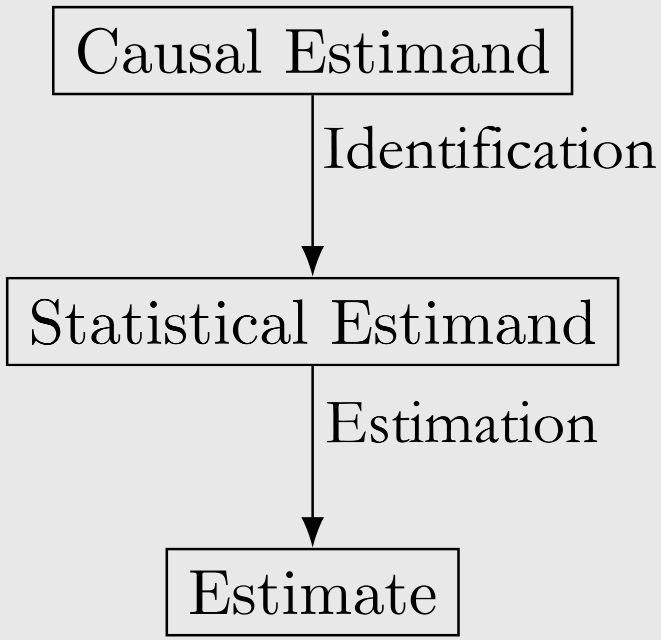
Causal Models

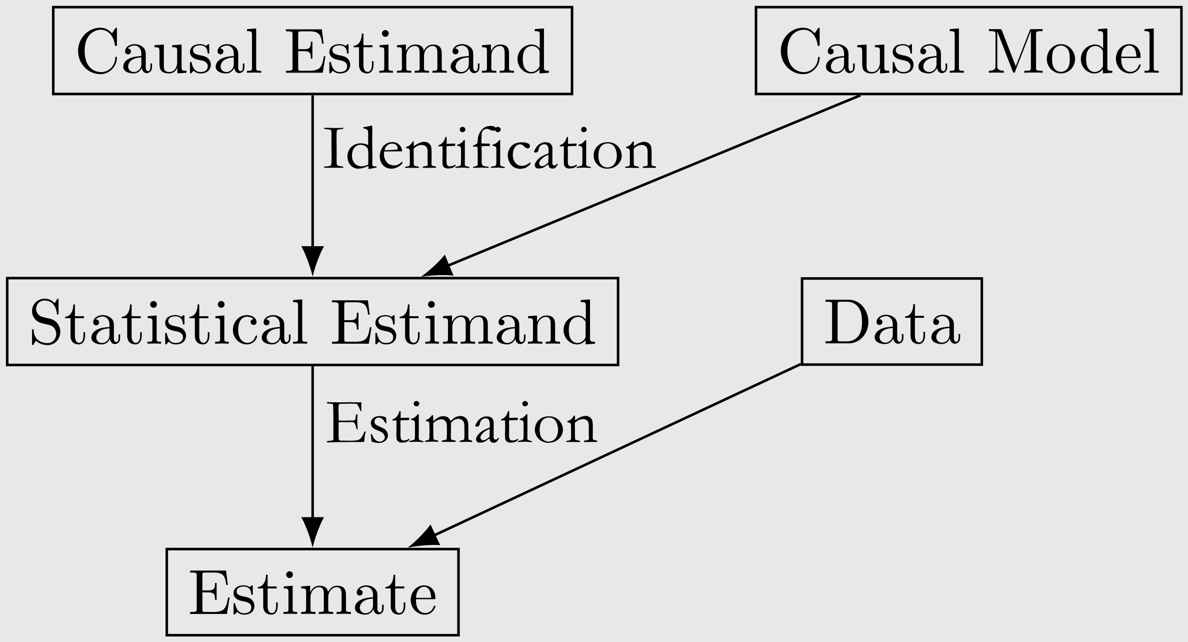
Brady Neal

causalcourse.com

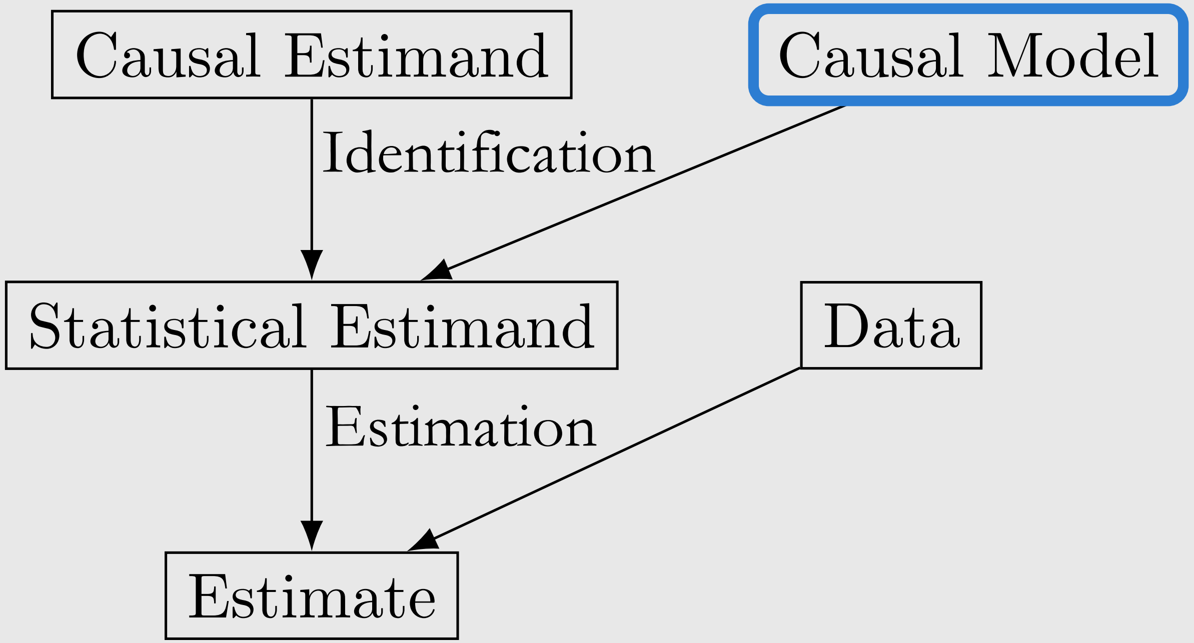
The Identification-Estimation Flowchart



The Identification-Estimation Flowchart



The Identification-Estimation Flowchart



The *do*-operator

Main assumption: modularity

Backdoor adjustment

Structural causal models

A complete example with estimation

The *do*-operator

Main assumption: modularity

Backdoor adjustment

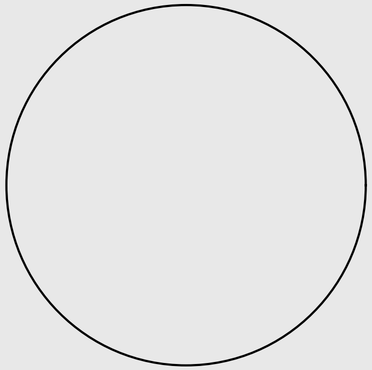
Structural causal models

A complete example with estimation

Conditioning vs. intervening

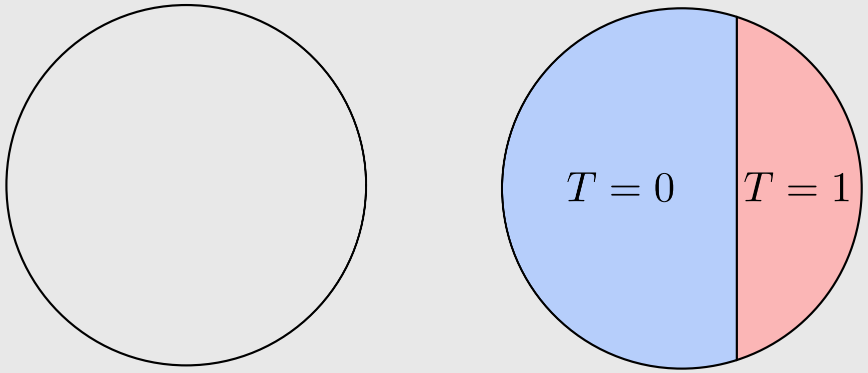
Conditioning vs. intervening

Population

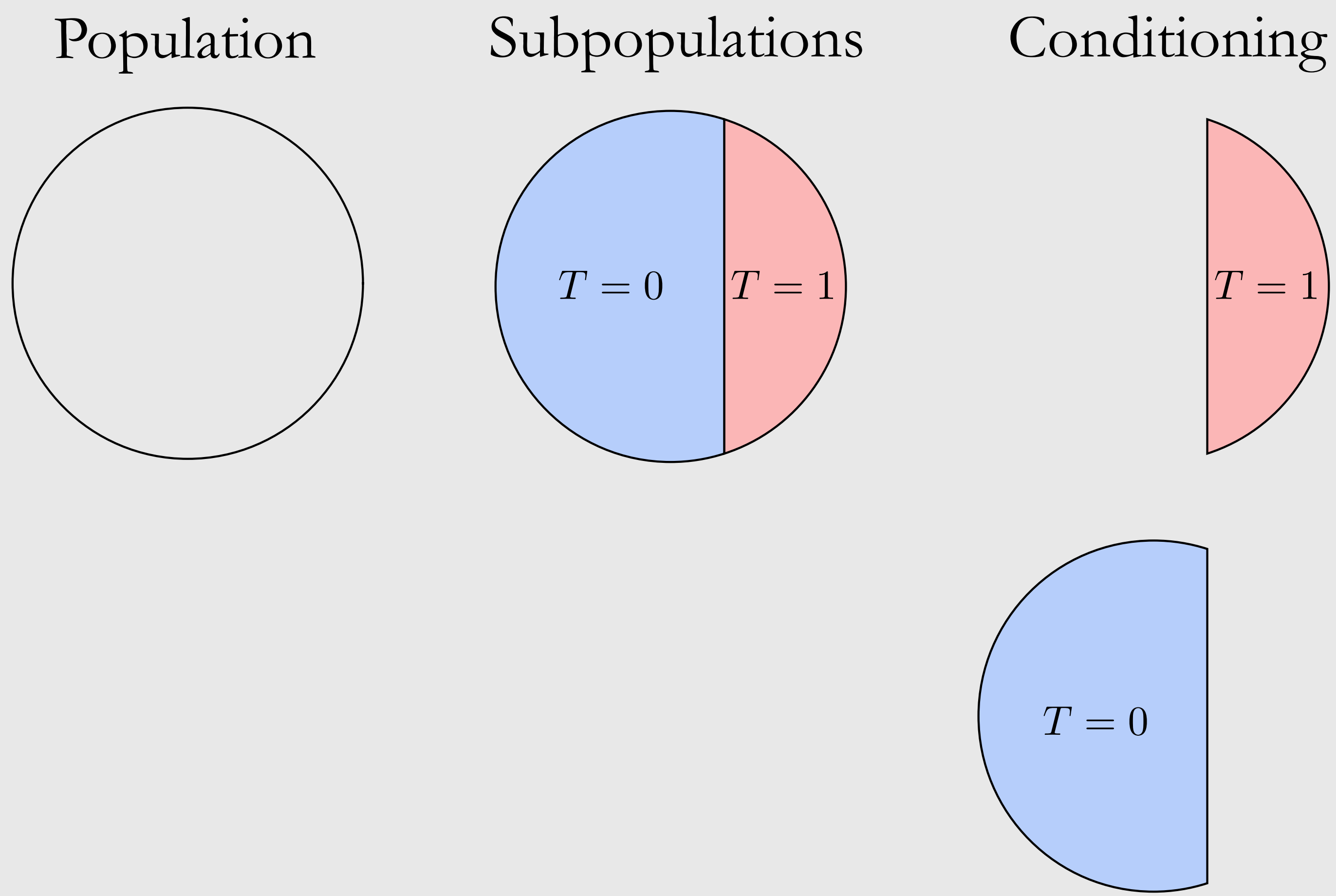


Conditioning vs. intervening

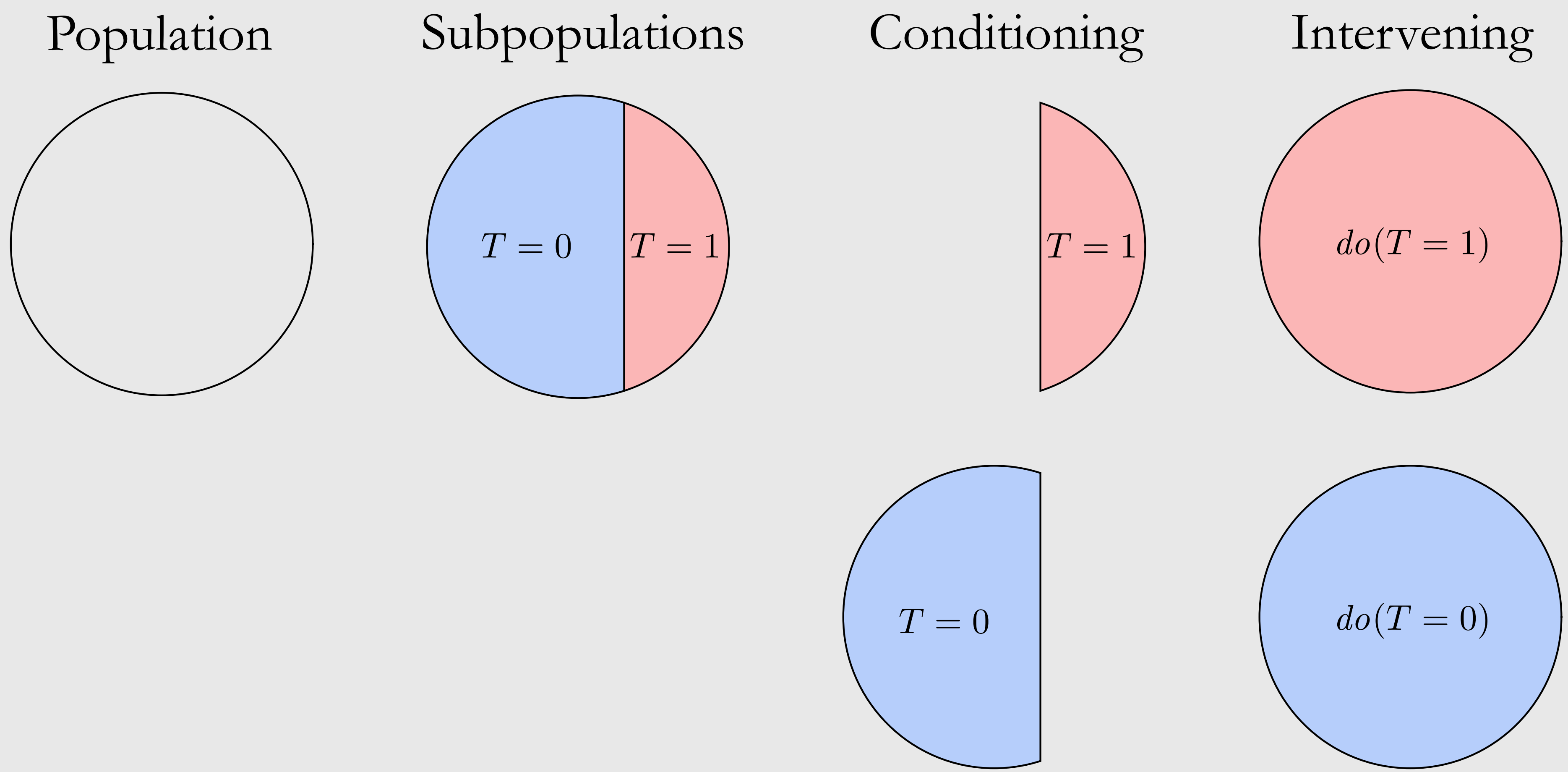
Population　　Subpopulations



Conditioning vs. intervening



Conditioning vs. intervening



Some notation and terminology

Some notation and terminology

Interventional distributions:

$$P(Y(t) = y)$$

Some notation and terminology

Interventional distributions:

$$P(Y(t) = y) \triangleq P(Y = y \mid do(T = t))$$

Some notation and terminology

Interventional distributions:

$$P(Y(t) = y) \triangleq P(Y = y \mid do(T = t)) \triangleq P(y \mid do(t))$$

Some notation and terminology

Interventional distributions:

$$P(Y(t) = y) \triangleq P(Y = y \mid do(T = t)) \triangleq P(y \mid do(t))$$

Average treatment effect (ATE):

$$\mathbb{E}[Y \mid do(T = 1)] - \mathbb{E}[Y \mid do(T = 0)]$$

Some notation and terminology

Interventional distributions:

$$P(Y(t) = y) \triangleq P(Y = y \mid do(T = t)) \triangleq P(y \mid do(t))$$

Average treatment effect (ATE):

$$\mathbb{E}[Y \mid do(T = 1)] - \mathbb{E}[Y \mid do(T = 0)]$$

Observational Interventional

Some notation and terminology

Interventional distributions:

$$P(Y(t) = y) \triangleq P(Y = y \mid do(T = t)) \triangleq P(y \mid do(t))$$

Average treatment effect (ATE):

$$\mathbb{E}[Y \mid do(T = 1)] - \mathbb{E}[Y \mid do(T = 0)]$$

Observational	Interventional
$P(Y, T, X)$	

Some notation and terminology

Interventional distributions:

$$P(Y(t) = y) \triangleq P(Y = y \mid do(T = t)) \triangleq P(y \mid do(t))$$

Average treatment effect (ATE):

$$\mathbb{E}[Y \mid do(T = 1)] - \mathbb{E}[Y \mid do(T = 0)]$$

Observational	Interventional
$P(Y, T, X)$	$P(Y \mid do(T = t))$

Some notation and terminology

Interventional distributions:

$$P(Y(t) = y) \triangleq P(Y = y \mid do(T = t)) \triangleq P(y \mid do(t))$$

Average treatment effect (ATE):

$$\mathbb{E}[Y \mid do(T = 1)] - \mathbb{E}[Y \mid do(T = 0)]$$

Observational	Interventional
$P(Y, T, X)$	$P(Y \mid do(T = t))$
$P(Y \mid T = t)$	

Some notation and terminology

Interventional distributions:

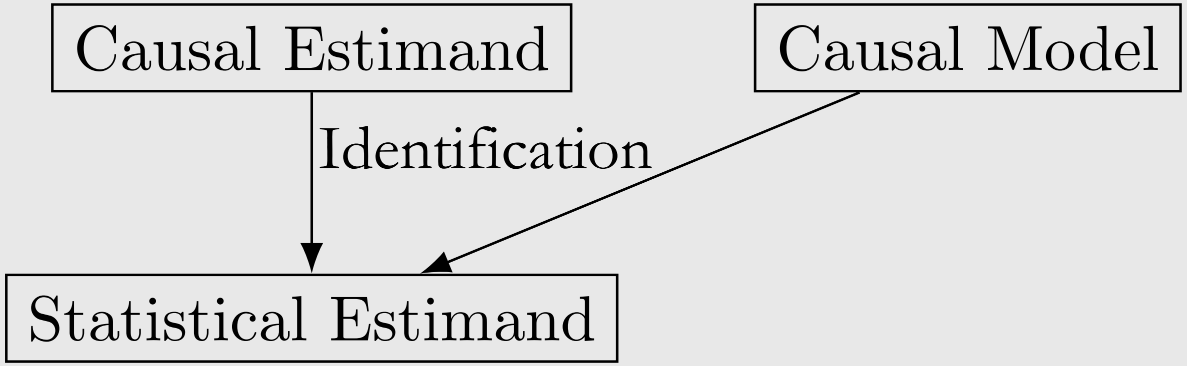
$$P(Y(t) = y) \triangleq P(Y = y \mid do(T = t)) \triangleq P(y \mid do(t))$$

Average treatment effect (ATE):

$$\mathbb{E}[Y \mid do(T = 1)] - \mathbb{E}[Y \mid do(T = 0)]$$

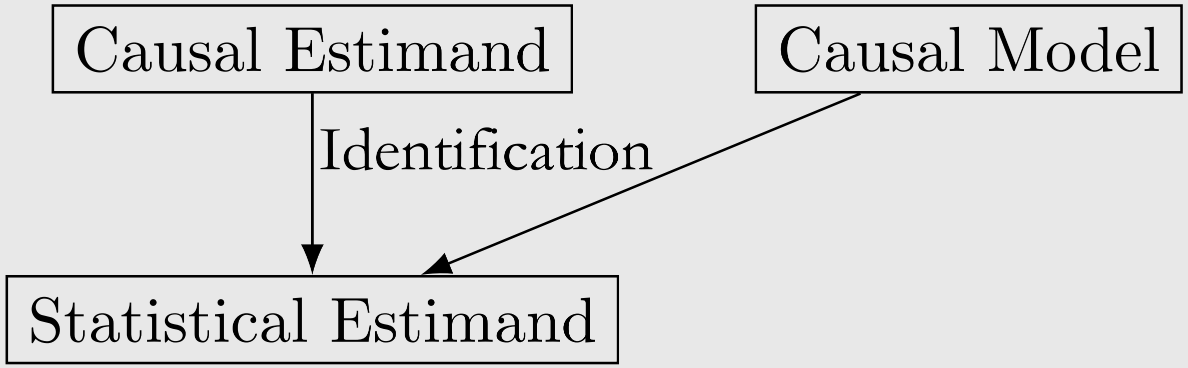
Observational	Interventional
$P(Y, T, X)$	$P(Y \mid do(T = t))$
$P(Y \mid T = t)$	$P(Y \mid do(T = t), X = x)$

