Average Treatment Effect (again)

$$\boxed{ATE = \frac{1}{N}\sum_{i=1}^{N}\tau_i = \frac{\sum_{i=1}^{N} Y_i(1)}{N} + \frac{\sum_{i=1}^{N} Y_i(0)}{N} = E[Y_i(1)] + E[Y_i(0)]}$$

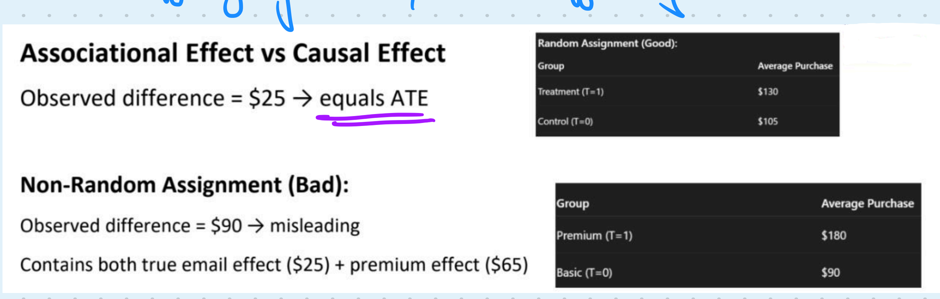
We can represent the effect of treatment as follows:

$$\boxed{Y_i = \underbrace{Y_i(0)}_{\text{before treatment}} + \underbrace{(Y_i(1) - Y_i(0))}_{\text{change due to treatment}} \underbrace{D_i}_{\text{whether } Y_i \text{ is treated or not}}}$$

Conditional ATE (CATE)

$$\boxed{CATE = E[Y_i(1) - Y_i(0) \mid D_i, \underbrace{X_i}_{\text{covariate characteristics of } Y_i}]}$$

Diff. groups respond differently to same treatment



Associational Effect vs Causal Effect

Observed difference = $25 → equals ATE

Random Assignment (Good):	
Group	Average Purchase
Treatment (T=1)	$130
Control (T=0)	$105

Non-Random Assignment (Bad):

Observed difference = $90 → misleading

Contains both true email effect ($25) + premium effect ($65)

Group	Average Purchase
Premium (T=1)	$180
Basic (T=0)	$90

Adjustment Formula

$$\boxed{\underbrace{E[Y(x) \mid Z=z]}_{\substack{\text{could set treatment} \\ \text{to } X=x, \\ \text{w.r.t. } Z=z}} = \underbrace{E[Y \mid do(X=x, Z=z]}_{\substack{\text{actively set } X=x \\ Z=z}} = \underbrace{E[Y \mid X=x, Z=z]}_{\substack{\text{given } X=x \\ Z=z}}}$$

$$\boxed{E(Y(x)) = E[Y \mid do(X=x)] = E[Y \mid X=x, Z]}$$

Calculate over every level of Z then average across Z distribution

Eg: $E(\text{Health} \mid do(\text{Exercise})) = E[\text{Health} \mid \text{Exercise} = \text{yes}]\, P(\text{Age} = \text{age}) = E[\text{Health} \mid \text{Exercise} = \text{yes}, \text{Age} = \text{age}]$

Variable selection: domain knowledge + knowledge of causal structure

Overadjustment: too many variables conditioned on → increased bias

Methods of Adjustment

Stratification

1. Divide Y into strata based on Z values
2. For each stratum, calculate $E(Y \mid X=x)$
3. Weight each:
   $E(Y \mid X=x) \cdot P(Z=z)$
4. Sum weighted estimates

Regression

1. Train a regression model for Y on X and Z
2. Use the model to predict for every X and Z combination } Frequencies of Y predicted become the weight
3. Average over predictions

Assumptions of Adjustment Formula

1. Causal sufficiency
   Basically unconfoundedness → untestable
   No variable beyond Z that has a causal relation with X, Y
2. No model misspecification [regression]
   Essentially consistency
   Model formulation should match data

Choosing Variables

- Choose all confounders
- Do not choose any $\underbrace{\text{mediators/colliders}}_{\text{intermediate link between treatment and effect}}$

  $X \rightarrow A \rightarrow Y$

  part of causal relationship b/w X, Y.
  should not be ignored.

