GROWTH THAT COMPOUNDS

Marketing mix modeling is a calibration problem

Marketing mix modeling can organise an aggregate budget, but it cannot certify causality by itself. Calibrate the model before trusting the allocation.

Sinan Isoglu, MBA (Quantic)
Commercial growth leader, lecturer and doctoral researcher

PUBLISHED
20 August 2026

READING TIME
5 min

WORDS
1,163

https://isoglu.com/insights/journal/marketing-mix-modeling-is-a-calibration-problem/

CONTENTS

MANAGEMENT SUMMARY

Marketing mix modeling is often asked to do two jobs at once: describe how aggregate demand moved and prove which channel caused the movement. Those are different jobs. A model can be useful for planning while remaining vulnerable to seasonality, targeting, spend response and the choice of transformations. The practical question is therefore not whether to use marketing mix modeling or an experiment as if they were rival religions. It is what each instrument can identify, how the model is calibrated, and which decision is allowed to rely on the output. This article compares attribution, aggregate modeling and holdout tests, then turns the comparison into a calibration log that a commercial team can maintain without pretending the model has seen a counterfactual.

Keywords: Marketing mix modeling · Incrementality testing · Advertising measurement · Budget allocation

Marketing mix modeling is a calibration problem, not a dashboard problem. It can organise an aggregate budget decision, show how demand and spend moved together, and expose a set of scenarios. It cannot certify causal lift just because the regression is sophisticated.

The distinction is easy to lose because every instrument returns a number. A multi-touch attribution model assigns credit to observed touches. A marketing mix model relates aggregate outcomes to aggregate inputs. An incrementality test asks what changed relative to a counterfactual. The numbers look comparable only after the question has been silently changed.

Which three distinct measurement questions do attribution, experimentation, and MMM answer?

The first discipline is to name the output before choosing the model. If the decision is whether a channel created demand that would not otherwise have existed, the model must contain a credible counterfactual. If the decision is how to plan a portfolio under several demand scenarios, an aggregate model may be the right instrument even when it cannot identify causal lift on its own.

Table 1 What each measurement instrument can carry

Instrument	What it observes	Strongest defensible output	Invalid leap
Touchpoint attribution	Recorded exposures and conversions	Which touches are present in the observed path	Those touches caused the conversion
Marketing mix modeling	Aggregate outcomes, spend and controls over time	A scenario relationship under stated model assumptions	The coefficient is experimental lift
Holdout or lift test	Treated and untreated units under a designed intervention	Incremental outcome under the tested conditions	The result transfers unchanged to every channel
Calibrated portfolio view	Model scenarios plus experiments and operating bounds	A decision with a stated confidence and downside	One dashboard is the truth

Author's synthesis of Gordon et al. (2019), Blake, Nosko & Tadelis (2015), Lewis & Rao (2015), and Johnson, Lewis & Nubbemeyer (2017).

This is the narrower claim behind the incrementality illusion. The problem is not that an observed association is useless. It is that an organisation starts making a counterfactual decision from an instrument that never observed the counterfactual.

Why do marketing mix models require continuous experimental calibration against lift tests?

An aggregate model sees time. Time carries seasonality, promotions, product launches, macroeconomic shocks and the company’s own decision to spend more when demand looks promising. The model must separate those movements with data and assumptions that are never perfectly observed.

The transformations matter too. Adstock, saturation, lag and interaction terms are not decoration. They are claims about how exposure becomes demand. Different assumptions can produce different allocations while fitting the past in a similar way.

Calibration is the discipline that keeps this from becoming a debate about whose chart looks more plausible. It uses experiments, external shocks, stable control series, known pricing changes or other observations to constrain the model. It does not make the model experimental. It tells the reader which parts of the model have an anchor outside the model.

Gordon and colleagues show why this matters when different advertising measurement approaches are compared against field experiments. Blake, Nosko and Tadelis show how paid search can capture inframarginal buyers. Lewis and Rao show how noisy outcomes make return measurement expensive. The lesson is not to discard every model. It is to stop treating model fit as proof of lift.

