

# Online Advertising Incrementality Testing: Practical Lessons And Emerging Challenges

Joel Barajas  
Yahoo Research, Verizon Media  
Sunnyvale, CA, USA  
joel.barajas@verizonmedia.com

Narayan Bhamidipati  
Yahoo Research, Verizon Media  
Sunnyvale, CA, USA  
narayanb@verizonmedia.com

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

## ABSTRACT

Online advertising has historically been approached as an ad-to-user matching problem within sophisticated optimization algorithms. As the research and ad-tech industries have progressed, advertisers have increasingly emphasized the causal effect estimation of their ads (incrementality) using controlled experiments (A/B testing). With low lift effects and sparse conversion, the development of incrementality testing platforms at scale suggests tremendous engineering challenges in measurement precision. Similarly, the correct interpretation of results addressing a business goal requires significant data science and experimentation research expertise.

We propose a practical tutorial in the incrementality testing landscape, including:

- The business need
- Literature solutions and industry practices
- Designs in the development of testing platforms
- The testing cycle, case studies, and recommendations

We provide first-hand lessons based on the development of such a platform in a major combined DSP and ad network, and after running several tests for up to two months each over recent years.

## CCS CONCEPTS

• **Applied computing** → **Marketing; Economics; • Information systems** → **Computational advertising; Display advertising; • General and reference** → **Experimentation; • Mathematics of computing** → **Probability and statistics.**

## KEYWORDS

Marketing Incrementality, Controlled Randomized Experiments, A/B Testing, Computational Advertising, Causal Inference

### ACM Reference Format:

Joel Barajas, Narayan Bhamidipati, and James G. Shanahan. 2021. Online Advertising Incrementality Testing: Practical Lessons And Emerging Challenges. In *Proceedings of the 30th ACM CIKM Conference on Information and Knowledge Management (CIKM '21), November 1–5, 2021, Gold Coast, Queensland, Australia*. ACM, New York, NY, USA, 4 pages. <https://doi.org/10.1145/1122445.1122456>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

*CIKM '21, November 1–5, 2021, Gold Coast, Queensland, Australia*

© 2021 Association for Computing Machinery.  
ACM ISBN 978-1-4503-XXXX-X/18/06...\$15.00  
<https://doi.org/10.1145/1122445.1122456>

## 1 OFFERING AND TOPICS

Length: half-day, i.e., 3 hours plus breaks.

### 1.1 Intended Audience

Practitioners and researchers in the field of online advertising and marketing. Familiarity with basic statistics, hypothesis testing, confidence intervals is desirable. General knowledge of online advertising, familiarity with A/B testing is a plus.

Even though there are currently solutions to evaluate the advertising effectiveness with randomized experiments, many details and recommendations rarely appear in papers. This tutorial provides a 360-degree view of the topic, from engineering designs to experiment planning and business use cases.

The key benefits to participants include:

- Specific recommendations to the correct execution of A/B testing
- Online advertising testing engineering designs and econometric evaluation approaches
- Marketing use cases for online advertising incrementality testing

We estimate that a significant amount of the CIKM community would benefit from either or all of these topics.

### 1.2 Learning Outcomes

Topics Reviewed:

- Experiment design and A/B testing
- Causal inference
- Ad tech architectures
- Business use cases in advertising and marketing analytics
- Practical recommendations to successful significant incrementality experiments

Participants will:

- (1) Identify and formulate key approaches to measuring the effectiveness of online advertising.
- (2) Execute relevant statistics for hypothesis testing, power analysis in experiment planning and simulate experiment scenarios.
- (3) Be able to define key ingredients of an operational incrementality testing platform and their trade-offs.
- (4) Understand the business need for incrementality testing.
- (5) Identify the necessary conditions to increase the likelihood of successful test given minimum detectable lift, conversion type, test duration among others.

