Professor: Chiara Binelli
Academic year: 2018/2019
Course: Economics for Data Science

Daniele Livio

IDENTIFICATION OF CAUSAL EFFECTS IN AMAZON'S RECOMMENDER SYSTEM

Split-Door Criterion

(Sharma, Hofman & Watts, 2016)

1. Recommender System
2. CTR naive method and related problems
3. Classical techniques for causal estimation
4. Split-Door criterior
5. Split-Door assumptions
6. Application on Amazon recommender system
7. Further applications
8. Comparison with Bayesian Structural Time Series

Would you like to buy also... ?

User visits product i, which suggests product j

How much activity comes from recommendation systems?

How much activity comes *because* of the recommendation system?

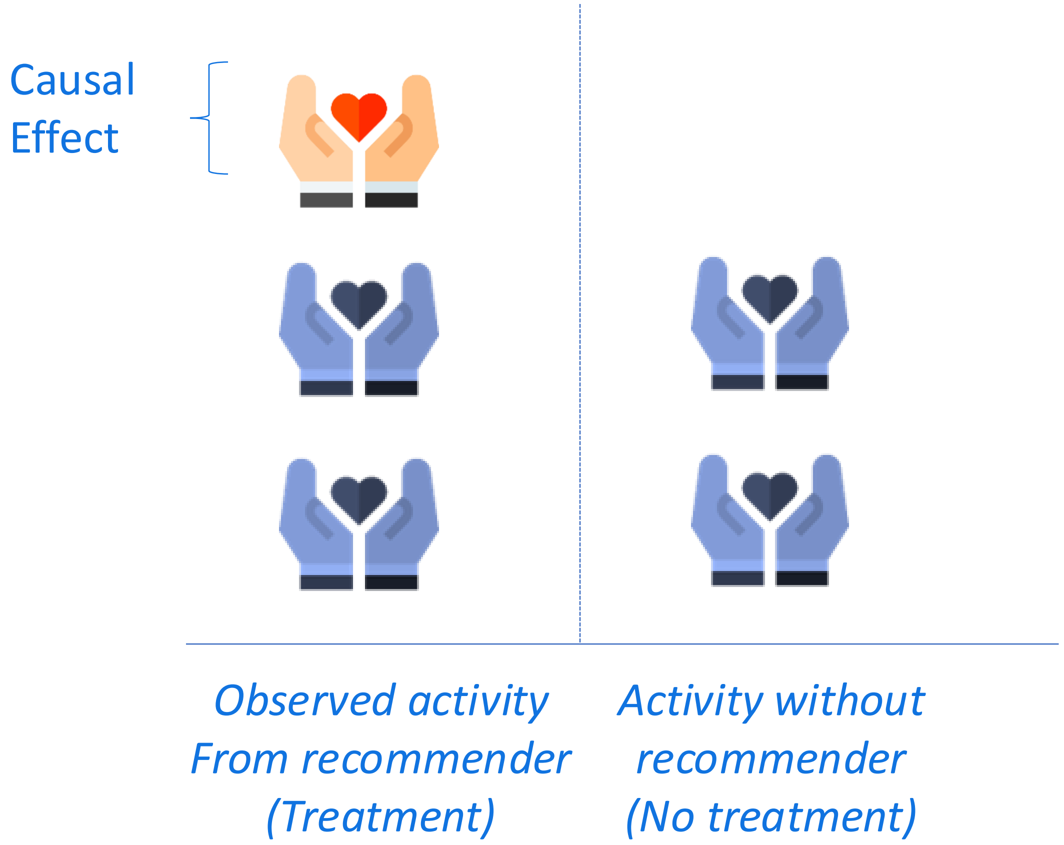
Overestimation of the impact of an advertising campaign

$$CTR_{i,j} = \frac{Clicks\ on\ recommendation\ j}{Visit\ on\ focal\ product\ i} * 100$$

Unrealistic assumption:

"If not suggested through *The Road*,
users would never click
No Country For Old Men"

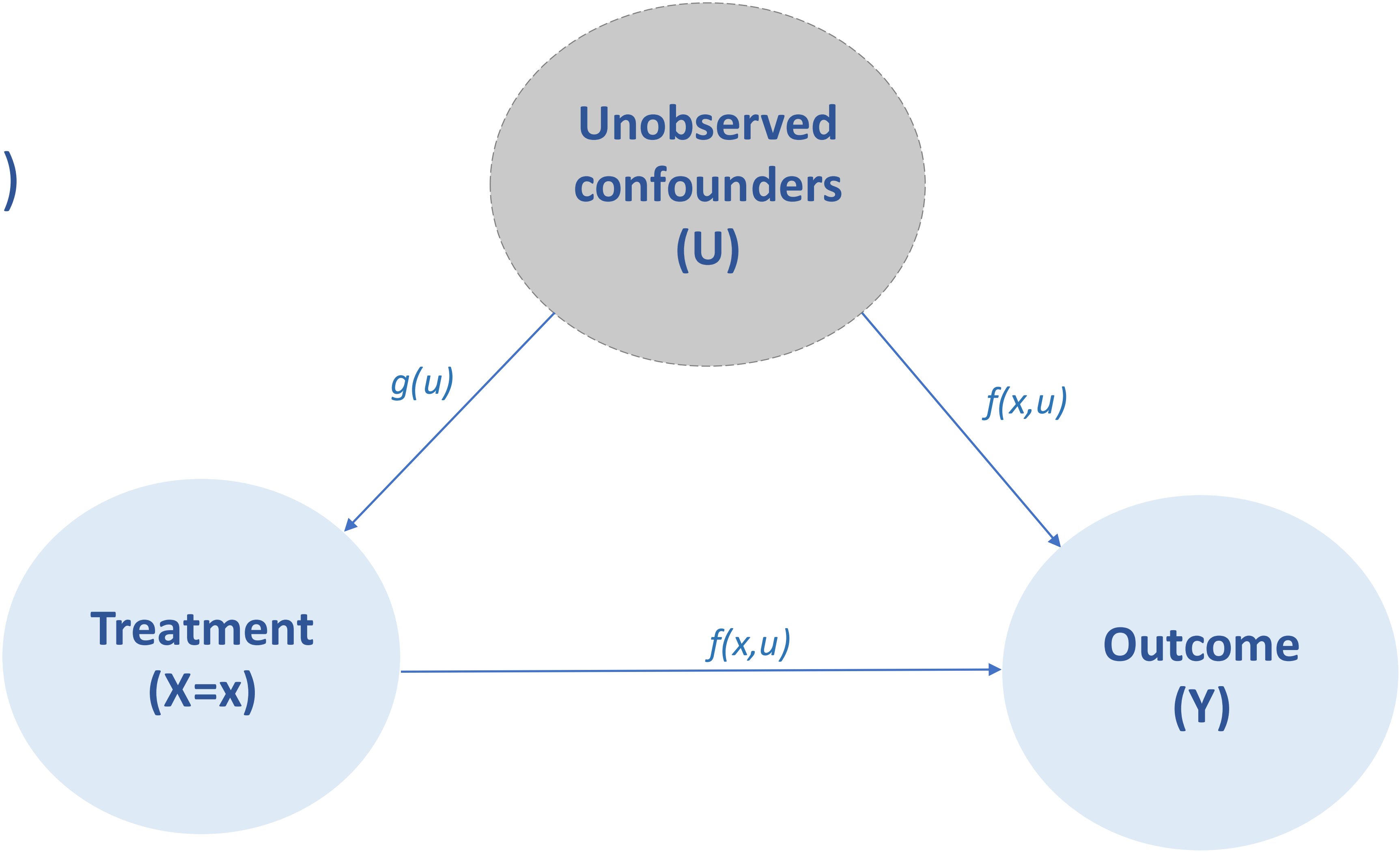
It can only answer question 1!



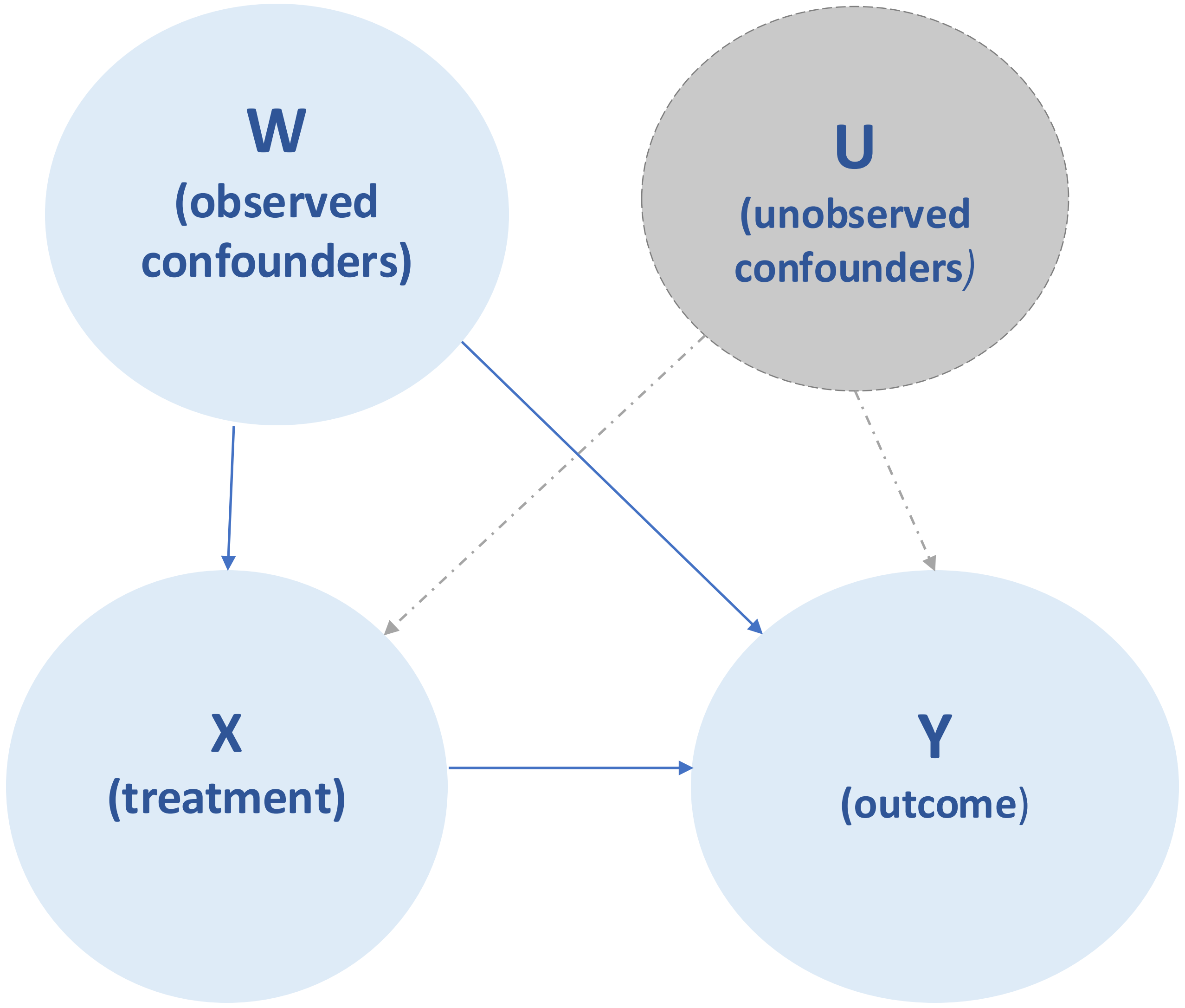
Unobserved correlated demand (confounder)

