Causal Inference and Uplift Modeling on Integrated Customer 360 Data for Targeted Personalization Strategies in B2C Digital Retail Platforms

Nimal Perera[1], Sajith Fernando[2], Roshan Abeysekera[3]

Abstract
Business-to-consumer digital retail platforms generate extensive observational traces of browsing, search, transaction, messaging, and support interactions, which can be integrated into customer 360 representations. These representations combine identifiers, event histories, inferred preferences, and contextual attributes into longitudinal profiles capable of supporting targeted personalization. Despite the availability of such data, many operational strategies remain based on response prediction or heuristic segmentation, which can systematically conflate correlation with causal impact and lead to inefficient use of incentives, exposure, and capacity. This paper examines a technical framework for causal inference and uplift modeling built directly on integrated customer 360 data with the objective of estimating heterogeneous treatment effects and deploying stable, auditable targeting policies. The discussion focuses on definition of exposure units, temporal alignment of features and outcomes, assumptions for identification in mixed experimental and observational regimes, and the use of orthogonal, doubly robust, and policy-learning methods that operate under budget and operational constraints. Attention is given to the interaction between model structure, identity resolution strategies, and multi-channel treatment assignment, as well as to mechanisms for drift detection, overlap monitoring, and fairness-aware analysis. The framework is intended to be implementable in production environments that require strict latency, governance, and privacy controls, while remaining explicit about assumptions and sensitivities. The paper is descriptive rather than promotional, outlining a set of consistent design choices and analytical components that can be combined to support cautious deployment of causal personalization in B2C digital retail platforms.

[1]University of Colombo, Department of Business Administration, 94 Cumaratunga Munidasa Mawatha, Colombo 00300, Sri Lanka
[2]University of Kelaniya, Department of Finance, Dalugama, Kelaniya 11600, Sri Lanka
[3]University of Sri Jayewardenepura, Department of Management Studies, Gangodawila, Nugegoda 10250, Sri Lanka

Contents

1. Introduction

Business-to-consumer digital retail platforms have evolved into complex socio-technical systems in which each interaction, from a homepage visit to a purchase confirmation or support contact, is instrumented, timestamped, and archived [1]. These systems observe sequences of page views, search queries, product detail inspections, cart

Concept	Definition	Challenge	Example
Customer 360	Unified profile integrating events	Identity uncertainty	Cross-device linkage
Predictive Models	Estimate purchase probability	Confound outcome vs. effect	Over-targeting high-propensity users
Causal Models	Estimate treatment effects	Identification limits	Targeting uplift-positive users
Uplift Objective	Incremental impact of treatment	Overlap, bias	Discount allocation

Table 1. Contrast between predictive and causal personalization paradigms.

Data Layer	Input Source	Identifier	Risk
Raw Events	Web/mobile/app logs	Cookies, Device IDs	Fragmentation
Identity Resolution	Cross-source mapping	Hashed emails, Loyalty IDs	Mis-linkage
Feature Layer	Aggregated histories	Stable person key	Temporal leakage
Outcome Layer	Post-treatment metrics	Session/event time	Censoring bias

Table 2. Customer 360 architecture layers and associated risks.

edits, payments, returns, loyalty events, email opens, notification interactions, and helpdesk resolutions. Customer 360 architectures emerged as a pragmatic response to the fragmentation of this information across devices, channels, and internal services. They integrate heterogeneous identifiers and event streams into longitudinal profiles intended to provide a unified view of customer behavior, preferences, and value. In practice, this integration is constrained by identity uncertainty, partial observability, regulatory limits, and engineering tradeoffs, yet it remains sufficiently informative to support a wide range of personalization tasks, including ranking products, allocating discounts, selecting creatives, and orchestrating contact strategies. [2]

Despite the richness of integrated customer 360 data, many personalization and marketing systems are optimized using objectives and methodologies that are primarily predictive rather than causal. Conventional uplift-insensitive approaches train models to estimate quantities such as conversion probability, short-term revenue, or engagement under existing policies, and then treat high-probability or high-revenue predictions as high-value targets. This design conflates outcomes that would occur without intervention with outcomes induced by intervention. Customers with high baseline purchase propensity are often prioritized for expensive promotions, while customers whose behavior might be meaningfully shifted by targeted incentives may be overlooked if they are less predictable or historically underexposed [3]. Over time, such strategies can systematically distort resource allocation, skew apparent campaign performance metrics, and create feedback loops wherein model-driven policies reinforce the very patterns on which they were trained.

Causal inference and uplift modeling provide a formal framework for separating correlation from incremental impact. Instead of asking whether a customer is likely to purchase, these methods seek to estimate how that probability would change if the customer were treated versus not treated, under clearly defined interventions. The central object of interest is the heterogeneous treatment effect, defined at the level of an exposure unit that couples a decision opportunity with a specific customer and context [4]. This perspective acknowledges that not all customers respond similarly to the same intervention and that the value of personalization is realized when treatments are directed where their causal effect is positive and sufficiently large relative to cost and constraints. However, estimating such heterogeneous effects reliably in operational retail environments requires resolving a series of modeling, identification, and systems challenges that arise from the use of mixed experimental and observational data, from complex assignment logic, and from the dynamics of evolving platforms.