Identifiability



Identifiability

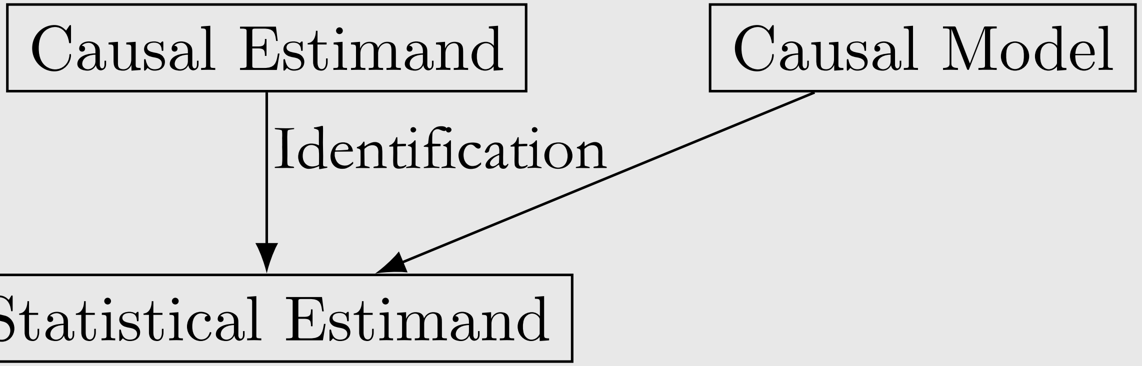
$P(y \mid do(t))$



Identifiability

$P(y \mid do(t))$

$P(y \mid t)$

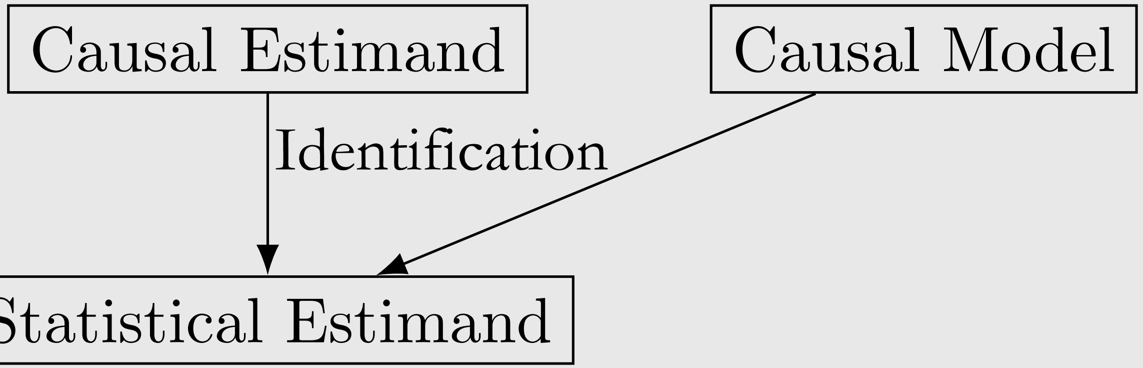


Identifiability

Accessible via experiment

$P(y \mid do(t))$

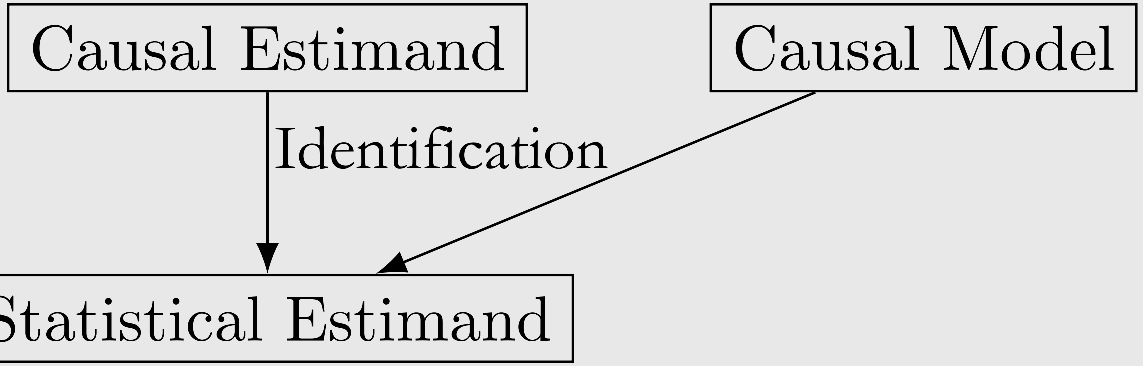
$P(y \mid t)$



Identifiability

Accessible via experiment

$P(y \mid do(t))$



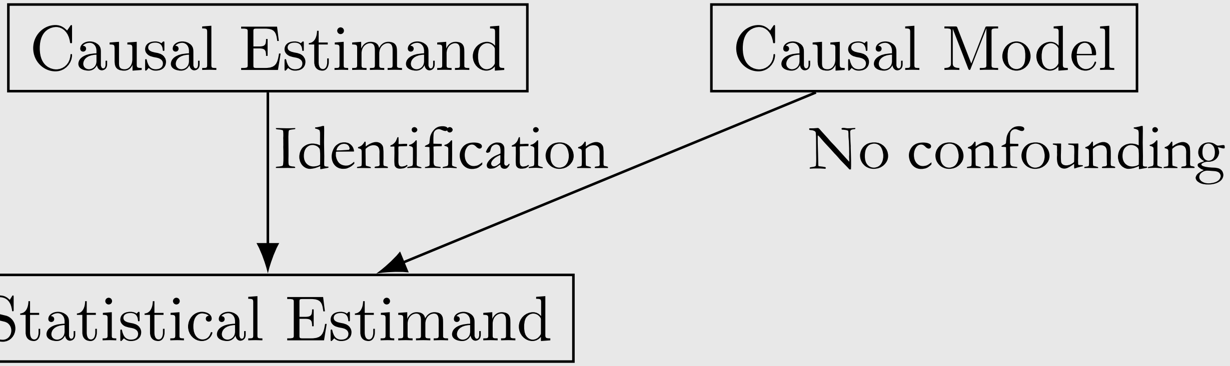
$P(y \mid t)$

Accessible via observational data

Identifiability

Accessible via experiment

$P(y \mid do(t))$



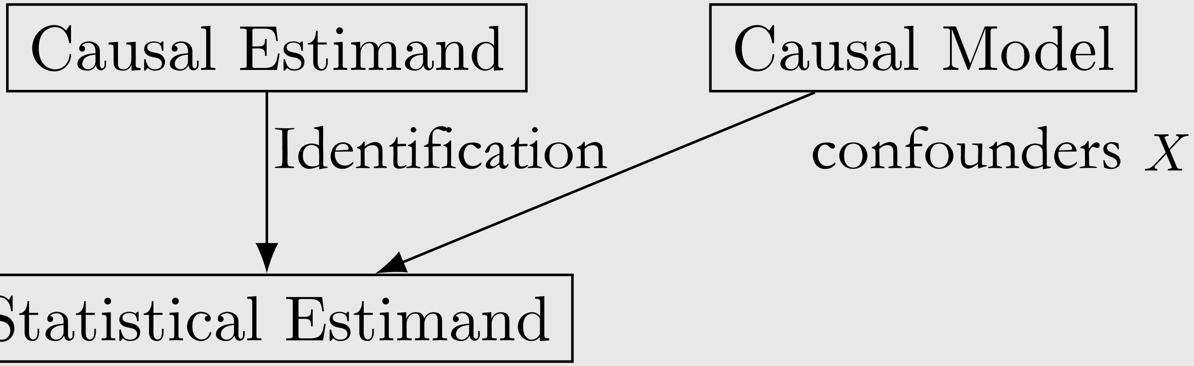
$P(y \mid t)$

Accessible via observational data

Identifiability

Accessible via experiment

$P(y \mid do(t))$



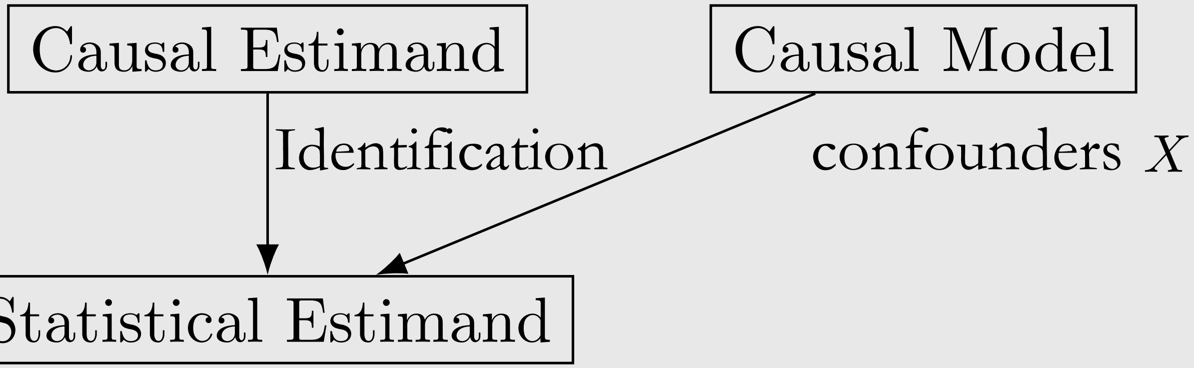
$P(y \mid t)$

Accessible via observational data

Identifiability

Accessible via experiment

$P(y \mid do(t))$



$\mathbb{E}_X[P(y \mid t, X)]$

Accessible via observational data

The *do*-operator

Main assumption: modularity

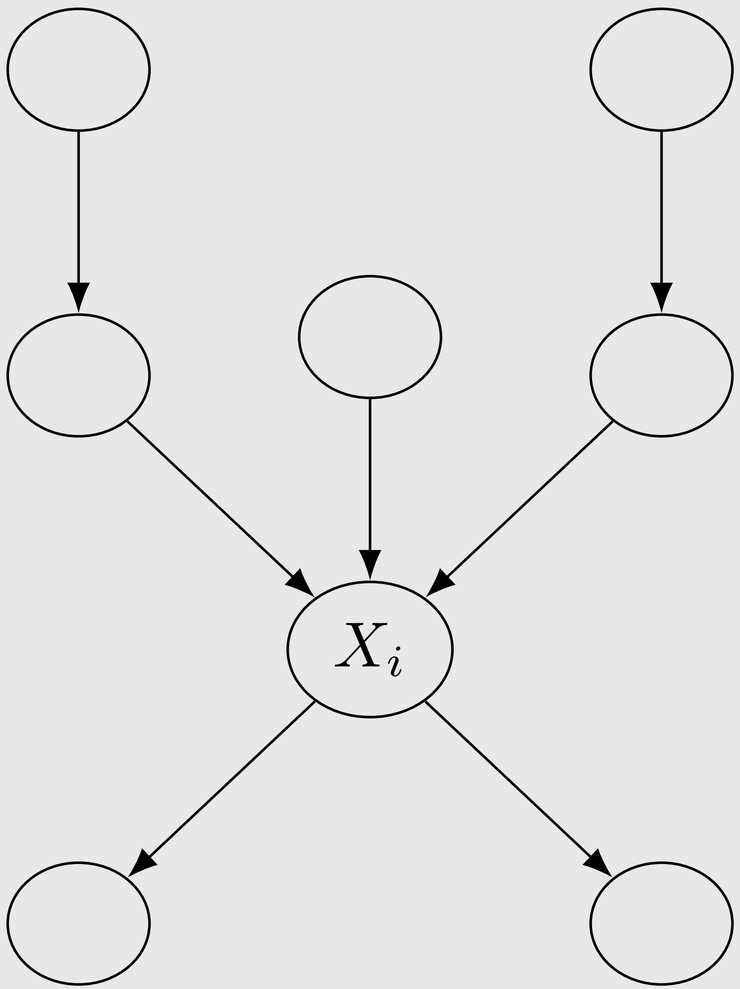
Backdoor adjustment

Structural causal models

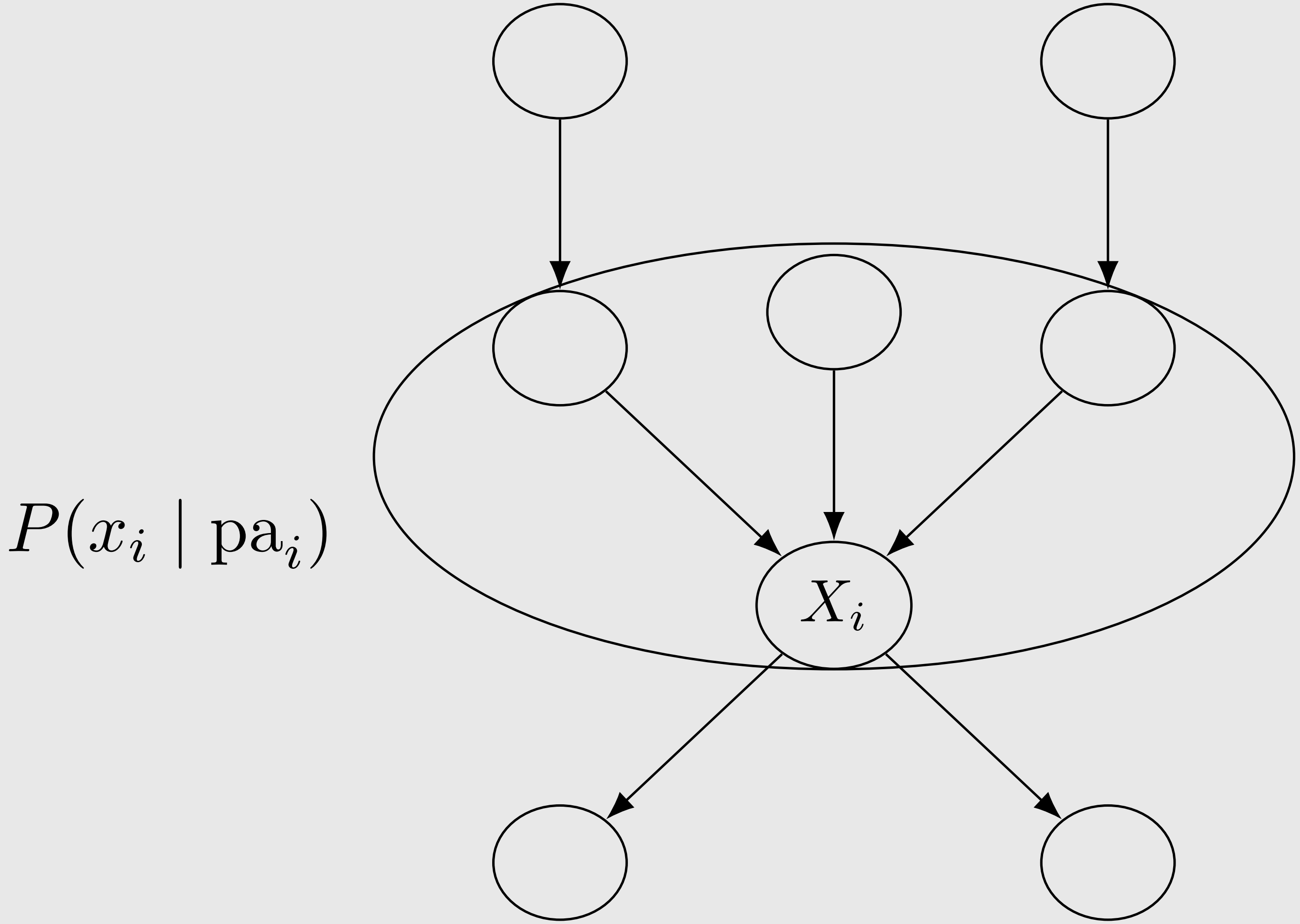
A complete example with estimation

Causal mechanism

$P(x_i \mid \text{pa}_i)$



Causal mechanism



Modularity assumption

Modularity assumption

If we intervene on a node X_i, then only the mechanism $P(x_i \mid \mathrm{pa}_i)$ changes. All other mechanisms $P(x_j \mid \mathrm{pa}_j)$ where $i \neq j$ remain unchanged.

Modularity assumption

If we intervene on a node X_i, then only the mechanism $P(x_i \mid \mathrm{pa}_i)$ changes. All other mechanisms $P(x_j \mid \mathrm{pa}_j)$ where $i \neq j$ remain unchanged.

In other words, the causal mechanisms are **modular**.

Modularity assumption

If we intervene on a node X_i, then only the mechanism $P(x_i \mid \mathrm{pa}_i)$ changes. All other mechanisms $P(x_j \mid \mathrm{pa}_j)$ where $i \neq j$ remain unchanged.

In other words, the causal mechanisms are **modular**.

Many names: independent mechanisms, autonomy, invariance, etc.

Modularity assumption: more formal

Modularity assumption: more formal

If we intervene on a set of nodes $S \subseteq [n]$, setting them to constants, then for all i, we have the following:

Modularity assumption: more formal

If we intervene on a set of nodes $S \subseteq [n]$, setting them to constants, then for all i, we have the following:

1. If $i \notin S$, then $P(x_i \mid \mathrm{pa}_i)$ remains unchanged.

Modularity assumption: more formal

If we intervene on a set of nodes $S \subseteq [n]$, setting them to constants, then for all i, we have the following:

1. If $i \notin S$, then $P(x_i \mid \mathrm{pa}_i)$ remains unchanged.
2. If $i \in S$, then $P(x_i \mid \mathrm{pa}_i) = 1$ if x_i is the value that X_i was set to by the intervention; otherwise, $P(x_i \mid \mathrm{pa}_i) = 0$.

Modularity assumption: more formal

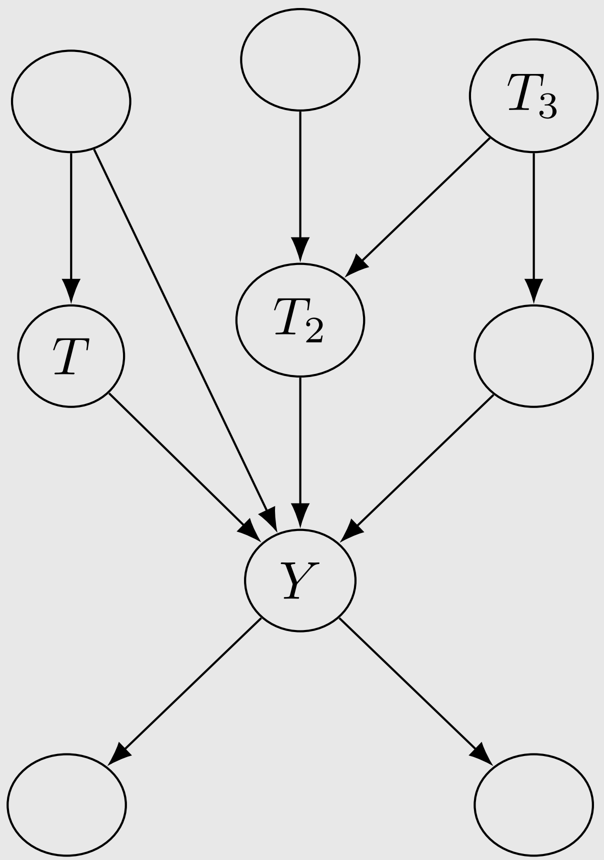
If we intervene on a set of nodes $S \subseteq [n]$, setting them to constants, then for all i, we have the following:

1. If $i \notin S$, then $P(x_i \mid \text{pa}_i)$ remains unchanged.
2. If $i \in S$, then $P(x_i \mid \text{pa}_i) = 1$ if x_i is the value that X_i was set to by the intervention; otherwise, $P(x_i \mid \text{pa}_i) = 0$.

