Online Advertising Incrementality Testing

Industry Practical Lessons And Emerging Challenges

Joel Barajas, **Narayan Bhamidipati, James G. Shanahan**

November, 2021

yahoo!

CIKM 2021
1-5 NOVEMBER

Who we are ...

Joel Barajas
Sr Research Scientist
Yahoo! Research,
Sunnyvale, CA, USA

Linkedin

Narayan Bhamidipati
Sr. Director, Research
Yahoo! Research,
Sunnyvale, CA, USA

Linkedin

James G. Shanahan
Church and Duncan
Group Inc
UC Berkeley, CA, USA

Linkedin

yahoo!

CIKM 2021
1–5 NOVEMBER

Tutorial Parts

1. The basics: context and challenges

2. Incrementality Testing: concepts, solutions and literature

3. From concept to production: platform building, challenges, case studies

4. Deployment at Scale: test cycle and case studies

5. Emerging trends: identity challenges, industry trends and solutions

yahoo!

Part 4

Deployment at Scale: test cycle and case studies

Experimentation: Incrementality Testing Cycle

Experiment Design and Planning

Minimum **Detectable Lift**, Total **Audience** Available, **Hold-out** group size, **Exposed Audience** Needed, Experiment **Duration**, etc

Intervention Execution

Start **holding out users randomly** from showing Ads. **Log "Ghost" Ad** impressions for these users

User Tracking and Metrics

Track **user cohorts**. Compute **aggregate metrics** to guide **high-level decisions** and campaign tuning

Incremental Performance and Recommendations

Document incrementality metrics, **CR lift, Cost per Incremental Action**. Provide **recommendations**, side effects and learnings

yahoo!

Testing Cycle

From the beginning to the end, step by step

yahoo!

Experimentation: Design and Planning

Experiment Design and Planning

Minimum **Detectable Lift**, Total **Audience** Available, **Hold-out** group size, **Exposed Audience** Needed, Experiment **Duration**, etc

Run statistical power analysis to manage expectations, budgets and duration

- In a period of time, given **existing advertiser data**:
  - **Number of users** exposed to impressions, conversions, impressions and total spend
  - **Metrics derived:** **converter rate, total audience, eCPM, ad frequency**

- Parameters to tune:
  - **Holdout size** (*never 50%!!*), minimum detectable lift, test significance

yahoo!

Power Analysis Simulation: 2 months of data

Aggregate Statistics:

Converter rate: **XXX%**

Conversions per 100 users: **XXX**

Conversions per converter: **XXX**

Impressions per user: **XXX**
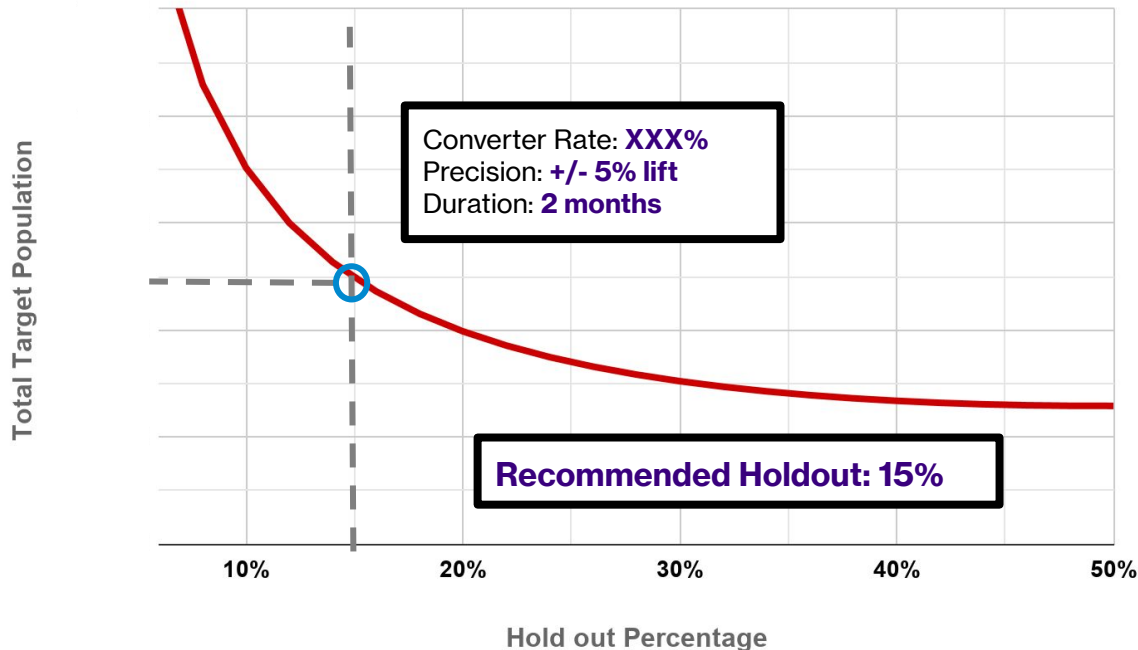
Experiment Users: **XXXM**

