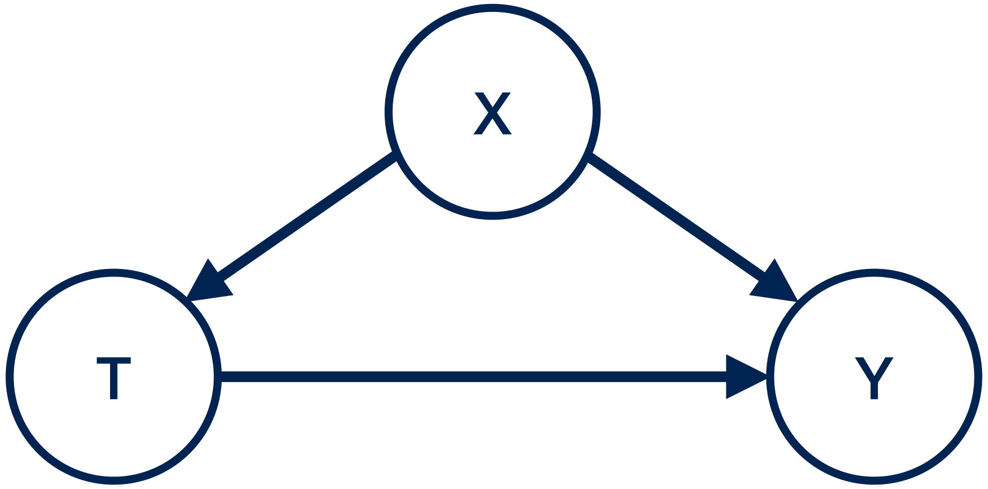
Methods for Causal Inference
Lecture 6

Ava Khamseh
School of Informatics

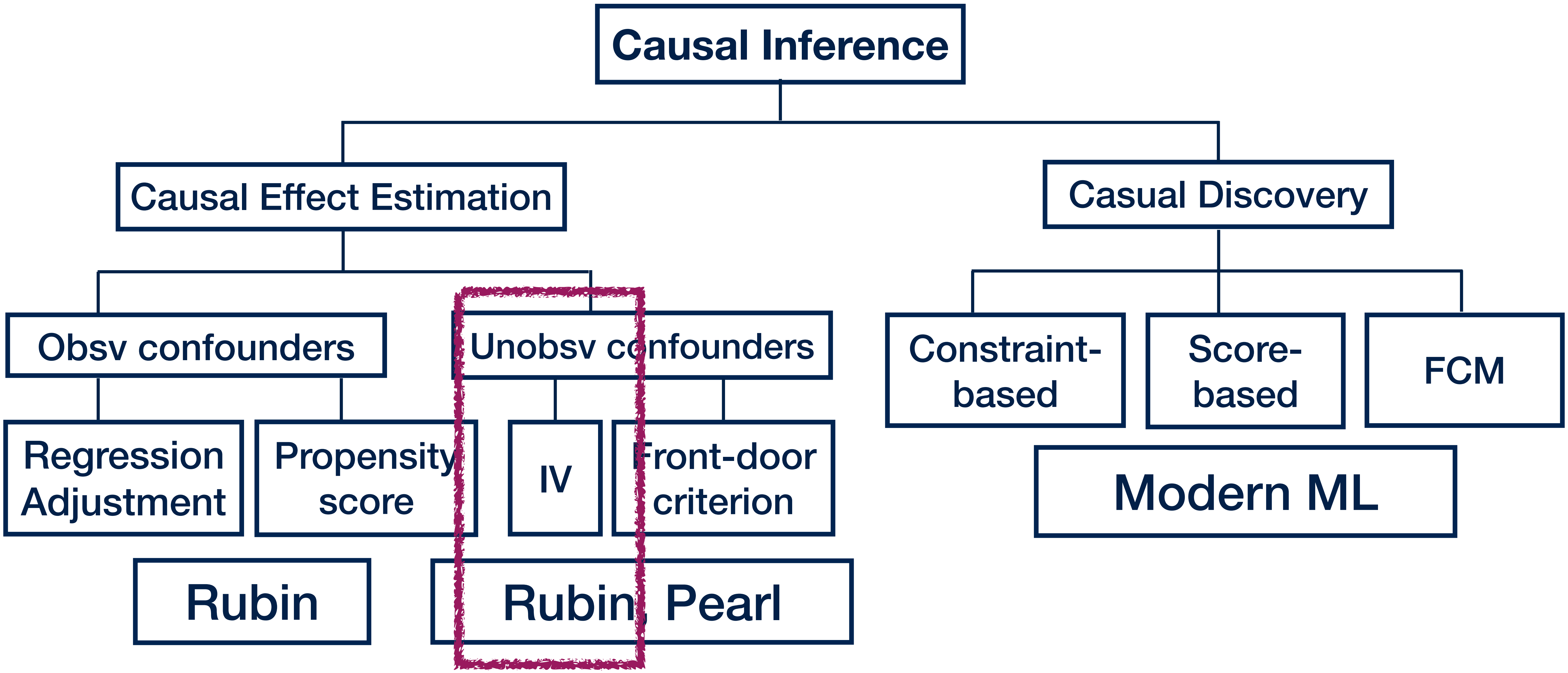
2021-2022

Causal inference with observed confounders



Overview of the course

- **Lecture 1**: Introduction & motivation, why do we care about causality?
- **Lecture 2**: Recap of probability theory, e.g., variables, events, conditional probabilities, independence, law of total probability, Bayes' rule
- **Lecture 3**: Recap of regression, multiple regression, graphs, SCM
- **Lectures 4-20**:



RCTs

Randomised Control Trials (RCT): Subjects are assigned at random to various groups (treatment or control)

RCTs are sometimes referred to ‘gold standard’ of scientific research, used in biological, medical and behavioural sciences

RCTs

Randomised Control Trials (RCT): Subjects are assigned at random to various groups (treatment or control)

RCTs are sometimes referred to ‘gold standard’ of scientific research, used in biological, medical and behavioural sciences

But RCT’s can be impossible, imperfect or unethical:

- Can be very costly and difficult to organise (demanding resources)

RCTs

Randomised Control Trials (RCT): Subjects are assigned at random to various groups (treatment or control)

RCTs are sometimes referred to ‘gold standard’ of scientific research, used in biological, medical and behavioural sciences

But RCT’s can be impossible, imperfect or unethical:

- Can be very costly and difficult to organise (demanding resources)
- Perfect control is hard to achieve (imperfect compliance): Adverse reaction to an experimental drug means dose has to be reduce no avoid harm

RCTs

Randomised Control Trials (RCT): Subjects are assigned at random to various groups (treatment or control)

RCTs are sometimes referred to ‘gold standard’ of scientific research, used in biological, medical and behavioural sciences

But RCT’s can be impossible, imperfect or unethical:

- Can be very costly and difficult to organise (demanding resources)
- Perfect control is hard to achieve (imperfect compliance): Adverse reaction to an experimental drug means dose has to be reduce no avoid harm
- Unethical: Asking pregnant women to smoke to observe child birth weight Denying the control subjects a drug, e.g. treatment could have been potentially life saving for cancer patients

RCTs

Randomised Control Trials (RCT): Subjects are assigned at random to various groups (treatment or control)

RCTs are sometimes referred to 'gold standard' of scientific research, used in biological, medical and behavioural sciences

But RCT's can be impossible, imperfect or unethical:

- Can be very costly and difficult to organise (demanding resources)
- Perfect control is hard to achieve (imperfect compliance): Adverse reaction to an experimental drug means dose has to be reduce no avoid harm
- Unethical: Asking pregnant women to smoke to observe child birth weight Denying the control subjects a drug, e.g. treatment could have been potentially life saving for cancer patients
- Randomisation may influence participation and behaviour

