

WHICH ANALYTICAL METHOD **BEST MEETS** THE ATTRIBUTION CHALLENGE

By **Robert Stratton** | Senior Group Director | MarketShare: A Neustar Solution





ADDRESSING THE ATTRIBUTION QUESTION

Advertisers don't always have a clear idea about how much value each channel or campaign adds to their companies' success.

Despite the availability of big data, advertisers don't always have a clear idea about how much value each channel or campaign adds to their companies' success. To solve this problem, marketers need a clear methodology to attribute credit to the different marketing channels that correctly measures the value of each marketing contact. While this statement may seem obvious, it is hardly a zero-risk proposition. The wrong attribution methods can lead advertisers to incorrectly assess the value of marketing efforts—driving them to move dollars away from successful channels and toward less-effective ones.

To address the attribution question, MarketShare, a Neustar Solution, conducted a major big data study to identify which of the most common attribution models are the most accurate. The study simulated a year's worth of advertising on a population the size of the United States (316 million people; hence, in the study, 316 million software "entities"). The aim of the project was to create a complete population in which it was clear how advertising influenced purchase decisions—and then to run various attribution approaches to see how close these approaches came to actually measuring the real influence of ads.

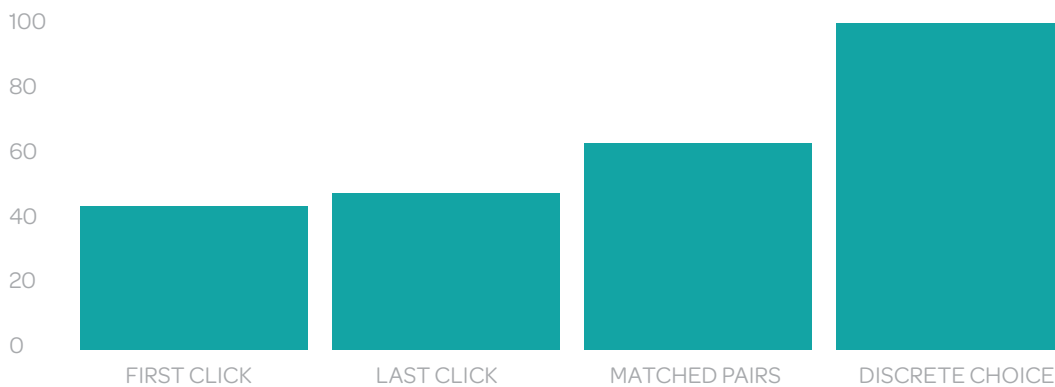
The simulated individuals were exposed to online and offline advertising; and were programmed to increase their likelihood to buy after every ad exposure—up to a saturation point. The entities were programmed with varying degrees of innate interest in the product; and many of them went in and out of market to shop over the course of the year. The resulting analysis was huge—generating 2 terabytes of simulated tracking data over two weeks’ time.

Once we had created the simulation, we ran attribution analyses on the ways advertising influenced the simulated population’s purchase decisions. We analyzed four of the most common digital attribution models: first click, last click, matched pairs, and discrete choice models. Since the simulation already “knew” exactly how advertising had influenced purchase decisions, we were able to match the various attribution analyses with the actual influence that ads had had on purchases. We then scored the attribution methodologies to see which matched most closely with the ads’ actual influence.* The result showed us which attribution models were the most—and least—effective at understanding the customer’s path to purchase.

As shown in the ranking below, there was one clear winner amongst the attribution approaches—discrete choice models (nearly 100% accurate); with first-click coming in last among the approaches (roughly 45% accurate).

Read on to learn more about the study findings and the strengths and weaknesses of the attribution methodologies we analyzed.

ACCURACY OF ATTRIBUTION MODELS



* The specific rating was determined by the sum of the differences between actual influence of ads on purchases and the attribution that each method reported, divided by the total number of sales.