Scenario Simulation:

Converter Lift precision: **5%**

Expected Converter Rate:

Control: **<XXX%**

Test: **>XXX%**

Test Significance: **5%**

Converter Rate: **XXX%**
Precision: **+/- 5% lift**
Duration: **2 months**

Recommended Holdout: 15%

Total Target Population

Hold out Percentage

10% 20% 30% 40% 50%

Response: User signup/acquisition

yahoo!

Experimentation: Intervention Execution

Intervention Execution

Start **holding out users randomly** from showing Ads. **Log "Ghost" Ad** impressions for these users

The test actually runs

- Run quick **sanity checks**:
  - Verify **impressions and ghost impressions** are generated
  - Is **percentage of control users** as expected?
  - **Define guidelines** to hit a button and stop the test if needed

yahoo!

Experimentation: User Tracking and Metrics

Track user groups, metrics and mid-flight recommendations

User Tracking and Metrics

Track **user cohorts**. Compute **aggregate metrics** to guide **high-level decisions** and campaign tuning

- Keep track of **top-level metrics**
  - Lift, incremental conversions, user ad frequency, CPIA, iROAS

- Pay special attention to **holidays** and fundamental user behaviour changes
  - For relevant holidays, **cumulative effects are realized** in these special periods (*ad stock* **realization**)
  - For irrelevant holidays sometimes **inventory becomes more expensive**

yahoo!

Experimentation: Performance and Recommendations

Incremental Performance and Recommendations

Document incrementality metrics, **CR lift, Cost per Incremental Action**. Provide **recommendations**, side effects and learnings

Testing is expensive, thus document learnings and insights

- Deliver **top-level metrics**
  - Including statistical tests and intervals

  These metrics will compared and used to calibrate advertiser's financial models

- Document lessons and **make recommendations**
  - What **can be learned from the test** beyond top-level metrics?
  - **Why** the test was (or was not) successful?
  - Make **business recommendations** and campaign management practices

yahoo!

Case Studies

How success
looks like

Business Use Cases: Recap
Barajas and Bhamidipati (2021)

- **New advertiser** wants to *test waters* before fully deploying budgets

- Existing advertiser wants to **scientifically prove incremental value**

- Incrementality of **tactical practices** with external validity for future deployment

- Quarterly regular **strategic adjustments** (retrospectively)

yahoo!

Case Studies

New advertiser wants to test waters before fully deploying budgets

yahoo!

Onboarding Funnel: Key Findings

+200%

Statistically Significant Lift Effect on **Landing Page Visits (funnel entry)**

12.4%

Consistently Positive Average Lift Effect on **Purchase Confirmation (successful end of the funnel)**

Incremental Funnel Entries and Successful Endings

yahoo!

Lift Results: Funnel Entries, Stages and Successful Endings

Lift on **Landing Page**, Funnel-entry Conversions **225% +/- 17%**

Lift on **User Logins**, incremental quality (sticky) users **16.45% +/- 12%**

Lift on **Complete Purchase Page Visit**, less funnel transitions **-6.58% +/- 8%**

Takeaways:
- ➔ Incremental Funnel Entries, with **sticky more valuable users** (more logins)
- ➔ More straight Purchases (end of funnel) with **less transitions** (complete purchase page)

Lift on **Purchase Confirmation**, end-of-funnel conversions **12.4% +/- 10%**

yahoo!

95% Confidence Intervals. Effects on funnel occurrences, not transitions.

Marketing Effectiveness goes beyond one conversion metric

yahoo!

Case Studies

Existing advertiser wants to scientifically prove incremental value

yahoo!

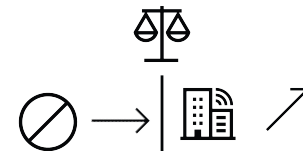# Insurance quotes: Key Findings

10.5%
Converter rate lift
(Statistically significant)

$XX
**Cost per Incremental Conversion** over 8 weeks of experiment

9.5%
Lift in Avg conversions per user (Statistically significant)

87%
Lower Cost per Incremental Conversion than post-click CPA

yahoo!