The practical environment of B2C digital retail complicates naive applications of causal methods in several ways. Treatments are not limited to single, cleanly randomized campaigns; instead, customers may simultaneously encounter multiple overlapping interventions such as banners, recommendations, email offers, mobile push notifications, and loyalty nudges, often governed by handcrafted rules, prioritization policies, or machine-learned rankers [5]. Assignment mechanisms are frequently only partially documented and can change over time in response to business priorities. Furthermore, observed logs encode the results of past optimization regimes, meaning that the data are generated under policies that selectively expose certain subpopulations to certain actions. This induces selection effects that violate the assumptions of simple observational estimators unless explicitly addressed. At the same time, large-scale platforms do run randomized experiments, but these experiments may focus on narrow variants or short windows, and their results must be carefully integrated with broader observational histories to obtain useful coverage of the feature space. [6]

Customer 360 representations interact with causal objectives in nontrivial ways. Identity resolution procedures define which events are assigned to which units, and any mis-linkage or fragmentation alters both estimated propensities and outcomes. Feature engineer-

Exposure Type	Outcome Window	Metric	Decision Unit
Email Promotion	7 days	Conversion rate	Customer-email pair
Onsite Banner	Same session	Click/Add-to-cart	Page view
Push Notification	24 hours	Engagement	App session
Loyalty Offer	14 days	Redemption	Customer-period

Table 3. Examples of exposure units and temporal outcome definitions.

Model Stage	Input	Output	Key Operation
Nuisance Estimation	X, T, Y	$\hat{e}(x), \hat{\mu}_t(x)$	Cross-validation folds
Orthogonalization	$\hat{e}, \hat{\mu}$	ψ pseudo-outcome	Variance truncation
Uplift Learning	ψ, X	$\hat{\tau}(x)$	Ensemble training
Policy Derivation	$\hat{\tau}, c(x)$	$\pi_\kappa(x)$	Threshold tuning

Table 4. Stages of causal uplift model training pipeline.

ing pipelines compress behavior, demographics, context, and history into covariates used both for treatment assignment and for effect estimation, making them central to any claim of conditional exchangeability. Choices about time windows, aggregation functions, and inclusion of cross-channel interactions determine whether important confounders are captured or omitted [7]. Temporal alignment rules that constrain features to pre-exposure information and outcomes to post-exposure windows must be strictly enforced to avoid subtle forms of data leakage that can inflate apparent performance or bias effect estimates. A coherent introduction to uplift modeling for digital retail must, therefore, treat these representational and temporal design decisions as foundational rather than peripheral.

The motivation for integrating causal inference with customer 360 data is not to assert that every personalization decision can be perfectly optimized, but to enable a disciplined treatment of what can be learned from available data under transparent assumptions. In particular, the use of heterogeneous treatment effect estimates for targeting must be grounded in explicit identification strategies [8]. In purely randomized settings, the link between treatment and outcomes is straightforward, and uplift learning reduces to exploiting experimental variation to uncover segments with different responses. In observational or hybrid settings, plausible causal interpretation depends on capturing the main drivers of assignment in the feature set and on monitoring violations of overlap, where certain customers are almost always or almost never treated. When such violations are severe, it may be more appropriate to report partial or local effects, introduce additional exploration to restore support, or restrict policies to regions of the feature space with adequate coverage. The introduction of these distinctions is essential for avoiding unqualified use of uplift scores as if they were universally valid. [9]

The targeting problem becomes more intricate once costs, budgets, capacity limits, and exposure constraints are considered. Digital retailers often operate under explicit or implicit limits on promotional spend, number of contacts per customer over given intervals, available impressions in key placements, and inventory levels for promoted items. Under such conditions, uplift modeling is coupled with a policy learning problem in which the aim is to allocate treatments where expected incremental benefit per unit cost is highest, while satisfying a set of structural constraints. The presence of constraints converts uplift estimation from a purely statistical exercise into an input for optimization, requiring that estimates be sufficiently stable and calibrated to support ranking decisions [10]. It also suggests the need to quantify uncertainty, so that policies can be tuned conservatively and adjusted as more evidence accumulates.

Deployment and lifecycle considerations influence the design of causal personalization systems from the outset. Real-time decision services face latency and reliability requirements that limit the complexity of models and feature computations that can be executed synchronously. This motivates the separation of concerns between offline pipelines that estimate uplift functions using flexible approaches such as ensemble learners or orthogonal methods, and online services that apply distilled versions of these models using a constrained and well-monitored feature set. Furthermore, continuous experimentation and exploration are needed to preserve the ability to learn from new data, detect drift, and reassess the validity of identification assumptions [11]. Exploration, in turn, must be integrated with the policy that exploits existing uplift estimates, so that the combined system maintains sufficient overlap without introducing uncontrolled variability in customer experience.