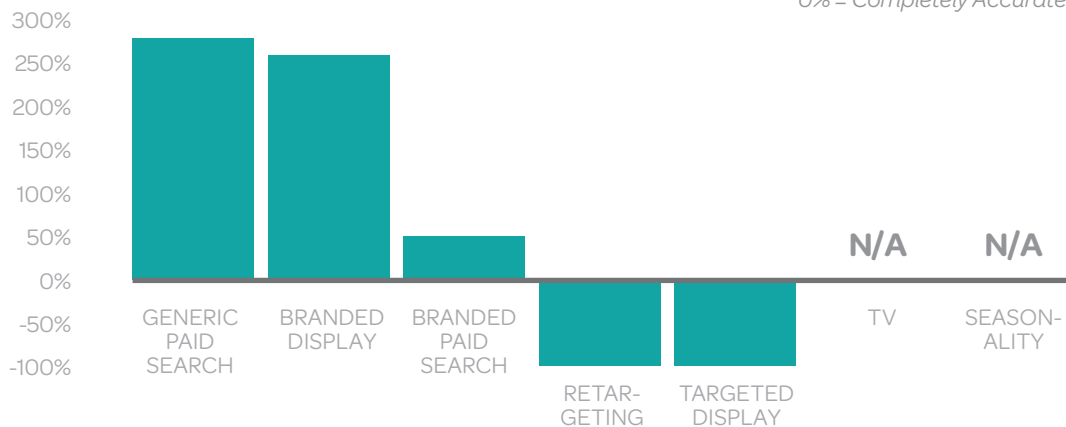
FIRST CLICK

First Click attribution gives all credit to the first touch along the conversion chain—based partially on the premise that the first consumer/brand interaction was the one to “kick off” the entire path to conversion.

Comparing first-click attribution with the real influence of ads reveals a number of “fatal flaws” inherent to the approach. The method clearly over-estimates the impact of search and branded display—both which drive high levels of initial awareness—but underestimates lower-funnel channels like retargeting and targeted display. In addition, because first-click is based purely on cookie level information recorded in digital log files, it fails to capture the influence of offline influences—like TV; as well as factors outside of marketing that can have an enormous influence on sales, such as seasonality.

FIRST CLICK: PERCENTAGE ERROR BY CHANNEL

0% = Completely Accurate





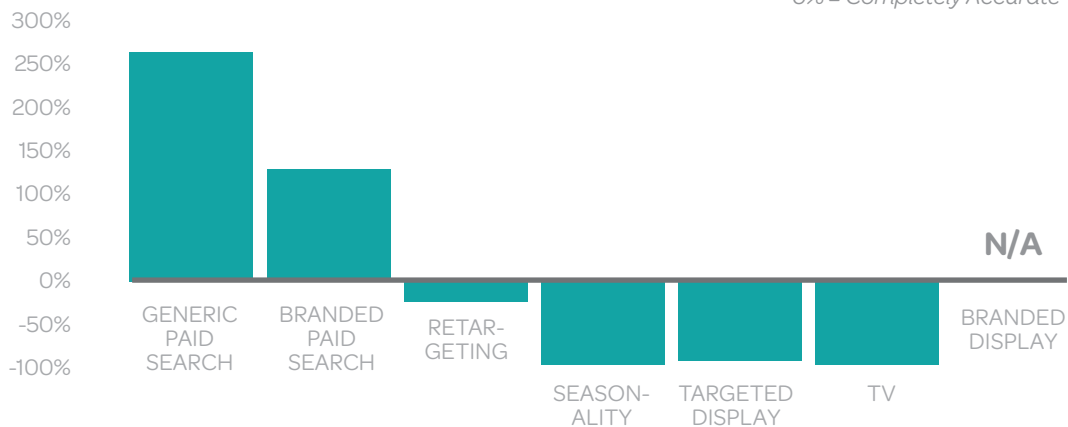
LAST CLICK

A “close cousin” of first-click, last click attribution gives complete credit to the touch that immediate precedes conversion. If a customer saw a TV ad, ran a web search that drove her to a site, and then saw a retargeting ad, last-click will give credit to the retargeting ad—and ignore the influence of the TV ad or the search. Both last click and first-click are widely viewed as overly simplified, but are commonly used for a quick “snapshot” of sources of marketing leads.

Comparing last-click attribution with the real influence of ads shows that last click strongly overemphasizes the impact of search—not surprising, as search tends to be the final marketing engagement before a consumer arrives at a site. Last click also was shown to overlook the enormous impact of TV and seasonality—once again unsurprising, as last-click is both bottom-funnel and cookie-based, and so is likely to fail in addressing non-addressable and/or top-funnel influences.

