Causal Mediation Analysis for Online Products

Xuan Yin

xyin100@gmail.com

Version: August 2, 2022

Xuan Yin and Liangjie Hong. 2019. The Identification and Estimation of Direct and Indirect Effects in A/B Tests through Causal Mediation Analysis. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (KDD '19). ACM, New York, NY, USA, 2989-2999. DOI: https://doi.org/10.1145/3292500.3330769

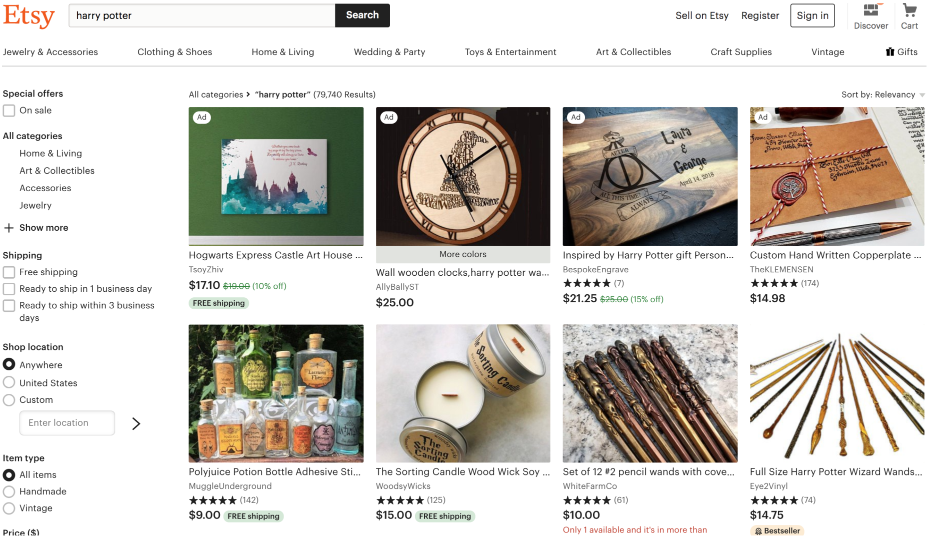
About Xuan Yin

- **Education**:
  - PhD in Economics
  - MS in Mathematics
- **Research Areas**:
  - Causal Inference
  - Quasi-Experiment Design
  - Applied Econometrics
- **Working Experiences**:
  - Person of Contact of Experimentation, Udemy Inc.
  - Tech Lead in Inference and Marketplace Team, Etsy Inc.
  - Research Intern of Time Series Forecasting in Data Mining Research Center, Bosch

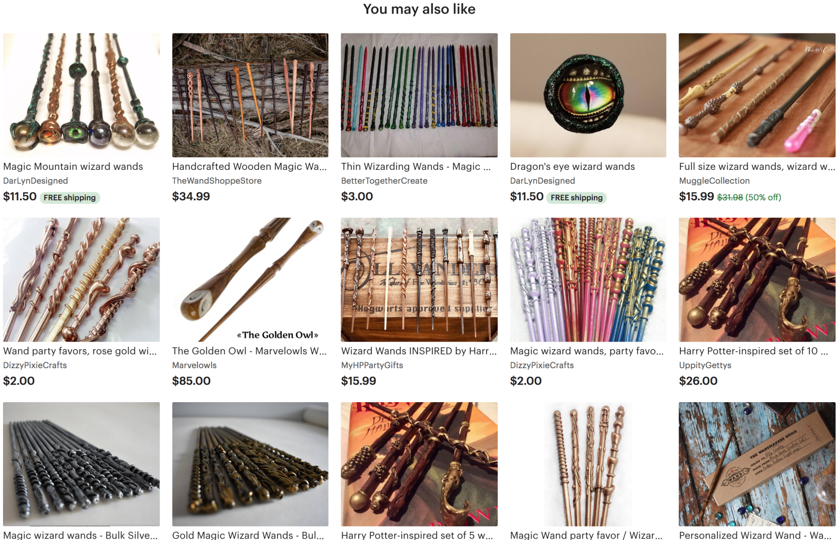
Overview: What is the research about?

