When Temporal Features Reuse Data: Adjacent-Window Bias in Customer Panels

Dr. Suraj Chandrakant Kamble[1], Dr. Sanmath Shetty[2], Dr. Sonali Khurjekar[3]

[1]Associate Professor, Deccan Education Society's Institute of Management Development and Research (DES IMDR), Pune, India, Email ID : suraj.kamble03@gmail.com , ORCID ID : 0000-0002-1080-6724.
[2]Associate Professor, Deccan Education Society's Institute of Management Development and Research (DES IMDR), Pune, India, Email ID : sanmath.shetty@imdr.edu , ORCID ID : 0009-0008-6633-5950
[3]Associate Professor, Deccan Education Society's Institute of Management Development and Research (DES IMDR), Pune, India, Email: sonali.khurjekar@imdr.edu ORCID ID : 0000-0001-8039-0288

Abstract
Customer histories are often converted into rolling or adjacent-window features for segmentation and prediction. The same period-level estimate may then appear as an outcome in one panel row and as a predictor in the next, creating apparent dynamics even when no carryover exists. We develop and apply a validation framework combining a no-carryover benchmark, fully disjoint temporal construction, an identical-sample estimator comparison and strict temporal holdout testing. The application uses 102 weeks of grocery transactions and 9,014 household-category relationships observed under both adjacent and disjoint designs. On this identical sample, the adjacent estimator implies a 1.78-percentage-point movement toward full-price purchasing for a 10-point increase in promotional share, whereas the fully disjoint estimator implies a 1.40-point movement in the opposite direction. A 36-cell Monte Carlo grid shows that adjacent estimators can remain positive under zero carryover across panel lengths and event frequencies, while disjoint estimators recover the null; under true persistence, adjacent construction attenuates or reverses the expected sign. Recent promotional share adds only .0063 to holdout AUC. The study shows that temporal feature validation must distinguish predictive usefulness from dynamic interpretation and intervention. The proposed audit supports model governance by identifying when customer-history features are informative for forecasting but unsafe for behavioural or policy claims.

Keywords: temporal feature validation; adjacent-window bias; marketing research methods; customer analytics; panel data; simulation; model governance

1. Introduction

Customer-level data are routinely compressed into behavioural features. Analysts calculate purchase shares, response rates, average discounts, channel proportions and category-specific frequencies over a recent period, then use those summaries to explain or predict what customers do next. This practice is sensible for forecasting when the feature is available at the prediction point and validated on genuinely future cases. It becomes more problematic when the same constructed history is interpreted as evidence of behavioural change. A feature can discriminate future outcomes because it captures stable customer differences, because it contains a genuine temporary state, or because its temporal construction mechanically links adjacent panel rows.

The third possibility receives less attention than the first two. In a short customer panel, an analyst may divide the observation horizon into windows and build repeated rows. A share estimated in one window can enter the outcome of one row and then reappear as the predictor of the following row. If the outcome is its complement, as with promotional share and full-price share, the same noisy period estimate changes roles and sign. Relationship fixed effects do not automatically remove this problem. Demeaning a short sequence can convert the cross-row reuse into a positive within association even when successive windows are independent.

This issue matters for market research because method choices govern the claims that can be attached to customer histories. A lagged feature may be useful for segmentation while being invalid as evidence of a dynamic mechanism. It may also add statistically detectable but commercially negligible predictive information. Yet these three questions—whether a feature predicts, whether it represents a behavioural state, and whether intervening on it would change behaviour—are often blurred. When they are blurred, descriptive labels such as deal prone or promotion dependent can become causal diagnoses, and model coefficients can be converted into offer rules that the design cannot support.

The empirical application concerns promotion history in grocery purchasing. Promotion-heavy purchasing may reflect persistent price responsiveness, offer access, category needs or retailer targeting. It may also reflect

reference-price learning, stockpiling, purchase acceleration or other short-run processes (Gupta, 1988; Lattin and Bucklin, 1989; Mela et al., 1997; Neslin et al., 1985; Seetharaman, 2004; Winer, 1986). These mechanisms are substantively important, but the present paper does not attempt to identify them causally. Instead, the application is used to test a measurement-validity question: can adjacent reuse of a constructed promotional share change the sign of an estimated dynamic relationship?

The study develops a four-stage audit. First, a formal within-transformation argument shows why shared-window measurement can induce a positive coefficient under a no-carryover null. Second, a 36-cell Monte Carlo grid varies panel length, event frequency, true persistence and latent serial correlation. Third, the empirical analysis compares adjacent and fully disjoint estimators on exactly the same 9,014 household-category relationships, holding sample composition fixed. Fourth, a strict temporal holdout assesses whether recent promotional share adds useful prediction for previously unseen households.

The exact-sample contrast is the central empirical result. A 10-percentage-point increase in current promotional share is associated with a 1.78-point increase in later full-price share under adjacent construction, but a 1.40-point decrease under fully disjoint construction. The sign reversal occurs without changing the analysed relationships. In the broader application, the conventional adjacent estimate is +2.49 points, while a no-carryover simulation produces an even larger +3.52 points. The primary fully disjoint estimate is -1.20 points. At the same time, historical promotional share prospectively predicts later price composition, and recent promotional share adds only .0063 to holdout AUC.

The contribution is methodological rather than a new theory of promotion response. It identifies a finite-panel validity risk created by reusing estimated window summaries, explains the sign of the resulting within association, and organizes a practical audit around null benchmarking, temporal disjointness, identical-sample comparison and future holdout validation. The contribution to market-research practice is a disciplined separation among prediction, explanation and intervention. The results do not imply that all lagged features are biased or that promotions create dependence. They show when a useful customer-history feature is unsafe for dynamic or policy interpretation.

2 Temporal feature validity in customer panels

2.1 Prediction is not dynamic explanation

A customer-history feature can be predictive for several reasons. It may summarize stable preferences, constraints, search costs, retailer exposure or category conditions. Promotion research documents substantial heterogeneity in price sensitivity and promotional response, as well as systematic variation by brand, category and store conditions (Ailawadi et al., 2006; Anderson and Fox, 2019; Bell et al., 1999; van Heerde et al., 2003). Related evidence identifies enduring differences in deal proneness, shopping patterns, price search and within-household choice (Bell and Lattin, 1998; Bruno et al., 2018; Gauri et al., 2008; Lichtenstein et al., 1990, 1993, 1995). A household-category relationship can therefore remain promotion heavy without any promotion-created state. The history is informative because it identifies where the relationship already sits in a persistent distribution.