Unobserved correlated demand (confounder)

Y=f(x,u)
X=g(u)



Backdoor Criterion



DESCRIPTION

Condition on observed confounders W to isolate the treatment effect

LIMITATIONS

Unlikely that there are no unobserved confounders U

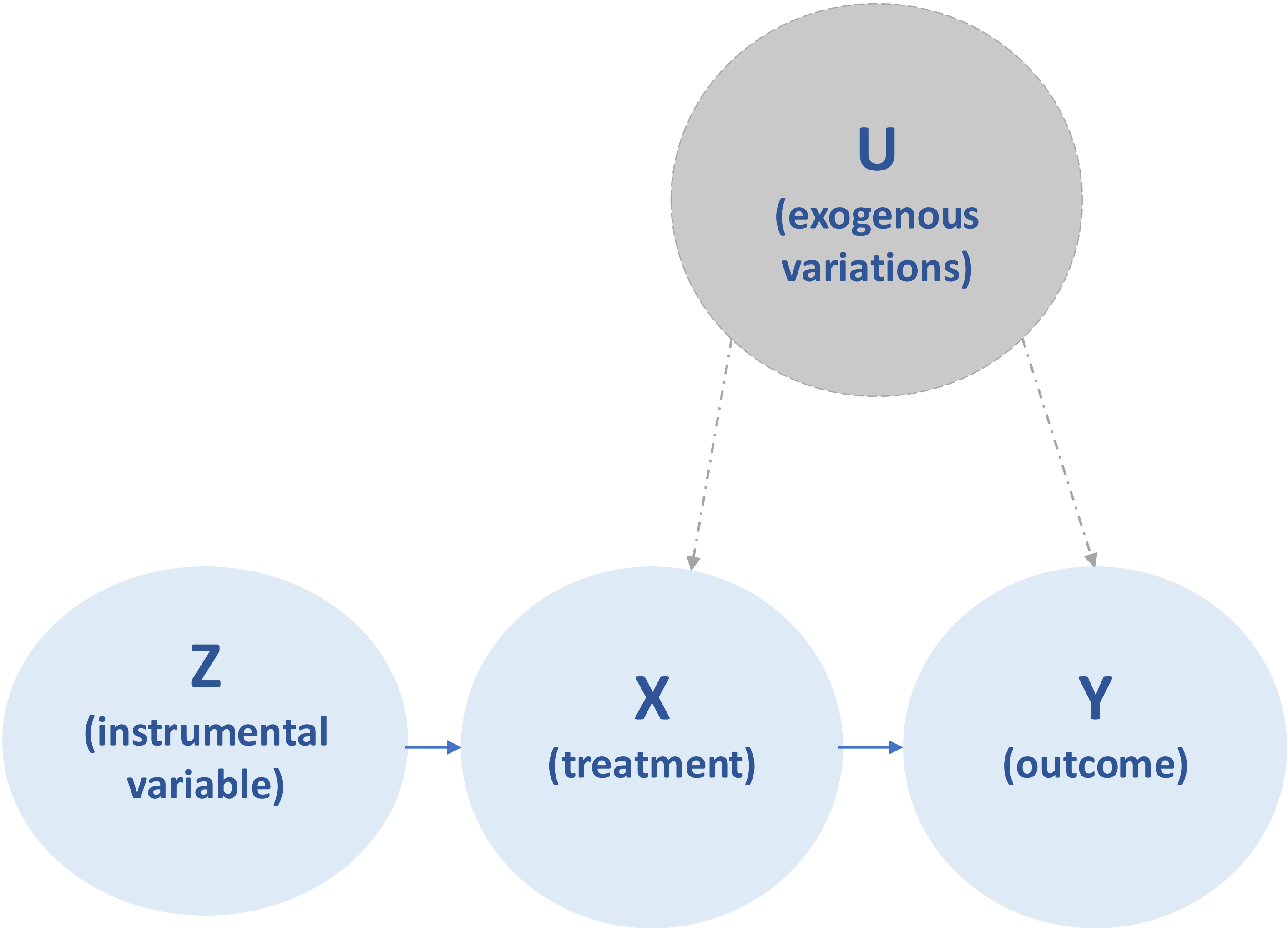
RECOMMANDATION EXAMPLE

Regress click-throughs on product attributes and direct visits to recommended product.

UNTESTABLE CONDITIONS

$X \perp\!\!\!\perp U$ or $Y \perp\!\!\!\perp U$

Instrumental Variable



DESCRIPTION

Analyze subset of data that has independent variation in the treatment.

LIMITATIONS

Difficult to find a source of exogenous variation in the treatment.

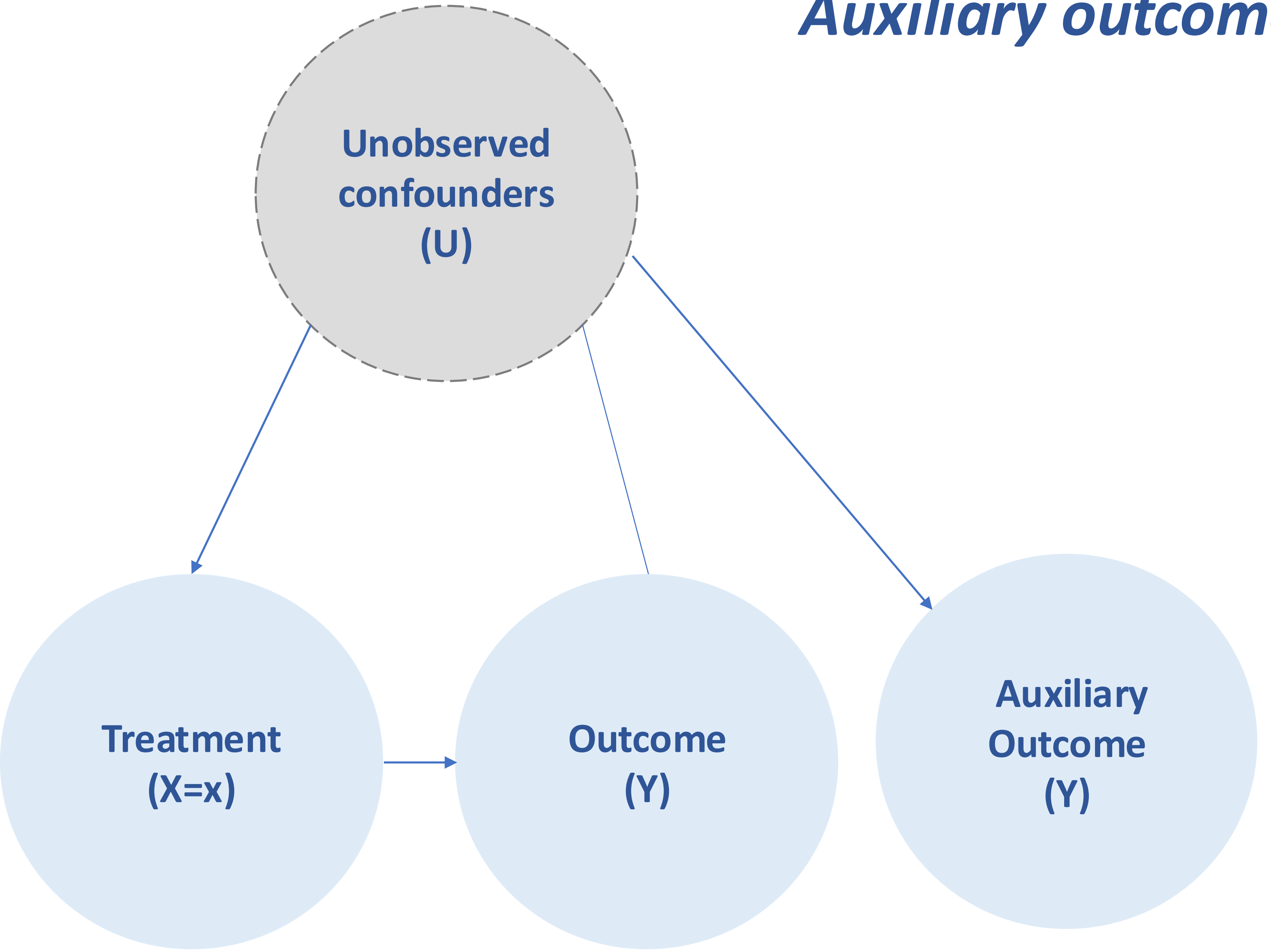
RECOMMANDATION EXAMPLE

Measure marginal click-throughs on products that experience large, sudden shocks in traffic.

