Estimating Ad Impact on Clicker Conversions for Causal Attribution: A Potential Outcomes Approach

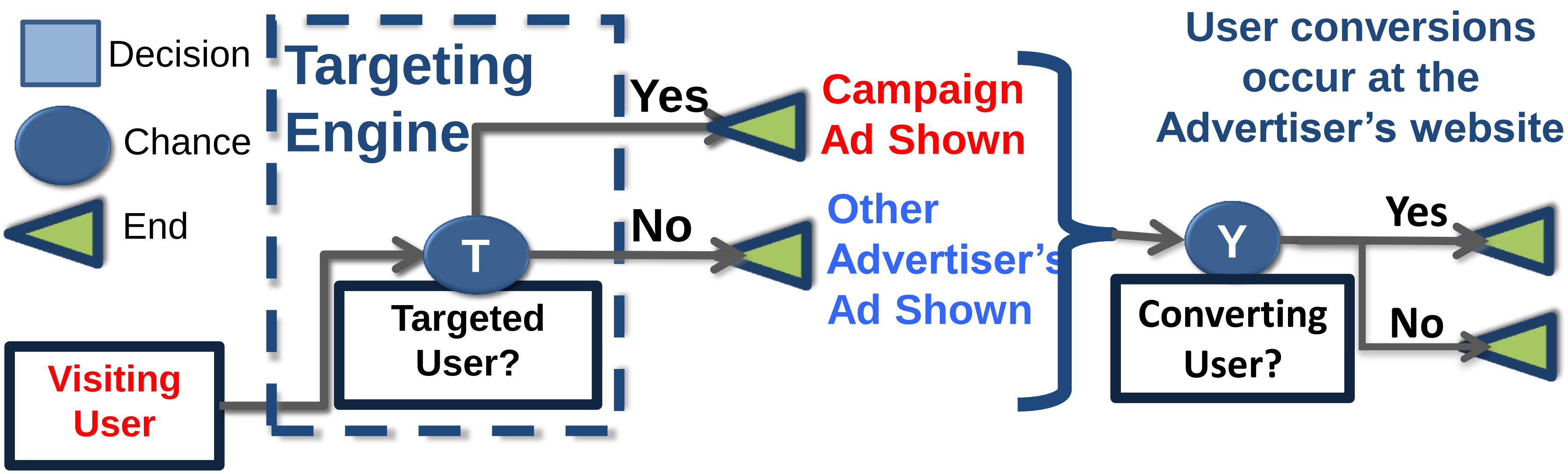
Joel Barajas*[1], Ram Akella[1, 2],
Marius Holtan[3], Aaron Flores[3]

[1]UC Santa Cruz, Santa Cruz CA, USA
[2]UC Berkeley, Berkeley, CA USA
[3]AOL Research, Palo Alto CA, USA

Introduction and Motivation

User Clicks and Causality: Why We Care about Them

Online Display Advertising Framework



- In performance-based online advertising (CPA campaigns), visiting users are targeted based on how likely they are to convert.
- Targeted and non-targeted users convert in the advertiser's website potentially (observed by tracking cookies)
- Converting users regardless of the ad exposure have motivated the causal analysis of campaign effect

Introduction and Motivation

- Are user clicks to online ads informative of campaign success?
  - **Advertiser**: **"All I care about are user conversions"**

- Is it a good idea to target users based on user clicks?
  - **Issues**: low click-through-rates, noisy clicks
  - **Benefits**: conversion rates are even lower

- Dalessandro *et al.* 2012: "Evaluate" click-based user targeting using prediction metrics (e.g. Area Under the Curve AUC)
  - **Conclusion: User targeting based on clicks is statistically indistinguishable from random guessing**
  - **Consequence:** User clicks are often ignored in the user targeting unless they appear to be "effective" to Machine Learning techniques (conversion prediction)

Our Proposal

- Use randomized experiments to measure the ad exposure effect on the user conversion probability of the clickers
  - Running randomized experiments (a.k.a A/B testing) is becoming **the standard practice** to measure the causal ad attribution effectively