LAST CLICK: PERCENTAGE ERROR BY CHANNEL

0% = Completely Accurate



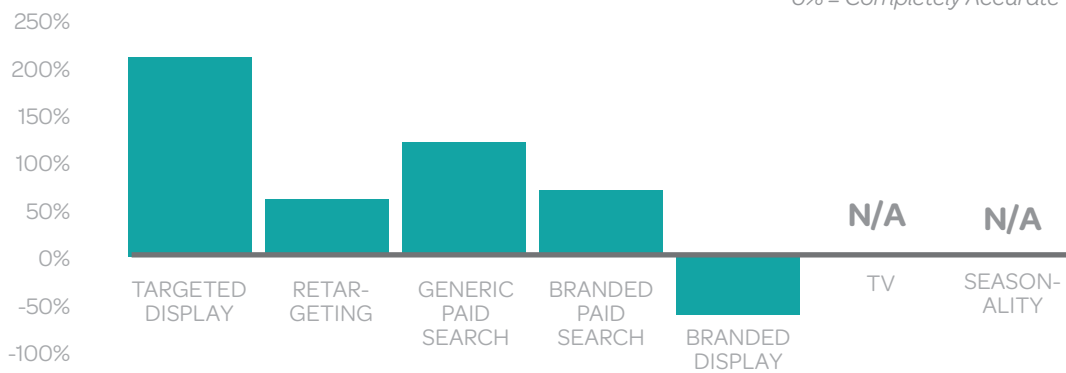


MATCHED PAIRS

Essentially a large post-hoc test and control study, matched pairs finds sets that are identical except for one attribute, and compares outcomes both with and without that attribute. For instance, to understand the impact of targeted display, matched pairs attribution might look at conversion rates for consumers who have been exposed to all media including targeted display, as compared to the conversion rate of those who have not been exposed to targeted display but who have been exposed to all other media. According to this approach, the difference in conversion rates between the groups should reveal the effectiveness of targeted display.

MATCHED PAIRS: PERCENTAGE ERROR BY CHANNEL

0% = Completely Accurate



There were several areas of discrepancy that we observed between the matched pairs approach and the true answer that the simulation provided. Investigating the discrepancies suggested four main problems:

- 1 Matched pairs relies on measuring the impact of consumers' different experiences—for instance, the impact of being exposed to retargeting versus not being exposed. However, many of the strongest influences on ad effectiveness, such as seasonality and TV exposure, are nearly universal, and so there is simply no difference to measure. For this reason, matched pairs is not designed to measure the impact of these shared experiences. This would largely explain why TV and seasonality are both completely overlooked by the matched pairs analysis.
- 2 On a related note, matched pairs relies almost exclusively on cookie-level data—which provides insight into who has or has not been exposed to a given type of advertising. Cookie level data alone, however, will not identify the influence of events that are not recorded at the cookie level, such as much of offline advertising and seasonality.
- 3 Matched pairs judges ad effectiveness by comparing conversion rates of populations who have been exposed to an ad with populations that have not. There is an inherent bias problem in this analysis, however—as advertisers naturally tend to target groups who are highly likely to convert. For this reason, it can be extremely difficult to separate the impact of the ad on a group of consumers from the natural tendency of that group to buy the product, regardless. In other words, it's easy to over-estimate the impact of an ad on a population that needs very little convincing to buy. To address these problems in our matched pairs analysis, we adjusted the data for intention bias—matching groups with like inherent intent to buy, and factoring for behavior that indicates an imminent purchase.

However, it is impossible to weed out all such biases from population-level studies. This likely explains why, even with our adjustments of the sample, the approach still significantly overestimates the impact of targeted display, which obviously favors buyers who are likely to purchase. (It is also worth noting that, in practice, many digital attribution matched pairs practitioners do not account for these biases at all. To learn more about what this study revealed about attribution at consumers' pre-existing intent to purchase, see [our separate brief](#) on attribution and customer intent.)

