MARKETING EVOLUTION
Customer Connections, ROI Optimized

Flaws in Attribution Are Costing the Industry Billions in Profits

5 Mistakes Every Marketer Should Avoid

by Rex Briggs

A revelation has recently come to light– that multi-touch attribution (MTA) models have fatal flaws that are costing companies billions in sales and profits.

An informal consortium of companies came together to look into the matter. The consortium included a few marketers, including a prominent insurance company, a leading advertising software company, a couple ad agencies, and a few academics. Each had independently found fatal flaws in the recommendations coming out of popular attribution solutions. These budget re-allocations were so contrary to effective spending that, when applied, they actually reduce the amount of sales per dollar of marketing spend. One retailer reported faithfully applying the findings of the MTA for two-years, yet, business growth didn't materialize.

Some questioned if the attribution vendors owned by media companies were self-serving in their inaccuracy. They suggested either fraud or system gaming. Others were more deferential, noting that applying math to a problem can create an illusion of good science, but may blur obvious logic. They pointed to over-reporting the value of impressions delivered to people that would have bought anyway. Whatever the reason for the flaws in attribution recommendations, the consortium members all believe in the value of analytics and want to ensure the industry gets the analytics right. As one member explained, "The consortium's desire is to shine a bright light into these dark, black boxes of attribution. This will help the industry make better decisions to create more value."

The consortium reached out to me for help because my history of using analytics to evolve marketing toward better performance. For those reading this that I have not had the pleasure of working with, the credentials include a twenty-plus year history of work in marketing analytics:

- First Director of Research for Wired Digital in 1995 (where attribution modeling began)
- First IAB Research Committee Chair
- Created XMOS, the first cross media studies measuring online and offline, side-by-side
- ANA School of Marketing ROI Trainer
- Author of a recent book, *SIRFs-Up*, on how software and algorithms are changing marketing
- Co-author of *What Sticks*, a book encouraging analytics, which earned Ad Age's #1 spot on their top
- 10 list of book marketers should read in 2006
- Current Research World Magazine Editorial Board
- Former Journal of Advertising Research (JAR) Editorial Review Board
- University guest lecturer on Big Data and analytics

The purpose of this brief paper is to outline the findings from the consortium and to open a dialogue in the industry so we can ensure we get the science of attribution right. There are billions at stake, and we have obligation to the professionalism of the industry to identify problems and promote best practices. The remainder of this paper is organized into two sections: First, a review of three independent validation studies that identified the problems with attribution. Second, a discussion of the biases we identified, along with advice on how to avoid the attribution biases.

Major Insurance Company Finds Problems Found With Adometry

According to their website, "Adometry by Google solves the complex challenge of integrating, measuring, and optimizing marketing data across all channels — both online and offline — so you can generate actionable insights that improve ROI."

The company promises that marketers can, "Gain a new level of visibility into the customer journey with sophisticated algorithms that process first- and third-party data, analyzing each and every conversion sequence to reveal the true value, or "lift", of each one. This advanced methodology enables you to make cross-channel decisions and drive greater returns on marketing investments."

But that is not what is found through independent testing from three separate sources.

The insurance company's first independent verification was the application of design of experiment at a large scale. Randomized experiments are considered the gold standard in science. Field experiments can be expensive to execute, and carry with them complexity in execution. The insurance company's research team deserves kudos for implementing a large scale validation experiment. They calculated what was at stake in terms of the size of their marketing budget, and scale of the sales their marketing influenced, and successfully made the case to management to validate the attribution recommendations through gold standard field experiments. However, the results of the attribution did not validate. Applying the recommendations from the attribution model would have done damage to the business.

The second experiment was conducted by their advertising software partner, TubeMogul. TubeMogul was concerned that the

findings from Adometry by Google were self-serving, or just wrong. They implemented a clever and inexpensive experiment within digital media. They placed blank placebo ads in their ad rotation, and trafficked it to the attribution vendor as if they were regular ads.

A placebo ad, by definition, should add no value to the marketer. If the attribution model is accurate, it should yield zero sales contribution. They didn't tell the attribution vendor the ads were blank placebos. Instead, they waited to see what recommendations came back from the attribution vendor. If the attribution was working properly, zero value should have been attributed to the blank placebos. However, Adometry reported hundreds of sales attributed to this placebo ad. This was a major red flag.