Sensitivity Analysis

assess how robust our causal effect estimates are to potential violations of assumptions, especially unmeasured confounding

Methods:

- E-values: quantify how strong an unmeasured confounder would have to be to explain away your causal result.
  high E-value → robust result
  low E-value → fragile result
- Bias formulae: Bias = f(unmeasured confounders)
- Simulations

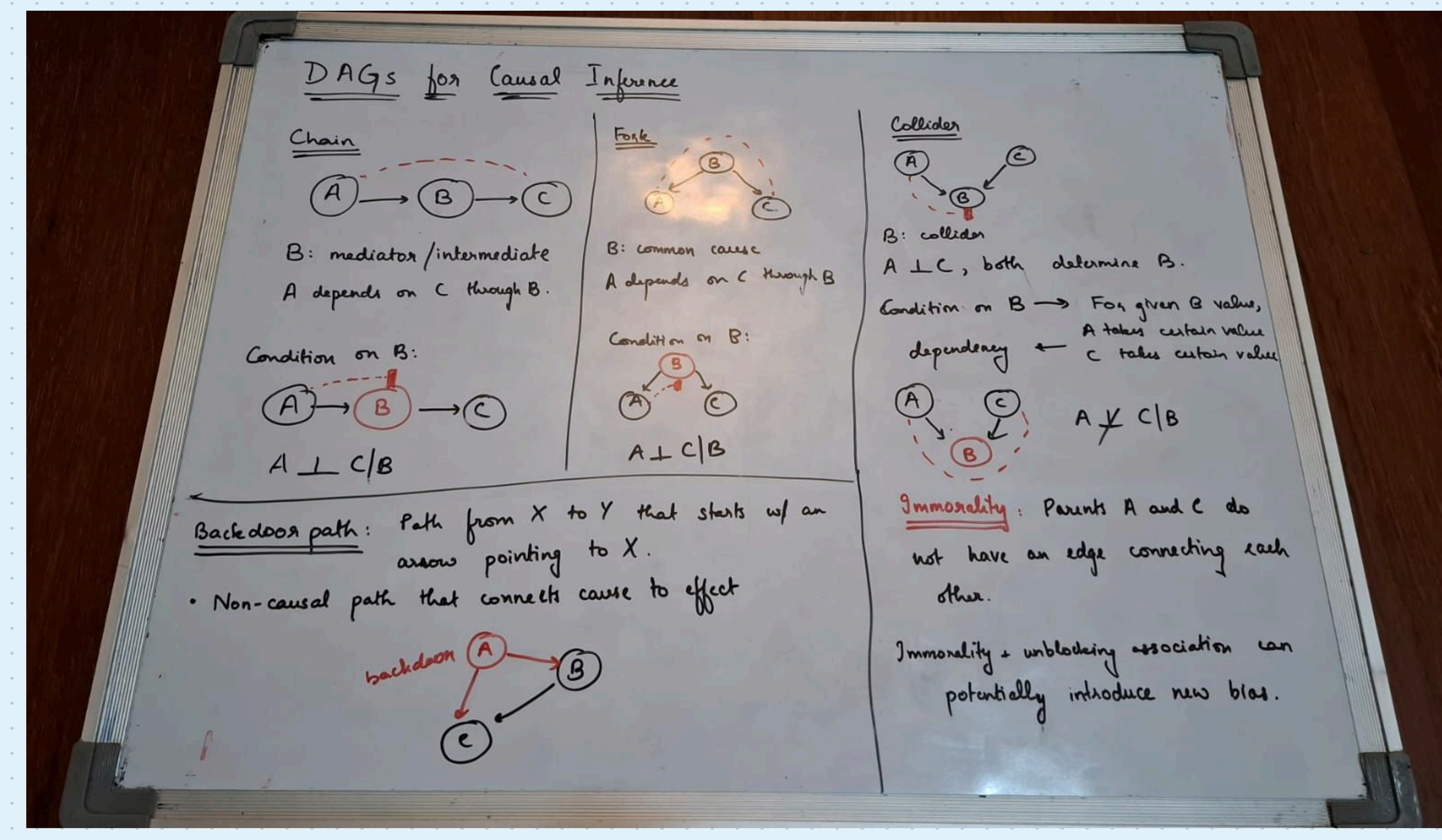
Causal Discovery

Algorithms that try to discover causal structure and hence inform choice of variables in adjustment

Many compatible structures for given data

Usually assumes all conditional independencies ∈ causal structure

Computationally expensive



Chains / Forks: conditioning blocks explaining power

Collider: " " "explaining away" dependence