Pearmill

A Practitioner's Guide to Paid Search Incrementality

How to measure what your ads actually cause and what to do with the answer

By Het Mavani
Paid Search Lead @ Pearmill

A Note From the Author

Over the years our team has been inside a lot of Google Ads accounts, fintech giants, scrappy DTC brands, marketplaces, and almost every time we sit down with a leadership team, the same quiet worry is in the room: *are these conversions even real?*

I wrote this because that question deserves a real answer, not just a reassuring one.

This report is built off of a year long incrementality program, run on a single DTC account, with four independent experiments. Some of what we found was uncomfortable to deliver. All of it made the strategy conversation better.

This is the guide I wish we'd had when we started. Here's what you'll get:

- Why platform attribution measures correlation, not causation
- Four independent measurement methods, tested on one account
- The counterintuitive finding about lowering bids
- How to turn your incrementality number into a bid strategy
- Where incrementality fits alongside MTA and MMM

Hope you enjoy the methodology and the numbers. If you ever want to chat, please don't hesitate to reach out!

Het Mavani
Paid Search Lead, Pearmill
het@pearmill.com

Table of Contents

01 The Question Every Performance Team Is Afraid to Ask

Our team at Pearmill has seen a lot of Google Ads accounts over the years: fintech giants, scrappy DTC brands, marketplaces, you name it. And almost every time we sit down with a leadership team, it feels like the same quiet worry is in the room, even when no one wants to say it out loud: *are these conversions even real?*

We get why that question is uncomfortable. Most marketers are stretched thin, the dashboard says the CPA looks good, and the pressure to keep spend efficient is relentless. It's easier to take Google's numbers at face value and keep moving. We've been there, both as operators and as a team helping clients out of it.

But here's the truth our paid search team keeps coming back to: every dollar in performance marketing carries a question that should keep you up at night. *Did the ad actually cause this conversion, or would the customer have come anyway?*

> ***That is the question of incrementality.*** *It's arguably the most important question in performance marketing today, and it's also one of the hardest to face honestly.*

02 What Platforms Tell You vs. What's True

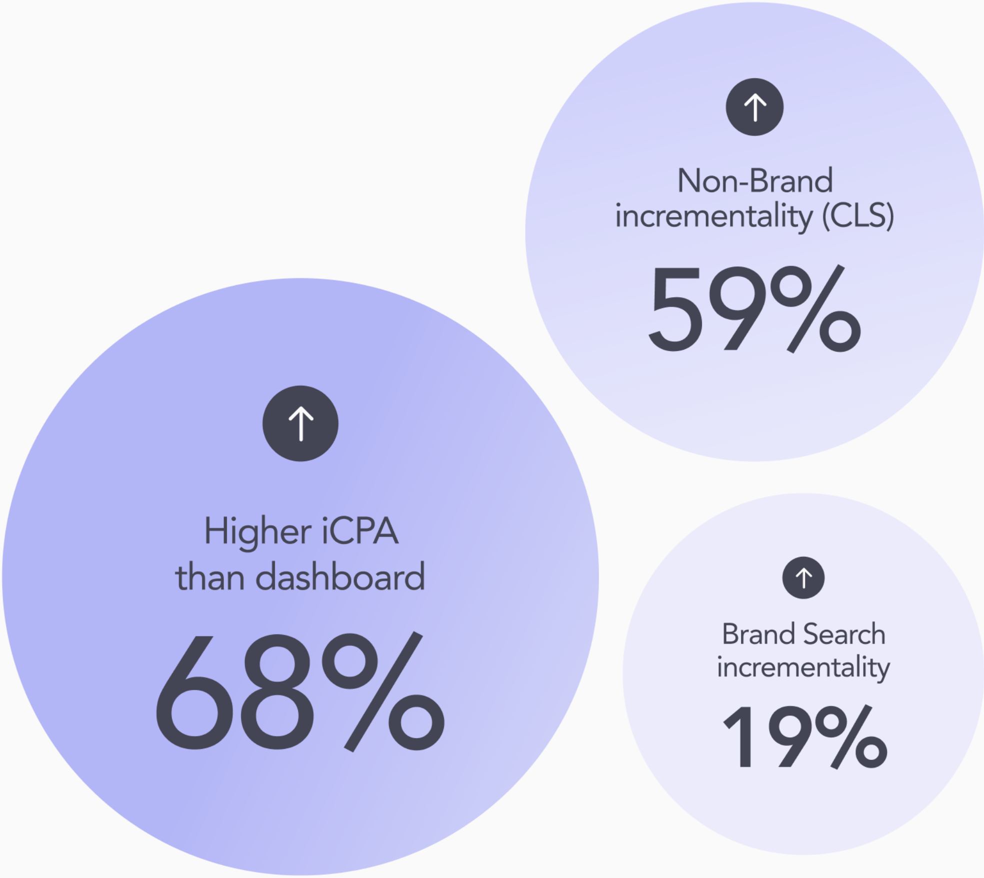
Google's attribution, whether you're on last-click or data-driven, measures **correlation, not causation**. It tells you which ad a user touched before converting. It doesn't tell you whether the ad changed their behavior. That's not Google being sneaky. It's just what platform attribution is built to do.