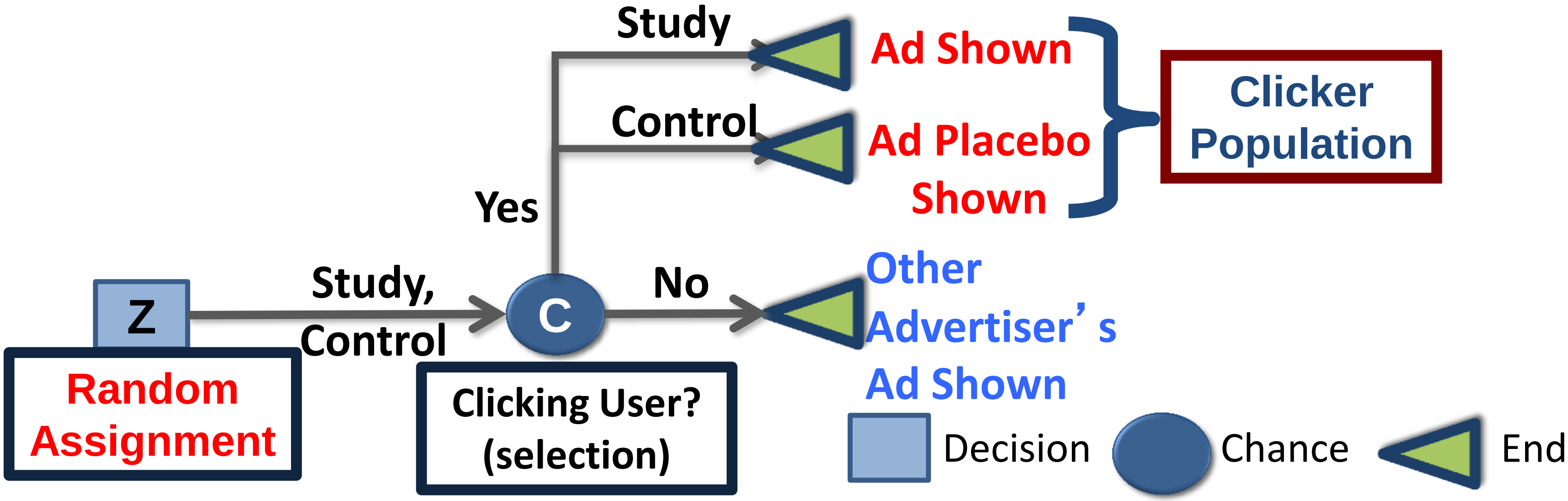
- Is it possible to do design a clean focused randomized experiment?
  - **Unfortunately NO!!**
  - **Reason:** We need to show the ad to know the population of interest

- We use the standard Ad evaluation practice and find the Local Average Treatment Effect on the clicker conversion probability
  - **Even though** we design an experiment, we need causal modeling to find this effect
  - **Reason:** We do **not observe** the user clicks in the control group

Randomized Design

Use A/B Testing: what's the big deal?

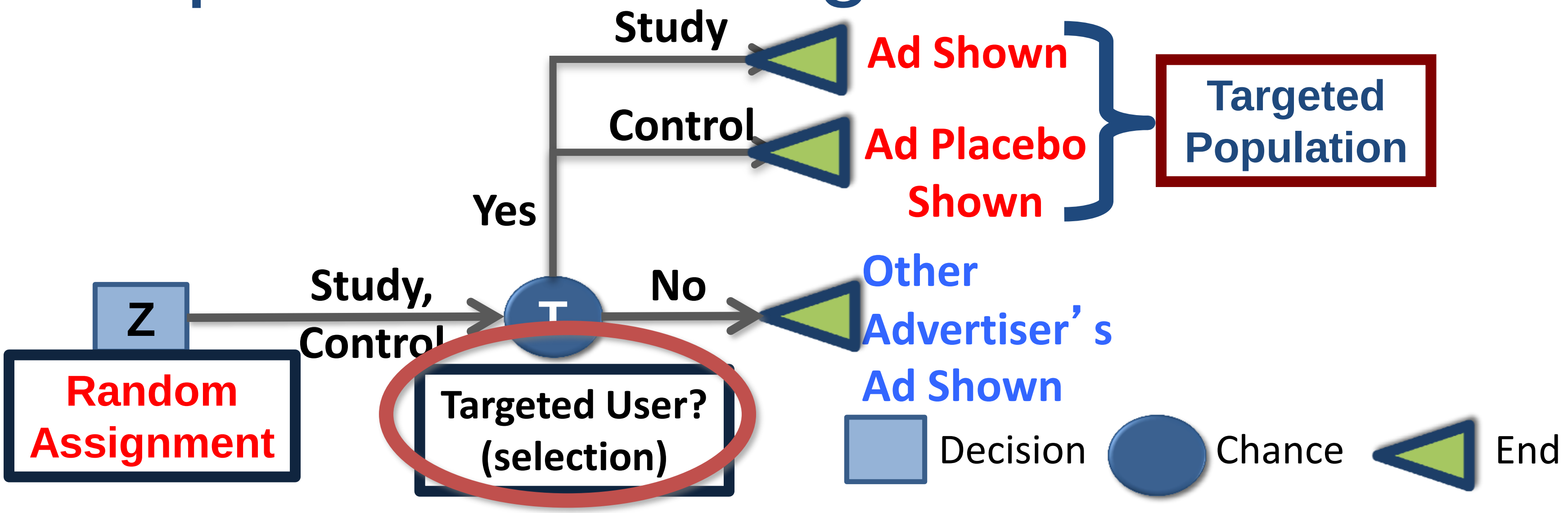
Ideal Experimental Design



- **We can not design a parallel placebo campaign for the control group**

BY THE TIME WE KNOW THE POPULATION OF INTEREST, WE HAVE SHOWN THE AD ALREADY!!