Table 2 The calibration log

Model element	Assumption	External anchor	If the anchor moves
Baseline demand	The non-media demand path has this shape	Control series, category data or known interruption	Re-estimate the base before reallocating
Carryover	Exposure persists for this long	Delayed response in an experiment or prior with a stated basis	Widen the scenario range
Saturation	Additional spend produces less response after this point	Spend variation with an independent shock	Do not use the point estimate as a ceiling
Channel interaction	Two channels reinforce or substitute	Designed test or a documented mechanism	Keep the interaction as a scenario, not a fact
Incremental lift	The allocation reflects causal contribution	Holdout, ghost-ad or lift result	Mark the model coefficient as uncalibrated

Author's worksheet for documenting what enters a marketing mix model and what anchors it outside the model.

How can commercial leadership apply MMM coefficients without econometric overreach?

Start with the portfolio question. Which channels need a directional bound? Which channels have enough scale for a designed holdout? Which decisions are defensive and should be governed by a budget ceiling rather than a claim of incremental demand?

Then show the range. A model that produces one allocation without a sensitivity view has hidden its uncertainty. Vary the baseline, the carryover and the saturation assumptions. If the recommended portfolio changes when a reasonable assumption moves, the decision is sensitive to the assumption. That is an output worth knowing.

The next step is to connect the model to a test. A holdout can calibrate a channel or a family of channels. It may not transfer to every market or period, but it gives the aggregate model an external point. The evidence over anecdote worksheet is useful here: write down what would count as evidence before the allocation is shown.

Finally, separate the model's scenario from the operating decision. The model can say that a portfolio looks better under one set of assumptions. Leadership still has to decide how much uncertainty it is willing to buy, what cash payback is acceptable and which learning matters more than this quarter's apparent efficiency.

Why is executive econometric governance distinct from dashboard monitoring?

The dashboard is a view. Governance is the rule that says what happens when the view conflicts with a test, a margin threshold or a customer constraint. The funnel bottleneck nobody's measuring shows the same pattern outside advertising: the visible metric is often the output, while the unmeasured interval decides whether the output can be trusted.

A marketing mix model is worth keeping when it makes assumptions explicit, supports scenarios, accepts outside calibration and changes a decision. It is not worth keeping as a decorative source of precision. The right question at the next review is not "What did the model say?" It is "Which part of the model has been tested, which part is a scenario, and what decision is allowed to rely on each?"

Gordon, B. R., Zettelmeyer, F., Bhargava, N., & Chapsky, D. (2019). A comparison of approaches to advertising measurement: Evidence from big field experiments at Facebook. Marketing Science, 38(2), 193–225. https://doi.org/10.1287/mksc.2018.1135

Blake, T., Nosko, C., & Tadelis, S. (2015). Consumer heterogeneity and paid search effectiveness: A large-scale field experiment. Econometrica, 83(1), 155–174. https://doi.org/10.3982/ECTA12423

Lewis, R. A., & Rao, J. M. (2015). The unfavorable economics of measuring the returns to advertising. The Quarterly Journal of Economics, 130(4), 1941–1973. https://doi.org/10.1093/qje/qjv023

Johnson, G. A., Lewis, R. A., & Nubbemeyer, E. I. (2017). Ghost ads: Improving the economics of measuring online ad effectiveness. Journal of Marketing Research, 54(6), 867–885. https://doi.org/10.1509/jmr.15.0297

END OF ESSAY

HOW TO CITE THIS ESSAY

Isoglu, S. (2026, August 20). Marketing mix modeling is a calibration problem. Sinan Isoglu. https://isoglu.com/insights/journal/marketing-mix-modeling-is-a-calibration-problem/

KEEP READING

Growth that compounds

A marketing attribution model needs a counterfactual

https://isoglu.com/insights/journal/marketing-attribution-model-needs-a-counterfactual/

Growth that compounds

What is marketing efficiency ratio (MER)? A blended signal, not a causal answer

https://isoglu.com/insights/journal/what-is-marketing-efficiency-ratio/

Growth that compounds

One customer model cannot predict every outcome

https://isoglu.com/insights/journal/one-customer-model-cannot-predict-every-outcome/

Sinan Isoglu, MBA (Quantic)
Commercial growth leader, lecturer and doctoral researcher

Online version: https://isoglu.com/insights/journal/marketing-mix-modeling-is-a-calibration-problem/