



AUS DER FORSCHUNG

Was ist Inkrementalität?

Inkrementalität isoliert den realen kausalen Mehrertrag kommerzieller Maßnahmen gegenüber einer unbehandelten Kontrollgruppe.

Sinan Isoglu, MBA (Quantic)

Führungskraft für Umsatzwachstum, Dozent und Doktorand

VERÖFFENTLICHT

2. September 2026

LESEZEIT

9 Min.

WÖRTER

1.889

<https://isoglu.com/de/insights/journal/was-ist-inkrementalitaet/>

Management Summary	2
Wie wird Inkrementalität formal definiert und berechnet?	3
Warum weisen gängige Attributions-Dashboards Scheinerfolge aus?	5
Praxisbeispiel: Holdout-Test bei bezahlten Markensuchanzeigen	6
Welche typischen Messfehler untergraben Inkrementalitätstests?	7
Welches auditierbare Protokoll etabliert verlässliche Inkrementalitätsmessung?	7
Wo liegen die empirischen Grenzen der Inkrementalitätsmessung?	8
References	9

MANAGEMENT SUMMARY

Inkrementalität misst den tatsächlichen kausalen Mehrwert einer Marketing- oder Vertriebsintervention gegenüber dem kontrafaktischen Basisszenario ohne diese Maßnahme. Während gängige Attributionsmodelle Umsätze beobachteten Kontaktpunkten zuschreiben, verwechseln sie systematisch Korrelation mit Kausalität und belohnen Werbung bei abschlussbereiten Stammkunden. Fundiert in der experimentellen Ökonomie und Kausalinferenz nutzt Inkrementalität randomisierte Kontrollgruppen, Ghost Ads und Geo-Tests zur Ermittlung des wahren Werbeeffekts. Dieser Leitfaden formalisiert inkrementelle Ertragsmetriken, vergleicht kausale Tests mit Attributionsmodellen, rechnet ein realistisches Suchmaschinen-Szenario durch und definiert ein auditierbares Protokoll zur Budgetallokation.

Schlagwörter: Inkrementalität · Kausalinferenz · Marketingattribution · Randomisierte Feldexperimente · Werbewirkungsmessung

Inkrementalität bezeichnet die Messung des tatsächlichen kausalen Mehrertrags, den eine Marketingkampagne, eine Vertriebsmaßnahme oder ein Preisnachlass gegenüber dem Zustand hervorruft, der ohne diese Maßnahme organisch eingetreten wäre. Fundiert in der experimentellen Wirtschaftsforschung und statistischen Kausalinferenz prüft das Konzept, ob ein Geschäftsabschluss durch den Werbeeinsatz verursacht wurde oder lediglich zeitlich mit einer ohnehin bestehenden Kaufabsicht zusammenfiel.

Standardmäßige Unternehmensberichte stützen sich primär auf beobachtende Attributionsmodelle wie Last-Touch-, First-Touch- oder Multitouch-Algorithmen. Diese Systeme schreiben Umsätze regelmäßig Kontaktpunkten zu, die bereits vorhandene Nachfrage lediglich abschöpfen. Ein Kunde, der den Markennamen bei Google eingibt oder ein Retargeting-Banner anklickt, ist häufig ein entschlossener Käufer. Diesen Umsatz bezahlten Kanälen zuzurechnen, erzeugt die Illusion hoher Marketingeffizienz, verschleiert jedoch eine massive Kapitalfehlallokation.

In der kaufmännischen Steuerung trennt die Inkrementalitätsmessung produktiven Kapitaleinsatz von teuren Scheinkorrelationen. Verlangt das Management keinen kontrafaktischen Nachweis, fließen Wachstumsbudgets bevorzugt in Maßnahmen, die bestehende organische Nachfrage vereinnahmen und ausgewiesene Akquisitionserrenditen künstlich aufblähen.

Wie wird Inkrementalität formal definiert und berechnet?

Kausale Inkrementalität wird durch den Vergleich einer behandelten Testgruppe mit einer unbehandelten Kontrollgruppe (dem kontrafaktischen Basisszenario) über einen identischen Zeitraum ermittelt.

