Instrumental variables

Ezequiel Smucler

August 22, 2025

Context

- We have an outcome Y, a binary treatment T and some covariates X.

Context

- We have an outcome Y, a binary treatment T and some covariates X.
- We are interested in estimating the causal effect of T on Y.

Context

- We have an outcome Y, a binary treatment T and some covariates X.
- We are interested in estimating the causal effect of T on Y.
- We suspect that there is unmeasured confounding.

What are instrumental variables?

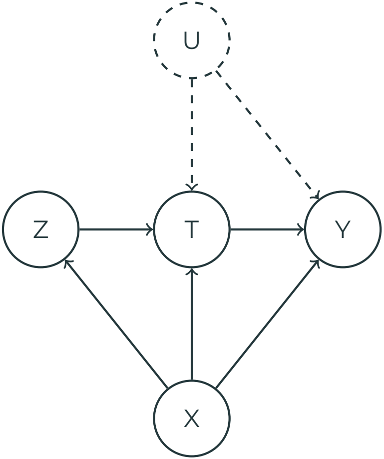
An instrumental variable is a variable that induces treatment, but has no direct effect on the outcome.

What are instrumental variables?

An instrumental variable is a variable that induces treatment, but has no direct effect on the outcome.

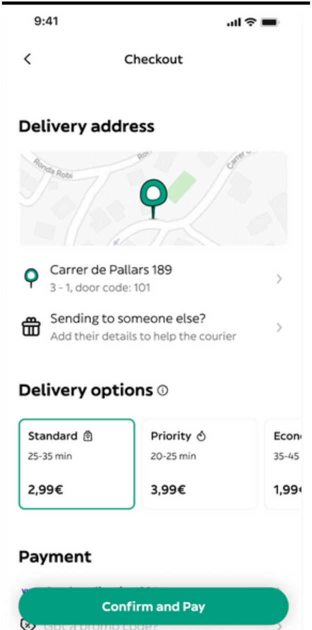
Instrumental variables can be used to estimate **the effect of the treatment on the outcome**, even when there is unmeasured confounding.

What are instrumental variables?



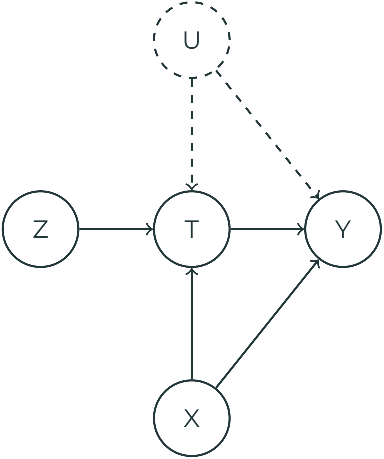
- Z is the instrument.
- T is the treatment.
- Y is an outcome.
- X are observed confounders.
- U are unobserved confounders.

What are instrumental variables?



Consider an experiment in which a random subset of users is offered the possibility of prioritizing their orders for faster delivery.

What are instrumental variables?

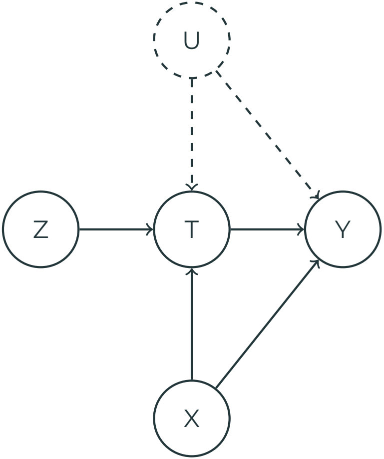


- *Z*: an indicator of a customer being offered priority delivery.
- *T*: an indicator of a customer placing a priority order.
- *Y*: order delivery time.
- *X*: saturation level, average DT, past orders,etc.
- *U*: customer employment status, mood during the experiment, etc.

What are instrumental variables?

Consider an experiment in which a random subset of users **is not** offered the possibility of subscribing to Prime.

What are instrumental variables?

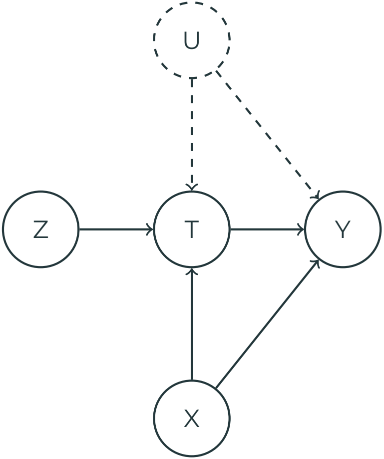


- *Z*: an indicator of a customer being offered Prime.
- *T*: an indicator of a customer signing up to Prime.
- *Y*: orders per customer.
- *X*: past orders, AOV, CM, etc.
- *U*: customer employment status, mood during the experiment, etc.