The gap between attributed performance and incremental performance is rarely small, and that's the part that catches teams off guard. One of our DTC clients recently ran a year-long incrementality program on their own paid search account, with four independent experiments. The answer wasn't subtle, and honestly, it wasn't an easy conversation to walk into.

$$\text{Incrementality Factor} = \text{Absolute Lift} / \text{Attributed Conversions}$$

If that number is 60%, then for every 100 conversions Google claims credit for, only 60 actually wouldn't have happened without the ad. The other 40 were going to convert through organic, direct, brand recall, or another channel anyway. You're paying retail for inventory you already owned.

The companion metric is **iCPA (Incremental CPA)**, calculated as Total Spend / Incremental Conversions. It's always higher than what the platform reports, sometimes meaningfully so. For our DTC client, the iCPA on Non-Brand Search came in roughly 68% higher than what Google was showing in the dashboard. That's not a rounding error. That's the difference between scaling and burning.



03 The Four Ways to Actually Measure It

Most teams pick one methodology and call it done. We understand the impulse. Measurement is hard, and any single test feels like a step forward. But each method has blind spots, and the picture gets sharper when you triangulate. Here's how we think about each, with what one of our DTC clients learned running all four on the same account.

1. Natural Experiments: The "Oh No" Pause

Sometimes a campaign accidentally pauses for a day. Sometimes a budget caps out. Sometimes a feed breaks. These moments feel awful in real time, and we've sat with plenty of teams in that "what just happened" panic. But they're actually gifts, even if they don't feel like it in the moment.

Spend drops to zero, and you get to watch whether conversions follow. If conversions hold steady when the ads go dark, the ads probably weren't doing as much work as they got credit for.

For our DTC client, a one-day accidental pause on Non-Brand Search produced one of the most striking pieces of evidence in the entire program. Total Google Search spend dropped sharply, and conversions barely moved. Even more telling, Facebook volume, CVR, and CPA all improved during the pause. Google had been quietly claiming credit for conversions Facebook was actually driving.

> ***Strengths: zero cost, real business-level signal.***
> ***Limitations: uncontrolled, single-day. Directional, not definitive.***

2. Conversion Lift Studies (CLS): User-Level Randomization

This is Google's in-platform randomized controlled trial. They split your audience into a test group that sees your ads and a holdout that's prevented from seeing them. The difference in conversion rates is the causal lift. It's the cleanest test most teams have available, and our team leans on it heavily as a starting point.

Conversion lift – how it works

MEASUREMENT · GOOGLE CLS

Exposed group

Sees ads on Google's platforms during the test window.

Control group

Held out – no exposure. Establishes the organic baseline.

INCREMENTAL LIFT

Exposed – Control

Once the test period starts, conversion behavior between test and control groups is analyzed to understand the impact ads on Google's platforms have on conversion volume – isolating the **incremental impact of advertising.**

For Non-Brand Search + PMax, the CLS landed on a **59% incrementality factor**, with a measured conversion lift of **182%** at 100% statistical confidence. Strong, real, and very much worth investing behind.

For Brand Search, it was a very different story. Just **19% incrementality**. Roughly 4 out of every 5 conversions Google attributed to branded ads would have come anyway, through organic search, direct, or another channel.

That 19% lines up with Google's own research: branded search is only ~50% incremental when a strong organic result exists, and 100% when one doesn't. A strong brand with strong organic presence will see lower brand search incrementality, not higher. Counterintuitive, but consistent.

> ***On competitor conquesting:*** *If rivals are actively bidding on your brand terms, the math shifts. Their ads can siphon off branded traffic that would otherwise have flowed to your organic result, which raises the defensive value of your own brand search ads. Before assuming your brand search spend is wasted, pull the auction insights report. Brand search is your center-back — it doesn't score, but it stops the other team from doing it on your turf.*

The CLS also surfaced upper-funnel lift: 42% on Add to Cart, 26% on Lead Form Submission, 10% on Membership Opt-In.

> ***Strengths: statistically rigorous, user-level randomization, Google-managed.***
> ***Limitations: platform-specific view, short measurement windows.***

3. Customer Surveys: Useful, But Biased

You email recent converters and ask them: had you heard of us before clicking? Would you have signed up without the ad? It's cheap, qualitative, and easy to scale, and our team likes having it in the mix as a sanity check.

Our DTC client's survey came back with a self-reported incrementality of 76%, the highest number in the program, and almost certainly the least reliable.

People are just genuinely bad at predicting their own counterfactual behavior. We underestimate how much advertising shapes our choices. Surveys are great for where customers first heard about you. In this case, the majority of brand-aware respondents first encountered the brand on TV, through word of mouth, or on Facebook. Not search.

> ***Strengths: cheap, useful for cross-channel discovery texture.***
> ***Limitations: self-reporting bias inflates the answer. Weight accordingly.***

4. Geo-Based Holdout Tests: Business-Level Truth

This is where it gets serious, and we won't pretend it's easy to set up. You group DMAs into matched cells based on historical conversion volume and trends, then vary spend across cells: one cell at BAU, one at reduced bids, one with spend cut entirely. The difference in total (not just attributed) conversions across cells reveals the causal impact.

Validation matters. You want your test and control groups balanced on conversions, revenue, active users, and daily trend correlation before you ever start. Skip this step and you're measuring noise.

For our DTC client, the geo test ran across multiple matched DMA groups split into three cells. The headline results:

• BAU cell incrementality: **~48%** (directionally consistent with the 59% CLS)
• Lower-bid cell (a 40% bid reduction): incrementality dropped to **~23%**

The counterintuitive finding: *Lowering bids, the thing most teams reach for when efficiency slips, reduced incrementality. At lower bids, Google's algorithm shifts toward users whose intent is already high — people who were planning to find you anyway. You haven't gotten more efficient. You've just bought more passengers.*

> ***Strengths: measures business-level conversions including organic, captures cross-channel effects, longer windows, more stable reads. Limitations: geo matching is hard, external noise creeps in, real opportunity cost in the holdout cells.***

04 Triangulating the Answer

Four methods, one account, one channel. Here's the range we landed in:

Method	Implied Incrementality
Natural experiment (single-day pause)	~0–20% (directional)
Conversion Lift Study	~59%
Customer survey	~76% (self-report inflated)
Geo holdout, BAU cell	~48%
Geo holdout, lower bids cell	~23%

The CLS is the most statistically rigorous and gets the most weight from our team. The geo test validates it at the business level. The survey overstates, as surveys do. The natural experiment is directional but compelling.

> *Our honest read on Non-Brand Search for this account: roughly **6 in 10** conversions are real. On Brand Search, it's closer to **2 in 10**. Those numbers were hard for the client to hear at first, and they were hard for us to deliver. But once everyone was looking at the same picture, the strategy conversation finally moved forward, and that's the part that matters.*

05 What You Do With This

This is the part most teams skip. It's the only part that matters. It's also where our paid search team spends the bulk of the work, because measurement only earns its keep when it changes a decision.

Recalibrate Your Bids

Once you've landed on an agreed-upon incrementality number (triangulated across your CLS, geo test, and other reads), you can turn it into a calibration factor and plug it straight into your bidding.

BIDDING · CALIBRATION

We adjust how we bid in-platform using a **calibration factor** to get closer to your iROAS goals.

EXAMPLE

1 STEP 1

Get the calibration factor

Absolute lift / Attributed conversions = Incrementality factor

2 STEP 2

Adjust in-UI bids to reach the efficiency goal

iCPA goal X Incrementality factor = tCPA

*If your true breakeven iCPA is $X and incrementality is 59%, your tCPA in Google Ads should sit at roughly **0.59 x $X**. That's the bid that delivers a profitable incremental customer, not a profitable platform-reported one.*

For Brand Search at 19% incrementality, the math gets even tighter. Most accounts are dramatically over-bidding on brand and paying Google to capture demand they already own.

