Sundar Dorai-Raj
Senior Quantitative Analyst
Google

Dan Zigmond
Engineering Manager
Google
• YouTube launched in May 2005

• Grown to the world’s most popular online video community
  – 3 billion watches every day
  – 48 hours of video uploaded every minute
  – 2 billion monetized views every week
Problem

• Deriving causation from passive data is challenging
  – Observational studies are subject to selection bias
  – Segmenting groups of users for statistical comparisons is difficult and error prone

• Large scale randomized experiments provide a powerful alternative
  – Run on live traffic
  – Allow for causal inferences
  – Smallest experiments yield about 200K unique cookies per day
Example

• Question: How do ads on YouTube impact usage?
  – Do ads cause viewers to use the site less?

• Naïve approach: Look for correlation between ad viewing and time on site
  – Do users who see lots of ads use YouTube less?
More ads lead to more playbacks? Or more playbacks lead to more ads?
What went wrong?

• Naïve analysis suffers from length-biased selection
  – Long sessions are more likely to have ads
  – Known issue in statistical sampling since at least 1969

• These issues are very common in practice
  – Thread length in textiles
  – Patient visit duration in hospitals
  – Vegetarians in business meetings
Better Methods

• Using cookies to divide the population of YouTube visitors
  – Expose some of the population to a new treatment (e.g. new ad format, withholding ads, throttling ad coverage)
  – Keep an equal sized sample of the population as a control

• Measure comparisons between the two groups to determine if the experiment changes user behavior:
  – More watches on YouTube
  – Longer session length
  – Reduced in-stream ad abandonment
Holdback experiments

- YouTube ad formats
  - In-stream video ads
  - Overlay ads
  - Mid-page companion units (MPUs)

- Holdback experiments
  - 6 experiments holding back combinations of the 3 ad formats
  - 1 additional experiment to holdback all ads
  - 1 additional experiment for the status quo (control)
  - Each experiment run on 0.1% of YouTube traffic

- Compare playbacks per visitor among the 8 groups
Watch impact in the U.S.

<table>
<thead>
<tr>
<th>Option</th>
<th>US</th>
<th>Global</th>
<th>View Impact (Index)</th>
</tr>
</thead>
<tbody>
<tr>
<td>no ads on watch</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>overlay only</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mpu only</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no instream</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>instream only</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no mpu</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no overlay</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>status quo</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ARF AUDIENCE MEASUREMENT 6.0
Further analysis: Impact of advertising on partners

- Partners control how many in-stream ads are shown on their content
- We can measure the partner-level impact from showing in-stream ads using the in-stream holdback experiment
  - Partners who show an in-stream on at least 1% of their views see a 5% decrease in watches
  - Approximately 1 view is lost for every 3 in-streams shown

Experiments provide necessary metrics partners can use to make decisions
Partner impact of instream ads

Partner Level Impact On Views From Instream Holdback

Partner Level Impact On Views From Instream Coverage

Impact on Views (Index)

Normal Quantiles

Percent Of Views Showing An Instream Ad

Impact on Views (Index)
Conclusions

• Retrospective analysis can be misleading
  – Direction of causation can be difficult to determine

• Randomized experiments can help
  – Provide causal connections rather than correlations

• Online media is uniquely suited to the experimental approach
  – Live traffic can be segmented at random
  – Changes in user behavior can be measured precisely
Next Steps

• Understand advertiser impact
  – Recent experiments focus on user and partner impact
  – New experiments should explore advertiser hypotheses as well

• Broaden our scope
  – Effectiveness of different ad formats
  – Relevant advertising to reduce ad impact
Thank You!

• Sundar Dorai-Raj (sdorairaj@google.com)

• Dan Zigmond (djz@google.com)