Die Gleichung des inkrementellen Lifts

Sei Y_T das beobachtete Ergebnis (Umsatz oder Abschlüsse) in der behandelten Testgruppe und Y_C das Ergebnis in der unbehandelten Kontrollgruppe, bereinigt um das Größenverhältnis der Gruppen (N_T / N_C):

$$\text{Inkrementeller Lift} = Y_T - (Y_C \times N_C / N_T)$$

Die Inkrementalitätsquote (I_R) beziffert den Anteil am gesamten gemessenen Umsatz, der tatsächlich durch die Maßnahme verursacht wurde:

$$I_R = Y_T / Y_T - (Y_C \times N_C / N_T)$$

Ein Wert von $I_R = 1,0$ bedeutet, dass 100 % der Umsätze kausal auf die Kampagne zurückzuführen sind. Ein Wert von $I_R = 0,0$ belegt, dass die Kampagne keinerlei zusätzliche Erträge schuf und reinen Mitnahmeeffekten unterlag.

Inkrementeller Return on Ad Spend (iROAS)

Während der nominale ROAS den Bruttoumsatz durch Werbekosten teilt, berücksichtigt der inkrementelle ROAS ausschließlich kausal belegten Mehrumsatz:

$$iROAS = \frac{\text{Werbeausgaben Inkrementeller Umsatz}}{\text{Werbeausgaben}} = \frac{Y_T - (Y_C \times N_C / N_T)}{Y_T}$$

Blake et al. (2015) führten „a series of large-scale field experiments done at eBay“ durch und berichten das Markensuchergebnis ausdrücklich als Extremfall, nicht als Regel: „as an extreme case, we show that brand keyword ads have no measurable short-term benefit“. Die Ökonomie kommt daher, wenn die Ausgaben erreichen: „more frequent users whose purchasing behavior is not influenced by ads account for most of the advertising expenses,

resulting in average returns that are negative.“ Die meist weggelassene Hälfte ist die Grenze: „for non-brand keywords, we find that new and infrequent users are positively influenced by ads.“

Gordon et al. (2019) vergleichen beobachtende Methoden mit randomisierten Benchmarks bei Facebook, und die Zahl gehört als Zahl genannt: „the point estimates in 7 of the 14 studies with a checkout-conversion outcome are consistently off by more than a factor of three“. In ihrer ausgearbeiteten Studie lag der randomisierte Benchmark bei 73 Prozent Lift mit einem Intervall von 49 bis 103 Prozent, der naive Vergleich von Kontaktierten und Nicht-Kontaktierten bei 316 Prozent, und „exact matching on age and gender alone performs poorly, yielding a lift of 222%“.

Lewis und Rao (2015) untersuchten 25 Feldexperimente und fassen die Schwierigkeit in eine Zahl: „The median confidence interval on return on investment is over 100 percentage points wide.“ Die Ursache ist Varianz, nicht Methode: „relative to the per capita cost of the advertising, individual-level sales are very volatile; a coefficient of variation of 10 is common.“ Ein so breites Intervall kann eine hochprofitable Kampagne nicht von einer verlustbringenden trennen.

Lewis et al. (2011) identifizieren die zeitbasierte Aktivitätsverzerrung in drei kontrollierten Experimenten: Markenrelevante Suchanfragen, „page views across diverse websites“ und Anmeldungen beim Wettbewerber, wo „the true “competitive effect” was minimal“. Werbekontakt signalisiert mehr von allem, was es schwer macht, „a suitable“ matched control „using prior behavior“ zu finden, sodass der „match is fundamentally different from the exposed group“. Das ist ein Matching-Fehler, keine Werbewirkung.

Johnson et al. (2017) stellen mit Ghost Ads ein Verfahren vor, das „facilitates this comparison by identifying the control-group counterparts of the exposed consumers in a randomized experiment“. Die genannten Vorteile sind relativ und nicht absolut: gegenüber PSA- und Intent-to-Treat-Tests könne es „reduce the cost of experimentation, improve measurement precision, deliver the relevant strategic baseline, and work with modern ad platforms that optimize ad delivery in real-time“. Vorliegend ist ein Preprint, also als Methode zu lesen und nicht als gesichertes Ergebnis.

