

PROGRAMMATIC GROCERY RETAIL BANNER ADS HAVE LASTING IMPACT ON SHOPPER BEHAVIORS



**PRECISION
MARKETING**

POWERED BY
34.51°

CONTRIBUTORS:

CLAYTON BECK

ILANA BLUMIN

NICK HAMILTON

LUIS LORET DE MOLA

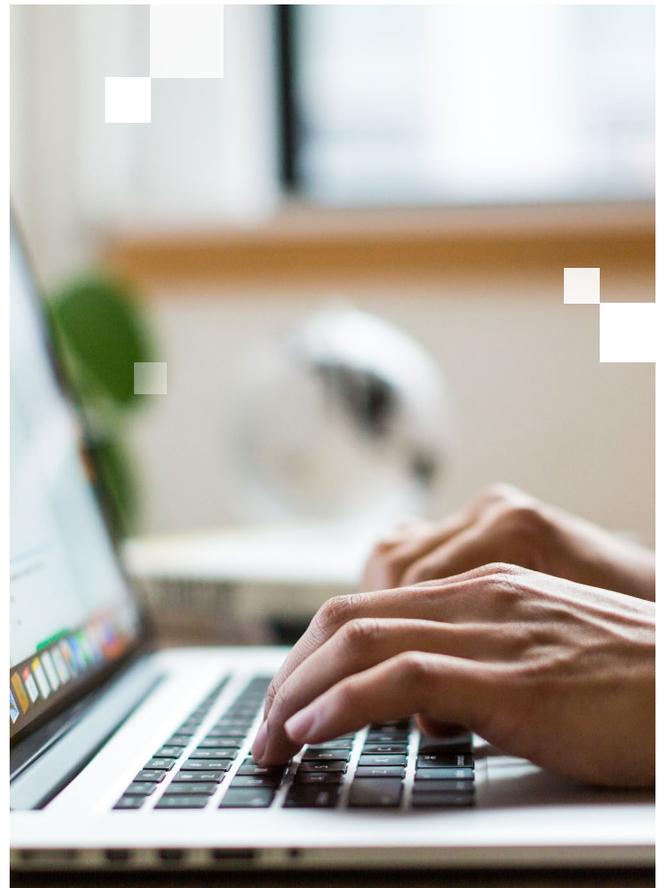
ADAM WARD

ABSTRACT

Traditionally, the media industry quantifies the impact of advertising by examining a relatively short period of time post-exposure. While advertising may have its largest impact soon after exposure, driving value to a brand is worth more than just a household's short-term purchase. 84.51° has the unique ability to isolate the true impact of advertising on households, and whether it is sustainable. We recently researched the measurement time frame of digital display media to evaluate our hypothesis that advertising has a larger, longer-term impact on household behavior. 84.51° examined 47 digital banner ad campaigns and assessed their impact on purchase behaviors at 2, 4, 6, 8, and 10 weeks post-exposure. This paper will provide the details of our research and insight into our decision to use a longer post-period for measuring digital media.

INTRODUCTION

The world of digital advertising measurement can be described as young and inconsistent, because in many cases data can be hard to obtain and is often incomplete. Many times, companies find themselves obtaining digital media performance metrics from highly aggregated marketing mix models, noisy matched market studies, overzealous attribution models, or test/control studies built from incomplete panels of receipt scanners. While these methods may be the only approaches available given data limitations, they can leave the marketer with inconsistent results that may drive disparate conclusions based on what data is available and which partner is executing the study.





At 84.51° we leverage a highly accurate, closed loop measurement approach. By passively capturing 97% of all transactions through Kroger's loyalty card, leveraging a randomized holdout control, and targeting only customers that can be measured, 84.51° is uniquely positioned to decipher the true impact of digital marketing at the household level. With our strict household level test/control methodology, 84.51° can tease out the true granular impact of a campaign by accounting for the fact that many purchases occur regardless of marketing exposure. Many companies rely on attributable sales to assess media which vastly exaggerates campaign performance because it's not able to differentiate purchases that occur as a direct result of media exposure and purchases that would have happened regardless of exposure. Calculating and reporting incremental sales allows 84.51° to tease out the portion of those attributable sales that were driven by the specific campaign being measured. The prior statement cannot be made by purveyors of marketing mix modeling and multi-touch attribution. Marketing mix lacks granularity, making household level impact impossible to assess. Highly aggregated weekly DMA level impressions leave much to be desired in attempting to analyze the micro-level impact of marketing at a campaign level, let alone at the household level. Multi-touch attribution tends to overstate the impact of media given it assigns full or partial credit to various media channels simply because someone was exposed, without regard to whether that purchase would have occurred anyways. Assumptions, highly varied algorithms, and projections from limited household data are all factors that should cause the marketer to look for a more accurate measurement solution. The ability to compare behaviors across a pure and complete test/control group separates 84.51°

in our ability to assess the efficacy of household level marketing. As we continue to explore and refine the impact marketing has on consumer behavior, leveraging our granular measurement approach has shown us that **digital media can directly impact households much longer than what the industry previously thought.**

The ability to compare behaviors across a pure and complete test/control group separates 84.51° in our ability to assess the efficacy of household level marketing.