UNTESTABLE CONDITIONS.

$Z \perp\!\!\!\perp U$ and $Z \perp\!\!\!\perp Y | X,U$

First step for Split-Door (SD) criterion: Auxiliary outcome

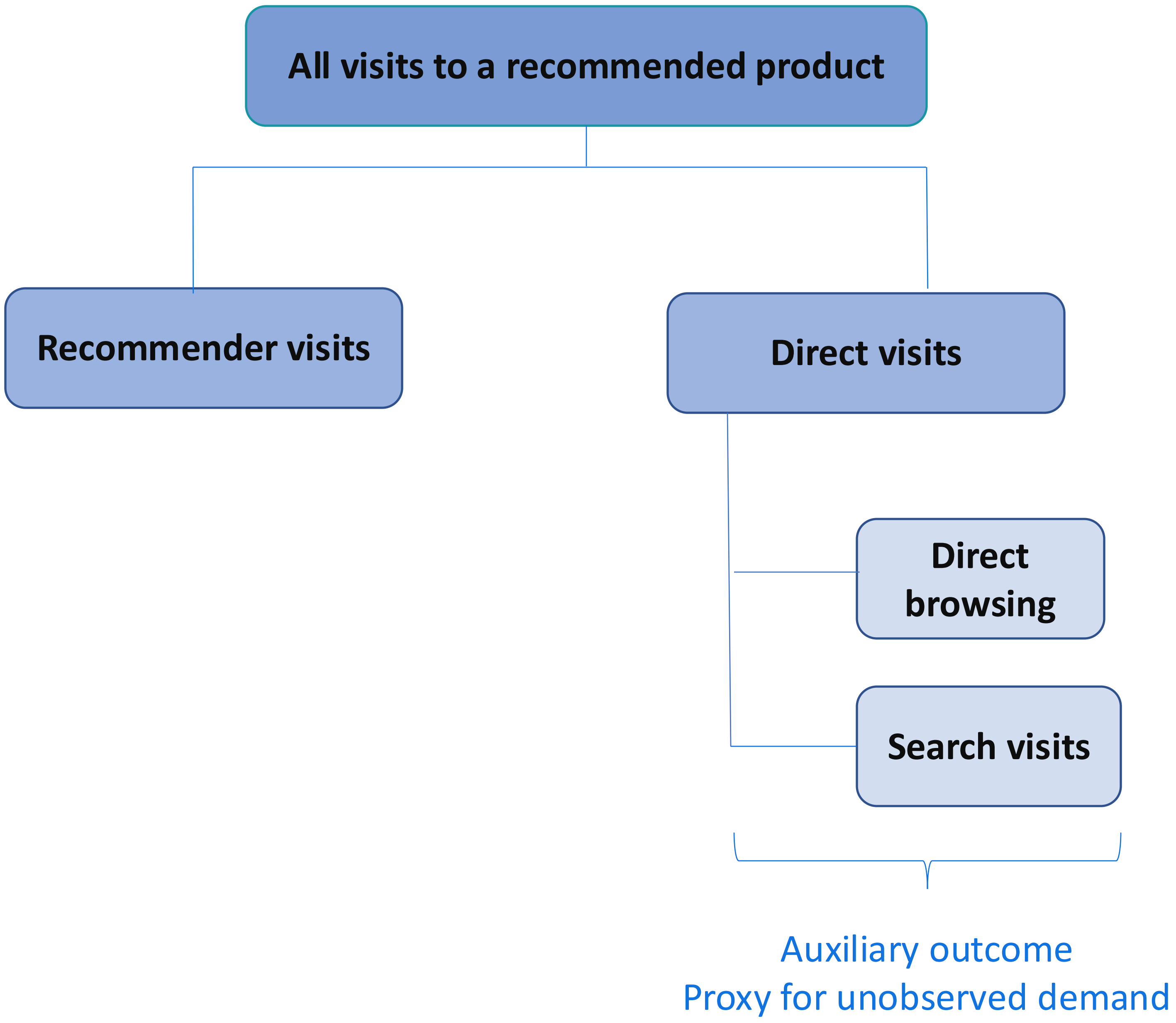


Outcome can be separated into two observable parts:

I) Primary Outcome (possibly) affected by the cause

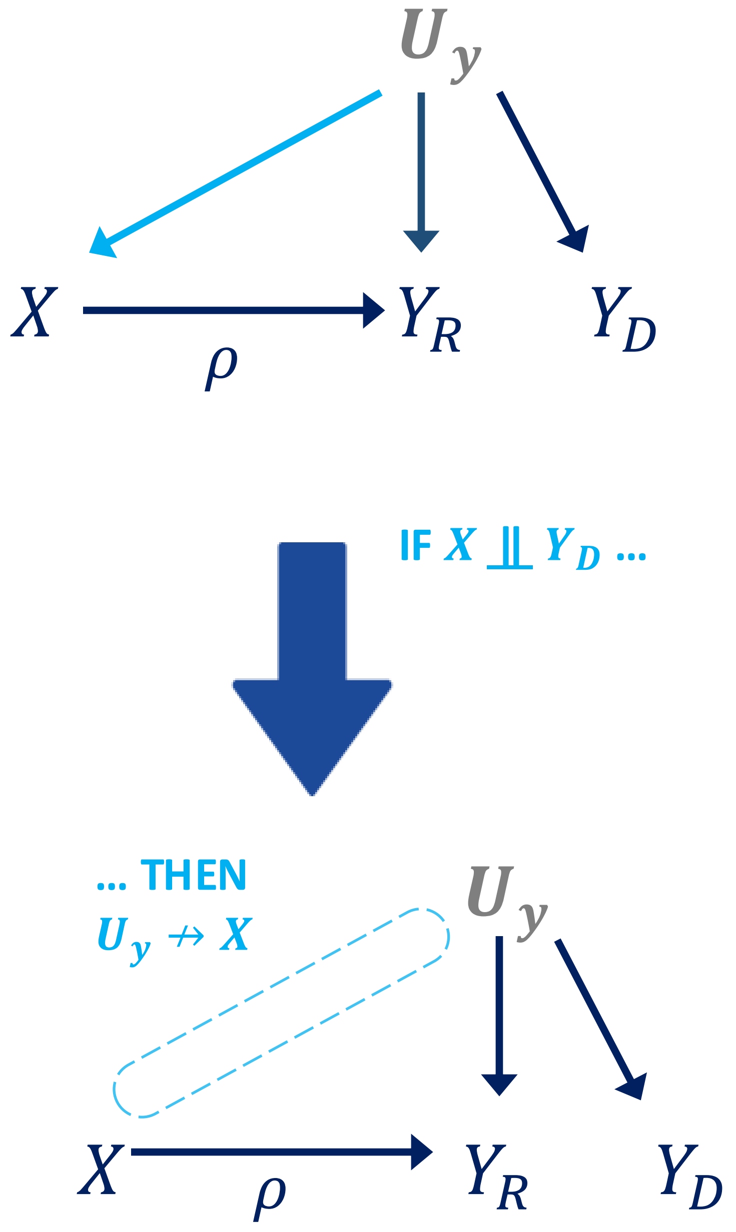
II) Auxiliary Outcome unaffected by the cause

First step for Split-Door (SD) criterion: Auxiliary outcome



Outcome can be separate into two parts

Split-Door (SD) criterion



Problem:

Causal effect still not identified

Solution:

If $X \perp\!\!\!\perp Y_D$ and SD's assumptions hold (next slides)

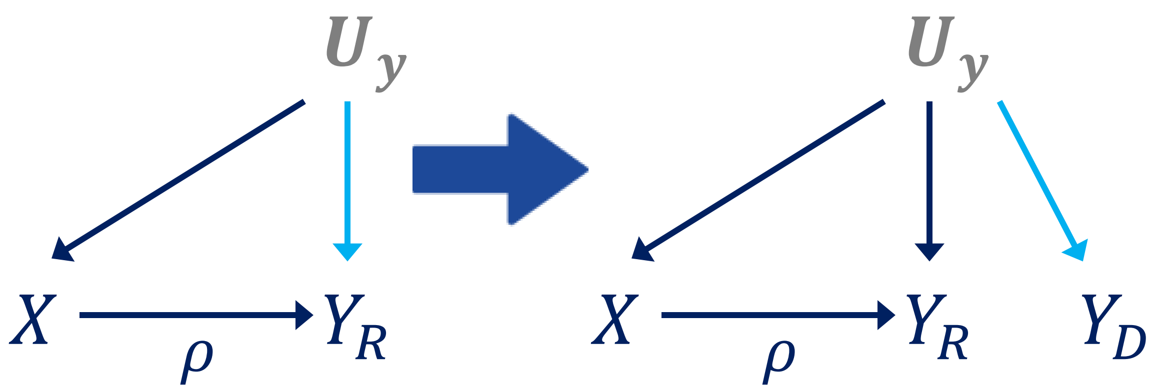
$\Longrightarrow$ **causal effect identified**

i) Connectedness

$$\forall\, U_y:\ \ U_y \to Y_R \ \ \Longrightarrow\ \ U_y \to Y_D$$

i.e.: Any cause of Y_R

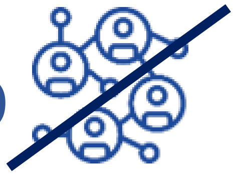
is also cause of Y_D

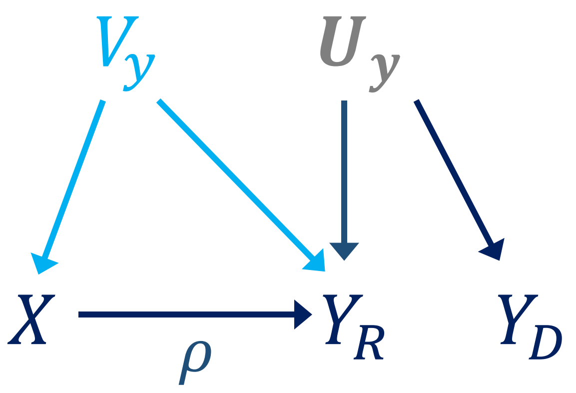


Connectedness reasonable in RS

- ✔ User's demand for suggested product (U_y) manifests itself through *both* direct searches *and* recommendations