Randomising an instrument

Causal inference from studies in which subject have a final choice

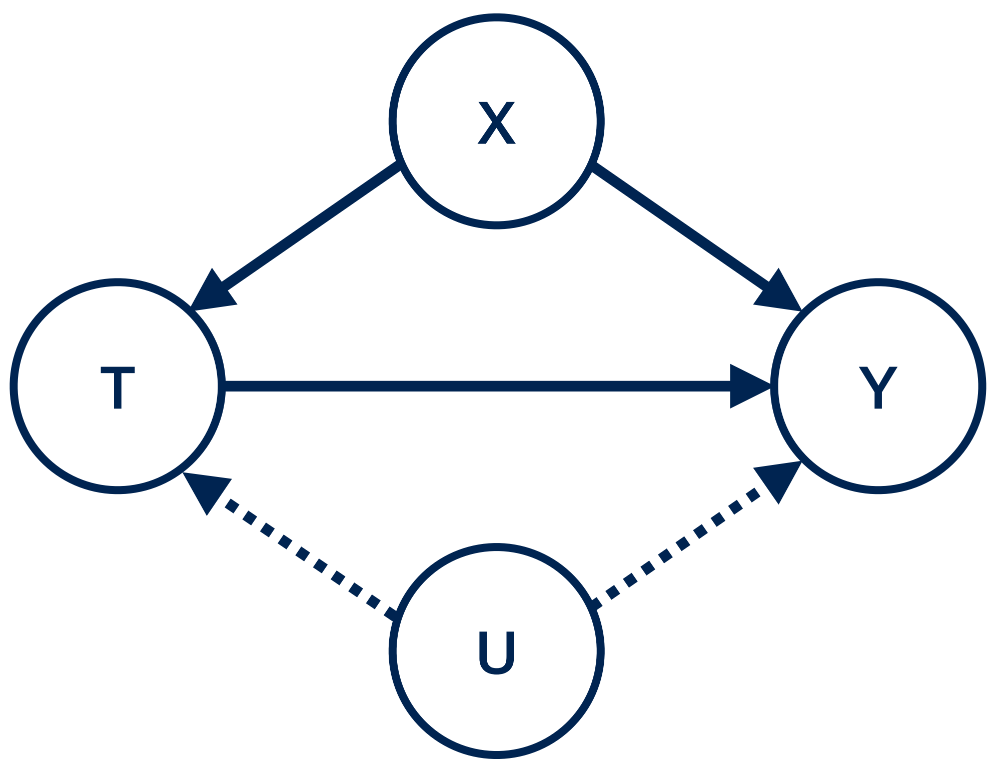
Randomisation is confined to an indirect **instrument** that encourages or discourage participation in treatment or control programmes.

(However, imperfect compliance poses a problem, e.g., subjects that declined taking the drug are precisely those who would have responded adversely. So experiment might conclude the drug is more effective than it actually is.
-> more complex methods, e.g. bounds)

Instrumental Variable

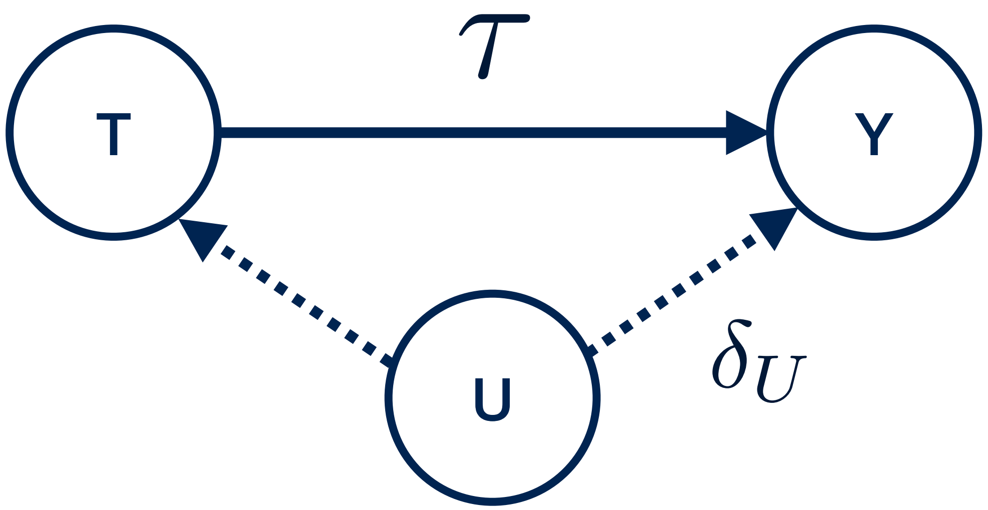
- Unobserved confounders (U), **violates unconfoundedness,** i.e. conditioning on X alone, would not results in a randomised treatment assignment

- Unconfoundedness is fundamentally unverifiable



Rubin 1996

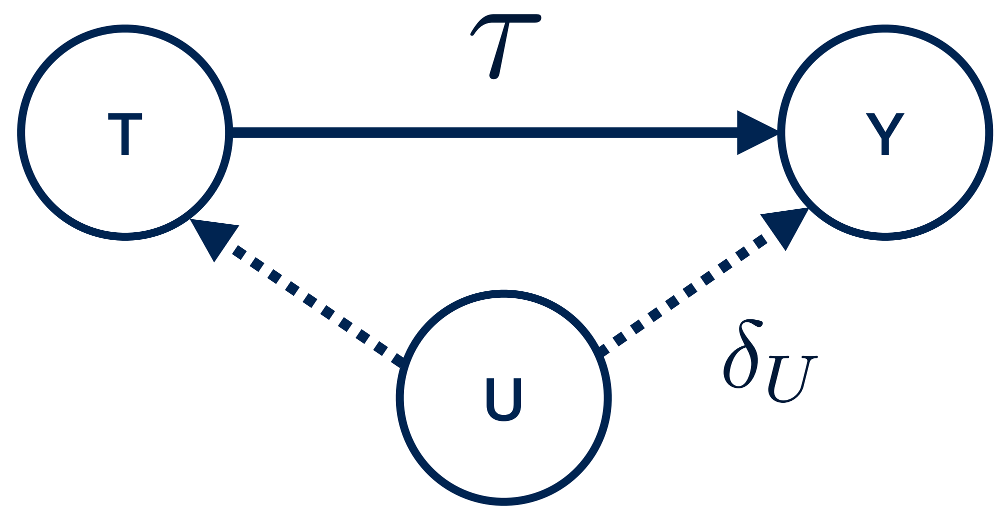
Naive regression lead to bias



$$Y = \tau T + \delta_U U$$

Naive regression lead to bias

why there is no X



What happens if we naively perform a regression of Y on T:

$$Y = \tau T + \delta_U U$$

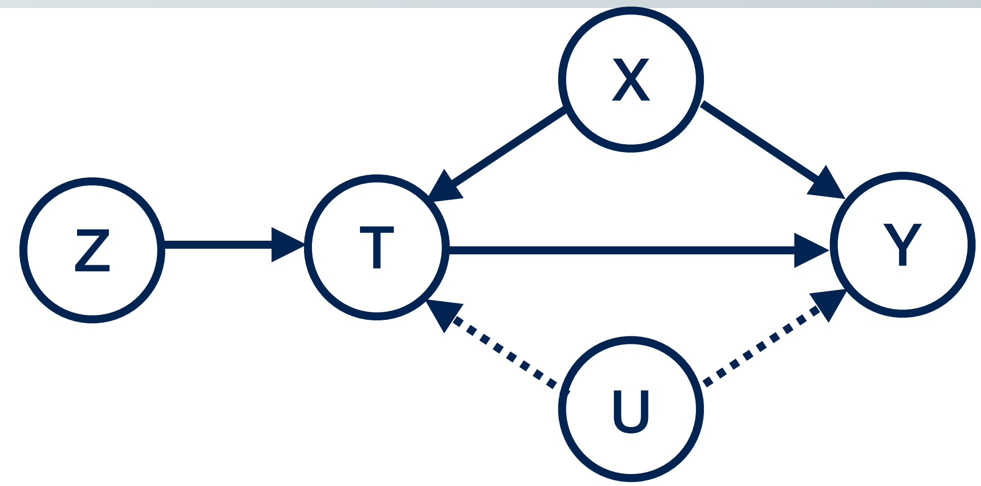
$$\frac{\mathrm{Cov}[T, Y]}{\mathrm{Var}[T]} = \frac{\tau \mathrm{Var}[T] + \tau\delta_U \mathrm{Var}[U]}{\mathrm{Var}[\mathrm{T}]} = \tau + \frac{\tau\delta_U \mathrm{Var}[U]}{\mathrm{Var}[\mathrm{T}]}$$