consistent with the intervention

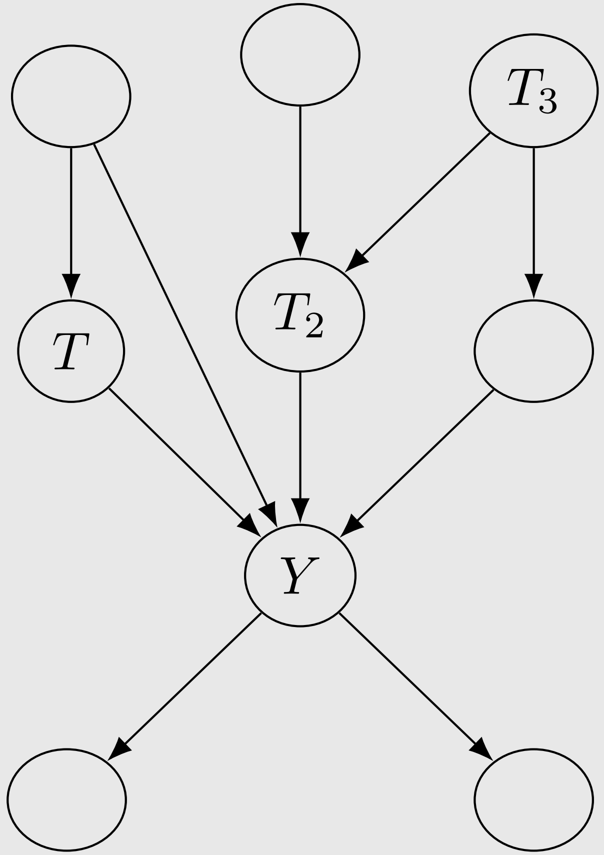
Manipulated graphs

Observational data

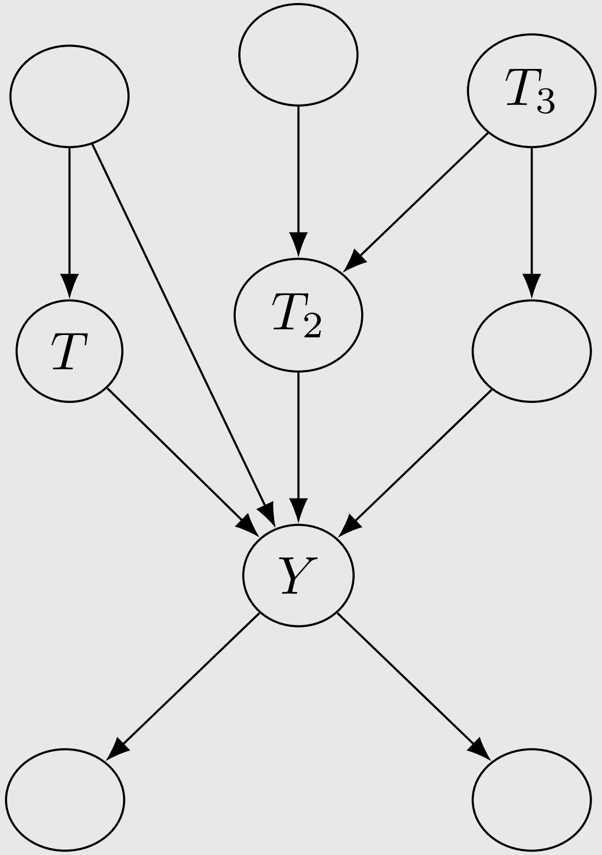


Manipulated graphs

Observational data

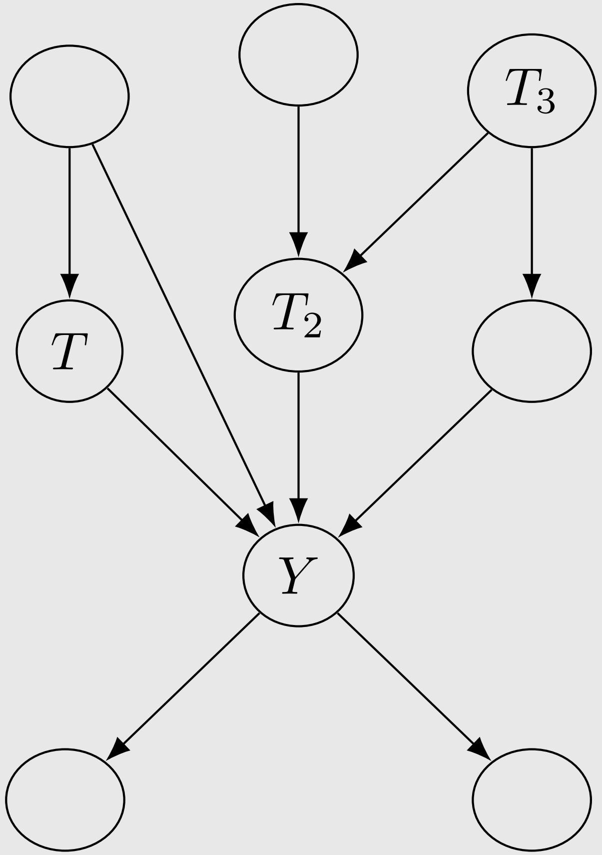


Interventional data

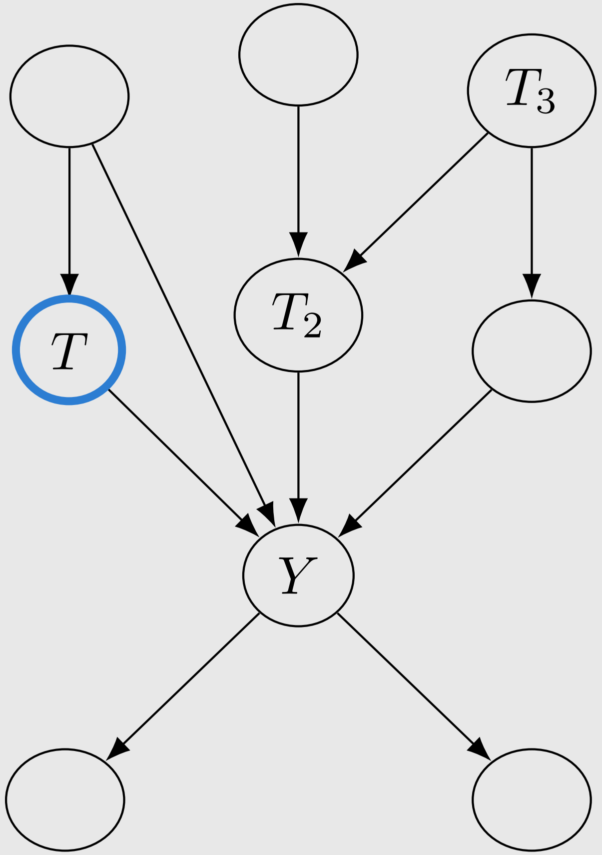


Manipulated graphs

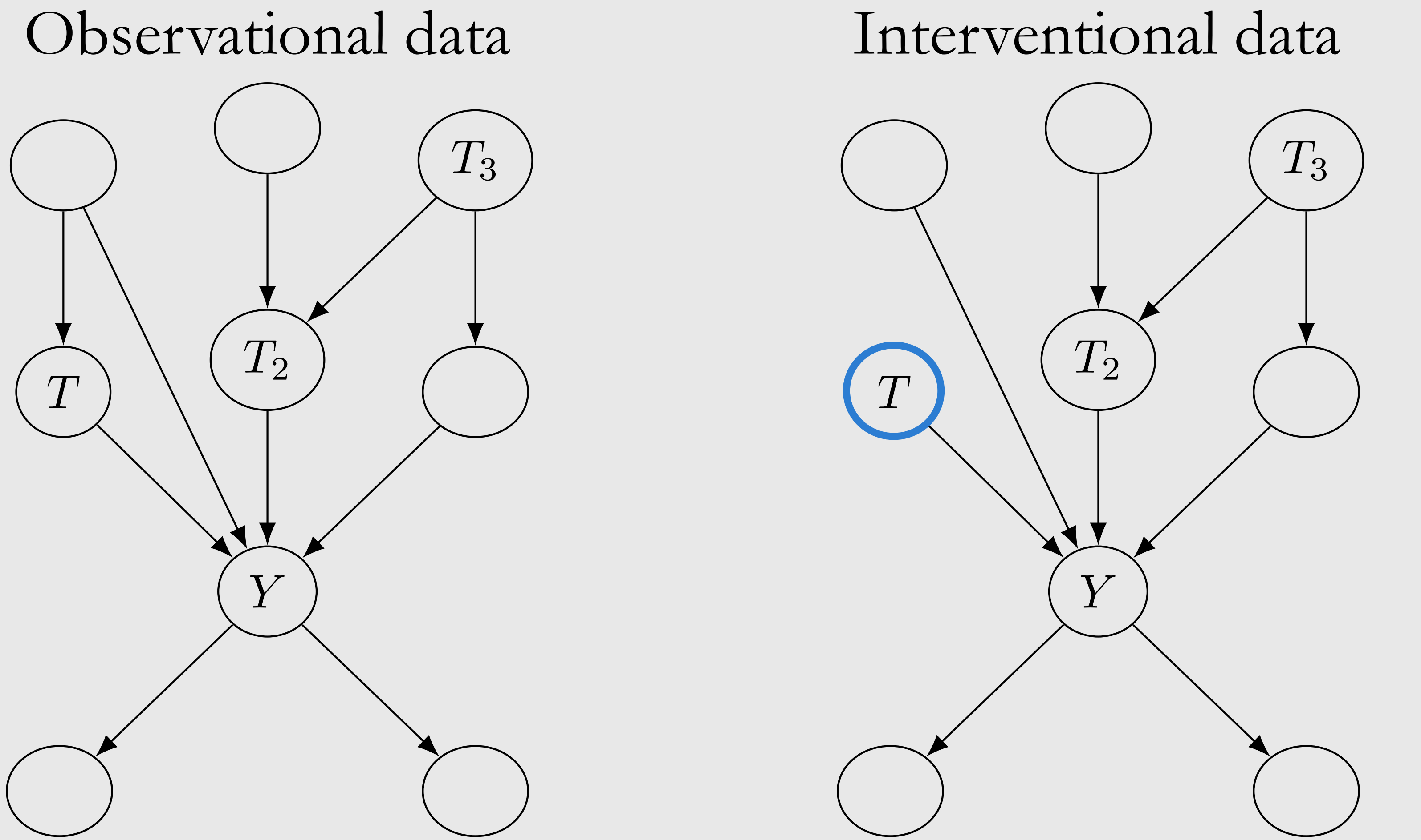
Observational data



Interventional data

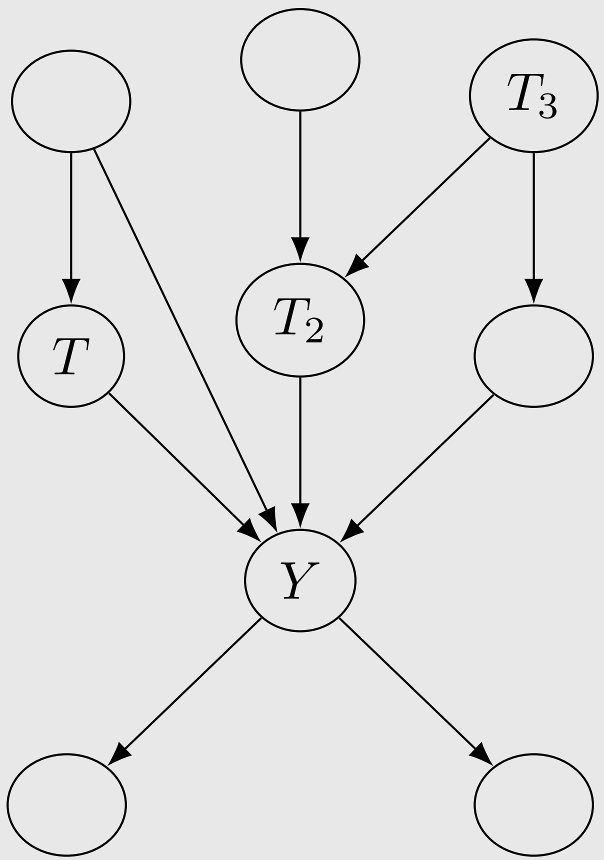


Manipulated graphs

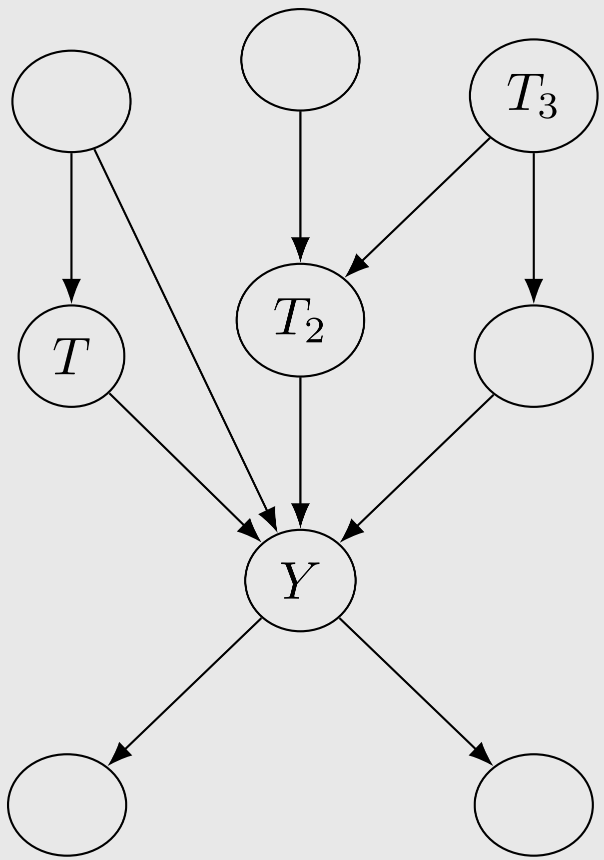


Manipulated graphs

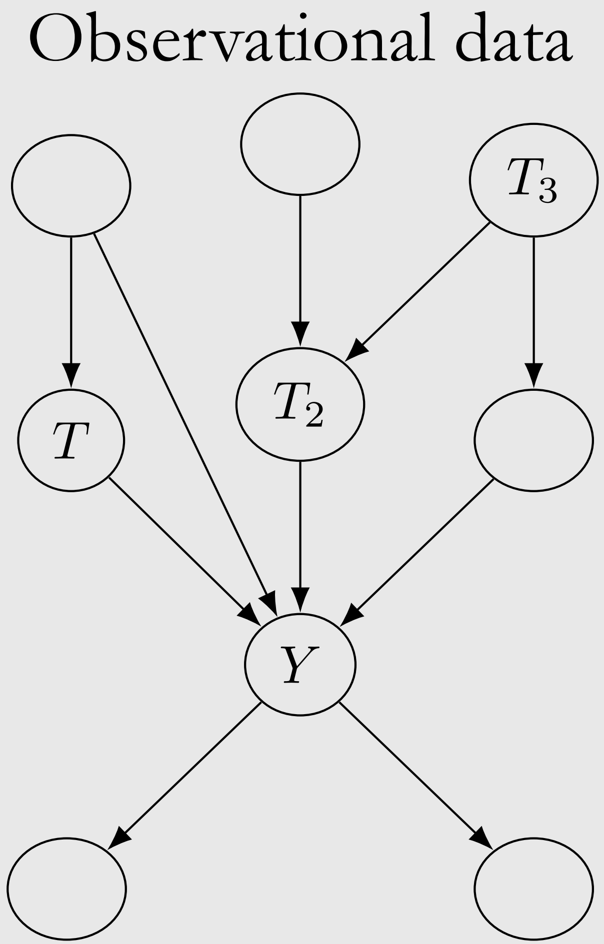
Observational data

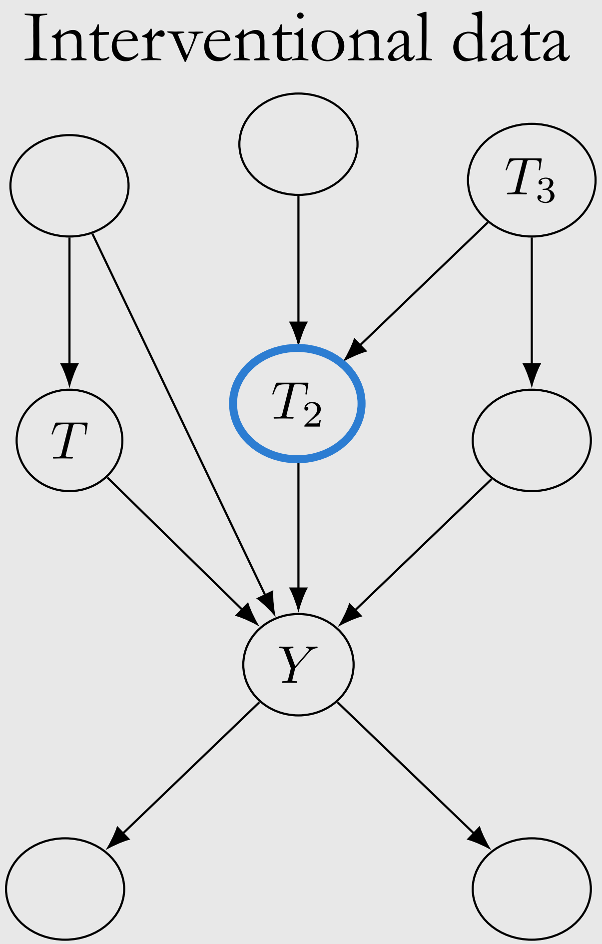


Interventional data

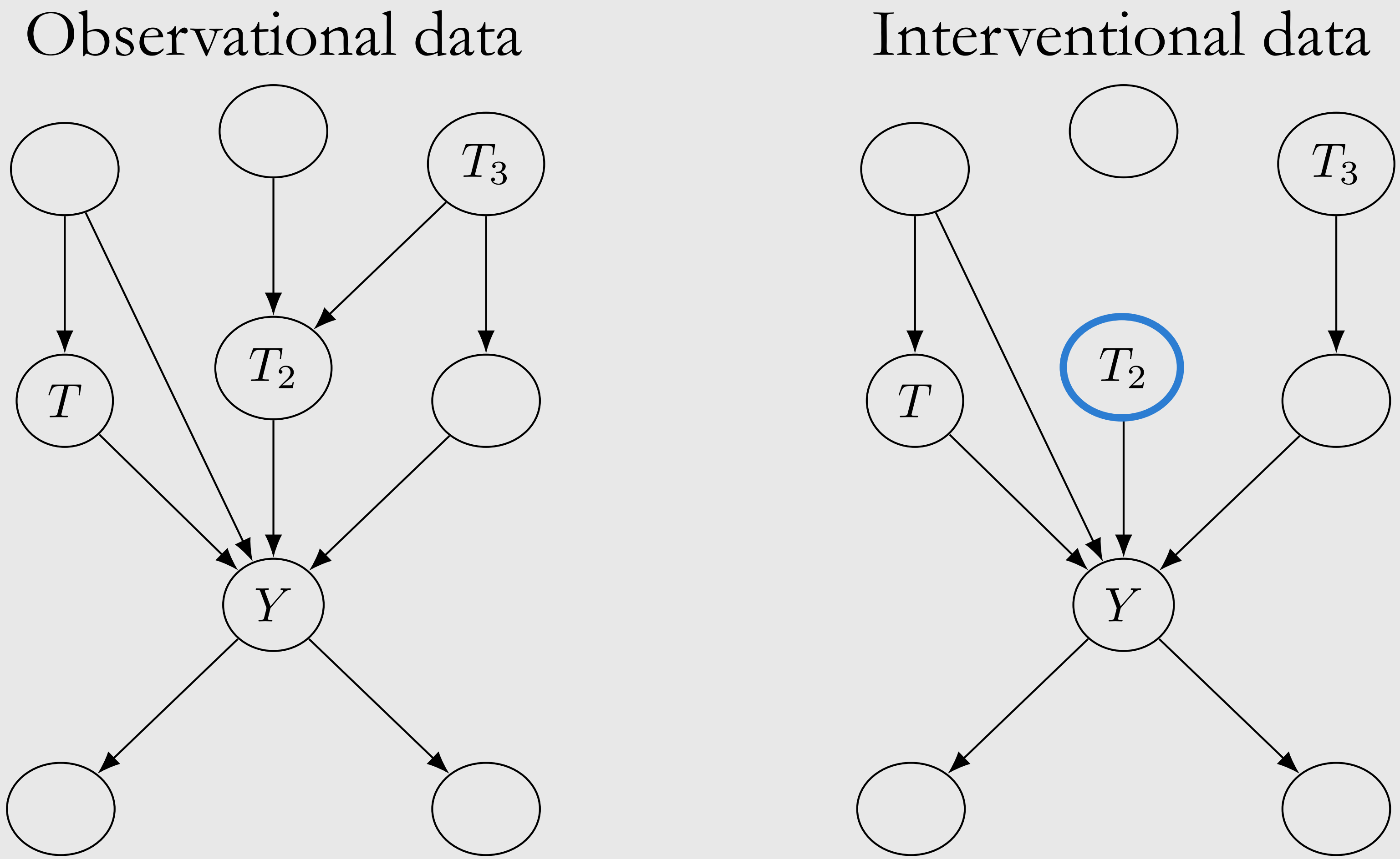


Manipulated graphs

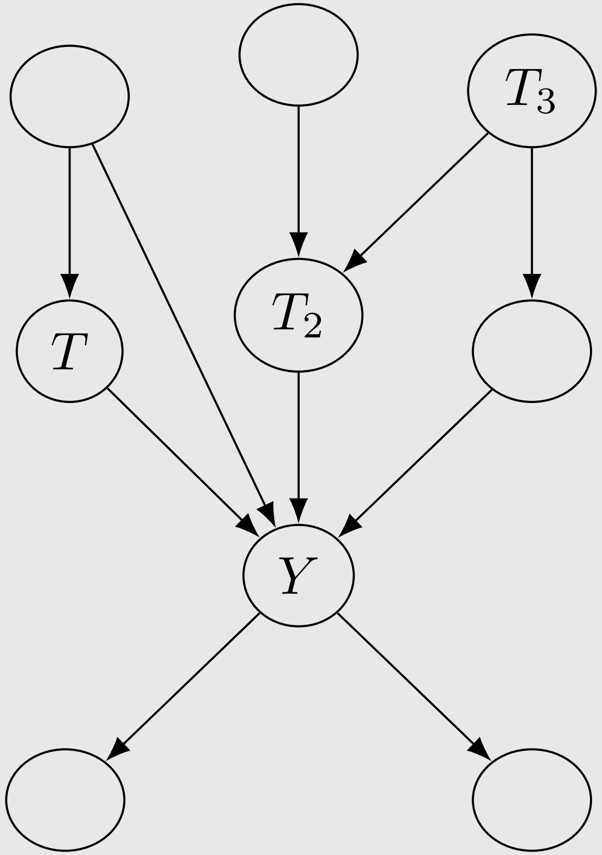




Manipulated graphs

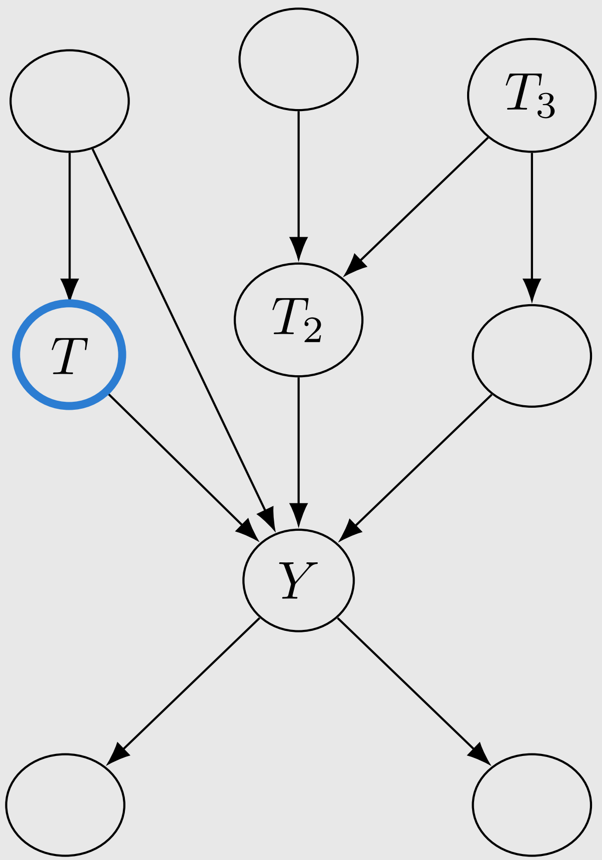


What would it mean if modularity is violated?



What would it mean if modularity is violated?

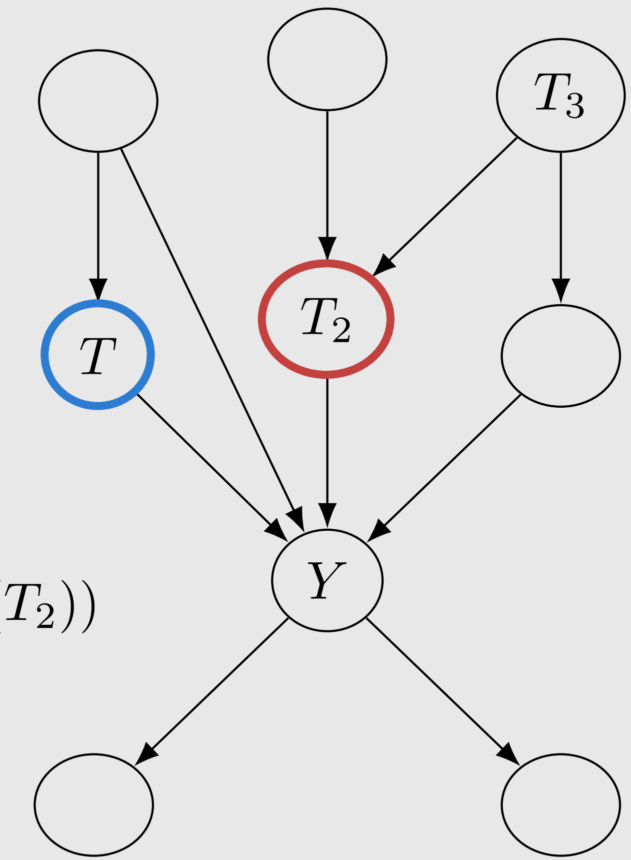
Intervention on T not only changes $P(T \mid \mathrm{pa}(T))$



What would it mean if modularity is violated?

Intervention on T not only changes $P(T \mid \text{pa}(T))$

but also changes other mechanisms such as $P(T_2 \mid \text{pa}(T_2))$



Truncated factorization

Recall the Bayesian network factorization:

$$P(x_1, \ldots, x_n) = \prod_i P(x_i \mid \mathrm{pa}_i)$$

Truncated factorization

Truncated factorization:

$$P(x_1, \ldots, x_n \mid do(S = s)) = \prod_i P(x_i \mid \text{pa}_i)$$

Truncated factorization

Truncated factorization:

$$P(x_1, \ldots, x_n \mid do(S = s)) = \prod_{i \notin S} P(x_i \mid \text{pa}_i)$$

Truncated factorization

Truncated factorization:

$$P(x_1, \ldots, x_n \mid do(S = s)) = \prod_{i \notin S} P(x_i \mid \text{pa}_i)$$

if x is consistent with the intervention.

Truncated factorization

Truncated factorization:

$$P(x_1, \dots, x_n \mid do(S = s)) = \prod_{i \notin S} P(x_i \mid \text{pa}_i)$$

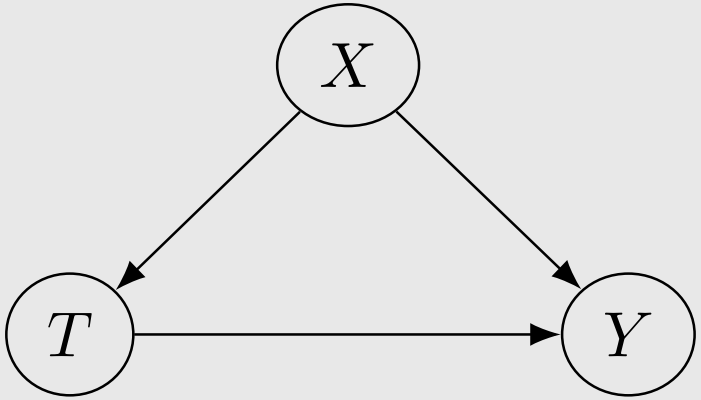
if x is consistent with the intervention.

Otherwise,

$$P(x_1, \dots, x_n \mid do(S = s)) = 0$$

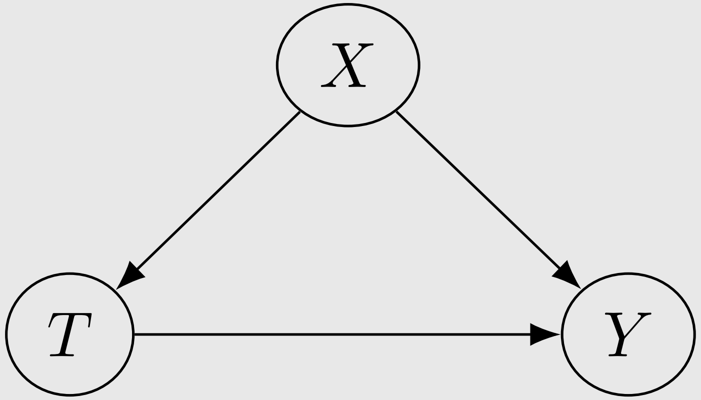
Simple identification via truncated factorization

Goal: identify $P(y \mid do(t))$



Simple identification via truncated factorization

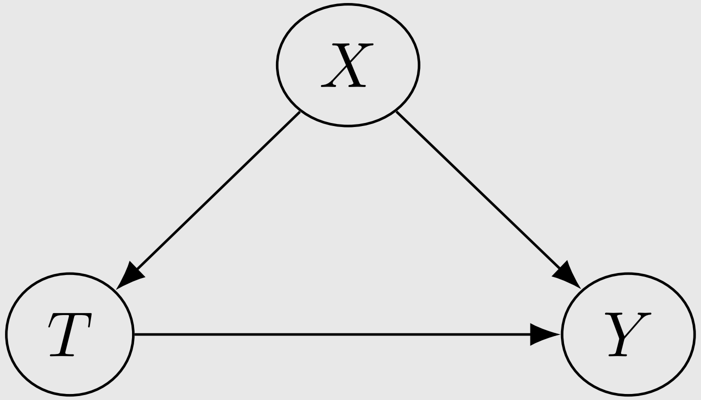
Goal: identify $P(y \mid do(t))$



Bayesian network factorization: $P(y, t, x) = P(x)\, P(t \mid x)\, P(y \mid t, x)$

Simple identification via truncated factorization

Goal: identify $P(y \mid do(t))$

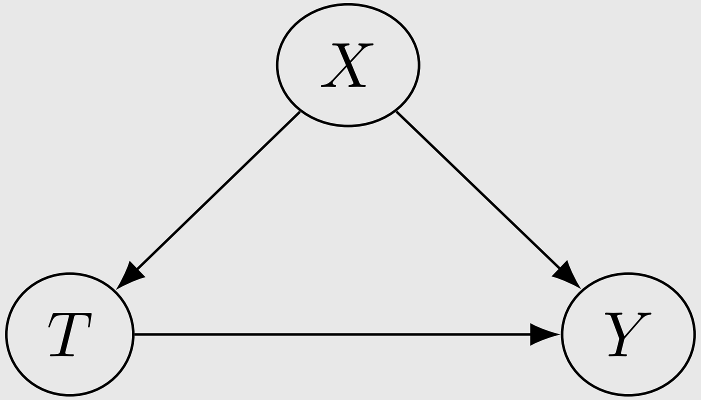


Bayesian network factorization: $P(y, t, x) = P(x)\, P(t \mid x)\, P(y \mid t, x)$

Truncated factorization: $P(y, x \mid do(t)) = P(x)\, P(y \mid t, x)$

Simple identification via truncated factorization

Goal: identify $P(y \mid do(t))$



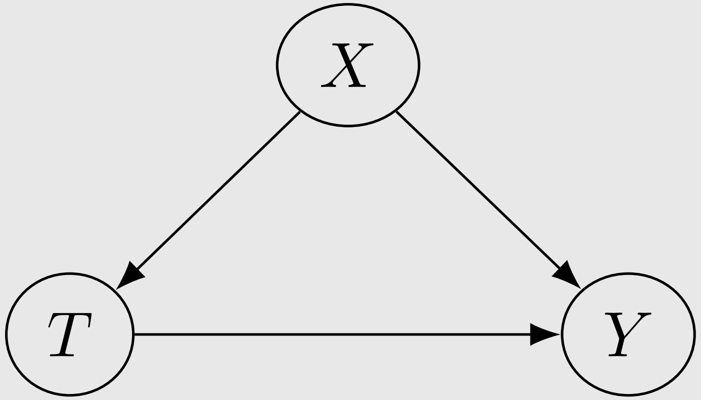
Bayesian network factorization: $P(y, t, x) = P(x)\, P(t \mid x)\, P(y \mid t, x)$

Truncated factorization: $P(y, x \mid do(t)) = P(x)\, P(y \mid t, x)$

Marginalize: $P(y \mid do(t)) = \sum_x P(y \mid t, x)\, P(x)$

Association vs. causation revisited

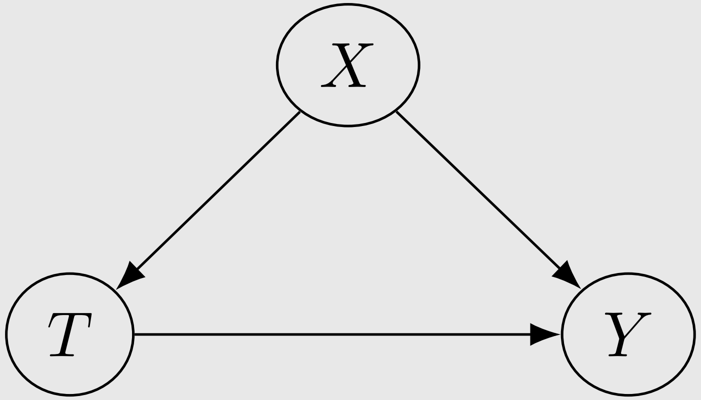
$$P(y \mid do(t)) = \sum_x P(y \mid t, x)\, P(x)$$



Association vs. causation revisited

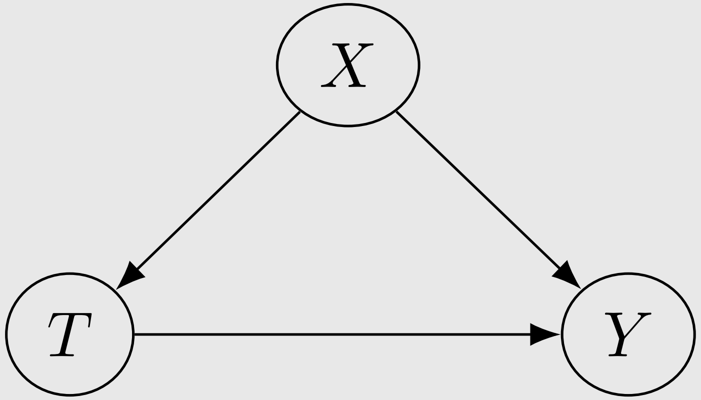
$$P(y \mid do(t)) = \sum_x P(y \mid t, x)\, P(x)$$

$$P(y \mid do(t)) \neq P(y \mid t)$$



Association vs. causation revisited

$$P(y \mid do(t)) = \sum_x P(y \mid t, x)\, P(x)$$



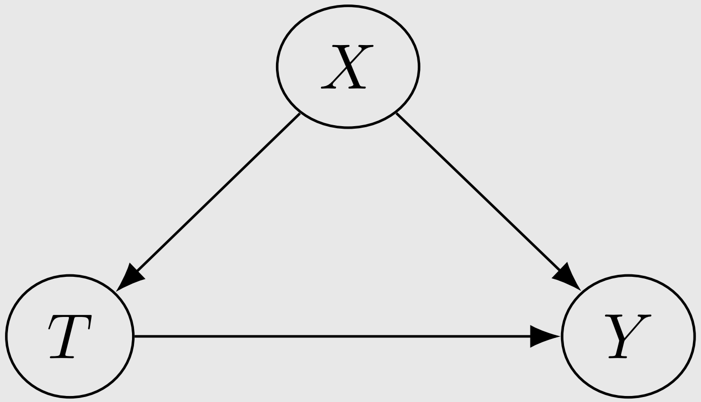
$$P(y \mid do(t)) \neq P(y \mid t)$$

$$\sum_x P(y \mid t, x)\, P(x)$$

Association vs. causation revisited

$$P(y \mid do(t)) = \sum_x P(y \mid t, x)\, P(x)$$

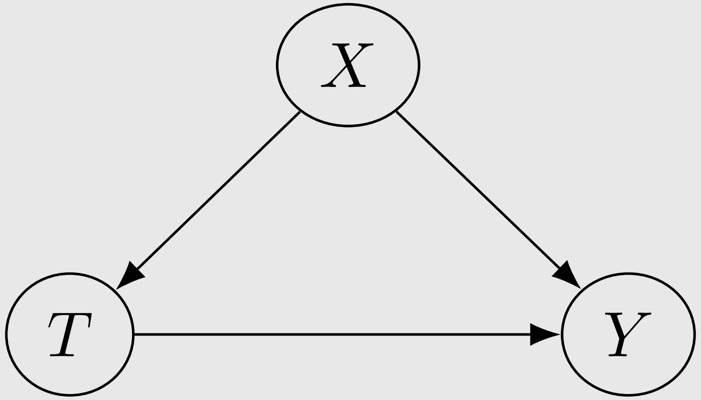
$$P(y \mid do(t)) \neq P(y \mid t)$$



$$\sum_x P(y \mid t, x)\, P(x \mid t)$$

Association vs. causation revisited

$$P(y \mid do(t)) = \sum_x P(y \mid t, x)\, P(x)$$



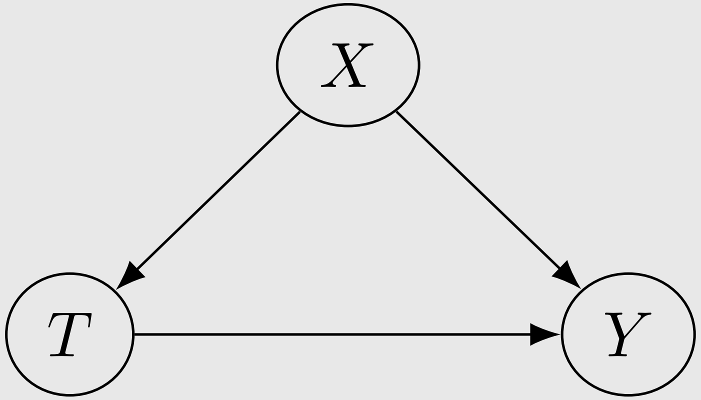
$$P(y \mid do(t)) \neq P(y \mid t)$$

$$\sum_x P(y \mid t, x)\, P(x \mid t) = \sum_x P(y, x \mid t)$$

Association vs. causation revisited

$$P(y \mid do(t)) = \sum_x P(y \mid t, x)\, P(x)$$

$$P(y \mid do(t)) \neq P(y \mid t)$$

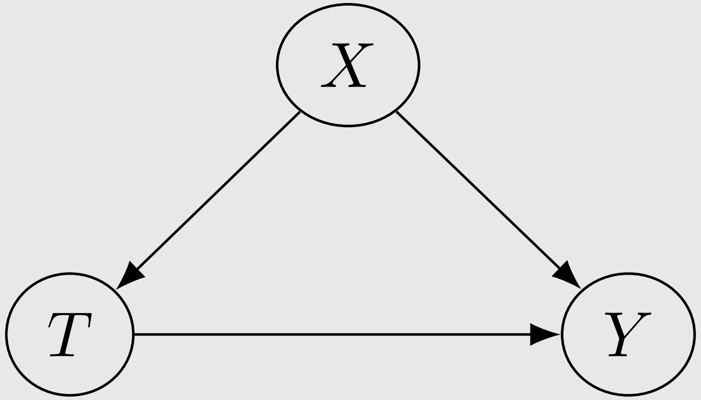


$$\sum_x P(y \mid t, x)\, P(x \mid t) = \sum_x P(y, x \mid t)$$
$$= P(y \mid t)$$