- ✔ Unlikely that users have demand for the recommended product only if arrived via recommendation link (multiple channels)

i) Connectedness VIOLATED 



There are causes of X and Y_R not in common with Y_D

Even if $X \perp\!\!\!\perp Y_D$, causal effect still *not identified*

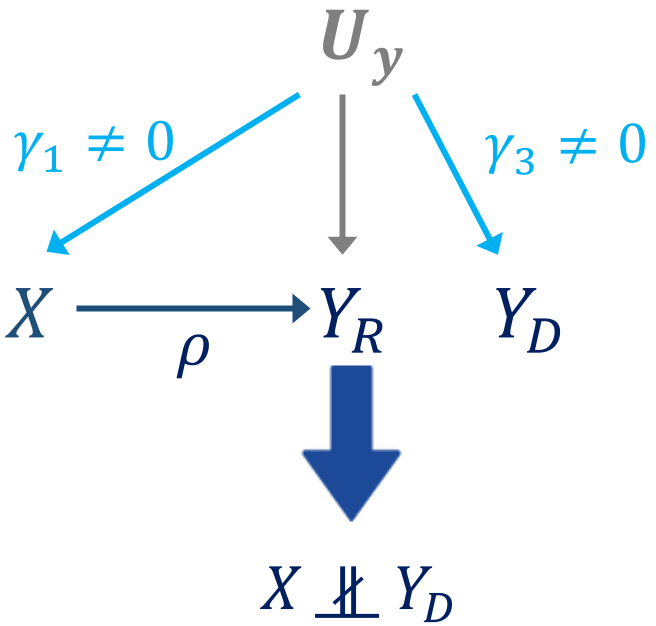
Connectedness reasonable in RS

- ✔ User's demand for suggested product (U_y) manifests itself through *both* direct searches *and* recommendations
- ✔ Unlikely that users have demand for the recommended product only if arrived via recommendation link (multiple channels)

ii) Independence

X and Y_D show dependence

when they share common cause U_y



Independence reasonable in RS

✔ $\gamma_3 = -\gamma_1$ rare event

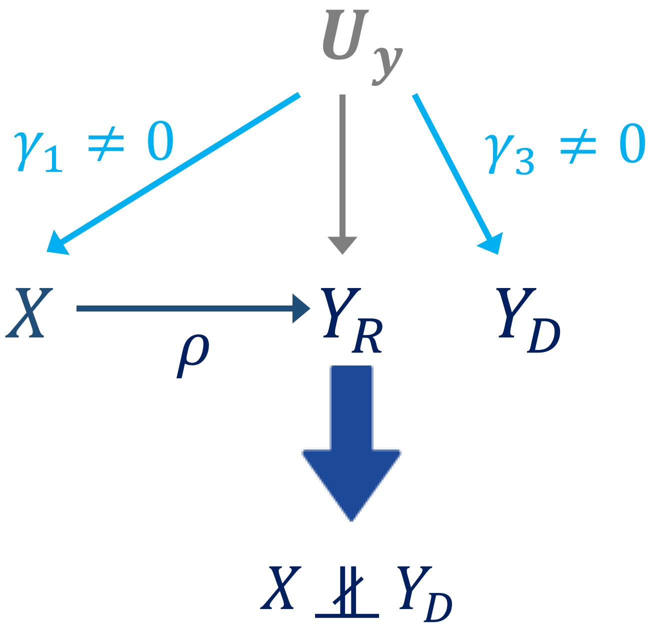
✔ Moreover, in RS: *complementary products* (e.g.: i book & j bookmark) $\Longrightarrow$
$\Longrightarrow$ unobserved demand U_y affects X and Y_D in *same* direction:

$$\begin{cases} \gamma_1 > 0 \\ \gamma_3 > 0 \end{cases} \quad \text{OR} \quad \begin{cases} \gamma_1 < 0 \\ \gamma_3 < 0 \end{cases}$$

ii) Independence VIOLATED

γ_1 and γ_3 **exactly counterbalance** one another …

… **Possible, but unlikely!**



Independence reasonable in RS

- ✔ $\gamma_3 = -\gamma_1$ rare event

- ✔ Moreover, in RS: *complementary products* (e.g.: i book & j bookmark) $\Longrightarrow$
$\Longrightarrow$ unobserved demand U_y affects X and Y_D in *same* direction:

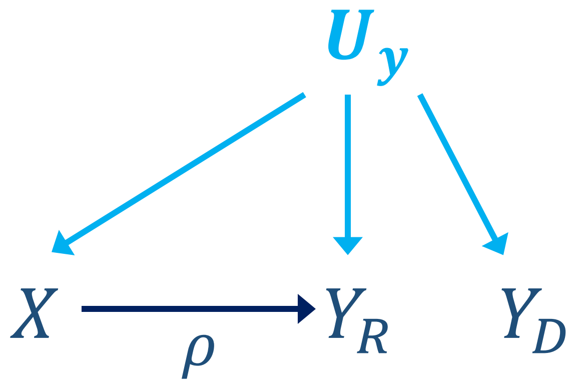
$$\begin{cases} \gamma_1 > 0 \\ \gamma_3 > 0 \end{cases} \quad \text{OR} \quad \begin{cases} \gamma_1 < 0 \\ \gamma_3 < 0 \end{cases}$$

Demonstration of SD mechanism

1) For *Independence (ii):*

IF $X \perp\!\!\!\perp Y_D \implies$ Something cancels out on $X \leftarrow U_y \rightarrow Y_D$

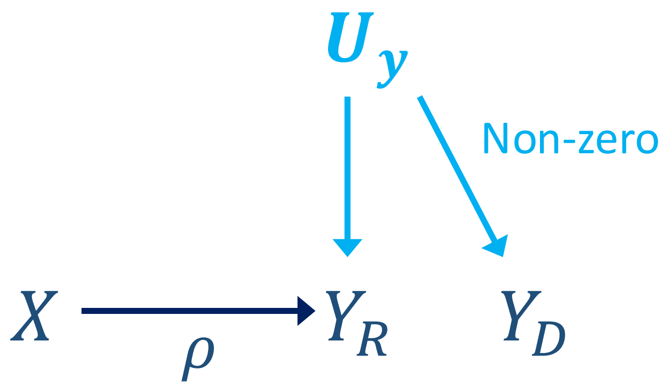
OR U_y constant (trivial)



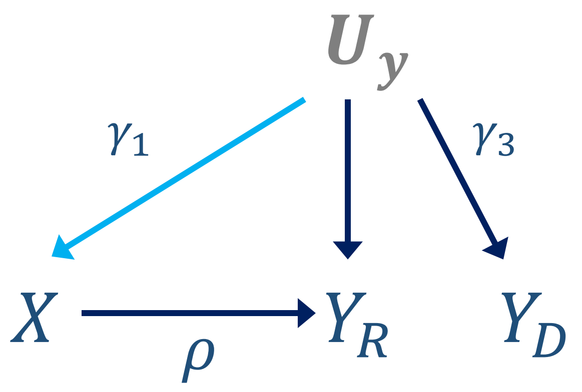
2) For *Connectedness (i):*

Non-zero effect of U_y on Y_D

3) Only alternative: the $X \leftarrow U_y$ edge does not exist, leading to unconfounded causal model



Demonstration of SD mechanism, linear structural equations



- For who is not convinced:

$$x = \eta u_x + \gamma_1 u_y + \epsilon_x$$

$$y_r = \rho x + \gamma_2 u_y + \epsilon_{yr}$$

$$y_d = \gamma_3 u_y + \epsilon_{yd}$$

$$0 = Cov(X, Y_D)$$

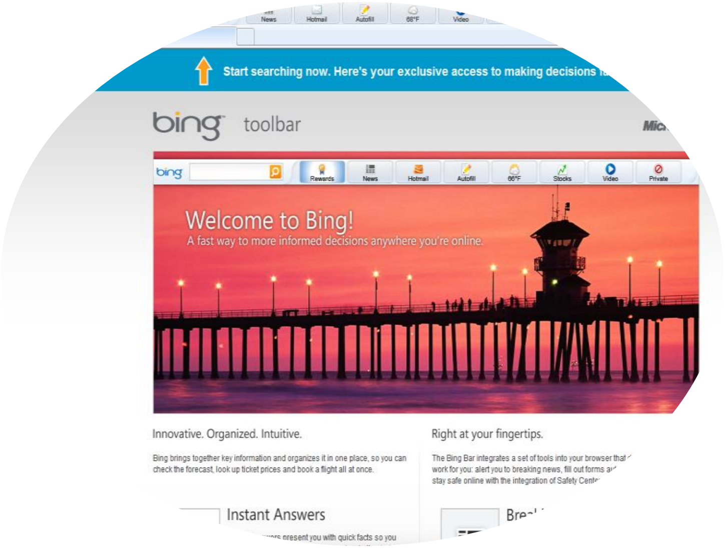
$$= E[XY_D] - E[X]E[Y_D]$$

$$= E\big[(\eta u_x + \gamma_1 u_y + \epsilon_x)(\gamma_3 u_y + \epsilon_{yd})\big] - E[\eta u_x + \gamma_1 u_y + \epsilon_x]E[\gamma_3 u_y + \epsilon_{yd}]$$