Bias term

something wrong here

Instrumental Variable example

- Example 1:
  - T: smoking during pregnancy
  - Y: birthweight
  - X: parity, mother's age, weight, …
  - U: Other unmeasured confounders



- Randomise Z (intention-to-treat): either receive encouragement to stop smoking (Z=1), or receive usual care (Z=0)
- Intention-to-treat analysis gives causal effect estimator of encouragement z on outcome y:

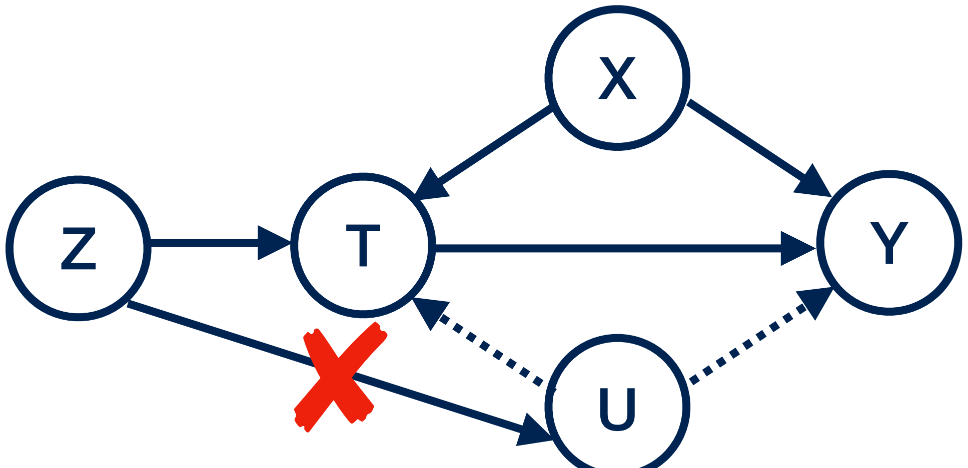
$$\mathbb{E}(y|z=1) - \mathbb{E}(y|z=0)$$

- What can we say about the causal effect of smoking itself?

Instrumental Variable assumptions

- **SUTVA**: Potential outcomes for each individual i are unrelated to the treatment status of other individuals:

$$Y^{(i)}(\mathbf{Z}, \mathbf{T}) = Y^{(i)}(Z^{(i)}, T^{(i)}) \ , \ |\mathbf{Z}| = |\mathbf{T}| = N \text{ individuals}$$



Rubin 1996

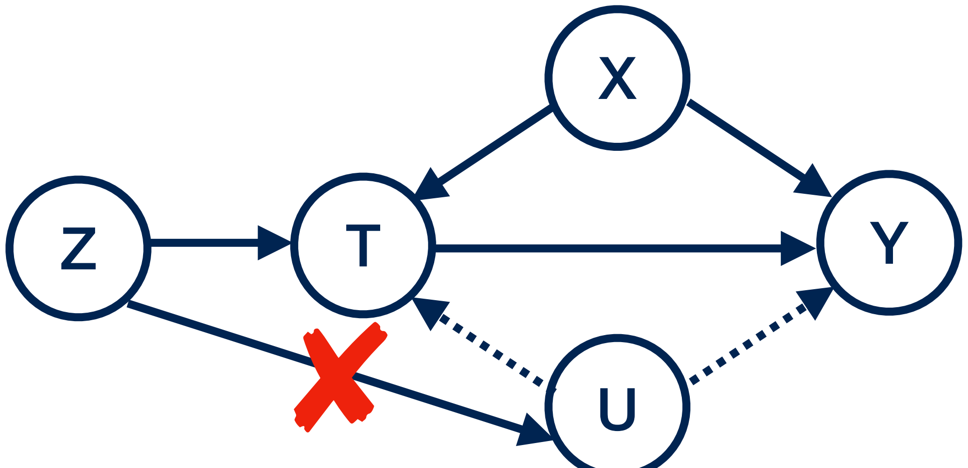
Instrumental Variable assumptions

- **SUTVA**: Potential outcomes for each individual i are unrelated to the treatment status of other individuals:

$$Y^{(i)}(\mathbf{Z},\mathbf{T}) = Y^{(i)}(Z^{(i)},T^{(i)}) \ , \ |\mathbf{Z}| = |\mathbf{T}| = N \text{ individuals}$$

- Treatment assignment Z associated with the treatment is random:

$$P(Z^{(i)}=0) = P(Z^{(i)}=1) \ , \ \forall i$$



Rubin 1996

Instrumental Variable assumptions

- **SUTVA**: Potential outcomes for each individual i are unrelated to the treatment status of other individuals:

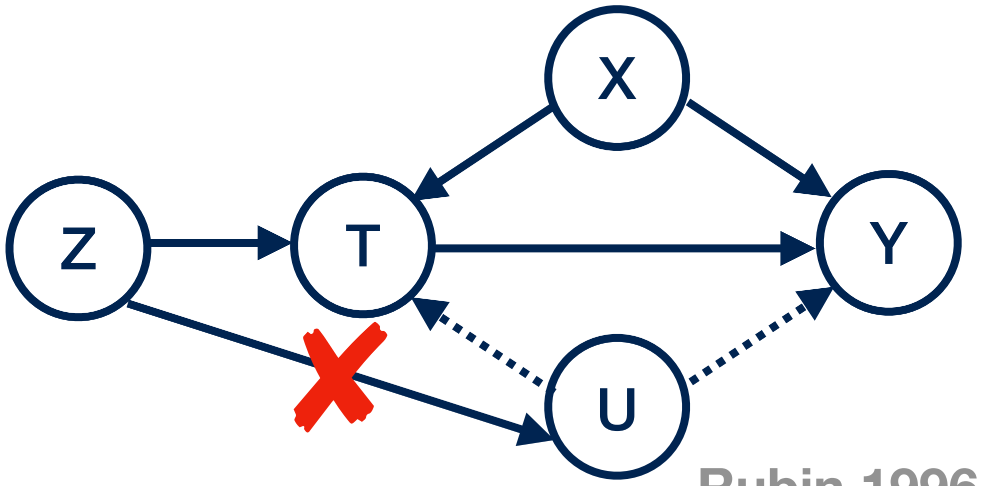
$$Y^{(i)}(\mathbf{Z},\mathbf{T}) = Y^{(i)}(Z^{(i)},T^{(i)}) \ , \ |\mathbf{Z}| = |\mathbf{T}| = N \text{ individuals}$$

- Treatment assignment Z associated with the treatment is random:

$$P(Z^{(i)} = 0) = P(Z^{(i)} = 1) \ , \ \forall i$$

- **Exclusion Restriction**: Any effect of Z on Y is via an effect of Z on T, i.e., Z should not affect Y when T is held constant $\left(Y^{(i)}|z=1,t\right) = \left(Y^{(i)}|z=0,t\right)$

z can only affect y through t



Rubin 1996

Instrumental Variable assumptions

- **SUTVA**: Potential outcomes for each individual i are unrelated to the treatment status of other individuals:

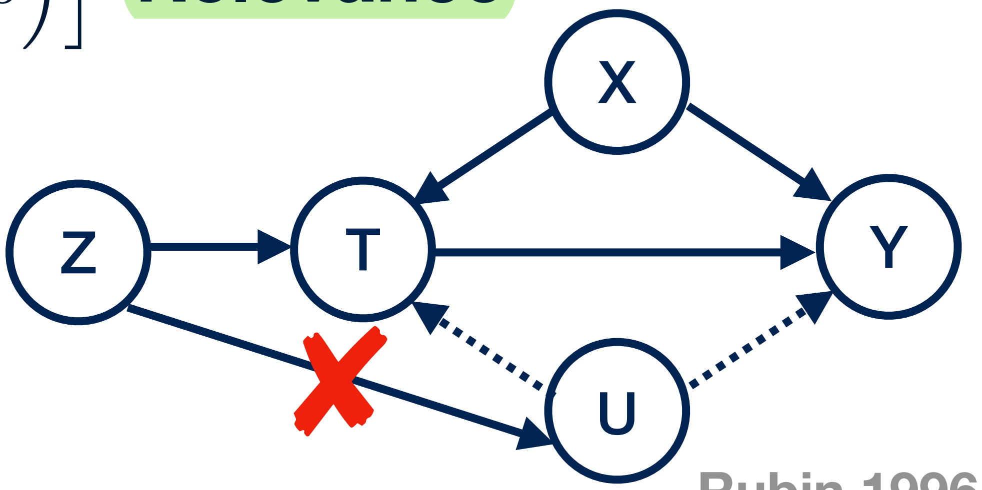
$$Y^{(i)}(\mathbf{Z}, \mathbf{T}) = Y^{(i)}(Z^{(i)}, T^{(i)}) \ , \ |\mathbf{Z}| = |\mathbf{T}| = N \text{ individuals}$$

- Treatment assignment Z associated with the treatment is random:

$$P(Z^{(i)} = 0) = P(Z^{(i)} = 1) \ , \ \forall i$$

- **Exclusion Restriction**: Any effect of Z on Y is via an effect of Z on T, i.e., Z should not affect Y when T is held constant $\left(Y^{(i)} | z = 1, t\right) = \left(Y^{(i)} | z = 0, t\right)$

- **Non-zero Average:** $\mathbb{E}\left[\left(T^{(i)} | z = 1\right) - \left(T^{(i)} | z = 0\right)\right]$ Relevance



Rubin 1996

Instrumental Variable assumptions

- **SUTVA**: Potential outcomes for each individual i are unrelated to the treatment status of other individuals:

$$Y^{(i)}(\mathbf{Z}, \mathbf{T}) = Y^{(i)}(Z^{(i)}, T^{(i)}) \ , \ |\mathbf{Z}| = |\mathbf{T}| = N \text{ individuals}$$

- Treatment assignment Z associated with the treatment is random:

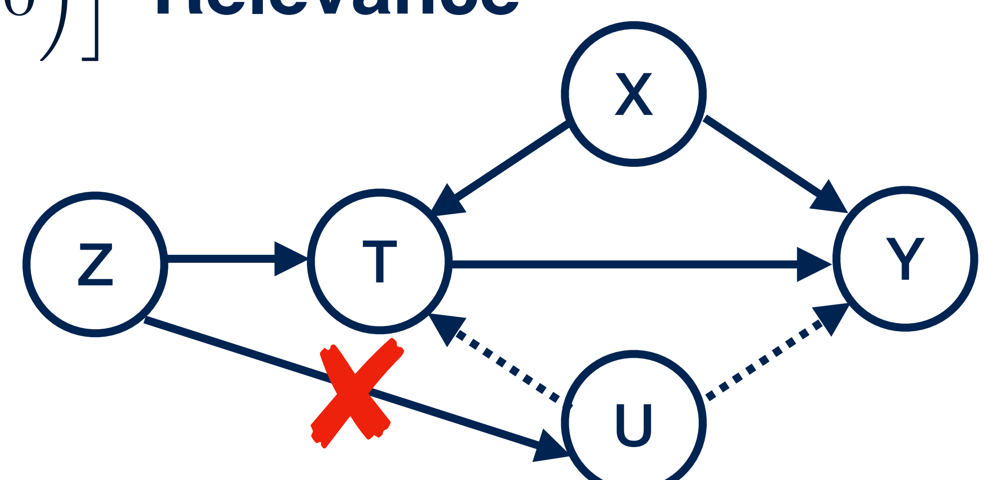
$$P(Z^{(i)} = 0) = P(Z^{(i)} = 1) \ , \ \forall i$$

- **Exclusion Restriction**: Any effect of Z on Y is via an effect of Z on T, i.e., Z should not affect Y when T is held constant $\left(Y^{(i)} | z = 1, t\right) = \left(Y^{(i)} | z = 0, t\right)$

- **Non-zero Average:** $\mathbb{E}\left[\left(T^{(i)} | z = 1\right) - \left(T^{(i)} | z = 0\right)\right]$ **Relevance**

- **Monotonicity** (increasing encouragement "dose" increases probability of treatment, no defiers)**:**

$$\left(T^{(i)} | z = 1\right) \geq \left(T^{(i)} | z = 0\right)$$



Rubin 1996

Instrumental Variable: Potential values of T

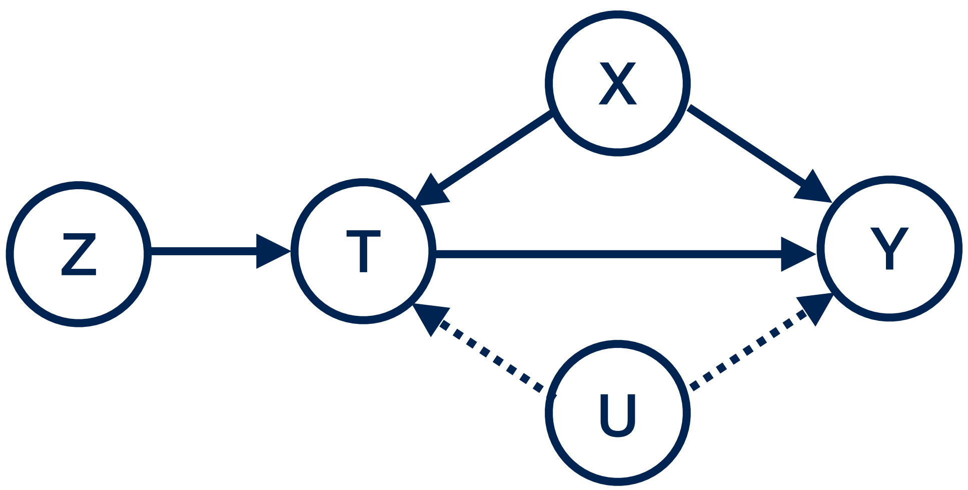
Population	T|z=0	T|z=1	Description
Never-takers	0	0	Causal effect of Z on T is zero, since $\left(T^{(i)} \vert z=1\right) - \left(T^{(i)} \vert z=0\right) = 0$
Compliers	0	1	$\left(T^{(i)} \vert z=1\right) - \left(T^{(i)} \vert z=0\right) = 1$ Treatment received is randomised **causal effect inference**: $\left(Y^{(i)} \vert T^{(i)}=1\right) - \left(Y^{(i)} \vert T^{(i)}=0\right)$
Defiers	1	0	Rule out by **monotonicity**, since $\left(T^{(i)} \vert z=1\right) - \left(T^{(i)} \vert z=0\right) = -1$
Always-takers	1	1	Causal effect of Z on Y is zero, since $\left(T^{(i)} \vert z=1\right) - \left(T^{(i)} \vert z=0\right) = 0$

Notation: T=1 is not smoking

Rubin 1996

Instrumental Variable: The estimand

Want ATE: $$\mathbb{E}\left[(Y|t=1)-(Y|t=0)\right]$$



"Almost"

Will estimate: $$\boxed{\tau = \frac{\mathbb{E}\left[(Y|z=1)-(Y|z=0)\right]}{\mathbb{E}\left[(T|z=1)-(T|z=0)\right]}}$$

Rubin 1996

Instrumental Variable: The estimand

Want ATE: $$\mathbb{E}\left[\left(Y^{(i)}|t^{(i)}=1\right)-\left(Y^{(i)}|t^{(i)}=0\right)\right]$$

$$\boxed{\tau=\frac{\mathbb{E}\left[(Y|z=1)-(Y|z=0)\right]}{\mathbb{E}\left[(T|z=1)-(T|z=0)\right]}}$$

Derivation:

$$\left(Y^{(i)}|T^{(i)}(z=1)\right)-\left(Y^{(i)}|T^{(i)}(z=0)\right)$$ t is either t=0 or t=1, and exclusion restriction

$$=\left[Y^{(i)}\left(t^{(i)}=1\right)\cdot\left(t^{(i)}|z=1\right)+Y^{(i)}\left(t^{(i)}=0\right)\cdot\left(1-\left(t^{(i)}|z=1\right)\right)\right]$$

$$-\left[Y^{(i)}\left(t^{(i)}=1\right)\cdot\left(t^{(i)}|z=0\right)+Y^{(i)}\left(t^{(i)}=0\right)\cdot\left(1-\left(t^{(i)}|z=0\right)\right)\right]$$

$$=\left(Y^{(i)}\left(t^{(i)}=1\right)-Y^{(i)}\left(t^{(i)}=0\right)\right)\cdot\left(\left(t^{(i)}|z=1\right)-\left(t^{(i)}|z=0\right)\right)$$

Hence, the causal effect of Z on Y for individual i, is the product of the causal effect of Z on T, and, the casual effect of T on Y.