METHODOLOGY

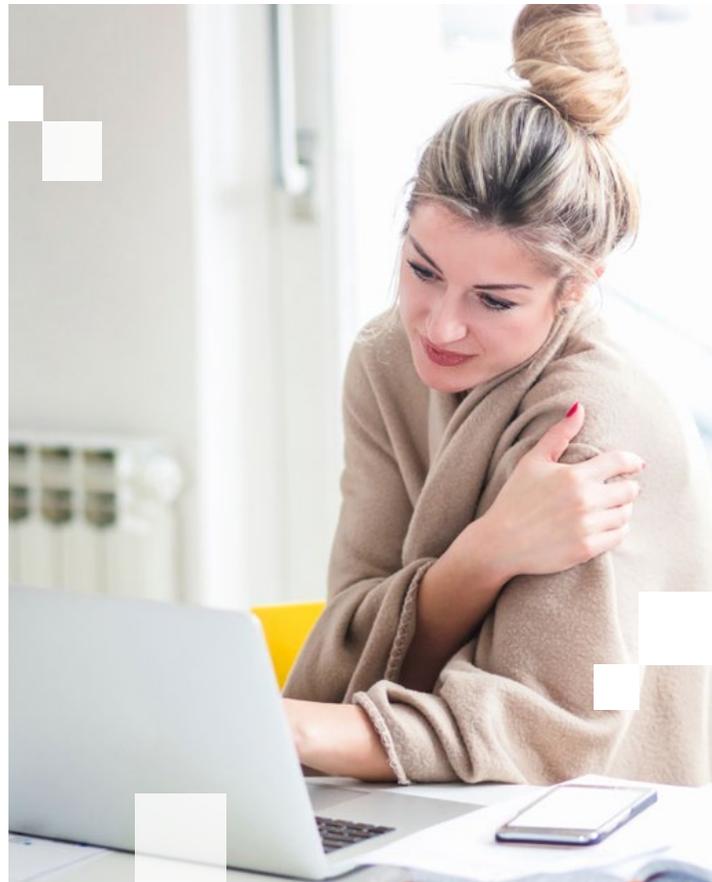
84.51°'s vast shopper data assets and our robust targeting and measurement approaches give 84.51° the unique ability to assess long-term digital banner impact. The Kroger loyalty program captures 97% of all purchases which are used to target households for digital media. This granular level of targeting also allows us to holdout an unbiased, stratified, random 10% control group. Through our relationship with a data onboarder, we can track exposures across devices back to Kroger households. Because not all targeted households end up being exposed in a digital media campaign, we leverage a thorough household matching process to ensure each final exposed household is aligned to its most comparable control household. As a final step, we apply ANCOVA models to measure lift to account for any minor differences in pre-campaign KPIs. Additional research was completed to ensure that our final matched exposed & control populations did not naturally degrade over time, ensuring an accurate longer-term assessment of incremental performance. Appendix A contains further details on our targeting, measurement, and research into match strength over time.

RESEARCH DETAILS & RESULTS

SAMPLE CAMPAIGN SELECTION

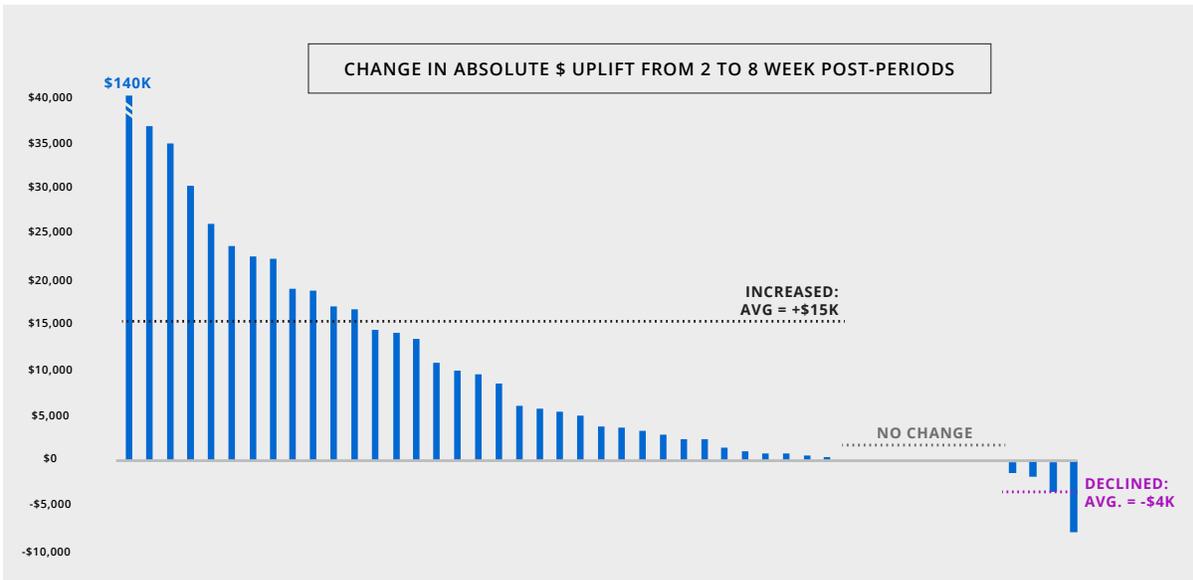
For this study we selected a representative sample of 47 Kroger/CPG co-branded programmatic banner ads and ran measurement using 2, 4, 6, 8, and 10-week post-period lengths. Various campaigns were selected for this analysis across multiple CPGs (and Kroger branded products), grocery departments, campaign durations, & time periods:

- **Number of campaign measurements:**
47 measurements
- **Range of campaign start dates:**
12/13/2017 – 11/11/2018
- **Range on campaign lengths:**
11 – 62 days
- **Sample categories:** ready-to-eat cereal, refrigerated beverage, cookies, crackers, snacks, nuts, carbonated soft drinks, yogurt, frozen, breakfast foods, bread, granola/snack bars, coffee, frozen pizza, candy, health/beauty/cosmetics, over-the-counter Rx
- **Range of households contacted per campaign:**
187K – 1.6MM
- **Range of total exposures per campaign:**
3MM – 21MM

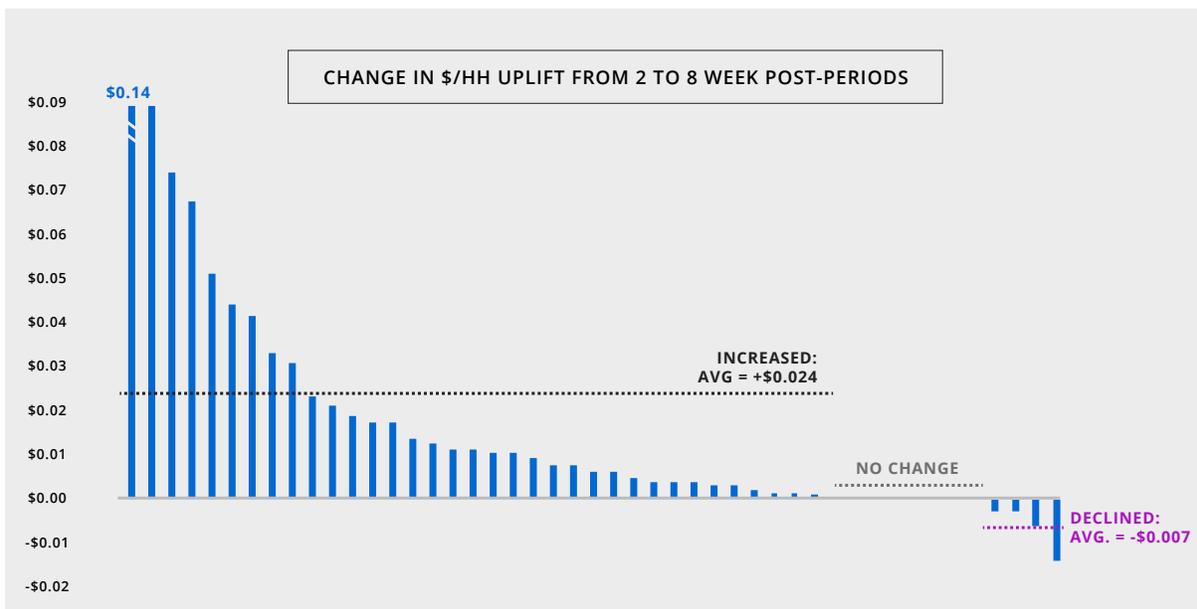


CAMPAIGN RESULTS

35 of the 47 campaign measurements exhibited an increase in lift from 2 to 8 weeks (8 were flat, 4 decreased). On average, campaigns that showed a positive change increased their absolute incremental sales by +\$15k, a 179% increase over 2-week results:

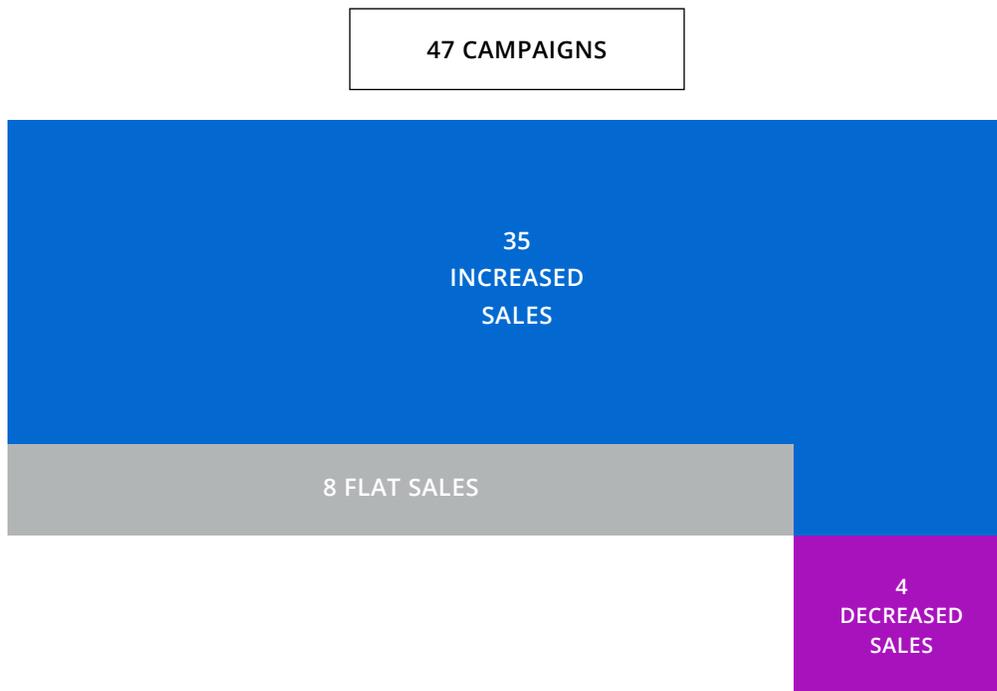


Campaigns that showed a positive change averaged an additional +\$0.024 of per-household incremental sales, an increase of 162% on average compared to 2-week results. We show per-household lift so that campaigns with varied overall household reaches and sales volume could be compared on equal footing.



All individual campaign results were tested for significance via ANCOVA, and any result that fell below 80% confidence were zeroed out (p-value > 0.2 = results not directional or statistically significant).

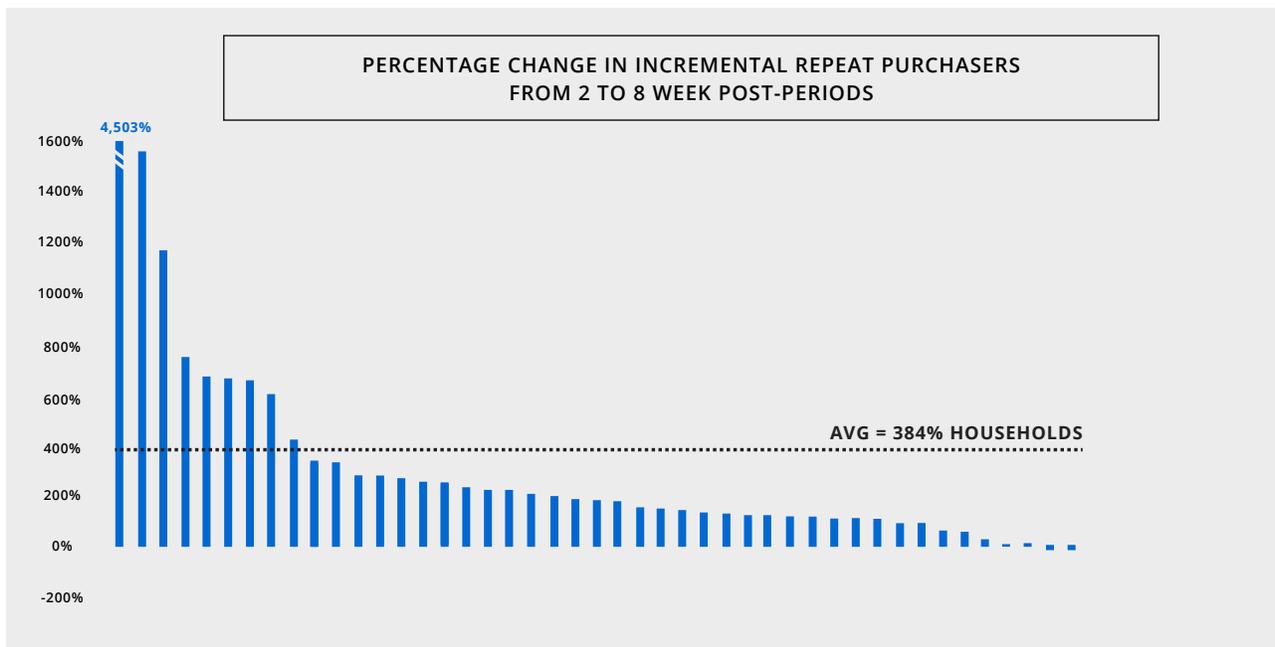
Of the 47 sample campaigns examined, there were 8 that remained flat and 4 that saw decreased sales lift from 2 to 8-week post-periods. The campaign information at our disposal did not provide clear evidence suggesting that there was a specific cause for this. Further research would be needed on a larger sample size to determine if there are specific campaign factors that are correlated with this small decrease in lift.