- **Background**:
  - Many recommendation experiments did not have statistically significant lifts in conversion and GMV for some time, which impacts the roll-out decision of recommendation features and the roadmap of recommendation system.
  - In recommendation module A/B tests, changes in recommendation module caused reduction in user engagement with the organic search
- **Goal**: To show: In recommendation module A/B tests, the reduction in user engagement with the organic search cannibalized the statistically significant lifts in conversion and GMV that changes in recommendation module should have brought.
- **How**: Through the causal identification and estimation of direct and indirect effects using data of A/B tests

Introduction: Examples of Online Products: Organic Search and Promoted Listings



Introduction: Examples of Online Products: Recommendation Module



Introduction

We see changes in user engagement with other products from A/B test results of one product.

- A change in one product can cause users to change their behaviors in other products.
- Examples I:

Table: **Recommendation Module** A/B Test Average Treatment Effect (**ATE**)

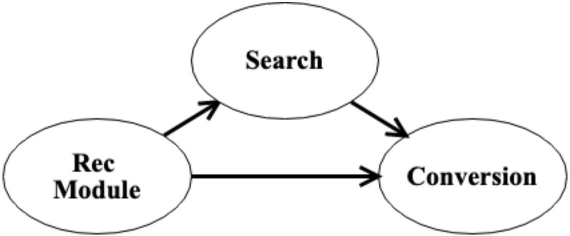
Number of clicks on **Recommendation Modules**	Signficant $\uparrow$
Number of clicks on **Organic Search** results	Significant $\downarrow$
Conversion/Gross merchandise value (GMV)	**Insignificant** Change

Background

- Recommendation team did not have statistically significant lifts in conversion and GMV in A/B tests for consecutive two quarters.
- A typical decision-making criteria is that we will not roll-out the treatment unless we see statistically significant lifts in business KPI (conversion and GMV).
  - Recommendation team roll out few features for consecutive two quarters.
- What is the future of recommendation module? What is the direction of development for recommendation module?
- This study lights up the roadmap for recommendation module.

Fact and Hypothesis

- **Fact**: In recommendation module A/B tests, changes in recommendation module caused reduction in user engagement with the organic search.
- **Hypothesis**: Is it that changes in recommendation module should have brought us statistically significant lifts in conversion and GMV but it has been cannibalized by the reduction in user engagement with the organic search?
  - Given users purchase plan unchanged, users can easily find what they need through the improved recommendation module so that they don't need to search as much as usual.



Problems of Funnel Analysis I

Many e-commerce companies use **Tight Attribution Metric as KPI**.

- purchase funnel: click A in rec module$\Rightarrow$purchase A

Problems: Too Heuristic, No Foundation

- **Ambiguous**

  click A in rec module

  $\Rightarrow$click A in search results

  $\Rightarrow$click A in many different places

  $\Rightarrow$purchase A

  Which place shall get the point?

- **Too Narrow**

  view the rec module, dwell time $\uparrow$, but never click it

  $\Rightarrow$purchase sth elsewhere

  Shall rec module get any point?

Problems of Funnel Analysis II

The severest problem of funnel analysis in A/B tests:

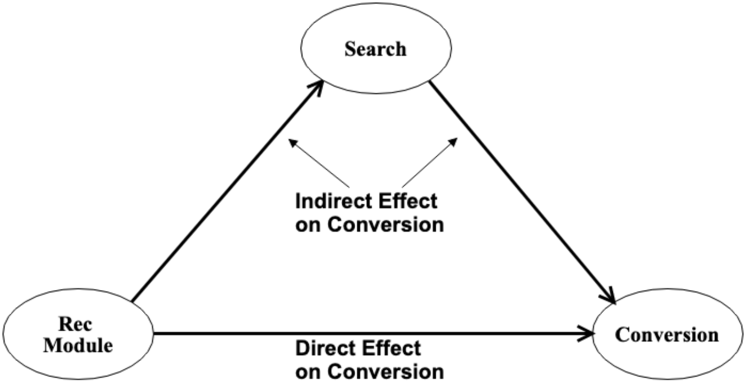
It may destroy the causal interpretation of experimental results.

Because

- It subsets the experimental results based on post-treatment criteria (the pre-defined purchase funnel).
- Conditional on post-treatment variable, the randomization of treatment assignment may no longer hold.
  (i.e., it could break **ignorability** of the identification of **ATE**)
  See, e.g., Montgomery et al. (2018)

An Idea: Direct and Indirect Effects

How about we split **ATE** to two parts: **Direct Effect** and **Indirect Effect**?