Stop Treating All Conversions as Equal

The aggregate factor hides variance underneath. Customers who came in via Google last-click and said "Google" on a "How did you hear about us?" survey are highly incremental. Customers who came in via Google last-click but said "TV" are barely incremental at all. Bucket conversions by these signals in your internal attribution and weight them accordingly.

Don't Assume Incrementality Is Static

Incrementality moves with your bids, your budgets, your audiences, and your creative. A 40% bid reduction in the client's geo test cut incrementality nearly in half. Whenever you change a major lever, remeasure. Otherwise you're optimizing on a number from a state of the account that no longer exists.

Even when nothing major has changed on your end, incrementality still drifts. Competitor activity, seasonality, organic SERP changes, audience saturation, and platform algorithm updates all quietly move the number underneath you. That's why we recommend a recurring testing cadence. For most accounts, a quarterly CLS on your largest spend channels plus an annual geo holdout is a healthy baseline.

Use Multiple Methods

No single test is definitive. The conviction in our DTC client's program didn't come from any one experiment. It came from four very different methodologies pointing in the same direction. When your CLS, your geo test, and your accidental pause all agree, you can act with confidence. When they disagree, that's a signal too — and our team treats it as a prompt to dig deeper rather than pick a favorite.

06 Where This Fits in Your Measurement Stack

MTA is fast and tactical but measures correlation. MMM is broad and privacy-safe but slow and expensive. Incrementality is the only one that gives you causal ground truth, and you can use it to recalibrate the other two.

	MTA	MMM	Incrementality
Speed	Real-time	Quarterly+	Test duration
Signal	Correlation	Statistical model	Causal
Scope	Digital touchpoints	All channels	Per-channel
Privacy	Requires user data	Privacy-safe	Varies by method
Role	Tactical optimization	Budget allocation	Calibration layer

**Incrementality doesn't replace MTA or MMM.
It calibrates them.**

Per BCG and Google's research on 2,135 global companies, only **46%** of organizations actually use all three together, despite individual adoption rates of around 80% for each. Only **9%** of companies consider their measurement AI capabilities "leading."

We don't share those numbers to shame anyone. We share them because most teams we meet are doing their best with the tools they have, and the gap is structural, not personal.

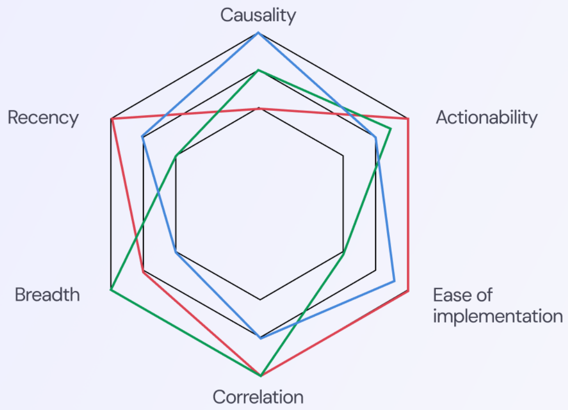
Measurement approaches by strength.

MEASUREMENT • COMPARISON

LEGEND

Each approach trades off causal rigor, breadth and how quickly it acts on results.

- **MMM** Marketing mix modeling
- **Incrementality** Geo & conversion lift tests
- **Attribution** Platform & multi-touch



07 The Top-Tier Mindset

Here's what our team tells every client at the start of an engagement: the platform dashboard is a measurement layer, not the answer. Google can tell you what it observed. It cannot tell you what it caused. Those are different questions, and you need different instruments to answer them.

The teams that win don't ask whether to do incrementality testing. They ask how often, with which methods, and what to do with the answer. They expect Brand Search to look worse under the microscope. They expect Non-Brand to look more expensive than the dashboard implies. They build their bid strategy, their attribution model, and their channel mix around the actual answer, not the convenient one.

That's how you stop fighting the algorithm and start architecting around it. And it's the work our paid search team at Pearmill genuinely loves doing alongside clients who are ready to look at the numbers honestly.

Ready to find your real number?

Our paid search team runs incrementality programs like the one in this guide alongside clients who are ready to look at the numbers honestly. If you want to know what your ads are actually causing (and build a bid strategy around the real answer), cwe'd love to talk.

Book a call with our team

paidsearch@pearmill.com