Key Insights: Incrementality Real Study

Industry standard last-touch attribution undervalued native display programmatic advertising by 87%

Lift: 10.5% +/- 1.56%

Converter Lift

Cool-off Period | November | December
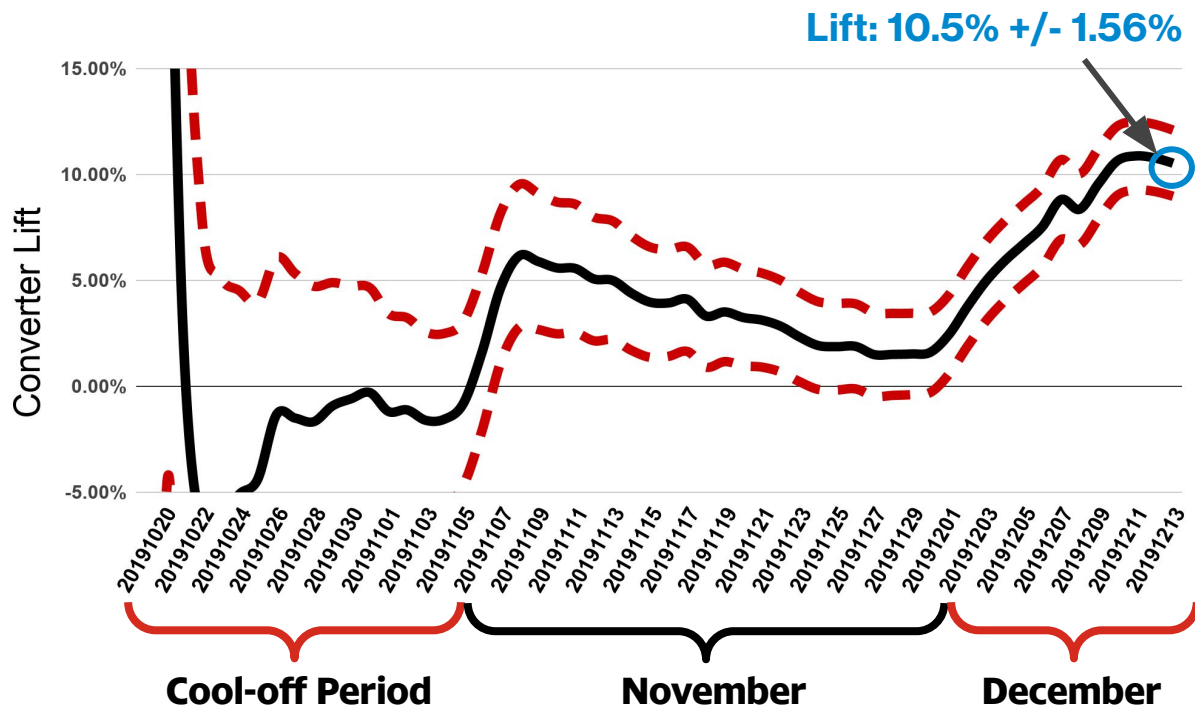
Cumulative converter lift progression. 95% confidence intervals are displayed. Conversion definition: insurance quote. Duration: 8 weeks, 10/18/2019 - 12/12/2019. Holdout size: 10%.

yahoo!

