Testing identification in mediation & dynamic treatment models

Martin Huber[1], Kevin Kloiber[2] and Lukáš Lafférs[3]

[1]University of Fribourg, Dept. of Economics
[2]University of Munich, Dept. of Economics
[3]Matej Bel University, Dept. of Mathematics
[3]NHH, Dept. of Economics

EALE Bergen 2024

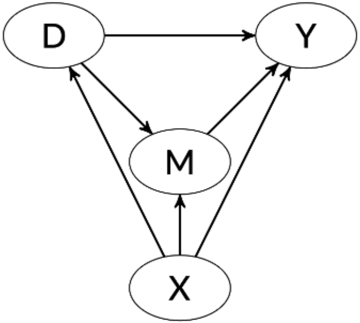
Test for identification

in mediation and dynamic treatment models

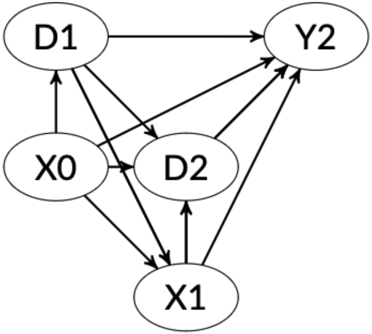
based on

jointly testing sequential ignorability and instrument validity in data.

Mediation



Dynamic treatment effects



Motivation

Motivation

- Identification relies on assumptions that are deemed to be intestable.
- sequential ignorability imposes that the treatment and the mediator is as good as randomly assigned after controlling for observed covariates.
- Whether the set of covariates is sufficient is typically motivated by theory, intuition, domain knowledge or previous empirical findings.
- plausibility of sequential ignorability is often subject to debate.

$\rightarrow$ statistical test for the identifying assumptions.

Contribution

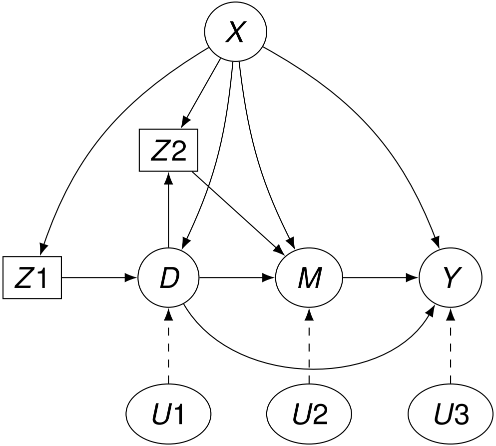
Contribution

- This study introduces a test for conditions that imply sequential ignorability.
- based on Huber and Kueck (2022)
- The testable conditions rely on two types of observables:
  - covariates X to be controlled for,
  - separate suspected instruments for the treatment Z_1 and the mediator Z_2.
- The testable conditions arise if...
  - there is no reverse causality, e.g. $Y \not\to D$, $Y \not\to X$, $Y \not\to Z_1$
  - the respective instruments are relevant (first stage). e.g. $Z_1 \not\perp\!\!\!\perp D|X$

- D: Treatment.
- Y: Outcome.
- M: Mediator.
- X: Covariates.
- Z_1: Suspected instrument for treatment.
- Z_2: Suspected instrument for mediator.
- U: Unobservables.
- $Y(d, m), M(d)$: Potential outcomes and mediators.
- $f(A = a|B = b)$: Cond. density/probability of $A = a$ given $B = b$.

Identification

Causal structure in line with Theorem 1



Assumption 1 (causal structure):

$$M(y) = M, D(m,y,z_2) = D, X(d,m,y,z_2) = X,$$
$$Z_1(d,m,y,z_2) = Z_1, Z_2(m,y) = Z_2,$$

Assumption 2 (common support for D and Z_1):

$$f(D = d, Z_1 = z_1 | M = m, X = x) > 0$$

Assumption 3 (common support for for M and Z_2):

$$f(M = m, Z_2 = z_2 | D = d, X = x) > 0$$

Assumption 4 (conditional dependence of D and Z_1):

$$D \not\perp\!\!\!\perp Z_1 | X = x$$

Assumption 5 (conditional dependence of M and Z_2):

$$M \not\perp\!\!\!\perp Z_2 | D = d, X = x$$

Assumptions 1, 2, 3, 4, 5 will be assumed to hold. We will condition on them being true.