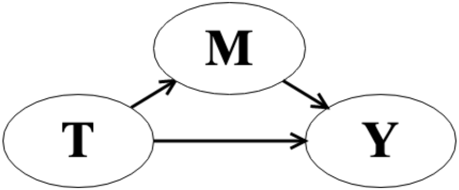
Note that, online experimentation literature messes up two different concepts: "causal effects" and "outcomes" by one word "metric".

Task: Formalize the Idea With Formal Causal Inference Language

- A/B tests cannot give us **Direct Effect** or **Indirect Effect**.
- It can only identify **ATE**.
- To conduct analysis, We need to formalize the idea using formal causal inference language.

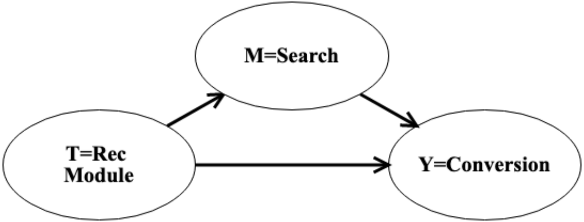
Introduction to Potential Outcome Framework

- In an A/B test, a user i is randomly assigned to either treatment group ($T_i = 1$) or control group ($T_i = 0$).
- Let $M_i(t)$ denote her potential mediator under treatment t.
- Let $Y_i(t, m)$ denote her potential outcome under the treatment t and the mediator m.
- Only one of potential mediators and only one of potential outcomes can be observed for each user.



Examples of Potential Outcomes

In recommendation module A/B tests,



- $M_i(1)$ is her numbers of clicks on search results if she was presented with the new recommendation module.
- $Y_i(1, M_i(0))$ is her conversion status if she was presented with the new rec module and clicked on search results as if she had been presented with the old one.

Causal Identification

The Fundamental Research Question of Causal Inference Is **Identification**.

- Causal effect: the difference between potential outcomes.

Causal Identification

Assumptions $\Rightarrow$ **Causal Effects**

Example: Identification in Rubin Causal Model
The Model Behind A/B Tests

Identification of ATE

Strong Ignorability **and** ***SUTVA*** $\Rightarrow$ **ATE**

- **ATE** on $Y := \mathbb{E}(Y_i(1, M_i(1))) - \mathbb{E}(Y_i(0, M_i(0)))$
- **ATE** on $M := \mathbb{E}(M_i(1)) - \mathbb{E}(M_i(0))$
- ***Strong Ignorability***: $\{Y_i(0), Y_i(1)\} \perp\!\!\!\perp T_i$ and $0 < \mathbb{P}(T_i = t) < 1$
- ***SUTVA***: *Stable Unit-Treatment-Value Assumption*

Causal Mediation Analysis (**CMA**)

Causal Effect Definition

Average Direct Effect (ADE)

$$\mathbf{ADE}(t) := \mathbb{E}(Y_i(1, M_i(t))) - \mathbb{E}(Y_i(0, M_i(t)))$$

- $\mathbf{ADE}(0)$ is the **direct effect** of the rec module change on conversion **leaving aside the induced change**.
- Because mediator is fixed at $M(t)$, the difference between the two potential outcomes can only be attributed to the two different treatments.

Average Causal Mediation Effect (ACME, Indirect Effect)

$$\textbf{ACME}(t) := \mathbb{E}(Y_i(t, M_i(1))) - \mathbb{E}(Y_i(t, M_i(0)))$$

- $\textbf{ACME}(1)$ is the average effect of the *induced change* in organic search clicks upon conversion given users were presented with the new rec module all the time.
- Because treatment is fixed at t, the difference between the two potential outcomes can only be attributed to the two different potential mediators, which are *induced* by different treatments.

Identification of Direct and Indirect Effects in **CMA** (Imai et al., 2010)

Sequential Ignorability (SI) **and** ***SUTVA*** $\Rightarrow$ **ACME and ADE**

SI: add two extra conditions to ***Strong Ignorability***:

$$Y_i(t', m) \perp\!\!\!\perp M_i(t) | T_i = t$$
$$0 < \mathbb{P}(M_i(t) = m | T_i = t) < 1$$

Conditional on the treatment, each potential mediator behaves like the treatment and is ignorable to any potential outcomes.

Causal Mediation Analysis (**CMA**)

Estimation

- The parametric identification of ACME and ADE implies the estimation of them.