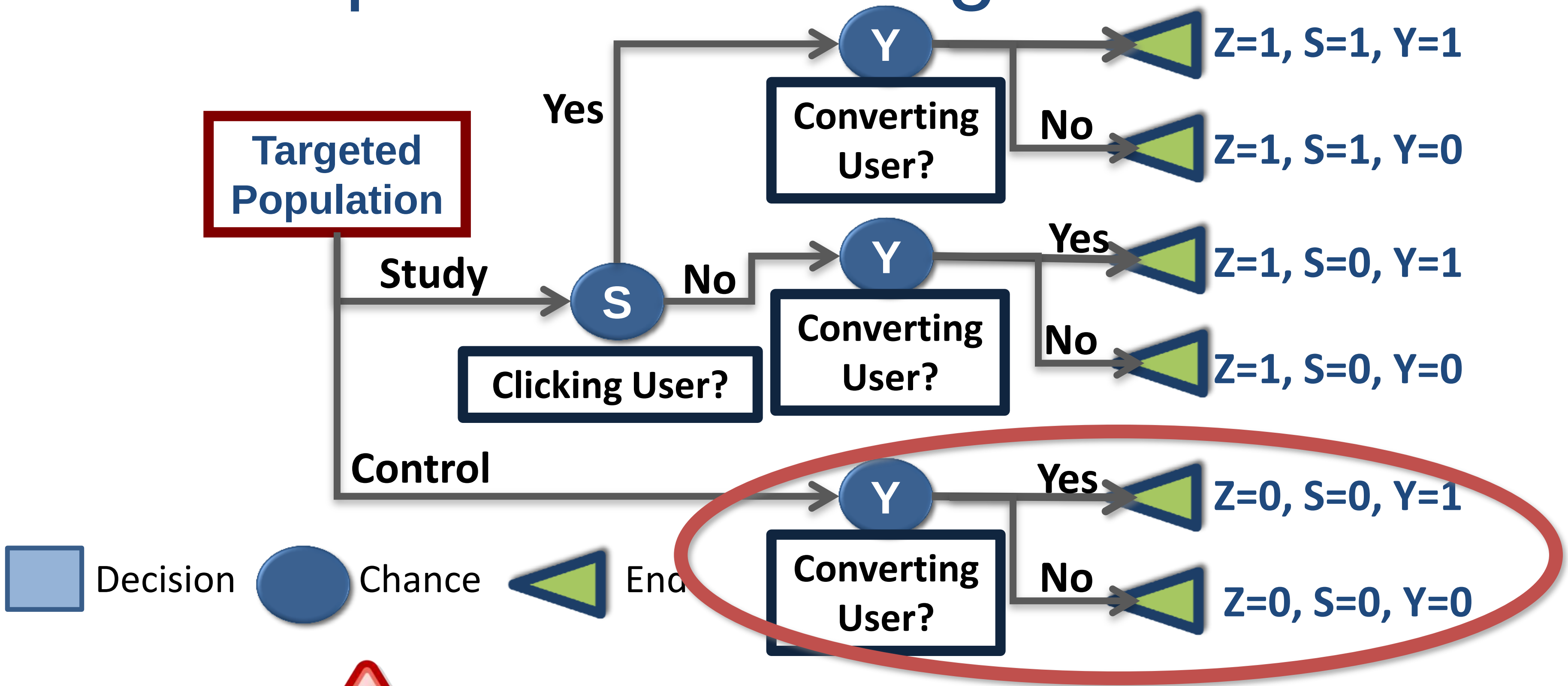
Experimental Design: Ad Evaluation



- **Targeted user sub-population is possible to replicate (under certain conditions)**

WE FIND THE LOCAL AD EFFECT ON THE SUB-POPULATION OF CLICKERS

Experimental Design Data



- **Nobody click on the ad in the control group by definition.**

FIND THE AD EFFECT CONDITIONAL ON THE CLICKER POST-TREATMENT VARIABLE

Effect Estimation

Wait a minute: you ran an experiment, why you need to infer causality?