REPEAT RATE TRENDS

We examined the change in incremental repeat purchasers when moving from a 2-week to an 8-week post-period. Nearly all the 8-week campaign reads drove more incremental repeat purchasers than the 2-week reads. This supports our hypotheses that the initial campaign exposure may motivate a short-term purchase, but many of those households will make additional repeat purchases over time, driving an increase in incremental sales.

The campaigns averaged a nearly 400% increase in incremental repeat households when moving from a 2 week to an 8-week post-period:



This increase in repeat households is a significant driver behind sales lift growth over time. Based on previous research we know the sales lift growth does not stop at 8 weeks, significant growth has been observed up to a 14-week post period.

CONCLUSION

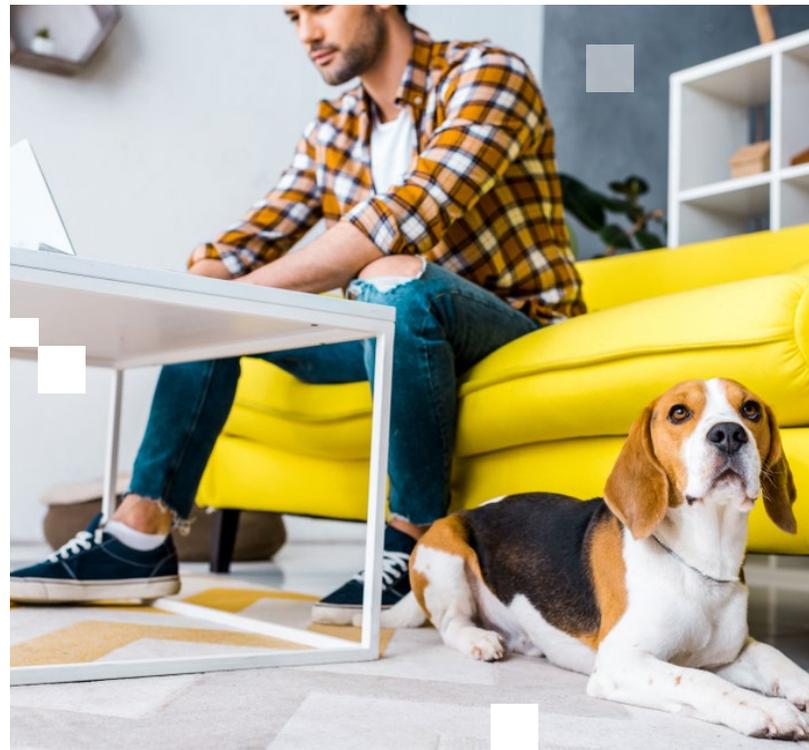
Our research shows that many of our retail banner ad campaigns provide significant sales lift well beyond the traditional 2-week attribution window typically used by the ad tech industry. By leveraging a longer measurement window, we can more precisely quantify the true sales impact of media for brands. It also gives 84.51° the ability to understand if the household impact is sustainable. This longer-term lift will only be enhanced by campaigns with quality messaging/creative and relevant targeted populations.

If an advertiser can leverage complete household-level sales data to measure media impact in a closed-loop environment, a longer-term post-period is the most accurate way to assess campaign performance. Otherwise, retailers and suppliers are leaving incremental sales on the table (179% difference on average) and are not taking full credit for lasting changes in household behavior due to online ad exposures.

After this research we have chosen to measure Kroger/CPG banner ads leveraging an 8-week post-period. This strikes a balance between maximizing true long-term incremental lift and the practicality of delivering results within a reasonable timeframe for our CPG partners to action against.

84.51°'s data assets and strict methodology enables better understanding of marketing products and their different effects at the household level. This understanding is crucial in determining budget allocations for our CPG partners and their specific objectives. Without truly understanding what each

marketing product accomplishes, CPGs are subject to misinformation and bias that can potentially result in substantial waste in marketing dollars and effort. 84.51°'s household level measurement methodology at an 8-week post period allows the marketer to get the most accurate view of how each marketing campaign performs without being forced to rely on the variability of multi-touch attribution or the macro-level aggregations of marketing mix modeling.



METHODOLOGY

84.51°'s vast shopper data assets and our robust targeting and measurement approaches give 84.51° the unique ability to assess long-term digital banner impact. This next section provides more detail on our industry-unique methodologies that allows for an extremely precise understanding of the impact advertising has on actual household behavior.

TARGETING & CONTROL SELECTION

84.51° has a uniquely robust dataset. 97% of all transactions and 100% of all online orders are tracked back to the household. For each campaign, we leverage this data to exclusively target verified Kroger households using machine learning algorithms to optimize household selection and performance.