Estimation of Direct and Indirect Effects in **CMA** (Imai et al., 2010)

Two Linear Regression Equations

$$
\begin{aligned}
M_{i1} &= \theta_{M_1 0} + \theta_{M_1 1} T_i + \mu_{M_1} \\
Y_i &= \theta_{Y0} + \theta_{Y1} T_i + \theta_{Y2} M_{i1} + \theta_{Y3} M_{i1} T_i + \mu_Y \\
\textbf{ADE}(t) &= \theta_{Y1} + \theta_{Y3}(\theta_{M_1 0} + \theta_{M_1 1} t) \\
\textbf{ACME}(t) &= \theta_{M_1 1}(\theta_{Y2} + \theta_{Y3} t)
\end{aligned}
$$

- Easy to implement in practice (just two linear regression equations!)
- Just use the data from existing A/B test
- No requirements on extra randomization/intervention

Threats to Identification of **CMA** in Practice

- Multiple unmeasured causally-dependent mediators in A/B tests break ***SI*** and invalidate **CMA**.
- **Fat Hand** (Peysakhovich and Eckles, 2018)

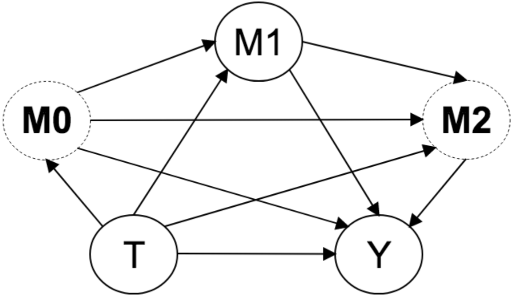


Figure: **M0** and **M2** are unmeasured upstream and downstream mediators of M1

New Problem in Practice:

- **Question**: If we stick with those two linear regression equations for estimation in practice, given the identification of ACME and ADE fail in reality, will we still have estimates of any causal effects? What are the assumptions we need to make to justify they are causal effects?
- **Our Contribution**: Yes, we can still get estimates of some causal effects from the two linear regression equations but with different assumptions.

What We Do: Define New Direct Effect

Generalized Average Direct Effect (GADE)

$$
\begin{aligned}
\mathbf{GADE}(t) \;=\; & \mathbb{E}[Y_i(1, \boldsymbol{M}_{i0}(1), M_{i1}(t, \boldsymbol{M}_{i0}(t)), \\
& \qquad \boldsymbol{M}_{i2}(1, \boldsymbol{M}_{i0}(1), M_{i1}(t, \boldsymbol{M}_{i0}(t))))] \\
& - \mathbb{E}[Y_i(0, \boldsymbol{M}_{i0}(0), M_{i1}(t, \boldsymbol{M}_{i0}(t)), \\
& \qquad \boldsymbol{M}_{i2}(0, \boldsymbol{M}_{i0}(0), M_{i1}(t, \boldsymbol{M}_{i0}(t))))]
\end{aligned}
$$

- It captures the causal effect of the treatment T_i that goes through all the channels that do not have M_{i1}:
  $T \to Y$
  $T \to \boldsymbol{M}_0 \to Y$
  $T \to \boldsymbol{M}_0 \to \boldsymbol{M}_2 \to Y$
  $T \to \boldsymbol{M}_2 \to Y$

What We Do: Define New Indirect Effect

Generalized Average Causal Mediation Effect (GACME, Indirect Effect)

$$\begin{aligned}\textbf{GACME}(t) &= \mathbb{E}[Y_i(t, \boldsymbol{M}_{i0}(t), M_{i1}(1, \boldsymbol{M}_{i0}(1)), \\ &\qquad \boldsymbol{M}_{i2}(t, \boldsymbol{M}_{i0}(t), M_{i1}(1, \boldsymbol{M}_{i0}(1))))] \\ &\quad - \mathbb{E}[Y_i(t, \boldsymbol{M}_{i0}(t), M_{i1}(0, \boldsymbol{M}_{i0}(0)), \\ &\qquad \boldsymbol{M}_{i2}(t, \boldsymbol{M}_{i0}(t), M_{i1}(0, \boldsymbol{M}_{i0}(0))))]\end{aligned}$$

- It captures the causal effect of the treatment T_i that goes through all the channels that have M_{i1}:
  $T \to M_1 \to Y$
  $T \to M_1 \to \boldsymbol{M}_2 \to Y$
  $T \to \boldsymbol{M}_0 \to M_1 \to Y$
  $T \to \boldsymbol{M}_0 \to M_1 \to \boldsymbol{M}_2 \to Y.$

What We Do: The Identification Assumptions

- ***Generalized SI***: Each potential mediator, conditional on the treatment and its upstream mediators, behave like the treatment and is ignorable to all the potential outcomes and all the potential downstream mediators.
- ***LSEM***: *Linear Structural Equation Model*. Potential mediators, potential outcomes, and treatment have linear relationships.