A dynamic interpretation is stronger. It states that a temporary deviation within the same relationship predicts a subsequent deviation. Reference-price updating, learning, inventory management and purchase timing can generate such relationships (Briesch et al., 1997; Chandon and Wansink, 2002; Chan et al., 2008; Dickson and Sawyer, 1990; Gupta and Cooper, 1992; Kalwani et al., 1990; Kopalle et al., 1999; Mayhew and Winer, 1992; Mela et al., 1998; Papatla and Krishnamurthi, 1996; Winer, 1986). Changes in category demand may also appear through purchase incidence or quantity, and these margins need not move together over time (Nijs et al., 2001; Pauwels et al., 2002). Even so, the within-relationship coefficient cannot tell us which process is operating. Exposure to promotions and redemption are both selected rather than random. The retailer uses past purchases to decide which offers to send, and households then choose whether to redeem them. At the same time, budgets, category needs, inventories and purchases at competing retailers can change from one period to the next.

There is also a basic measurement problem. Promotional share is calculated from observed purchases; it is not an underlying household characteristic that can be read directly. In a short window, two relationships with the same propensity may produce different shares simply because one happens to contain more promoted purchases. The number of events in the denominator therefore matters, and that number changes across relationships and windows. A longer window will usually give a steadier share. It also leaves fewer separate windows for testing whether an apparent state carries forward. The practical choice is between more precise measurement within a window and more independent time variation across windows.

Ordinary predictive validation does not settle this issue. If constructed panel rows are split at random, records from the same household-category relationship may appear in both samples. Source periods can also overlap across the split. A strong test score may then reflect a familiar relationship or partly reused information rather than performance for a future customer. Validation must therefore identify two boundaries: the date at which information is cut off and the observational unit kept out of training. In our test, the outcome comes from a later transition and every test household is absent from the training sample. The incremental prediction result cannot therefore be driven by another record for the same customer relationship.

Marketing measurement research makes the same broader point: conclusions depend on how a measure is constructed and checked (Churchill, 1979; Diamantopoulos and Winklhofer, 2001). The measure examined here is not a psychometric scale. It is a behavioural index calculated repeatedly from a limited set of purchase events. Its meaning depends on the records included, the denominator, the time boundary and the part it plays in the model. With shared windows, a value intended to describe an earlier state also enters the data used to measure change. We refer to this problem as temporal construct contamination.

The problem also bears on state dependence and on the claimed value of marketing analytics. State-dependence models interpret earlier behaviour as changing later utility. Persistent-heterogeneity models explain repeated choices through stable differences (Dubé et al., 2010; Levine and Seiler, 2023). If the construction of the panel can determine the sign, an adjacent-window coefficient offers no sound basis for choosing between these accounts. Organizations use marketing analytics to sense customer changes and respond quickly, but that promise assumes valid data transformations and a reproducible workflow (Agag et al., 2024; Menukhin et al., 2025). Temporal validity is therefore a condition that analytics must satisfy before it can yield defensible managerial knowledge. It is not another theory of response to promotions.

Recent work on reproducibility points to a closely related concern. Leakage may enter through temporal ordering, repeated units or preprocessing that crosses the boundary used for validation. Its importance also depends on where and to whom the model is meant to generalize (Kapoor and Narayanan, 2023; Kapoor et al., 2024; Rosenblatt et al., 2024). Work in business analytics likewise argues that preprocessing choices should be tested because they are not neutral steps (Jimenez Garcia and Gil Herrera, 2026). The present study extends that logic from train-test leakage to within-panel interpretation: even when prediction is evaluated honestly, reuse inside the estimating panel can manufacture an explanatory coefficient.

Intervention is stronger again. An offer-policy rule assumes that changing promotion exposure will change a later outcome. Contemporary work shows that coupon design, beneficiary framing, automation, targeting and promotion predictability can materially change observed responses (Daljord et al., 2023; Fisher et al., 2023; Gabel and Guhl, 2022; Haviv, 2022; Lamei et al., 2025; Nigam et al., 2022; Osuna et al., 2016; Park et al., 2018; Venkatesan and Farris, 2012; Zhang et al., 2025). Such evidence does not remove the identification problem in an observational customer panel: an intervention claim still requires exogenous variation or a credible causal design. The central methodological principle is therefore hierarchical: prospective association may support description; future holdout performance may support prediction; dynamic explanation requires valid temporal construction; intervention requires causal identification.

2.2 How shared-window reuse creates an apparent dynamic

Let X_it denote the observed behavioural share for relationship i in window t. In the promotion application, X_it is the promotional-purchase share, and the next-window full-price share is defined in Equation (1). Under a simplified no-carryover null, successive observed shares fluctuate independently around a stable relationship propensity after conditioning on the relationship effect; hence the row-wise population covariance between X_it and Y_it is zero.

$$Y_{it} = 1 - X_{i,t+1} \tag{1}$$

$$M_T = I_T - \frac{1}{T} 1_T 1_T^{'} \tag{2}$$

$$\mathrm{Cov}\left(X_{it}, Y_{i,t-1}\right) = -\sigma_\varepsilon^2 \tag{3}$$

$$\mathrm{E}\left(X_i^{'} M_T Y_i\right) = \frac{T-1}{T}\sigma_\varepsilon^2 > 0 \tag{4}$$

$$\mathrm{E}\left(X_i^{'} M_T X_i\right) = (T-1)\sigma_\varepsilon^2 \tag{5}$$

$$\mathrm{E}\left(\hat{\beta}_{FE}\right) \approx \frac{1}{T} \tag{6}$$

An adjacent fixed-effects panel does not use each period estimate only once. X_it is the predictor in row t, while the same X_it appears with a negative sign inside Y_i,t-1. Equation (2) defines the within-demeaning matrix for T estimating rows. Although the same-row covariance is zero under the null, Equation (3) shows the negative cross-row covariance created by reuse. Applying the within transformation yields the positive expected numerator in Equation (4) and the expected denominator in Equation (5). Their ratio gives the approximate induced coefficient in Equation (6).

The result is not behavioural mean reversion. It is a finite-panel consequence of demeaning a sequence in which a noisy estimated share is reused across outcome and predictor roles. The induced association is largest when the number of estimating rows is small. Unequal event counts, missing windows, controls and serially persistent latent propensities change its magnitude, so the expression is a sign argument rather than a closed-form description of the empirical estimator. It nevertheless yields a falsifiable expectation: under zero carryover, adjacent construction can produce a positive coefficient, whereas a design that prevents period reuse should recover approximately zero.

2.3 A four-stage validation framework

As summarized in Table 1, the framework distinguishes estimator validity from predictive value. A feature may fail the dynamic audit yet remain useful in a prospective model. Conversely, an apparently strong dynamic coefficient may be unsafe if it is reproduced under the null or disappears after temporal disjointness is imposed.

Table 1. Four-stage validation framework

Stage	Question	Diagnostic
1 Null benchmark	Can the observed dynamic sign arise when carryover is absent?	Preserve relevant panel structure but simulate independent window outcomes.
2 Temporal disjointness	Does the sign survive when no source window is reused across estimating rows?	Construct fully disjoint predictor-control-outcome triples.
3 Identical-sample contrast	Is any change due to construction rather than eligibility or composition?	Estimate adjacent and disjoint models on exactly the same relationships.
4 Future validation	Does the feature add useful prospective information?	Evaluate incremental performance in a future period and on previously unseen customers.