$$= \gamma_1\gamma_3 E[U_Y U_Y] - \gamma_1\gamma_3 E[U_Y]E[U_Y]$$

$$= \gamma_1 \gamma_3\, Var(U_Y)$$

Data



- Browsing logs from Bing Toolbar, aggregated by day
- Selected 22 000 focal products with at least 10 visits per day
- Cover a period of 9 months (Sept.2013-May 2014)
- Focal products and recommended products pairs (tseries_id)
- Y_r, Y_d counted for each recommended product each day (ref in Amazon URLs)

date	treatment_tseries_id	treatment_group	outcome_tseries_id	treatment_val	outcome_val	aux_outcome_val
2013-09-01	15493	Book	4521	4.166667	0.000000	0.000000
2013-09-01	15493	Book	10911	4.166667	0.000000	0.000000
2013-09-01	15493	Book	15586	4.166667	0.000000	0.000000
2013-09-01	15493	Book	48250	4.166667	0.000000	0.000000
2013-09-01	15493	Book	48922	4.166667	0.000000	0.000000
2013-09-01	15493	Book	11810	4.166667	0.000000	0.000000
2013-09-01	15493	Book	53333	4.166667	0.000000	0.000000
2013-09-01	21714	Book	21671	31.250000	6.250000	6.250000
2013-09-01	21714	Book	36021	31.250000	0.000000	0.000000
2013-09-01	21714	Book	15586	31.250000	0.000000	0.000000
2013-09-01	21714	Book	24935	31.250000	6.250000	0.000000
2013-09-01	21714	Book	50690	31.250000	0.000000	6.250000
2013-09-01	21714	Book	13447	31.250000	0.000000	0.000000

Algorithm

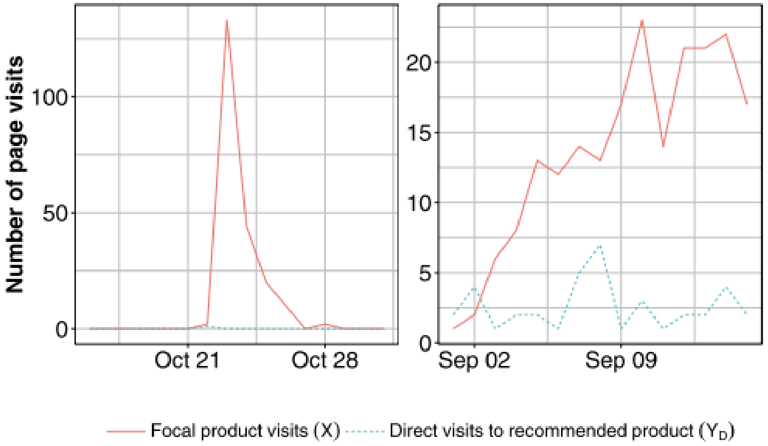
01100
10110
11110

For each *focal product*, split entire dataset in 15-days periods

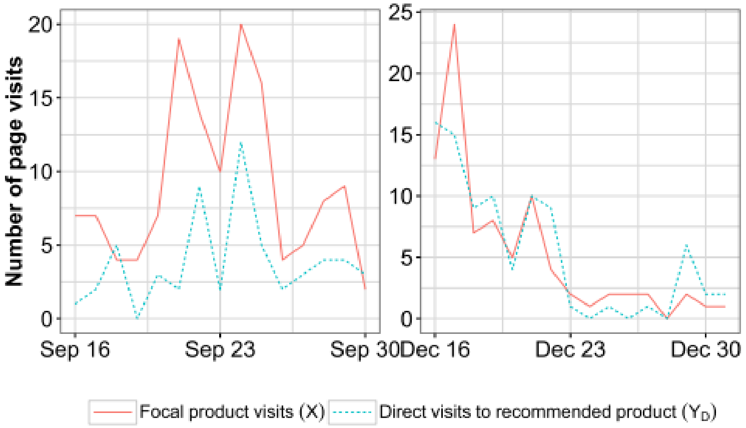
1. Find all valid time periods for which:

H_0: X ⫫ Yd is **«accepted»**

2. On each valid time period S, compute $\widehat{p_s} = \mathrm{CTR(S)}$ and average them



(a) Accepted at $\alpha = 0.95$

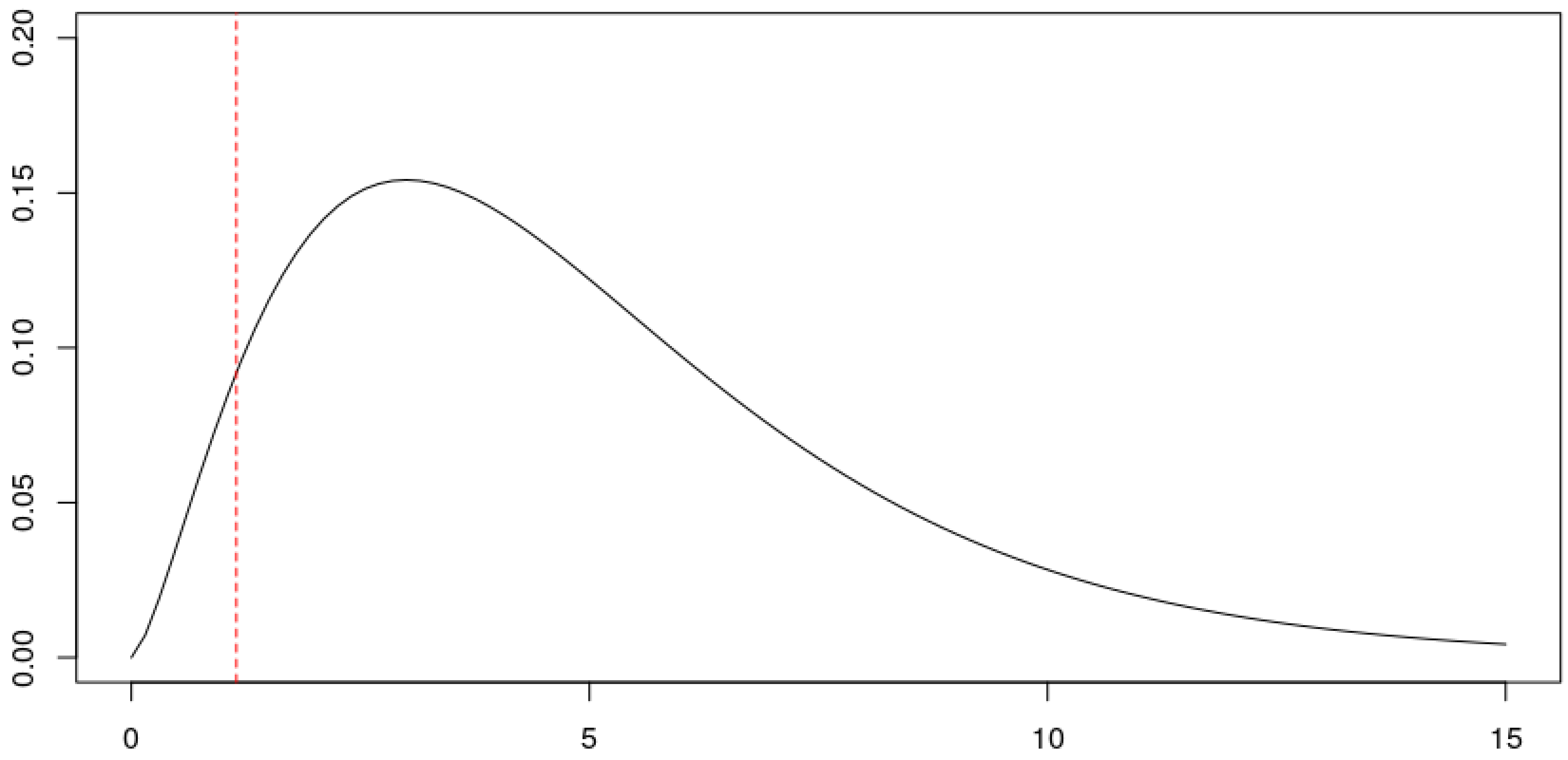


(b) Rejected at $\alpha = 0.95$

Details about independence test

We want to "accept" independence, rather than "not-refusing" it, so cutoff must be *low* (avoid False Negatives, i.e. "dependent that are declared independent")

Cutoff 0.05



Details about independence test

Fisher's exact tests are preferable, due to low statistical power (15-days periods)

P-value is obtained simulating distribution under null hypothesis

Permute X several times

X	Y_D
x_3	y_{D1}
x_5	y_{D2}
x_1	y_{D3}
x_6	y_{D4}
x_2	y_{D5}
x_4	y_{D6}

Vector of test statistics who measure dependence under Ho

$$\begin{pmatrix} S_1 \\ S_2 \\ \dots \end{pmatrix}$$

P-value

Details about independence test

High number of hypothesis test holds to acceptance of independence just by chance

We control the type II errors using the False Non Discovery Rate.