What are instrumental variables?

Consider an experiment in which a random subset of users is offered an incentive (free delivery) for placing their third food order.

What are instrumental variables?



- *Z*: an indicator of a customer being offered special pricing on her third food order.
- *T*: an indicator of a customer placing an order.
- *Y*: orders per customer.
- *X*: past orders, AOV, CM, etc.
- *U*: customer employment status, mood during the experiment, etc.

Example HVA

	customer_id	Z	A	Y	X
101	85457447	0	1	3	0
714	170142187	1	0	0	0
912	83135127	0	0	0	0
3401	138743257	0	1	1	2
6406	84556306	0	0	0	0

Estimating the effect of HVAs

More generally, estimating the effect of HVAs is important, for at least two reasons:

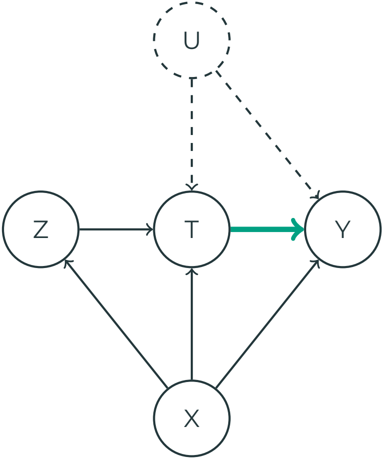
- Calculating ROIs and doing budget allocation.

Estimating the effect of HVAs

More generally, estimating the effect of HVAs is important, for at least two reasons:

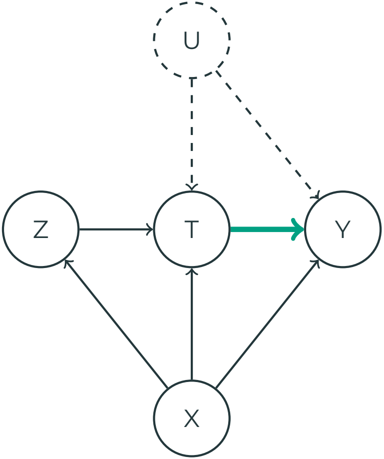
- Calculating ROIs and doing budget allocation.
- Content optimization: which incentives/ads/assets do we show and to whom.

What are instrumental variables?



Interest lies in estimating the causal effect, in some sense, of T **on** Y.

What are instrumental variables?



Interest lies in estimating the causal effect, in some sense, of T **on** Y.

Since there is unmeasured confounding, just adjusting for measured confounders via IPW, DML, etc, won't work.

What are instrumental variables?

In this setting, the average treatment effect

$$ATE = E[Y(t = 1) - Y(t = 0)]$$

is not identified.

What are instrumental variables?

In this setting, the average treatment effect

$$ATE = E[Y(t = 1) - Y(t = 0)]$$

is not identified.

The **effect of the instrument on the outcome** is easy to estimate

$$ITT = E[Y(z = 1) - Y(z = 0)] = E[Y \mid Z = 1] - E[Y \mid Z = 0].$$

This is called the ‘intention to treat’ effect.

What are instrumental variables?

In this setting, the average treatment effect

$$ATE = E[Y(t=1) - Y(t=0)]$$

is not identified.

The **effect of the instrument on the outcome** is easy to estimate

$$ITT = E[Y(z=1) - Y(z=0)] = E[Y \mid Z=1] - E[Y \mid Z=0].$$

This is called the 'intention to treat' effect.

The local average treatment effect is:

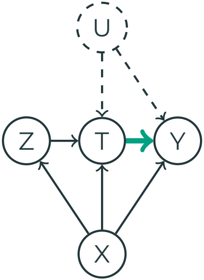
$$LATE = E[Y(t=1) - Y(t=0) \mid T(z=1) > T(z=0)]$$

The local average treatment effect

$$LATE = E[Y(t=1) - Y(t=0) \mid T(z=1) > T(z=0)]$$

This is the effect of treatment on the subset of people that respond to the instrument; these are called the compliers.