The stages are intentionally cumulative. A null benchmark alone can show that a coefficient is not diagnostic of the proposed mechanism, but a poorly calibrated simulation might omit relevant persistence. A disjoint estimator removes the direct reuse channel, but it often changes the eligible sample. An identical-sample comparison addresses that composition problem, although it may reduce precision. A temporal holdout answers a different question: whether the feature contributes to prediction when used as it would be operationally available. Agreement across stages provides stronger evidence than any single robustness check.

The framework also makes non-confirmation informative. If an adjacent sign disappears under disjoint construction, the correct response is not to search across specifications until the original sign returns. The result identifies a boundary on interpretation. The feature can still be retained for prediction if it improves future performance, but claims about behavioural recovery should be withdrawn. This logic turns robustness analysis from a catalogue of alternative models into a sequence of decisions tied to distinct inferential purposes.

2.4 Research questions

RQ1. Under what temporal construction conditions can reuse of a period-level customer measure generate an apparent dynamic association when no carryover exists?

RQ2. Does fully disjoint construction recover the intended direction more reliably than adjacent construction when the true process contains no carryover or modest persistence?

RQ3. In an empirical customer panel, does changing only temporal construction reverse the conclusion while holding the analysed relationships fixed?

RQ4. Does the disputed temporal feature add meaningful predictive information in a genuinely future, previously unseen-household holdout?

3 Validation design

3.1 Monte Carlo generality grid

The simulation evaluates whether the reuse mechanism is confined to one empirical sample. The frozen grid contains 36 design cells and 100 replications per cell. It crosses six, eight and ten periods; mean event counts of three, six and twelve; a carryover parameter of zero or .15; and latent autoregression of zero or .40. Each design is estimated using an adjacent model, a fully disjoint lagged-control model and a current-only disjoint model. The sign is expressed in full-price-share points for a 10-point change in promotional share, matching the empirical interpretation.

Each replication follows the same five-step workflow. First, a relationship-specific propensity is generated. Second, the latent propensity either remains serially independent or follows the specified autoregressive condition. Third, finite event counts convert the latent propensity into noisy window-level promotional shares. Fourth, the zero-carryover condition generates subsequent windows without a structural dependence on the previous observed share, whereas the .15 condition introduces modest persistence in the direction of lower later full-price share. Fifth, the three estimators are fitted to the same generated panel and their promotional-share coefficients are rescaled to percentage-point effects for a 10-point change. No cell or specification was selected after inspecting the empirical sign.

The grid is the full Cartesian product 3 panel lengths × 3 mean event counts × 2 carryover settings × 2 latent-autoregression settings = 36 cells. One hundred independent replications per cell yield 3,600 simulated panels and 10,800 fitted models across the three estimators. Cell summaries report the replication mean and the 2.5th and 97.5th percentiles. Table 2 aggregates only across prespecified cells sharing the same generating-process class; it does not pool individual rows or re-estimate a fourth model.

The grid is diagnostic rather than a calibrated reproduction of every feature of the grocery panel. Event frequency controls measurement precision. Panel length changes the number of repeated rows over which the within transformation operates. Latent autoregression represents serial persistence in the underlying relationship propensity. The carryover parameter introduces a genuine negative relationship between current promotional share and later full-price share. The design therefore tests two requirements: whether adjacent construction creates a positive sign under a pure null, and whether it can mask a modest true negative relationship.

Three estimators are retained in each cell. The adjacent specification recreates the period reuse found in the empirical panel. The lagged-control disjoint specification uses separate previous, current and outcome periods in nonoverlapping blocks. The current-only disjoint specification removes the lagged control and therefore checks whether conclusions depend on conditioning on an earlier noisy share. Comparing these estimators is important because disjointness can be achieved in more than one way, and a result that appears only under one lag structure would provide weak evidence of a general construction problem.

The simulation outcome is the mean coefficient across 100 replications in each design cell, accompanied in the frozen output by 2.5th and 97.5th percentiles. The analysis emphasizes sign recovery rather than exact equality to the data-generating parameter. The simplified generated process, estimator controls and finite panels mean that exact numerical recovery is not expected in every cell. The relevant validity tests are whether the adjacent estimator is centred away from zero under the pure null, whether the disjoint estimator is centred near zero in the same condition, and whether true persistence retains its negative direction after period reuse is removed.

The grid is deliberately limited to dimensions that bear directly on the proposed mechanism. Adding many application-specific parameters would create the appearance of realism without improving identification of the reuse channel. Panel length, event frequency, latent persistence and true carryover are sufficient to show when the artifact weakens, persists or overwhelms the intended sign. Missing-data processes and more complex targeting rules remain subjects for future work rather than post hoc additions selected to strengthen the present conclusion.

3.2 Empirical data and sample

The application uses the dunnhumby Complete Journey source files, a household-level retail dataset released for research and teaching. It represents approximately two years of in-store purchases from 2,500 frequent-shopper households at one grocery retailer and includes transactions, products, campaigns, coupon redemption and selected household attributes (dunnhumby, n.d.). The transaction file reports relative day and week identifiers rather than calendar dates. The observed horizon is therefore described as 102 weeks without inferring specific calendar years or national representativeness.

The raw transaction file contains 2,595,732 item lines, 276,484 baskets, 92,339 purchased products and 582 store identifiers. The product master contains 92,353 records across 44 departments, 308 merchandise categories and 2,383 subcategories. Demographic descriptors are available for 801 households. These households are more active than those without demographic records, so demographic information is used only for descriptive profiling and not as a basis for population inference or principal moderation tests.

Records with invalid identifiers, non-positive quantities, infeasible reconstructed regular prices, negative paid values or implausible discount depths are removed. Coupon and miscellaneous lines that do not represent ordinary category demand are excluded. The cleaned transaction file contains 2,551,997 lines. Item lines are aggregated by household, basket, day, store and merchandise category to form 1,876,500 household-category purchase events.

The primary promotion indicator equals one when at least one line in an event carries a retail discount, manufacturer coupon or coupon-match discount. Sensitivity definitions use majority-discounted lines, all-discounted lines, retailer-funded discounts, coupon-funded discounts and value-weighted discount share. Historical promotional share is computed only from events before the current window. Current promotional share is the share in the current window. Conditional on purchasing the category in the next window, the main outcome is the proportion of next-window purchase events made at full price.

3.3 Adjacent and fully disjoint constructions

The primary empirical construction divides 102 weeks into six 17-week windows. Current windows two through five are linked prospectively to the next window. A reliable transition requires at least three current-window purchase events. Categories must provide at least 300 reliable transitions, a promotional-share standard deviation of at least .10 and non-degenerate continuation. These rules yield 123,304 reliable transitions from 2,236 households across 108 categories. Of these, 110,470 continuing transitions contribute 580,337 later purchase events.

