International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

The Self-Correcting Budget: Removing an Allocator's Own Footprint from Attribution

Sameer Siddiqui

Manager Data Science, Product and Platforms, Accenture

Abstract

Automated advertising systems increasingly close the loop between measurement and action: an attribution model estimates channel contribution, and an optimizer reallocates budget toward the apparent contributors. When such a loop operates on raw attribution, it confounds the consequences of its own reallocation with the underlying merit of the channels it evaluates. A channel that harvests pre-existing demand appears more efficient precisely because the system spends into it, producing a self-reinforcing bias toward the least incremental channels. We characterize this failure mode from first principles using event-level data, and evaluate a mechanism that treats each reallocation as a randomized experiment the system performs on itself: a within-channel holdout supplies an incrementality estimate, and the debiased signal—rather than raw attribution—drives the next allocation. In a controlled simulation with known ground truth across forty randomized environments, the debiased loop reduces efficiency-estimation error, avoids the adversarial "trap" channel, and—most importantly—retains a near-constant share of optimal return as confounding strengthens, whereas the raw-signal loop degrades monotonically. We report magnitudes as illustrative of the mechanism rather than as benchmarks from live systems, and delineate the boundary between established incrementality measurement and the novel closure it enables.

Keywords: Marketing Attribution; Incrementality; Causal Inference; Budget Allocation; Closed-Loop Control; Confounding

1. Introduction

Modern advertising systems increasingly close the loop between measurement and action. An attribution model estimates what each channel contributed, and an optimizer moves budget toward the contributors. Performed in real time by autonomous agents, this promises faster and better allocation. It also introduces a failure mode that is easy to overlook and costly to encounter: a system that confounds the consequences of its own decisions with the merit of the channels it evaluates.

This article makes three contributions. First, we explain from raw event data how attribution and true causal efficiency diverge (Section 2). Second, we show that a closed loop ignoring this divergence drifts toward its least incremental channels (Section 4). Third, we evaluate a corrective mechanism—treating each reallocation as an experiment the system runs on itself, and subtracting that experiment's effect before feeding the signal forward (Section 5). All claims are grounded in a simulation with known ground truth, so estimates can be validated against parameters fixed in advance.

2. Attribution versus efficiency

The terms are frequently conflated. We distinguish them precisely, as the gap between them motivates the

entire analysis. Attribution denotes the conversions a tracking system credits to a channel; it counts sales that follow an advertisement, whether or not the advertisement caused them. Incremental efficiency denotes the conversions per unit spend that would not have occurred absent the advertisement; it is causal, and it is the quantity an allocator should maximize.

The distinction is visible only below the level of aggregate dashboards. Consider one channel—retargeting on a social platform, a canonical case. The event log records one row per user event: the user reached, whether an advertisement was shown, its cost, and whether a purchase followed. A randomized holdout withholds advertising from a slice of the audience, providing a control group.

Table 1. Representative rows from one day's event log. Highlighted rows are holdout users, shown no advertisement.

Time	User	Ad	Purch.
09:00	u_10000	yes	—
09:03	u_10003	yes	✓
09:04	u_10004	no	—
09:08	u_10008	no	✓
09:13	u_10013	yes	✓

The highlighted holdout rows carry the central idea. If users who saw no advertisement nonetheless purchase at some rate, then a corresponding fraction of exposed purchasers would also have purchased regardless. The holdout estimates that baseline rate directly.

2.1 From event rows to incremental lift

Grouping the day's log into treated and holdout arms and counting: among a treated audience of approximately 170,000 users, 7,367 purchased; among approximately 30,000 holdout users, 3.02% purchased without exposure. Applying the baseline rate to the treated population yields the conversions expected absent advertising, and the remainder is incremental:

$$\text{incremental} = 7{,}367 - (0.0302 \times 170{,}184) = \mathbf{2{,}230} \quad (1)$$

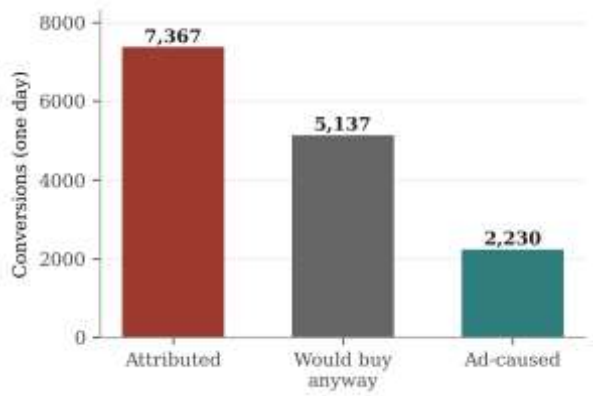

Fig. 1. Last-click attribution credits 7,367 conversions; a holdout reveals only 2,230 as ad-caused. The remainder is demand the channel harvested but did not create.

Efficiency at this spend level is a single quotient: 2,230 incremental conversions per $2,042 spent, or roughly 1.1 conversions per dollar. Raw attribution reports over three times that value, misdirecting any optimizer that consumes it.

