the Advertising Research Foundation

AUDIENCE xSCIENCE 2024

A Clean Room Incrementality Experiment – An Indeed Case Study

Isaac Dinner

Indeed

Joe Zucker

Indeed

#ARFAxS

Data Clean Room Experimentation: Assessing Incrementality

Joe Zucker
Marketing Analytics Lab

Isaac Dinner
Marketing Analytics Lab

indeed

Outline

- Indeed and Our Goals
- Bias in Online Experiments
- Managing Bias with Experimental Design
- Conclusion: Comparing Tests and Results

What is Indeed?

Indeed is the world's number one job site and leading matching and hiring platform. We strive to make a positive impact on society by connecting people to better work to create better lives.

We must regularly engage with new employers *and* jobseekers to balance the market

Employers → ← ***Jobseekers***

What is Indeed?

Indeed is the world’s number one job site and leading matching and hiring platform. We strive to make a positive impact on society by connecting people to better work to create better lives.

We must regularly engage with new employers *and* jobseekers to balance the market

What are we trying to achieve?

1. Measurement of the *incremental* impact of advertising on KPIs.

2. High quality continuous reporting with actionable analytics and insights.

Major Considerations for Experimentation at Indeed

- **New customer focus**

  For current customers, retargeting would be straightforward

- **Cookie-free**

  Deprecation in process

- **Privacy forward**

  Doesn't allow us to see individual customer data

- **Ability to verify advertiser data**

  Allows for "checks" on delivery and efficacy

  Rules out most "walled gardens"

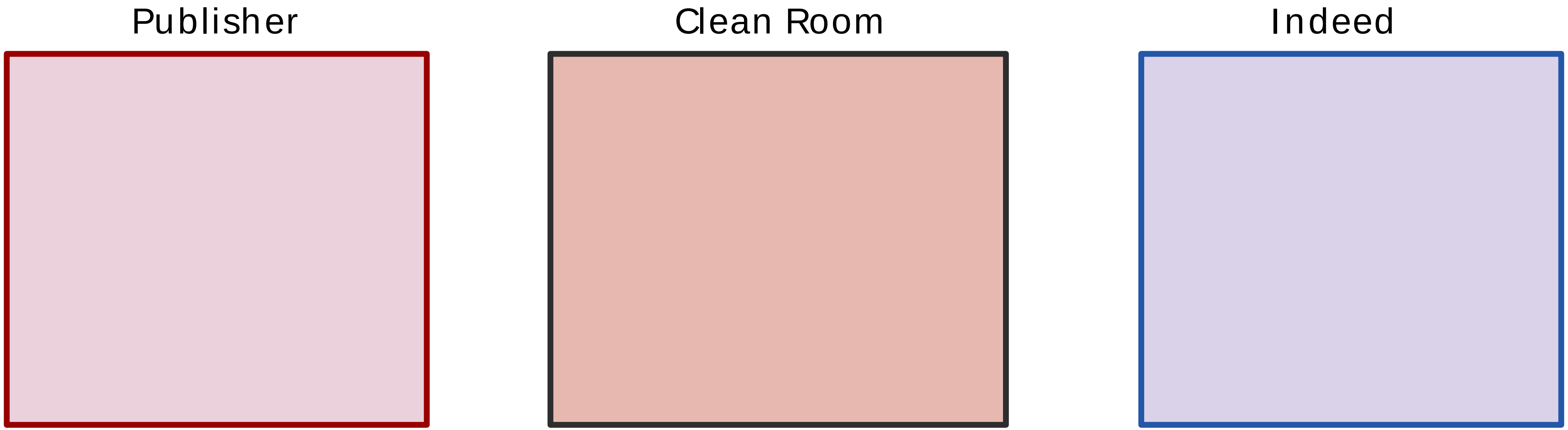
Experimentation through Data Clean Rooms

Fragmented Landscape: Indeed uses multiple vendors to measure incrementality, but the process remains the same.

Experimentation through Data Clean Rooms

Fragmented Landscape: Indeed uses multiple vendors to measure incrementality, but the process remains the same.

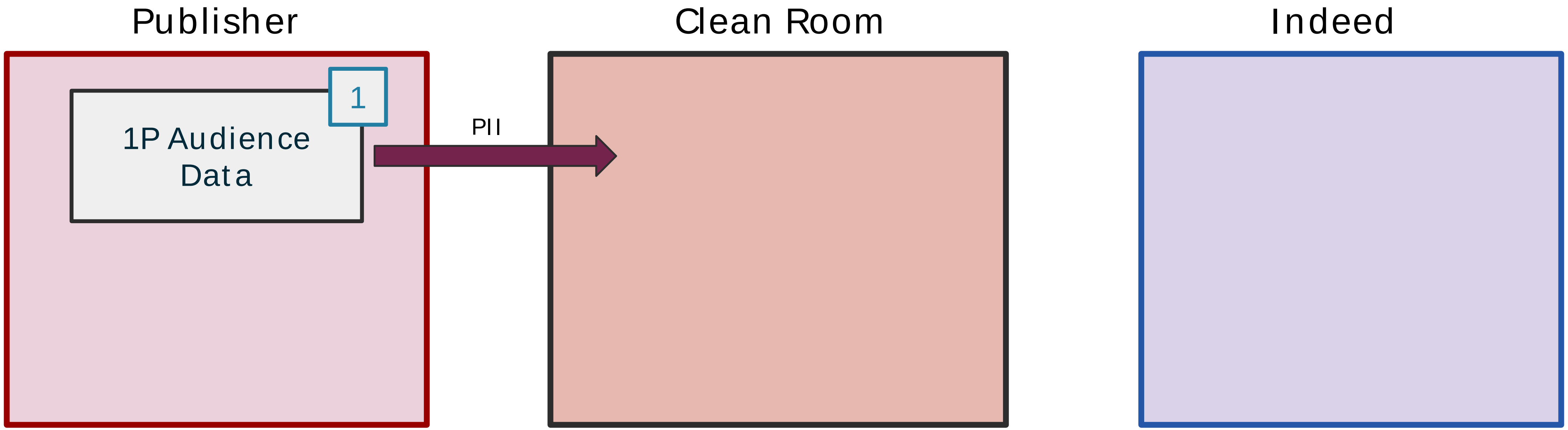
snowflake samooha amazon ads LiveRamp Habu