Generalized SI and ***LSEM*** $\Rightarrow$ **GADE** and **GACME**

What We Do: How to Estimate Using Real Data

- We estimate **GACME** and **GADE** by General Method of Moments (GMM) on the same two linear regression equations.
  - GMM can give us estimators that have theoretically minimum standard errors.
- We still have those goods from the two linear regression equations.
  - Easy to implement in practice (just two linear regression equations!)
  - Just use the data from existing A/B test
  - No requirements on extra randomization/intervention

What We Do: How to do Hypothesis Testing

- We estimate the asymptotic variances of estimators by Delta method.
- We test H_0: **GADE** $= 0$ and H_0: **GACME** $= 0$ based on asymptotic normality.

What We Do: The Relationship to The Literature

Case 1: No Unmeasured Upstream and Downstream Mediators

GADE and **GACME** collapse to **ADE** and **ACME**.

Case 2: Unmeasured Upstream or Downstream Mediator

We cannot identify **ADE** and **ACME**.

However, we can identify **GADE** and **GACME**

In practice, it is difficult and costly to know or to estimate the upstream or downstream mediators.

Estimates of Causal Effects for Recommendation Module A/B Test

Mediator is Organic Search Clicks

	Outcome: Conversion	
Effect	% Change	Std Error
GADE(0)	0.4959%*	0.000272
GADE(1)	0.4905%*	0.000271
GACME(0)	-0.2703%***	0.000047
GACME(1)	-0.2757%***	0.000049
ATE	0.2202%	0.000275

1) % Change = Effect/Mean of Control

2) '***' $p < 0.001$, '**' $p < 0.01$, '*' $p < 0.05$, '.' $p < 0.1$. Two-tailed p-value is derived from z-test for H_0: the effect is zero, which is based on asymptotical normality.

Take-Aways

- The causal mediation analysis confirms our hypothesis that search and recommendation systems compete for users' attention and the competition cannibalized the revenue the improved recommendation module should have brought to us.
- It lights up the road map of search and recommendation: search team and recommendation team should collaborate more in the future or maybe even have a joint OKR.
  - Product Design of Complement Goods: search toothpaste; recommend tooth brushes
  - Machine Learning Literature: joint optimization of search and recommendation
- To unblock the roll out of the recommendation feature, direct effect seems to be a better KPI than ATE for decision-making criteria in A/B tests.
  - Note that, online experimentation literature messes up two different concepts: "causal effects" and "outcomes" by one word "metric".

Introduction

We see changes in user engagement with other products from A/B test results of one product.

- Examples II:

Table: **Promoted Listing** A/B Test Average Treatment Effect (**ATE**)

Promoted Listing	
click-through-rate	Significant $\uparrow$
number of clicks	Significant $\uparrow$
advertising revenue	Significant $\uparrow$
Number of clicks on **Organic Search** results	Significant $\downarrow$
Conversion/Gross merchandise value (GMV)	**Insignificant** Change

Estimates of Causal Effects for Promoted Listing A/B Test

Mediator is Organic Search Clicks

	Outcome: Conversion	
Effect	% Change	Std Error
GADE(0)	-0.1448%	0.000203
GADE(1)	-0.1472%	0.000202
GACME(0)	-0.2237%***	0.000034
GACME(1)	-0.2261%***	0.000034
ATE	-0.3709%	0.000205

1) % Change = Effect/Mean of Control

2) '***' $p < 0.001$, '**' $p < 0.01$, '*' $p < 0.05$, '.' $p < 0.1$. Two-tailed p-value is derived from z-test for H_0: the effect is zero, which is based on asymptotical normality.

Imai, K., L. Keele, and T. Yamamoto (2010). Identification, inference and sensitivity analysis for causal mediation effects. *Statistical Science*.

Montgomery, J. M., B. Nyhan, and M. Torres (2018). How Conditioning on Posttreatment Variables Can Ruin Your Experiment and What to Do about It. *American Journal of Political Science 62*(3), 760–775.

Peysakhovich, A. and D. Eckles (2018). Learning causal effects from many randomized experiments using regularized instrumental variables. In *The Web Conference 2018 (WWW 2018)*, New York, NY. ACM.