The conventional dynamic model uses adjacent household-category rows. Equation (7) gives the common specification: Y_it is the later full-price share associated with row t as defined in Equation (1); X_it and X_i,t-1 are current and previous promotional shares; α_i denotes household-category fixed effects; λ_t denotes category-by-transition effects in the empirical application; and Z_it contains purchase-frequency controls. The coefficient of interest is β. Standard errors are clustered as described below. Because a window can enter the outcome of one row and later become a predictor, the adjacent estimate is treated as a diagnostic rather than a preferred dynamic effect. Short fixed-effects panels also require caution, but standard panel corrections do not repair a regressor-outcome construction that reuses the same noisy summary (Dhaene and Jochmans, 2015; Nickell, 1981; Wooldridge, 2010).

$$Y_{it} = \alpha_i + \lambda_t + \beta X_{it} + \gamma X_{i,t-1} + \boldsymbol{\delta}'\mathbf{Z}_{it} + \varepsilon_{it} \quad (7)$$

The primary fully disjoint lagged-control design uses previous-current-outcome triples 1-2-3 and 4-5-6. No window is reused between the two estimating rows. It contains 23,236 rows representing 11,618 matched pairs from 1,292 households across 108 categories. The main estimate gives equal weight to each relationship; an event-weighted estimate is reported separately. Two current-only disjoint designs omit the lagged promotional-share control. A 13-week construction provides an alternative window-length check.

The central common-sample design identifies 9,014 household-category relationships that can be estimated under both approaches. These relationships contribute 36,056 adjacent rows and 18,028 fully disjoint rows. Re-estimating both models on this identical set removes sample composition as an explanation for a sign change.

3.4 Prospective models and predictive validation

The prospective linear model uses 110,470 continuing transitions and equal relationship weights. Later full-price share is regressed on historical promotional share, current promotional share, current and previous purchase frequency, and category-by-transition fixed effects. Standard errors are two-way clustered by household and category (Cameron and Miller, 2015). A grouped-binomial logit models full-price and promotional event counts among 580,337 later events. Continuation is modelled over all 123,304 reliable transitions, and inverse-probability weights test whether selection into the later-purchase sample drives the estimates.

For prediction, the base model contains historical and earlier promotional shares, purchase frequency, discount depth, category and transition. The augmented model adds current promotional share. The final transition is excluded from training, and the strict evaluation additionally uses households not seen during model fitting. Event-weighted AUC is the primary metric; Brier score, log loss, average precision and calibration are supporting diagnostics. Paired intervals use household-level bootstrap resampling.

Reproducibility is organized around an explicit analysis contract: source windows, eligibility thresholds, outcome horizon, fixed effects, controls, weighting, clustering and scaling are specified above before results are compared. The source transactions are publicly released and de-identified but remain subject to the distributor's access terms and therefore are not redistributed with the manuscript. The accompanying technical archive contains the frozen cell-level Monte Carlo summaries, exact-sample comparisons, disjoint-model estimates and equivalence checks needed to audit every number reported in Tables 2 and 3. Reproduction from raw data requires users to obtain the Complete Journey files directly from dunnhumby and implement the stated cleaning and construction rules.

4 Results

4.1 The simulation isolates the construction artifact

As shown in Table 2, under zero carryover and no latent serial persistence, the adjacent estimator is positive in every summarized design, with cell means ranging from +1.41 to +3.11 percentage points. The fully disjoint cell means remain at the null, from -0.01 to 0.00 points. These are ranges of the 100-replication cell means; they are not replication-level percentile limits.

Table 2. Monte Carlo evidence across generating processes

Generating process	Adjacent mean range	Fully disjoint mean range	Inference
Zero carryover and zero latent AR	1.41 to 3.11	-0.01 to 0.00	Adjacent model creates a positive effect; disjoint model recovers the null.
Zero structural carryover and latent AR .40	1.10 to 2.86	-0.31 to -0.28	Adjacent bias remains positive; disjoint estimate reflects the latent serial component.

Generating process	Adjacent mean range	Fully disjoint mean range	Inference
True persistence parameter .15	-0.35 to 1.80	-1.83 to -1.49	Adjacent construction attenuates, masks or reverses persistence; disjoint construction retains the negative direction.

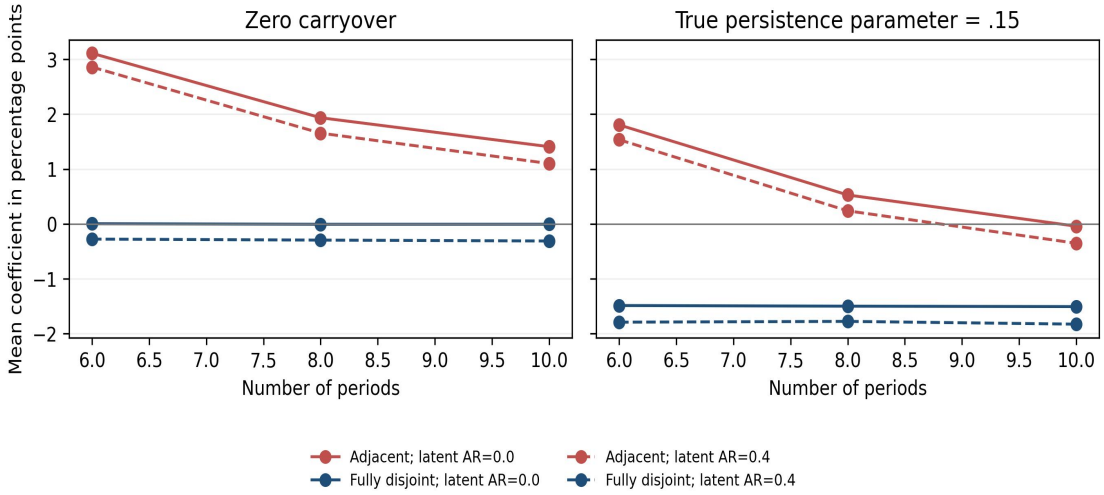

Figure 1. Adjacent and fully disjoint estimates across the frozen Monte Carlo grid

When the structural carryover parameter remains zero but the latent propensity follows an autoregressive process of .40, adjacent cell means remain positive, ranging from +1.10 to +2.86 points. Fully disjoint cell means are small and negative, ranging from -0.31 to -0.28 points, reflecting the latent serial component without the shared-window channel.

When true persistence is introduced, the fully disjoint cell means consistently recover a negative relationship, ranging from -1.83 to -1.49 points across the design grid. The adjacent cell means range from -0.35 to +1.80 points. Adjacent construction therefore attenuates the negative relationship, moves it toward zero or reverses its sign, depending on panel length and latent persistence. The general lesson is not merely that the adjacent model is biased upward under a null. In short panels, the reuse channel can be large enough to conceal the direction of a modest true dynamic.