Experimentation through Data Clean Rooms

Fragmented Landscape: Indeed uses multiple vendors to measure incrementality, but the process remains the same.

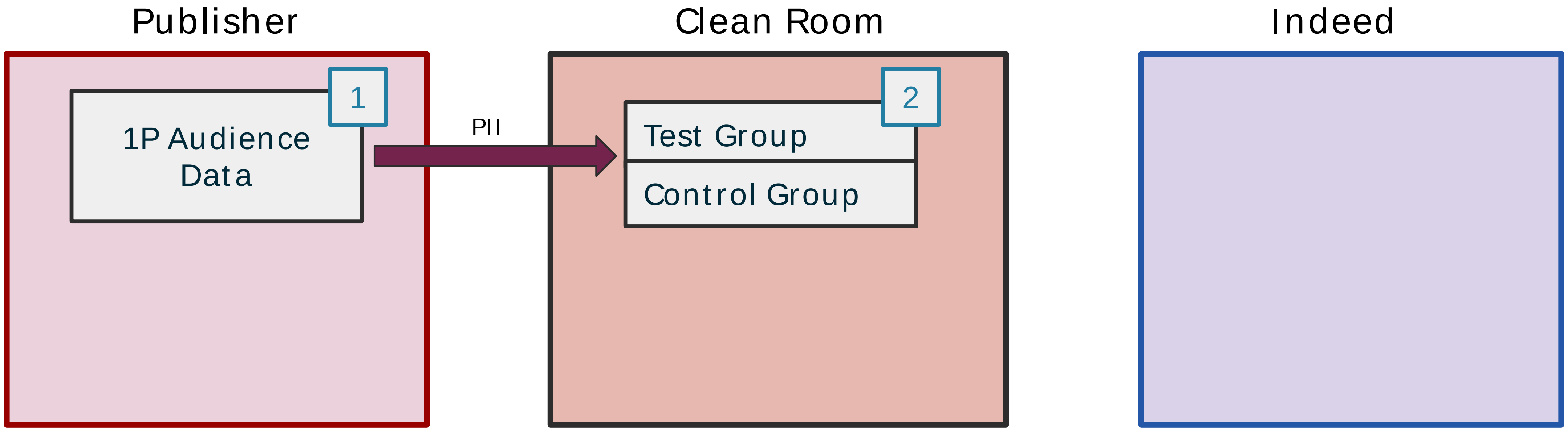
snowflake samooha amazon ads LiveRamp Habu



Experimentation through Data Clean Rooms

Fragmented Landscape: Indeed uses multiple vendors to measure incrementality, but the process remains the same.

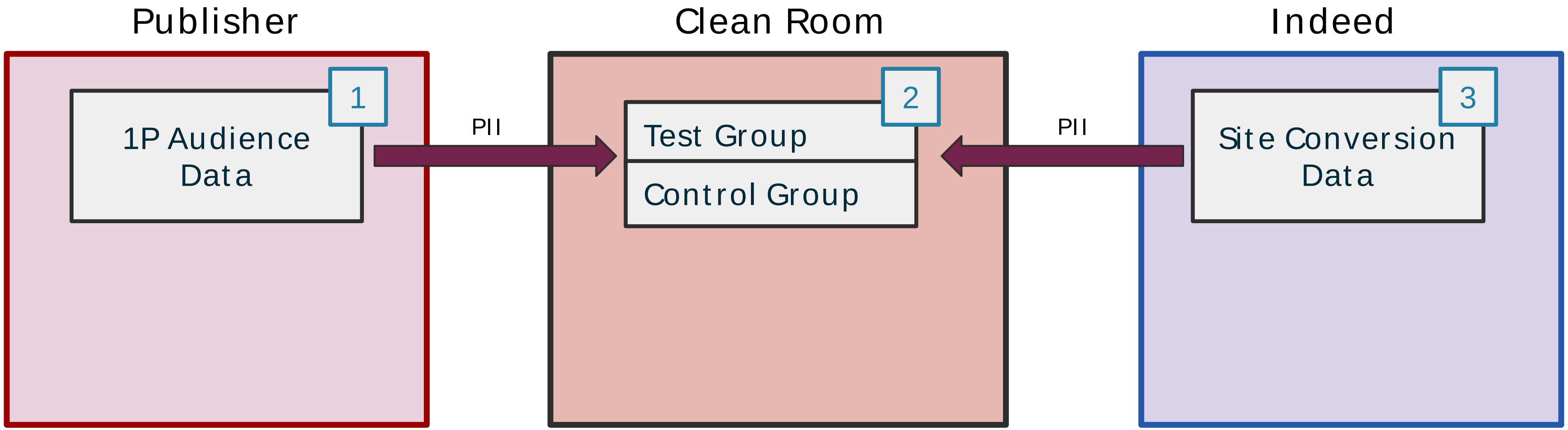
snowflake samooha amazon ads LiveRamp Habu



