$ generates $(Y, M, U) \perp\!\!\!\perp X$. By weak union and decomposition, we obtain $Y \perp\!\!\!\perp X|M$. Thus, by Rule 3 of the do-calculus, we obtain that $\mathbf{P}(Y|do(X), do(M)) = \mathbf{P}(Y|do(M))$. Thus, $\mathbf{P}(Y|M, do(X)) = \mathbf{P}(Y|do(M), do(X)) = \mathbf{P}(Y|do(M))$.
2. We collect the results from the three previous steps to identify $\mathbf{P}(Y|do(X))$:

$$
\begin{aligned}
&\mathbf{P}(Y|do(X) = x) \\
&= \sum_{m \in \text{supp}(M)} \mathbf{P}(Y|M, do(X) = x)\, \mathbf{P}(M|do(X) = x) \\
&= \sum_{m \in \text{supp}(M)} \underbrace{\mathbf{P}(Y|do(M) = m, do(X) = x)}_{\textbf{Step 3}}\, \mathbf{P}(M = m|do(X) = x) \\
&= \sum_{m \in \text{supp}(M)} \underbrace{\mathbf{P}(Y|do(M) = m)}_{\textbf{Step 3}}\, \mathbf{P}(M = m|do(X) = x) \\
&= \sum_{m \in \text{supp}(M)} \underbrace{\left(\sum_{x' \in \text{supp}(X)} \mathbf{P}(Y|X = x', M)\, \mathbf{P}(X = x') \right)}_{\textbf{Step 2}} \underbrace{\mathbf{P}(M = m|X = x)}_{\textbf{Step 1}}.
\end{aligned}
$$

- We use the do-calculus to identify the desired causal parameter, using the approach inspired by Haavelmo's ideas.
- We replace the relationship of X on M by a hypothetical variable $\tilde{X}$ that causes M.
- We use $\mathbf{P}_{\mathrm{E}}$ to denote the probability of the Front-Door model that generates the data and $\mathbf{P}_{\mathrm{H}}$ for the hypothetical model.

Lemma 1

In the Front-Door hypothetical model,
(1) $Y \perp\!\!\!\perp \tilde{X} | M$,
(2) $X \perp\!\!\!\perp M$, *and*
(3) $Y \perp\!\!\!\perp \tilde{X} | (M, X)$

Proof

By LMC for X, we obtain $(Y, M, \tilde{X}) \perp\!\!\!\perp X | U$. By LMC for Y we obtain $Y \perp\!\!\!\perp (X, \tilde{X}) | (M, U)$. By Contraction applied to $(Y, M, \tilde{X}) \perp\!\!\!\perp X | U$ and $Y \perp\!\!\!\perp (X, \tilde{X}) | (M, U)$ we obtain $(Y, X) \perp\!\!\!\perp \tilde{X} | (M, U)$. By LMC for U we obtain $(M, \tilde{X}) \perp\!\!\!\perp U$. By Contraction applied to $(M, \tilde{X}) \perp\!\!\!\perp U$ and$(Y, M, \tilde{X}) \perp\!\!\!\perp X | U$ we obtain$(X, U) \perp\!\!\!\perp (M, \tilde{X})$. The second relationship of the Lemma is obtained by Decomposition. In addition, by Contraction on $(Y, X) \perp\!\!\!\perp \tilde{X} | (M, U)$ and $(M, \tilde{X}) \perp\!\!\!\perp U$ we obtain $(Y, X, U) \perp\!\!\!\perp \tilde{X} | M$. The two remaining conditional independence relationships of the Lemma are obtained by Weak Union and Decomposition.

