International Journal of Business Quantitative Economics and Applied Management Research

Volume-6, Issue-8, 2020 ISSN No: 2349-5677

# IMPACT OF MATCH RATES ON COST BASIS METRICS IN PRIVACY-PRESERVING DIGITAL ADVERTISING

*Varun Chivukula*
*University of California, Berkeley*
*varunvenkatesh88@berkeley.edu*

*Abstract*

***This study explores the role of match rates in determining cost basis metrics for randomized controlled trials (RCTs) conducted through Privacy Enhancing Technologies (PETs), such as multi-party computation (MPC), Trusted Execution Environments (TEEs), and cleanroom-based protocols. These methods are used to measure the causal effectiveness of digital advertising campaigns while maintaining data privacy. We specifically analyze how variations in match rates influence incremental Return on Ad Spend (iROAS) and other cost basis metrics, emphasizing the challenges in deriving accurate estimates within privacy-preserving frameworks. Alternative estimators of match rates and their computability under PETs are discussed. Practical issues, including the inability to account for unmatched conversions, and potential proxies for mitigating these limitations, are examined [1][2].***

***IndexTerms—Privacy enhancing technologies, Multi-party computation , Causal inference, Digital Ad platforms, Auction based RTB's, Data encryption and retrieval, Randomized control trials***

## I. INTRODUCTION

Randomized controlled trials (RCTs) are widely regarded as the gold standard for evaluating the causal impact of digital advertising[3]. With the increasing emphasis on privacy-preserving advertising measurement, privacy-enhancing technologies (PETs) such as cleanrooms, MPC, and TEEs have emerged as vital tools for conducting such studies while adhering to stringent data privacy regulations, such as the General Data Protection Regulation (GDPR)[1] and the California Consumer Privacy Act (CCPA)[2].

A critical challenge in these privacy-centric environments is the inability to perfectly match all conversions to users in the test group. This issue, termed match rate bias, can distort cost-related metrics such as incremental Return on Ad Spend (iROAS). For example, a reported iROAS of 2x based on a 50% match rate could correspond to a true iROAS of 4x, but this correction is unachievable without accurate match rate data [4].

International Journal of Business Quantitative Economics and Applied Management Research
Volume-6, Issue-8, 2020 ISSN No: 2349-5677

This paper provides a structured framework for understanding match rates in PET environments and examines their implications for cost basis metrics. By categorizing relevant metrics and assessing their computational feasibility within these frameworks, we aim to provide actionable insights for advertisers navigating the trade-offs between privacy and measurement accuracy.

## II. PRIVACY-ENHANCING TECHNOLOGIES (PETS): CONTEXT AND CHALLENGES

Match rates are fundamental to PET-based advertising measurement as they represent the proportion of uploaded users or conversions that can be linked to the platform's user base [5]. Variability in match rates directly impacts lift metrics and cost basis metrics. For instance, unmatched conversions distort accuracy, and for cost basis metrics like iROAS, unmatched conversions result in an underestimation of incremental outcomes.

Bias in incremental outcomes remains a concern, as differential matchability between treatment and control groups can affect lift metrics, even when randomization is applied. Understanding and mitigating these biases is critical for ensuring accurate RCT outcomes [6].
Match rates are fundamental to PET-based advertising measurement as they represent the proportion of uploaded users or conversions that can be linked to the platform's user base. Variability in match rates has direct implications for both lift metrics and cost basis metrics as seen in Fig.1