Rubin 1996

Instrumental Variable: The estimand

Want ATE: $$\mathbb{E}\left[\left(Y^{(i)}|t^{(i)}=1\right)-\left(Y^{(i)}|t^{(i)}=0\right)\right]$$

$$\boxed{\tau=\frac{\mathbb{E}\left[(Y|z=1)-(Y|z=0)\right]}{\mathbb{E}\left[(T|z=1)-(T|z=0)\right]}}$$

Derivation:

$$\boxed{\left(Y^{(i)}|T^{(i)}(z=1)\right)}-\left(Y^{(i)}|T^{(i)}(z=0)\right)$$

t is either t=0 or t=1, and exclusion restriction

$$=\boxed{\left[Y^{(i)}\left(t^{(i)}=1\right)\cdot\left(t^{(i)}|z=1\right)+Y^{(i)}\left(t^{(i)}=0\right)\cdot\left(1-\left(t^{(i)}|z=1\right)\right)\right]}$$

$$-\left[Y^{(i)}\left(t^{(i)}=1\right)\cdot\left(t^{(i)}|z=0\right)+Y^{(i)}\left(t^{(i)}=0\right)\cdot\left(1-\left(t^{(i)}|z=0\right)\right)\right]$$

$$=\left(Y^{(i)}\left(t^{(i)}=1\right)-Y^{(i)}\left(t^{(i)}=0\right)\right)\cdot\left(\left(t^{(i)}|z=1\right)-\left(t^{(i)}|z=0\right)\right)$$

Hence, the causal effect of Z on Y for individual i, is the product of the causal effect of Z on T, and, the casual effect of T on Y.

Rubin 1996

Instrumental Variable: The estimand

Want ATE: $$\mathbb{E}\left[\left(Y^{(i)}|t^{(i)}=1\right)-\left(Y^{(i)}|t^{(i)}=0\right)\right]$$

$$\boxed{\tau=\frac{\mathbb{E}\left[(Y|z=1)-(Y|z=0)\right]}{\mathbb{E}\left[(T|z=1)-(T|z=0)\right]}}$$

Derivation:

$$\left(Y^{(i)}|T^{(i)}(z=1)\right)-\boxed{\left(Y^{(i)}|T^{(i)}(z=0)\right)}$$

t is either t=0 or t=1, and exclusion restriction

$$=\left[Y^{(i)}\left(t^{(i)}=1\right)\cdot\left(t^{(i)}|z=1\right)+Y^{(i)}\left(t^{(i)}=0\right)\cdot\left(1-\left(t^{(i)}|z=1\right)\right)\right]$$

$$-\boxed{\left[Y^{(i)}\left(t^{(i)}=1\right)\cdot\left(t^{(i)}|z=0\right)+Y^{(i)}\left(t^{(i)}=0\right)\cdot\left(1-\left(t^{(i)}|z=0\right)\right)\right]}$$

$$=\left(Y^{(i)}\left(t^{(i)}=1\right)-Y^{(i)}\left(t^{(i)}=0\right)\right)\cdot\left(\left(t^{(i)}|z=1\right)-\left(t^{(i)}|z=0\right)\right)$$

Hence, the causal effect of Z on Y for individual i, is the product of the causal effect of Z on T, and, the casual effect of T on Y.

Rubin 1996

Instrumental Variable: The estimand

Want ATE: $$\mathbb{E}\left[\left(Y^{(i)}|t^{(i)}=1\right)-\left(Y^{(i)}|t^{(i)}=0\right)\right]$$

$$\boxed{\tau=\frac{\mathbb{E}\left[(Y|z=1)-(Y|z=0)\right]}{\mathbb{E}\left[(T|z=1)-(T|z=0)\right]}}$$

Derivation:

$$\left(Y^{(i)}|T^{(i)}(z=1)\right)-\left(Y^{(i)}|T^{(i)}(z=0)\right)$$ t is either t=0 or t=1, and exclusion restriction

$$=\left[Y^{(i)}\left(t^{(i)}=1\right)\cdot\left(t^{(i)}|z=1\right)+Y^{(i)}\left(t^{(i)}=0\right)\cdot\left(1-\left(t^{(i)}|z=1\right)\right)\right]$$

$$-\left[Y^{(i)}\left(t^{(i)}=1\right)\cdot\left(t^{(i)}|z=0\right)+Y^{(i)}\left(t^{(i)}=0\right)\cdot\left(1-\left(t^{(i)}|z=0\right)\right)\right]$$

$$=\left(Y^{(i)}\left(t^{(i)}=1\right)-Y^{(i)}\left(t^{(i)}=0\right)\right)\cdot\left(\left(t^{(i)}|z=1\right)-\left(t^{(i)}|z=0\right)\right)$$

Hence, the causal effect of Z on Y for individual i, is the product of the causal effect of Z on T, and, the casual effect of T on Y.

Rubin 1996

Instrumental Variable: The estimand

To continue the derivation, we use the fact that:

$$\mathbb{E}\left[XY\right] = \int\int xy\ p(x,y)dxdy = \int dy\ y\ p(y) \int dx\ x\ p(x|y) = \int dy\ y\ p(y)\mathbb{E}[x|y]$$

and write,

$$\mathbb{E}\left[\left(Y^{(i)}|T^{(i)}(z=1)\right) - \left(Y^{(i)}|T^{(i)}(z=0)\right)\right]$$

$$= \mathbb{E}\left[\left(Y^{(i)}\left(t^{(i)}=1\right) - Y^{(i)}\left(t^{(i)}=0\right)\right) \cdot \left(\left(t^{(i)}|z=1\right) - \left(t^{(i)}|z=0\right)\right)\right]$$

0, 1, -1

Rubin 1996

Instrumental Variable: The estimand

To continue the derivation, we use the fact that:

$$\mathbb{E}\left[XY\right] = \int\int xy\ p(x,y)dxdy = \int dy\ y\ p(y) \int dx\ x\ p(x|y) = \int dy\ y\ p(y)\mathbb{E}[x|y]$$

and write,

$$\mathbb{E}\left[\left(Y^{(i)}|T^{(i)}(z=1)\right) - \left(Y^{(i)}|T^{(i)}(z=0)\right)\right]$$