Applying these results,

$$
\begin{aligned}
&\mathbf{P}_{\mathrm{H}}(Y|\tilde{X}=x) \\
&\quad= \sum_{m\in\mathrm{supp}(M)} \mathbf{P}_{\mathrm{H}}(Y|M=m,\tilde{X}=x)\,\mathbf{P}_{\mathrm{H}}(M=m|\tilde{X}=x) \\
&\quad= \sum_{m\in\mathrm{supp}(M)} \mathbf{P}_{\mathrm{H}}(Y|M=m)\,\mathbf{P}_{\mathrm{H}}(M=m|\tilde{X}=x) \\
&\quad= \sum_{m\in\mathrm{supp}(M)} \left(\sum_{x'\in\mathrm{supp}(X)} \mathbf{P}_{\mathrm{H}}(Y|X=x',M=m)\,\mathbf{P}_{\mathrm{H}}(X=x'|M=m)\right)\mathbf{P}_{\mathrm{H}}(M=m|\tilde{X}=x) \\
&\quad= \sum_{m\in\mathrm{supp}(M)} \left(\sum_{x'\in\mathrm{supp}(X)} \mathbf{P}_{\mathrm{H}}(Y|X=x',M=m)\,\mathbf{P}_{\mathrm{H}}(X=x')\right)\mathbf{P}_{\mathrm{H}}(M=m|\tilde{X}=x) \\
&\quad= \sum_{m\in\mathrm{supp}(M)} \left(\sum_{x'\in\mathrm{supp}(X)} \mathbf{P}_{\mathrm{H}}(Y|X=x',\tilde{X}=x',M=m)\,\mathbf{P}_{\mathrm{H}}(X=x')\right)\mathbf{P}_{\mathrm{H}}(M=m|\tilde{X}=x) \\
&\quad= \sum_{m\in\mathrm{supp}(M)} \left(\sum_{x'\in\mathrm{supp}(X)} \underbrace{\mathbf{P}_{\mathrm{E}}(Y|M,X=x')}_{\text{by T1}}\,\underbrace{\mathbf{P}_{\mathrm{E}}(X=x')}_{\text{by Lemma1}}\right)\underbrace{\mathbf{P}_{\mathrm{E}}(M=m|X=x)}_{\text{by M1}}.
\end{aligned}
$$

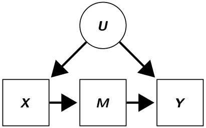
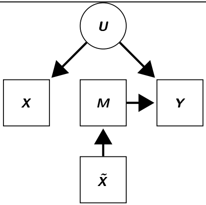
THE UNIVERSITY OF CHICAGO

- The second equality comes from relationship (1) $Y \perp\!\!\!\perp \tilde{X}|M$ of Lemma 1.
- The fourth equality comes from relationship (2) $X \perp\!\!\!\perp M$ of Lemma 1.
- The fifth equality comes from relationship (3) $Y \perp\!\!\!\perp \tilde{X}|(M, X)$ of Lemma 1.
- The last equality links the distributions of the hypothetical model with the ones of the empirical model.

- The first term uses Theorem 1 to equate $\mathbf{P}_{\mathrm{H}}(Y|X=x',\tilde{X}=x',M=m)=\mathbf{P}_{\mathrm{E}}(Y|M,X=x')$.
- The second term uses the fact that X is not a child of $\tilde{X}$, thus by Lemma, $\mathbf{P}_{\mathrm{H}}(X=x')=\mathbf{P}_{\mathrm{E}}(X=x')$.
- Finally, the last term uses Matching applied to M. Namely, LMC for M generates $M \perp\!\!\!\perp X|\tilde{X}$ in the hypothetical model.
- Then, by Matching, $\mathbf{P}_{\mathrm{H}}(M|\tilde{X}=x)=\mathbf{P}_{\mathrm{E}}(M|X=x)$.

- Both frameworks produce the same final identification formula.
- The methods underlying them differ greatly.
- Concept in the framework inspired by Haavelmo is the notion of a hypothetical model.

"Front-Door" Empirical and Hypothetical Models

1. Pearl's "Front-Door" Empirical Model	2. Our Version of the "Front-Door" Hypothetical Model
$\mathcal{T} = \{U, X, M, Y\}$ $\boldsymbol{\epsilon} = \{\epsilon_U, \epsilon_X, \epsilon_M, \epsilon_Y\}$ $Y = f_Y(M, U, \epsilon_Y)$ $X = f_X(U, \epsilon_X)$ $M = f_M(X, \epsilon_M)$ $U = f_U(\epsilon_U)$	$\mathcal{T} = \{U, X, M, Y, \tilde{X}\}$ $\boldsymbol{\epsilon} = \{\epsilon_U, \epsilon_X, \epsilon_M, \epsilon_Y\}$ $Y = f_Y(M, U, \epsilon_Y)$ $X = f_X(U, \epsilon_X)$ $M = f_M(\tilde{X}, \epsilon_M)$ $U = f_U(\epsilon_U)$
	