Effect Estimation

Because the effect is conditional on a post-treatment (non-ignorable) variable

- **Pre-treatment variable**: Observed **before** the treatment assignment is made (e.g. demographics)
- **Post-treatment variable**: Observed **after** the treatment assignment is made (e.g. behavioral features, and user clicks)

Causal Estimation: Potential Outcomes

- We use the **Principal Stratification** framework, developed in **Potential Outcomes Causal Model**, to model the user selection as a post-treatment variable.

- In Potential Outcomes, everything framed in terms of: **Units, Treatments, and Potential Outcomes.**

- In Principal Stratification, the effect estimation is conditioned on user classes characterized by the value of the post-treatment variable under **both treatment groups,** control and study.

The Estimation Model: Mechanics

- $Z_i \in \{0,1\}$: Treatment assignment {control,study}
- $S_i \in \{0,1\}$: Non-clicking/clicking ad user
- $Y_i \in \{0,1\}$: Non-converting/converting user

$$S_i^P = \left\{ \begin{pmatrix} S_i(0) \\ S_i(1) \end{pmatrix} \right\} = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$$

Define

$$C_i = \begin{cases} 0 \; if \; S_i^P = (0,0)' \\ 1 \; if \; S_i^P = (0,1)' \end{cases}$$

- $C_i = 1$ defines the principal stratum of users who click on the ad **when assigned** to the study group (**clicker-if-assigned**)

- $C_i = 0$ defines the principal stratum of users who **never click** on the ad regardless of the treatment group (**never-clicker**)

What we do and do not Know

User Counts	Potential Outcomes				Treatment Assignment	Principal Stratum	
	Control		Study				
N^{y}_{cz}	$S_i(0)$	$Y_i(0)$	$S_i(1)$	$Y_i(1)$	Z_i	$(S_i(0), S_i(1))$	C_i
$N^{0}_{\{0,1\}0}$	0	0	*	*	0	(0,*)	*
$N^{1}_{\{0,1\}0}$	0	1	*	*	0	(0,*)	*
N^{0}_{01}	0	*	0	0	1	(0,0)	0
N^{1}_{01}	0	*	0	1	1	(0,0)	0
N^{0}_{11}	0	*	1	0	1	(0,1)	1
N^{1}_{11}	0	*	1	1	1	(0,1)	1

We know all necessary data of the study group

What we do and do not Know

User Counts	Potential Outcomes				Treatment Assignment	Principal Stratum	
	Control		Study				
N^y_{cz}	$S_i(0)$	$Y_i(0)$	$S_i(1)$	$Y_i(1)$	Z_i	$(S_i(0), S_i(1))$	C_i
$N^0_{\{0,1\}0}$	0	0	*	*	0	(0,*)	*
$N^1_{\{0,1\}0}$	0	1	*	*	0	(0,*)	*
N^0_{01}	0	*	0	0	1	(0,0)	0
N^1_{01}	0	*	0	1	1	(0,0)	0
N^0_{11}	0	*	1	0	1	(0,1)	1
N^1_{11}	0	*	1	1	1	(0,1)	1

We need to infer the principal stratum of the control group

The Inference Problem

By **randomization**, the probability of *clicker-if-assigned* (the principal stratum) is the same for **both treatment groups**

$$P(Y, Z, D, \Theta) = P(\Theta) \times \prod_{\forall i} P(C_i|\pi)P(Z_i)P(Y_i|C_i = c, Z_i = z, \theta_{cz})$$

The Inference Problem

By **randomization**, the probability of *clicker-if-assigned* (the principal stratum) is the same for **both treatment groups**

- The user conversion probability for both principal strata in the control group is **not identifiable**

$$P(Y, Z, D, \Theta) = P(\Theta) \times \prod_{\forall i} P(C_i|\pi)P(Z_i)P(Y_i|C_i = c, Z_i = z, \theta_{cz})$$

The Inference Problem

By **randomization**, the probability of *clicker-if-assigned* (the principal stratum) is the same for **both treatment groups**

- The user conversion probability for both principal strata in the control group is **not identifiable**

- We assume **POSITIVE campaign effect**

$$\boxed{\theta_{c1} \geq \theta_{c0} \quad \forall c \in \{0,1\}}$$

$$P(Y,Z,D,\Theta) = P(\Theta) I_{[0,\theta_{01})}(\theta_{00}) I_{[0,\theta_{11})}(\theta_{10}) \times \prod_{\forall i} P(C_i|\pi) P(Z_i) P(Y_i|C_i = c, Z_i = z, \theta_{cz})$$