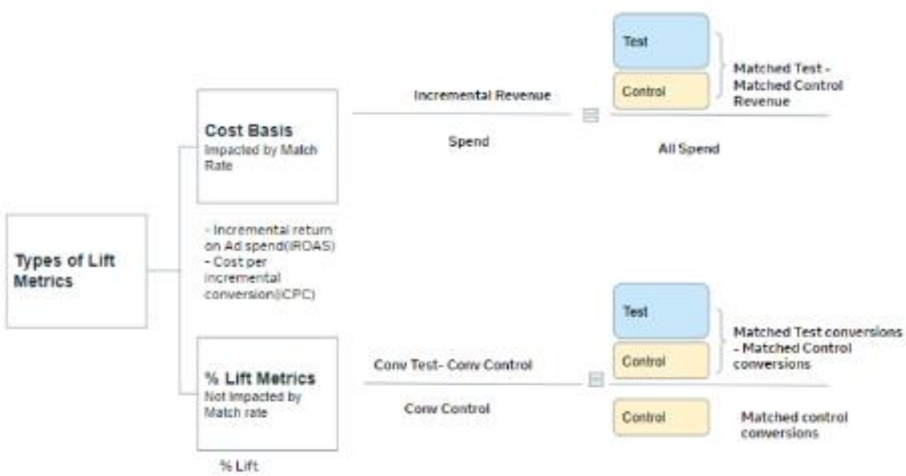


Fig 1.

- **Unmatched conversions distort accuracy:** For cost basis metrics like iROAS, unmatched conversions result in an underestimation of incremental outcomes.
    - **Example:** A reported iROAS of 2x at a 50% match rate could imply a true iROAS of 4x, assuming all unmatched conversions are incremental. This discrepancy underscores the importance of understanding match rates in interpreting results.

International Journal of Business Quantitative Economics and Applied Management Research
Volume-6, Issue-8, 2020 ISSN No: 2349-5677

- **Bias in incremental outcomes:** Lift metrics, which compare treatment and control groups, may remain unbiased under randomization but are still affected by differential matchability between groups.

## III. MATCH RATE METRICS IN PETS

**Match Rate (Test)** represents the theoretical match ability of conversions under ideal conditions. However, this metric is generally unknowable in PETs due to privacy constraints [7].

**Match Rate (MAU)** is the proportion of users in the uploaded dataset that can be matched to the platform's user base. It is directly computable using platforms with large user bases, such as Google and Meta, which cover over 90% of the US and Canadian adult population[8].

**Intersection Rate (Advertiser in Denominator)** measures the proportion of matched conversions mappable to the test group. This metric is challenging to extrapolate to broader cost basis metrics due to biases introduced by segment-specific conversion data[9].

In Fig 2. The Match Rate(Test) metric represents the theoretical matchability of conversions under ideal conditions. This is the required measure for estimating the true value of cost basis metrics.

$$\text{Match Rate (Test)} = \frac{T_m}{T_m + T_{un}}$$

where $T_{un}$ represents unmatched conversions.

**Computability under PETs:**
Match Rate (Test) is generally unknowable, as $T_{un}$ cannot be determined without violating privacyconstraints.

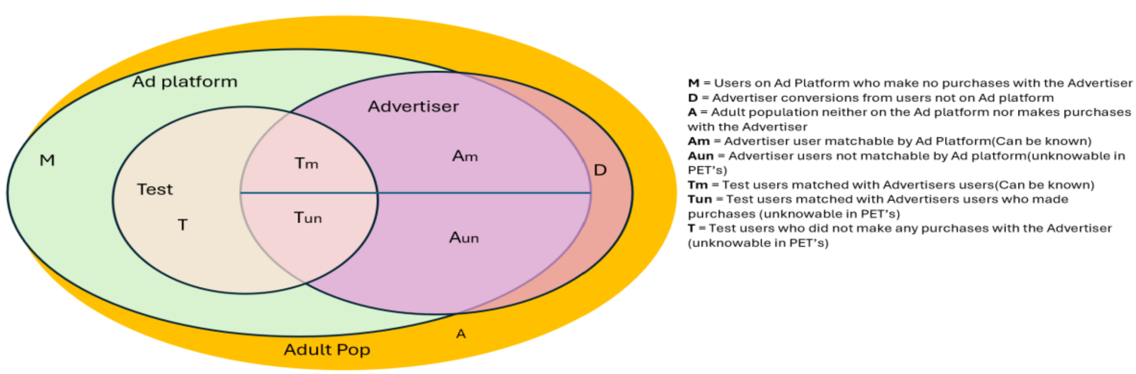


Fig 2.