Experimentation through Data Clean Rooms

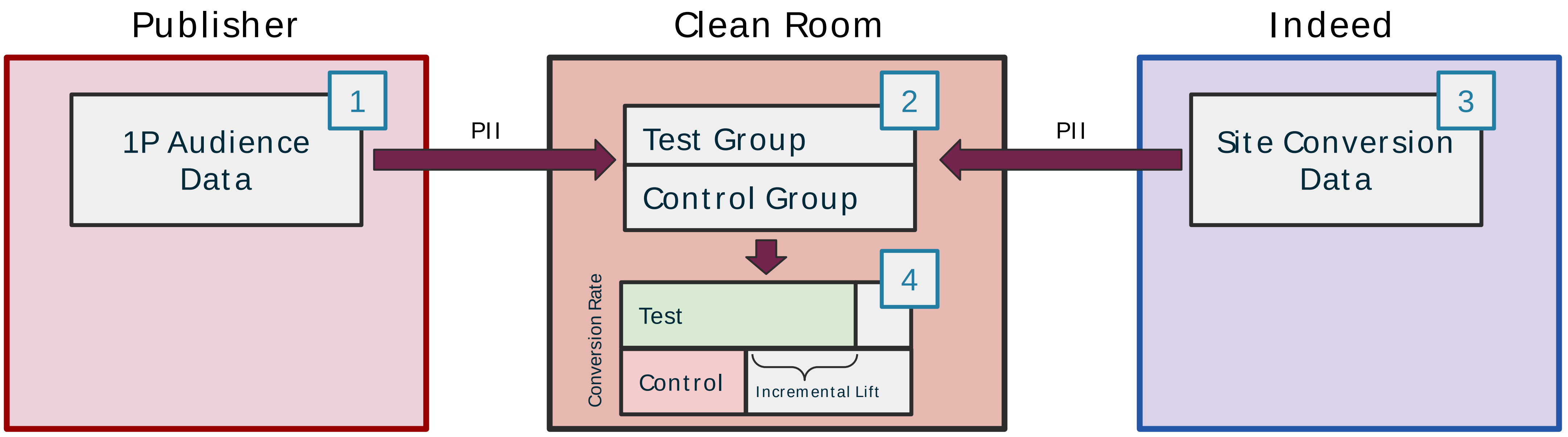
Fragmented Landscape: Indeed uses multiple vendors to measure incrementality, but the process remains the same.

snowflake samooha amazon ads LiveRamp Habu



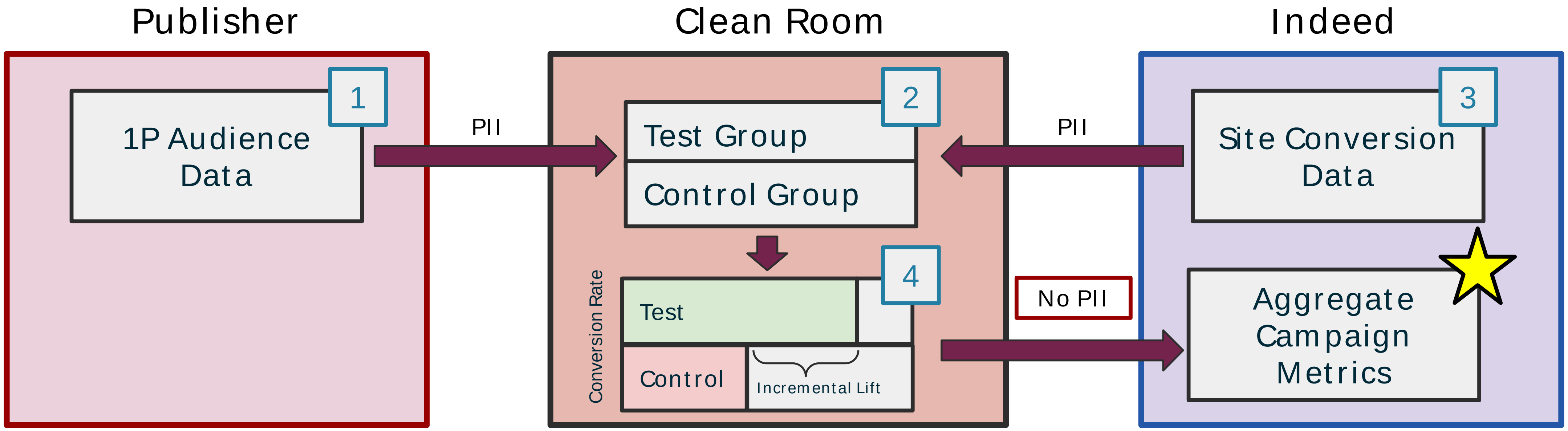
Experimentation through Data Clean Rooms

Fragmented Landscape: Indeed uses multiple vendors to measure incrementality, but the process remains the same.