Fairness, interpretability, and governance provide another dimension that shapes introductory formulations of causal uplift modeling in retail. When treatments include discounts, differential service levels, or visibility advantages, systematic disparities in allocation across demographic or behavioral groups may raise regulatory, ethical, or reputational concerns. While uplift optimiza-

Lifecycle Phase	Goal	Mechanism	Artifact
Data Ingestion	Stable event capture	Append-only logs	Raw tables
Feature Store	Temporal correctness	Deterministic mapping	Reusable features
Model Registry	Reproducibility	Versioned training runs	Model metadata
Decision Service	Real-time scoring	Stateless execution	Log records

Table 5. End-to-end system integration and model lifecycle management.

Governance Aspect	Focus	Implementation	Outcome
Fairness	Group parity	Monitor uplift by group	Disparity alerts
Traceability	Decision logging	Versioned policy IDs	Auditable history
Privacy	Data minimization	Tokenized identifiers	Controlled joins
Monitoring	Drift detection	Calibration diagnostics	Stable overlap

Table 6. Governance and monitoring dimensions in uplift system operation.

tion naturally directs resources where estimated effects are higher, this optimization must be examined for unintended group-level patterns. In addition, because uplift models are built upon customer 360 features that may correlate with sensitive attributes, it is necessary to understand how effect heterogeneity is being captured and whether certain signals should be excluded or constrained. The introduction of governance structures, including documentation of model assumptions, logging for decision traceability, and procedures for auditing group-level outcomes, is therefore integral to the practical framing of uplift modeling in this context.

2. Customer 360 Data Architecture and Temporal Design

Customer 360 data architectures are typically constructed as layered models, beginning with raw events captured from instrumentation in web and mobile applications, transactional systems, email and notification services, loyalty and subscription platforms, and support tools. These events are associated with identifiers that may include cookies, device IDs, hashed emails, account IDs, and in-store loyalty tokens [12]. The identity resolution layer produces a mapping from these source identifiers to a stable person or household key. Each mapping decision introduces the risk of linking distinct individuals or fragmenting a single individual across multiple keys, and these risks propagate into subsequent estimation.

A central requirement for causal analysis is that the unit of decision, the features used for decisioning, and the outcome definition are temporally consistent and do not leak post-treatment information. Let the basic exposure unit be defined as a tuple indicating that an individual was eligible for and considered for a treatment at a specific decision time [13]. For each exposure unit, features are derived from events strictly preceding that time. Outcomes are measured in a fixed or variable window following the exposure, depending on the nature of the intervention. For example, for a promotional email treatment, the outcome might be binary conversion within seven days; for an on-site recommendation banner, it might be click or add-to-cart within the same session or a defined short window.

Formally, consider a set of individuals indexed by i and exposure events indexed by j [14]. Let X_{ij} be the feature vector constructed from the history of individual i up to exposure time t_{ij}, let T_{ij} be an indicator or categorical variable encoding the treatment assigned at that exposure, and let Y_{ij} be the outcome measured after a specified lag. The temporal design requires that X_{ij} excludes any event with timestamp greater than or equal to t_{ij} and that Y_{ij} is insensitive to censoring choices that depend on post-treatment behavior unrelated to the outcome of interest.

$$X_{ij} = \Gamma\big(\mathscr{H}_i(s) : s < t_{ij}\big) \tag{1}$$

Here $\mathscr{H}_i$ denotes the event history for individual i, and Γ is a deterministic feature mapping. Outcomes are then defined by an aggregation operator over a window $(t_{ij}, t_{ij} + \Delta]$ that is held fixed during modeling for comparability across treatments.

$$Y_{ij} = \Lambda\big(\mathscr{H}_i(s) : t_{ij} < s \le t_{ij} + \Delta\big) \tag{2}$$

Multiple exposures for the same individual introduce dependence structures. A simple approach is to restrict analysis to first exposures within non-overlapping windows or to define rules that censor later exposures for modeling purposes. An alternative is to model each exposure as conditionally independent given the history, acknowledging that this is an approximation when dynamic feedback is present. In either case, the customer 360 representation must retain sufficient granularity so that the selection and censoring rules can be reconstructed and audited. [15]

Feature engineering within this architecture includes counts of visits and transactions, recency metrics, category- and brand-level affinities, price sensitivity indicators derived from response to historical promotions, and signals extracted from session paths. For causal applications, the key consideration is that these features summarize information that plausibly influenced treatment assignment and future outcomes without encoding future responses. To maintain consistency between offline training and online deployment, features are implemented in a feature store abstraction so that both environments invoke the same transformations with respect to event time and time zones.

Cross-channel effects complicate design [16]. A single individual may receive a sequence of emails, see personalized banners, and interact with recommendations in the mobile app. Treatments defined at an exposure unit should capture the relevant competing or complementary interventions in a way that approximates the environment seen by the decisioning system. For example, exposure units corresponding to email sends may include as covariates the recent onsite treatment history, while onsite recommendation exposures may include recent promotional communications. This coupling allows the estimation machinery to condition on observable factors that influence both assignment and outcomes, which is essential when some treatments are deterministic functions of past behavior. [17]

3. End-to-End System Integration and Model Lifecycle Management

Deploying causal uplift models on integrated customer 360 data requires that statistical methodology, data architecture, and software infrastructure form a coherent pipeline from logging through decisioning. The central objective is not only to obtain estimates of heterogeneous treatment effects but to ensure that these estimates are reproducible, interpretable under stable assumptions, and traceable to the operational environment in which decisions are made. This section develops an end-to-end perspective on how data flows, model artifacts, and policy configurations can be organized across their lifecycle, starting from raw behavioral streams and extending to production services that assign treatments in real time. The discussion emphasizes deterministic feature definitions, explicit representations of propensities and exploration, isolation between training and serving logic, and mechanisms for rolling updates and rollback [18]. The intent is to outline a structure in which causal uplift modeling is embedded as a controlled subsystem of the broader personalization stack rather than an ad hoc layer added on top of predictive scoring engines.

The lifecycle begins with event ingestion and normalization. Web and mobile clients, transactional databases, and messaging systems emit events that are captured in append-only logs with timestamps, identifiers, and minimal contextual attributes. Identity resolution maps these identifiers into stable customer keys under deterministic rules that can be versioned; every resolution decision is stored along with the rule configuration that produced it, enabling reconstruction of historical keys when rules evolve. Exposure definitions for treatments, such as a promotional impression or eligibility for a recommendation module, are generated by applying consistent predicates on these logs [19]. Crucially, each exposure record captures not only whether the customer was ultimately shown a treatment but also whether they were eligible and whether randomization or policy logic governed assignment. The resulting exposure table is the anchor for causal analysis, connecting pre-exposure features and post-exposure outcomes while retaining metadata about assignment mechanisms. To preserve consistency, this table is produced by a pipeline that is immutable over a given analysis window and updated via explicit version increments rather than silent modifications.

Feature computation is managed via a feature store abstraction that enforces temporal correctness and cross-environment parity [20]. For each feature, a definition specifies its input events, aggregation logic, and time indexing relative to an exposure timestamp. The same code paths, with identical parameterization, are used in offline training jobs and in online scoring services, ensuring that any uplift model relies on consistent numerical representations. When backfilling historical exposures for training, the feature store reconstructs the feature values as of the exposure time using only events with earlier timestamps. In streaming contexts, incrementally maintained aggregates approximate the same definitions [21]. Discrepancies between batch and streaming implementations are tracked by periodic reconciliations on overlapping time windows, and deviations beyond tolerance thresholds trigger investigation. This alignment is foundational because any divergence between training and serving distributions can introduce systematic biases into estimated treatment effects and invalidate offline policy evaluation that assumes stable mapping from raw logs to model inputs.

The model training stage consumes exposure-level datasets with linked features, treatments, outcomes, and, where applicable, known propensities. Training pipelines are expressed as declarative graphs that separate nuisance estimation from effect learning [22]. One subset of components fits models for $\hat{e}(x)$ and $\hat{\mu}_t(x)$ using cross-validation folds that respect temporal ordering, while another constructs orthogonal pseudo-outcomes for uplift learning. To keep expressions compact and facilitate inspection, the key transformation for each exposure can be summarized as

$$\psi = \hat{\mu}_1(X) - \hat{\mu}_0(X) + w(Y - \hat{\mu}_T(X)), \tag{3}$$

where w is a folded function of $\hat{e}(X)$ and T with truncation to control variance. The uplift learner fits $\hat{\tau}(x)$ to approximate $\mathbb{E}(\psi \mid X = x)$ using tree ensembles or other flexible models. Every training run logs configurations, random seeds, data intervals, model hyperparameters, and transformation versions so that any deployed artifact is exactly reproducible. Instead of overwriting models in place, each candidate is stored as a versioned object with its own metadata and evaluation reports. The lineage from raw logs through feature generations and nuisance models to final uplift estimators is maintained in a registry accessible for audits. [23]

Policy derivation operates on the outputs of uplift models under explicit constraints. Given an estimated effect $\hat{\tau}(x)$ and cost $c(x)$, a parametric policy class is instantiated, for example threshold rules of the form $\pi_\kappa(x) = \mathbf{1}\{\hat{\tau}(x)/c(x) \geq \kappa\}$ for an adjustable parameter κ. Offline evaluation procedures, using randomized or exploration logs, estimate the value of each candidate π_κ together with uncertainty bands. Conceptually, the optimization searches over κ to find a configuration that respects budget constraints and yields acceptable risk characteristics. This step can be framed as a constrained empirical risk minimization problem in which the empirical objective approximates expected incremental outcome while penalties encode variance, fairness metrics, or operational cost. For instance, a conservative calibration adjusts κ so that the lower bound of an estimated interval for incremental value remains non-negative [24]. Policies selected under this process are exported as explicit parameter sets bound to a particular model version and accompanied by an evaluation dossier.

The online decisioning service is implemented as a stateless component that, for each incoming request, assembles a context from the feature store, scores uplift using the bound model, and applies the associated policy configuration along with exploration and guardrail logic. To ensure traceability, every decision yields a compact log record that includes a decision identifier, hashed features, model version, policy parameters, propensity of the chosen action under the active mixture of targeting and exploration rules, and treatment outcome when available. Exploration is implemented as a controlled randomization overlay, for example by assigning a small fraction of traffic to alternative actions with pre-specified probabilities that are logged precisely [25]. This preserves overlap for future off-policy evaluation without destabilizing operational metrics. The decisioning service also enforces eligibility and frequency caps, ensuring that uplift-based recommendations do not violate channel limits or customer-level saturation rules.

Model governance requires structured procedures for promotion, rollback, and retirement. Before a new uplift model and its derived policy are rolled out beyond a small experimental cohort, they are subject to checks for data leakage, degradation in calibration, and sensitivity to weight truncation choices [26]. A shadow deployment pattern can be used where the new model scores traffic in parallel without influencing decisions, allowing comparison of recommended actions, predicted effects, and realized outcomes with the incumbent system. If observed discrepancies fall within predetermined tolerance intervals and no systematic adverse patterns emerge across key segments, the policy may be incrementally ramped. At each ramp step, logging and monitoring verify that empirical treatment rates, budget consumption, and outcome distributions align with projected values. Rollback is pre-configured as a revert to a prior stable model and policy version without altering downstream schemas, enabling rapid mitigation of unforeseen behavior. [27]

An integrated monitoring layer spans technical and causal diagnostics. On the technical side, monitors track latencies, error rates, feature availability, and consistency between online and offline feature distributions. On the causal side, monitors estimate realized incremental effects in regions where randomized or exploration data is available, examine uplift calibration across deciles of predicted effects, and quantify changes in propensity distributions indicating loss of overlap. For example, if the live policy yields nearly deterministic treatment decisions for a large portion of the population, leading to propensities that concentrate near zero or one, alerts can propose increasing exploration or relaxing thresholds in selected regions of the feature space. These operations are framed as adjustments to preserve evaluability rather than as ad hoc corrections to performance metrics. [28]

Fairness, compliance, and privacy concerns are integrated into each lifecycle stage instead of treated as afterthoughts. During feature design, attributes with direct or strong proxy relationships to protected characteristics are identified, and their roles in treatment decisions are bounded or excluded according to applicable policy. During model training, group-based summaries of estimated uplift and recommended treatment rates are computed, and large disparities are flagged for review. At deployment, per-group monitoring uses the logged propensities and outcomes to estimate realized incremental effects with confidence intervals, while acknowledging that smaller groups may yield higher uncertainty [29]. Throughout, privacy controls limit exposure of raw identifiers by relying on hashed or tokenized representations and by enforcing strict access policies around joining sensitive datasets. These measures are aligned with the requirement that causal uplift systems operate in a transparent and accountable manner, where decision logic can be inspected without revealing individual-level details beyond what is necessary.

Lifecycle management also addresses nonstationarity in customer behavior and platform operations. Regular retraining schedules are defined, but their triggering is conditioned on observed drift in feature distributions, outcome processes, or calibration [30]. Instead of retraining on a fixed cadence without regard to stability, diagnostics determine when the current uplift model deviates from recent data beyond acceptable ranges. When retraining is initiated, historical windows are selected to balance recency with sample size, and exploration data is incorporated explicitly to stabilize propensity estimates. Competing model candidates can be generated under alternative feature subsets or regularization strengths to assess robustness. Model selection criteria consider not only point estimates of policy value but also variance, fairness indicators, and sensitivity to choices in the treatment of extreme weights and outliers [31]. Accepted models become new versions in the registry; unselected models remain as documented experiments, providing a record of design space exploration.

Finally, the integration of uplift modeling into the broader personalization ecosystem benefits from clear separation of concerns between effect estimation and content generation. The uplift system concerns itself with estimating which customer-exposure pairs are more likely to exhibit incremental response to classes of interventions, while other components of the platform determine concrete creative, ranking, or messaging variants. This separation permits the reuse of effect estimators across multiple campaigns of similar type and simplifies validation, as the causal layer focuses on relative impact of being treated versus not treated within defined intervention categories [32]. Over time, this modular structure can extend to multi-action settings where treatments correspond to families of interventions, with uplift models producing vector-valued effect estimates and policies selecting among them under resource constraints. Throughout, the lifecycle framework aims to maintain a stable mapping from assumptions and algorithms to observable behavior in production, making it possible to interpret observed outcomes in light of explicit causal reasoning rather than opaque optimization dynamics.

4. Causal Identification in Multi-Channel Retail Environments

The potential outcomes framework provides a formal language to reason about the causal effects of treatments defined over the customer 360 data. For binary treatment at an exposure unit, define $Y_{ij}(1)$ and $Y_{ij}(0)$ as the potential outcomes under treatment and control. The observed outcome satisfies the consistency relation $Y_{ij} = T_{ij}Y_{ij}(1) + (1-T_{ij})Y_{ij}(0)$. The conditional average treatment effect given features $X_{ij} = x$ is

$$\tau(x) = \mathbb{E}\{Y_{ij}(1) - Y_{ij}(0) \mid X_{ij} = x\}. \tag{4}$$

Identification of $\tau(x)$ requires conditions that connect observed data to these potential outcomes [33]. When treatments are assigned through randomized experiments, the assignment mechanism ensures that T_{ij} is independent of $(Y_{ij}(1), Y_{ij}(0))$ conditional on design variables, often even unconditionally. In this setting, unbiased estimation is straightforward, and heterogeneity can be explored by conditioning on X_{ij} without invoking strong structural assumptions. However, many operational decisions in retail systems result from algorithms or rules driven by features of the customer and context, yielding observational data where treatment is not random.

A standard assumption for identification in observational settings is conditional exchangeability, also described as unconfoundedness, which states that the potential outcomes are independent of treatment given features. Formally,

$$(Y_{ij}(1), Y_{ij}(0)) \perp\!\!\!\perp T_{ij} \mid X_{ij}. \tag{5}$$

Combined with positivity, which requires that for all x in the support of X_{ij} the propensity to treat satisfies $0 < e(x) < 1$, these conditions permit identification of $\tau(x)$ based on outcome regression, inverse probability weighting, or doubly robust estimators. In multi-channel personalization, verifying these conditions is challenging, as assignment mechanisms may depend on unobserved variables, latent preferences, or internal optimization states that are not captured in X_{ij}.

Instrumental variables and natural experiments can support identification in specific contexts [34]. For example, randomized subject lines or layout variations can serve as instruments that shift treatment probability without directly affecting revenue beyond the mediated effect through engagement. Similarly, logistic or inventory constraints that exogenously suppress some promotions may generate variation that is plausibly as good as random for certain segments. An instrument Z_{ij} must satisfy relevance (affecting treatment) and exclusion (affecting outcome only through treatment); under these conditions, local average treatment effects can be recovered in the subpopulation whose treatment status responds to the instrument. These effects are local in nature and must be interpreted relative to the inducing mechanism.

Interference is a structural concern in retail environments [35]. One customer's treatment can influence another's experience through shared inventory, social influence, or platform-level constraints. Partial interference assumptions, where interference is confined within pre-defined groups such as region or campaign cohort, are more realistic than global independence. Under partial interference, potential outcomes are indexed by both own treatment and group-level treatment pattern, complicating identification and estimation but allowing for structured models when group definitions align with system architecture.

Temporal dynamics pose additional identification issues [36]. The outcome for one exposure can influence future treatment through updating of personalization models or eligibility thresholds, leading to feedback loops. In such cases, classical single-stage causal estimands may not capture long-run consequences of policies. Sequential causal inference frameworks define potential outcome trajectories under dynamic treatment regimes. While full estimation of these regimes can be complex, in many retail use cases it is still informative to define intermediate estimands, such as the effect of a treatment on outcomes within a limited horizon, under the assumption that subsequent treatments follow observed policies. [37]

Throughout, the framework treats identification claims as conditional on documented assignment rules, feature sets, and interference assumptions. The goal is to delineate regimes in which uplift estimates can be interpreted causally with reasonable robustness and regimes where estimates remain associative and should be used with caution in policy learning.

5. Heterogeneous Treatment Effect Estimation and Uplift Modeling

Once identification conditions are specified, estimation of heterogeneous treatment effects proceeds by combining flexible predictive models with constructions that separate outcome and propensity components. Define the propensity score $e(x) = \mathbb{P}(T_{ij} = 1 \mid X_{ij} = x)$ and the outcome regressions $\mu_t(x) = \mathbb{E}(Y_{ij} \mid X_{ij} = x, T_{ij} = t)$ for $t \in \{0,1\}$. Under unconfoundedness and positivity, the conditional effect $\tau(x)$ can be expressed in multiple equivalent forms. Doubly robust estimators exploit these equivalences, providing consistency if either the propensity or outcome model is correctly specified. [38]

For each observation, construct residuals and weights as

$$r_{ij} = Y_{ij} - \mu_{T_{ij}}(X_{ij}), \tag{6}$$

$$w_{ij} = \frac{T_{ij}}{e(X_{ij})} - \frac{1 - T_{ij}}{1 - e(X_{ij})}. \tag{7}$$

An orthogonalized pseudo-outcome for learning $\tau(x)$ is

$$\psi_{ij} = \mu_1(X_{ij}) - \mu_0(X_{ij}) + w_{ij} r_{ij}. \tag{8}$$

Regressing ψ_{ij} on X_{ij} using a flexible learner yields an estimate of $\tau(x)$ that is less sensitive to small errors in the nuisance functions e and μ_t. Cross-fitting, where nuisance models are trained on folds disjoint from the fold used to fit the final regressor for $\tau(x)$, further reduces overfitting-induced bias.

Common meta-learners implement alternative decompositions [39]. In the T-learner, separate models approximate μ_1 and μ_0, and uplift is obtained by subtraction; in the S-learner, a single model approximates $\mathbb{E}(Y_{ij} \mid X_{ij}, T_{ij})$, and uplift is the difference in predictions at $T_{ij} = 1$ and $T_{ij} = 0$. X- and R-learners construct individual-level effect estimates by re-weighting residualized outcomes with propensity scores. For large retail datasets, tree-based ensembles and gradient boosting models are often used as base learners due to their ability to capture nonlinearities and interactions without strong parametric assumptions.

An uplift tree model partitions the feature space into regions where the difference between treated and control outcomes is relatively homogeneous. For a candidate split, a splitting criterion evaluates the gain in estimated uplift heterogeneity [40]. Honest variants reserve part of the data for evaluating treatment effects in leaves to mitigate adaptive bias. While such models can be interpretable at small depth, deeper trees trade interpretability for fit. Hybrid approaches combine global models with localized diagnostic trees to describe uplift patterns.

Regularization practices recognize that in many feature dimensions, treatment effect heterogeneity may be limited [41]. Penalized regressions or sparsity-inducing priors on effect modifiers enforce parsimony. A simple linear heterogeneous effect model decomposes the outcome into a baseline and an interaction with treatment as

$$Y_{ij} = \alpha^\top X_{ij} + T_{ij} \beta^\top X_{ij} + \varepsilon_{ij}, \tag{9}$$

which implies $\tau(x) = \beta^\top x$. Suitable penalties on β encourage the model to identify a limited set of features that drive heterogeneity.

Representation learning can be used to compress high-dimensional behavioral traces into embeddings that serve as inputs for uplift models [42]. When such embeddings are learned without using post-treatment outcomes from the target intervention, they can capture stable patterns while preserving the validity of subsequent causal estimation. Nevertheless, embeddings can introduce opacity, and their training procedures must be aligned with identification assumptions to avoid inadvertently conditioning on post-treatment variables.

Model selection for uplift focuses on predictive performance with respect to treatment effect ranking and calibration rather than only on outcome prediction. Cross-validation folds aligned with time help assess temporal stability [43]. The resulting uplift estimates form the basis for policy learning and must be accompanied by diagnostics on variance, overlap, and sensitivity to nuisance specifications.

6. Policy Learning, Budget-Constrained Targeting, and Sequential Decisions

The primary operational use of uplift estimates is to guide treatment assignment policies that respect resource and constraint structures. Let a policy be a function

Estimator	Key Components	Robustness Property	Example Formula
Outcome Regression	$\mu_t(x)$ models	Consistent if μ_t correct	$\hat{\tau}(x) = \hat{\mu}_1(x) - \hat{\mu}_0(x)$
Inverse Propensity	$e(x)$ scores	Consistent if e correct	$w_{ij}Y_{ij}$ weighting
Doubly Robust	$\mu_t(x), e(x)$ both	Consistent if either correct	$\psi_{ij} = \mu_1 - \mu_0 + w_{ij}r_{ij}$
Cross-Fitted	Fold-split estimation	Reduces overfitting bias	Orthogonal learners

Table 7. Heterogeneous treatment effect estimation strategies.

Meta-Learner	Structure	Estimation Idea	Notes
T-Learner	Two separate μ_t models	Predict, then subtract	Simple, may ignore imbalance
S-Learner	Single model with T as feature	Difference in predictions	Compact, may underfit effect
X-Learner	Residualized outcomes	Weighted re-estimation	Low variance in large data
R-Learner	Orthogonal loss	Regularized effect fitting	Efficient and general

Table 8. Common meta-learners for uplift modeling.

$\pi(x) \in \{0,1\}$ indicating whether to treat an exposure with features x. Under a single-period static view and known conditional effect $\tau(x)$, an unconstrained value-maximizing policy would treat if and only if $\tau(x) > 0$. In practice, treatments carry heterogeneous costs and are subject to budgets, frequency caps, or inventory limits [44]. Let $c(x)$ denote the cost of treating a unit with features x, and let the decision-maker be constrained by an expected total cost not exceeding a budget B.

The optimization problem can be written as

$$\max_{\pi} \ \mathbb{E}\{\pi(X)\tau(X)\} \tag{10}$$

subject to

$$\mathbb{E}\{\pi(X)c(X)\} \leq B. \tag{11}$$

Introducing a Lagrange multiplier $\lambda \geq 0$ yields an unconstrained objective [45]

$$L(\pi) = \mathbb{E}\{\pi(X)(\tau(X) - \lambda c(X))\}, \tag{12}$$

which is maximized by treating whenever $\tau(x) - \lambda c(x) > 0$. The multiplier λ is adjusted so that the induced treatment rate satisfies the budget at equilibrium. In empirical implementations, this results in ranking exposures by estimated uplift per unit cost and selecting the top fraction consistent with available resources.