Take-away material:

**Cheatsheet of statistical tools** including python notebooks and spreadsheets for statistical power analysis

**Summary of recommendations handout** document for operationalization addressing the challenges and trade-offs of operationalizing testing

**Categorization of business use cases handout** document addressing typical needs and focus of incrementality testing

## 2 TUTORIAL OUTLINE

**Part 1** The basics: context and challenges [7, 10, 13, 17]

- The problem
  - Online Advertising spend trends between performance and brand
  - Big picture problem: quarterly/yearly budget allocation
  - Budget allocation practices based on financial models
  - The need for testing combined with industry attribution practices
- How channel-level testing fits within other forms of testing
  - Real-time decision making in targeting engines
  - Tactic testing: A/B testing with last-touch attribution
  - Multi-cell testing A/B testing + Incrementality testing
  - CMO decision making at the end of the quarter/semester/year
- Business Use cases
  - Advertiser joining new partners
  - Testing to calibrate and rebase financial models
  - Media Mix Models calibration
  - Last-touch attribution multipliers
  - The Marketing component: Growth marketing vs CRM marketing

**Part 2** Incrementality Testing: concepts, solutions and literature [2, 3, 12, 19]

- Literature and Industry practices
  - Placebo based testing: practice and issues
  - Intention to treat testing
  - Ghost ads testing proposal
- Estimation Frameworks
  - Econometric causality
  - Potential Outcomes Causal Framework
  - Pitfalls

**Part 3** From concept to production: platform building, challenges, case studies [3, 14]

- Building the experiment platform journey
- The identity graph and treatment groups
  - Cookie-based experiments
  - Device-based experiments
  - Logged-in users based experiments
  - Household-level experiments
- User holdout design within modern Ad tech serving systems
  - The hashing functions
  - The challenge with targeting and scoring algorithms
  - How to avoid targeting bias
  - The role of look-back windows in last-touch attribution engines
- Data Logging and Analysis

**Part 4** Deployment at Scale: test cycle and case studies

- Experiment execution cycle
  - Experiment Design and Planning
  - Intervention Execution
  - Experiment Tracking and Metrics
  - End of experiment readout
- Case Studies
  - Insurance quotes and comparison with post-click conversions
  - Online food ordering revenue: CRM versus New audiences
  - Online acquisition signup

**Part 5** Emerging trends: identity challenges, industry trends and solutions [1, 5, 6]

- Advertisers Testing without Ad Network holdouts
  - Spend as experiment intervention
  - Methodologies: Time series based testing
- Geo-testing
  - Geo units specification
  - Geo unit treatment assignment
  - The power of A/A tests in the experiment design
- Emerging challenges with user ids
  - User group randomization as a proxy of user level splits
  - Identity fragmentation challenges
  - Test precision challenges

## 3 LIST OF REFERENCES BY TOPIC

### 3.1 The need for Incrementality Testing Solutions

- *A comparison of approaches to advertising measurement: Evidence from big field experiments at Facebook* by Gordon et al. (2019) [10]
- *Do display ads influence search? Attribution and dynamics in online advertising* by Kireyev et al. (2016) [13].
- *Attributing conversions in a multichannel online marketing environment: An empirical model and a field experiment* by Li and Kannan (2014) [17].
- *Evaluating online ad campaigns in a pipeline: causal models at scale* by Chan et al. (2010) [7].

### 3.2 Incrementality Testing Solutions

- *Incrementality Testing in Programmatic Advertising: Enhanced Precision with Double-Blind Designs* by Barajas and Bhamidipati (2021) [3]
- *Ghost ads: Improving the economics of measuring online ad effectiveness* by Johnson et al. (2017) [12].
- *Experimental designs and estimation for online display advertising attribution in marketplaces* by Barajas et al. (2016) [2].
- *Here, there, and everywhere: correlated online behaviors can lead to overestimates of the effects of advertising* by Lewis et al. (2011) [16].

### 3.3 Causal Inference

- *Causal inference using potential outcomes: Design, modeling, decisions* by Rubin (2005) [19]
- *Principal stratification in causal inference* by Frangakis and Rubin (2002) [9].
- *Bayesian inference for causal effects in randomized experiments with noncompliance* by Imbens and Rubin (1997) [11].

### 3.4 Operationalization and Practical Recommendations

- *Incrementality Testing in Programmatic Advertising: Enhanced Precision with Double-Blind Designs* by Barajas and Bhamidipati (2021) [3]
- *Trustworthy online controlled experiments: A practical guide to a/b testing* by Kohavi et al. (2020) [14].
- *The unfavorable economics of measuring the returns to advertising* by Lewis et al. (2015) [15].

### 3.5 Geo-testing and Synthetic Control and Identity Challenges

- *Advertising Incrementality Measurement using Controlled Geo-Experiments: The Universal App Campaign Case Study* by Barajas et al. (2020) [5]
- *Consumer heterogeneity and paid search effectiveness: A large-scale field experiment* by Blake et al. (2015) [6].
- *Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program* by Abadie et al. (2010) [1].
- *The identity fragmentation bias* by Lin and Misra (2020) [18].