Experimentation through Data Clean Rooms

Fragmented Landscape: Indeed uses multiple vendors to measure incrementality, but the process remains the same.



Clean Room Experimentation in Practice

Bias in Online Experimentation

Clean experiments in an online marketplace is difficult due to a complex marketplace.

- **Activity Bias**

  Test users may be more active on the publisher's platform than control users

  Experiment may measure "Activity", rather than ad effects

- **Ad Server Bias**

  Ad servers optimize to specific behaviour (i.e. clicks) and may serve PSA/House ads to a different type of user than the test group

- **Base Rate Bias**

  Baseline conversion rates of different methods (Ghost Ads, Suppression, etc.) may be dissimilar, causing incorrect lift estimates against a test population in unpredictable ways

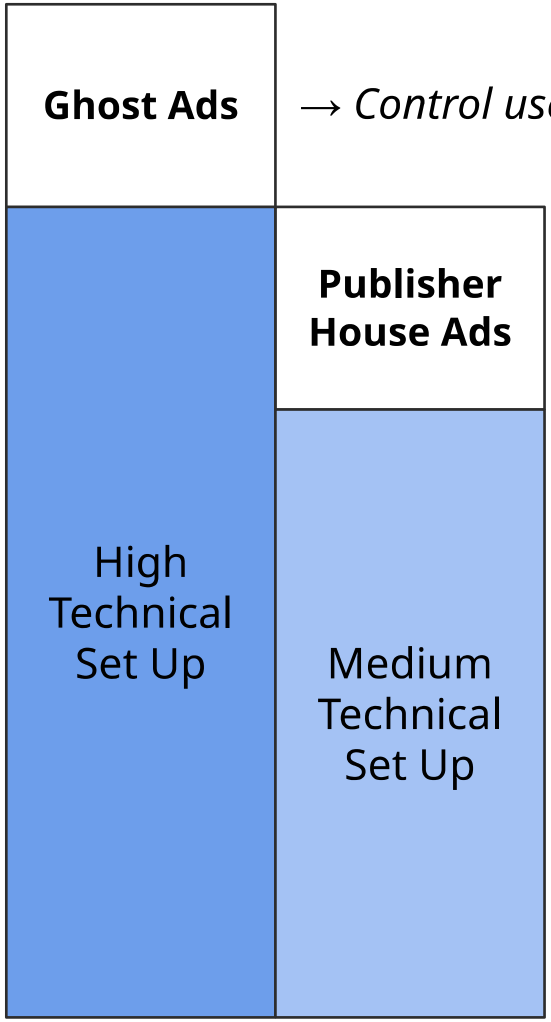
Experimental Design: Five Variations

Creating of the control group can be achieved in multiple ways

Ghost Ads	→ *Control users are served ads normally, but log when test ads would have won the auction*
High Technical Set Up	