Abbildung 1 Die kausale Messhierarchie der Inkrementalität

Hierarchiestufe	Methodischer Ansatz	Wirkungsmechanismus	Kausale Validität	Risiko für Selektionsverzerrung
Stufe 1 (Goldstandard)	Randomisierte RCTs / Ghost Ads	Randomisierte Holdouts mit synthetischen Tags	Hoch	Null (statistisch eliminiert)
Stufe 2 (Regionale Tests)	Regionale Geo-Experimente	Synthetische Kontrollregionen ohne Werbedruck	Hoch bis Moderat	Gering (Abgleich der Vorperioden)
Stufe 3 (Ökonometrie)	Kalibrierte Marketing-Mix-Modelle	Bayesianische MMMs mit experimentellen Priors	Moderat	Moderat (benötigt Test-Anker)
Stufe 4 (Statistik)	Propensity Score Matching	Statistische Kontrollgruppen auf Basis historischer Daten	Gering bis Moderat	Hoch (blendet verborgene Absichten aus)
Stufe 5 (Klick-Attribution)	Multi-Touch-Attribution (MTA)	Algorithmische Gewichtung beobachteter Klicks	Vernachlässigbar	Kritisch (vereinbart Stammkunden)
Stufe 6 (Fehlleitend)	Last-Touch / First-Touch	Schreibt 100 % des Erfolgs dem letzten Klick zu	Null	Vollständig (belehnt reine Mitnahmeeffekte)

Causal-Measurement-Framework des Autors, fundiert in experimentellen Feldstudien von Blake et al. (2015), Gordon et al. (2019), Lewis und Rao (2015), Johnson et al. (2017) sowie Lewis et al. (2011).

Warum weisen gängige Attributions-Dashboards Scheinerfolge aus?

Die Kluft zwischen gemeldeten Kennzahlen und realem Geschäftserfolg entsteht durch strukturelle Verzerrungen in der digitalen Werbeauslieferung:

Tabelle 2 Warum weisen gängige Attributions-Dashboards Scheinerfolge aus?

Verzerrungsmechanismus	Verhaltensökonomische Ursache	Folge im Dashboard	Strategische Fehlsteuerung
Abschöpfung bestehender Absichten	Retargeting bietet aggressiv auf Nutzer im Warenkorb	Meldet extrem hohen ROAS (z. B. 1.500 %)	Entzieht Neukundenkampagnen das Budget
Aktivitätsverzerrung	Nutzer surfen vor Käufen intensiver im Netz	Zufällig eingeblendete Banner beanspruchen Kausalität	Verwechselt Surfintensität mit Werbewirkung
Kannibalisierung der Suche	Markensuchanzeigen fangen Markensucher ab	Klicks ersetzen kostenfreie Klicks auf organische Treffer	Verwandelt freie Suchtreffer in dauerhafte Kosten
Geräte- und Cookie-Brüche	Gerätewechsel verhindern die Nachverfolgung	Langfristiger Beziehungsaufbau wird ignoriert	Bevorzugt kurzfristige digitale Werbeformen
Eigenbewertung der Plattformen	Netzwerke messen ihren Erfolg über View-Throughs	Summe der Plattform-Umsätze übersteigt Firmenumsatz	Täuscht nicht existierende Nachfrage vor

Tabelle aus diesem Essay. Quellen und Einordnung stehen im Beitrag.

Die Verknüpfung mit den Wirtschaftlichkeitskennzahlen ist essenziell. Wie in Was sind Kundenakquisitionskosten? und Was ist der Deckungsbeitrag? verdeutlicht, führt die Attribution nicht-inkrementeller Umsätze zu geschönten Akquisitionskosten und verfälscht die Deckungsbeitragsrechnung.

Praxisbeispiel: Holdout-Test bei bezahlten Markensuchanzeigen

Ein E-Commerce-Unternehmen investiert monatlich 100.000 € in Markensuchanzeigen (Brand Search) bei Google. Das Dashboard der Werbepattform meldet hervorragende Zahlen: 800.000 € zugeordneten Umsatz (ein ROAS von 8{,}0).