4.2 Identical-sample construction reverses the empirical conclusion

Table 3 reports the common-sample contrast, which provides the most direct empirical test. On the same 9,014 household-category relationships, the adjacent estimator indicates +1.78 percentage points of apparent recovery, whereas the fully disjoint estimator indicates -1.40 points of persistence.

Because relationship membership is identical, the reversal cannot be attributed to different households, categories or baseline relationship composition. Temporal construction is the changed element. The result answers RQ3: the substantive conclusion can reverse when the same customer relationships are represented through adjacent rather than disjoint panel rows.

Table 3. Identical-sample comparison of adjacent and fully disjoint estimators

Estimator	Pairs	Rows	Effect per 10-point increase	95% confidence interval
Adjacent, exact common sample	9,014	36,056	+1.78 points	[+1.56, +2.00]
Fully disjoint, exact common sample	9,014	18,028	-1.40 points	[-1.72, -1.08]

4.3 Null benchmarking and disjoint robustness support the same diagnosis

In the full adjacent fixed-effects analysis, a 10-point increase in current promotional share is followed by a 2.49-point higher later full-price share (95% CI [2.31, 2.67]). The sign grows as the constructed windows lengthen:

+1.26 points at 13 weeks, +2.49 at 17 weeks and +5.33 at 26 weeks. This pattern is inconsistent with treating the coefficient as a stable behavioural recovery parameter.

Two hundred no-carryover simulations preserve shrunken relationship propensities, purchase counts, missing-window patterns and category-window conditions while generating purchases independently across windows. They produce an average adjacent coefficient of +3.52 points, with a 95% simulation interval [3.43, 3.62]. A positive coefficient larger than the observed estimate therefore arises in data deliberately constructed without temporal carryover.

The primary 17-week fully disjoint lagged-control estimate is -1.20 points per 10-point increase (95% CI [-1.51, -.89]). Event weighting yields -1.49 points (95% CI [-1.84, -1.14]). Current-only disjoint checks are -.78 and -.61 points. Fully disjoint 13-week lagged-control estimates are -1.54 and -1.22 points. The frozen specifications differ in weighting, windows and lag controls, yet retain the negative direction. The evidence supports modest persistence, not the positive recovery suggested by adjacent construction.

4.4 Promotion history remains descriptively informative

The construction audit does not make promotion history useless. Among continuing relationships, a 10-point higher historical promotional share is associated with a 1.59-point lower later full-price share (95% CI [-1.72, -1.46]). Holding history constant, a 10-point higher current promotional share is associated with a 3.34-point lower later full-price share (95% CI [-3.56, -3.12]). The grouped-binomial model gives the same direction: the odds ratio is .921 (95% CI [.915, .927]) for historical share and .833 (95% CI [.821, .846]) for current share.

Historical share remains negative under 13-, 17- and 26-week constructions. In the continuation model, historical and current promotional-share coefficients are -.01 and +.04 percentage points per 10-point increase, respectively, and neither is statistically significant. Inverse-probability weighting leaves the pooled price-composition estimates nearly unchanged. The feature therefore contains durable descriptive information about later price composition, but the observational relationship should not be interpreted as an effect of promotion exposure.

4.5 Incremental prediction is statistically stable but practically small

In the strict final-period holdout on households excluded from training, adding current promotional share raises AUC from .7548 to .7611. The increase is .0063, with a 95% household-bootstrap interval [.0030, .0099]. Average precision rises from .7072 to .7120, and Brier score and log loss improve modestly. Calibration remains close to one.

The result answers RQ4 with an important qualification. Recent promotional share adds information, but it does not transform predictive performance. The feature may be retained as one input in a broader model; the evidence does not justify a stand-alone targeting rule or a claim of substantial commercial value.

5 Discussion

5.1 Contribution to market-research methodology

The paper identifies a validity problem that arises before substantive interpretation. Researchers often recognize measurement error, short-panel bias and unobserved heterogeneity as separate concerns. Shared-window reuse links them operationally. The same estimated period share is embedded in adjacent rows, and the within transformation converts that reuse into an apparent dynamic. Fixed effects can therefore remove stable relationship levels while preserving a construction-induced sign.

The formal argument, Monte Carlo grid and common-sample reversal provide complementary evidence. The formal argument establishes why a positive sign is possible under an independent-window null. The simulation shows that the pattern survives variation in panel length and event frequency, and that it can conceal modest persistence. The common-sample analysis demonstrates that the problem is consequential in observed customer data rather than only algebraically possible.

The resulting four-stage audit is the principal methodological contribution. A null benchmark asks whether the observed sign can arise without the proposed process. Fully disjoint construction removes direct period reuse. An identical-sample contrast prevents changing eligibility from being confused with changing estimators. A future holdout then determines whether the feature remains useful for prediction. Together, these stages prevent a feature from moving silently from descriptive association to dynamic mechanism and then to intervention.

This contribution differs from a generic recommendation to perform robustness tests. Each stage targets a named failure mode and has a corresponding decision rule. If the no-carryover benchmark reproduces the sign, the coefficient is not specific to the hypothesized dynamic. If disjoint construction changes the sign, the adjacent result is construction sensitive. If the reversal remains on the same relationships, sample composition cannot explain it. If the feature adds little in a future unseen-customer holdout, statistical significance should not be translated into strong operational value. The framework therefore links diagnostics to permissible claims.

5.2 What the grocery application does and does not show

The application demonstrates that promotion history is mainly a persistent marker of price composition at the focal retailer. This is compatible with heterogeneous value consciousness, search, price knowledge, offer access

and category-specific deal responsiveness (Bell and Lattin, 1998; Gauri et al., 2008; Lichtenstein et al., 1990, 1993, 1995). It also accords with evidence that promotional buyers need not become more loyal after a promotional purchase (Dawes, 2018). The unit of analysis is a household-category relationship, so the signal should not be treated as a general psychological trait.

The negative fully disjoint association is compatible with short-run persistence, including reference-price or repeated-offer processes. It does not identify them. Offer availability and redemption are jointly selected; eligibility is not observed; competitor purchases, budgets and household inventory are missing. The study therefore cannot distinguish consumer learning from retailer targeting or changing category need. Thus, the positive adjacent estimate is not acceptable evidence of mean reversion. The negative disjoint estimate should not be recast as proof that promotions created dependence.

Inventory effects and purchase acceleration remain established parts of promotional response (Chandon and Wansink, 2002; Gupta, 1988; Mela et al., 1998). Their existence, however, does not make a positive coefficient from adjacent windows evidence of post-promotion recovery. In these data, a positive sign appears even in simulated panels with no carryover. Once the windows are kept separate, the sign turns negative.