$Parent(U) = \varnothing,$ $Parent(X) = \{U\}$ $Parent(M) = \{X\}$ $Parent(Y) = \{M, U\}$	$Parent(U) = Parent(\tilde{X}) = \varnothing,$ $Parent(X) = \{U\}$ $Parent(M) = \{\tilde{X}\}$ $Parent(Y) = \{M, U\}$
$Y \perp\!\!\!\perp X \| (M, U)$ $M \perp\!\!\!\perp U \| X$	$Y \perp\!\!\!\perp (\tilde{X}, X) \| (M, U)$ $M \perp\!\!\!\perp (U, X) \| \tilde{X}$ $X \perp\!\!\!\perp (M, \tilde{X}, Y) \| U$ $U \perp\!\!\!\perp (M, \tilde{X})$ $\tilde{X} \perp\!\!\!\perp (X, U)$
$\mathbf{P}_\mathrm{E}(Y, M, X, U) =$ $\mathbf{P}_\mathrm{E}(Y \| M, U)\, \mathbf{P}_\mathrm{E}(X \| U)\, \mathbf{P}_\mathrm{E}(M \| X)\, \mathbf{P}_\mathrm{E}(U)$	$\mathbf{P}_\mathrm{H}(Y, M, X, U, \tilde{X}) =$ $\mathbf{P}_\mathrm{H}(Y \| M, U)\, \mathbf{P}(X \| U)\, \mathbf{P}_\mathrm{H}(M \| \tilde{X})\, \mathbf{P}_\mathrm{H}(U)\, \mathbf{P}_\mathrm{H}(\tilde{X})$
$\mathbf{P}_\mathrm{E}(Y, M, U \| do(X) = x) =$ $\mathbf{P}_\mathrm{E}(Y \| M, U)\, \mathbf{P}_\mathrm{E}(M \| X = x)\, \mathbf{P}_\mathrm{E}(U)$	$\mathbf{P}_\mathrm{H}(Y, M, U, X \| \tilde{X} = x) =$ $\mathbf{P}_\mathrm{H}(Y \| M, U)\, \mathbf{P}(X \| U)\, \mathbf{P}_\mathrm{H}(M \| \tilde{X} = x)\, \mathbf{P}_\mathrm{H}(U)$

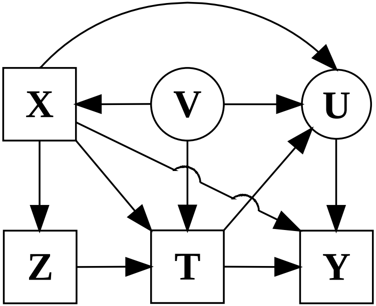
Comparing Pearl and Haavelmo

Summarizing the Do-Calculus of Pearl (2009) and Haavelmo's Inspired Framework

- **Common Features** of Haavelmo and Do Calculus:
  1. **Autonomy** (Frisch, 1938)
  2. **Errors Terms:** ϵ mutually independent
  3. **Statistical Tools:** LMC and GA apply
  4. **Counterfactuals:** Fixing or Do-operator is a Causal, not statistical, Operation.
- **Distinct Features** of Haavelmo and Do Calculus:

	Haavelmo	**Do-calculus**
Approach:	Thinks Outside the Box	Applies Complex Tools
Introduces:	Hypothetical Model	Graphical Rules
Identification:	Connects P_H and P_E	Iteration of Rules
Versatility:	Basic Statistics Apply	Extra Notation/Tools

Generalized Roy Model



This figure represents causal relationships of the Generalized Roy Model. Arrows represent direct causal relationships. Circles represent unobserved variables. Squares represent observed variables

Key Aspects of the Generalized Roy Model

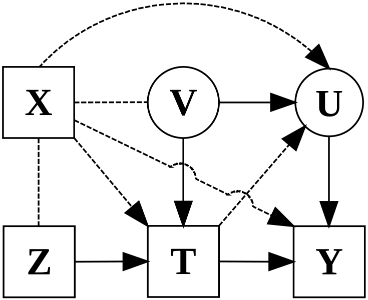
1. T is caused by Z, V;
2. U mediates the effects of V on Y (that is V causes U);
3. T and U cause Y and
4. Z (instrument) not caused by V, U and does not directly cause Y, U.

We are left to examine the cases whether:

1. V causes X (or vice-versa),
2. X causes Z (or vice-versa),
3. X causes T,
4. X causes U,
5. T causes U, and
6. X causes Y.

The combinations of all these causal relationships generate 144 possible models (Pinto, 2013).

Key Aspects of the Generalized Roy Model (Pinto, 2013)



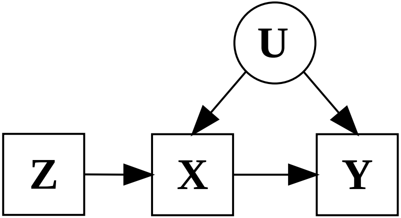
Dashed lines denote causal relationships that may not exist or, if they exist, the causal direction can go either way. Dashed arrows denote causal relationships that may not exist, but, if they exist, the causal direction must comply the arrow direction.