0, 1, -1

$$= \mathbb{E}\left[\left(Y^{(i)}\left(t^{(i)}=1\right) - Y^{(i)}\left(t^{(i)}=0\right)\right)\cdot\left(\left(t^{(i)}|z=1\right) - \left(t^{(i)}|z=0\right)\right)\right]$$

$$= \mathbb{E}\left[\left(Y^{(i)}\left(t^{(i)}=1\right) - Y^{(i)}\left(t^{(i)}=0\right)\right)|\left(\left(t^{(i)}|z=1\right) - \left(t^{(i)}|z=0\right)\right)=1\right]\cdot$$

$$P\left(\left(t^{(i)}|z=1\right) - \left(t^{(i)}|z=0\right)=1\right)$$

$$-\mathbb{E}\left[\left(Y^{(i)}\left(t^{(i)}=1\right) - Y^{(i)}\left(t^{(i)}=0\right)\right)|\left(\left(t^{(i)}|z=1\right) - \left(t^{(i)}|z=0\right)\right)=-1\right]\cdot$$

$$P\left(\left(t^{(i)}|z=1\right) - \left(t^{(i)}|z=0\right)=-1\right)$$

because 0 will eliminate the term

0, by monotonicity

Instrumental Variable: The estimand

$$\frac{\mathbb{E}\left[\left(Y^{(i)}|T^{(i)}(z=1)\right)-\left(Y^{(i)}|T^{(i)}(z=0)\right)\right]}{\mathbb{E}\left[\left(t^{(i)}|z=1\right)-\left(t^{(i)}|z=0\right)\right]}$$

$$=\mathbb{E}\left[\left(Y^{(i)}\left(t^{(i)}=1\right)-Y^{(i)}\left(t^{(i)}=0\right)\right)\middle|\boxed{\left(\left(t^{(i)}|z=1\right)-\left(t^{(i)}|z=0\right)\right)=1}\right]$$

i.e. restricting to *compliers*, the average causal effect of Z on Y is proportional to the average causal effect of T on Y. Rubin 1996

$$\boxed{\tau=\frac{\mathbb{E}\left[(Y|z=1)-(Y|z=0)\right]}{\mathbb{E}\left[(T|z=1)-(T|z=0)\right]}}$$

- In this example, Z was randomly assigned as part of the study
- IV can also be randomised in nature (nature randomiser):
  - Mendelian randomisation
  - Quarter of birth (T=education, Y=earning)

Instrumental Variable: Mendelian Randomisation

Population genetics:

Z = a DNA variant associated with a particular exposure T

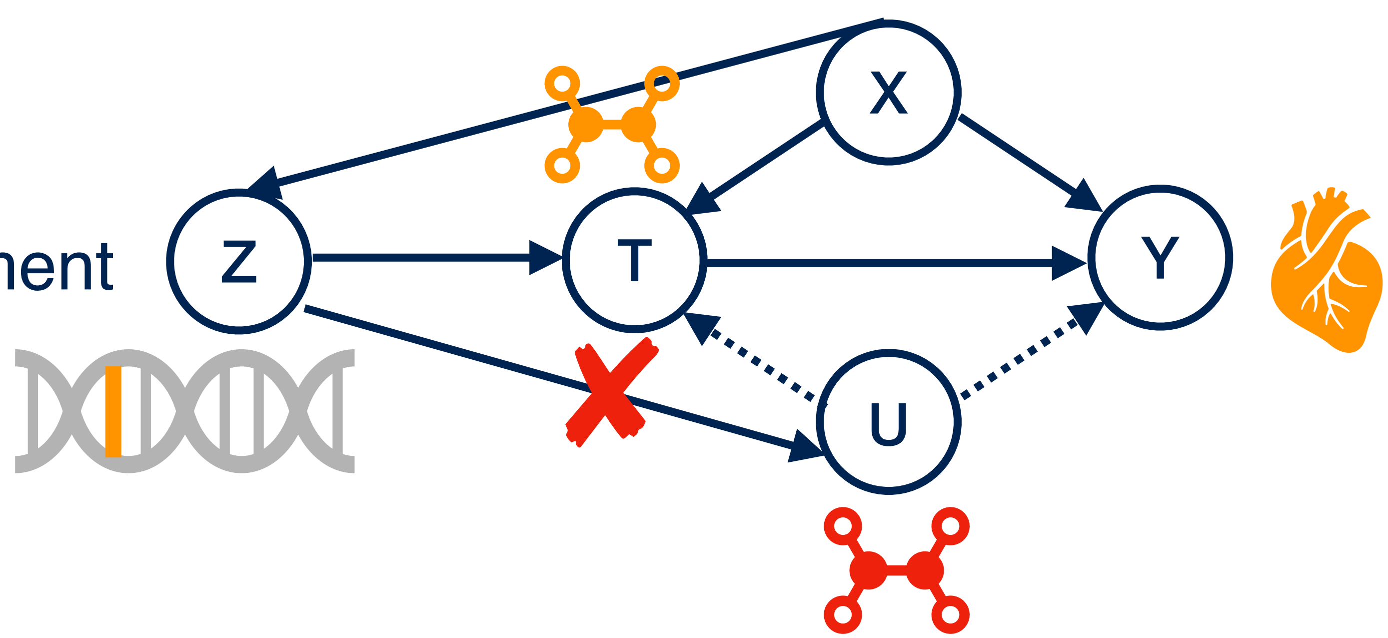
T = exposure, e.g. lipid levels in the blood

Y = heart disease

X = population stratification (might affect Z, need to adjust)

U = unobserved variables affecting both lipid levels and disease

Conditional instrument



Instrumental Variable: Economics

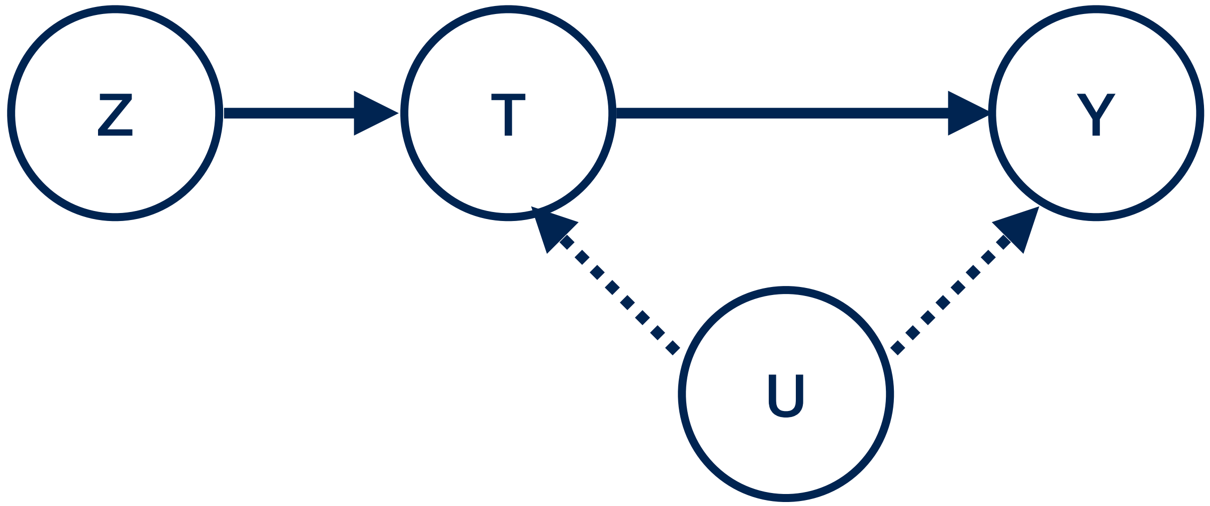
How does price of a product casually affect demand?

Z = Market supply

T = Price

Y = Demand

U = Factors confounding influencing price and demand
(e.g. tax imposed)



Instrumental Variable: Economics

How does price of a product casually affect demand?