## 4 PREVIOUS RELATED OFFERINGS

*Econometric Analysis and Digital Marketing: How to Measure the Effectiveness*[8]. This tutorial was given at the time ad effectiveness testing was gaining popularity in the industry. Our current tutorial shows the eight-year evolution of the methods, real-world practices, operationalization practices, and emerging challenges within user privacy implications in experimentation.

*Online Advertising Incrementality Testing And Experimentation: Industry Practical Lessons*[4]. Our current tutorial proposal is an evolution of this tutorial at SIG KDD 2021<sup>1</sup>. This tutorial is similar to the first four parts of the outline, including deliverables and objectives. In contrast, in our current submission, we elaborate more on the user privacy implications in online experimentation and incrementality testing. We aim to motivate the research community to focus on solutions under these emerging constraints.

## 5 AUTHORS BIOGRAPHY

Joel Barajas, Sr Research Scientist, has over 11 years of experience in the online advertising industry with research contributions at the intersection of Ad tech, Marketing Science, and Experimentation. He has experience with Ad load personalization and experimentation in a publisher marketplace. Within Marketing Data Science, he has supported regular budget allocation and Media Mix Models

in multi-channel advertising. With a PhD dissertation focussed on ad incrementality testing, his published work has appeared in top outlets including INFORMS Marketing Science Journal, ACM CIKM, ACM WWW, SIAM SDM. He led the science development and marketing analytics of the incrementality testing platform in a multidisciplinary team. He currently oversees most incrementality tests in Verizon Media ad network (previously yahoo!) and DSP (previously AOL advertising.com). Joel also leads the science development in CTV and linear TV measurement modeling. He holds a B.S. (with honors) in Electrical and Electronics Engineering from the Tecnológico de Monterrey, and a PhD in Electrical Engineering (with emphasis on statistics) from UC Santa Cruz.

Narayan Bhamidipati, Sr Director of Research, has over 14 years of experience in Computational Advertising and Machine Learning. He currently leads a team of researchers focused on providing state-of-the-art ad targeting solutions to help ads be more effective and relevant. This includes creating various contextual targeting products to reduce the company's reliance on user profiles and help improve monetization in a more privacy aware world. Alongside that, Narayan ensures that the user profile based ad targeting products continue to improve despite the decline of tracking data. In addition, Narayan is keen on developing the most accurate ad effectiveness measurement platform which would help the company attract more revenue by proving the true value of the ad spend on our platforms. He holds B.Stat(Hons), M.Stat and PhD(CS) degrees, all from the Indian Statistical Institute, Kolkata.

Dr. James G. Shanahan has spent the past 30 years developing and researching cutting-edge artificial intelligence systems, splitting his time between industry and academia. For the academic year 2019-2020, Jimi held the position of Rowe Professor of Data Science at Bryant University, Rhode Island. He has (co) founded several companies that leverage AI/machine learning/deep learning/computer vision in verticals such as digital advertising, web search, local search, and smart cameras. Previously he has held appointments at AT&T (Executive Director of Research), NativeX (SVP of data science), Xerox Research (staff research scientist), and Mitsubishi. He is on the board of Anvia, and he also advises several high-tech startups including Aylien, ChartBoost, DigitalBank, LucidWorks, and others. Dr. Shanahan received his PhD in engineering mathematics and computer vision from the University of Bristol, U. K. Jimi has been involved with KDD since 2004 as an author, as a tutorial presenter, and as a workshop co-chair; he has actively been involved as a PC/SPC member over the years also.

## REFERENCES

- [1] Alberto Abadie, Alexis Diamond, and Jens Hainmueller. 2010. Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American statistical Association* 105, 490 (2010), 493–505.
- [2] Joel Barajas, Ram Akella, Marius Holtan, and Aaron Flores. 2016. Experimental designs and estimation for online display advertising attribution in marketplaces. *Marketing Science* 35, 3 (2016), 465–483.
- [3] Joel Barajas and Narayan Bhamidipati. 2021. Incrementality Testing in Programmatic Advertising: Enhanced Precision with Double-Blind Designs. In *Proceedings of the Web Conference 2021* (Ljubljana, Slovenia) (WWW '21). Association for Computing Machinery, New York, NY, USA, 2818–2827. <https://doi.org/10.1145/3442381.3450106>
- [4] Joel Barajas, Narayan Bhamidipati, and James G. Shanahan. 2021. Online Advertising Incrementality Testing And Experimentation: Industry Practical Lessons (KDD '21). Association for Computing Machinery, New York, NY, USA, 2 pages. <https://doi.org/10.1145/3447548.3470819>