$$FNDR = E[\frac{False\ Negative(T)}{False\ Negative(T) + True\ Negative(U)}]$$

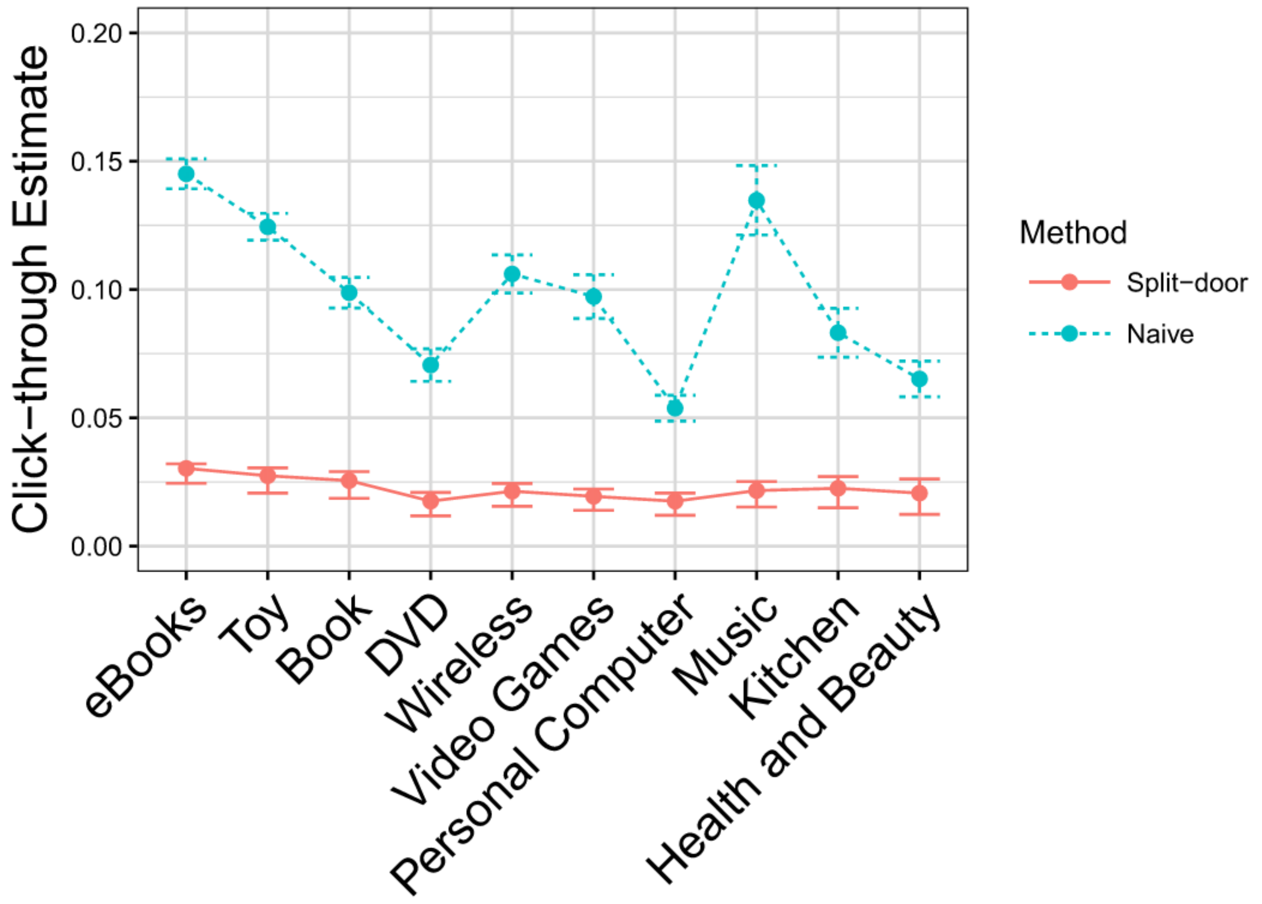
	Null hypothesis is true (H_0)	Alternative hypothesis is true (H_A)	Total
Test is declared significant	V	S	R
Test is declared non-significant	U	T	$m - R$
Total	m_0	$m - m_0$	m

RESULTS

Observed CTR versus Causal CTR, by product

Causal CTR overall:
2.6%

Observed CTR
overstates causal
impact by 2x or 3x times



ARE THE RESULTS ROBUST?

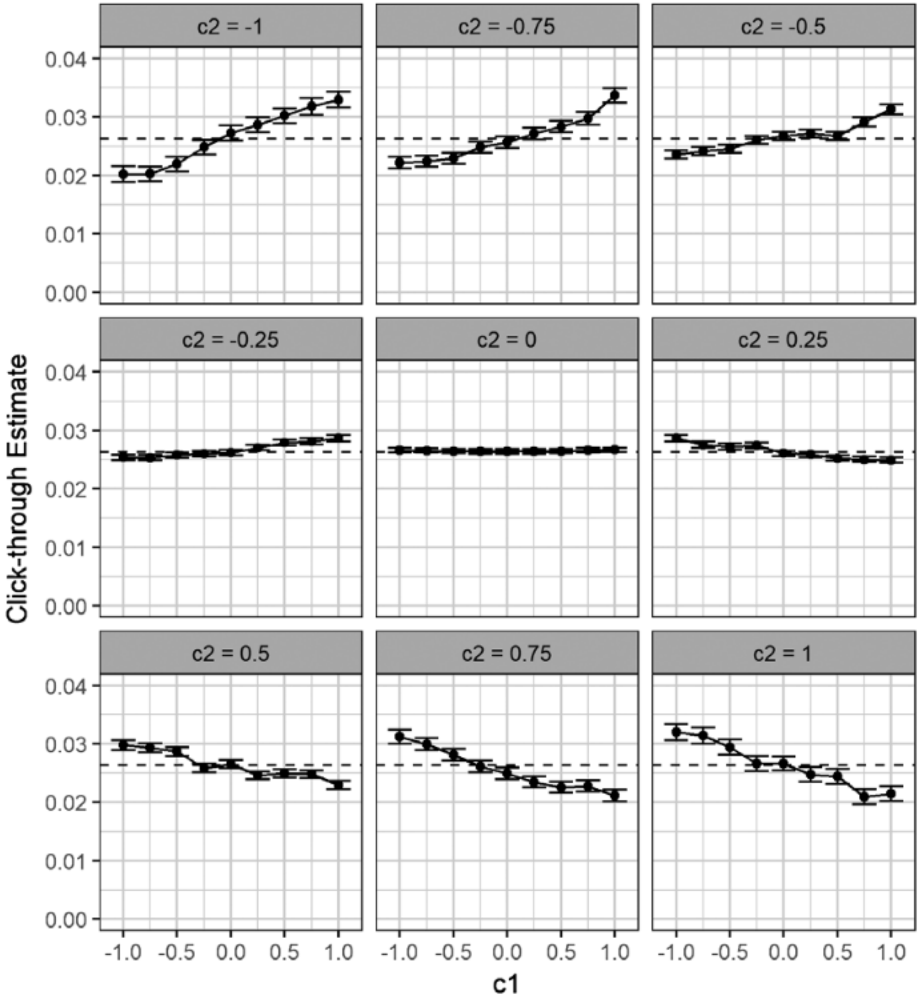
Sensitivity to connectedness violation (internal validity)

Simulation under the worst case scenario:
ALL SPLIT-DOOR INSTANCES ARE INVALID

Simulate V_y from a standard normal and imagine that it affects X and Y_R on all valid time periods:

$$x = x + c_1 v_y$$
$$y_R = y_R + c_2 v_y$$

CTR affected ± 0.05

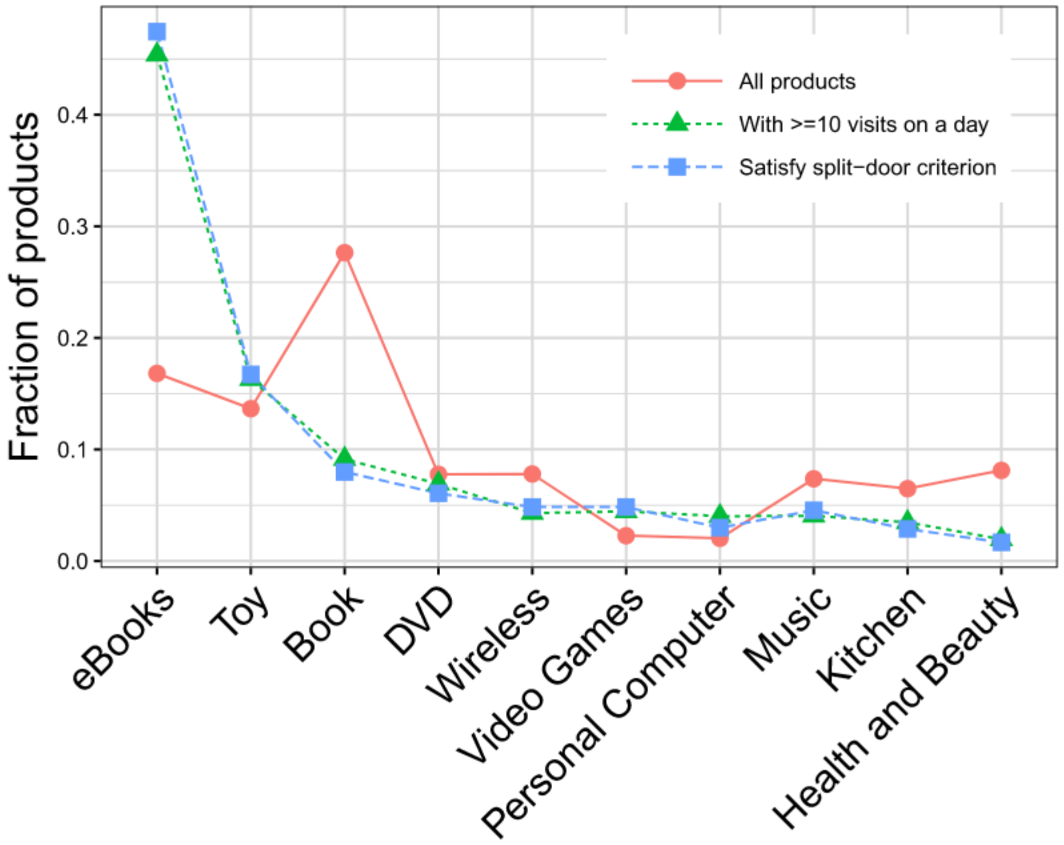


ARE THE RESULTS ROBUST?