Experimental Design: Five Variations

Creating of the control group can be achieved in multiple ways

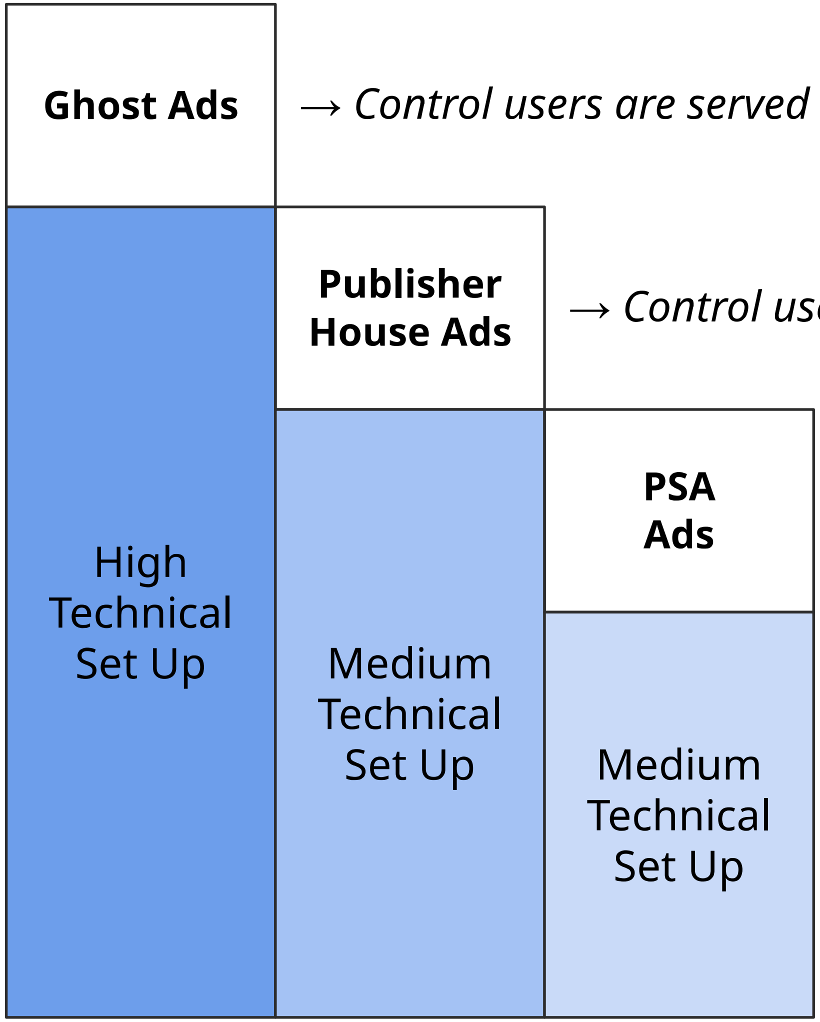


→ *Control users are served ads normally, but log when test ads would have won the auction*

→ *Control users are identified by serving a house ad for the publisher’s platform*

Experimental Design: Five Variations

Creating of the control group can be achieved in multiple ways



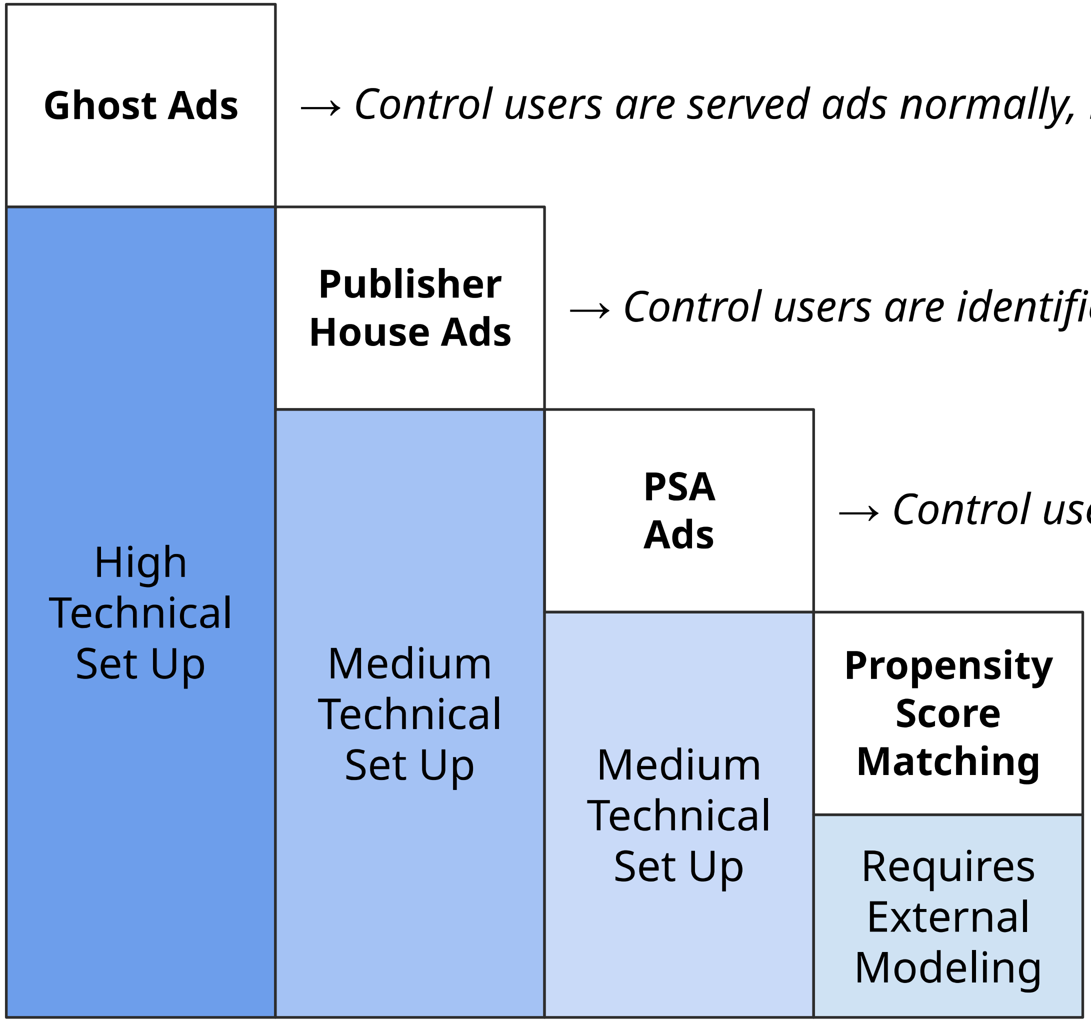
→ *Control users are served ads normally, but log when test ads would have won the auction*

→ *Control users are identified by serving a house ad for the publisher’s platform*

→ *Control users are served a PSA ad paid for my the advertiser*

Experimental Design: Five Variations

Creating of the control group can be achieved in multiple ways



→ *Control users are served ads normally, but log when test ads would have won the auction*