Marginalizing the Generalized Roy Model

- We examine the identification of causal effects of the Generalized Roy Model using a simplified model w.l.o.g.
- Suppress variables X and U.
- This simplification is usually called marginalization in the DAG literature (Koster (2002), Lauritzen (1996), Wermuth (2011)).

Marginalizing the Generalized Roy Model

$$G = G_{\overline{Z}}$$



This figure represents causal relationships of the Marginalized Roy Model. Arrows represent direct causal relationships. Circles represent unobserved variables. Squares represent observed variables

Note: Z is exogenous, thus conditioning on Z is equivalent to fixing Z.

Examining the Marginalized Roy Model – 1/4

- $Y \perp\!\!\!\perp Z$ in $G_{\overline{X}}$, by **Rule 1**

$$\Pr(Y|do(X), Z) = \Pr(Y|do(X))$$

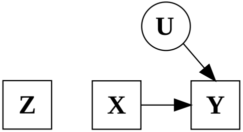
- $Y \perp\!\!\!\perp Z$, in $G_{\overline{X},\overline{Z}}$, by **Rule 3**

$$\Pr(Y|do(X), Z) = \Pr(Y|do(X))$$

- $Y \perp\!\!\!\perp Z|X$ in $G_{\overline{X},\underline{Z}}$, by **Rule 2**

$$\Pr(Y|do(X), do(Z)) = \Pr(Y|do(X), Z)$$

$$G_{\overline{X}} = G_{\overline{X},\overline{Z}} = G_{\overline{X},\underline{Z}}$$

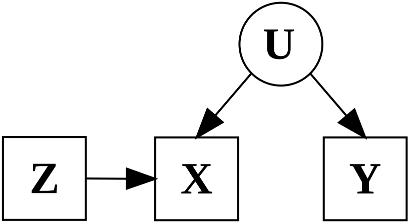


THE UNIVERSITY OF CHICAGO

Examining the Marginalized Roy Model – 2/4

- Under $G_{\overline{X}}$, $Y \not\perp\!\!\!\perp X$, thus **Rule 2** does not apply.
- Under $G_{\underline{X},\overline{Z}}$, $Y \not\perp\!\!\!\perp X|Z$, thus **Rule 2** does not apply.

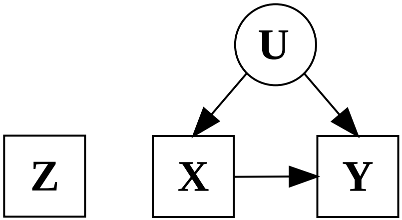
$$G_{\underline{X}} = G_{\underline{X},\overline{Z}}$$



Examining the Marginalized Roy Model – 3/4

- $G_{\underline{Z}} \Rightarrow Y \perp\!\!\!\perp Z$, thus by **Rule 2** $\Pr(Y|do(Z)) = \Pr(Y|Z)$.

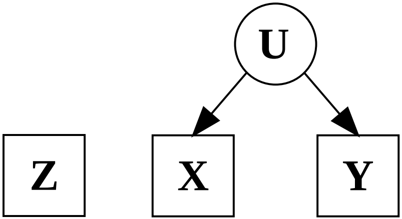
$G_{\underline{Z}}$



Examining the Marginalized Roy Model – 4/4 Modifications

- Under $G_{\underline{X},\underline{Z}}$, $Y \not\perp\!\!\!\perp (X, Z)$, thus **Rule 2** does not apply.

$G_{\underline{X},\underline{Z}}$



Conclusion of Do-Calculus and the Roy Model

The Do-Calculus applied to the Marginalized Roy Model generates:

1. $\Pr(Y|do(X), do(Z)) = \Pr(Y|do(X), Z) = \Pr(Y|do(X))$,
2. $\Pr(Y|do(Z)) = \Pr(Y|Z)$

These relationships only corroborate the exogeneity of the instrumental variable Z and are not sufficient to identify $\Pr(Y|do(X))$.

Identification of the Roy Model

To identify the Roy Model, we make assumption on how Z impacts X, i.e. monotonicity/separability.
These assumptions **cannot** be represented in a DAG.
These assumptions are associated with properties of **how** Z causes X and not only **if** Z causes X.

Return to main text