The Inference Problem

- By **randomization**, the probability of *clicker-if-assigned* (the principal stratum) is the same for **treatment groups**

- both

We use Gibbs sampling with truncated distributions to fit the model

- We assume **POSITIVE cam**

$$\boxed{\theta_{c1} \geq \theta_{c0} \quad \forall c \in \{0,1\}}$$

$$P(Y,Z,D,\Theta) = P(\Theta) I_{[0,\theta_{01})}(\theta_{00}) I_{[0,\theta_{11})}(\theta_{10}) \times \prod_{\forall i} P(C_i|\pi) P(Z_i) P(Y_i | C_i = c, Z_i = z, \theta_{cz})$$

The Inference Problem

By **randomization**, the probability of *clicker-if-assigned* (the principal stratum) is the sa... for ...h treatment groups

We use G...
with tr...
...g
...butions
...odel

See the paper for the Gibbs sampling expressions and algorithm

...SITIVE cam...

... $\forall c \in \{0,1\}$

$$P(Y,Z,D,\Theta) = P(\Theta) I_{[0,\theta_{01})}(\theta_{00}) I_{[0,\theta_{11})}(\theta_{10}) \times \prod_{\forall i} P(C_i|\pi)P(Z_i)P(Y_i|C_i = c, Z_i = z, \theta_{cz})$$

Causal Effect Metrics

- Causal **average ad effect** on conversion rate

$$LATE_{Click} = E[Y_i|C_i = 1, Z_i = 1, \theta_{11}] - E[Y_i|C_i = 1, Z_i = 0, \theta_{10}] = \theta_{11} - \theta_{10}$$

$$LATE_{NoClick} = E[Y_i|C_i = 0, Z_i = 1, \theta_{01}] - E[Y_i|C_i = 0, Z_i = 0, \theta_{00}] = \theta_{01} - \theta_{00}$$

- Observational and post-treatment **biased effects**
- ATE^{obs}_{Click} is similar to a **prediction** "evaluation"

$$ATE^{obs}_{Click} = E[Y_i|S_i(1) = 1, Z_i = 1] - E[Y_i|S_i(1) = 0, Z_i = 1]$$

$$ATE^{post}_{Click} = E[Y_i|S_i(1) = 1, Z_i = 1] - E[Y_i|S_i(0) = 0, Z_i = 0]$$

- **Attributed converting users** respect to the observed study conversions

$$ATRB_{Click} = LATE_{Click} \times (N^0_{11} + N^1_{11}) \,/\, (N^1_{01} + N^1_{11})$$

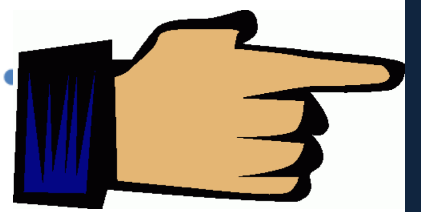
$$ATRB_{NoClick} = LATE_{NoClick} \times (N^0_{01} + N^1_{01}) \,/\, (N^1_{01} + N^1_{11})$$

Causal Effect Metrics

- Causal **average ad effect** on conversion rate

$$LATE_{Click} = E[Y_i|C_i = 1, Z_i = 1, \theta_{11}] - E[Y_i|C_i = 1, Z_i = 0, \theta_{10}] = \theta_{11} - \theta_{10}$$

$$LATE_{NoClick} = E[Y_i|C_i = 0, Z_i = 1, \theta_{01}] - E[Y_i|C_i = 0, Z_i = 0, \theta_{00}] = \theta_{01} - \theta_{00}$$



Confidence intervals are trivial to obtain from the posterior distribution Gibbs samples

- ATE^{obs}_{Click} is s
  prediction “

- **Attributed converting users** respect to the observed study conversions

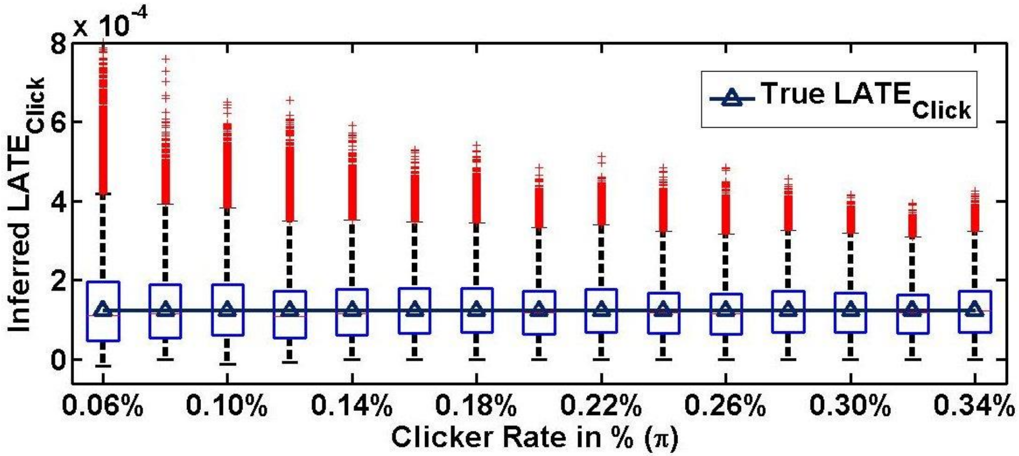
$$\times (N_{11} + N_{11}) / (N_{01} + N_{11})$$