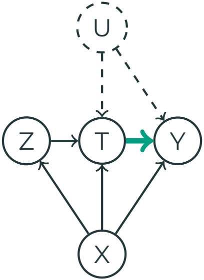
The local average treatment effect



The DAG is saying:

1. No direct effect of Z on Y.
2. No unmeasured confounders between Z and T (**holds if Z is randomized**).

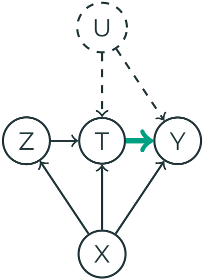
What are instrumental variables?



Under, essentially, the assumptions in the DAG, that there are 'no defiers' and that $cov(T, Z) > 0$,

$$LATE = \frac{E\left\{E\left(Y \mid Z=1, X\right) - E\left(Y \mid Z=0, X\right)\right\}}{E\left\{E\left(T \mid Z=1, X\right) - E\left(T \mid Z=0, X\right)\right\}}$$

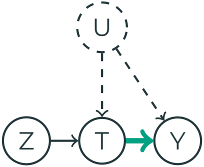
What are instrumental variables?



Under, essentially, the assumptions in the DAG, that there are 'no defiers' and that $cov(T, Z) > 0$,

$$
\begin{aligned}
LATE &= \frac{E\left\{E\left(Y \mid Z=1, X\right) - E\left(Y \mid Z=0, X\right)\right\}}{E\left\{E\left(T \mid Z=1, X\right) - E\left(T \mid Z=0, X\right)\right\}} \\
&= \frac{ATE_{Z \to Y}}{ATE_{Z \to T}}
\end{aligned}
$$

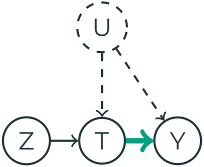
The local average treatment effect



If there is no X, the formula simplies to

$$LATE = \frac{cov(Y, Z)}{cov(T, Z)}$$

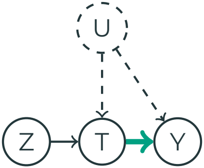
The local average treatment effect



If there is no X, the formula simplies to

$$LATE = \frac{cov(Y, Z)}{cov(T, Z)}$$

This is the ratio of the population regression coefficients of Y on Z and T on Z.

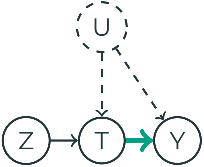


If there is no X, the formula simplies to

$$LATE = \frac{cov(Y, Z)}{cov(T, Z)}$$

This is the ratio of the population regression coefficients of Y on Z and T on Z.

In what follows, φ is the true LATE, and $\widetilde{\varphi}$, $\widehat{\varphi}$, etc, are estimators of it.

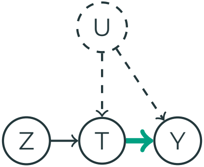


When there is no X, the standard estimator is just

$$\widetilde{\varphi} = \frac{\widehat{cov}(Y, Z)}{\widehat{cov}(T, Z)},$$

the ratio of the sample regression coefficients of Y on Z and T on Z. This is the so-called two stage least squares estimator.

Estimators of the LATE



When there is no X, the standard estimator is just

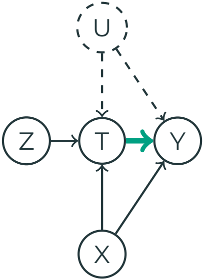
$$\widetilde{\varphi} = \frac{\widehat{cov}(Y, Z)}{\widehat{cov}(T, Z)},$$

the ratio of the sample regression coefficients of Y on Z and T on Z. This is the so-called two stage least squares estimator.

In our running example,

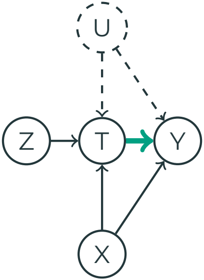
$$\widetilde{\varphi} = 1.89$$

Estimators of the LATE



If the instrument Z is fully randomized, there is no arrow from X to Z, and we don't **need** to use X.

Estimators of the LATE



If the instrument Z is fully randomized, there is no arrow from X to Z, and we don't **need** to use X.

However, using X can lead to more precise estimations. How can we use it?

Estimators of the LATE

Recall that

$$LATE = \frac{E\left\{E\left(Y \mid Z=1, X\right) - E\left(Y \mid Z=0, X\right)\right\}}{E\left\{E\left(T \mid Z=1, X\right) - E\left(T \mid Z=0, X\right)\right\}} = \frac{ATE_{Z \to Y}}{ATE_{Z \to T}}$$

This is the ratio of the treatment effect of Z on Y and the treatment effect of Z on T.