- 4 Matched pairs asks a binary question—were consumers exposed to a type of advertising, or not?—and measures ad influence accordingly. That approach fails to account for the incremental impact of additional ad exposures. A consumer will respond very differently to one, two, or fifteen ad exposures; since these are not binary distinctions, matched pairs struggles to address that difference.

A second result of this binary approach is that it leaves recency out of the calculation, even though an ad exposure from today is very likely to have a different influence than an ad exposure from last week. Since matched pairs looks primarily at who was exposed to an ad versus who was not, the approach gives the same treatment to all ad exposures, regardless of when they were delivered.

This binary approach likely accounts for the fact that matched pairs under-attributes branded display by a great deal, and also underestimates branded paid search and retargeting. In their own ways, all of these channels serve a branding function—and so they stand to become more effective the more the consumer is exposed, and the more recently the exposure has been. Matched pairs, however, is not designed to register the impact of those factors.



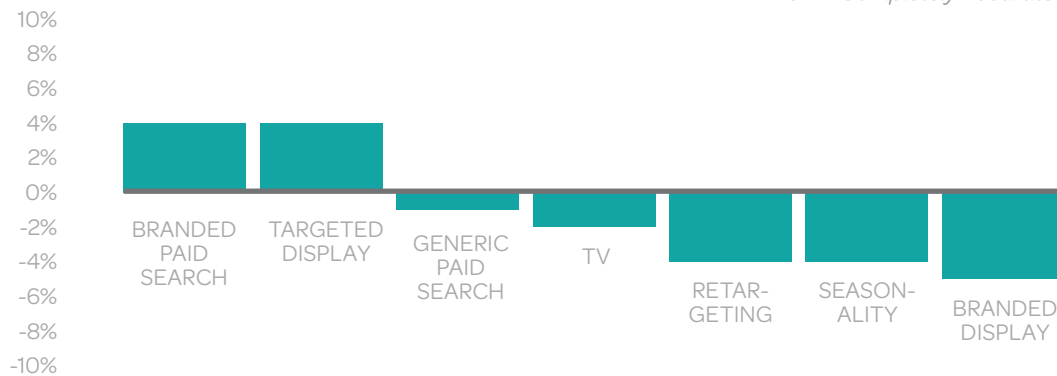
DISCRETE CHOICE MODELS

Widely used in machine learning applications, discrete choice models ask which attributes predict a given action, as well as how changes in those attributes might bring an action closer. Discrete choice models ask “who” the customer is, what actions that customer has taken in the past, and how these may work as predictors of an imminent conversion. The answers to these questions might lie in path data—such as prior interactions with the brand’s advertising and website—that suggest a logical next step in a journey. It could also use segmentation data, such as demographics and psychographics—as different populations will be more or less likely to take specific actions.

The discrete choice model was close to 100% effective in capturing the real underlying process of ad influence:

DISCRETE CHOICE: PERCENTAGE ERROR BY CHANNEL

0% = Completely Accurate



The high degree of accuracy that we see in the discrete choice model can be attributed to a number of factors:

- Because the model takes offline factors into account, the models are not inaccurately biased toward online media alone.
- Because discrete choice models operate at the individual consumer level, they can fully control for sampling and intention. As mentioned above, matched pairs can adjust sampling, but only at the segment level. Inevitably, accuracy will be lost in the blunter approach. Discrete choice modelling examines individual consumer paths, not group behaviors—and so can create a much more precise picture of ad effectiveness.
- The model takes both the frequency and recency of exposures into account—again, both enabled by tracking and analysing customer journeys at the individual level.

In short, discrete choice modelling provides a more complete and exact understanding of what drives individuals toward conversion. This, in turn, produces a more accurate understanding of which marketing investments work, and how. These are amongst the reasons MarketShare employs the discrete choice approach to digital attribution.

Discrete choice modelling provides a more complete and exact understanding of what drives individuals toward conversion.

Learn more at www.marketshare.com



ABOUT MARKETSHARE

MarketShare, a Neustar Solution, combines advanced analytics technology, scientific leadership and deep domain expertise to enable large companies to measure, predict and dramatically improve marketing's impact on revenue.

MarketShare's clients use its analytics and tools to direct tens of billions of marketing investment dollars globally.

Learn more at www.MarketShare.com



MarketShare[™]
a Neustar[®] Solution