<sup>1</sup>The online version of this tutorial will be available at <https://joel-barajas.github.io/kdd2021-incrementality-testing/>

- [5] Joel Barajas, Tom Zidar, and Mert Bay. 2020. Advertising Incrementality Measurement using Controlled Geo-Experiments: The Universal App Campaign Case Study. (2020).
- [6] Thomas Blake, Chris Nosko, and Steven Tadelis. 2015. Consumer heterogeneity and paid search effectiveness: A large-scale field experiment. *Econometrica* 83, 1 (2015), 155–174.
- [7] David Chan, Rong Ge, Ori Gershony, Tim Hesterberg, and Diane Lambert. 2010. Evaluating Online Ad Campaigns in a Pipeline: Causal Models at Scale. In *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (Washington, DC, USA) (KDD '10). ACM, New York, NY, USA, 7–16. <https://doi.org/10.1145/1835804.1835809>
- [8] Ayman Farahat and James Shanahan. 2013. Econometric analysis and digital marketing: how to measure the effectiveness of an ad. In *Proceedings of the sixth ACM international conference on Web search and data mining*. 785–785.
- [9] C. E. Frangakis and D. B. Rubin. 2002. Principal stratification in causal inference. *Biometrics* 58, 1 (2002), 21–29. <https://doi.org/10.1111/j.0006-341X.2002.00021.x>
- [10] Brett R Gordon, Florian Zettelmeyer, Neha Bhargava, and Dan Chapsky. 2019. A comparison of approaches to advertising measurement: Evidence from big field experiments at Facebook. *Marketing Science* 38, 2 (2019), 193–225.
- [11] Guido W. Imbens and Donald B. Rubin. 1997. Bayesian Inference for Causal Effects in Randomized Experiments with Noncompliance. *The Annals of Statistics* 25, 1 (1997), 305–327. <http://www.jstor.org/stable/2242722>
- [12] Garrett A. Johnson, Randall A. Lewis, and Elmar I. Nubbemeyer. 2017. Ghost Ads: Improving the Economics of Measuring Online Ad Effectiveness. *Journal of Marketing Research* 54, 6 (2017), 867–884. <https://doi.org/10.1509/jmr.15.0297> arXiv:<https://doi.org/10.1509/jmr.15.0297>
- [13] Pavel Kireyev, Koen Pauwels, and Sunil Gupta. 2016. Do display ads influence search? Attribution and dynamics in online advertising. *International Journal of Research in Marketing* 33, 3 (2016), 475–490.
- [14] Ron Kohavi, Diane Tang, and Ya Xu. 2020. *Trustworthy Online Controlled Experiments: A Practical Guide to A/B Testing*. Cambridge University Press.
- [15] Randall A. Lewis and Justin M. Rao. 2015. The Unfavorable Economics of Measuring the Returns to Advertising \*. *The Quarterly Journal of Economics* 130, 4 (2015), 1941–1973.
- [16] Randall A. Lewis, Justin M. Rao, and David H. Reiley. 2011. Here, There, and Everywhere: Correlated Online Behaviors Can Lead to Overestimates of the Effects of Advertising. In *Proceedings of the 20th International Conference on World Wide Web* (Hyderabad, India) (WWW '11). ACM, New York, NY, USA, 157–166. <https://doi.org/10.1145/1963405.1963431>
- [17] Hongshuang (Alice) Li and P.K. Kannan. 2014. Attributing Conversions in a Multichannel Online Marketing Environment: An Empirical Model and a Field Experiment. *Journal of Marketing Research* 51, 1 (2014), 40–56. <https://doi.org/10.1509/jmr.13.0050> arXiv:<https://doi.org/10.1509/jmr.13.0050>
- [18] Tesary Lin and Sanjog Misra. 2020. The Identity Fragmentation Bias. arXiv:2008.12849 [econ.EM]
- [19] Donald B Rubin. 2005. Causal Inference Using Potential Outcomes. *J. Amer. Statist. Assoc.* 100, 469 (2005), 322–331. <https://doi.org/10.1198/016214504000001880> arXiv:<https://doi.org/10.1198/016214504000001880>