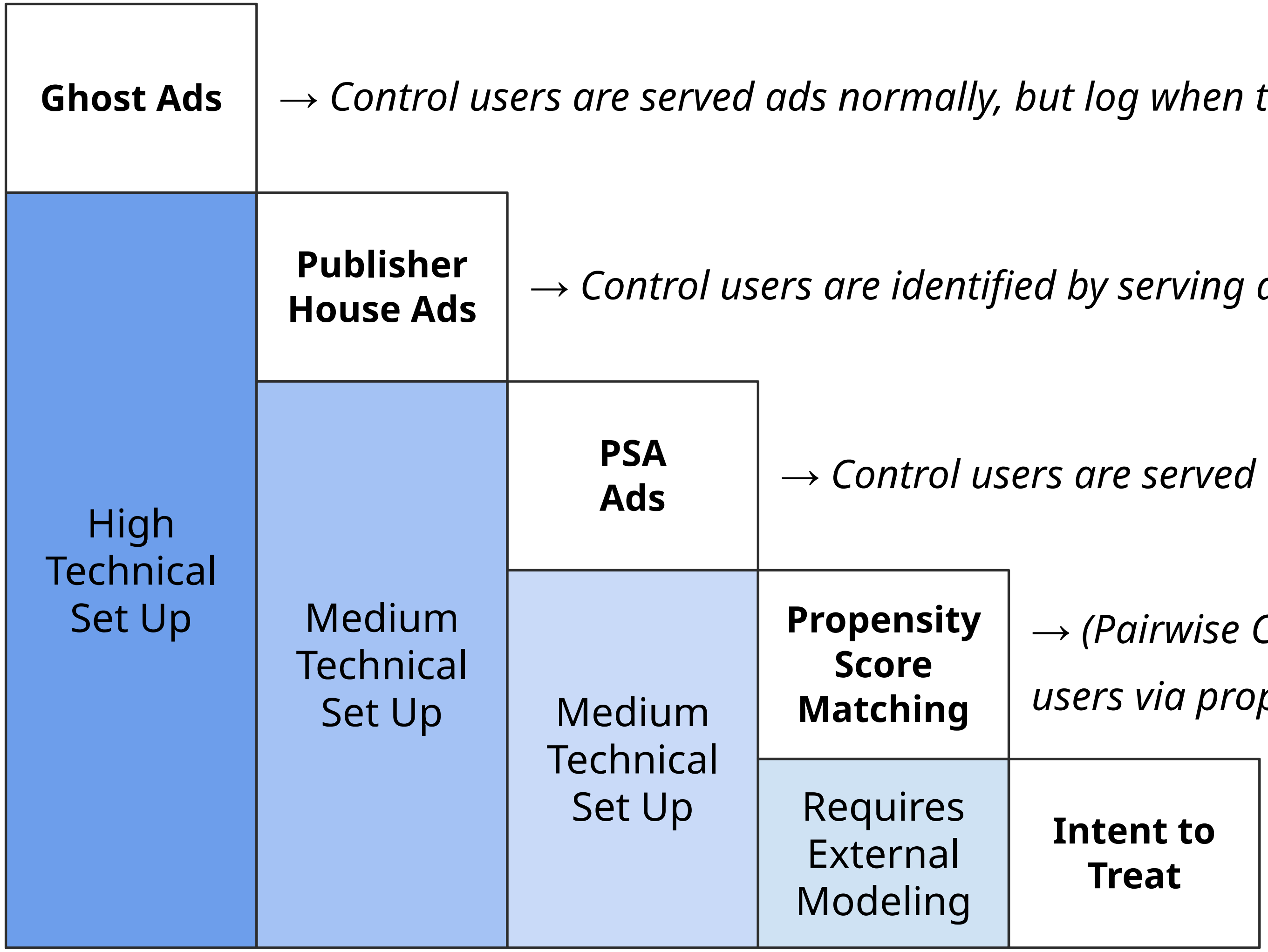
→ *Control users are identified by serving a house ad for the publisher’s platform*

→ *Control users are served a PSA ad paid for my the advertiser*

→ *(Pairwise Comparison) Control users matched to test users via propensity score modeling*

Experimental Design: Five Variations

Creating of the control group can be achieved in multiple ways



→ *Control users are served ads normally, but log when test ads would have won the auction*

→ *Control users are identified by serving a house ad for the publisher’s platform*

→ *Control users are served a PSA ad paid for my the advertiser*

→ *(Pairwise Comparison) Control users matched to test users via propensity score modeling*

→ *(Suppression) Control users are identified and suppressed from the test campaign*

Interpreting Experiment Results

Comparing the Options: There is No Free Lunch

Experiment Design	Activity Bias	Ad/Context Bias	Base Rate Bias	Pros	Cons
Ghost Ads	✅	✅	✅	Best mimics real world for control users No cost to advertiser	Few publishing partners have this capability Low control over the analysis

Comparing the Options: There is No Free Lunch

Experiment Design	Activity Bias	Ad/Context Bias	Base Rate Bias	Pros	Cons
Ghost Ads	✅	✅	✅	• Best mimics real world for control users • No cost to advertiser	• Few publishing partners have this capability • Low control over the analysis
Publisher House Ads	✅	❎	✅	• Partially accounts for activity bias	• Requires publisher support • Unknown behaviour of control users receiving house ads