TubeMogul led an effort to dig deeper. They trafficked a set of placebo ads following the common recommendations of attribution vendors such as Convertro, VisualIQ and others. If the attribution was working properly, zero value should have been attributed to the blank placebos. However, the conclusion is that the problems with Adometry apply to the broader class of multi-touch attribution vendors, for reasons explained in the "known biases" section of this paper.

The third test was conducted by my firm. Prior to agreeing to author this paper, I wanted to see firsthand what we would find from our own investigation. We used my firm's methodology, which has been independently reviewed by Advertising Research Foundation (ARF). The method has been named a best practice by the Corporate Executive Board, and written up in ESOMAR's book of best practices as well. The design uses advertisements and placebo advertisements to measure the incremental difference in sales as well as brand perceptions. Where placebos cannot be cost effectively implemented, the design uses heavier and lighter exposure levels to measure differences in advertising impact. Sales are measured behaviorally. Brand perceptions are measured through a sample of the population that is surveyed. This combination of brand perceptions and behavior give a holistic view of advertising influence. Brand perceptions are useful in understanding how marketing is influencing consumers. Certain brand perceptions can also be a leading indicator of sales, and therefore useful to include in optimization. What we found supported the Insurance company's and TubeMogul's findings.[1] Our findings confirmed the other two independent measurements that the attribution model was producing recommendations that were damaging to sales and profits. Furthermore, the brand perception data gave us some insight.

Known Bias & How To Avoid Them

Based on these three independent studies, we examined the damage caused by misattribution. We have summarized a list of five biases to watch out for:

1. In-Market/Targeting Bias
2. Cheap Inventory Bias
3. Short Term Digital Signal Bias
4. Missing Leading Indicators
5. Myopic Measurement Design as to why the MTA models were missing their mark.

1. In-Market/Targeting Bias

Imagine for a moment you are standing at the bar and see a line of people with their money out, trying to get the bartender's attention. If you suggest to them they should buy a beer, and they do, should you get credit for selling the beer? Or, were a meaningful number of the people already going to make the purchase – and you would be over-stating your contribution to take credit for all the beers sold that night?

As TubeMogul research revealed, the attribution models analyzed over-count the contribution of advertising targeted to in-market consumers. We believe this bias occurs because the attribution vendors capture the customer journey, but aren't able to accurately tell how advertising within the journey incrementally increases sales. The easy fix is to implement continuous experiments – in other words, use a control group.

Here's an example of how an audience's predisposition to make a purchase can skew correlation-based measurement systems. Tactic A is not creating any new buyers, but appears to have lower cost per action because it was delivering ads to people with a higher propensity to purchase in the first place.

Keep in mind, marketers have been moving budget to the lowest Cost Per Action (CPA), but in fact, that means losing sales in this case.

Source: TubeMogul

2. Targeting Cheap Inventory Bias

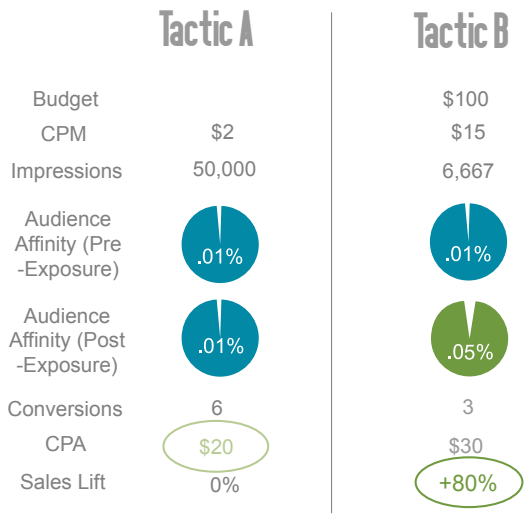
Because one can buy more cheap impressions than expensive ones, the cheap inventory has a higher chance of showing up in a buyer's exposure pool – but that doesn't mean the cheap impression caused the sale. With a control group, we can properly credit the contribution of the impressions, and remove the cheap inventory bias. The consortium found the current crop of attribution modeling credit lower cost inventory with more sales than they deserved. As TubeMogul found, the extreme example of this is the significant amount of sales credit given to a blank placebo ad that was purposely trafficked in low cost placements.

Here's an example of how lack of a control mechanism for cost of media can be problematic in correlation-based measurement systems. Similar to the In-market /Targeting Bias, the CPA reported from the MTA moves dollars away from sales, reducing marketing ROI.

3. Targeting Short Term Digital Signal Bias

If you are trying to create a new customer, you are likely to under-count the contribution of upper-funnel brand development because these people are less likely to click on the advertisements. By the time they are in buying mode, the cookies dropped when delivering the upper-funnel message are likely to already have been deleted. This results in under-counting brand building, and over-counting the short-term activities – because the digital signals (cookie, clicks, etc.) are more likely to be present for the most recent advertising only.