3. Saturation and the response curve

Efficiency is not a fixed channel property; it declines with spend, as the most persuadable users are reached first. This decline—saturation—is invisible in a single day. It emerges only when the holdout experiment is repeated across spend levels and the incremental return is observed to bend.

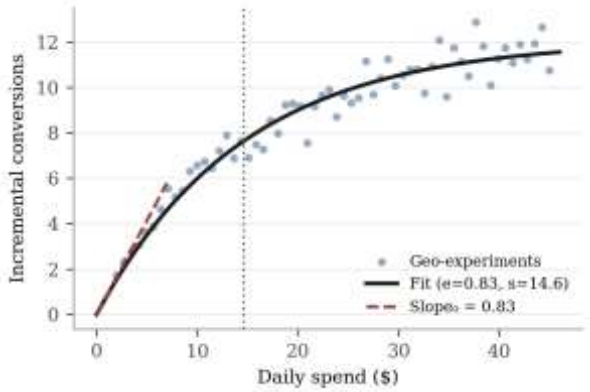

Fig. 2. Each point is a daily holdout experiment at a distinct spend level. The fitted response curve recovers peak efficiency (slope at origin, ≈ 0.83) and saturation (the bend, near $15/day).

Efficiency and saturation are precisely the parameters a marketing-mix model estimates, and the inputs an allocator requires. Both are properties of the incremental signal; estimated from raw attribution instead, every demand-harvesting channel appears more efficient and less saturated than it is.

A practical caveat governs estimation. Incremental conversions are a small fraction of baseline, so day-to-day baseline noise can exceed the entire signal. A naive fit on few small daily tests produced an efficiency estimate five times too high in our runs. Reliable estimation requires large randomized holdouts, many experiments, and pooling—making **holdout sizing a first-class design parameter** for any system dependent on the debiased signal.

4. Self-reinforcement in a naive loop

We place an autonomous allocator atop this measurement. Each cycle it reads a signal, shifts budget toward apparently efficient channels, and repeats. The question is the behavior induced when the signal is raw attribution rather than incremental lift.

We simulate five channels with controlled response curves and confounding, so ground truth is known. One channel is adversarial by construction: least efficient in truth, yet strongest at harvesting demand it did not create—the profile of branded search or aggressive retargeting. We term it the trap.

Table 2. Simulated channels. Efficiency and saturation are the planted true values; the trap combines the lowest efficiency with the strongest confounding.

Channel	Eff.	Sat.	Confound
Branded search	0.50	$12	high

Channel	Eff.	Sat.	Confound
Meta retargeting	0.92	$13	low
TikTok prosp.	2.16	$8	low
YouTube prosp.	1.68	$16	low
Display	0.61	$8	low

Once converged, the loops' beliefs about efficiency can be compared against the planted truth. The raw-signal loop does not merely add noise; it inverts the ranking (Fig. 3), ranking the least efficient channel as the most efficient.

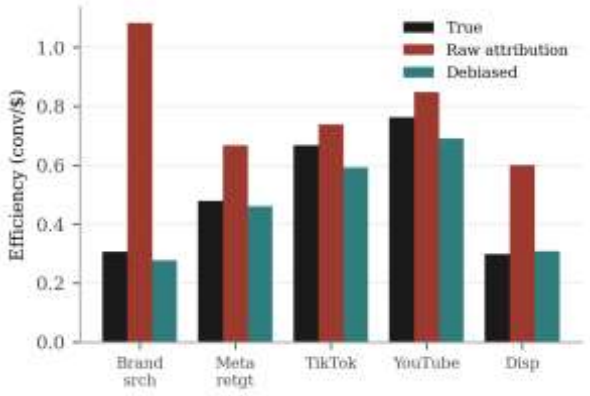

Fig. 3. Estimated versus true efficiency by channel at convergence. The debiased estimate tracks truth; raw attribution over-credits uniformly and, at the trap, inverts the ranking.

Because the trap appears excellent under raw attribution, the naive loop continues to fund it; because funding it manufactures further harvested conversions, it continues to appear excellent. The loop pursues its own tail (Fig. 4).

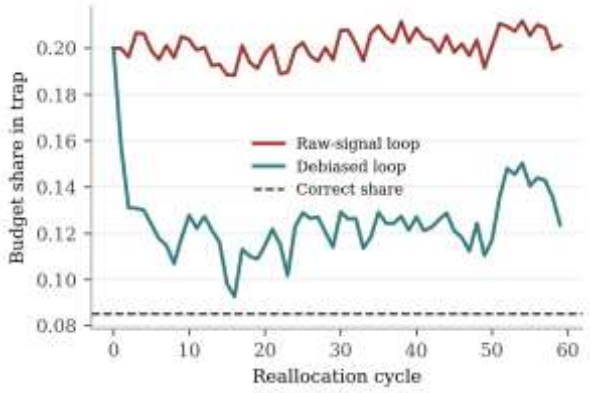

Fig. 4. Budget share directed to the trap over 60 cycles. The debiased loop converges to the level warranted by true efficiency; the raw-signal loop does not learn, as its own spend sustains the confounded signal.