Comparing the Options: There is No Free Lunch

Experiment Design	Activity Bias	Ad/Context Bias	Base Rate Bias	Pros	Cons
Ghost Ads	✅	✅	✅	• Best mimics real world for control users • No cost to advertiser	• Few publishing partners have this capability • Low control over the analysis
Publisher House Ads	✅	❎	✅	• Partially accounts for activity bias	• Requires publisher support • Unknown behaviour of control users receiving house ads
PSA Ads	✅	❎	✅	• Partially accounts for activity bias	• Additional cost to advertiser • Unknown behaviour of control users receiving PSA ads

Comparing the Options: There is No Free Lunch

Experiment Design	Activity Bias	Ad/Context Bias	Base Rate Bias	Pros	Cons
Ghost Ads	✅	✅	✅	• Best mimics real world for control users • No cost to advertiser	• Few publishing partners have this capability • Low control over the analysis
Publisher House Ads	✅	❎	✅	• Partially accounts for activity bias	• Requires publisher support • Unknown behaviour of control users receiving house ads
PSA Ads	✅	❎	✅	• Partially accounts for activity bias	• Additional cost to advertiser • Unknown behaviour of control users receiving PSA ads
Propensity Score Matching	tbd	❎	✅	• Potentially strips away noise in unexposed users	• Requires additional data to model high quality matches

Comparing the Options: There is No Free Lunch

Experiment Design	Activity Bias	Ad/Context Bias	Base Rate Bias	Pros	Cons
Ghost Ads	✅	✅	✅	• Best mimics real world for control users • No cost to advertiser	• Few publishing partners have this capability • Low control over the analysis
Publisher House Ads	✅	☒	✅	• Partially accounts for activity bias	• Requires publisher support • Unknown behaviour of control users receiving house ads
PSA Ads	✅	☒	✅	• Partially accounts for activity bias	• Additional cost to advertiser • Unknown behaviour of control users receiving PSA ads
Propensity Score Matching	tbd	☒	✅	• Potentially strips away noise in unexposed users	• Requires additional data to model high quality matches
Intent to Treat	tbd	✅	tbd	• Clean analysis IF there is access to unexposed test population data	• Requires publisher support • Lower statistical power due to additional “noise”