Once the most relevant households for targeting are finalized, we then take a 10% random sample stratified by targeting cell, division, and model scores to form the control group. The control group is a subset of the finalized households targeted, meaning that both groups share characteristics across the array of variables used to specify targeting cells and apply model scores. The only difference is that the control group will not receive media exposure, while the targeted group will. This is important as our test and control groups have no bias between them. In the industry, many companies will simply leverage households who were not exposed during the campaign as their "control" group (if a control group is used at all), without stratifying on key household characteristics that can affect purchase behavior. The issue with this approach is the inherent bias between the exposed and unexposed populations. While approaches can be taken to reduce this bias, it is impossible to remove all bias without a pure holdout. Proper control selection is imperative to truly isolating the impact of advertising on customers and is not consistently practiced in the industry.

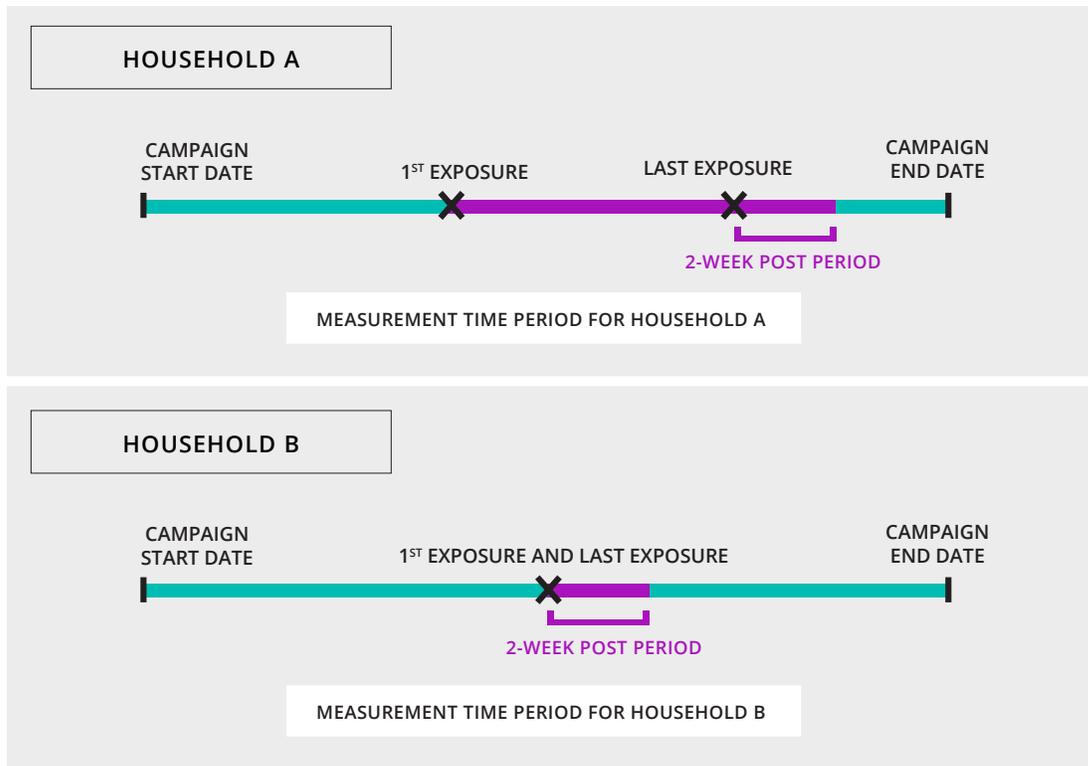
EXPOSURE TRANSPARENCY

Through our partnership with a data onboarder, we can target Kroger households across their multiple devices and match back exposures at the household level. Because of this we know exactly which households were exposed, how frequently, and when those exposures occurred. Access to this granular, shopper-level exposure data means we do not need to rely on highly aggregated matched market or marketing mix modeling studies to assess performance.

Exposure data plays a critical role in determining the measurement period for a campaign. Digital banner ad campaigns usually have a fixed start and end date, but not all households receive their first exposure at the start of the campaign period. A household can receive their first exposure at any point during the campaign. Therefore, if we measure all exposed households based on the same fixed start date of the campaign, we will be erroneously measuring the impact of media on households before they are ever exposed. Similarly, not all households are exposed on the fixed end date of the campaign.

So, while all households must have their first and last exposure within the range of fixed start and end dates for any given campaign, there is significant variation of first and last exposures for each household.

MEASUREMENT WINDOWS FOR TWO SAMPLE HOUSEHOLDS:



Our granular exposure data enables us to define a unique measurement window for each household, depending on when they were first and last exposed to the digital banner ad. This is unique within the industry.

Prior to this research, the measurement end date for each household was based on the last exposure to media plus an additional 2 weeks. This paper specifically examines the impact to lift results caused by extending this post-period past a standard 2-week attribution window the industry typically relies on.