5.3 Implications for analytics governance

For a marketing manager, the risk is a wrong campaign decision. The positive adjacent estimate might be read as a reason to stop offers to promotion-heavy households on the assumption that they will soon pay full price. Yet the estimate from the same 9,014 relationships changes direction when the periods do not reuse data. The negative estimate presents a different danger: it might be treated as proof that more promotions will create continued deal buying. This study does not establish that causal effect. The audit therefore asks managers to state what they want to claim before using a coefficient to change an offer rule. That discipline matters in customized promotion systems, where the useful test is profit and customer response, not the size of an isolated coefficient (Chaudhry et al., 2024).

The first task is to trace the data used to build each feature. Feature documentation should state which source periods enter each constructed predictor and outcome. Validation should flag any period that changes role across adjacent rows, particularly when predictor and outcome are complements or share a common denominator. This check belongs upstream of significance testing.

Second, model reports should identify the claim level being evaluated. A prospective association supports description. Incremental future performance supports prediction. A dynamic explanation additionally requires construction-valid temporal evidence. An intervention claim requires treatment identification. Presenting these levels separately would reduce the tendency to label customers in causal language merely because a historical feature predicts an outcome.

Third, alternative temporal designs should be compared on the same observational units where possible. A disjoint design often sacrifices data and changes eligibility. Without a common-sample contrast, an analyst cannot know whether a changed estimate reflects temporal construction or a different population. The 9,014-pair reversal shows why this control is substantive rather than cosmetic.

Fourth, small performance increments require restraint. An AUC improvement of .0063 may be statistically stable but does not establish economic value. Before release, decision owners should prespecify the business metric and minimum material increment; compare expected margin with contact, discount and opportunity costs; test whether the rule changes who receives an offer; and monitor realized uplift and customer outcomes. Promotion history may assist segmentation, but it should be combined with eligibility, inventory, availability and profitability information rather than used as an automatic suppression or escalation rule.

Finally, causal language should be treated as a governance risk. Terms such as promotion dependent or trained to wait for deals imply a mechanism and can lead to consequential decisions. Unless exposure is randomized or otherwise identified, a safer description is observed promotional orientation. This wording preserves the feature's descriptive value without claiming that the retailer created the behaviour.

An implementable governance workflow can be organized around five records. The first is a feature-lineage map showing source events, aggregation windows and information cutoffs. The second is an overlap matrix identifying periods shared by predictors, outcomes and repeated rows. The third is a claim register classifying each intended use as descriptive, predictive, explanatory or intervention oriented. The fourth is a validation record matching each claim to its required evidence and documenting same-sample sensitivity. The fifth is a deployment record stating the business metric, decision threshold and monitoring plan. This workflow complements broader calls for reproducible model documentation and validated preprocessing (Kapoor et al., 2024; Jimenez Garcia and Gil Herrera, 2026).

Such review is especially important when models are updated frequently. A feature definition may change from a six-month history to a three-month rolling window, or the prediction horizon may move closer to the feature window. These changes can alter overlap and reuse without changing the variable name. Version control for feature definitions should therefore be treated as part of research validity rather than only engineering maintenance. Revalidation is necessary whenever the source window, outcome horizon or row-construction rule changes.

The framework also clarifies responsibilities. Data engineers can document lineage and overlap. Analysts can run null and disjoint diagnostics. Domain experts can assess whether a descriptive label implies an unsupported mechanism. Decision owners can determine whether the predictive increment is economically material. Separating these responsibilities reduces the chance that a technically significant coefficient becomes a policy rule without scrutiny at the explanation and intervention levels.

5.4 Relationship to existing methodological concerns

The findings extend established warnings about separating state dependence from persistent heterogeneity. Models can overstate loyalty, learning or price sensitivity when preference information is incomplete (Dubé et al., 2010; Horsky et al., 2006; Shin et al., 2012). Levine and Seiler (2023) show that a conventional temporal model can imply state dependence where plausibly exogenous stockout variation does not. The present study identifies a more elementary but practically common vulnerability: the construction of the panel itself can generate a dynamic sign before the analyst confronts endogeneity of exposure.

The paper also complements work on temporal aggregation. Aggregation can change marketing-response estimates because it alters variation and timing (Franses, 2021). Shared-window reuse is more specific. It concerns the repeated use of the same estimated summary across adjacent row roles. The remedy is therefore not simply shorter or longer windows. It is explicit nonreuse, null benchmarking and same-sample comparison.

6 Limitations and future research

The analytic expression is deliberately simplified. It assumes a balanced sequence, independent window errors and a direct complement relationship. The empirical estimators contain controls, unequal event counts, missing windows and category-time conditions. The expression establishes the sign mechanism, while the simulation evaluates selected departures. Future methodological work should derive finite-sample properties under unbalanced panels, varying denominators and multivariate feature construction.

The simulation grid is broader than a single calibrated null but is not exhaustive. The grid changes panel length, event frequency, carryover and latent autoregression. It does not model every form of missing data, nonlinear response, retailer targeting or machine-learning workflow. Further tests should consider rolling features that share part of a window, alternative denominators and outcomes that are not exact complements.

The empirical application uses transactions from a single grocery retailer. Purchases made at competing retailers, online or through other channels are absent. The estimated promotion effects, the observed distribution of promotional orientation and the .0063 gain in AUC may therefore differ elsewhere; none should be transferred without replication. In this study, full-price share refers only to purchases recorded by the focal retailer. Promotion status is inferred from purchased lines, and the data contain no information on unredeemed offers, competitor prices, household budgets, search effort or home inventory. These gaps rule out causal claims and restrict generalization to a wider customer population.

The mechanism is not restricted to grocery data, but its reach has limits. It becomes relevant when the same noisy period estimate serves in adjacent rows as both an outcome component and a predictor. The risk is greater in short fixed-effects panels and when variables are complements or use the same denominator. This does not mean that every lagged variable is biased, nor that long panels or genuinely non-overlapping designs have the same problem. The algebra and the audit travel across settings; the frequency and size of the distortion do not. Useful replications would include several retailers, digital and omnichannel histories, outcomes that are not complements, and rolling windows with partial overlap.

Household demographics are available for 801 of the 2,500 households. These households also shop more actively, so the descriptors cannot support population-wide claims about heterogeneity. A further constraint is the 102-week observation period, which yields only a small number of non-overlapping panel rows. A longer panel would permit more precise estimates of relationship-level dynamics without recycling a source window.

Our outcome concerns the mix of full-price and promoted purchases. It does not measure profit, category incidence, quantity or switching between brands. The gain of .0063 in AUC is small enough that any use in practice needs a separate economic test. Future research should test whether construction-valid features improve contribution margin or customer outcomes, ideally within randomized offer policies. A second independent retailer panel would also provide a stronger test of external generality.

7 Conclusion

Temporal customer features can be useful and still be unsafe to interpret dynamically. When the same noisy window estimate enters an outcome and later reappears as a predictor, fixed-effects demeaning in a short panel can create a positive association under zero carryover. A 36-cell simulation grid shows that the problem persists across panel lengths and event frequencies and can mask modest true persistence.

