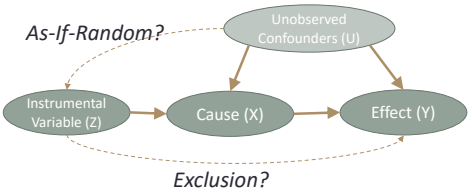
Split-door criterion for causal identification: Natural experiments with testable assumptions

Amit Sharma, Jake M. Hofman, Duncan J. Watts
Microsoft Research, New York
amshar@microsoft.com, jmh@microsoft.com , duncan@microsoft.com

NATURAL EXPERIMENTS RELY ON UNTESTABLE ASSUMPTIONS

Canonical causal inference problem
Estimate the effect of cause *X* on outcome *Y* using observational data.

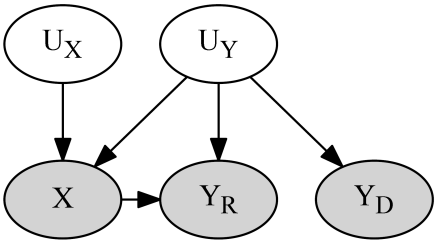
Natural experiments are a popular way of estimating the effect of X on Y. For example, using *Z* as an auxiliary instrumental variable.



However, assumptions for valid identification are untestable:

- **Exclusion:** *Z* does not directly cause *Y*, except through *X*.
- **As-if-random:** *Z* is not caused by the unobserved confounders *U*.

SPLIT-DOOR: USING ADDITIONAL DATA TO REMOVE CONFOUNDING



Split-door criterion: Y split up as Y_R+Y_D
If X, Y_R and Y_D are random variables generated by the process shown in the above graphical model, then $X \perp\!\!\!\perp Y_D$ implies that the effect of X on Y is not confounded by U_Y.

Assumptions

- **Connectedness:** Any unobserved confounder U_Y that causes both X and Y_R also causes Y_D and the effect of such U_Y on Y_D cannot be zero.
- **Statistical Independence:** If X and Y_D are statistically independent, then they are also causally independent in the graphical model.

Causal estimate: $P(Y|\ do(X = x)) = P(Y|X = x)$.

ALGORITHM FOR FINDING NATURAL EXPERIMENTS

Assume X, Y_R and Y_D are observational time-series.
Divide the data into equally-spaced time-intervals *t*.

Split-door algorithm
For each time interval t,

- Check $X \perp\!\!\!\perp Y_D$ using an empirical independence test, such as mutual information.
- If $X \perp\!\!\!\perp Y_D$, then use the observed conditional probability $P(Y|X = x)$ to estimate the causal effect in the interval *t*.
- Average over all time-intervals where $X \perp\!\!\!\perp Y_D$ to obtain the mean causal effect of X on Y.

HOW DOES SPLIT-DOOR COMPARE WITH OTHER IDENTIFICATION STRATEGIES?

Weaker assumptions than instrumental variables.

Instrumental Variable	Split-door criterion
Assumption: Z ⫫ Y | X, U and Z ⫫ U.	*Assumption:* $Y_D \not\perp\!\!\!\perp U$ for each confounder U.
Requires independence between observed and unobserved variables.	Requires dependence between observed and unobserved variables.
Best suited for arguably randomized Z.	Best suited for time-series data.

Generalizes notion of a large, sudden shock in time-series data.

Shock-IV criterion	Split-door criterion
Assumption: $Y_D \not\perp\!\!\!\perp U$	*Assumption:* $Y_D \not\perp\!\!\!\perp U$
Criterion: X has a large, sudden spike and Y_D is constant w.r.t. time.	*Criterion:* $X \perp\!\!\!\perp Y_D$. More general, admits diverse variations in data.

Provides control for data selection, rather than conditioning on observed variables.

Back-door criterion	Split-door criterion
Assumption: Y_D is a perfect proxy for U. Constant Y_D => Constant U	*Assumption:* $Y_D \not\perp\!\!\!\perp U$ for each confounder U.
Unlikely to be true, because U may have unknown confounders.	Plausible assumption when Y_D and Y_R are components of same Y.

APPLICATION: CAUSAL IMPACT OF A RECOMMENDER SYSTEM

Observational log data
Anonymized toolbar logs for *Amazon.com*

- Dates: September 2013 to May 2014.
- 23.4 million visits by 2.1 million users.
- 1.38 million unique products.

Restrict analysis to products with at least 10 page views on any single day during the nine month period.

- Filtered set: over 22,000 products.

Causal question
How much traffic does the recommender system cause that would not have happened in its absence?

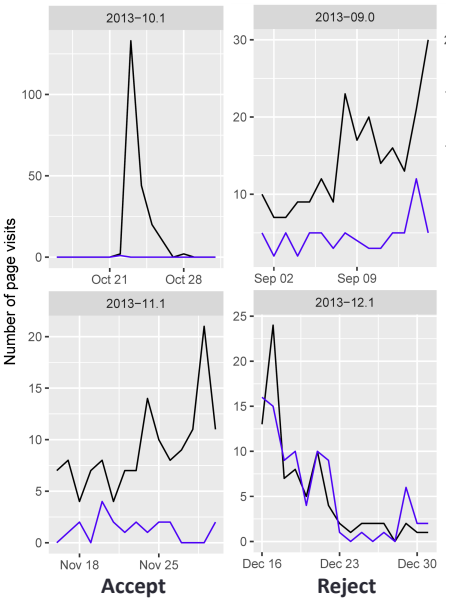
X: Visits to focal product's webpage.

Y_R: Number of click-throughs from focal to recommended product.

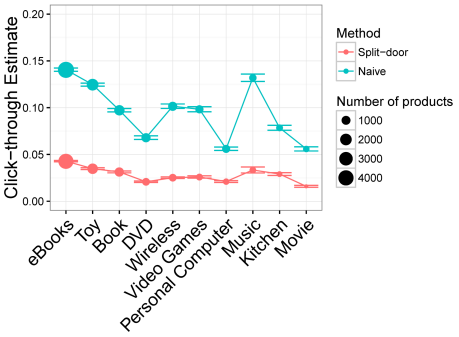
Y_D: Direct search visits to recommended product.

FINDING NATURAL EXPERIMENTS

Use t=15 days and Fisher's exact test for checking independence of X and Y_D.



NAÏVE CTR OVERESTIMATES EFFECT OF RECOMMENDATIONS



DISCUSSION: DATA-DRIVEN IDENTIFICATION STRATEGIES

- Split-door criterion allows us to find natural experiments for 12K products, over half of all products.
- Distribution of products that satisfy Split-door criterion similar to overall distribution of products w.r.t. page visits, and product category.

Data-driven strategies hold promise as more fine-grained data becomes available.

REFERENCES

- **Instrumental variable:** Angrist, Joshua D., and Jörn-Steffen Pischke. Mostly harmless econometrics: An empiricist's companion. Princeton University Press, 2008.
- **Shock-IV criterion:** Sharma, Amit, Jake M. Hofman, and Duncan J. Watts. Estimating the causal impact of recommendation systems from observational data. Proc. Sixteenth ACM Conference on Economics and Computation. ACM, 2015.
- **Back-door criterion:** Pearl, Judea. Causality. Cambridge University Press, 2009.