Generalizability (external validity)

SD-sample is *not* selected at random!

Distribution of product categories for
SD-sample is similar to that of the left-out products,
So we can at least generalize on it.



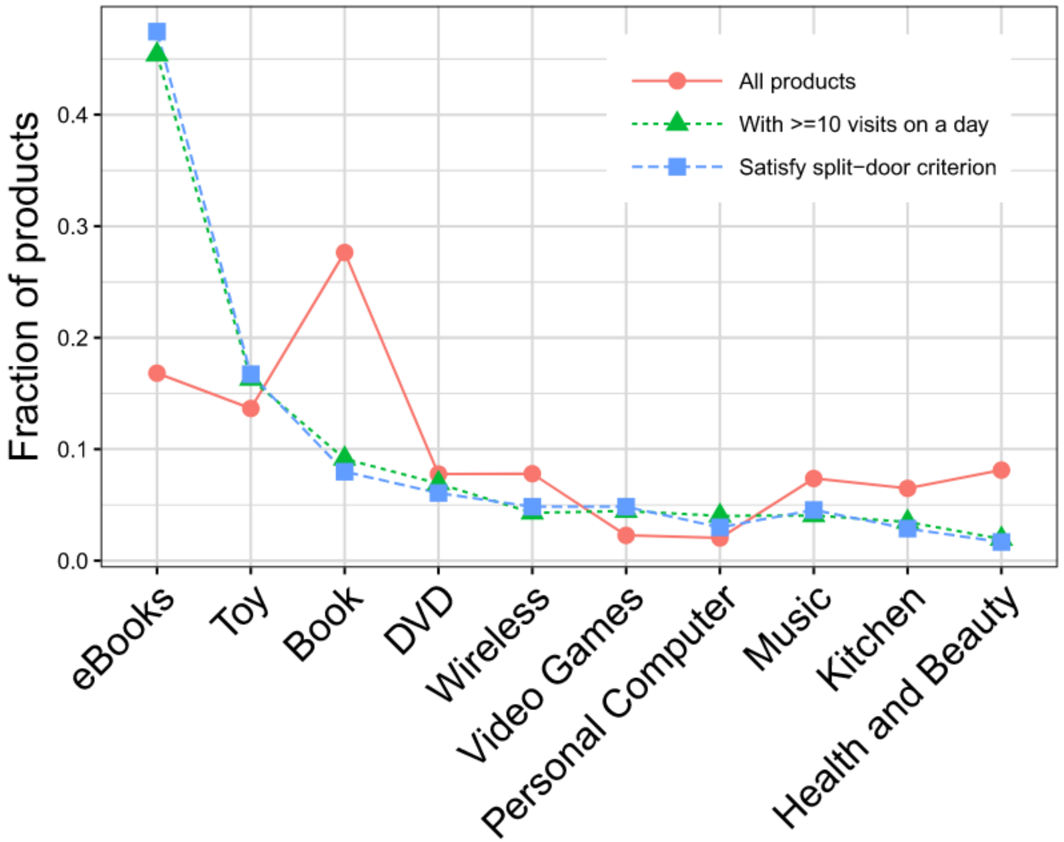
ARE THE RESULTS ROBUST?

Generalizability (external validity)

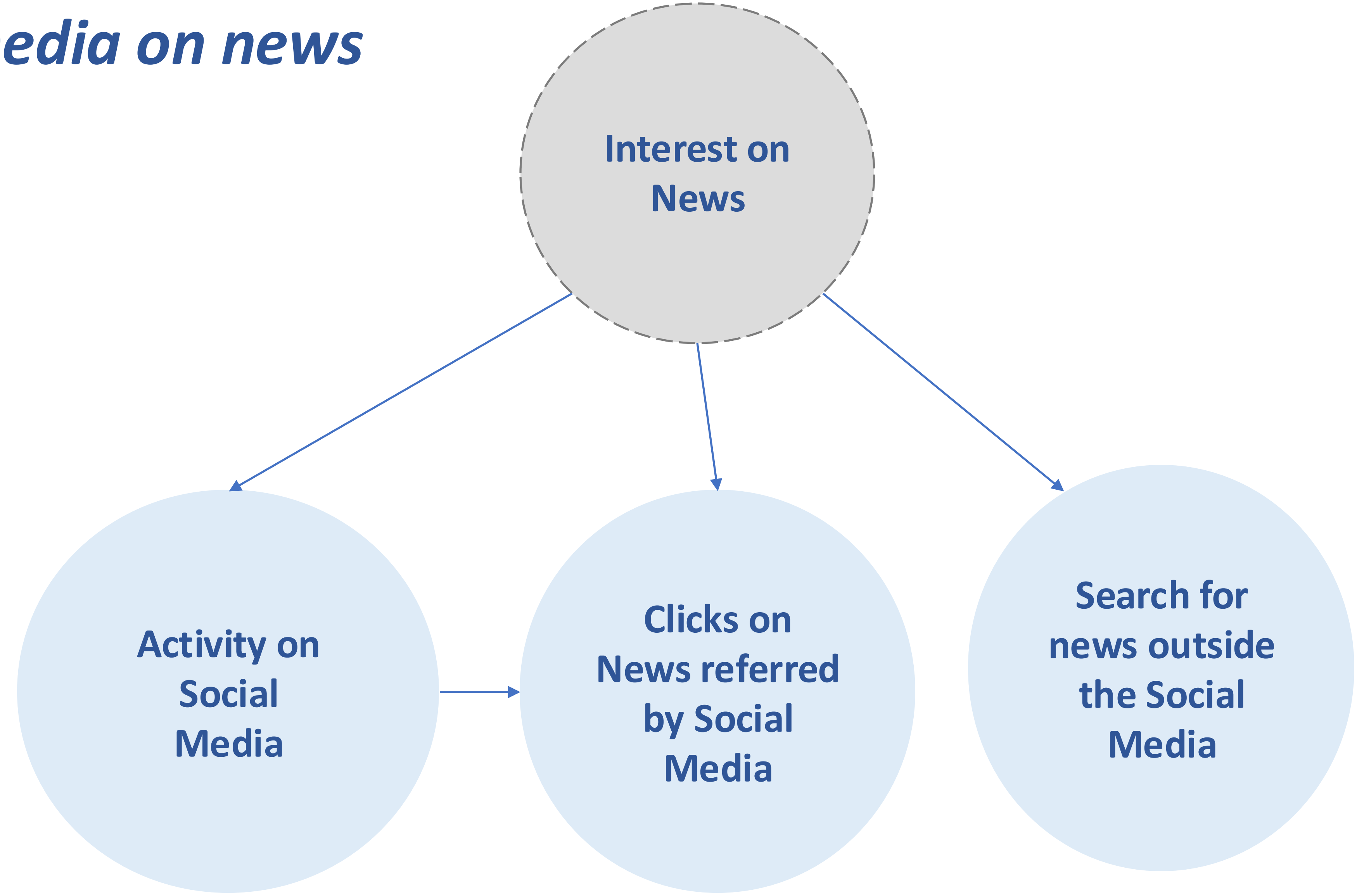
SD-sample is *not* selected at random!

Distribution of product categories for
SD-sample is similar to that of the left-out products,
So we can at least generalize on it.

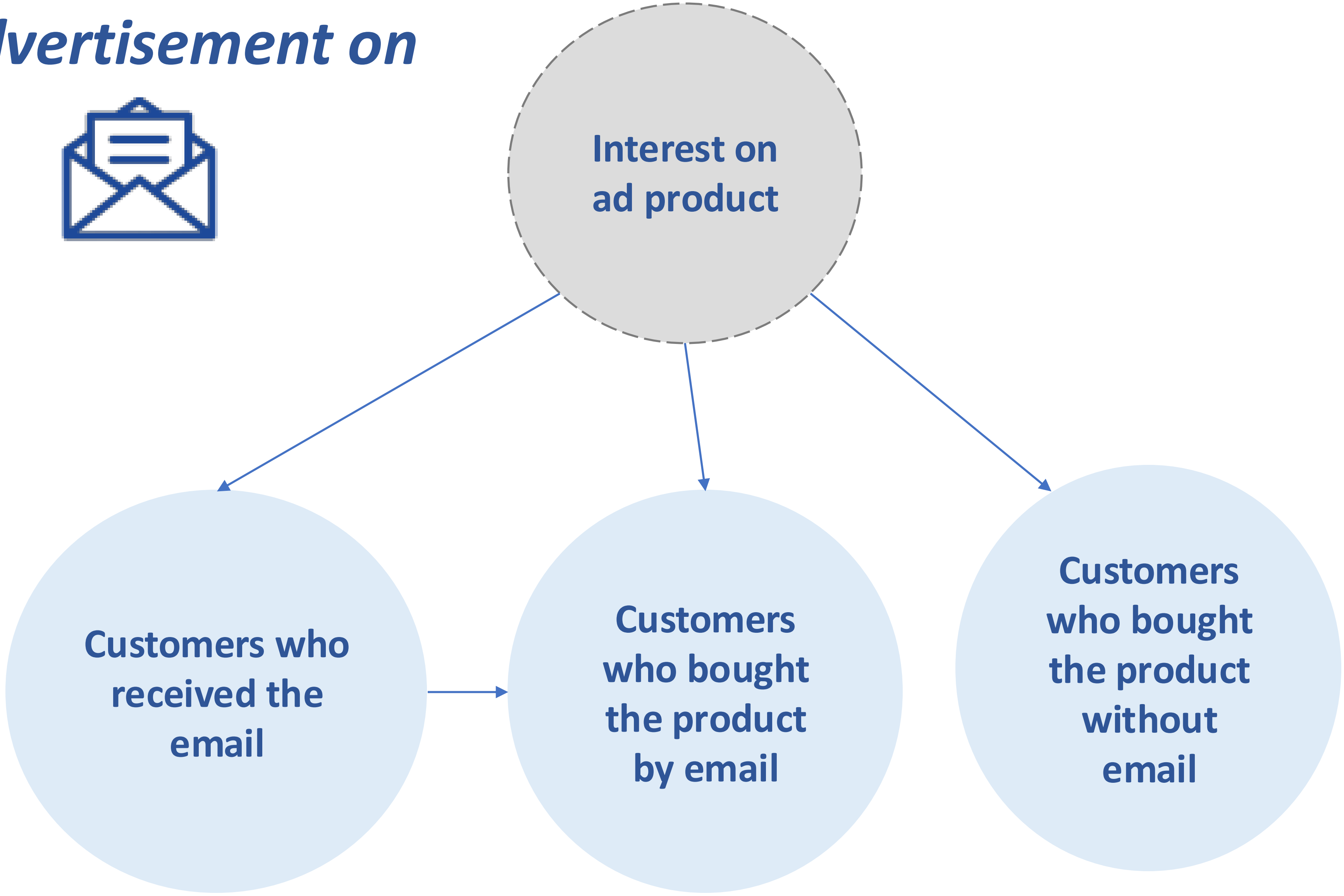
Same number of page visits.