Now I will list assumptions that we construct a test for:

Sequential ignorability $+$ Instruments

Assumptions - sequential ignorability

Assumption 6a

$$Y(d,m) \perp\!\!\!\perp D | X = x$$

Assumption 6b

$$M(d) \perp\!\!\!\perp D | X = x$$

- conditional on covariates X, there exist no confounders jointly affecting D on the one hand and Y or M on the other hand.
- $D \rightarrow M$ and $D \rightarrow Y$

Assumptions - sequential ignorability

Assumption 7

$$Y(d, m) \perp\!\!\!\perp M | D = d, X = x$$

- conditional on D and X, there exist no confounders jointly affecting the mediator M and the outcome Y.
- $M \rightarrow Y$

Assumptions - instruments

Assumption 8a

$$Y(d,m) \perp\!\!\!\perp Z_1 | X = x$$

Assumption 8b

$$M(d) \perp\!\!\!\perp Z_1 | X = x$$

- these rule out confounders jointly affecting Z_1 on the one hand and Y or M on the other hand given X.
- they require that conditional on X, Z_1 does not directly affect M or Y other than through D

Assumptions - instruments

Assumption 9

$$Y(d', m) \perp\!\!\!\perp Z_2 | D = d, X = x$$

- Assumption 9 rules out confounders jointly affecting Z_2 and Y conditional on D and X.
- Assumption 9 requires that Z_2 does not directly affect Y (other than through M) such that $Y(d, m, z_2) = Y(d, m)$ for any value z_2 of Z_2.

Testable implications

$$Y \perp\!\!\!\perp Z_1 | D = d, X = x, \tag{TIa}$$
$$M \perp\!\!\!\perp Z_1 | D = d, X = x, \tag{TIb}$$
$$Y \perp\!\!\!\perp Z_2 | D = d, M = m, X = x \tag{TIc}$$

Theorem 1:

$$\text{Under } \underbrace{1,4,5}_{\substack{\text{causal structure} \\ + \\ \text{relevance conditions}}} : \underbrace{6a, 6b, 7}_{\text{Sequential ignorability}}, \underbrace{8a, 8b, 9}_{\text{Instruments}} \iff \underbrace{(TIa), (TIb), (TIc)}_{\text{Testable implications}}.$$

Limitations

- Counterfactual values $d' \neq d$ or $m' \neq m$ cannot be tested for subjects with $D = d$ and $M = d$.
- $\rightarrow$ violations exclusively concerning counterfactual rather than (f)actual outcomes and mediators cannot be detected.
- However, it seems unlikely that violations exclusively occur among counterfactual, but never among factual outcomes and mediators, because this would imply very specific models.

Identified causal effects

- $D \rightarrow Y$ by Assumption 6a (see de Luna and Johansson, 2014, or Huber and Kueck, 2022).
- $D \rightarrow M$ by Assumption 6b.
- $M \rightarrow Y$ by Assumption 7.
- $(D, M) \rightarrow Y$, e.g. $E[Y(d,m) - Y(d',m')]$, including the controlled direct effect $E[Y(d,m) - Y(d',m)]$, by Assumptions 6a and 7 (see e.g. Robins and coauthors).
- $D \rightarrow Y|M = 1$ The effect of D on Y in sample selection models, where M indicates the observability of Y (but does not affect Y such that $Y(d,m)$ is $Y(d)$), by Assumptions 6a and 7 (as assumed by Bia, Huber, and Lafférs, 2023).

Natural direct and indirect effects

- Assumption 6a, 6b, and 7 are not sufficient for identifying natural direct and natural indirect effects, like $E[Y(d, M(d)) - Y(d', M(d))]$ and $E[Y(d, M(d)) - Y(d, M(d'))]$.
- Pearl (2001) suggests an additional counterfactual assumption yielding identification:

$$Y(d, m) \perp\!\!\!\perp M(d') | X = x$$

- The latter assumption and Assumptions 6a and 6b are implied by the following assumption of Imai, Keele, and Yamamoto (2010):

$$\{Y(d, m), M(d')\} \perp\!\!\!\perp D | X = x$$

- We cannot test this conditional independence for joint counterfactuals, but testing Assumptions 6a and 6b for actual outcomes arguably has **nontrivial power** against its violation.

Proof of Theorem 1

Analytical approach

- follows Huber and Kueck (2022).

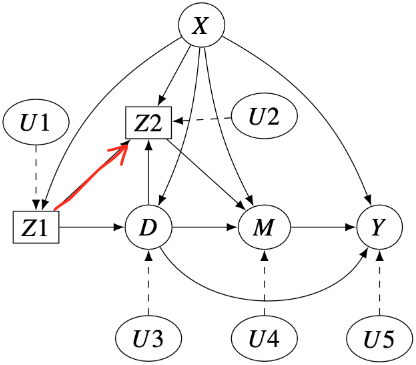
Computational approach

- We translate assumptions into DAG semantics.
- Conduct an exhaustive search in the space of DAGs.
- Verify the theorem directly.