EXPOSED VS. CONTROL MATCHING

Thanks to the exposure data from LiveRamp, we can identify who from our original targeted group was exposed to the media and filter down to exposed households for the measurement of campaign performance.

Due to the nature of digital media, all targeted households do not end up being exposed to the campaign. Because of this, we need to ensure that our holdout control is representative of those ultimately exposed to the media. We do this by matching each exposed household to its most similar control household. These households are matched on their behavior prior to their first exposure, ensuring each exposed household is precisely matched with a customer withheld from the campaign after removing any outlier households. We use a number of different variables to find similar exposed and control household pairs. Variables such as sales for advertised products and their respective categories, store segmentation, and total store shopping habits.

This additional household matching step ensures that any pre-campaign differences between the exposed and control group are minimized. For example, utilizing this type of match prohibits us from targeting a heavy buyer of product X and then comparing their post-campaign sales to a light buyer of product X.

ANCOVA

Although the match between exposed and control households are extremely tight, we recognize that the match will never be exact and there may remain a small difference between the exposed and control's purchase behavior before the campaign starts (because we match on spend and spend is a continuous variable). To adjust for this, we calculate KPI lift using the analysis of covariance (ANCOVA) model displayed below using pre-campaign spend/units/visits as covariates.

$$Post_i - B_0 + B_1 (Group_i) + B_2 (Pre_i) + \epsilon_i$$

ANCOVA measures the KPI difference between the exposed and control group after the campaign completes, while controlling for any innate KPI differences between those groups prior to the campaign starting.

In the equation above, Pre and Post refer to the household-level KPIs before and after exposure to the campaign (e.g. sales, units, visits). Group refers to whether a household was in the exposed or control group.

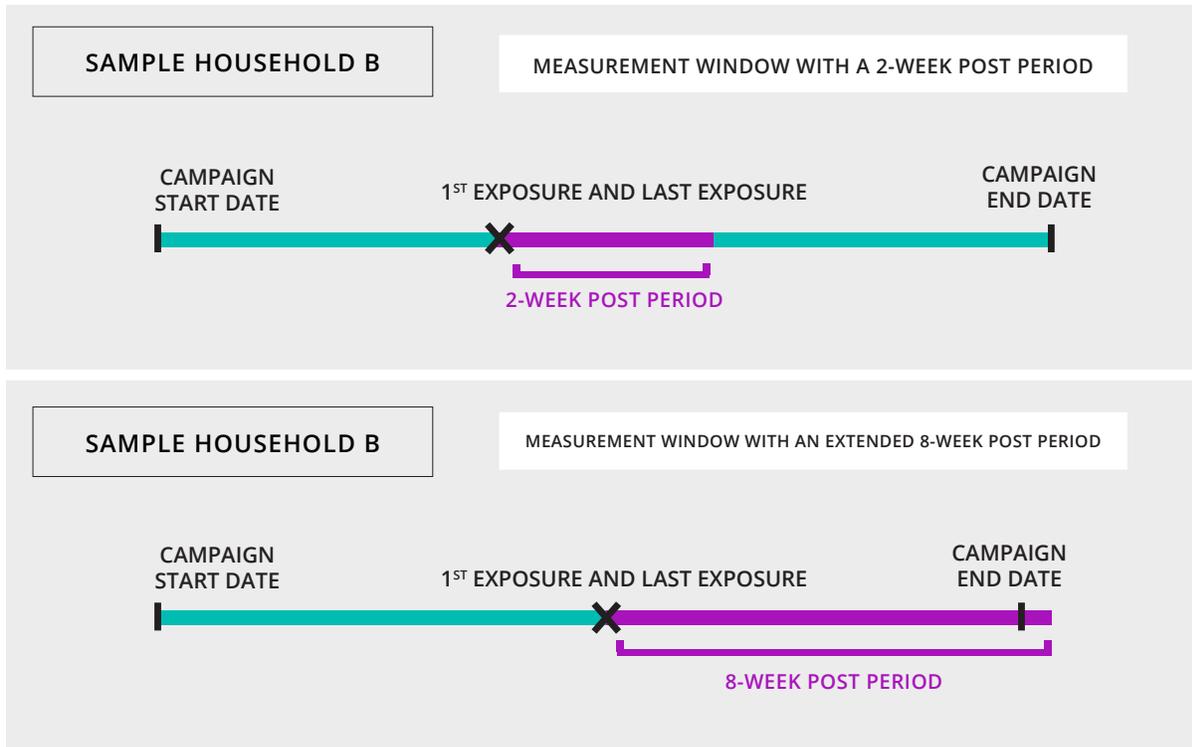
So, if we were calculating sales lift for a campaign, the coefficient B₁ would convey the difference in exposed and control group post campaign sales (sales lift), while holding pre-campaign sales constant (B₂ represents pre-campaign sales).