Um zu überprüfen, ob diese Umsätze echt oder vorgetäuscht sind, initiiert die Geschäftsleitung einen vierwöchigen randomisierten Geo-Holdout-Test. Der Gesamtmarkt wird in zwei demografisch gleichwertige Regionen aufgeteilt, die jeweils 50 % des Kundenstamms abbilden.

1. Parameter des Testdesigns

- Testregion (Werbedruck unverändert): 50.000 € monatliche Werbeausgaben.
- Kontrollregion (Anzeigen vollständig pausiert): 0 € Werbeausgaben (100 % Holdout).
- Die historischen Umsätze beider Regionen lagen vor dem Test stabil bei jeweils 400.000 € monatlich.

2. Beobachtete Ergebnisse während des Testmonats

- Ergebnisse der Testregion (Anzeigen aktiv): Umsatz über bezahlte Suche: 400.000 € Umsatz über organische Suche: 80.000 € Direktaufrufe und sonstige Kanäle: 20.000 € Realisierter Gesamtumsatz Testregion: 500.000 €

- Ergebnisse der Kontrollregion (Anzeigen pausiert): Umsatz über bezahlte Suche: 0 € Umsatz über organische Suche: 440.000 € (Nutzer klickten auf organische Treffer) Direktaufrufe und sonstige Kanäle: 25.000 € Realisierter Gesamtumsatz Kontrollregion: 465.000 €

3. Berechnung der Kausaleffekte

Das Experiment belegt, dass 91,25 % der von der Plattform ausgewiesenen Umsätze nicht inkrementell waren (365.000 € der gemeldeten 400.000 €). Ohne Anzeigen wählten die Kunden einfach das organische Suchergebnis direkt darunter. Statt eines achtfachen Gewinns erzeugte die Kampagne einen realen Verlust von 15.000 € pro Monat in der Testregion.

Die Auswirkung auf den Unternehmenswert ist gravierend. Wie in Was ist Customer Lifetime Value? dargelegt, beschädigt die Finanzierung nicht-kausaler Akquisitionskanäle die gesamte Kohortenökonomie nachhaltig.

Welche typischen Messfehler untergraben Inkrementalitätstests?

Tabelle 3 Welche typischen Messfehler untergraben Inkrementalitätstests?

Messfehler	Ursache	Experimentelles Scheitern	Korrekturprotokoll
Zu geringe Stichprobengröße	Unterschätzung der natürlichen Umsatzvarianz	Zu breite Konfidenzintervalle machen Tests unbrauchbar	Power-Berechnungen vor Teststart nach Lewis und Rao (2015) durchführen
Vorher-Nachher-Vergleiche	Zeitweises nationales Abschalten ohne Kontrollgruppe	Saisoneffekte werden fälschlich als Werbewirkung gedeutet	Stets zeitgleiche Kontrollgruppen oder Geo-Zwillinge nutzen
Missachtung von Übertragungseffekten	Werbung strahlt aus Test- in Kontrollregionen ab	Verletzt die SUTVA-Annahme für unverzerrte Kausalschätzungen	Geografische Pufferzonen um Testgebiete definieren
Zu kurze Beobachtungsfenster	Messung endet unmittelbar mit Kampagnenschluss	Verzögerte Kaufentscheidungen und Ad-Stock werden übersehen	Nachlaufperioden zur Erfassung verzögerter Wirkungen etablieren
Pauschale Inkrementalitätsannahmen	Gleiche Quoten für Brand Search und Prospecting	Überfinanzierung von Brand Search bei Vernachlässigung von Neukunden	Inkrementalität für jeden Kanal separat ermitteln

Tabelle aus diesem Essay. Quellen und Einordnung stehen im Beitrag.

Welches auditierbare Protokoll etabliert verlässliche Inkrementalitätsmessung?

- 1 Kanäle nach Selektionsrisiko klassifizieren: Unterteilen Sie das Budget in hochgradig anfällige Maßnahmen (Brand Search, Retargeting) und echte Neukundenakquise (breite Markenwerbung).