Association vs. causation revisited

$$\underline{P(y \mid do(t))} = \sum_x P(y \mid t, x)\, P(x)$$

$$\underline{P(y \mid do(t))} \neq P(y \mid t)$$

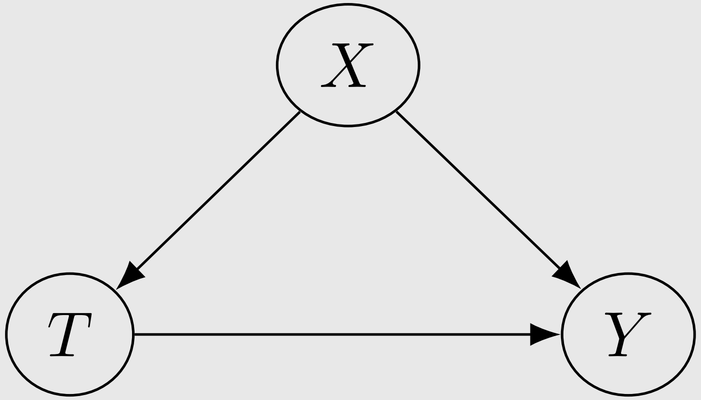


$$\sum_x P(y \mid t, x)\, P(x \mid t) = \sum_x P(y, x \mid t)$$
$$= P(y \mid t)$$

Association vs. causation revisited

$$\underline{P(y \mid do(t))} = \sum_x P(y \mid t, x)\, P(x)$$

$$\underline{P(y \mid do(t))} \neq \underline{P(y \mid t)}$$

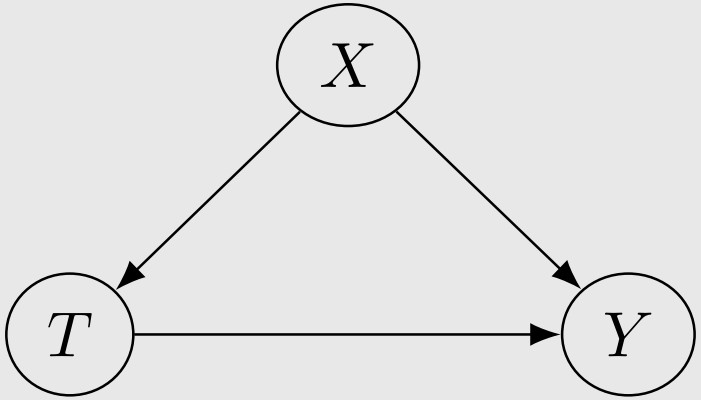


$$\sum_x P(y \mid t, x)\, P(x \mid t) = \sum_x P(y, x \mid t)$$
$$= \underline{P(y \mid t)}$$

Association vs. causation revisited

$$\underline{P(y \mid do(t))} = \sum_x P(y \mid t, x)\, \underline{P(x)}$$

$$\underline{P(y \mid do(t))} \neq \underline{P(y \mid t)}$$

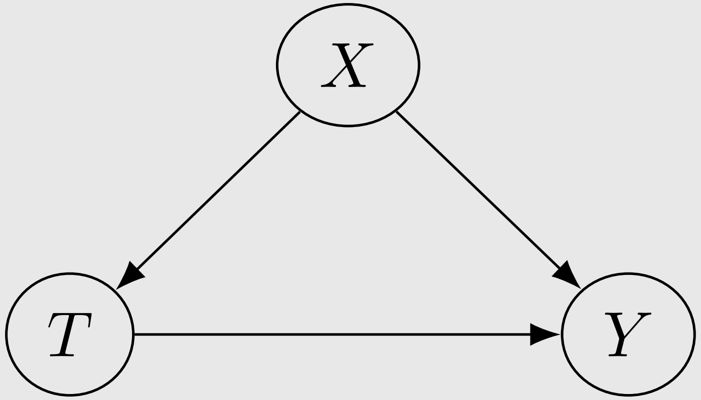


$$\sum_x P(y \mid t, x)\, P(x \mid t) = \sum_x P(y, x \mid t)$$

$$= \underline{P(y \mid t)}$$

Association vs. causation revisited

$$\underline{P(y \mid do(t))} = \sum_x P(y \mid t, x)\, \underline{P(x)}$$



$$\underline{P(y \mid do(t))} \neq \underline{P(y \mid t)}$$

$$\sum_x P(y \mid t, x)\, \underline{P(x \mid t)} = \sum_x P(y, x \mid t)$$

$$= \underline{P(y \mid t)}$$

The *do*-operator

Main assumption: modularity

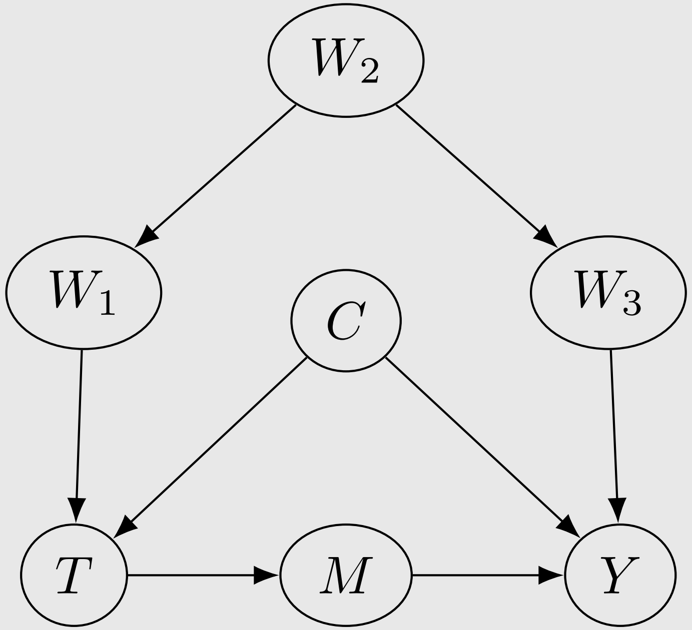
Backdoor adjustment

Structural causal models

A complete example with estimation

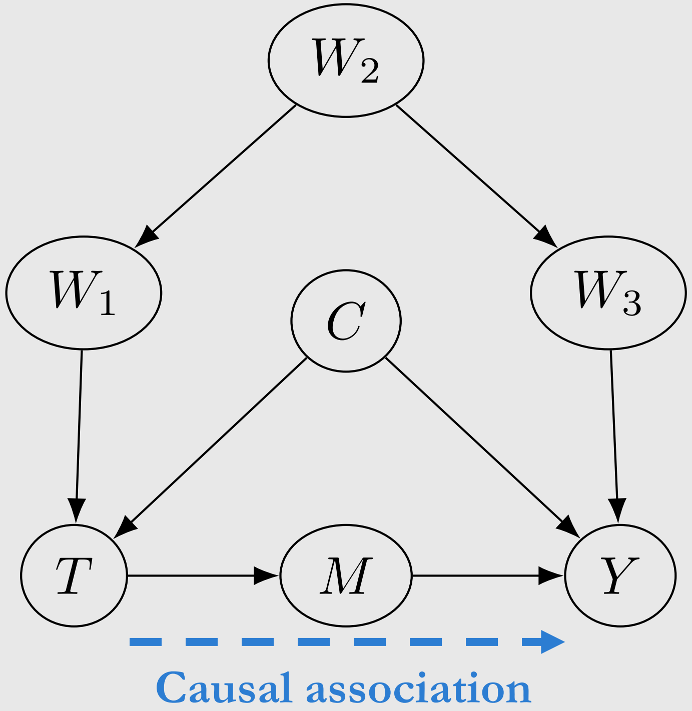
Blocking backdoor paths

$P(Y \mid t)$



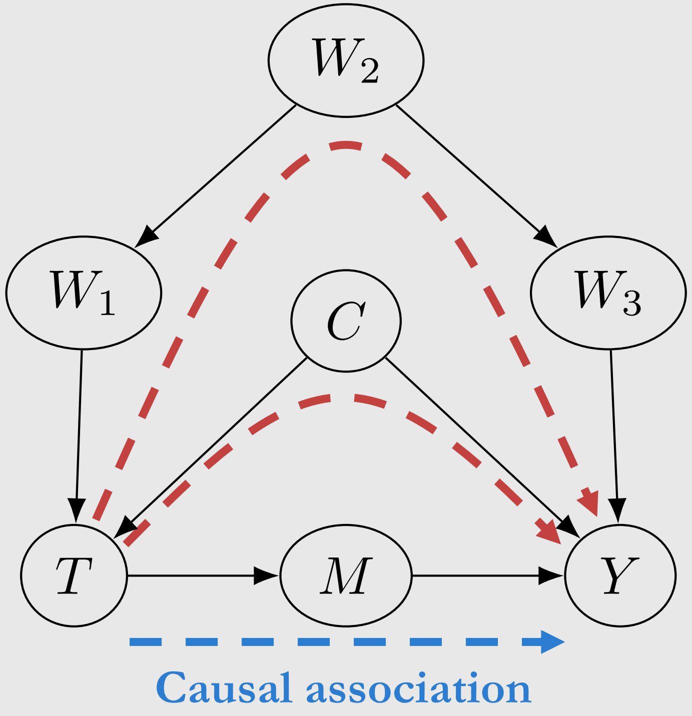
Blocking backdoor paths

$P(Y \mid t)$



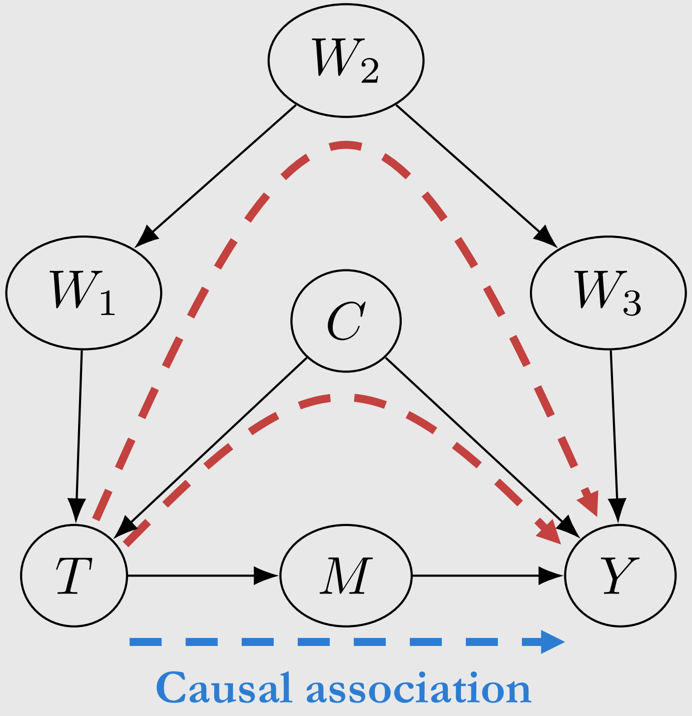
Blocking backdoor paths

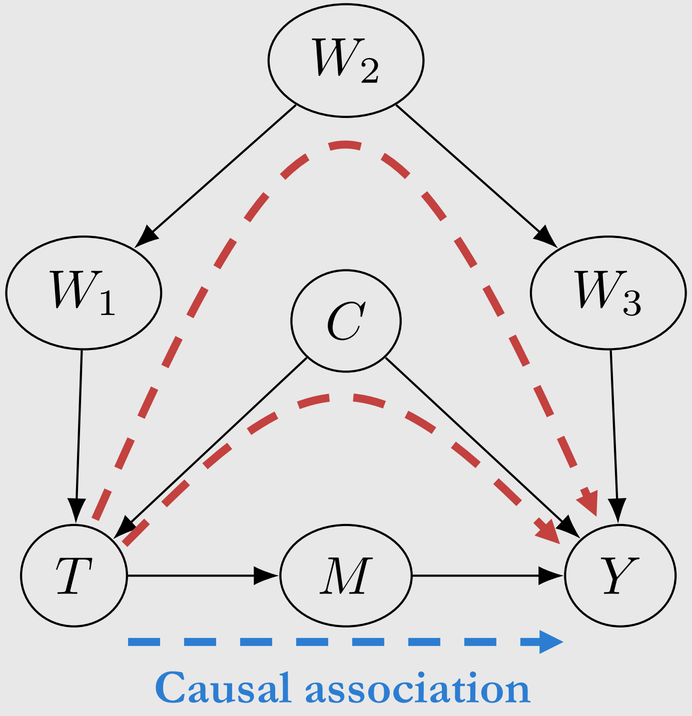
$P(Y \mid t)$



Blocking backdoor paths

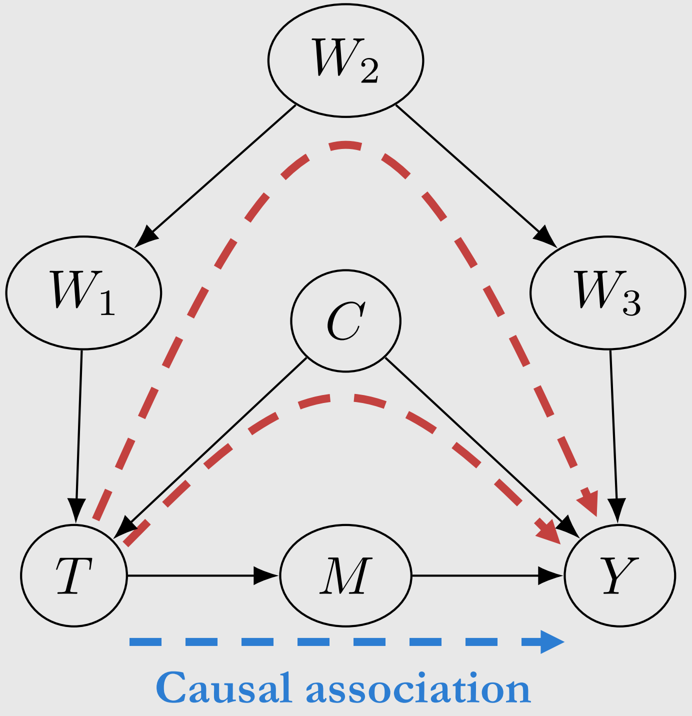
$P(Y \mid t)$



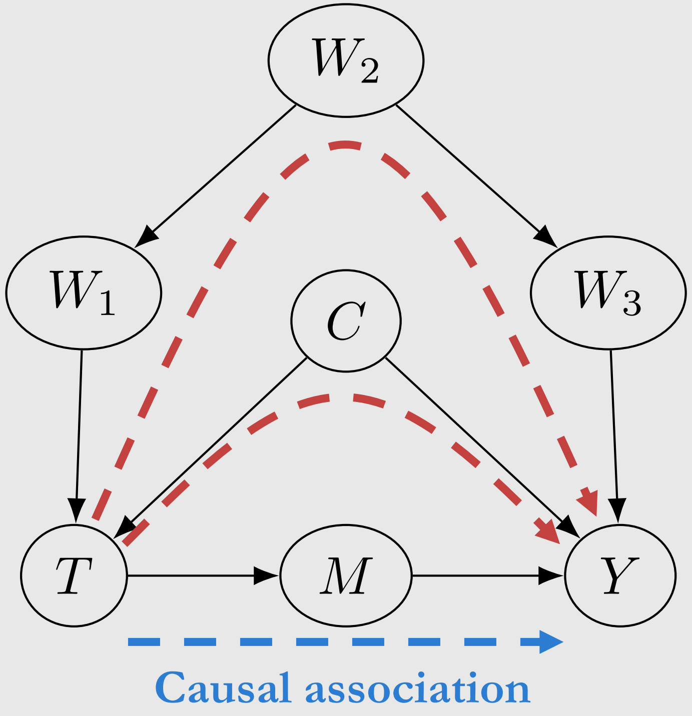


Blocking backdoor paths

$P(Y \mid t)$

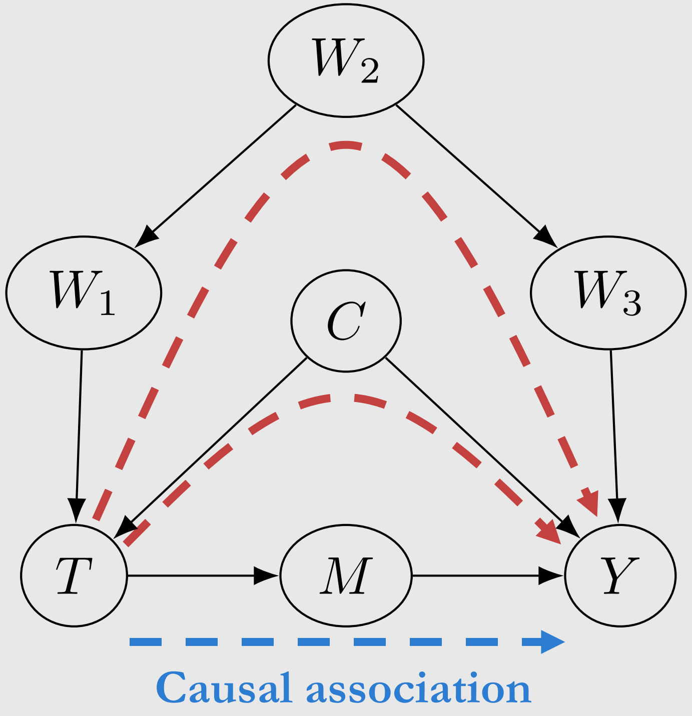


$P(Y \mid do(t))$

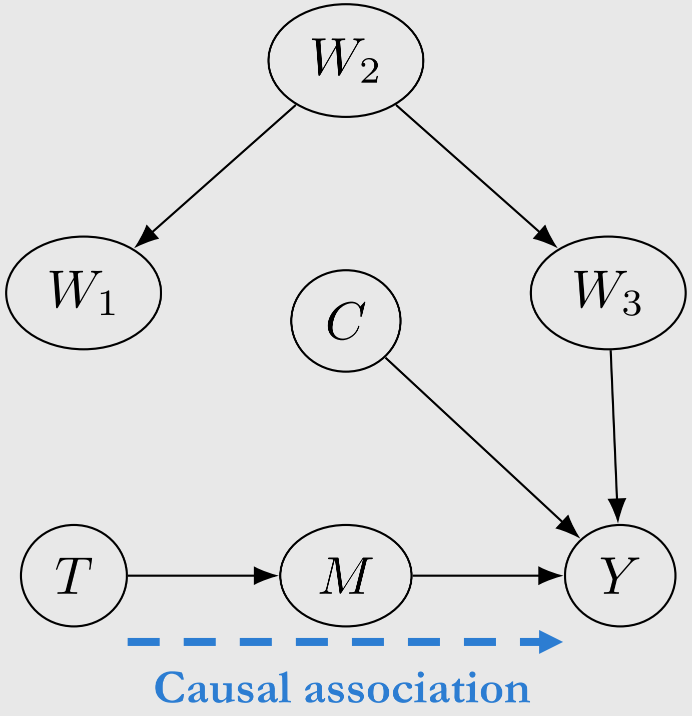


Blocking backdoor paths

$P(Y \mid t)$

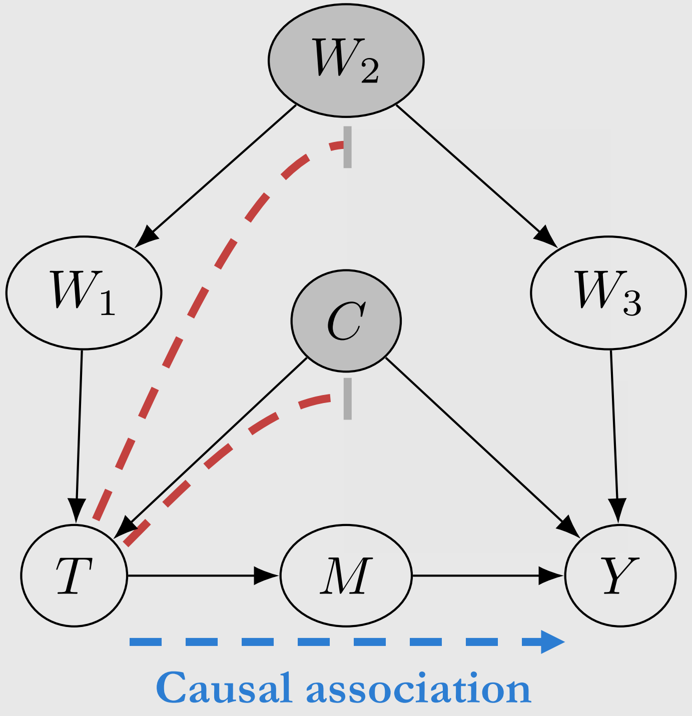


$P(Y \mid do(t))$

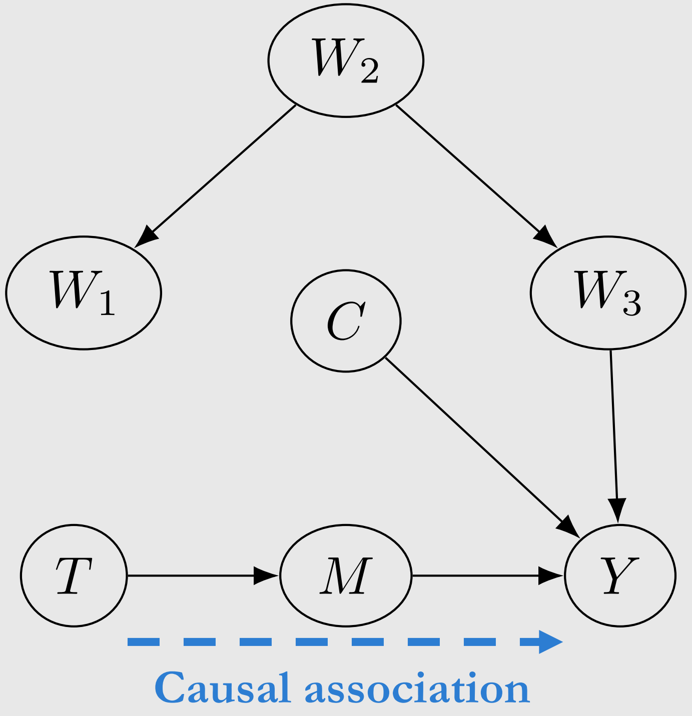


Blocking backdoor paths

$P(Y \mid t, c, w_2)$

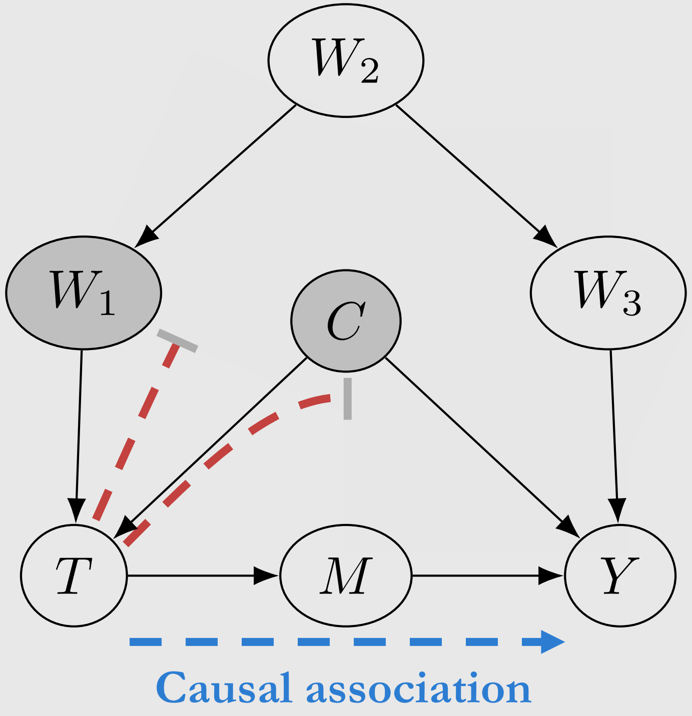


$P(Y \mid do(t))$

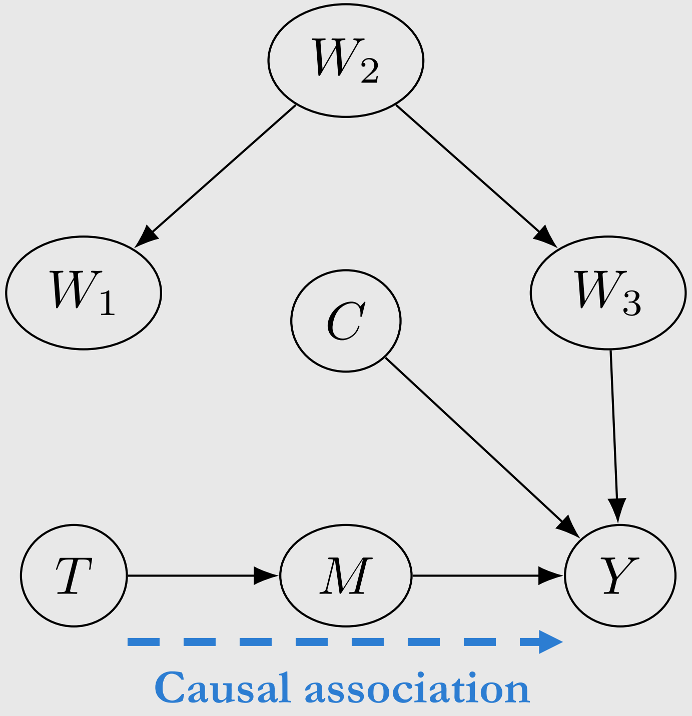


Blocking backdoor paths

$$P(Y \mid t, c, w_1)$$

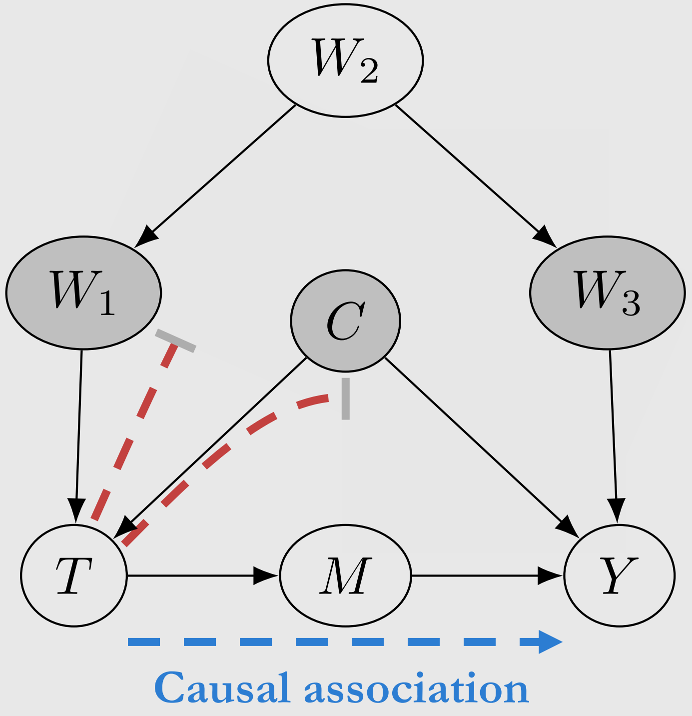


$$P(Y \mid do(t))$$

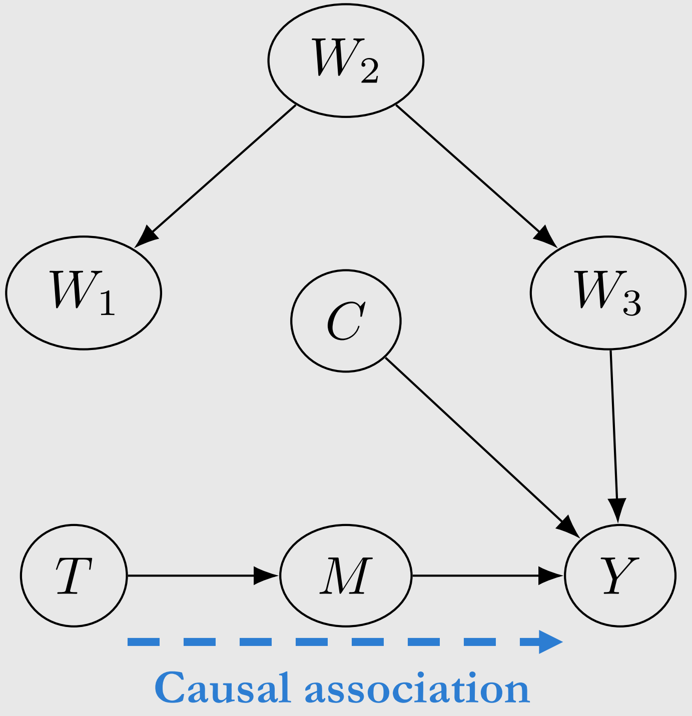


Blocking backdoor paths

$P(Y \mid t, c, w_1, w_3)$



$P(Y \mid do(t))$



Backdoor criterion and backdoor adjustment

Backdoor criterion and backdoor adjustment

A set of variables W satisfies the backdoor criterion relative to T and Y if the following are true:

Backdoor criterion and backdoor adjustment

A set of variables W satisfies the backdoor criterion relative to T and Y if the following are true:

1. W blocks all backdoor paths from T to Y
2.

Backdoor criterion and backdoor adjustment

A set of variables W satisfies the backdoor criterion relative to T and Y if the following are true:

1. W blocks all backdoor paths from T to Y
2. W does not contain any descendants of T

Backdoor criterion and backdoor adjustment

A set of variables W satisfies the backdoor criterion relative to T and Y if the following are true:

1. W blocks all backdoor paths from T to Y
2. W does not contain any descendants of T

Given the modularity assumption and that W satisfies the backdoor criterion, we can identify the causal effect of T on Y:

$$P(y \mid do(t)) = \sum_{w} P(y \mid t, w)\, P(w)$$

Proof of backdoor adjustment

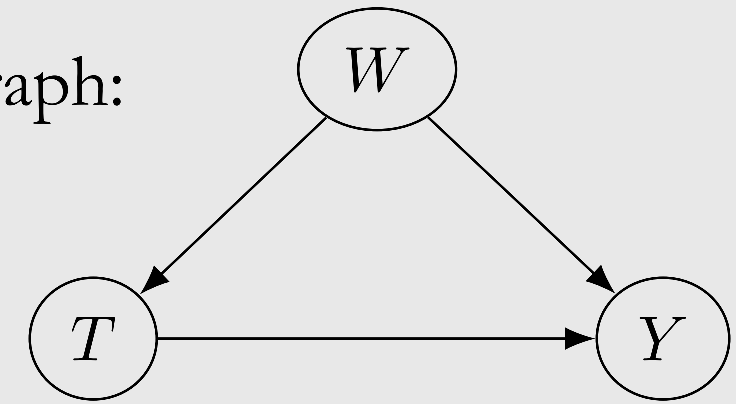
$$P(y \mid do(t))$$

$$= \sum_w P(y \mid t, w) \, P(w)$$

Proof of backdoor adjustment

$$P(y \mid do(t))$$

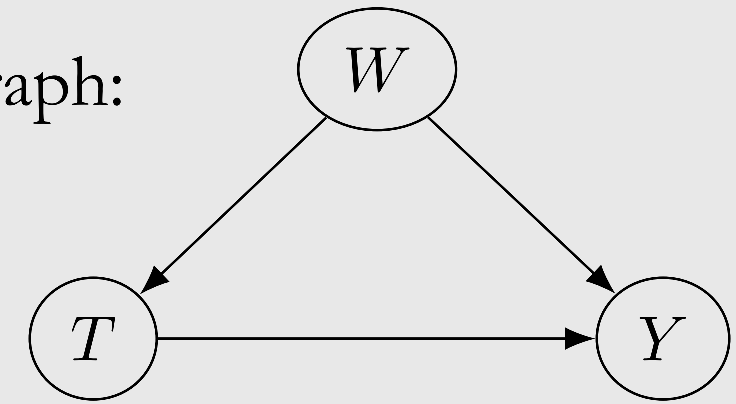
Example graph:



Proof of backdoor adjustment

$$P(y \mid do(t)) = \sum_w P(y \mid do(t), w) \, P(w \mid do(t))$$

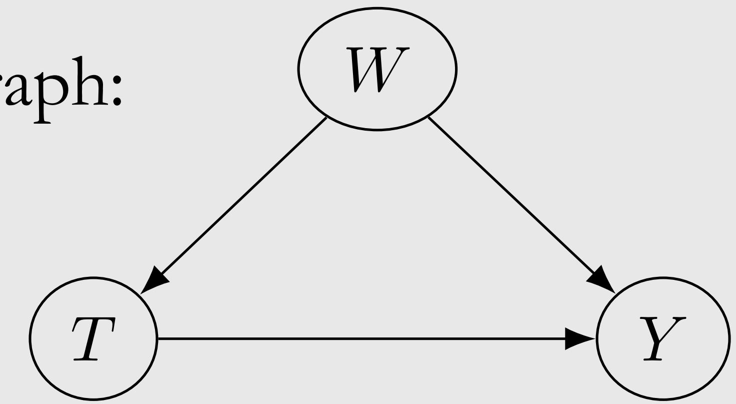
Example graph:



Proof of backdoor adjustment

$$P(y \mid do(t)) = \sum_w P(y \mid do(t), w)\, P(w \mid do(t))$$
$$= \sum_w P(y \mid t, w)\, P(w \mid do(t))$$

Example graph:



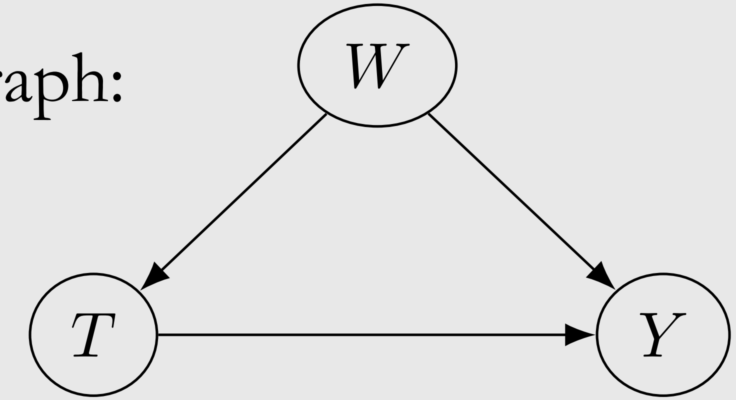
Proof of backdoor adjustment

$$P(y \mid do(t)) = \sum_w P(y \mid do(t), w)\, P(w \mid do(t))$$

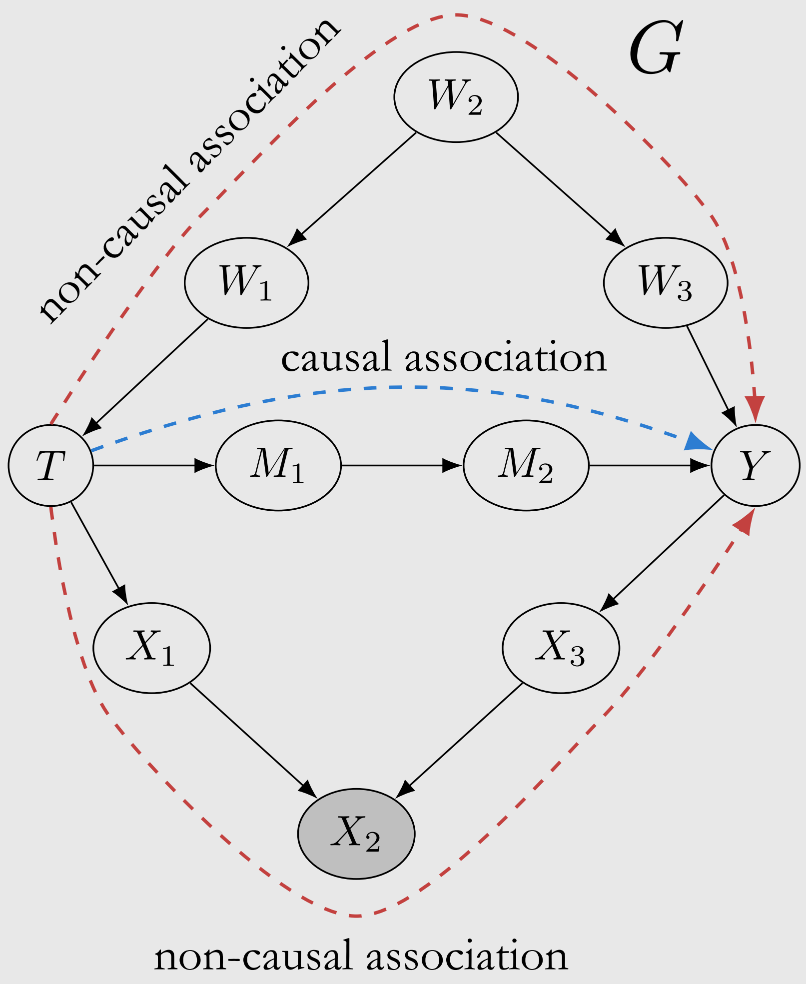
$$= \sum_w P(y \mid t, w)\, P(w \mid do(t))$$

$$= \sum_w P(y \mid t, w)\, P(w)$$

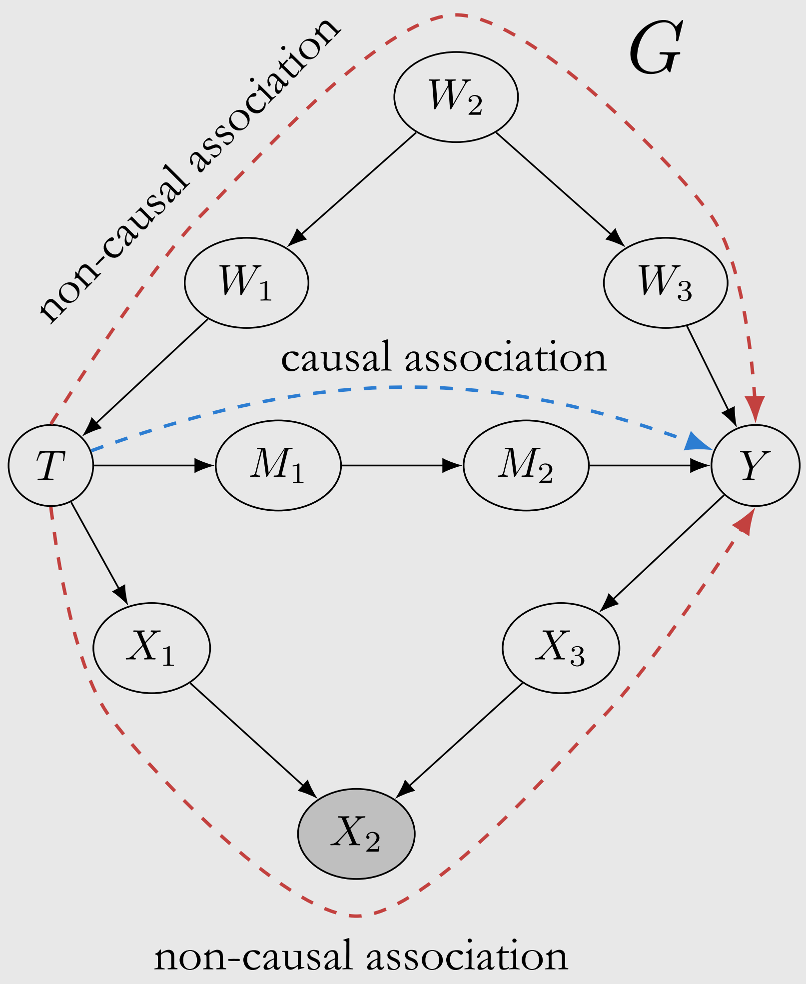
Example graph:



Backdoor criterion as d-separation

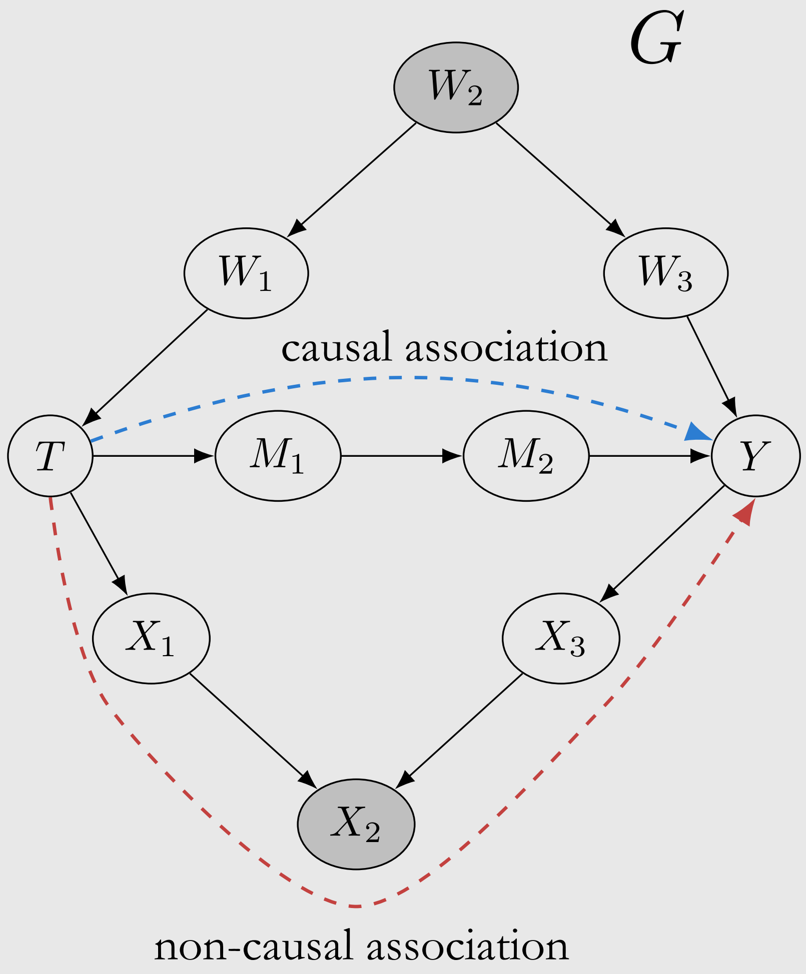


Backdoor criterion as d-separation



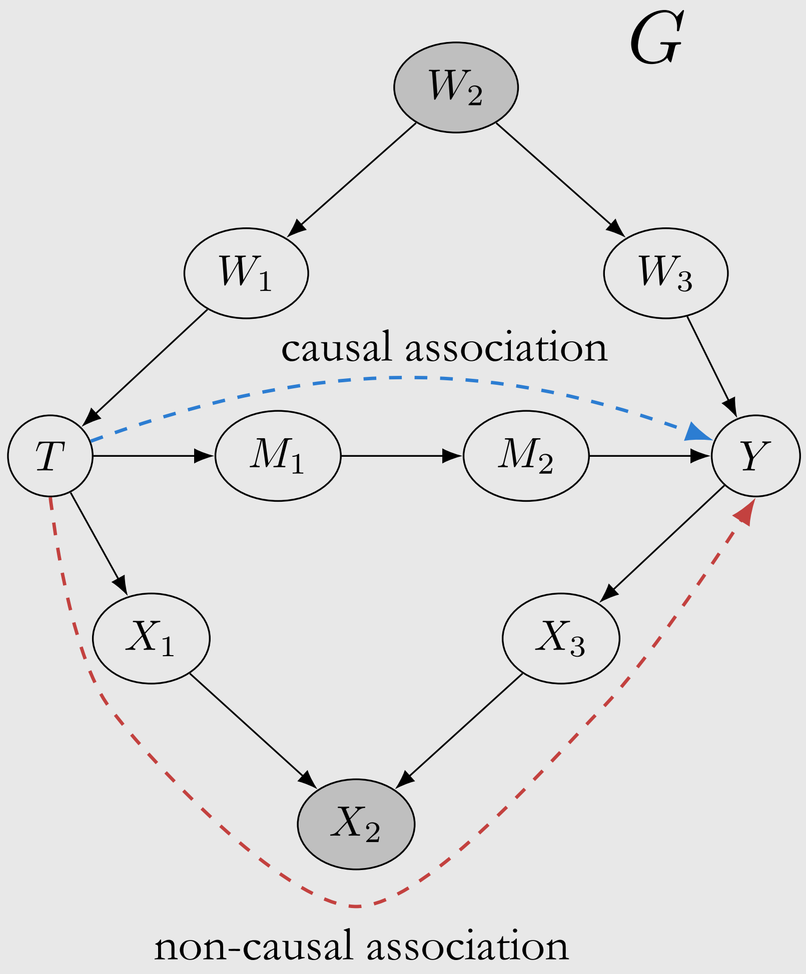
1. W blocks all backdoor paths from T to Y
2.

Backdoor criterion as d-separation



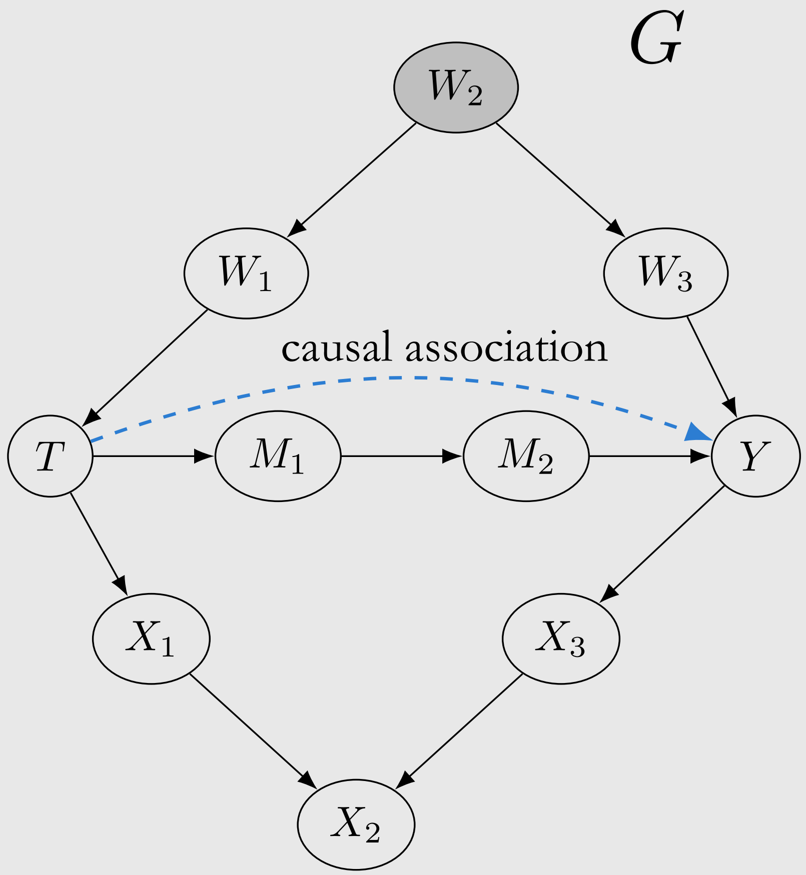
1. W blocks all backdoor paths from T to Y
2.

Backdoor criterion as d-separation



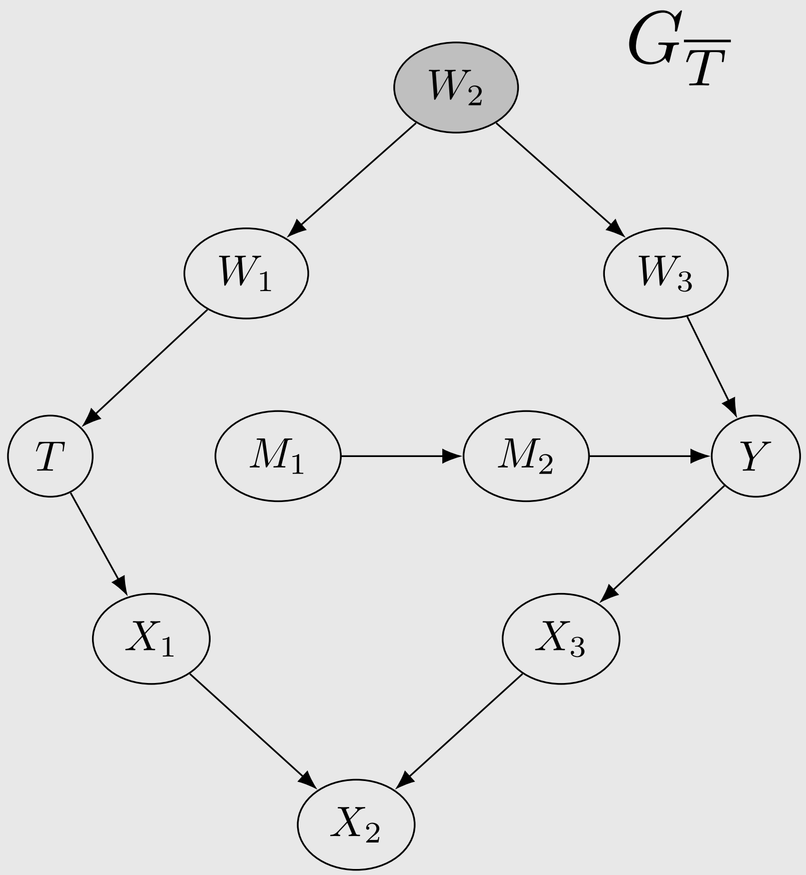
1. W blocks all backdoor paths from T to Y
2. W does not contain any descendants of T

Backdoor criterion as d-separation



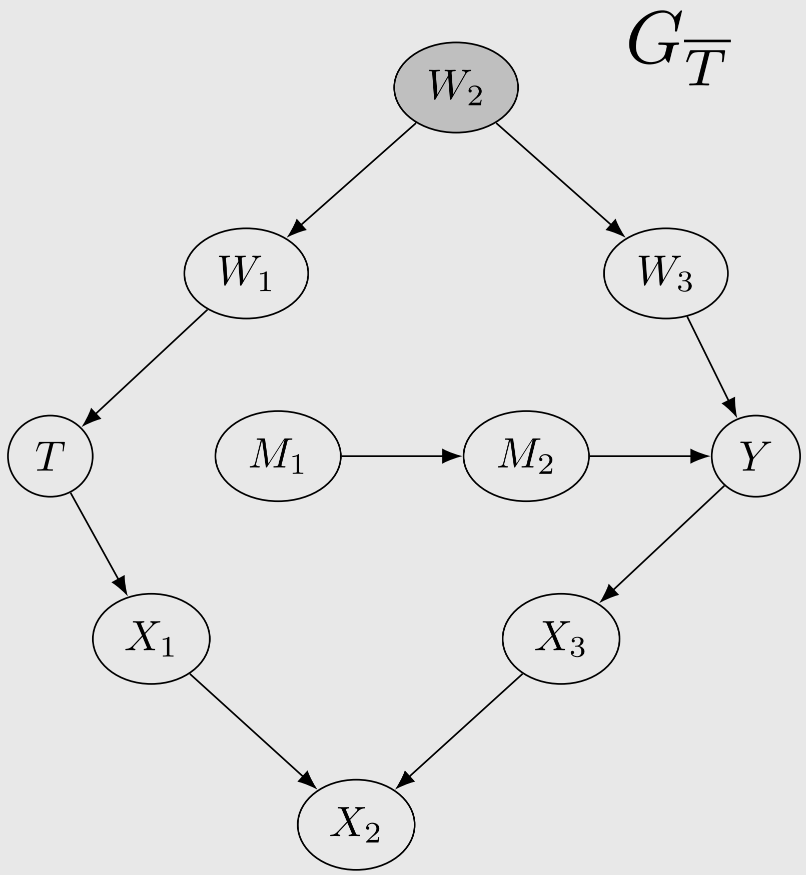
1. W blocks all backdoor paths from T to Y
2. W does not contain any descendants of T

Backdoor criterion as d-separation



1. W blocks all backdoor paths from T to Y
2. W does not contain any descendants of T

Backdoor criterion as d-separation



1. W blocks all backdoor paths from T to Y
2. W does not contain any descendants of T

$$Y \perp\!\!\!\perp_{G_{\overline{T}}} T \mid W$$

Question:
How does this backdoor adjustment relate to the adjustment formula we saw in the potential outcomes lecture?

Backdoor adjustment:

$$P(y \mid do(t)) = \sum_w P(y \mid t, w)\, P(w)$$

Adjustment formula from before:

$$\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_W \left[\mathbb{E}[Y \mid T=1, W] - \mathbb{E}[Y \mid T=0, W]\right]$$

Question:
How does this backdoor adjustment relate to the adjustment formula we saw in the potential outcomes lecture?

Section 4.4.1 of the ICI book

Backdoor adjustment:

$$P(y \mid do(t)) = \sum_w P(y \mid t, w)\, P(w)$$

Adjustment formula from before:

$$\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_W \left[\mathbb{E}[Y \mid T = 1, W] - \mathbb{E}[Y \mid T = 0, W]\right]$$

The *do*-operator

Main assumption: modularity

Backdoor adjustment

Structural causal models

A complete example with estimation

Structural equations

Structural equations

The equals sign does not convey any causal information.