By leveraging ANCOVA, we can be certain that any lift over the control group is due to the campaign itself and not any minor inherent differences between exposed and control populations prior to the campaign even starting.

Most companies extrapolate campaign results based on a small sample of households where data is available. This causes retailers and advertisers to make multiple assumptions, often causing performance to be overstated.

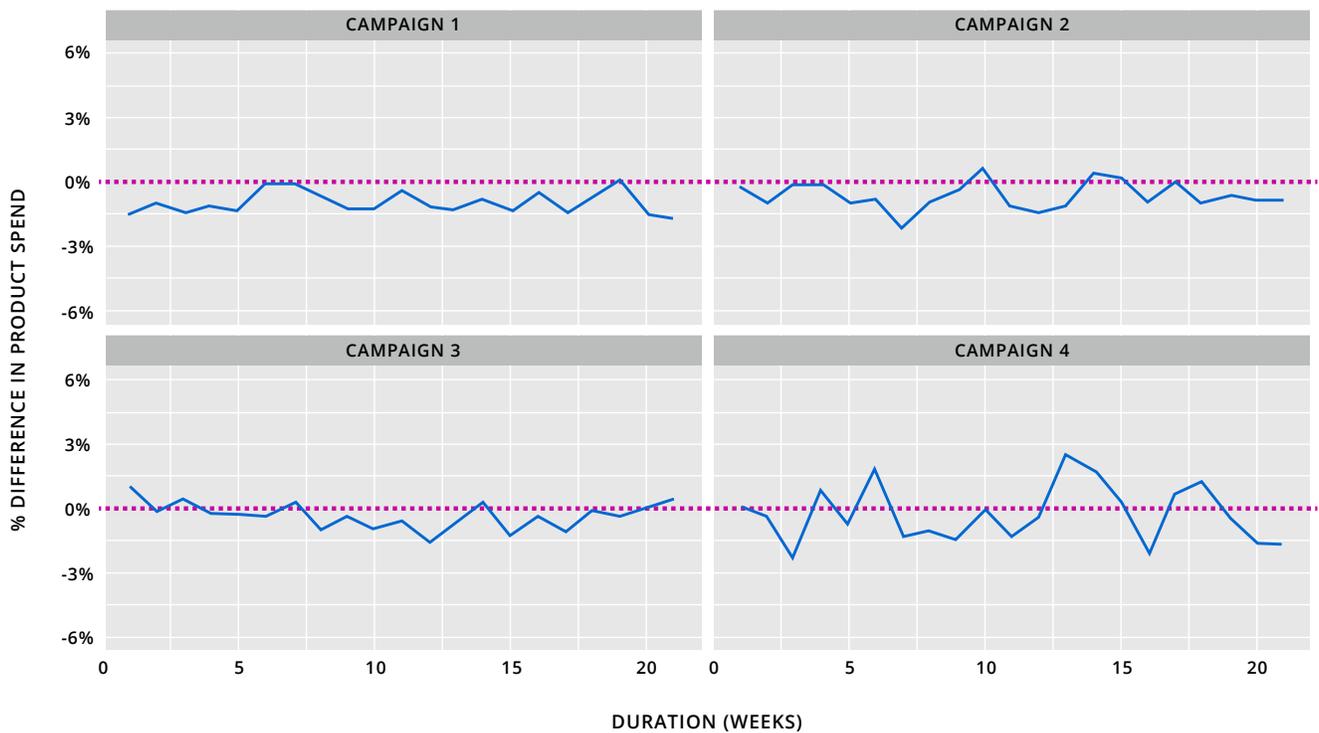
TEST AND CONTROL MATCH THROUGH TIME

In order to properly test our hypothesis that digital banner ads have a lasting impact on household behavior, we had to ensure that any sales lift captured throughout the extended post period was not simply due to the exposed and control match degrading over time.



To test this, we ran 4 “dummy” campaigns through our test/control matching and measurement processes (using different product groups). There was no media being exposed to our test group, so that we could examine how long our household match remained intact. If we observe an increase in variance between the test and control group as time increases, then we know that there is a match degradation through time since the test group was not actually exposed to anything that would impact purchase behavior.

TEST AND CONTROL MATCH THROUGH TIME



We observed no increase in product spend variance between the test and control groups through the course of 21 weeks. The duration of 21 weeks was chosen to capture the campaign period plus longer-term post period of all 47 digital banner ad campaigns researched. This gave us enough evidence to confidently infer that the paired test and control households remained comparable throughout the course of the extended measurement window, and the observed KPI lifts are not due to a match degradation over time.

CHANGE IN LIFT BY CAMPAIGN DURATION

A natural question that arises, is whether this longer-term impact can be captured by simply extending the time in which the digital banner ad is live. Based on the campaigns sampled for this research, we found no meaningful relationship between campaign duration and change in 2 to 8-week sales lift:

