

The Reconciliation Layer: Reconciling Marketing Mix Model Estimates with Financial Actuals

Karan Dhir

301 W Washington Avenue #501, Sunnyvale, California, USA

ARTICLE INFO

Received: 03 July 2026

Revised: 08 Aug 2026

Accepted: 20 Aug 2026

ABSTRACT

Privacy rules, platform controls, and identifier loss have weakened user-level attribution. Marketing mix modeling (MMM) has therefore regained attention as a strategic measurement approach. Enterprise use can stall when model estimates meet financial reporting. In this case, the MMM estimates incremental net sales against a modeled counterfactual. Finance reports booked outcomes that may include accruals, reserves, allocations, and restatements. This paper develops and examines a Reconciliation Layer in an 18-month participatory case at a large U.S. biopharmaceutical enterprise. The layer has four parts: aligned definitions, a decomposition bridge with a disclosed residual, experimental checks, and governance tied to the financial calendar. The documentary record includes six quarterly release cycles, four governance versions, a 13-item defect register, and a reconciliation workbook covering 16 brands. The case identified four recurring reconciliation problems: source selection, timing, scope, and silent upstream changes. During implementation, the enterprise adopted an intake gate, a data-drop tracker, cutoff rules, and a bridge workbook. Finance also joined hierarchy mapping and long-term source-path work. These observations are descriptive. The case does not establish causal effects on allocation decisions or firm performance.

Keywords: marketing mix modeling; marketing accountability; return on marketing investment; incrementality testing; marketing-finance interface; measurement governance

1. INTRODUCTION

Digital advertising measurement has changed. Third-party identifiers have declined, platform privacy controls have expanded, and more media activity sits inside walled gardens. Those shifts weaken the user-level tracking used by multi-touch attribution (Chan & Perry, 2017; Gordon et al., 2019). They have renewed interest in marketing mix modeling. MMM is an aggregate econometric method rooted in decision calculus (Little, 1970) and market-response models (Hanssens et al., 2001). Bayesian specifications can represent carryover and saturation directly (Jin et al., 2017). Open-source frameworks such as Meridian and Robyn have also made implementation more accessible within their respective model classes (Google, 2026; Meta Platforms, 2024).

MMM still faces an older operating problem: finance needs to trace what a marketing number means and how it relates to reported results. Prior studies link marketing's standing to financial accountability (O'Sullivan & Abela, 2007; Verhoef & Leeflang, 2009). Other work shows that metric use depends on management context and governance (Homburg et al., 2012; Mintz & Currim, 2013). The CMO Survey continues to report difficulty demonstrating marketing impact quantitatively (Moorman, 2024). These studies define the accountability problem. They do not specify how to reconcile a model-based causal estimate with a finance reporting process.

In this case, the model estimates incremental net sales attributable to marketing relative to a modeled counterfactual. Finance reports net sales under its controlling profit-and-loss hierarchy. Those reported values can include accounting estimates, allocations, reserves, and later adjustments. The two quantities therefore require definition before comparison. A model estimate has no direct ledger entry. Finance can still review source data, scope, timing, hierarchy, calculation steps, and any unexplained difference. The case records show requests for source designation, hierarchy alignment, restatement policy, allocation timing, and file-level change control before recurring governance review. The records summarized here do not show that reconciliation alone changed a budget decision.

Prior work recommends expressing marketing results in financial terms (Stewart, 2009; Srinivasan & Hanssens, 2009). It also documents conditions under which measurement systems enter decisions (Homburg et al., 2012; Mintz & Currim, 2013). In the literature reviewed for this paper, no recurring operating mechanism was identified for aligning a counterfactual MMM estimate with finance actuals. The paper treats reconciliation as a design problem. It defines the quantities, maps scope and time, traces adjustments, discloses the remainder, and sets a recurring governance cadence.

Three research questions guide the case. RQ1 asks which reconciliation failures appear when MMM outputs meet financial reporting. RQ2 asks what operating artifact can make that reconciliation repeatable. RQ3 asks what changes are visible during implementation. The study uses design science methods (Hevner et al., 2004; Peffers et al., 2007) within an 18-month participatory case.

The case produced a four-part operating artifact and a descriptive account of its implementation. The evidence supports tentative propositions about source ambiguity, residual disclosure, shared ownership, and experimental checks. It does not support population-level or causal claims.

2. LITERATURE REVIEW

2.1 Marketing Accountability and the Marketing-Finance Interface

Marketing accountability research asks whether functional metrics connect to financial results (Kaplan & Norton, 1992; Rust et al., 2004). O'Sullivan and Abela (2007) associate measurement ability with firm performance and marketing's internal standing. Verhoef and Leeflang (2009) link marketing influence to accountability. Other work connects marketing actions to firm value (Edeling & Fischer, 2016; Srinivasan & Hanssens, 2009). Separate frameworks formalize marketing performance measurement (Katsikeas et al., 2016; Farris et al., 2010).

Metric availability does not guarantee metric use. Mintz and Currim (2013) show variation in managers' use of marketing and financial metrics. Homburg et al. (2012) find that measurement-system results depend on organizational conditions, including senior-management involvement. Marketing accountability work also calls for financial translation (Stewart, 2009; Ambler, 2003). Hanssens and Pauwels (2016) separate measurement of marketing value from managerial acceptance of that value. The reviewed literature does not specify a recurring reconciliation process for MMM estimates and finance actuals.

2.2 Marketing Mix Modeling

Marketing mix modeling grew from decision calculus and aggregate market-response modeling (Little, 1970; Hanssens et al., 2001). Later research modeled persistent effects and long-horizon response (Dekimpe & Hanssens, 1995; Mela et al., 1997). Meta-analyses and experiments provide empirical reference points for advertising effects (Lodish et al., 1995; Sethuraman et al., 2011). These bodies of work address estimation and response. They do not resolve how an enterprise should map model output to a finance reporting hierarchy.

Modern MMM includes several model classes. Bayesian specifications can encode carryover, saturation, and prior information (Jin et al., 2017). Meridian uses a Bayesian framework, while Robyn's R implementation uses regularized regression with hyperparameter optimization (Google, 2026; Meta Platforms, 2024). The frameworks differ in assumptions, workflow, and outputs. Strategic interpretation of long- and short-term effects also draws on effectiveness research (Binet & Field, 2013). Their existence does not establish which model class dominates enterprise practice. For reconciliation, the narrower point matters: MMM output remains model-based and uncertain regardless of implementation (Chan & Perry, 2017).

2.3 Experimental Measurement and Causal Inference

The potential-outcomes framework defines a causal effect as the difference between treated and untreated states for the same unit (Rubin, 1974; Holland, 1986). Only one state is observed. Field experiments estimate that contrast directly through random assignment. Television and online-to-offline studies demonstrate both feasibility and modest effect sizes in advertising settings (Lodish et al., 1995; Lewis & Reiley, 2014). Power requirements can be large relative to expected returns (Lewis & Rao, 2015). Available designs include ghost ads (Johnson et al., 2017) and randomized geo experiments (Vaver & Koehler, 2011; Kerman et al., 2017). Bayesian structural time-series and

synthetic controls provide quasi-experimental counterfactuals when randomization is unavailable (Brodersen et al., 2015; Abadie et al., 2010). Test design should also account for decision economics (Feit & Berman, 2019).

Experimental evidence sets a limit on observational confidence. A branded-search experiment found a large gap between attributed and incremental returns (Blake et al., 2015). Gordon et al. (2019, 2023) likewise report divergence between non-experimental estimates and randomized benchmarks. Sustained experimentation also requires operating infrastructure and protocols (Kohavi et al., 2020; Thomke, 2020). For this paper, experimental anchoring means using comparable causal evidence to challenge or calibrate selected model components. The strength of that evidence depends on design.

2.4 Research Gap

Three literatures supply separate pieces of the problem. Marketing accountability defines the demand for financial linkage. MMM supplies model-based incremental estimates. Experimentation supplies external checks on selected causal estimates. The reviewed literature did not identify a recurring enterprise process that connects these pieces to finance reporting. Adjacent practitioner evidence documents verification and definition problems in advertising measurement (Association of National Advertisers & K2 Intelligence, 2016; Media Rating Council, 2014). The CMO Survey reports continued difficulty quantifying marketing impact (Moorman, 2024). These sources motivate the governance problem. They do not establish the Reconciliation Layer as an industry-wide solution. Here, reconciliation means traceability and comparability. It does not mean finance can audit causal identification through accounting procedures.

3. METHODOLOGY

3.1 Research Design

The study uses design science as its organizing method (Hevner et al., 2004; Peffers et al., 2007). It tracks an operating artifact through one participatory case. The author served as product owner of the measurement platform throughout the observation window. That role provided direct access to release records, defect logs, governance documents, finance correspondence, and meeting records. It also creates selection and interpretation risk. The analysis is therefore documentary and descriptive. It does not claim an independent or complete record of every relevant interaction. No second coder or inter-rater reliability test is reported. The dual role also creates an incentive to favor the artifact, so the paper separates implemented controls, proposals, and participant interpretations where the records allow. The manuscript does not document an IRB or exemption determination, a participant-consent protocol, or employer authorization for research use of internal records. Those items require confirmation before external submission. The single-case structure follows Yin (2018). Theory-building use is treated cautiously, consistent with Eisenhardt (1989).

3.2 Research Setting

The setting is a large U.S. biopharmaceutical commercial organization, referred to as the Enterprise. It runs a Bayesian MMM on a quarterly release cycle for a multi-brand portfolio. Brand, commercial analytics, and finance teams use the outputs. Finance closes each fiscal period before final sales allocations land, often in the following one or two weeks. A third party aggregates media inputs. During the study, an enterprise sales-data migration changed the upstream source path. The paper omits brand names, individual names, exact release dates, company codes, and several numeric details to reduce deductive identification risk. External release still requires confirmation of the Enterprise's disclosure and research-use requirements.

3.3 Data Sources

Evidence covers roughly 18 months, from early 2025 through mid-2026. The documentary set includes four versions of release governance materials and reconciliation workbooks. One workbook compares 16 brands with linked formulas. The set also includes a 13-item upstream defect register and records from six quarterly release cycles. Finance correspondence and meeting records cover source designation, restatement, allocation timing, and change control. Sales-data migration records are also included. The manuscript does not include a separate codebook, a complete correspondence census, or independent coding. The analysis therefore treats these records as a bounded documentary set. Claims below are limited to what those records show.

3.4 Evidence Criteria

The analysis organizes implementation evidence under four descriptive criteria. Adoption means a control appears in standing governance, not only in an incident response. Defect handling means an issue receives a named owner, a bounded handoff, and a tracked record. Decision retention means model output continues to appear in planning during disruption. The records do not show whether a specific allocation decision changed because of it. Finance participation is assessed through observable actions, including hierarchy confirmation, file-design work, source-path ownership, and acceptance or rejection of change-control rules. These criteria are not presented as pre-registered thresholds. They structure the case evidence.

4. THE RECONCILIATION LAYER FRAMEWORK

The Reconciliation Layer is a recurring operating process that maps model-estimated incremental contribution to a finance-defined comparison. It uses four elements. Definitional alignment fixes the metric, scope, hierarchy, and time window. The decomposition bridge records each adjustment and the remaining discrepancy. Experimental anchoring checks selected model components against stronger causal evidence when comparable tests exist. Cadence governance sets file ownership, cutoffs, restatement rules, and review timing. In the case, these elements did not arrive at once. The bridge workbook and source discussions came first. Intake controls and change rules followed. A longer-term data-product contract remained an agreement in principle, and experimental coverage was partial.

Component	Purpose	Primary Owner	Representative Artifacts
Definitional alignment	Fix metric, scope, hierarchy, and time-window correspondence	Joint (measurement and finance)	Shared glossary; scope and hierarchy maps; fiscal window rules
Decomposition bridge	Trace aligned model terms, documented adjustments, and the residual	Measurement, reviewed by finance	Bridge schedule; residual log; audit workbook
Experimental anchoring	Check selected components against comparable causal evidence	Measurement / experimentation	Randomized geo tests; quasi-experimental checks; calibration records
Cadence governance	Set source ownership, cutoffs, restatement rules, and review timing	Joint	Release governance; intake gate; data-drop tracker; change-control rules

Table 1. Components of the Reconciliation Layer.

4.1 Definitional Alignment

Reconciliation starts by fixing what each number measures. The case exposed four recurring mismatches. Metric mismatch arose when model outcome definitions differed from the finance net-sales hierarchy. Scope mismatch appeared when contract sales sat under a separate company code. Timing mismatch arose because sales allocations landed after close and because model carryover crossed fiscal boundaries. Taxonomy mismatch appeared when model channels did not map cleanly to finance cost centers. These categories describe the observed case. The paper does not claim saturation or frequency ranking because the records were not independently coded.

4.2 The Decomposition Bridge

The bridge compares terms only after units, scope, hierarchy, and time window are fixed. Let F be the finance-reported movement in the aligned outcome. Let M be the MMM-estimated marketing contribution on that same basis. Let N be the contribution from non-marketing drivers needed to explain the same observed movement. Let A

be signed adjustments required to align scope, timing, or interaction treatment. The bridge residual is $R = F - (M + N + A)$. Each term must use the same outcome unit and reporting window. Model uncertainty remains attached to M and to any model-derived part of N or A. The bridge traces comparability. It does not turn the estimate into an accounting identity or validate causal identification.

A residual is the portion of financial movement not explained by the aligned model terms and documented adjustments. It can arise from definition gaps, timing, omitted drivers, model error, source error, or accounting changes. Financial actuals can also change through accruals, reserves, allocations, and restatements. The process records R, assigns an investigation status, and does not force it into another category. Exact agreement is possible when terms are aligned by construction. It still requires review when agreement depends on judgmental reallocations or model-derived adjustments. Case records associated residual disclosure with greater finance participation, but the study cannot separate that change from executive sponsorship, data repair, or repeated cross-functional work.

Term	Definition
F	Finance-reported movement in the aligned outcome, hierarchy, and reporting window.
M	MMM-estimated marketing contribution on the same aligned basis.
N	Non-marketing contribution used to explain the same observed movement.
A	Signed adjustments required to align scope, timing, or interaction treatment.
R	Residual: $R = F - (M + N + A)$. Investigate and report rather than allocate it away.

Table 2. Reconciliation identity and review terms.

4.3 Experimental Anchoring

Experimental anchoring applies only when the external estimate is comparable to the MMM component. The comparison must align on estimand, population, channel definition, geography, time horizon, and interference assumptions. Randomized geo experiments provide the strongest evidence among the methods considered here (Vaver & Koehler, 2011; Kerman et al., 2017). Bayesian structural time-series and synthetic control analyses can provide quasi-experimental checks when randomization is not feasible (Brodersen et al., 2015; Abadie et al., 2010). They should not be labeled independent experimental verification. If a comparable experiment conflicts with the model, the result triggers model and data review. The case did not establish a fixed recency window or automatic calibration rule, and portfolio coverage remained partial.

4.4 Cadence Governance

Cadence governance turns a one-off bridge into a repeatable control. Four rules appeared in the case. Review followed the financial close calendar. Each inbound file needed a named source, delivery expectation, and cutoff. Closed-period restatements required an explicit exception. A pre-model intake gate checked coverage and magnitude before modeling, with a data-drop tracker for failures. These controls govern handoffs and change. They do not validate the causal model.

5. CASE ANALYSIS AND RESULTS

5.1 Baseline Failure State

Early in the observation window, one quarterly release was paused after model-input net sales did not match finance-reported sales. Contemporaneous records described brand-level discrepancies ranging from roughly 10% to 50%. Because the manuscript does not document a single denominator for those percentages, the range is treated as descriptive rather than a formal statistic. Most affected brands required ROI recalculation; a smaller set required

model rebuilds. Investigation traced one major issue to an integration process that dropped sales months even though the source file itself reconciled. The records also describe one to two brands per quarter requiring remodeling because of upstream data problems. This paper reports those counts as case records, not population estimates.

5.2 Reconciliation Failure Categories

The 13-item defect register and release records contain incidents that fit four recurring categories. Source ambiguity involved multiple candidate feeds or hierarchies for the same financial quantity. Timing mismatch arose when sales allocation and model ingestion used different vintages. Scope divergence involved company codes and indication splits that did not match the model boundary. Silent upstream change included mapping changes and dropped data during the platform transition. These categories are descriptive. No independent codebook or second-coder review is available, so the paper does not rank them by prevalence. Within the documented reconciliation incidents, the largest problems described in the records originated upstream of model estimation. The case does not test model specification, prior sensitivity, or causal identification.

5.3 Artifact Implementation

Implementation was uneven. The audit workbook, pre-model intake gate, and data-drop tracker were implemented. Cutoff and restatement rules were documented in governance, including finance's no-restatement position for closed periods. An interim manually produced sales file was used as a bridge and scoped to one company code. A split reconciliation contract for the future enterprise data product was agreed in principle. A three-way governance forum was proposed but is not counted as adopted. Experimental anchoring used calibration studies and a design calendar without full portfolio coverage. Section 5.4 uses these status distinctions.

5.4 Observed Changes

Four changes are visible in the records. First, several controls moved into standing governance: the intake gate, data-drop tracker, cutoff rules, and bridge workbook. Second, the split reconciliation contract assigned different parts of the source path to named owners. The records do not establish a measured decline in recurrence, resolution time, escalation volume, or rework. Third, model outputs continued to appear in quarterly planning after upstream repair. The materials summarized here do not show that a specific allocation changed because of a reconciled output. Fourth, finance confirmed hierarchy-level correspondence, joined interim-file design, and identified the enterprise data-product owner. Those actions show deeper participation. They do not isolate the cause of that change.

5.5 Related Experimental Evidence

Advertising experiments provide a useful analogy for one part of the framework. Blake et al. (2015) and Gordon et al. (2019, 2023) show that attributed or observational estimates can differ from randomized benchmarks. That evidence supports external checks on model estimates. It does not validate the Reconciliation Layer's governance design or its generalizability across firms.

6. DISCUSSION

6.1 Provisional Propositions

Three propositions come from the case. First, traceable source, scope, timing, hierarchy, and adjustment rules may reduce ambiguity in finance review. Second, an explicit residual gives both teams a specific discrepancy to investigate. Third, split ownership can clarify responsibility across the data path. The case cannot isolate these mechanisms. Executive sponsorship, data repair, repeated interaction, and the sales-platform migration changed during the same period. These propositions require testing outside this setting.

6.2 Operating Sequence and Staffing

The observed sequence suggests a practical order, not a universal recipe. Start by aligning the outcome metric, hierarchy, scope, and fiscal window. Build the first bridge with every adjustment traceable to a source. Keep the residual explicit. Add experimental checks only where the estimand and test design are comparable. Then place file ownership, cutoffs, restatement rules, and review timing into standing governance. The case also exposed a need for work that spans model structure and finance reporting. Tasks included metric mapping, source tracing, residual

investigation, and change-control design. One person could perform that work, or a paired measurement-finance team could split it. The observed sequence unfolded over roughly one year. Platform migration, data repair, staffing, and governance dependencies contributed to that elapsed time, so it should not be used as an implementation benchmark.

6.3 Proposed Boundary Conditions

Three possible boundary conditions follow from the case and prior work. Experimental anchoring requires enough scale and variation to support adequately powered tests (Lewis & Rao, 2015). Platform transitions create more opportunities for silent source and mapping changes, which raises the need for intake controls. Formal close processes provide a ready cadence for cutoffs and restatement rules. These are propositions, not tested moderators. The single case does not show how the framework performs in smaller portfolios, informal finance environments, or firms without a shared data platform.

7. LIMITATIONS AND FUTURE RESEARCH

Several limits are material. This is one case in one industry. The author was the product owner, and no independent coder or complete record audit is reported. The evidence is documentary and qualitative. It shows implementation actions and participation changes, but it does not estimate causal effects on budgets, decisions, or firm performance. Experimental anchoring remained partial. Several counts and dates are generalized or withheld for confidentiality. The manuscript does not document an IRB or exemption determination, a participant-consent process, or employer authorization for research use. Those matters must be resolved before external submission. Future work should replicate the framework across firms, pre-specify evaluation measures, and test whether reconciled outputs change specific allocation decisions.

7.1 Data and Materials Availability

Underlying defect records, reconciliation workbooks, finance correspondence, meeting records, and migration artifacts contain confidential enterprise information and are not included with this manuscript. The current materials do not include a public codebook or redacted evidence package. Any external release of source artifacts would require Enterprise approval and a disclosure review. A journal submission should state which redacted or synthetic materials can be shared and under what access conditions.

8. CONCLUSION

In this case, reconciliation failed when model outputs and finance reports used different sources, timing, scope, or mappings. The Reconciliation Layer addressed those issues through aligned definitions, a traceable bridge, selective external checks, and governance tied to financial close. Several controls entered standing use, and finance took on more of the source and hierarchy work. The records do not show a causal effect on allocation decisions or firm performance. Experimental anchoring remained partial. The case therefore supports the framework as a documented operating design for one enterprise. Replication and independent evaluation are still needed.

References:

- [1] Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American Statistical Association*, 105(490), 493-505. <https://doi.org/10.1198/jasa.2009.ap08746>
- [2] Ambler, T. (2003). *Marketing and the bottom line: The marketing metrics to pump up cash flow* (2nd ed.). FT Prentice Hall.
- [3] Association of National Advertisers & K2 Intelligence. (2016). *An independent study of media transparency in the U.S. advertising industry*. Association of National Advertisers.
- [4] Binet, L., & Field, P. (2013). *The long and the short of it: Balancing short and long-term marketing strategies*. Institute of Practitioners in Advertising.
- [5] 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>

- [6] Brodersen, K. H., Gallusser, F., Koehler, J., Remy, N., & Scott, S. L. (2015). Inferring causal impact using Bayesian structural time-series models. *Annals of Applied Statistics*, 9(1), 247-274. <https://doi.org/10.1214/14-AOAS788>
- [7] Chan, D., & Perry, M. (2017). Challenges and opportunities in media mix modeling. Google Research.
- [8] Dekimpe, M. G., & Hanssens, D. M. (1995). The persistence of marketing effects on sales. *Marketing Science*, 14(1), 1-21. <https://doi.org/10.1287/mksc.14.1.1>
- [9] Edeling, A., & Fischer, M. (2016). Marketing's impact on firm value: Generalizations from a meta-analysis. *Journal of Marketing Research*, 53(4), 515-534. <https://doi.org/10.1509/jmr.14.0046>
- [10] Eisenhardt, K. M. (1989). Building theories from case study research. *Academy of Management Review*, 14(4), 532-550. <https://doi.org/10.5465/amr.1989.4308385>
- [11] Farris, P. W., Bendle, N. T., Pfeifer, P. E., & Reibstein, D. J. (2010). *Marketing metrics: The definitive guide to measuring marketing performance* (2nd ed.). Wharton School Publishing.
- [12] Feit, E. M., & Berman, R. (2019). Test & roll: Profit-maximizing A/B tests. *Marketing Science*, 38(6), 1038-1058. <https://doi.org/10.1287/mksc.2019.1194>
- [13] Google. (2026). Meridian: Marketing mix modeling (Version 1.8.0; source commit 9369071) [Computer software]. GitHub. Retrieved August 19, 2026, from <https://github.com/google/meridian>
- [14] Gordon, B. R., Moakler, R., & Zettelmeyer, F. (2023). Close enough? A large-scale exploration of non-experimental approaches to advertising measurement. *Marketing Science*, 42(4), 768-793. <https://doi.org/10.1287/mksc.2022.1413>
- [15] 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>
- [16] Hanssens, D. M., Parsons, L. J., & Schultz, R. L. (2001). *Market response models: Econometric and time series analysis* (2nd ed.). Kluwer Academic Publishers.
- [17] Hanssens, D. M., & Pauwels, K. H. (2016). Demonstrating the value of marketing. *Journal of Marketing*, 80(6), 173-190. <https://doi.org/10.1509/jm.15.0417>
- [18] Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly*, 28(1), 75-105. <https://doi.org/10.2307/25148625>
- [19] Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, 81(396), 945-960. <https://doi.org/10.1080/01621459.1986.10478354>
- [20] Homburg, C., Artz, M., & Wieseke, J. (2012). Marketing performance measurement systems: Does comprehensiveness really improve performance? *Journal of Marketing*, 76(3), 56-77. <https://doi.org/10.1509/jm.09.0487>
- [21] Jin, Y., Wang, Y., Sun, Y., Chan, D., & Koehler, J. (2017). Bayesian methods for media mix modeling with carryover and shape effects. Google Research.
- [22] 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), 867-884. <https://doi.org/10.1509/jmr.15.0297>
- [23] Kaplan, R. S., & Norton, D. P. (1992). The balanced scorecard: Measures that drive performance. *Harvard Business Review*, 70(1), 71-79.
- [24] Katsikeas, C. S., Morgan, N. A., Leonidou, L. C., & Hult, G. T. M. (2016). Assessing performance outcomes in marketing. *Journal of Marketing*, 80(2), 1-20. <https://doi.org/10.1509/jm.15.0287>
- [25] Kerman, J., Wang, P., & Vaver, J. (2017). Estimating ad effectiveness using geo experiments in a time-based regression framework. Google Research.