Structural equations

The equals sign does not convey any causal information.

$$B = A \quad \text{means the same thing as} \quad A = B$$

Structural equations

The equals sign does not convey any causal information.

$$B = A \quad \text{means the same thing as} \quad A = B$$

Structural equation for A as a cause of B:

$$B := f(A)$$

Structural equations

The equals sign does not convey any causal information.

$$B = A \quad \text{means the same thing as} \quad A = B$$

Structural equation for A as a cause of B:

$$B := f(A)$$

$$B := f(A, U)$$

Structural equations

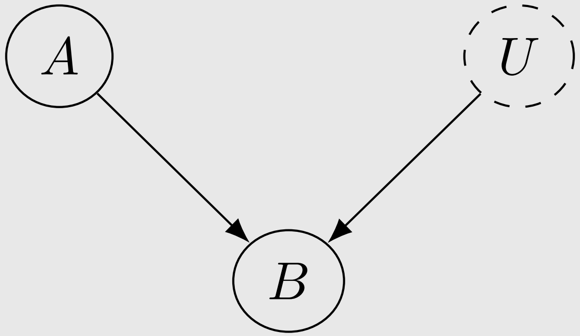
The equals sign does not convey any causal information.

$$B = A \quad \text{means the same thing as} \quad A = B$$

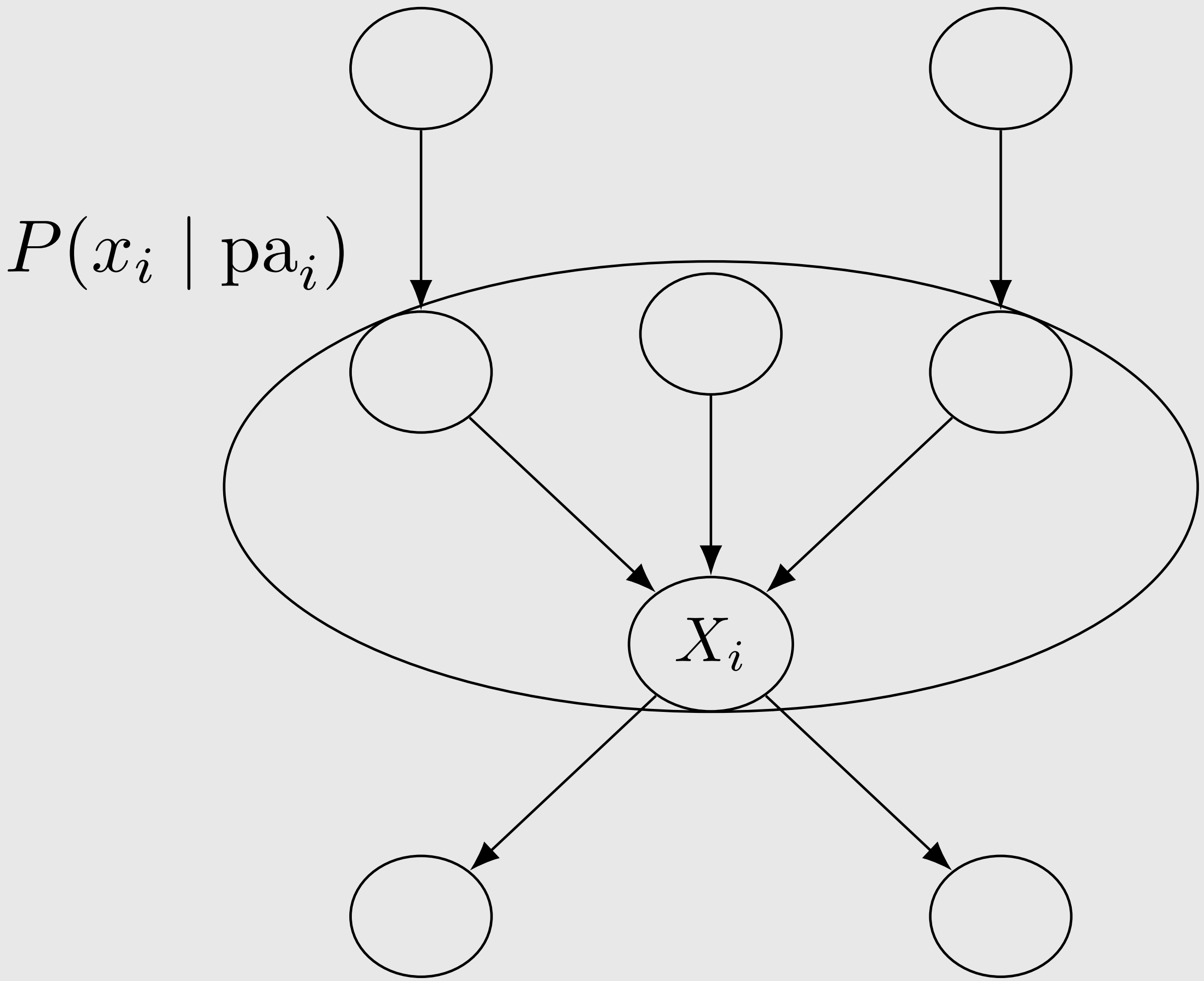
Structural equation for A as a cause of B:

$$B := f(A)$$

$$B := f(A, U)$$



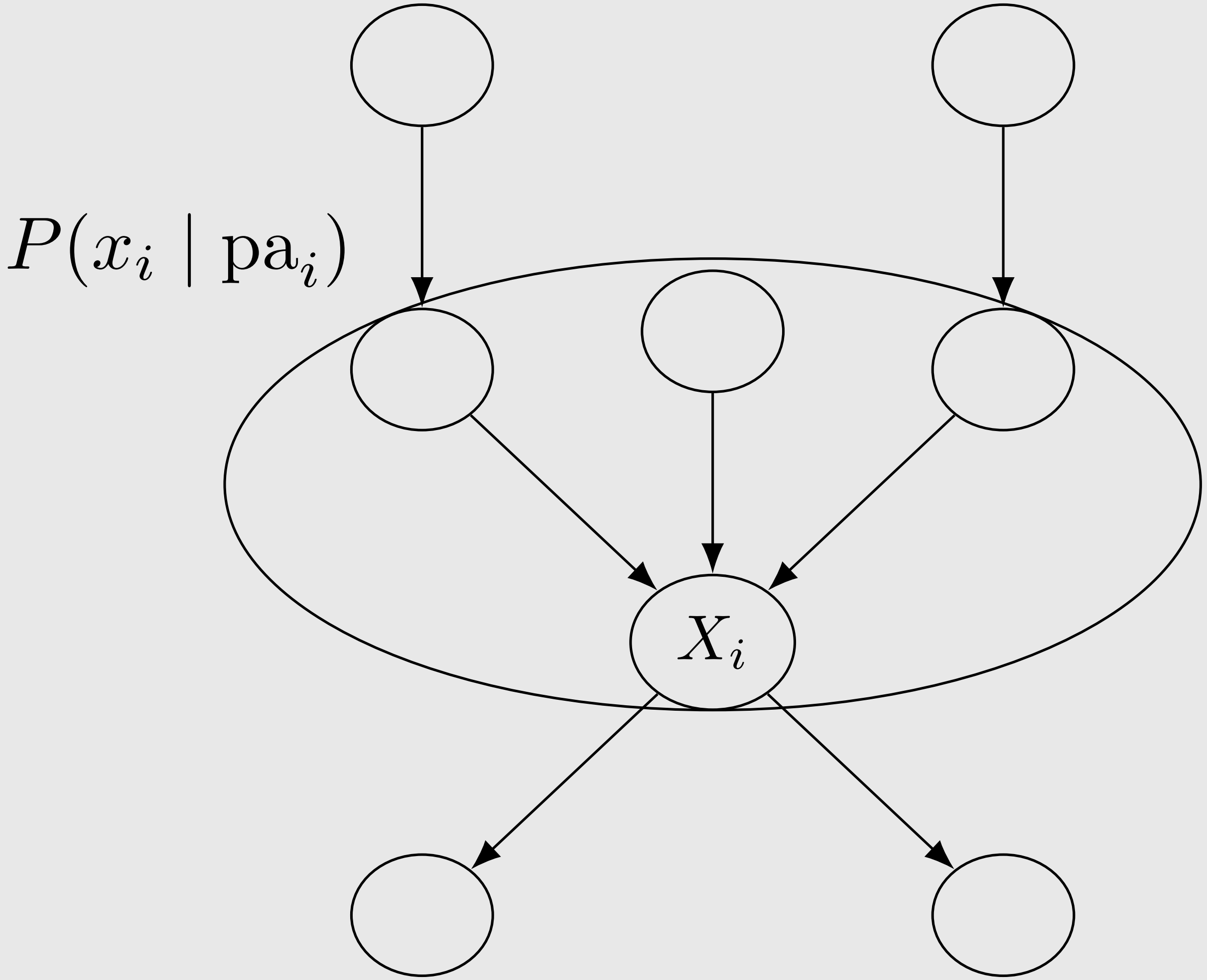
Causal mechanisms and direct causes revisited



Causal mechanisms and direct causes revisited

Causal mechanism for X_i

$$X_i := f(A, B, \ldots)$$

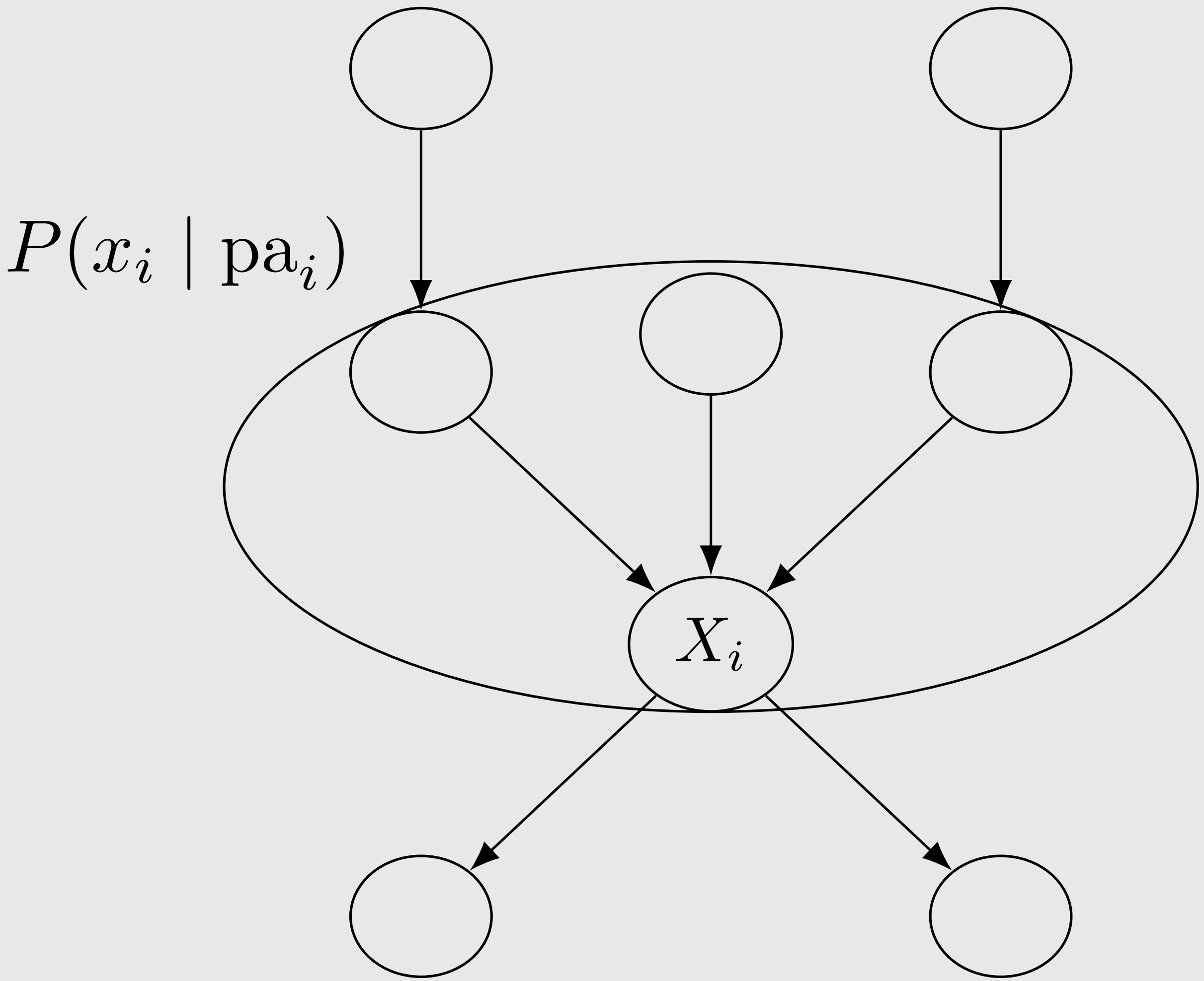


Causal mechanisms and direct causes revisited

Causal mechanism for X_i

$$X_i := f(A, B, \ldots)$$

Direct causes of X_i

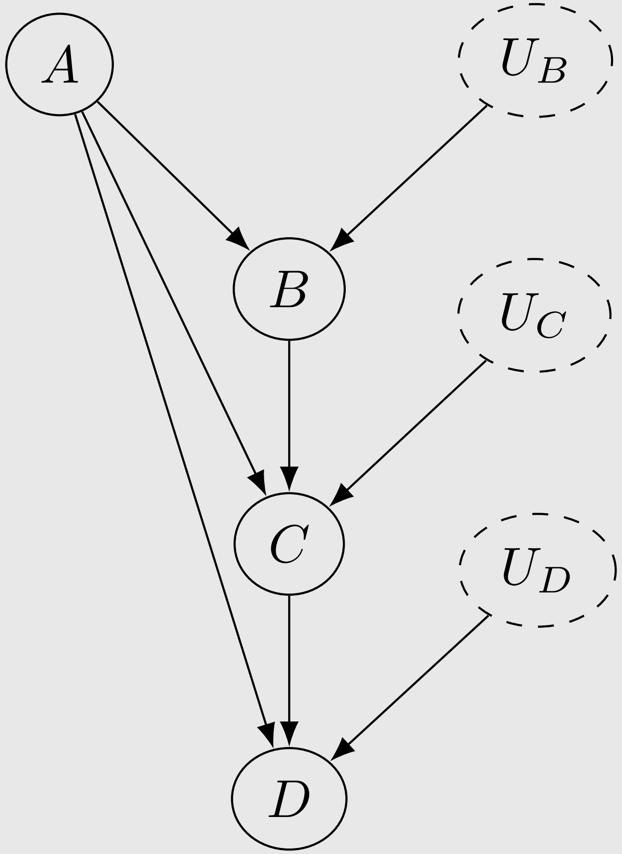


Structural causal models (SCMs)

$$M: \quad \begin{aligned} B &:= f_B(A, U_B) \\ C &:= f_C(A, B, U_C) \\ D &:= f_D(A, C, U_D) \end{aligned}$$

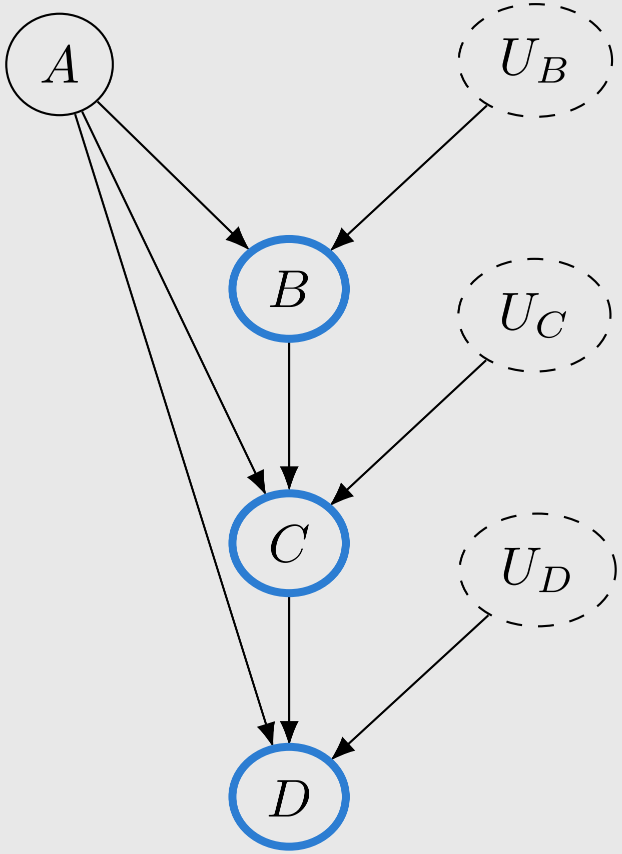
Structural causal models (SCMs)

$$M: \quad \begin{aligned} B &:= f_B(A, U_B) \\ C &:= f_C(A, B, U_C) \\ D &:= f_D(A, C, U_D) \end{aligned}$$



Structural causal models (SCMs)

$$M: \quad \begin{aligned} B &:= f_B(A, U_B) \\ C &:= f_C(A, B, U_C) \\ D &:= f_D(A, C, U_D) \end{aligned}$$

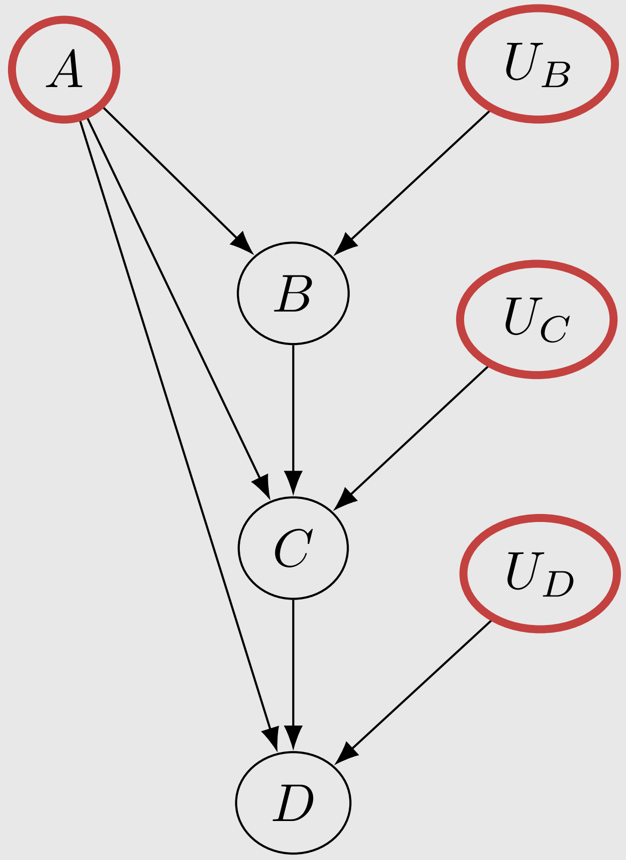


Endogenous variables

Structural causal models (SCMs)

$$M: \quad \begin{aligned} B &:= f_B(A, U_B) \\ C &:= f_C(A, B, U_C) \\ D &:= f_D(A, C, U_D) \end{aligned}$$

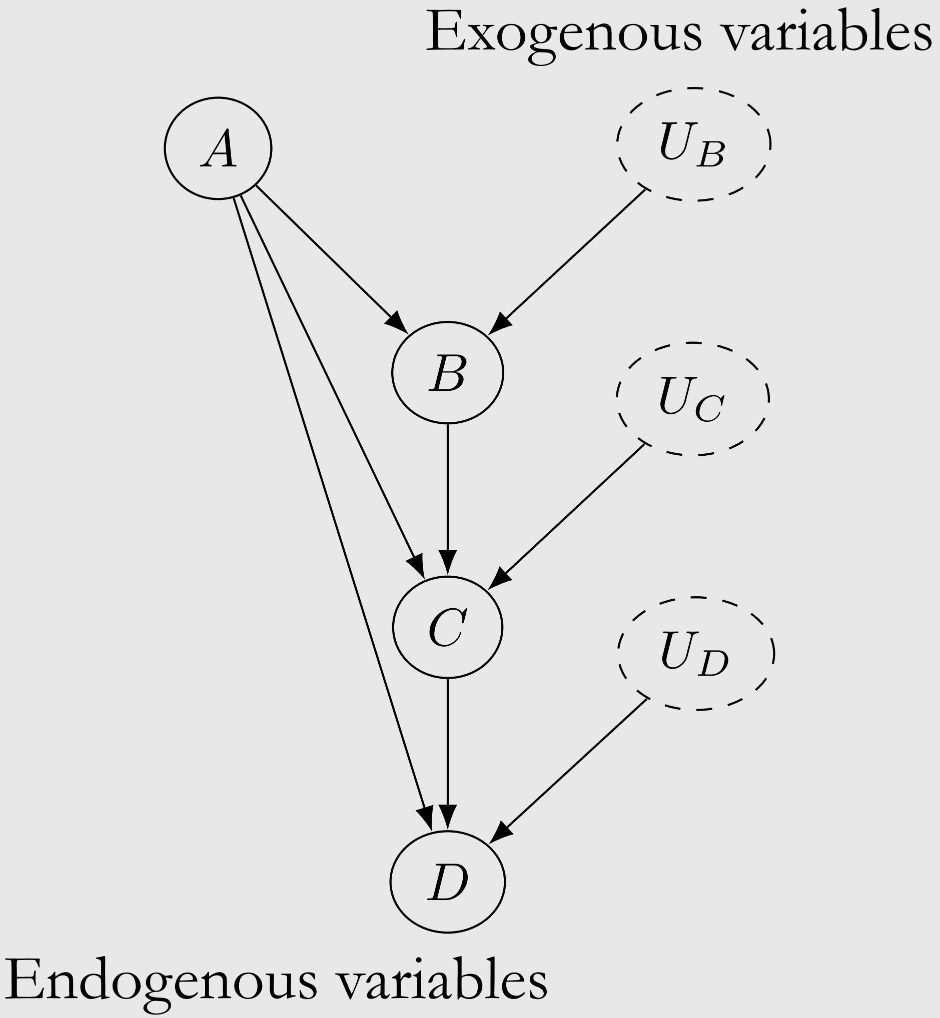
Exogenous variables



Endogenous variables

Structural causal models (SCMs)

$$M : \quad \begin{aligned} B &:= f_B(A, U_B) \\ C &:= f_C(A, B, U_C) \\ D &:= f_D(A, C, U_D) \end{aligned}$$



SCM Definition

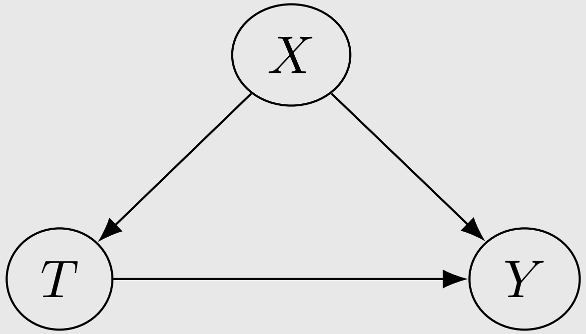
A tuple of the following sets:

1. A set of endogenous variables
2. A set of exogenous variables
3. A set of functions, one to generate each endogenous variable as a function of the other variables

Interventions

SCM
(model)

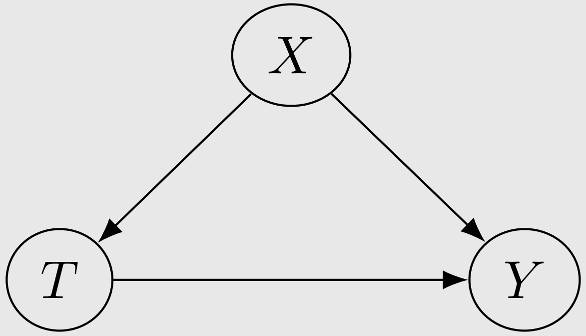
$$M: \quad \begin{aligned} T &:= f_T(X, U_T) \\ Y &:= f_Y(X, T, U_Y) \end{aligned}$$



Interventions

SCM
(model)

$$M: \quad \begin{aligned} T &:= f_T(X, U_T) \\ Y &:= f_Y(X, T, U_Y) \end{aligned}$$



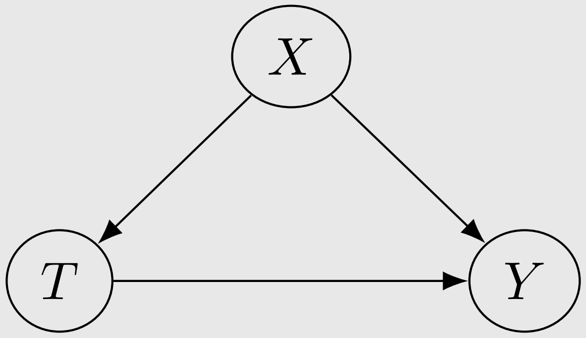
Interventional SCM
(submodel)

$$M_t: \quad \begin{aligned} T &:= t \\ Y &:= f_Y(X, T, U_Y) \end{aligned}$$

Interventions

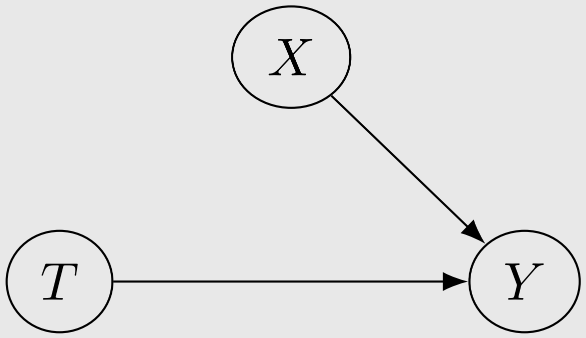
SCM
(model)

$$M: \quad \begin{aligned} T &:= f_T(X, U_T) \\ Y &:= f_Y(X, T, U_Y) \end{aligned}$$



Interventional SCM
(submodel)

$$M_t: \quad \begin{aligned} T &:= t \\ Y &:= f_Y(X, T, U_Y) \end{aligned}$$



Modularity assumption for SCMs

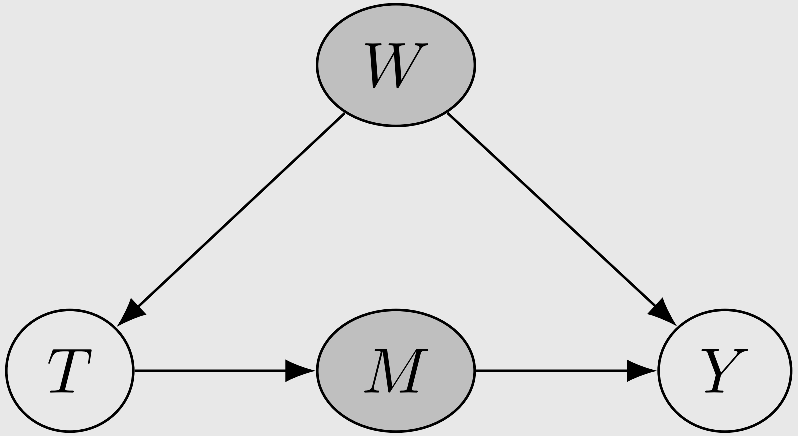
Consider an SCM M and an interventional SCM M_t that we get by performing the intervention $do(T = t)$. The modularity assumption states that M and M_t share all of their structural equations except the structural equation for T, which is $T := t$ in M_t.