▸ Details on the computational approach

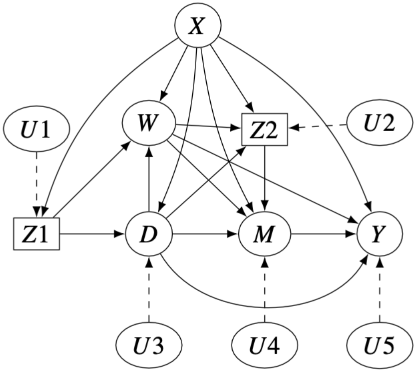
Theorem 2:

Controlling for the Second instrument



Theorem 3:

Post treatment covariates



Testing

Testing

Null hypothesis:

- Denote by $\mu_B(a) = E(B|A = a)$ the conditional mean of B given $A = a$.
- The null hypothesis is given by

$$H_0 : 0 = \theta := E\begin{pmatrix} (\mu_Y(D,X) - \mu_Y(D,X,Z_1))^2 \\ (\mu_M(D,X) - \mu_M(D,X,Z_1))^2 \\ (\mu_Y(D,M,X) - \mu_Y(D,M,X,Z_2))^2 \end{pmatrix}.$$

Details on Testing

Simulation

Main setup (Theorem 1)

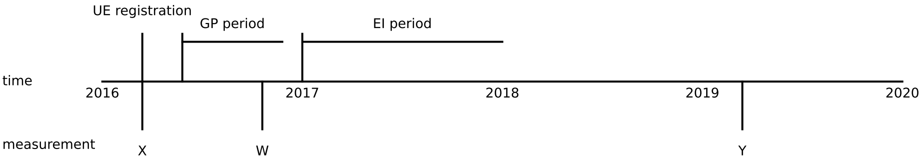
$$
\begin{aligned}
D &= I\{X'\beta + 0.5Z_1 + U_1 > 0\}, \\
M &= 0.5D + 0.5Z_2 + X'\beta + \delta U_1 + U_2, \\
Y &= D + 0.5M + X'\beta + \gamma Z_1 + \gamma Z_2 + \delta U_1 + U_3, \\
X &\sim \mathcal{N}(0, \sigma_X^2), Z_1 \sim \mathcal{N}(0,1), Z_2 \sim \mathcal{N}(0,1), \\
U_1 &\sim \mathcal{N}(0,1), U_2 \sim \mathcal{N}(0,1), U_3 \sim \mathcal{N}(0,1),
\end{aligned}
$$

δ – confounding
γ – exclusion restriction violation

sample size	rej. rate	mean pval
$\delta = 0$ & $\gamma = 0$		
1000	0.044	0.513
4000	0.047	0.510
$\delta = 1$ & $\gamma = 0$		
1000	0.688	0.122
4000	1.000	0.000
$\delta = 0$ & $\gamma = 0.2$		
1000	0.086	0.447
4000	1.000	0.000

▸ Simulations for Theorem 2

Empirical Illustration


UE registration
GP period
EI period
time
2016
2017
2018
2019
2020
measurement
X
W
Y

Dynamic treatment effects of Slovak labor market programs: administrative data on job seekers in Slovakia previously analyzed by Lafférs and Štefánik (2024).

- D is six-month training starting in 2016 named ***Graduate practice***.
- M is ***Employment incentives*** program (combines hiring incentives with subsidized employment) starting in 2017 (typically one year).
- Y is employment indicator in 2019.
- Z_1 is local availability of D and corresponds to the ratio of jobseekers enrolled in intervention D in the previous year (2015); analogous method is used to compute Z_2 related to M.
- Pre-treatment covariates X (264 variables): regional information, marital status, dependents, education and skills, employment histories, prior unemployment benefits, willingness to relocate for work, health information, and caseworker assessments of employability.
- Five post-treatment covariates (W) that might affect both M and Y: participation in programs other than D during treatment period, absence from the unemployment register, application for minimum subsistence benefits.

Application

teststat	se	pval	effect	effect_se	effect_pval	effect_ntrimmed
0.00042	0.00036	0.24189	0.0855	0.0249	0.0006	6,288

Results with limited X and without W

p-value $= 0.242 \rightarrow$ p-value $= 0.069$

Conclusion

Conclusion

- Joint test for instrument validity and sequential ignorability in dynamic treatment and mediation models.
- Machine learning-based procedure allowing for high-dimensional control variables.
- Application to labor market data from Slovakia.
- `testmedident()` in package `causalweight`

Thank you.