Estimators of the LATE

Recall that

$$LATE = \frac{E\left\{E\left(Y \mid Z=1, X\right) - E\left(Y \mid Z=0, X\right)\right\}}{E\left\{E\left(T \mid Z=1, X\right) - E\left(T \mid Z=0, X\right)\right\}} = \frac{ATE_{Z\rightarrow Y}}{ATE_{Z\rightarrow T}}$$

This is the ratio of the treatment effect of Z on Y and the treatment effect of Z on T.

We can use as an estimator of LATE ratios of any estimators we like of $ATE_{Z\rightarrow Y}$ and $ATE_{Z\rightarrow T}$.

Estimators of the LATE

It can be shown that, in a sense, the optimal estimator of the LATE is the one that is built by taking a ratio of the doubly robust/double ML (DRML) estimators of $ATE_{Z \to Y}$ and $ATE_{Z \to T}$.

Estimators of the LATE

It can be shown that, in a sense, the optimal estimator of the LATE is the one that is built by taking a ratio of the doubly robust/double ML (DRML) estimators of $ATE_{Z \to Y}$ and $ATE_{Z \to T}$.

In our running example, the DRML estimator gives

$$\widehat{\varphi} = 1.90,$$

very similar to the 2SLS estimator.

Building confidence intervals

Consider an estimator of the form

$$\widehat{\varphi} = \frac{\widehat{ATE}_{Z \to Y}}{\widehat{ATE}_{Z \to T}}.$$

Building confidence intervals

Consider an estimator of the form

$$\widehat{\varphi} = \frac{\widehat{ATE}_{Z \to Y}}{\widehat{ATE}_{Z \to T}}.$$

Using the delta method, we can show that, as long as $ATE_{Z \to T} \neq 0$, when the sample size $n \to \infty$,

$$\widehat{\varphi} \approx N\left(\varphi, \frac{\widehat{\sigma}^2}{n}\right),$$

for a certain $\widehat{\sigma}^2$.

Weak instruments

In particular, as long as $ATE_{Z \to T} \neq 0$, when the sample size $n \to \infty$,

$$\widehat{\varphi} \pm 1.96 \frac{\widehat{\sigma}}{\sqrt{n}}$$

will be a 95% level confidence interval.

Example HVA

Using the DoubleML package:

```
Double ML IIVM Model Fit:
      coef   std err        t     P>|t|    2.5 %    97.5 %
d  1.903817  0.742359  2.56455  0.010331  0.44882  3.358814
```

The confidence interval is a bit narrower than the 2SLS one.

Building confidence intervals

However, for any data generating process and a fixed n, the approximation

$$\widehat{\varphi} \approx N\left(\varphi, \frac{\widehat{\sigma}^2}{n}\right),$$

can be very bad if $ATE_{Z \to T} \approx 0$, and the corresponding confidence interval have low coverage.

Building confidence intervals

However, for any data generating process and a fixed n, the approximation

$$\widehat{\varphi} \approx N\left(\varphi, \frac{\widehat{\sigma}^2}{n}\right),$$

can be very bad if $ATE_{Z\to T} \approx 0$, and the corresponding confidence interval have low coverage.

When $ATE_{Z\to T} \approx 0$, we say that **the instrument is weak.**

Simulation

In fact, it can be shown that if a confidence interval for the LATE **does not** have infinite length with positive probability, it will have low coverage at some laws and sample sizes.

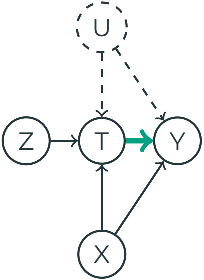
Our recent paper

In a recent paper, Ludovico Lanni, David Masip and myself showed how to construct a confidence **set** that is:

- robust to weak instruments.
- optimal asymptotically,

and contributed an implementation of it to the DoubleML Python package.

Simulation



$$U \sim N(0,1),$$

$$X \sim N(0,1),$$

$$Z \sim \text{Bernoulli}(0.5),$$

$$T = I\{\pi \times Z \times I\{X > 0\} + U\},$$

$$Y = 2 \times \text{sign}(U).$$

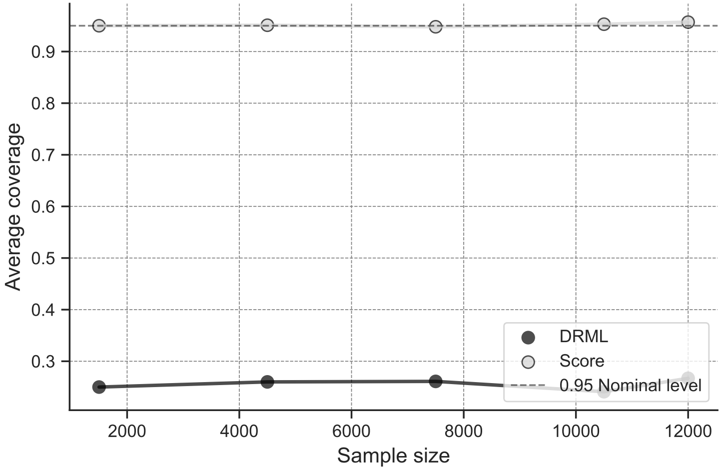


Figure 1: Empirical coverage of the score confidence set and the DRML Wald confidence interval in the weak instrument setting.