### 3.2 Match Rate (MAU)

Match rate (MAU) is the proportion of users in the uploaded dataset that can be matched to the platform's user base:

$$\text{Match Rate (MAU)} = \frac{A_m}{A_m + A_{un}}$$

where $A_m$ represents matched users, and $A_{un}$ represents unmatched users.

**Computability under PETs:**

Match rate (MAU) is directly computable in PETs by using the platform's user base as the computation universe. For instance, if the Ad-platform's user base represents approximately 90%+ of the US and Canadian adult population[5], implying that unmatched users are a minority. This estimate of the match rate measure is an option with the least amount of bias when the Ad platform accounts for 90%+ of the adult population (eg. Google, Meta).However this will not be a feasible method for other Ad platforms.

### 3.3 Intersection Rate (Advertiser in Denominator)

The intersection rate is the proportion of matched conversions mappable to the test group:

$$\text{Intersection Rate} = \frac{T_m}{A_m + A_{un}}$$

where $T_m$ is the number of matched conversions for the test group.

**Computability under PETs:**

While $T_m$ is computable, $A_{un}$ requires assumptions about the unmatched population. This metric however cannot be used for extrapolation of cost basis metrics since intersection rate depends heavily on the targetable audience used within the test. However, by controlling the type of conversions Advertiser shares it might be possible to use this metric as a way of estimating Match Rate (test) albeit with potential bias. For eg.if the advertiser is able to track conversions which are tied to click tags from the Ad platform, they can estimate Match rate through this measure. This measure will however have bias since it focusses on a narrow segment of users and hence representativeness (in this eg. non clicking conversions from the advertiser). For Ad platforms where the Monthly active users are less than 90% this might be the only mechanism to estimate match rate.

## IV. DISCUSSION AND PRACTICAL RECOMMENDATIONS

Ensuring randomization mitigates risks of imbalances in treatment and control group matchability. However, lift metrics remain sensitive to differential matchability[6].

Unmatched conversions create significant barriers to accuracy in cost basis metrics. Metrics like Match Rate (MAU) serve as practical proxies for estimating match rates under certain conditions[8][9].

Advertisers should leverage causal estimates from user-level RCTs to interpret metrics conservatively, acknowledging potential underestimations of iROAS. Investing in advanced PETs and cryptographic hashing can improve matchability and measurement accuracy [4][10].

## V. CONCLUSIONS

AThis study highlights the significant impact of match rates on cost basis metrics in privacy-preserving advertising measurement. By categorizing match rate metrics and analyzing their implications for lift and cost-based outcomes, we offer practical guidance for navigating these challenges. Future developments in PETs and secure computation techniques will further enhance measurement accuracy while maintaining privacy.

## REFERENCES

1. European Union. (2016). General Data Protection Regulation (GDPR).
2. California Legislature. (2018). California Consumer Privacy Act (CCPA).
3. Abadi, M., & Andersen, D. G. (2017). "Privacy-preserving machine learning." Communications of the ACM.
4. Li, X., et al. (2020). "Trusted execution environments for secure computation." Journal of Privacy Technology.
5. U.S. Census Bureau. (2020). Population Estimates for U.S. and Canada.
6. Tadelis, S. (2016). "The economics of digital advertising."
7. Bonawitz, K., et al. (2019). "Practical secure aggregation for privacy-preserving machine learning." ACM SIGSAC.
8. Papadimitriou, C., et al. (2020). "Data privacy in large-scale platforms." Journal of Computing.
9. Hardt, M., et al. (2019). "Addressing biases in differential matchability." Advances in Privacy Technology.
10. Zhang, X., & Zhang, L. (2021). "Efficiency in privacy-preserving RCTs." Journal of Advertising Research.

International Journal of Business Quantitative Economics and Applied Management Research
Volume-6, Issue-8, 2020 ISSN No: 2349-5677

11. Chaudhuri, K., et al. (2022). "Future directions in PET-based advertising." IEEE Transactions on Big Data.