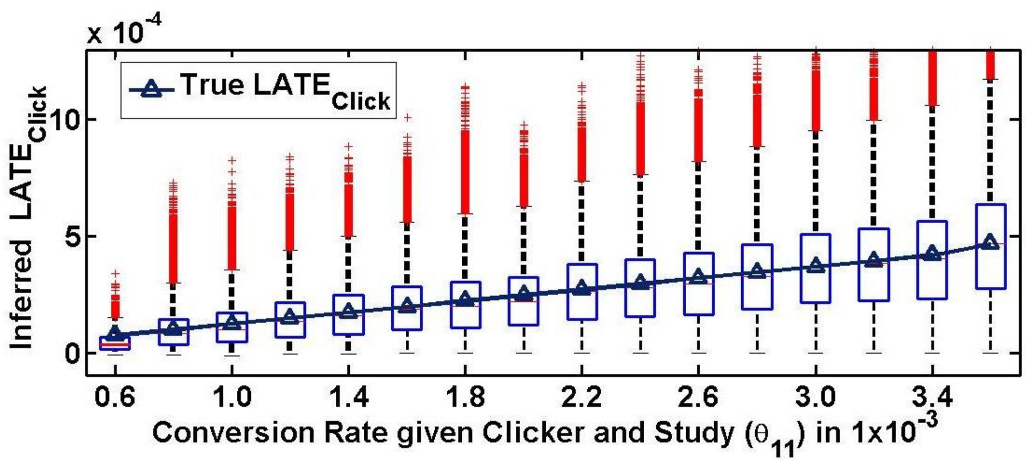
$$ATRB_{NoClick} = LATE_{NoClick} \times (N^0_{01} + N^1_{01}) / (N^1_{01} + N^1_{11})$$

Results

OK, OK, …, but show me the money

Results: Validation





- Given the model parameters
  - We generate 100 count sets
  - We fit our model: 3000 samples
  - Concatenate the posterior samples or each run
- Settings:
  - Population: 23 Million Users
  - P(Z=1)=0.5
  - P(Y=1|C=1,Z=1)=1e-3
  - P(Y=1|C=1,Z=1)=9.82e-5
  - Lift LATE_click = 14%
  - Lift LATE_NoClick = 14%
  - Clicker Rate = 0.068%

BOTTOM LINE: WE RECOVER THE TRUE EFFECT ON THE CLICKER CONVERSION RATE

Large-Scale Randomized Experiment Data

Treatment Assignment Z_i	Principal Stratum $(S_i(0), S_i(1))$	C_i	User Counts N^y_{cz}	Campaign 1	Campaign 2
0	(0,*)	*	$N^0_{\{0,1\}0}$	3,621,409	11,431,495
0	(0,*)	*	$N^1_{\{0,1\}0}$	314	961
1	(0,0)	0	N^0_{01}	3,535,571	11,328,649
1	(0,0)	0	N^1_{01}	347	1,014
1	(0,1)	1	N^0_{11}	2,414	9,799
1	(0,1)	1	N^1_{11}	2	8
Total Population:				**7.16 Million**	**22.7 Million**

- Pre-processing:
  - Use **tracking cookies** (users)
  - **Randomize users** based on the timestamp the cookie was born
  - Focus on the user cookies **born before** the campaign started
  - CPM (**non-optimized**) campaigns for the same advertiser
  - Use a **placebo campaign** to target users in the control group
  - **Objective**: perform an **exploratory** campaign effectiveness analysis

Large-Scale Randomized Experiment Data

Treatment Assignment Z_i	Principal Stratum $(S_i(0), S_i(1))$	C_i	User Counts N^y_{cz}		
0	(0,*)	*	$N^0_{\{0,1\}0}$		
0	(0,*)	*	$N^1_{\{0,1\}0}$		
				3,535,571	11,328,649
				347	1,014
				2,414	9,799
1	(0,1)	1	N^1_{11}	2	8

8.27e-4 vs 9.81e-5

8.16e-4 vs 8.95e-5

Clicker conversion rate in the study group is almost 10 times higher!!