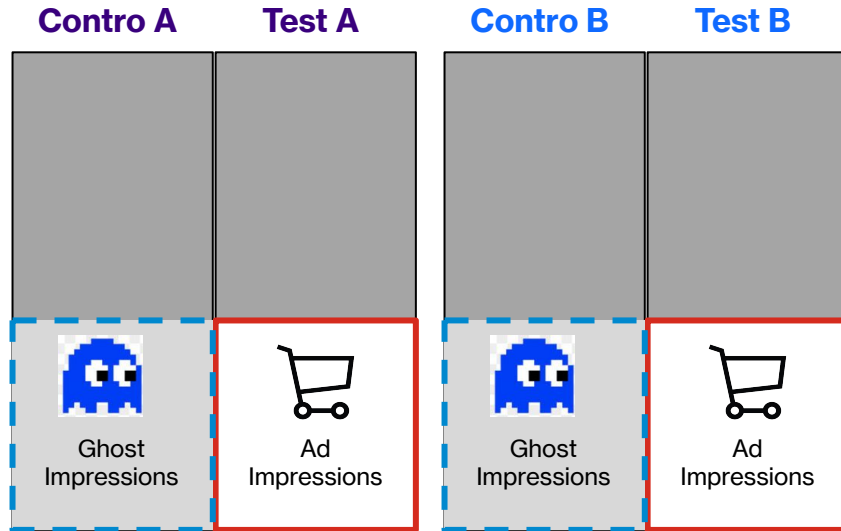Seasonalities are critical and change user behaviour fundamentally
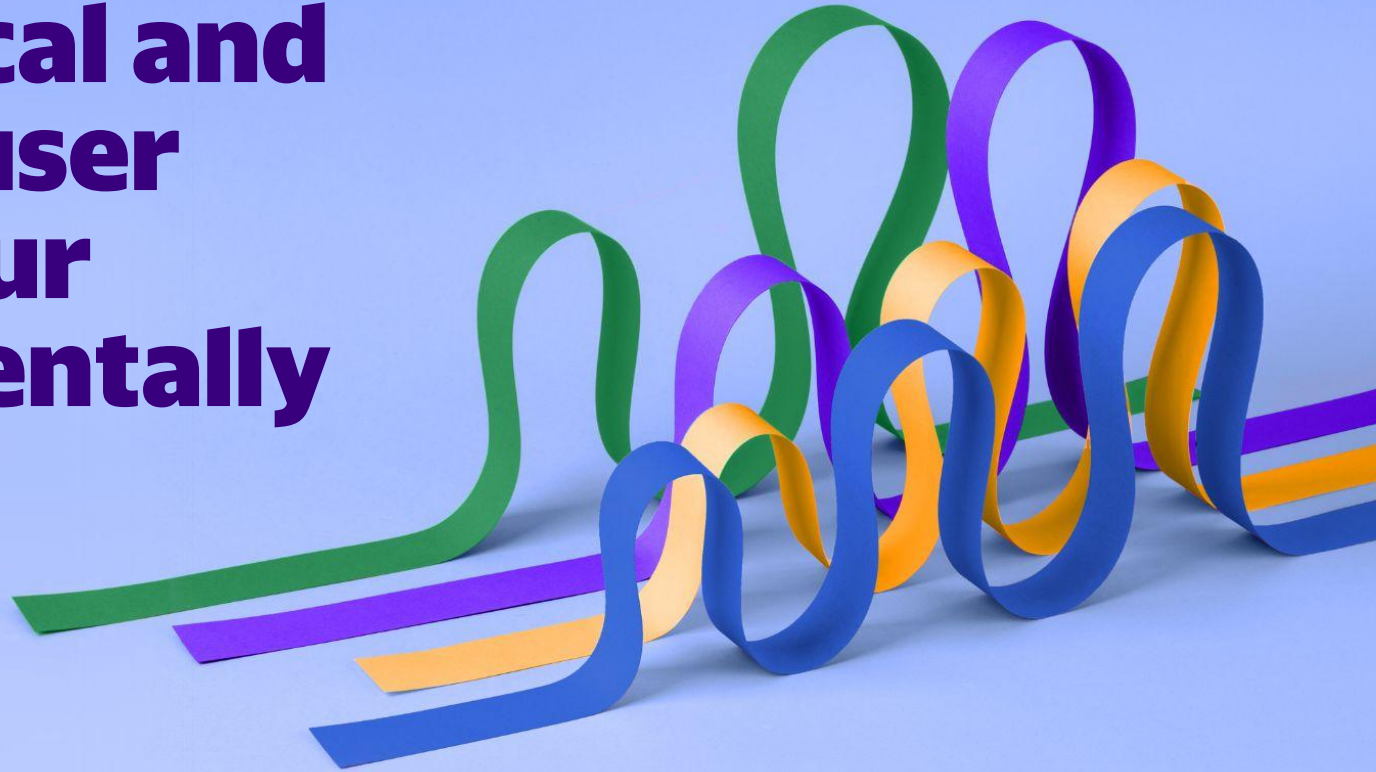
yahoo!

Case Studies

Incrementality of tactical practices

yahoo!

A/B test + Incrementality test: Multi-cell testing

Contro A	Test A	Contro B	Test B

Ghost Impressions | Ad Impressions | Ghost Impressions | Ad Impressions

A benefit from double blind designs, an A/B split test fits easily to run a multi-cell test

User split is required to test A and B tactical practices

yahoo!

Tactical Practices on user Signups: Key Findings

4.5%
Tactic A cell reached
Statistically Significant
Converter rate lift

2.2%
Tactic B cell reached
Statistical Significant
Converter rate lift

**Tactic A
outperforms
Tactic B**

$XX
**Cost per Incremental
Converter Tactic A**

VS

$XX
**Cost per Incremental
Converter Tactic B**

yahoo!

Experiment eligibility: first impression. All converters after first impression (any touch). 90% confidence intervals

Test your best 2 or 3 tactics maximum

Statistical mass is generally a constraint

yahoo!

What if we can't randomize user ids?

We'll review identity challenges and trends in the next and last part of the tutorial….

yahoo!