The grocery application demonstrates the consequence. On exactly the same 9,014 household-category relationships, adjacent construction implies +1.78 points of recovery toward full-price purchasing, while fully disjoint construction implies -1.40 points of persistence. The broader adjacent result is reproduced by a no-carryover simulation, whereas disjoint checks retain the negative direction. Promotion history remains

prospectively informative, but recent history adds only .0063 to holdout AUC.

The appropriate response is not to discard customer-history features. It is to govern their claims. Null benchmarking, temporal disjointness, identical-sample comparison and future holdout validation distinguish forecasting value from dynamic interpretation. Observational customer histories may support segmentation; they do not automatically explain behaviour, and they do not identify the effects of changing offers. Market-research practice should validate each claim at the level on which it will be used.

References

1. Ailawadi KL, Harlam BA, César J and Trounce D (2006) Promotion profitability for a retailer: The role of promotion, brand, category and store characteristics. Journal of Marketing Research 43(4): 518–535. https://doi.org/10.1509/jmkr.43.4.518
2. Agag G, Shehawy YM, Almoraish A et al. (2024) Understanding the relationship between marketing analytics, customer agility, and customer satisfaction: A longitudinal perspective. Journal of Retailing and Consumer Services 77: 103663. https://doi.org/10.1016/j.jretconser.2023.103663
3. Anderson ET and Fox EJ (2019) How price promotions work: A review of practice and theory. In: Dubé J-P and Rossi PE (eds) Handbook of the Economics of Marketing, vol. 1. Amsterdam: Elsevier, pp. 497–552. https://doi.org/10.1016/bs.hem.2019.04.006
4. Bell DR, Chiang J and Padmanabhan V (1999) The decomposition of promotional response: An empirical generalization. Marketing Science 18(4): 504–526. https://doi.org/10.1287/mksc.18.4.504
5. Bell DR and Lattin JM (1998) Shopping behavior and consumer preference for store price format: Why "large basket" shoppers prefer EDLP. Marketing Science 17(1): 66–88. https://doi.org/10.1287/mksc.17.1.66
6. Briesch RA, Krishnamurthi L, Mazumdar T et al. (1997) A comparative analysis of reference price models. Journal of Consumer Research 24(2): 202–214. https://doi.org/10.1086/209505
7. Bruno HA, Cebollada J and Chintagunta PK (2018) Targeting Mr. or Mrs. Smith: Modeling and leveraging intrahousehold heterogeneity in brand choice behavior. Marketing Science 37(4): 631–648. https://doi.org/10.1287/mksc.2018.1088
8. Cameron AC and Miller DL (2015) A practitioner's guide to cluster-robust inference. Journal of Human Resources 50(2): 317–372. https://doi.org/10.3368/jhr.50.2.317
9. Chan T, Narasimhan C and Zhang Q (2008) Decomposing promotional effects with a dynamic structural model of flexible consumption. Journal of Marketing Research 45(4): 487–498. https://doi.org/10.1509/jmkr.45.4.487
10. Chandon P and Wansink B (2002) When are stockpiled products consumed faster? A convenience-salience framework of postpurchase consumption incidence and quantity. Journal of Marketing Research 39(3): 321–335. https://doi.org/10.1509/jmkr.39.3.321.19111
11. Churchill GA Jr (1979) A paradigm for developing better measures of marketing constructs. Journal of Marketing Research 16(1): 64–73. https://doi.org/10.1177/002224377901600110
12. Chaudhry A, Heilman C and Seetharaman PB (2024) Measuring the effects of customized targeted promotions on retailer profits: Prescriptive analytics using basket-level econometric analysis. Journal of Marketing Analytics 12(4): 1028–1051. https://doi.org/10.1057/s41270-023-00253-3
13. Daljord Ø, Mela CF, Roos JMT et al. (2023) The design and targeting of compliance promotions. Marketing Science 42(5): 866–891. https://doi.org/10.1287/mksc.2022.1420
14. Dawes JG (2018) Price promotions: Examining the buyer mix and subsequent changes in purchase loyalty. Journal of Consumer Marketing 35(4): 366–376. https://doi.org/10.1108/JCM-03-2017-2134
15. Dhaene G and Jochmans K (2015) Split-panel jackknife estimation of fixed-effect models. Review of Economic Studies 82(3): 991–1030. https://doi.org/10.1093/restud/rdv007
16. Diamantopoulos A and Winklhofer HM (2001) Index construction with formative indicators: An alternative to scale development. Journal of Marketing Research 38(2): 269–277. https://doi.org/10.1509/jmkr.38.2.269.18845
17. Dickson PR and Sawyer AG (1990) The price knowledge and search of supermarket shoppers. Journal of Marketing 54(3): 42–53. https://doi.org/10.1177/002224299005400304
18. Dubé J-P, Hitsch GJ and Rossi PE (2010) State dependence and alternative explanations for consumer inertia. RAND Journal of Economics 41(3): 417–445. https://doi.org/10.1111/j.1756-2171.2010.00106.x
19. dunnhumby (n.d.) Source files: The Complete Journey. Available at: https://www.dunnhumby.com/source-files/ (accessed 13 August 2026).
20. Fisher G, McGranaghan M, Liaukonyte J et al. (2023) Price promotions, beneficiary framing and mental accounting. Quantitative Marketing and Economics 21(2): 147–181. https://doi.org/10.1007/s11129-023-09261-0
21. Franses PH (2021) Marketing response and temporal aggregation. Journal of Marketing Analytics 9: 111–117. https://doi.org/10.1057/s41270-020-00102-7
22. Gabel S and Guhl D (2022) Comparing the effectiveness of rewards and individually targeted coupons in