Modularity assumption for SCMs

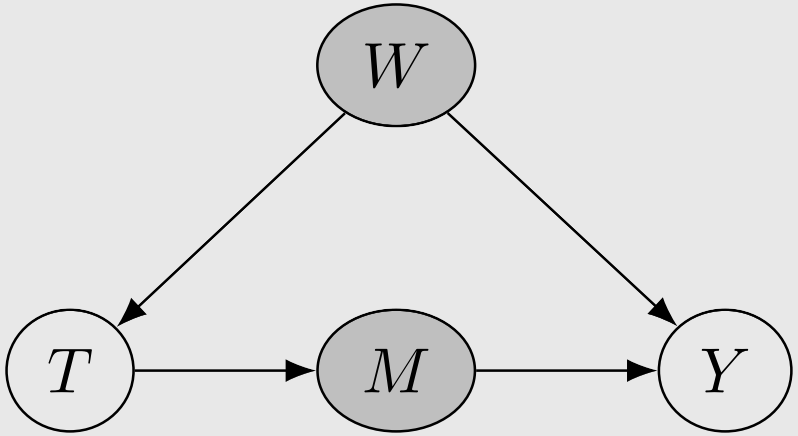
Consider an SCM M and an interventional SCM M_t that we get by performing the intervention *do*(T = t). The modularity assumption states that M and M_t share all of their structural equations except the structural equation for T, which is T := t in M_t.

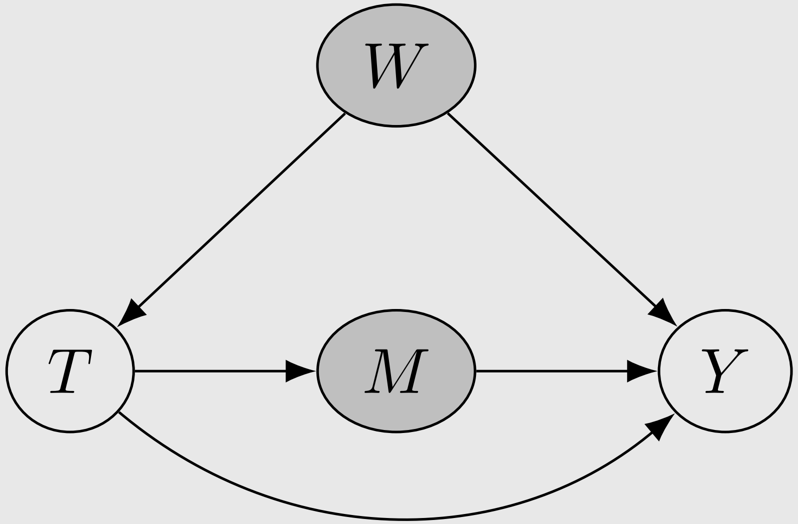
$$M : \quad \begin{aligned} T &:= f_T(X, U_T) \\ Y &:= f_Y(X, T, U_Y) \end{aligned} \qquad\qquad M_t : \quad \begin{aligned} T &:= t \\ Y &:= f_Y(X, T, U_Y) \end{aligned}$$

Why not condition on descendants of treatment: blocking causal association



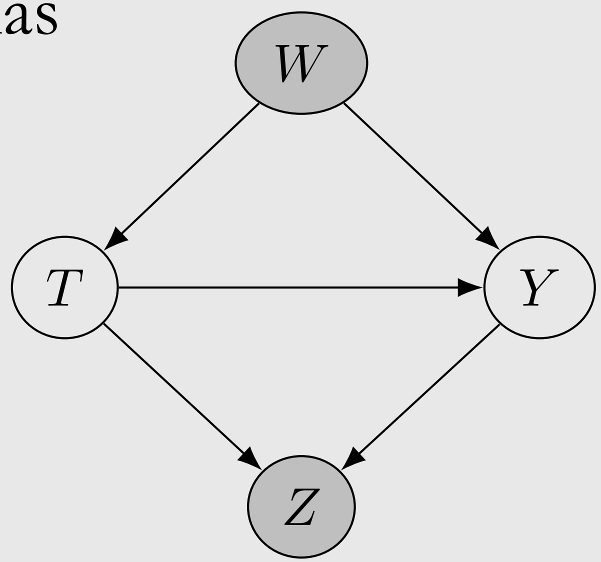
Why not condition on descendants of treatment: blocking causal association





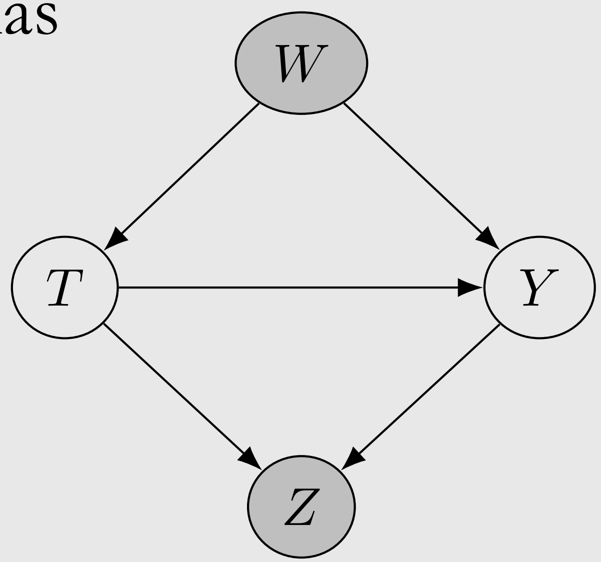
Why not condition on descendants of treatment: inducing new post-treatment association

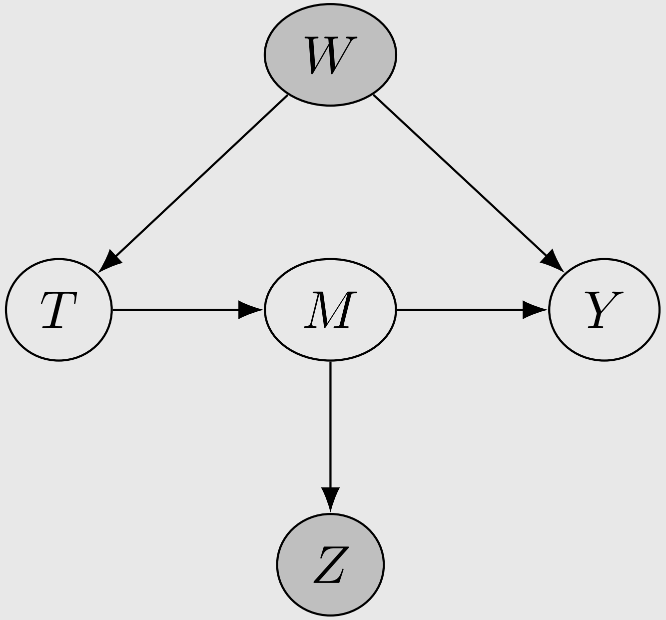
Collider bias



Why not condition on descendants of treatment: inducing new post-treatment association

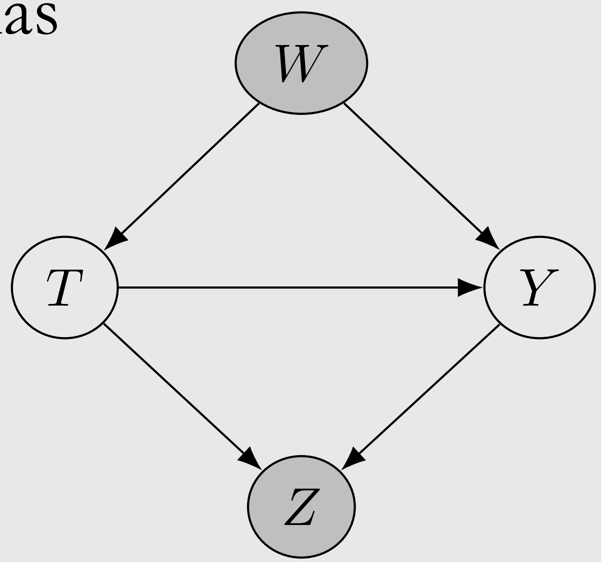
Collider bias

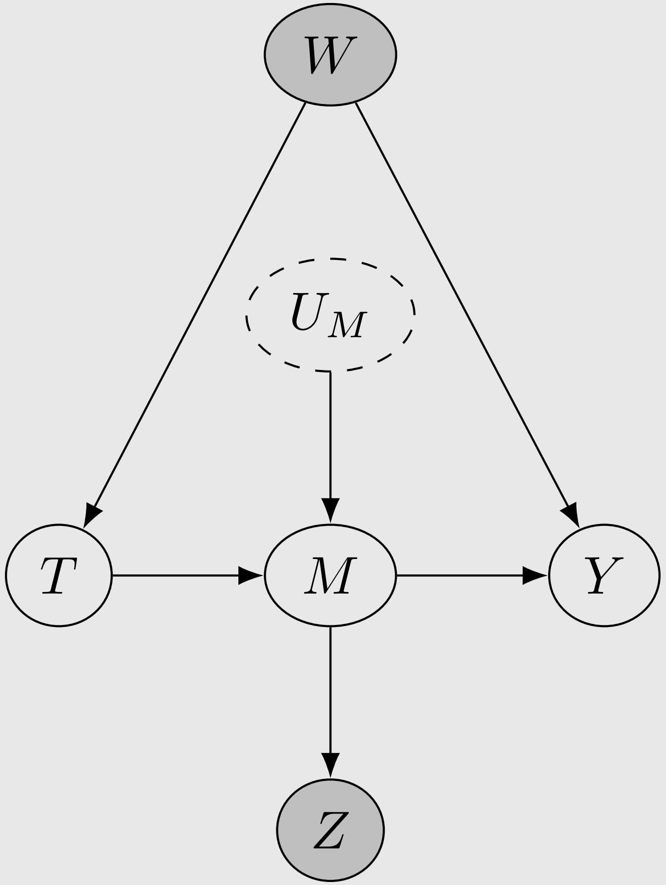




Why not condition on descendants of treatment: inducing new post-treatment association

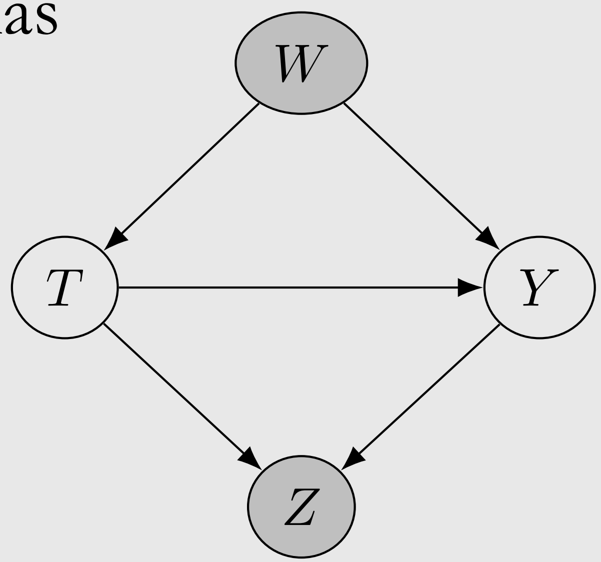
Collider bias

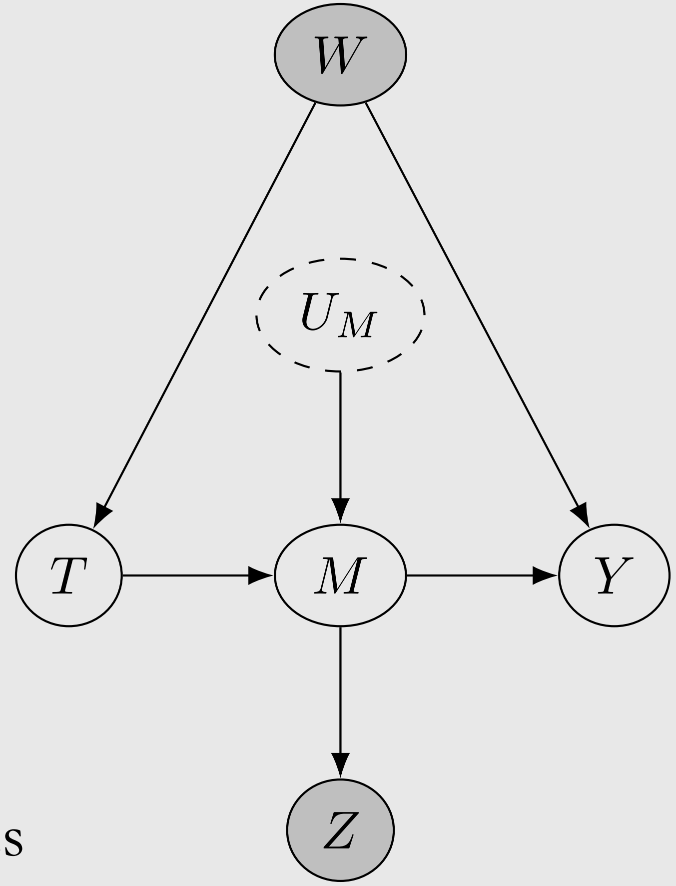




Why not condition on descendants of treatment: inducing new post-treatment association

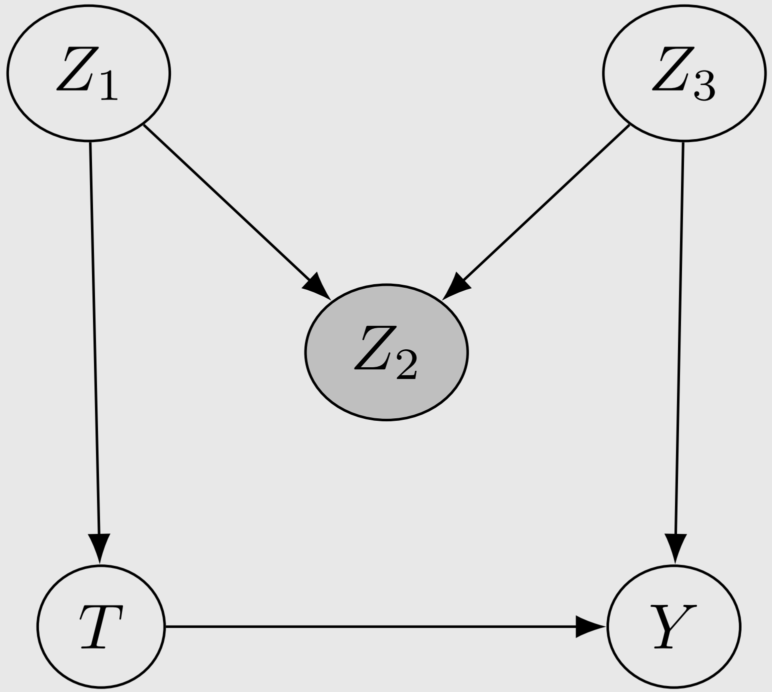
Collider bias



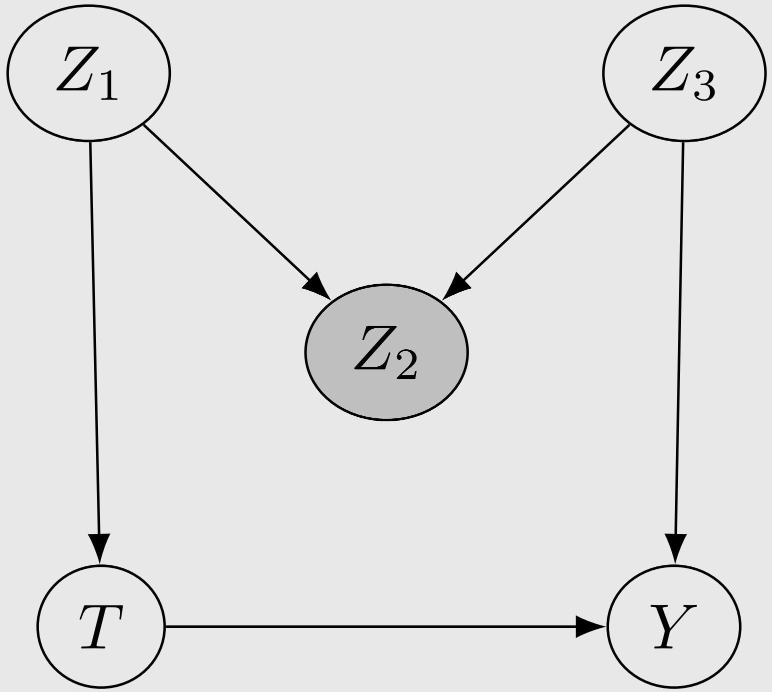


Rule: don't condition on post-treatment covariates

Inducing new **pre**treatment association (M-bias)



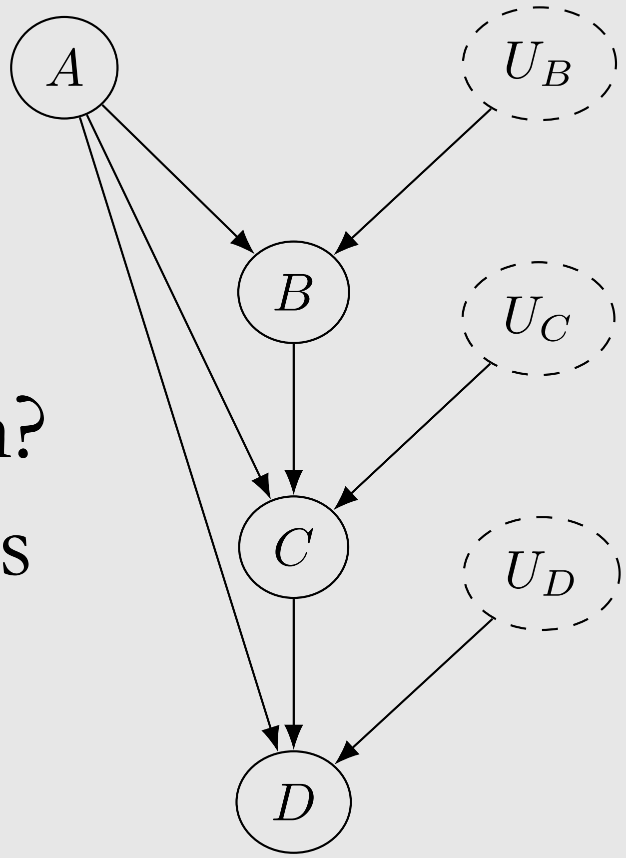
Inducing new **pre**treatment association (M-bias)



See Elwert & Winship (2014) for many real examples of collider bias

Questions:

1. What are are the nonparametric structural equations for this causal graph?
2. What are the endogenous and exogenous variables in this causal graph?
3. What is collider bias?



The *do*-operator

Main assumption: modularity

Backdoor adjustment

Structural causal models

A complete example with estimation

Problem: effect of sodium intake on blood pressure

Problem: effect of sodium intake on blood pressure

Motivation: 46% of Americans have high blood pressure and high blood pressure is associated with increased risk of mortality

Problem: effect of sodium intake on blood pressure

Motivation: 46% of Americans have high blood pressure and high blood pressure is associated with increased risk of mortality

Data:

- Epidemiological example taken from Luque-Fernandez et al. (2018)

Problem: effect of sodium intake on blood pressure

Motivation: 46% of Americans have high blood pressure and high blood pressure is associated with increased risk of mortality

Data:

- Epidemiological example taken from Luque-Fernandez et al. (2018)
- Outcome Y: (systolic) blood pressure (continuous)

Problem: effect of sodium intake on blood pressure

Motivation: 46% of Americans have high blood pressure and high blood pressure is associated with increased risk of mortality

Data:

- Epidemiological example taken from Luque-Fernandez et al. (2018)
- Outcome Y: (systolic) blood pressure (continuous)
- Treatment T: sodium intake (1 if above 3.5 mg and 0 if below)

Problem: effect of sodium intake on blood pressure

Motivation: 46% of Americans have high blood pressure and high blood pressure is associated with increased risk of mortality

Data:

- Epidemiological example taken from Luque-Fernandez et al. (2018)
- Outcome Y: (systolic) blood pressure (continuous)
- Treatment T: sodium intake (1 if above 3.5 mg and 0 if below)
- Covariates
  - W age
  - Z amount of protein excreted in urine

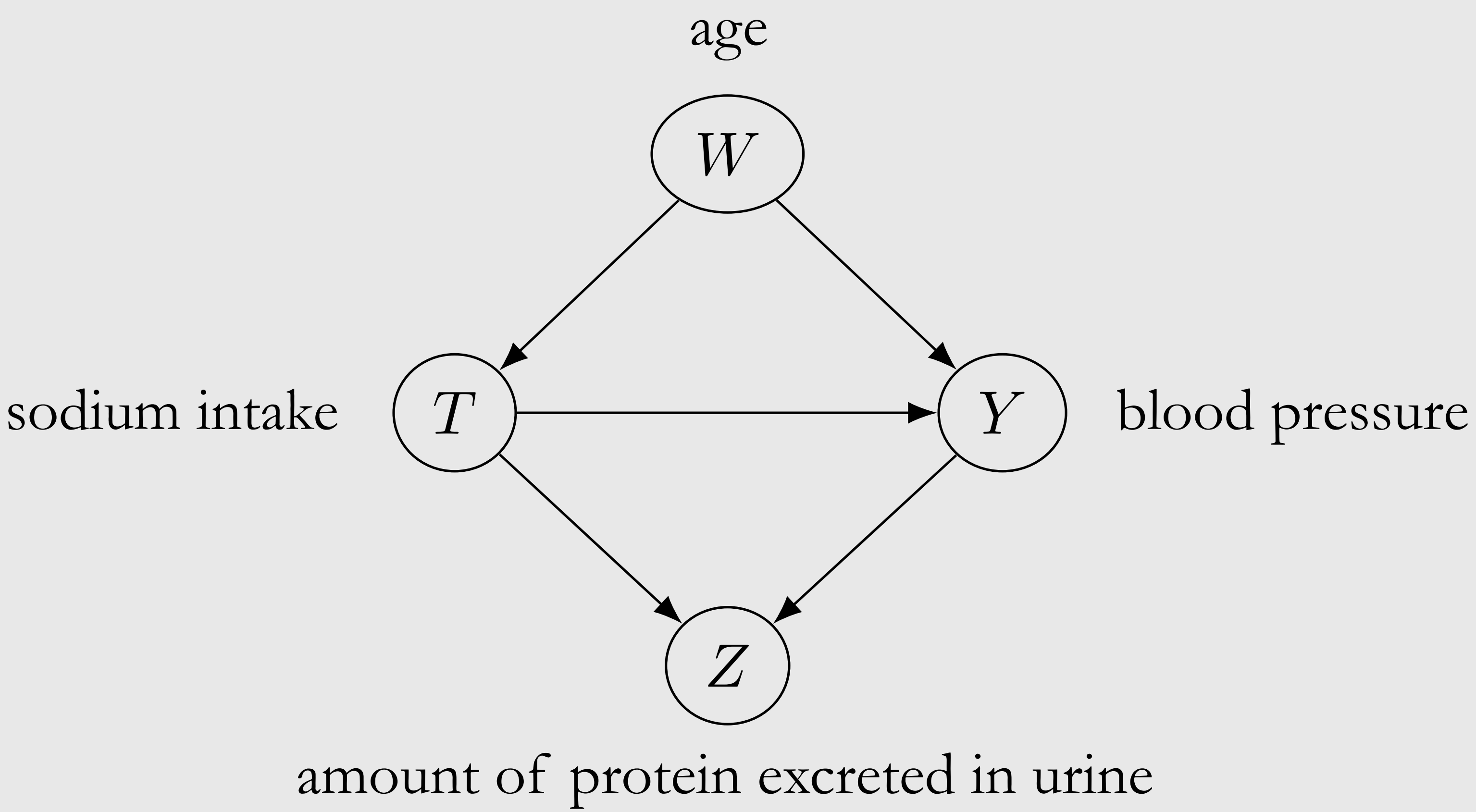
Problem: effect of sodium intake on blood pressure

Motivation: 46% of Americans have high blood pressure and high blood pressure is associated with increased risk of mortality

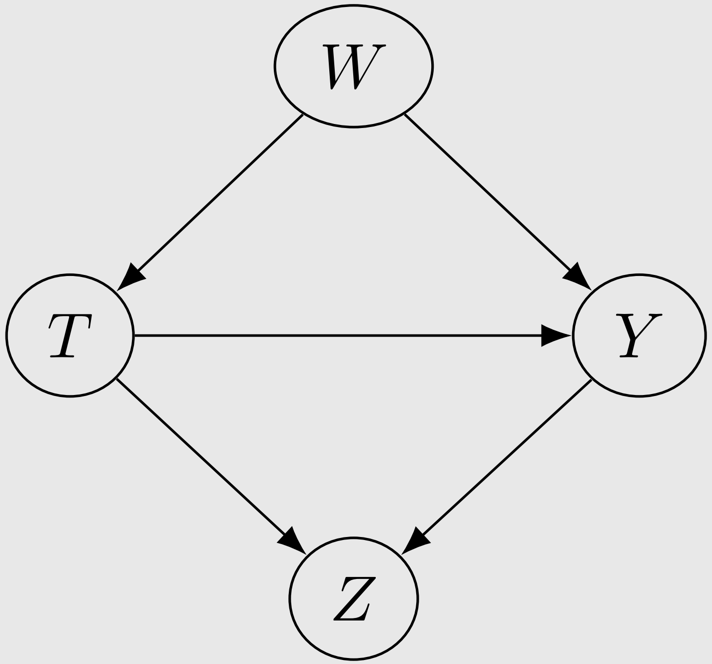
Data:

- Epidemiological example taken from Luque-Fernandez et al. (2018)
- Outcome Y: (systolic) blood pressure (continuous)
- Treatment T: sodium intake (1 if above 3.5 mg and 0 if below)
- Covariates
  - W age
  - Z amount of protein excreted in urine
- Simulation: so we know the "true" ATE is 1.05