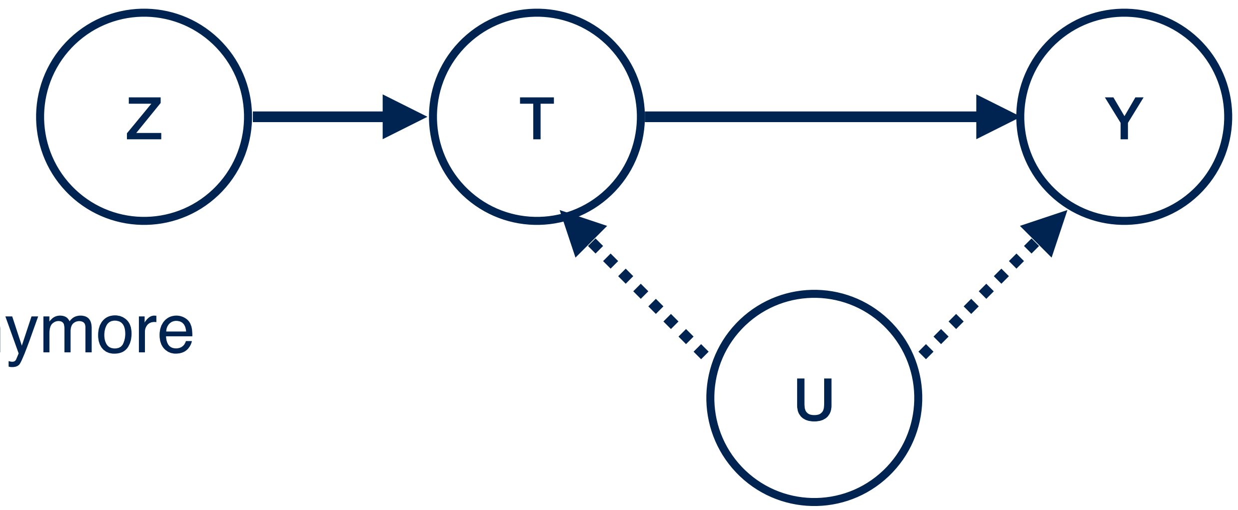
Z = Market supply

T = Price

Y = Demand

U = Factors confounding influencing price and demand
(e.g. tax imposed)

Exclusion restriction requires that market supply does not affect demand
(e.g. COVID-19 toilet paper fiasco!)
(e.g. Pokemon cards)
Also, individuals may not be independent anymore



The Wald Estimator (for binary variables)

$$\boxed{\tau = \frac{\mathbb{E}\left[(Y|z=1)-(Y|z=0)\right]}{\mathbb{E}\left[(T|z=1)-(T|z=0)\right]}}$$

$$\Downarrow$$

$$\boxed{\hat{\tau} = \frac{\frac{1}{n_{z=1}}\sum_{i\in z=1} Y^{(i)} - \frac{1}{n_{z=0}}\sum_{i\in z=0} Y^{(i)}}{\frac{1}{n_{z=1}}\sum_{i\in z=1} T^{(i)} - \frac{1}{n_{z=0}}\sum_{i\in z=0} T^{(i)}}}$$

IV Estimator: continuous variables case

Linear case:

$$\boxed{\tau = \frac{\mathrm{Cov}(Y,Z)}{\mathrm{Cov}(T,Z)}}$$

$$\boxed{\hat{\tau} = \frac{\hat{\mathrm{Cov}}(Y,Z)}{\hat{\mathrm{Cov}}(T,Z)}}$$

Two-Stage Least-squares Estimator

IV Estimator: continuous variables case

Linear case:

$$\boxed{\tau = \frac{\mathrm{Cov}(Y,Z)}{\mathrm{Cov}(T,Z)}}$$

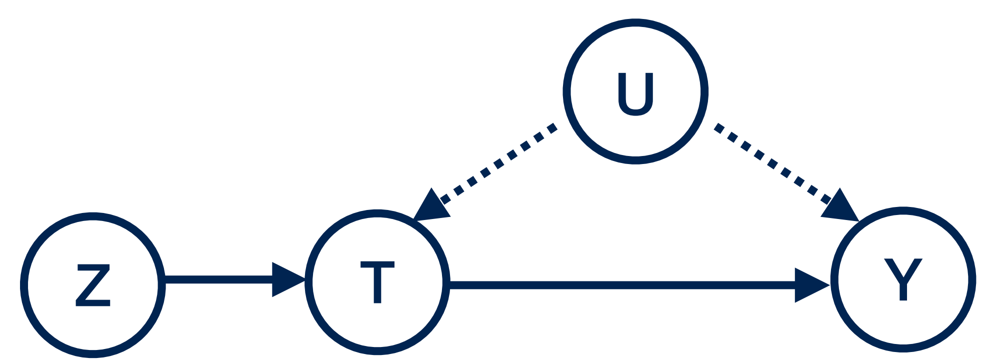
$$\boxed{\hat{\tau} = \frac{\hat{\mathrm{Cov}}(Y,Z)}{\hat{\mathrm{Cov}}(T,Z)}}$$

Two-Stage Least-squares Estimator

IV Estimator: continuous variables case

$$\mathrm{Cov}(Y, Z) = \mathbb{E}[YZ] - \mathbb{E}[Y]\mathbb{E}[Z]$$

$$\boxed{\hat{\tau} = \frac{\hat{\mathrm{Cov}}(Y, Z)}{\hat{\mathrm{Cov}}(T, Z)}}$$



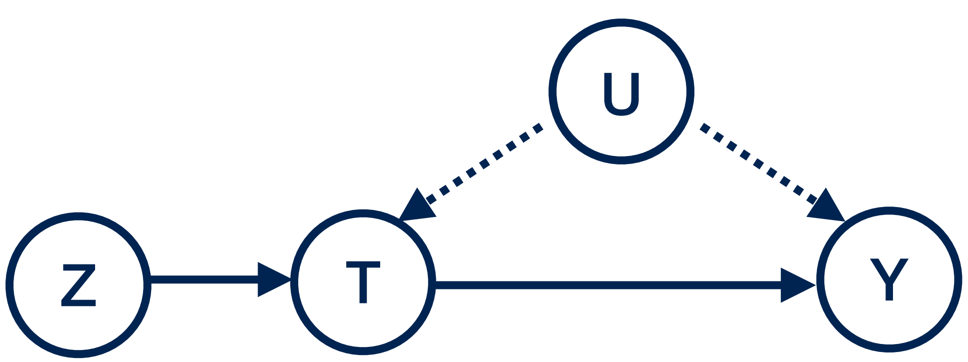
$$Y = \tau T + \delta_U U$$

IV Estimator: continuous variables case

$$\begin{aligned}\mathrm{Cov}(Y, Z) &= \mathbb{E}[YZ] - \mathbb{E}[Y]\mathbb{E}[Z] \\ &= \mathbb{E}(\tau T + \delta_u U)Z] - \mathbb{E}[\tau T + \delta_u U]\mathbb{E}[Z]\end{aligned}$$

By linearity and exclusion restriction

$$\boxed{\hat{\tau} = \frac{\hat{\mathrm{Cov}}(Y, Z)}{\hat{\mathrm{Cov}}(T, Z)}}$$

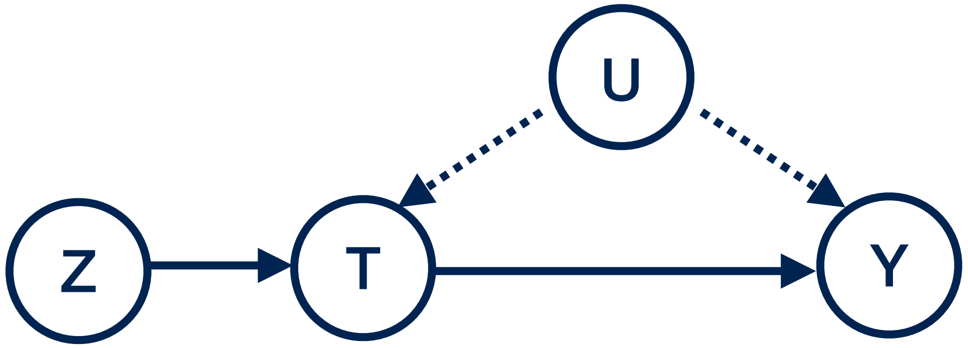


$$Y = \tau T + \delta_U U$$

IV Estimator: continuous variables case

$$\begin{aligned}
\mathrm{Cov}(Y,Z) &= \mathbb{E}[YZ] - \mathbb{E}[Y]\mathbb{E}[Z] \\
&= \mathbb{E}(\tau T + \delta_u U)Z] - \mathbb{E}[\tau T + \delta_u U]\mathbb{E}[Z] \\
&= \tau\mathbb{E}[TZ] + \delta_u\mathbb{E}[UZ] - \tau\mathbb{E}[T]\mathbb{E}[Z] - \delta_u\mathbb{E}[U]\mathbb{E}[Z]
\end{aligned}$$