Total Population:| 7.16 Million | 22.7 Million

- Pre-processing:
  - Use **tracking cookies** (users)
  - **Randomize users** based on the timestamp the cookie was born
  - Focus on the user cookies **born before** the campaign started
  - CPM (**non-optimized**) campaigns for the same advertiser
  - Use a **placebo campaign** to target users in the control group
  - **Objective**: perform an **exploratory** campaign effectiveness analysis

Analysis of Results

	Campaign 1				Campaign 2			
Measure		Low	Med	High		Low	Med	High
Clicker Rate, π	(%)	0.0667	[illegible]	0.0699	(%)	0.0851	[illegible]	0.0880
ATE^{obs}_{Click} (naive)	(1e-4)	-2.33	7.30	16.92	(1e-4)	2.54	7.26	12.00
lift ATE^{obs}_{Click}	(%)	-237	**743**	1720	(%)	282	**811**	1340
ATE^{post}_{Click} (biased)	(1e-4)	-2.21	7.41	17.04	(1e-4)	2.54	7.26	12.00
lift ATE^{post}_{Click}	(%)	-255	**855**	1960	(%)	306	**870**	**1400**
ATE_{Camp}	(1e-5)	0.37	[illegible]	1.99	(1e-6)	-0.33	[illegible]	12.50
lift ATE_{Camp}	(%)	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
$ATRB_{Camp}$	(%)	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
$LATE_{NoClick}$	(1e-5)	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
lift $LATE_{NoClick}$	(%)	3.81	13.46	24.21	(%)	1.04	6.97	25.13
$ATRB_{NoClick}$	(%)	3.48	11.78	20.22	(%)	0.99	6.45	13.53
$LATE_{Click}$	(1e-4)	**0.35**	4.61	13.72	(1e-4)	**0.43**	4.65	11.04
lift $LATE_{Click}$	(%)	7.28	150.77	874.20	(%)	7.21	145.85	813.12
$ATRB_{Click}$	(%)	0.02	**0.32**	0.95	(%)	0.04	**0.45**	1.06
C2C ATRB	(%)	-	**0.57**	-	(%)	-	**0.78**	-

Biased effects similar to the prediction based "effectiveness"

SEVERE OVER-ESTIMATION: UPPER BOUND LARGER THAN 1,400%

Analysis of Results

	Campaign 1				Campaign 2			
Measure		Low	Med	High		Low	Med	High
Clicker Rate, π	(%)	0.0667	0.0683	0.0699	(%)	0.0851	0.0865	0.0880
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	(1e-4)	2.54	[illegible]	12.00
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	(%)	282	**811**	1340
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	(1e-4)	2.54	7.26	12.00
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
lift ATE_{Camp}	[illegible]	[illegible]	[illegible]	24.02	[illegible]	[illegible]	[illegible]	[illegible]
$ATRB_{Camp}$	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
$LATE_{NoClick}$	[illegible]	[illegible]	[illegible]	**2.00**	(1e-6)	0.89	5.82	**12.20**
lift $LATE_{NoClick}$	[illegible]	[illegible]	[illegible]	24.21	(%)	1.04	6.97	[illegible]
$ATRB_{NoClick}$	(%)	3.48	11.78	20.22	[illegible]	[illegible]	6.45	13.53
$LATE_{Click}$	(1e-4)	**0.35**	4.61	13.72	(1e-4)	**0.43**	4.65	11.04
lift $LATE_{Click}$	(%)	7.28	150.77	874.20	(%)	7.21	145.85	813.12
$ATRB_{Click}$	(%)	0.02	**0.32**	0.95	(%)	0.04	**0.45**	1.06
C2C ATRB	(%)	-	**0.57**	-	(%)	-	**0.78**	-

Very large confidence intervals due to low clicker rates and non-optimized CPM campaigns: 0.0683%, 0.0865%

3.50e-5 vs 2.00e-5 =
75% increase

4.30e-5 vs 1.22e-5 =
252% increase

EVEN IN A PESSIMISTIC ANALYSIS THE CLICKERS ARE MORE VALUABLE THAN THE NON-CLICKERS

Analysis of Results

Small number of attributed converting users are clickers.
Reason: more than 900 times larger non-clicker volume!!