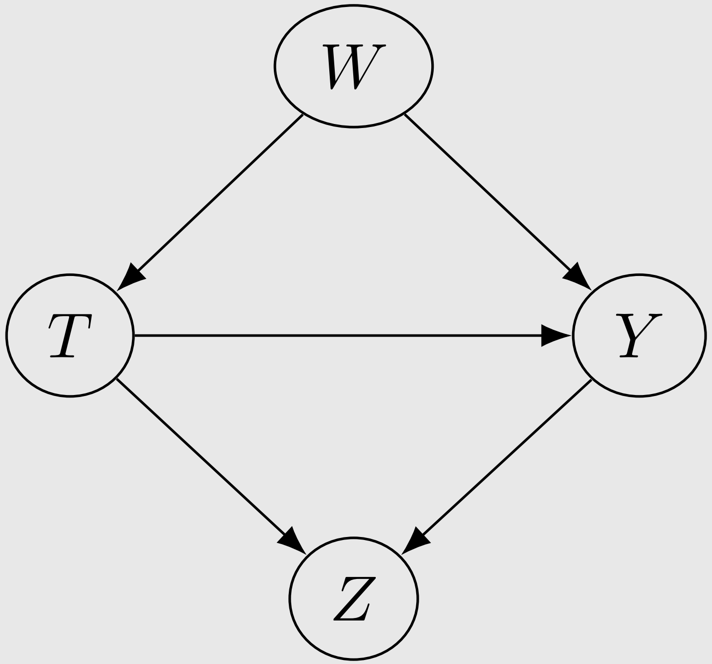
The causal graph



Identification

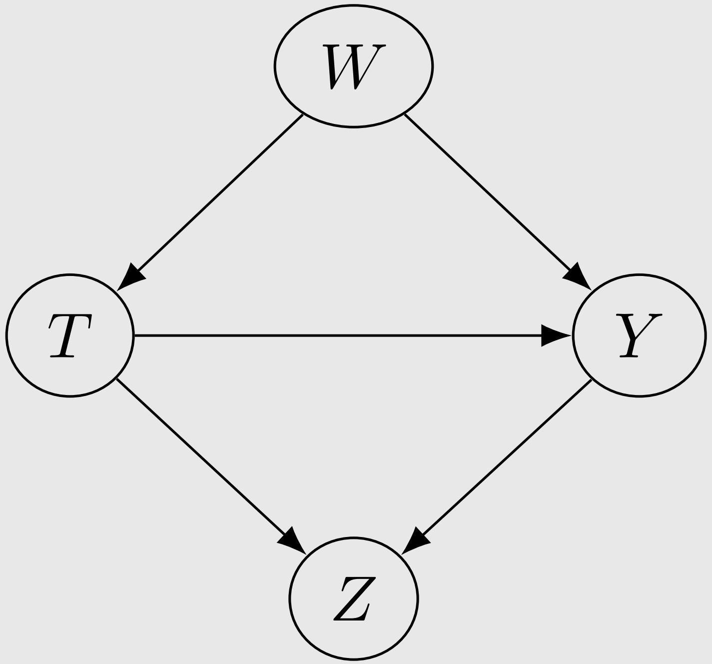


Identification



Causal estimand: $\mathbb{E}[Y \mid do(t)]$

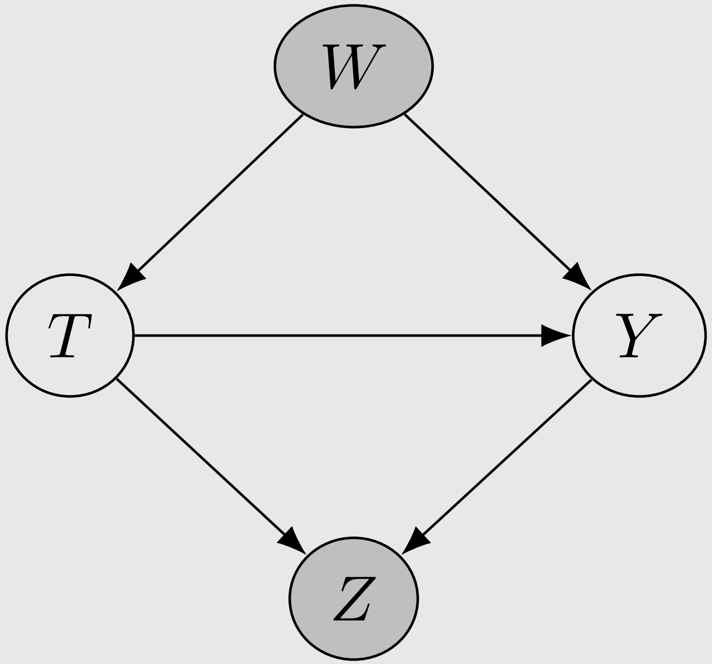
Identification



Causal estimand: $\mathbb{E}[Y \mid do(t)]$

Statistical estimand from last week: $\mathbb{E}_{W,Z}\mathbb{E}[Y \mid t, W, Z]$

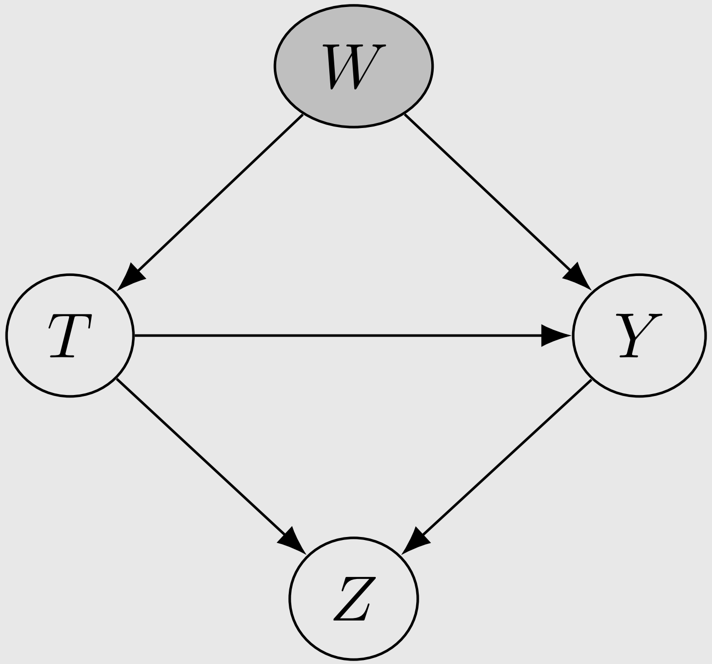
Identification



Causal estimand: $\mathbb{E}[Y \mid do(t)]$

Statistical estimand from last week: $\mathbb{E}_{W,Z}\mathbb{E}[Y \mid t, W, Z]$

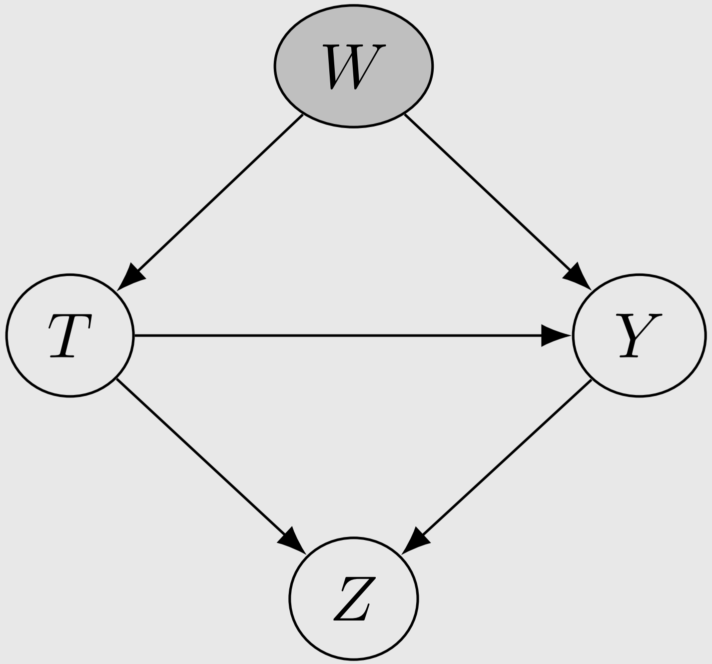
Identification



Causal estimand: $\mathbb{E}[Y \mid do(t)]$

Statistical estimand from last week: $\mathbb{E}_{W,Z}\mathbb{E}[Y \mid t, W, Z]$

Identification



Causal estimand: $\mathbb{E}[Y \mid do(t)]$

Statistical estimand from last week: $\mathbb{E}_{W,Z}\mathbb{E}[Y \mid t, W, Z]$

Statistical estimand from causal graph: $\mathbb{E}_W\mathbb{E}[Y \mid t, W]$

Estimation of ATE

True ATE: $\mathbb{E}[Y(1) - Y(0)] = 1.05$

Estimation of ATE

True ATE: $\mathbb{E}[Y(1) - Y(0)] = 1.05$

Identification: $\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_X \left[\mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X]\right]$

Estimation of ATE

True ATE: $\mathbb{E}[Y(1) - Y(0)] = 1.05$

Identification: $\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_X\left[\mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X]\right]$

Estimation:

Estimation of ATE

True ATE: $\mathbb{E}[Y(1) - Y(0)] = 1.05$

Identification: $\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_X \left[\mathbb{E}[Y \mid T = 1,\ X] - \mathbb{E}[Y \mid T = 0,\ X]\right]$

Estimation: $\frac{1}{n} \sum_i [\mathbb{E}[Y \mid T = 1, X = x_i] - \mathbb{E}[Y \mid T = 0, X = x_i]]$

Estimation of ATE

True ATE: $\mathbb{E}[Y(1) - Y(0)] = 1.05$

Identification: $\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_X \left[\mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X]\right]$

Estimation: $\frac{1}{n} \sum_i [\underbrace{\mathbb{E}[Y \mid T = 1, X = x_i]} - \underbrace{\mathbb{E}[Y \mid T = 0, X = x_i]}]$

Model (linear regression)

Estimation of ATE

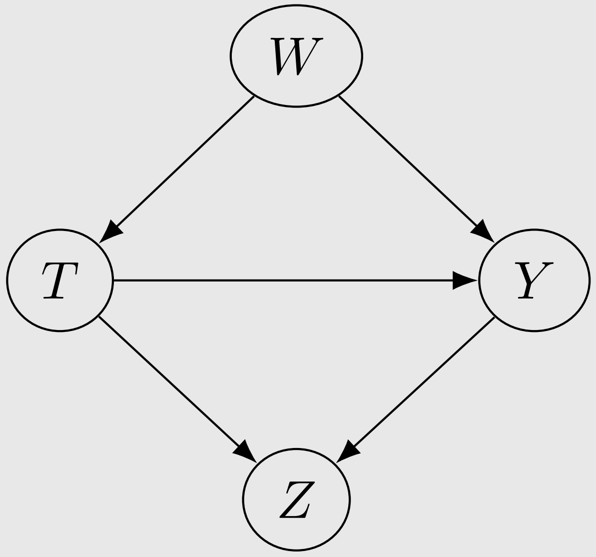
True ATE: $\mathbb{E}[Y(1) - Y(0)] = 1.05$

Identification: $\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_X\left[\mathbb{E}[Y \mid T = 1,\ X] - \mathbb{E}[Y \mid T = 0,\ X]\right]$

Estimation: $\frac{1}{n}\sum_i [\underbrace{\mathbb{E}[Y \mid T = 1, X = x_i]}_{} - \underbrace{\mathbb{E}[Y \mid T = 0, X = x_i]}_{}]$

Model (linear regression)

Estimates:



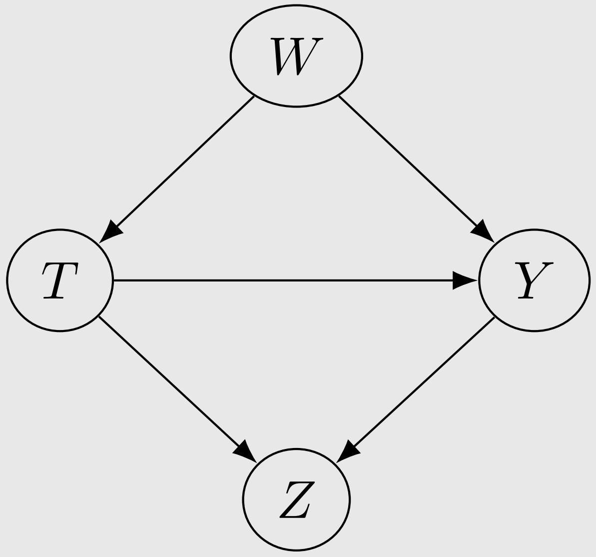
Estimation of ATE

True ATE: $\mathbb{E}[Y(1) - Y(0)] = 1.05$

Identification: $\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_X \left[\mathbb{E}[Y \mid T = 1,\ X] - \mathbb{E}[Y \mid T = 0,\ X]\right]$

Estimation: $\frac{1}{n} \sum_i [\underbrace{\mathbb{E}[Y \mid T = 1, X = x_i]} - \underbrace{\mathbb{E}[Y \mid T = 0, X = x_i]}]$

Model (linear regression)



Estimates:

$X = \{\}$ (naive): 5.33

Estimation of ATE

True ATE: $\mathbb{E}[Y(1) - Y(0)] = 1.05$

Identification: $\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_X\left[\mathbb{E}[Y \mid T = 1,\ X] - \mathbb{E}[Y \mid T = 0,\ X]\right]$

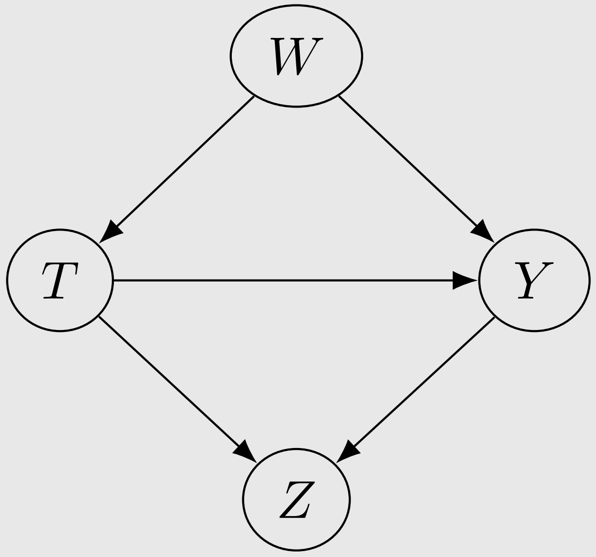
Estimation: $\frac{1}{n}\sum_i [\underbrace{\mathbb{E}[Y \mid T = 1, X = x_i]}_{} - \underbrace{\mathbb{E}[Y \mid T = 0, X = x_i]}_{}]$

Model (linear regression)

Estimates:

$X = \{\}$ (naive): 5.33

$\frac{|5.33 - 1.05|}{1.05} \times 100\% = 407\%$ error



Estimation of ATE

True ATE: $\mathbb{E}[Y(1) - Y(0)] = 1.05$

Identification: $\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_X\left[\mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X]\right]$

Estimation: $\frac{1}{n} \sum_i [\underbrace{\mathbb{E}[Y \mid T = 1, X = x_i]} - \underbrace{\mathbb{E}[Y \mid T = 0, X = x_i]}]$

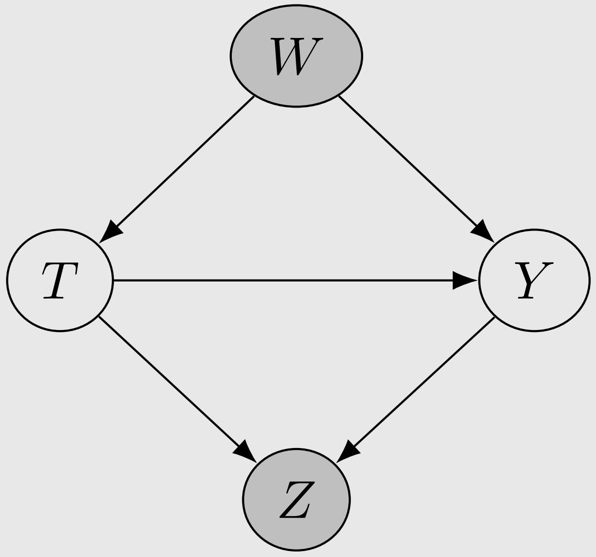
Model (linear regression)

Estimates:

$X = \{\}$ (naive): 5.33

$\frac{|5.33 - 1.05|}{1.05} \times 100\% = 407\%$ error

$X = \{W, Z\}$ (last week): 0.85



Estimation of ATE

True ATE: $\mathbb{E}[Y(1) - Y(0)] = 1.05$

Identification: $\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_X\left[\mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X]\right]$

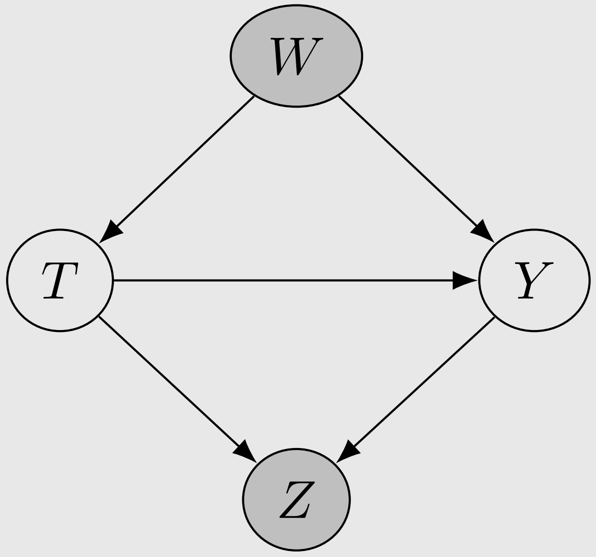
Estimation: $\frac{1}{n}\sum_i [\underbrace{\mathbb{E}[Y \mid T = 1, X = x_i]} - \underbrace{\mathbb{E}[Y \mid T = 0, X = x_i]}]$

Model (linear regression)

Estimates:

$X = \{\}$ (naive): 5.33 $\qquad \frac{|5.33 - 1.05|}{1.05} \times 100\% = 407\%$ error

$X = \{W, Z\}$ (last week): 0.85 $\qquad 19\%$ error



Estimation of ATE

True ATE: $\mathbb{E}[Y(1) - Y(0)] = 1.05$

Identification: $\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_X \left[\mathbb{E}[Y \mid T = 1,\ X] - \mathbb{E}[Y \mid T = 0,\ X]\right]$

Estimation: $\frac{1}{n} \sum_i [\underbrace{\mathbb{E}[Y \mid T = 1, X = x_i]}_{} - \underbrace{\mathbb{E}[Y \mid T = 0, X = x_i]}_{}]$

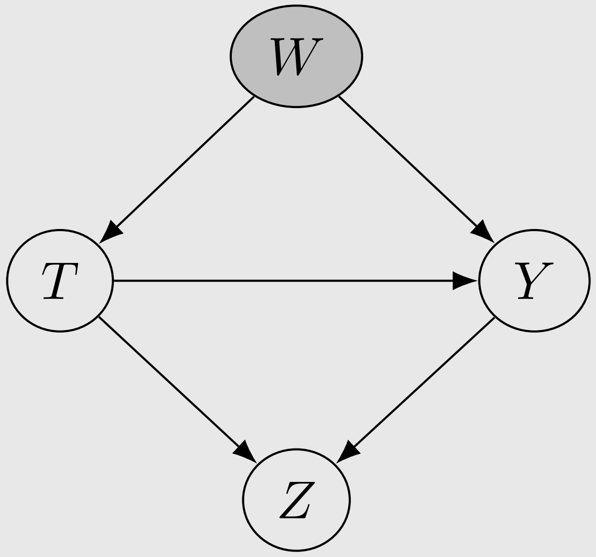
Model (linear regression)

Estimates:

$X = \{\}$ (naive): 5.33 $\qquad \frac{|5.33 - 1.05|}{1.05} \times 100\% = 407\%$ error

$X = \{W, Z\}$ (last week): 0.85 $\qquad 19\%$ error

$X = \{W\}$ (unbiased): 1.0502



Estimation of ATE

True ATE: $\mathbb{E}[Y(1) - Y(0)] = 1.05$

Identification: $\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_X\left[\mathbb{E}[Y \mid T = 1,\ X] - \mathbb{E}[Y \mid T = 0,\ X]\right]$

Estimation: $\frac{1}{n}\sum_i [\underbrace{\mathbb{E}[Y \mid T = 1, X = x_i]} - \underbrace{\mathbb{E}[Y \mid T = 0, X = x_i]}]$

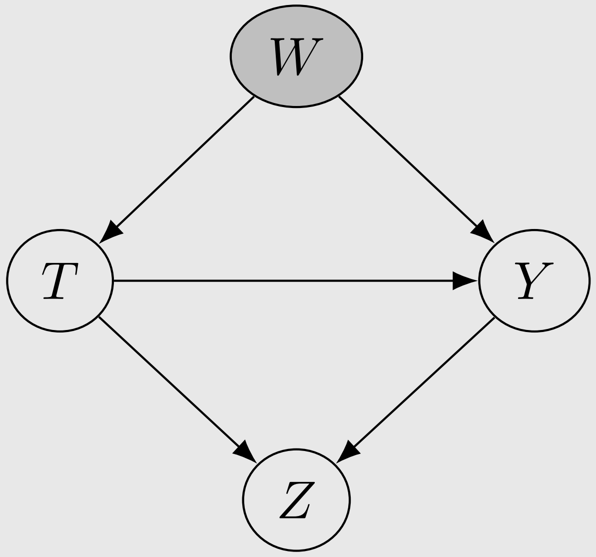
Model (linear regression)

Estimates:

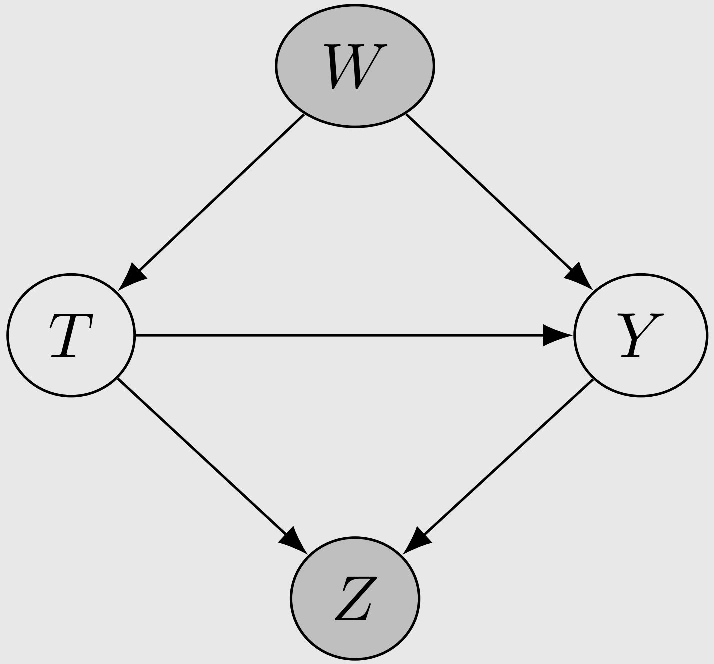
$X = \{\}$ (naive): 5.33 $\qquad \frac{|5.33 - 1.05|}{1.05} \times 100\% = 407\%$ error

$X = \{W, Z\}$ (last week): 0.85 $\qquad 19\%$ error

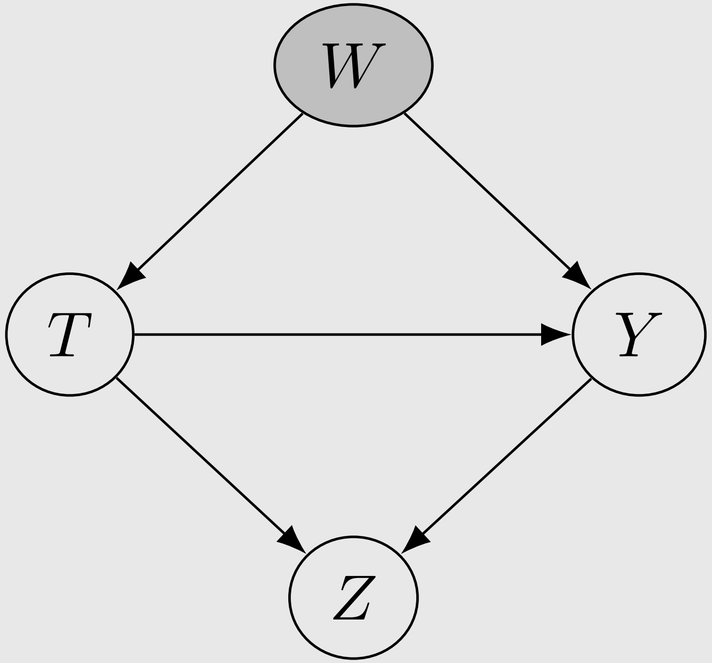
$X = \{W\}$ (unbiased): 1.0502 $\qquad 0.02\%$ error



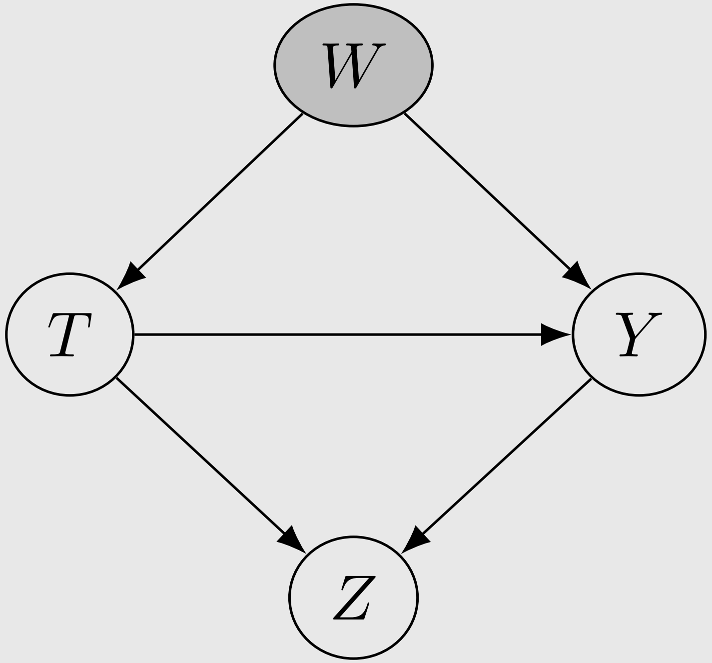
M-bias

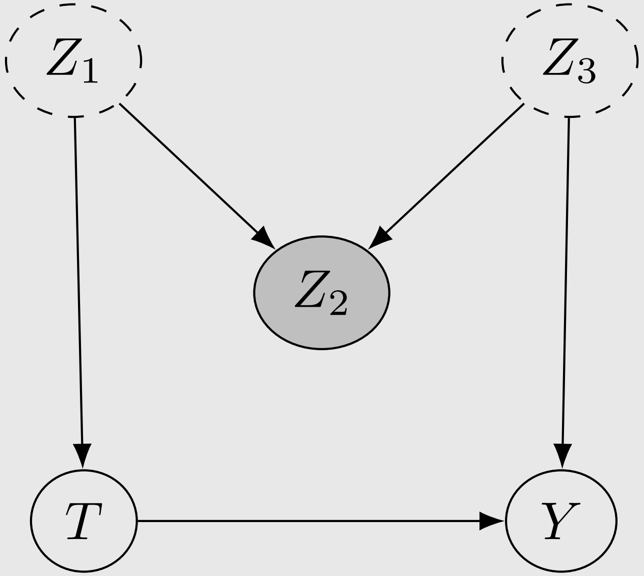


M-bias



M-bias





M-bias