$$\boxed{\hat{\tau} = \frac{\hat{\mathrm{Cov}}(Y,Z)}{\hat{\mathrm{Cov}}(T,Z)}}$$

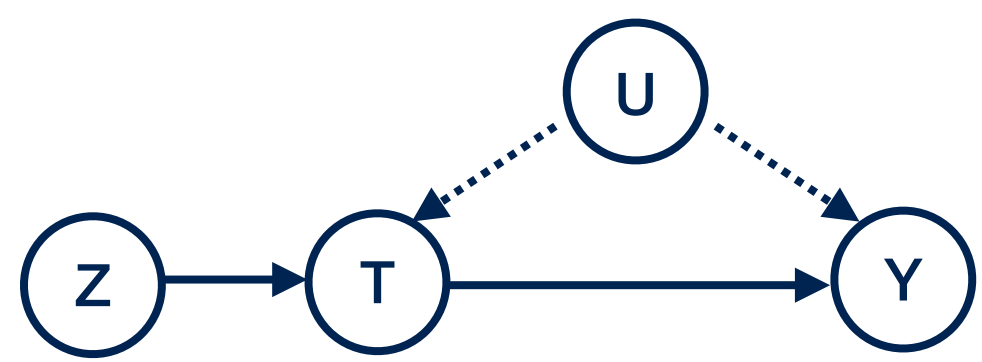


$$Y = \tau T + \delta_U U$$

IV Estimator: continuous variables case

$$
\begin{aligned}
\mathrm{Cov}(Y, Z) &= \mathbb{E}[YZ] - \mathbb{E}[Y]\mathbb{E}[Z] \\
&= \mathbb{E}(\tau T + \delta_u U)Z] - \mathbb{E}[\tau T + \delta_u U]\mathbb{E}[Z] \\
&= \tau \mathbb{E}[TZ] + \delta_u \mathbb{E}[UZ] - \tau \mathbb{E}[T]\mathbb{E}[Z] - \delta_u \mathbb{E}[U]\mathbb{E}[Z] \\
&= \tau \mathrm{Cov}(T, Z) + \delta_U \mathrm{Cov}(U, Z)
\end{aligned}
$$

$$
\boxed{\hat{\tau} = \frac{\hat{\mathrm{Cov}}(Y, Z)}{\hat{\mathrm{Cov}}(T, Z)}}
$$



$$Y = \tau T + \delta_U U$$

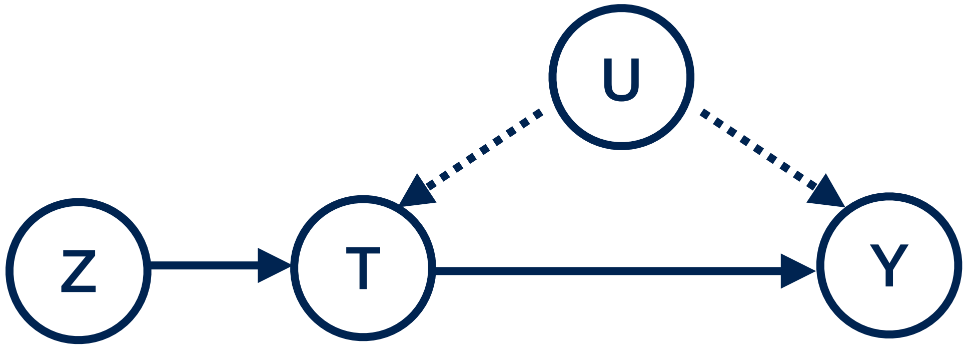
IV Estimator: continuous variables case

$$
\begin{aligned}
\mathrm{Cov}(Y, Z) &= \mathbb{E}[YZ] - \mathbb{E}[Y]\mathbb{E}[Z] \\
&= \mathbb{E}(\tau T + \delta_u U)Z] - \mathbb{E}[\tau T + \delta_u U]\mathbb{E}[Z] \\
&= \tau \mathbb{E}[TZ] + \delta_u \mathbb{E}[UZ] - \tau \mathbb{E}[T]\mathbb{E}[Z] - \delta_u \mathbb{E}[U]\mathbb{E}[Z] \\
&= \tau \mathrm{Cov}(T, Z) + \delta_U \mathrm{Cov}(U, Z) \\
&= \tau \mathrm{Cov}(T, Z)
\end{aligned}
$$

Instrument is not confounded by U

$$\boxed{\hat{\tau} = \frac{\hat{\mathrm{Cov}}(Y, Z)}{\hat{\mathrm{Cov}}(T, Z)}}$$

Non-zero denominator by relevance assumption



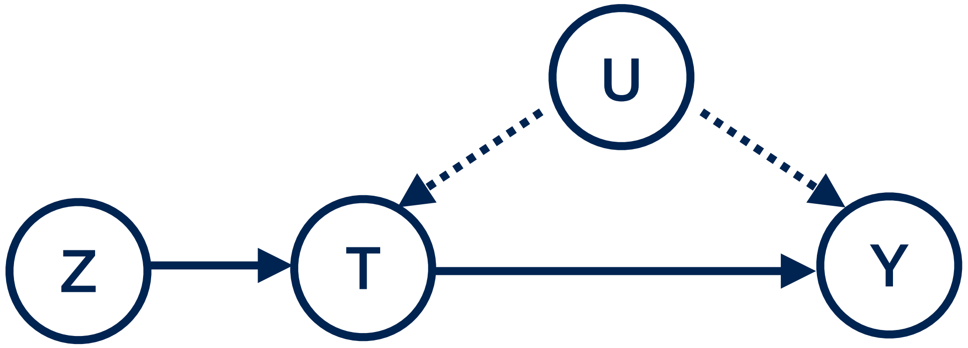
$$Y = \tau T + \delta_U U$$

IV Estimator: continuous variables case

Two-Stage Least Squares Estimator (linear regression):

1. Estimate $\mathbb{E}[T|Z]$, to obtain $\hat{T}$ in subspace Z

2. Estimate $\mathbb{E}[Y|\hat{T}]$, to obtain $\hat{\tau}$, which is the fitted coefficient in front of $\hat{T}$ in this regression.



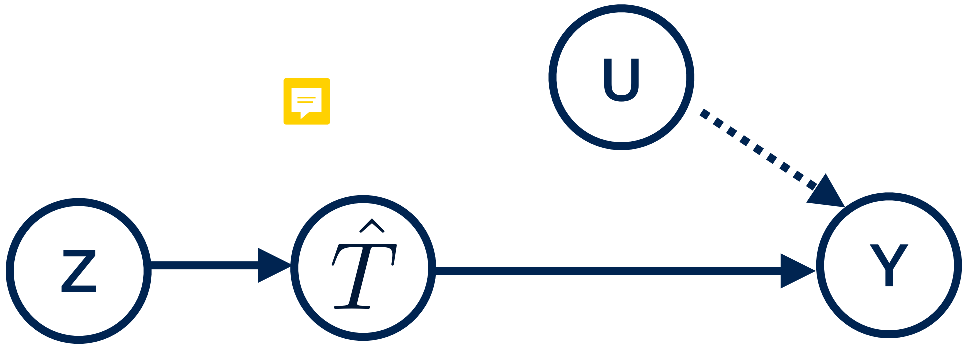
$$Y = \tau T + \delta_U U$$

IV Estimator: continuous variables case

Two-Stage Least Squares Estimator (linear regression):

1. Estimate $\mathbb{E}[T|Z]$, to obtain $\hat{T}$ in subspace Z

2. Estimate $\mathbb{E}[Y|\hat{T}]$, to obtain $\hat{\tau}$, which is the fitted coefficient in front of $\hat{T}$ in this regression.



$$Y = \tau T + \delta_U U$$

Methods for Causal Inference
Lecture 6

Ava Khamseh
School of Informatics

2021-2022