	Campaign 1				Campaign 2			
Measure		Low	Med	High		Low	Med	High
$ATE^{post}_{...}$ (biased)	(1e-4)	-2.21	7.41	17.04	(1e-4)	2.54	7.26	12.00
[illegible]	[illegible]	[illegible]	**855**	[illegible]	[illegible]	[illegible]	[illegible]	**1400**
[illegible]	[illegible]	[illegible]	1.20	[illegible]	[illegible]	[illegible]	[illegible]	12.50
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	15.43
$ATRB_{Camp}$	(%)	3.74	**12.17**	20.20	(%)	-0.37	**6.80**	13.87
$LATE_{NoClick}$	(1e-5)	0.34	[illegible]	**2.00**	(1e-6)	0.89	[illegible]	**12.20**
lift $LATE_{NoClick}$	(%)	3.81	[illegible]	24.21	(%)	1.04	[illegible]	25.13
$ATRB_{NoClick}$	(%)	3.48	11.78	20.22	(%)	0.99	6.45	13.53
$LATE_{Click}$	(1e-4)	**0.35**	[illegible]	13.72	(1e-4)	**0.43**	[illegible]	11.04
lift $LATE_{Click}$	(%)	7.28	[illegible]	874.20	(%)	7.21	[illegible]	813.12
$ATRB_{Click}$	(%)	0.02	**0.32**	0.95	(%)	0.04	**0.45**	1.06
C2C ATRB	(%)	-	**0.57**	-	(%)	-	**0.78**	-

Non-clickers: 0.32% vs Campaign: 12.17%

Non-clickers: 0.45% vs Campaign: 6.62%

EVEN THOUGH THE AD EFFECT IS HIGHER FOR CLICKERS, FEW CLICKERS ARE OBSERVED!!

Analysis of Results

		Campaign 1				Campaign 2		
Measure		Low	Med	High		Low	Med	High
Cl[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
AT[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
lift[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
AT[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
lift ATE_{Click}	(%)	[illegible]	[illegible]	[illegible]	(%)	[illegible]	[illegible]	[illegible]
ATE_{Camp}	(1e-5)	0.37	1.20	1.99	(1e-6)	-0.33	6.13	12.50
lift ATE_{Camp}	(%)	4.06	[illegible]	24.02	(%)	-0.38	[illegible]	15.43
$ATRB_{Camp}$	(%)	3.74	**12.17**	20.20	(%)	-0.37	**6.80**	13.87
[illegible]	[illegible]	[illegible]	1.16	[illegible]	[illegible]	[illegible]	[illegible]	**12.20**
[illegible]	[illegible]	[illegible]	13.46	[illegible]	[illegible]	[illegible]	[illegible]	25.13
[illegible]	[illegible]	[illegible]	11.78	[illegible]	[illegible]	[illegible]	[illegible]	13.53
$LATE_{Click}$	(1e-4)	**0.35**	4.61	13.72	(1e-4)	**0.43**	4.65	11.04
lift $LATE_{Click}$	(%)	7.28	[illegible]	874.20	(%)	7.21	[illegible]	813.12
$ATRB_{Click}$	(%)	0.02	**0.32**	0.95	(%)	0.04	**0.45**	1.06
C2C ATRB	(%)	-	**0.57**	-	(%)	-	**0.78**	-

A click-to-conversion attribution (C2C), where a user conversion is attributed to the last user click, tends to under-estimate the true campaign value.

C2C: 0.57% vs Campaign: 12.17%

C2C: 0.78% vs Campaign: 6.62%

MOST OF THE USERS WHO ARE AFFECTED BY THE CAMPAIGN DO NOT CLICK ON THE AD!!

Conclusion

Fine, this is too much detail… give me your recommendation

Business Implications

- Optimizing user clicks **optimizes** the **causally** generated conversions.
  - **Positive correlation** of clicks with **incremental** conversions
- Clicks **must be** a component of a combined user targeting policy, but **not the only objective**
  - Many positive affected users **do not click** on the ad
- C2C business model **under-estimates** the causal attribution, **contradicting** the general industry belief that C2C over-estimates the value
  - C2C **does not** consider the ad effects on the **non-clickers**

Conclusion

- We proposed a method to find the causal ad effect on the clicker conversions based on **randomized experiments**
  - The average campaign effect must **not be negative**.
- We found evidence suggesting a **higher ad effect** on the user conversion probability of the **clickers**
  - A **pessimistic** analysis supports this hypothesis
- Our method and analysis open a path for **more studies** of the user clicks
  - Why many users **affected** by the ad **do NOT click** on the ad is an open research problem

Joel Barajas*, Ram Akella, Marius Holtan, Aaron Flores

jbarajas_AT_soe.ucsc_DOT_edu

https://users.soe.ucsc.edu/~jbarajas/

Thank you!!