Some of the attribution solutions for bricks-and-clicks retail focus on the online retail as proxy for all sales– but that isn't a consistent yardstick. It is more common for a current customer to buy online than a new customer. TubeMogul found the solution to this bias requires more than applying control groups. One needs a holistic measurement of online and offline sales. Identity matching solutions help connect the dots, and when combined with control groups, marketers will get the right answers and avoid this bias.



4. Short Missing Leading Indicators

So many of the current attribution models report results after the campaign is over. They chase lagging indicators such as sales. If all a marketer wants is a report card after the fact, maybe that is fine. But, really, if all the marketer wants is a report card after the fact, they will not be successful because each round of creative influences different people in different ways. Marketers simply can't be successful in the increasingly real-time world if they are missing leading indicators.

Using leading indicators allows marketers to optimize while the campaign is live, and deliver more relevant messages to consumers, and more effective marketing for the business. In this investigation, we built leading indicator models using a combination of behavioral and attitudinal signals. We validate the leading indicator models against the lagging measures of sales at a consumer level, with use of heavy/lighter exposure or control group. Because we can validate the conversion rate, this allowed us to credit the brand building value of advertising with a specific dollar contribution to sales. We recommend the practice of leading indicators model to marketers to achieve more timely and actionable attribution.

5. Missing Myopic Measurement Design

Back in 1997, we found the power of combining behavioral data with attitudinal brand perception data and a person's profile when analyzing online advertising.[2] Specifically, it allows the marketer to know the "why behind the buy" as Welch's CMO Tom Dixon put it. Importantly, it also illuminates who is being influenced at specific touchpoints. Both "the who" and the when are critical to developing a feedback loop that will improve marketer's intuition about what sticks. As marketing moves increasingly to real-time, personal messaging (RTIM, as Forrester puts it), it is important that attribution design includes more than it does today.

Brand perception data at the personal level is left out of most attribution analysis. When brand perception data is included, it is often a separate aggregate trend, which fails to combine the individual's exposure with their attitudinal and behavioral data. It is cheaper to skip over person-level brand perception data. Some vendors argue that they are avoiding "sampling error" by leaving this data out, but we think they are introducing a myopia bias. We find including attitudinal and profile data is additive – it doesn't take anything away from using 100% of impressions and sales data for the analysis.

The inclusion of brand perception data makes it much easier for a marketer to evaluate the voracity of the attribution recommendations. It makes it easy to diagnose why some messages convert to sales and others don't. A marketer can see whom, specifically, certain ads most influence. Marketers can look at the creative execution and interpret if the attribution results appear logical.

Marketers should look at a cross-section of people. They should examine the specific next best messages for each person to

access the face-validity of attribution. They should examine the profile of whom different messages influence and confirm the logic of the analysis. They should diagnose certain messages that are failing to generate a sales impact to determine if it is a brand favorability problem, or a persuasiveness of the offer problem. An attitudinal + behavioral design helps a marketer to confirm the logic of the attribution and helps the marketer to develop better intuition as they learn more from each analysis.

Conclusion

Unfortunately, the evidence from the consortium is that these different decisions may not be the right decision to increase the contribution from marketing to sales and profits. Since marketing generally drives a minority of the total sales in the short-run, it may be difficult to tell if the recommendations are working or not. Three independent and carefully controlled experiments uncovered biases that marketers should keep in mind. If the results you are getting smack of the five problems mentioned above (In-Market/ Targeting Bias, Cheap Inventory Bias, Short Term Digital Signal Bias, Missing Leading Indicators, Myopic Measurement Design), treat the results with skepticism and seek a second opinion.

When getting that second opinion, make sure you are calculating the incremental contribution of each touch. For this, you need a control group, or a heavier/lighter exposure design. One may choose to do a comprehensive design of experiments across all media, as the insurance company did, or one may choose to concentrate the control groups on the digital ad inventory as a less complex, and less expensive spot check, as TubeMogul did. Either way, if the marketer is spending more than $10 million a year on advertising, the value of independent validation is well worth the investment.

End Notes:
[1] "Your Attribution Model is Broken" by Tom Riordan, Dec. 15, 2015: http://adexchanger.c 1 om/data-driven-thinking/your-attribution-model-is-broken/
[2] 1997 IAB Advertising Effectiveness Study and VOYAGER