5. Mechanism: subtracting the system's own effect

The remedy follows from the diagnosis. If the system confuses its own effect with channel merit, it should measure and remove that effect before acting. Concretely, the loop maintains a randomized within-channel

holdout and forms its per-cycle signal from the exposed–holdout difference rather than exposed conversions alone:

$$s^{debiased} = \text{exposed} - \text{holdout\_baseline} \quad (2)$$

All else is identical between the loops: channels, budget, and update rule. Only the driving signal differs, and that single substitution separates the two trajectories of Fig. 4. In a deployed system it is accompanied by the machinery Section 3 requires: adequately sized holdouts, cross-cycle pooling of the lift estimate to control variance, and a record of each reallocation permitting reconstruction of its effect.

5.1 Cost and conditions of benefit

Debiasing is not costless. A within-channel holdout parks measurable spend, and the subtraction introduces variance requiring smoothing. Where channels differ little in efficiency, or confounding is mild, the raw loop is already near-optimal and the correction yields little. The benefit concentrates where the hazard is greatest.

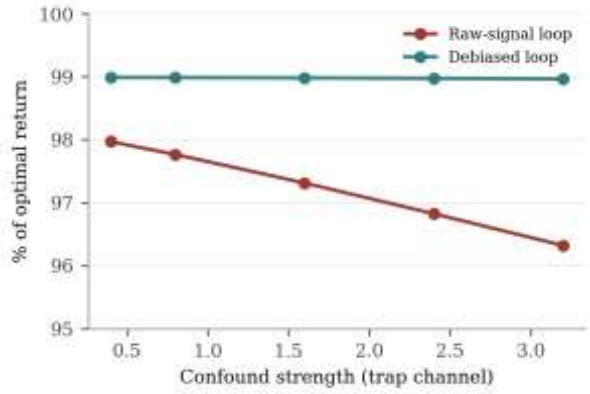

Fig. 5. Fraction of optimal return captured as trap confounding increases. The raw-signal loop degrades monotonically; the debiased loop is invariant, its immunity independent of confound strength.

Across forty randomized environments, the debiased loop recovered true efficiency substantially more accurately, parked far less budget in the trap, and produced higher genuine return in the large majority of cases. The reward gap was modest where response curves were flat near their optimum and widened with confounding, consistent with Fig. 5. The mechanism promises not a fixed lift but invariance: the loop cannot be misled by its own spending, a guarantee most valuable where a channel is most misleading.

6. Scope and relation to prior work

The measurement employed here—holdout-based incrementality and response-curve fitting—is established. Marketing-mix models fit saturation curves [1], [2]; incrementality platforms correct attribution against control groups [3]; and off-policy methods in auto-bidding address an agent's confounding by its own historical actions [4]. None of this is claimed as novel.

The contribution is the closure: wiring a debiased, self-referential estimate back into an autonomous, cross-channel allocation loop, so the system controls for the effect of its own reallocation rather than a single campaign's exposure. The governing observation—that an optimizer consuming raw attribution optimizes partly against its own shadow—is a property of the loop, not of any single measurement.

7. Limitations and future work

Our evidence is simulation with planted ground truth; reported magnitudes illustrate the mechanism and

are not benchmarks from a live account. The debiased signal is only as reliable as the incrementality estimate feeding it, so an undersized holdout degrades the loop. The natural next step is validation of the debiased estimate against a real geo-level holdout on live campaigns prior to any operational reliance.

8. Conclusion

An always-on automated optimizer is attractive, but a loop that reads its own footprints as signal will circle, and circle fastest toward the channels that flatter it. Measuring that footprint and subtracting it is a small change in implementation and a large change in behavior—the difference between a system that learns and one that merely reinforces.

References

1. Jin, Y., Wang, Y., Sun, Y., Chan, D., & Koehler, J. Bayesian methods for media mix modeling with carryover and shape effects. Technical report, 2017.
2. Chan, D., & Perry, M. Challenges and opportunities in media mix modeling. Technical report, 2017.
3. Gordon, B. R., Zettelmeyer, F., Bhargava, N., & Chapsky, D. A comparison of approaches to advertising measurement. Marketing Science, 38(2), 193–225, 2019.
4. Bottou, L., et al. Counterfactual reasoning and learning systems: the example of computational advertising. Journal of Machine Learning Research, 14, 3207–3260, 2013.
5. Lewis, R. A., & Rao, J. M. The unfavorable economics of measuring the returns to advertising. Quarterly Journal of Economics, 130(4), 1941–1973, 2015.
6. **Reproducibility & disclosure.** All figures derive from a single simulation with fixed random seeds and known ground-truth parameters; each result is regenerable and auditable. This article is a technical explainer and does not constitute legal, financial, or investment advice.