loyalty programs. Journal of Retailing 98(3): 395–411. https://doi.org/10.1016/j.jretai.2021.08.001
23. Gauri DK, Sudhir K and Talukdar D (2008) The temporal and spatial dimensions of price search: Insights from matching household survey and purchase data. Journal of Marketing Research 45(2): 226–240. https://doi.org/10.1509/jmkr.45.2.226
24. Gupta S (1988) Impact of sales promotions on when, what and how much to buy. Journal of Marketing Research 25(4): 342–355. https://doi.org/10.1177/002224378802500402
25. Gupta S and Cooper LG (1992) The discounting of discounts and promotion thresholds. Journal of Consumer Research 19(3): 401–411. https://doi.org/10.1086/209310
26. Haviv A (2022) Consumer search, price promotions and counter-cyclic pricing. Marketing Science 41(2): 294–314. https://doi.org/10.1287/mksc.2021.1327
27. Horsky D, Misra S and Nelson P (2006) Observed and unobserved preference heterogeneity in brand-choice models. Marketing Science 25(4): 322–335. https://doi.org/10.1287/mksc.1050.0192
28. Kalwani MU, Yim CK, Rinne HJ et al. (1990) A price expectations model of customer brand choice. Journal of Marketing Research 27(3): 251–262. https://doi.org/10.1177/002224379002700301
29. Jimenez Garcia MA and Gil Herrera RJ (2026) A systematic framework for preprocessing validation in business analytics. Journal of Business Analytics. Published online 7 April. https://doi.org/10.1080/2573234X.2026.2634271
30. Kopalle PK, Mela CF and Marsh L (1999) The dynamic effect of discounting on sales: Empirical analysis and normative pricing implications. Marketing Science 18(3): 317–332. https://doi.org/10.1287/mksc.18.3.317
31. Kapoor S and Narayanan A (2023) Leakage and the reproducibility crisis in machine-learning-based science. Patterns 4(9): 100804. https://doi.org/10.1016/j.patter.2023.100804
32. Kapoor S, Cantrell EM, Peng K et al. (2024) REFORMS: Consensus-based recommendations for machine-learning-based science. Science Advances 10(18): eadk3452. https://doi.org/10.1126/sciadv.adk3452
33. Lamei A, Mohammadi Darani M, Wiggins J et al. (2025) Can mobile promotion automation decrease overall purchases? The role of promotion pattern predictability in customer habit formation. Journal of Retailing and Consumer Services 82: 104138. https://doi.org/10.1016/j.jretconser.2024.104138
34. Lattin JM and Bucklin RE (1989) Reference effects of price and promotion on brand choice behavior. Journal of Marketing Research 26(3): 299–310. https://doi.org/10.1177/002224378902600304
35. Levine J and Seiler S (2023) Identifying state dependence in brand choice: Evidence from hurricanes. Marketing Science 42(5): 934–957. https://doi.org/10.1287/mksc.2022.1415
36. Lichtenstein DR, Netemeyer RG and Burton S (1990) Distinguishing coupon proneness from value consciousness: An acquisition-transaction utility theory perspective. Journal of Marketing 54(3): 54–67. https://doi.org/10.1177/002224299005400305
37. Lichtenstein DR, Netemeyer RG and Burton S (1995) Assessing the domain specificity of deal proneness: A field study. Journal of Consumer Research 22(3): 314–326. https://doi.org/10.1086/209453
38. Lichtenstein DR, Ridgway NM and Netemeyer RG (1993) Price perceptions and consumer shopping behavior: A field study. Journal of Marketing Research 30(2): 234–245. https://doi.org/10.1177/002224379303000208
39. Mayhew GE and Winer RS (1992) An empirical analysis of internal and external reference prices using scanner data. Journal of Consumer Research 19(1): 62–70. https://doi.org/10.1086/209286
40. Mela CF, Gupta S and Lehmann DR (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
41. Menukhin O, Mandungu C, Shahgholian A and Mehandjiev N (2025) Guiding the integration of analytics in business operations through a maturity framework. Annals of Operations Research 348(3): 2017–2047. https://doi.org/10.1007/s10479-023-05614-w
42. Mela CF, Jedidi K and Bowman D (1998) The long-term impact of promotions on consumer stockpiling behavior. Journal of Marketing Research 35(2): 250–262. https://doi.org/10.1177/002224379803500210
43. Neslin SA, Henderson C and Quelch J (1985) Consumer promotions and the acceleration of product purchases. Marketing Science 4(2): 147–165. https://doi.org/10.1287/mksc.4.2.147
44. Nickell SJ (1981) Biases in dynamic models with fixed effects. Econometrica 49(6): 1417–1426. https://doi.org/10.2307/1911408
45. Nigam A, Dewani P, Behl A et al. (2022) Consumer's response to conditional promotions in retailing: An empirical inquiry. Journal of Business Research 144: 751–763. https://doi.org/10.1016/j.jbusres.2022.02.051
46. Nijs VR, Dekimpe MG, Steenkamp J-BEM et al. (2001) The category-demand effects of price promotions. Marketing Science 20(1): 1–22. https://doi.org/10.1287/mksc.20.1.1.10197
47. Osuna I, González J and Capizzani M (2016) Which categories and brands to promote with targeted coupons to reward and to develop customers in supermarkets. Journal of Retailing 92(2): 236–251.

https://doi.org/10.1016/j.jretai.2015.12.002
48. Papatla P and Krishnamurthi L (1996) Measuring the dynamic effects of promotions on brand choice. Journal of Marketing Research 33(1): 20–35. https://doi.org/10.1177/002224379603300103
49. Park CH, Park YH and Schweidel DA (2018) The effects of mobile promotions on customer purchase dynamics. International Journal of Research in Marketing 35(3): 453–470. https://doi.org/10.1016/j.ijresmar.2018.05.001
50. Pauwels K, Hanssens DM and Siddarth S (2002) The long-term effects of price promotions on category incidence, brand choice and purchase quantity. Journal of Marketing Research 39(4): 421–439. https://doi.org/10.1509/jmkr.39.4.421.19114
51. Seetharaman PB (2004) Modeling multiple sources of state dependence in random utility models: A distributed lag approach. Marketing Science 23(2): 263–271. https://doi.org/10.1287/mksc.1030.0024
52. Rosenblatt M, Tejavibulya L, Jiang R et al. (2024) Data leakage inflates prediction performance in connectome-based machine learning models. Nature Communications 15: 1829. https://doi.org/10.1038/s41467-024-46150-w
53. Shin S, Misra S and Horsky D (2012) Disentangling preferences and learning in brand choice models. Marketing Science 31(1): 115–137. https://doi.org/10.1287/mksc.1110.0680
54. Shmueli G (2010) To explain or to predict? Statistical Science 25(3): 289–310. https://doi.org/10.1214/10-STS330
55. van Heerde HJ, Gupta S and Wittink DR (2003) Is 75% of the sales promotion bump due to brand switching? No, only 33% is. Journal of Marketing Research 40(4): 481–491. https://doi.org/10.1509/jmkr.40.4.481.19386
56. Venkatesan R and Farris PW (2012) Measuring and managing returns from retailer-customized coupon campaigns. Journal of Marketing 76(1): 76–94. https://doi.org/10.1509/jm.10.0162
57. Winer RS (1986) A reference price model of brand choice for frequently purchased products. Journal of Consumer Research 13(2): 250–256. https://doi.org/10.1086/209064
58. Wooldridge JM (2010) Econometric Analysis of Cross Section and Panel Data. 2nd edn. Cambridge, MA: MIT Press.
59. Yarkoni T and Westfall J (2017) Choosing prediction over explanation in psychology: Lessons from machine learning. Perspectives on Psychological Science 12(6): 1100–1122. https://doi.org/10.1177/1745691617693393
60. Zhang H, Zhou K, Zhu T et al. (2025) The effects of mobile coupons on consumer buying behavior: A dynamic perspective based on hidden Markov model. Journal of Consumer Behaviour 24(5): 2372–2383. https://doi.org/10.1002/cb.70012