Simulation

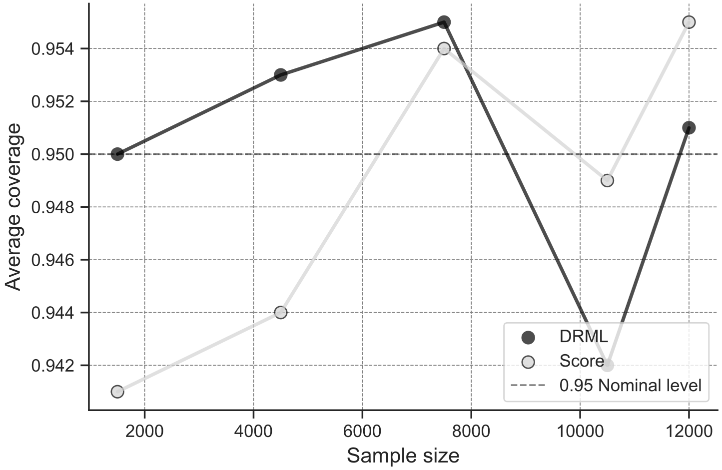


Figure 2: Empirical coverage of the score confidence set and the DRML Wald confidence interval in the strong instrument setting.

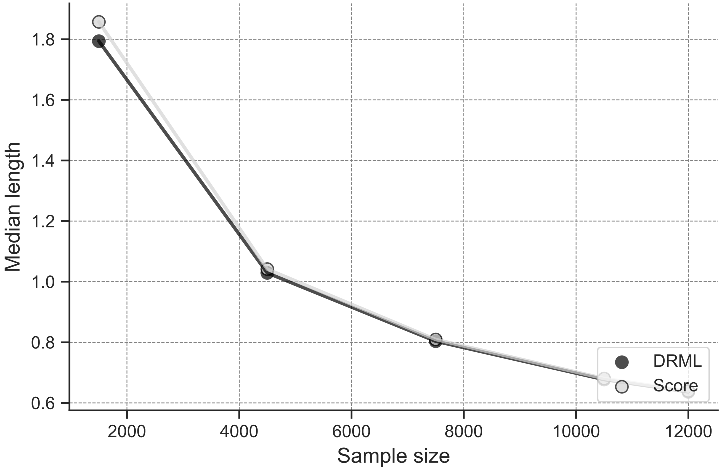


Figure 3: Median length of the score confidence set and the DRML Wald confidence interval in the strong instrument setting.

Example HVA

```
Double ML IIVM Model Fit:
        coef   std err        t     P>|t|    2.5 %    97.5 %
d   1.903817  0.742359  2.56455  0.010331  0.44882  3.358814

Robust score confidence set:
[(-inf, np.float64(-1.6540386736703014)), (np.float64(0.49814335191951703), inf)]
```

Wrap-up

1. Instrumental variables can be used to estimate causal effects even when there is unmeasured confounding between treatment and outcome.

Wrap-up

1. Instrumental variables can be used to estimate causal effects even when there is unmeasured confounding between treatment and outcome.
2. Especially useful for ‘nudging’ experiments.

Wrap-up

1. Instrumental variables can be used to estimate causal effects even when there is unmeasured confounding between treatment and outcome.
2. Especially useful for ‘nudging’ experiments.
3. It is easy to get estimations, confidence intervals, etc, using the DoubleML Python package.

Wrap-up

1. Instrumental variables can be used to estimate causal effects even when there is unmeasured confounding between treatment and outcome.
2. Especially useful for ‘nudging’ experiments.
3. It is easy to get estimations, confidence intervals, etc, using the DoubleML Python package.
4. If you are concerned about weak instruments, calculate the robust score confidence set and check if it has infinite length.

https://causalml-book.org/

https://github.com/DoubleML/doubleml-for-py

https://arxiv.org/abs/2506.10449

https://github.com/david26694/simulations-score-confidence-set