`www.lukaslaffers.com`

Computational approach

Construct all the DAGs with observed Y, D, Z_1, Z_2.
We don't need to consider

- unobserved colliders, as these paths are closed anyway,
- unobserved mediators, as these can be interpreted as direct paths,
- unobserved confounders for more than two observed variables, as these are equivalent to the existence of multiple pair-wise confounders from the point of view of existence of open paths and hence identification.
- X - because everything is conditional on X

Potential outcomes $\rightarrow \rightarrow \rightarrow$ DAG semantics

$$M(y) = M, D(m, y, z_2) = D, X(d, m, y, z_2) = X, Z_1(d, m, y, z_2) = Z_1, \quad (1)$$
$$Z_2(m, y) = Z_2$$
$$D \not\perp\!\!\!\perp Z_1 | X = x \quad (4)$$
$$M(d) \perp\!\!\!\perp D | X = x \quad (6b)$$
$$Y \perp\!\!\!\perp Z_1 | D = d, X = x \quad (TIa)$$

$\rightarrow \rightarrow \rightarrow$ translated into $\rightarrow \rightarrow \rightarrow$

There are no directed paths in the following directions: (1)

$Y \rightarrow M, Y \rightarrow X, Y \rightarrow Z_1, Y \rightarrow Z_2, M \rightarrow D, M \rightarrow X, M \rightarrow Z_1, M \rightarrow Z_2,$

$Z_2 \rightarrow D, Z_2 \rightarrow X, Z_2 \rightarrow Z_1, Z_2 \rightarrow D, D \rightarrow X, D \rightarrow Z_1$ in graph G

D and Z_1 are d-connected with conditioning set $\{X\}$ in graph G (4)

M and D are d-separated with conditioning set $\{X\}$ in graph G_D (6b)

Y and Z_1 are d-separated with conditioning set $\{X, D\}$ in graph G (TIa)

Computational approach - Theorem 1

Direct and principled way.

- There are 1048576 DAGs that satisfy (1).
- There are 735232 DAGs that satisfy assumptions (1), (4), (5). Out of these
  (i) 480 DAGs satisfy (6a), (6b), (8a), (8b), (7), (9) and at the same time, satisfy (TIa), (TIb), (TIc),
  (ii) 73043 (=73523-480) DAGs that do not satisfy (6a), (6b), (8a), (8b), (7), (9) and at the same time, do not satisfy (TIa), (TIb), (TIc).

Back

Testing

Score function for testing:

- Testing is based on the following score function (in analogy to Huber and Kueck, 2022), which is Neyman-orthogonal and asymptotically normal under the null:

$$\phi(V,\theta,\eta) = (\eta_1(V) - \eta_2(V))^2 - \theta + \zeta.$$

- $V = (Y, D, M, X, Z_1, Z_2)$,
- $\eta_1(V) = (\mu_Y(D,X), \mu_M(D,X), \mu_Y(D,M,X))'$,
  $\eta_2(V) = (\mu_Y(D,X,Z_1), \mu_M(D,X,Z_1), \mu_Y(D,M,X,Z_2))'$,
- ζ is an independent mean-zero random variable with variance $\sigma_\zeta^2 > 0$ to avoid the test statistic to be degenerate under the null.
- $\eta_1(V)$, $\eta_2(V)$ may be estimated by machine learning with cross-fitting (see e.g. Chernozhukov et al. 2018) if X is high-dimensional.

Back to Testing

$Z_1 \to Z_2$ (Theorem 2)

$$
\begin{aligned}
D &= I\{X'\beta + 0.5Z_1 + U_1 > 0\}, \\
M &= 0.5D + 0.5Z_2 + X'\beta + \delta U_1 + U_2, \\
Y &= D + 0.5M + X'\beta + \gamma Z_1 + \gamma Z_2 + \delta U_1 + U_3, \\
X &\sim \mathcal{N}(0, \sigma_X^2), Z_1 \sim \mathcal{N}(0,1), Z_2 = U_4 + 0.5Z_1 \\
U_1 &\sim \mathcal{N}(0,1), U_2 \sim \mathcal{N}(0,1), U_3 \sim \mathcal{N}(0,1), U_4 \sim \mathcal{N}(0,1)
\end{aligned}
$$

sample size	rej. rate	mean pval
$\delta = 0$ & $\gamma = 0$		
1000	0.042	0.514
4000	0.049	0.510
$\delta = 1$& $\gamma = 0$		
1000	0.297	0.286
4000	1.000	0.000
$\delta = 0$ & $\gamma = 0.2$		
1000	0.234	0.318
4000	1.000	0.000

Back to Simulation Theorem 1