- 2 Dauerhafte Kontrollgruppen bei Retargeting einrichten: Halten Sie permanent 5 % bis 10 % der Nutzer von Retargeting und Markensuche fern, um den organischen Basiseffekt zu überwachen.
- 3 Geo-Experimente für obere Trichterstufen nutzen: Bilden Sie statistische Zwillingsregionen auf Basis historischer Umsatzmuster, um über regionale Spend-Spikes echte Hebeleffekte zu testen.
- 4 Ghost Ads implementieren: Nutzen Sie synthetische Impression-Tags, um Auktionsgewinner ohne die Kosten für Null-Anzeigen zu isolieren (Johnson et al., 2017).
- 5 iROAS und iCAC als Führungsgrößen verankern: Ersetzen Sie Plattform-ROAS und gemischte CAC im Reporting durch kausal bereinigte Kennzahlen.
- 6 Budgets aus Scheineffekten abziehen: Beenden Sie Kampagnen mit Inkrementalitätsquoten unter 0,20 und lenken Sie die Mittel in Kanäle mit nachgewiesenem Mehrertrag.
- 7 Marketing-Mix-Modelle mit Testergebnissen kalibrieren: Speisen Sie experimentell gemessene Lift-Werte als feste Ankerwerte in ökonometrische Modelle ein, um Modellabweichungen zu verhindern.

Wo liegen die empirischen Grenzen der Inkrementalitätsmessung?

Inkrementalitätstests sind ein anspruchsvolles wissenschaftliches Verfahren, kein automatischer Selbstläufer. Im B2B-Geschäft mit mehrjährigen Vertriebszyklen, geringen Transaktionszahlen und komplexen Buying Centern stoßen randomisierte Feldexperimente an praktische Machbarkeitsgrenzen.

Lewis und Rao (2015) beziffern auch den Aufwand: Informative Werbeexperimente „can easily require more than 10 million person-weeks, making experiments costly and potentially infeasible for many firms“. Der Nachweis eines kleinen Effekts ist damit zuerst eine Budgetfrage und dann eine statistische, und Führungskräfte müssen die Kosten der Messung gegen die Kosten einer Fehlentscheidung abwägen.

Die wissenschaftlichen Grundlagen dieses Beitrags basieren auf der Fachliteratur zu digitaler Werbewirkung und Kausalinferenz, namentlich Blake et al. (2015), Gordon et al. (2019), Lewis und Rao (2015), Johnson et al. (2017) sowie Lewis et al. (2011). Die Strukturmodelle, Versuchsabläufe und Kontrollprotokolle stellen die methodische Synthese des Autors für eine belastbare Budgetallokation dar.

- Blake, T., Nosko, C., & Tadelis, S. (2015). Consumer heterogeneity and paid search effectiveness: A large-scale field experiment. *Econometrica*, 83(1), 155-174. DOI
- 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. DOI
- 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. DOI
- 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. DOI
- Lewis, R. A., Rao, J. M., & Reiley, D. H. (2011). Here, there, and everywhere: Correlated online behaviors can lead to overestimates of the effects of advertising. In *Proceedings of the 20th International Conference on World Wide Web* (pp. 157-166). DOI

ENDE DES BEITRAGS

ZITIERWEISE

Isoglu, S. (2026, 2. September). Was ist Inkrementalität? Sinan Isoglu.
<https://isoglu.com/de/insights/journal/was-ist-inkrementalitaet/>

WEITERLESEN

Aus der Forschung

Gestaffelte Difference-in-Differences braucht ein Kohorten-Zeit-Estimand

<https://isoglu.com/de/insights/journal/gestaffelte-difference-in-differences-braucht-ein-kohortENZEIT-estimand/>

Aus der Forschung

Was ist Claim-Entailment? Stützt die Quelle genau diesen Satz?

<https://isoglu.com/de/insights/journal/was-ist-claim-entailment/>

Aus der Forschung

Was ist Data Lineage? Die Transformationen hinter einem Ergebnis

<https://isoglu.com/de/insights/journal/was-ist-data-lineage/>



Sinan Isoglu, MBA (Quantic)

Führungskraft für Umsatzwachstum, Dozent und Doktorand