Challenges with Interpreting Experiment Results

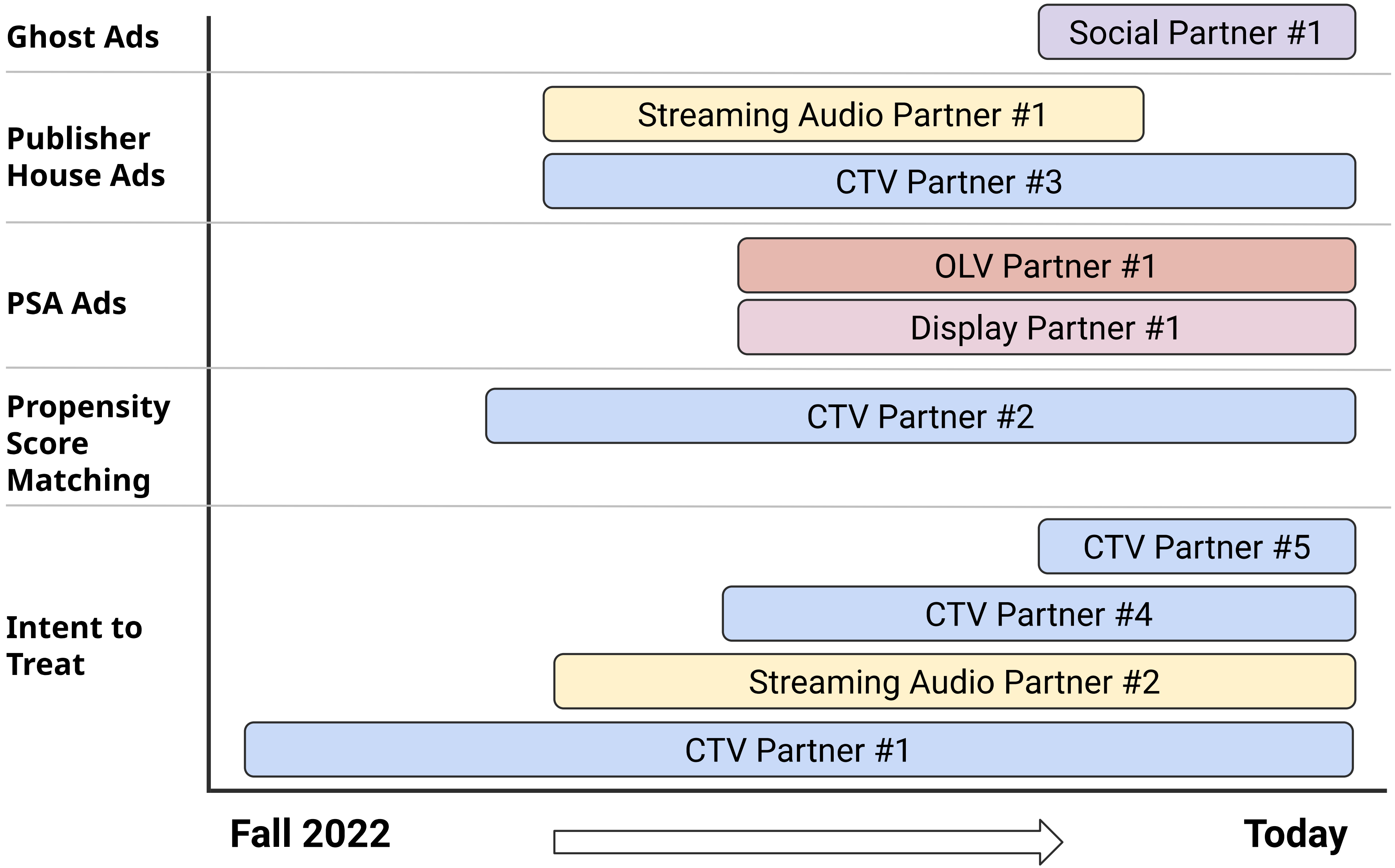
Low Match/Conversion Rates

1. **Narrow definition of a “match”**
2. **Match rates can be low and variable**
    a. Is the advertising ineffective or is there significant loss in matching?
3. **Resolving personal and corporate email accounts is a challenge**

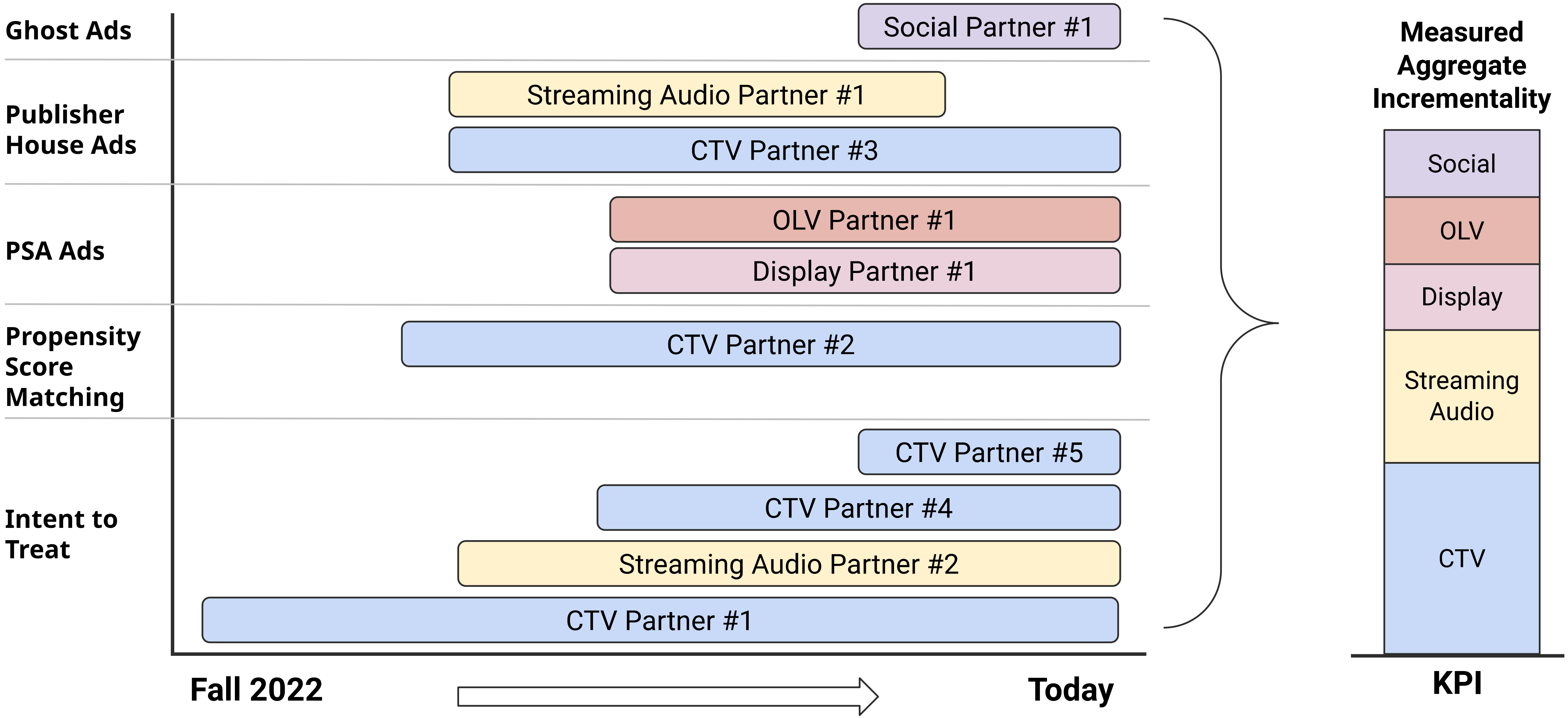
Comparing Results

1. **Household identity graphs (one to many) vs individual identity graphs (one to one)**
    a. Relative lift is comparable, but attributed conversions can be very different.
2. **How comparable are experiments with different control group construction?**

Clean Room Experimentation: A Success Story



Clean Room Experimentation: A Success Story



Q&A