- [26] Kohavi, R., Tang, D., & Xu, Y. (2020). Trustworthy online controlled experiments: A practical guide to A/B testing. Cambridge University Press.
- [27] Lewis, R. A., & Rao, J. M. (2015). The unfavorable economics of measuring the returns to advertising. *Quarterly Journal of Economics*, 130(4), 1941-1973. <https://doi.org/10.1093/qje/qjv023>
- [28] Lewis, R. A., & Reiley, D. H. (2014). Online ads and offline sales: Measuring the effect of retail advertising via a controlled experiment on Yahoo! *Quantitative Marketing and Economics*, 12(3), 235-266. <https://doi.org/10.1007/s11129-014-9146-6>
- [29] Little, J. D. C. (1970). Models and managers: The concept of a decision calculus. *Management Science*, 16(8), B466-B485. <https://doi.org/10.1287/mnsc.16.8.B466>
- [30] Lodish, L. M., Abraham, M., Kalmenson, S., Livelsberger, J., Lubetkin, B., Richardson, B., & Stevens, M. E. (1995). How T.V. advertising works: A meta-analysis of 389 real world split cable T.V. advertising experiments. *Journal of Marketing Research*, 32(2), 125-139. <https://doi.org/10.1177/002224379503200201>
- [31] Media Rating Council. (2014). Viewable ad impression measurement guidelines. Media Rating Council.
- [32] Mela, C. F., Gupta, S., & Lehmann, D. R. (1997). The long-term impact of promotion and advertising on consumer brand choice. *Journal of Marketing Research*, 34(2), 248-261. <https://doi.org/10.1177/002224379703400205>
- [33] Meta Platforms. (2024). Robyn (Version 3.12.0; source commit 8eb5171) [Computer software]. GitHub. Retrieved August 19, 2026, from <https://github.com/facebookexperimental/Robyn>
- [34] Mintz, O., & Currim, I. S. (2013). What drives managerial use of marketing and financial metrics and does metric use affect performance of marketing-mix activities? *Journal of Marketing*, 77(2), 17-40. <https://doi.org/10.1509/jm.11.0463>
- [35] Moorman, C. (2024). The CMO Survey: Highlights and insights report. Duke University, Fuqua School of Business.
- [36] O'Sullivan, D., & Abela, A. V. (2007). Marketing performance measurement ability and firm performance. *Journal of Marketing*, 71(2), 79-93. <https://doi.org/10.1509/jmkg.71.2.079>
- [37] Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems*, 24(3), 45-77. <https://doi.org/10.2753/MIS0742-1222240302>
- [38] Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5), 688-701. <https://doi.org/10.1037/h0037350>
- [39] Rust, R. T., Ambler, T., Carpenter, G. S., Kumar, V., & Srivastava, R. K. (2004). Measuring marketing productivity: Current knowledge and future directions. *Journal of Marketing*, 68(4), 76-89. <https://doi.org/10.1509/jmkg.68.4.76.42721>
- [40] Sethuraman, R., Tellis, G. J., & Briesch, R. A. (2011). How well does advertising work? Generalizations from meta-analysis of brand advertising elasticities. *Journal of Marketing Research*, 48(3), 457-471. <https://doi.org/10.1509/jmkr.48.3.457>
- [41] Srinivasan, S., & Hanssens, D. M. (2009). Marketing and firm value: Metrics, methods, findings, and future directions. *Journal of Marketing Research*, 46(3), 293-312. <https://doi.org/10.1509/jmkr.46.3.293>
- [42] Stewart, D. W. (2009). Marketing accountability: Linking marketing actions to financial results. *Journal of Business Research*, 62(6), 636-643. <https://doi.org/10.1016/j.jbusres.2008.02.005>
- [43] Thomke, S. (2020). Experimentation works: The surprising power of business experiments. Harvard Business Review Press.
- [44] Vaver, J., & Koehler, J. (2011). Measuring ad effectiveness using geo experiments. Google Research.

- [45] Verhoef, P. C., & Leeflang, P. S. H. (2009). Understanding the marketing department's influence within the firm. *Journal of Marketing*, 73(2), 14-37. <https://doi.org/10.1509/jmkg.73.2.14>
- [46] Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). Sage Publications.