1. Effect of social media on news consumption



2. Effect of email advertisement on product purchasing



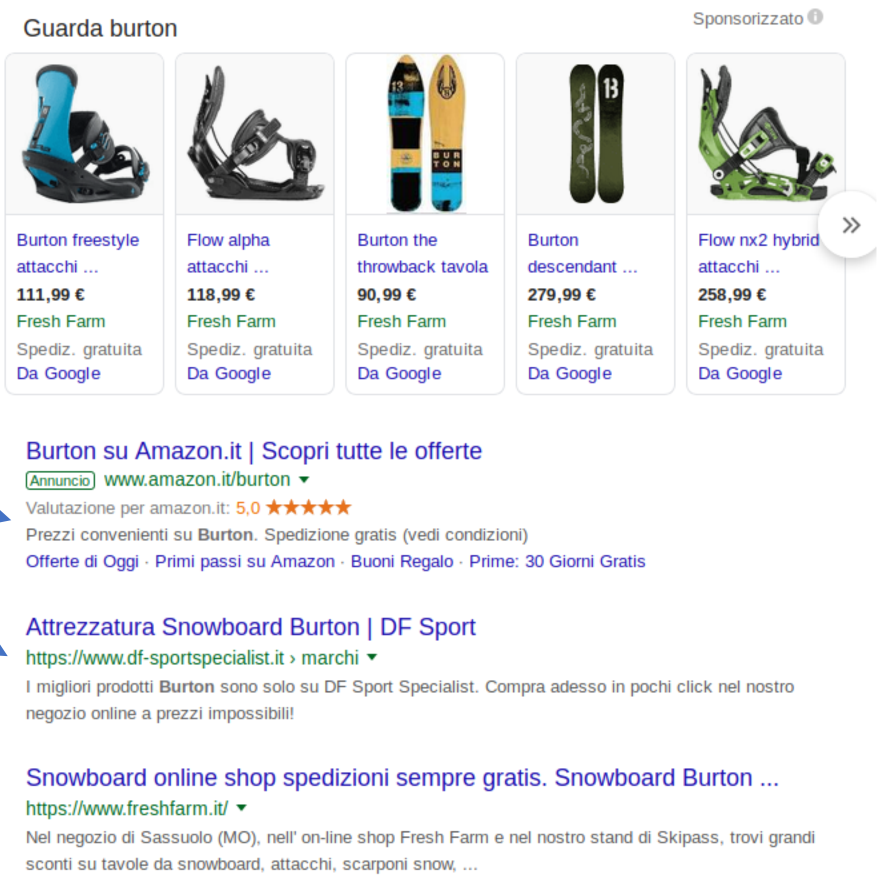
Advertisting on Google search engine

In Brodersen et al. (2015) estimated causal effect of an ads. campaign for firm F on clicks that F received using Bayesian Structural Time Series in which:

Outcome can be *splitted into*

- paid clicks: clicks on an ad next to results
- organic clicks: clicks on a search result

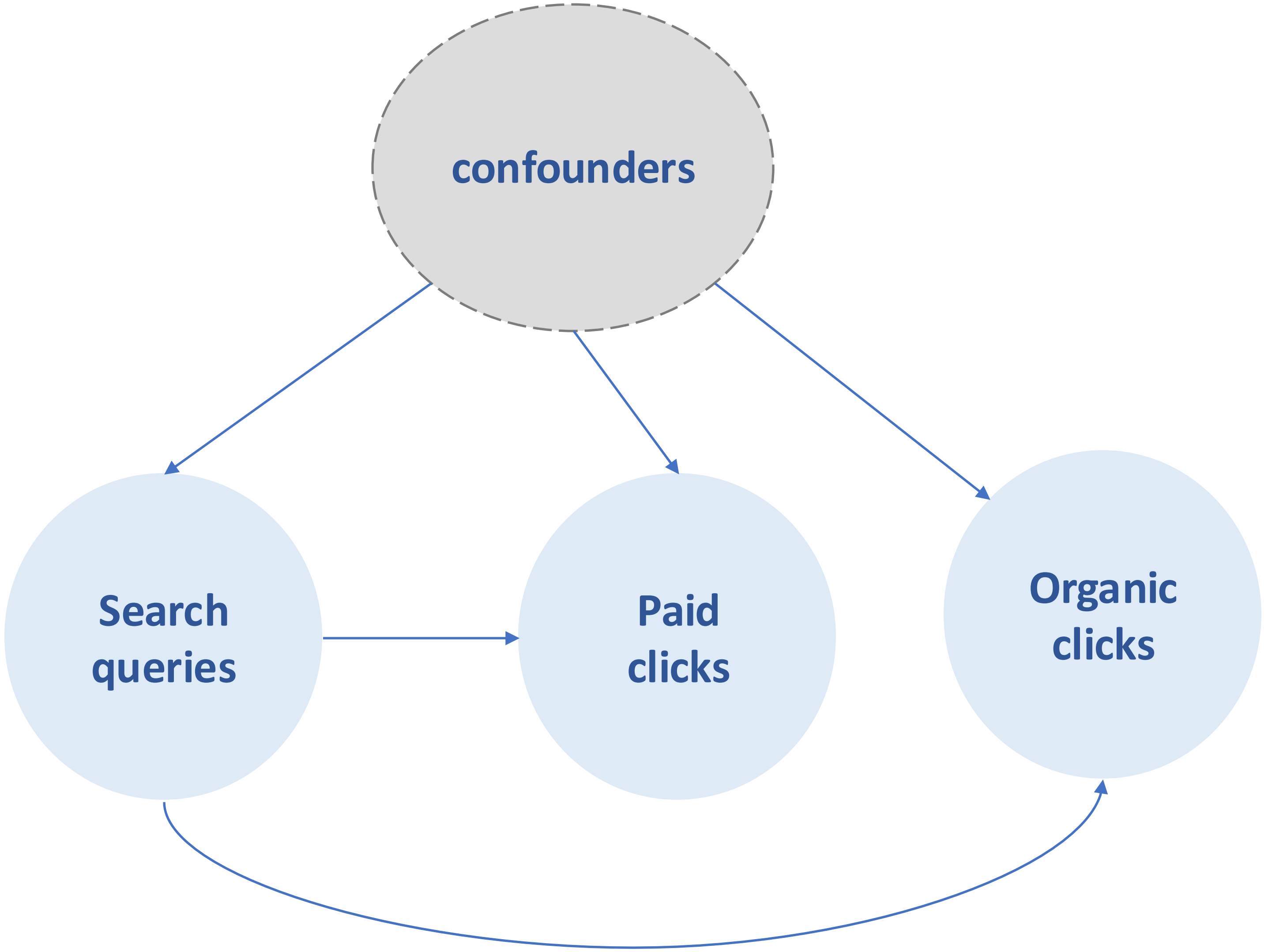
Paid clicks alone overestimate effect of campaign, because paid click cannibalise organic click



Why we can't apply Split-Door in this case?

Organic clicks is not a valid auxiliary outcome.

Search queries cause both paid and organic clicks.



Split-Door:

- Prior information not required
- Auxiliary Outcome not always available
- Needs time periods for indipendence test
- It needs to exclude time periods to ensure internal validity

Bayesian Structural Time Series:

- Needs «Prior information»: Behaviour before the intervention, behaviour of other time series predictive of the target preintervention, bayesian prior (pre-knowledge).
- Auxiliary outcome is not necessary
- Is defined for each point in the time serie
- No time periods excluded.

References

Brodersen, K. H., Gallusser, F., Koehler, J., Remy, N., & Scott, S. L. (2015). Inferring causal impact using Bayesian structural time-series models. *The Annals of Applied Statistics, 9*(1), 247-274.

Sharma, A., Hofman, J. M., & Watts, D. J. (2016). Split-door criterion for causal identification: Automatic search for natural experiments. *arXiv preprint arXiv:1611.09414*

Sharma, A. (2017). *Causal Data Mining indentifying causal effects at scale.* Retrieved from https://www.slideshare.net/AmitSharma315/causal-data-mining-identifying-causal-effects-at-scale

Sharma, A. (2018). Split-door causal criterion: Automatic search for natural experiments. Retrieved from https://github.com/amit-sharma/splitdoor-causal-criterion

Thanks for the attention!