When operational constraints include maximum exposure rates per user or per segment, policies must incorporate these caps explicitly [46]. For example, a recency-based cap might prohibit further promotional emails to a customer above a given recent count. Policies can then be defined in terms of both X_{ij} and state variables summarizing prior treatments. Estimation remains based on exposure-level data, but policy implementation enforces caps in real time, which modifies future logged propensities and must be captured in evaluation frameworks.

Sequential decision processes arise when the retailer repeatedly interacts with customers and the impact of a treatment depends on past treatment history. In such settings, static uplift policies may overlook long-term trade-offs, such as habituation to discounts or evolving category interests. A dynamic framework specifies a state vector S_t summarizing the history at time t, an action A_t representing treatment choice, and a reward R_t representing immediate outcome. A stochastic policy $\pi(A_t \mid S_t)$ maps states to action distributions. The objective is to maximize the expected discounted sum of rewards. [47]

$$V^{\pi} = \mathbb{E}^{\pi}\Big[\sum_{t=0}^{\infty} \gamma^t R_t\Big], \tag{13}$$

with discount factor $\gamma \in (0,1)$. Estimation of optimal dynamic policies from logged bandit or reinforcement learning data is complex and requires careful control of extrapolation. In practice, many retail implementations adopt restricted classes of policies that adjust uplift thresholds based on coarse state summaries, to preserve interpretability and facilitate evaluation.

Policy learning in the uplift context can be approached through direct optimization of estimators for V^{π}. Offline learning algorithms construct candidate policies and evaluate them using inverse propensity weighting or doubly robust techniques on historical data [48]. To maintain compatibility with underlying identification assumptions, policies are constrained to depend only on observed pre-treatment features and possibly known design variables. The search space for policies is often parameterized by thresholds or scores to reduce variance.

7. Evaluation, Monitoring, and Governance in Production Systems

Reliable deployment of causal uplift policies depends on evaluation protocols that quantify expected performance, detect deviations, and provide transparency. Offline evaluation uses logged data to approximate the value of candidate policies before deployment [49]. In randomized experiment settings, where treatment probabilities are known, inverse probability weighting can reweight observed outcomes to emulate counterfactual

Approach	Mechanism	Advantage	Limitation
Uplift Trees	Partition by treatment difference	Interpretability	Bias in adaptive splits
Honest Trees	Separate fit/eval sets	Reduced overfitting	Data inefficiency
Regularized Linear	Sparse β for heterogeneity	Parsimony	Linear assumption
Representation Learning	Embedding features	Capture nonlinearities	Potential opacity

Table 9. Modeling techniques for heterogeneous treatment effects.

Policy Type	Objective	Decision Rule	Constraint
Unconstrained	Maximize uplift	Treat if $\tau(x) > 0$	None
Budget-Constrained	Maximize uplift per cost	$\tau(x)/c(x) \geq \kappa$	$\mathbb{E}\{\pi c\} \leq B$
Frequency-Capped	Respect exposure limits	Enforce recency bounds	User-level caps
Sequential	Long-term optimization	$\pi(A_t\|S_t)$ stochastic	Markov constraints

Table 10. Policy learning and operational targeting frameworks.

policies that differ from the logging policy but rely on the same available features.

Let p_{ij} denote the propensity with which treatment T_{ij} was assigned in the logging environment, and let a candidate policy be $\pi(X_{ij})$. A basic importance-weighted estimator of the policy value is

$$\hat{V} = \frac{1}{N}\sum_{i,j}\frac{\mathbf{1}\{T_{ij} = \pi(X_{ij})\}Y_{ij}}{p_{ij}}, \tag{14}$$

where the sum is over exposure units. To reduce variance and incorporate outcome modeling, doubly robust estimators augment this with predictions under each action [50]. Let $\hat{\mu}_t(x)$ be estimated expected outcomes; a doubly robust estimator is

$$\hat{V}_{DR} = \frac{1}{N}\sum_{i,j} h_{ij}, \tag{15}$$

where

$$h_{ij} = \hat{\mu}_{\pi(X_{ij})}(X_{ij}) + \frac{\mathbf{1}\{T_{ij} = \pi(X_{ij})\}}{p_{ij}}\{Y_{ij} - \hat{\mu}_{T_{ij}}(X_{ij})\}. \tag{16}$$

This construction yields consistent estimates when either the propensity or outcome model is correct. Confidence intervals can be obtained using influence function approximations or bootstrap over exposure units with time-aware resampling.

Ranking-based metrics such as uplift curves and Qini indices assess how well uplift models order individuals by expected incremental effect. To construct an uplift curve, exposures are sorted by predicted $\hat{\tau}(X_{ij})$, and for each prefix proportion, the difference in observed outcomes between treated and control units is computed. The area between this curve and a baseline representing random targeting summarizes discriminatory power [51]. While such metrics are informative, they should be interpreted alongside budget and constraint-aware evaluations that reflect actual operational usage.

Monitoring in production addresses distributional shift, overlap degradation, and calibration drift. Distributional shift is assessed by comparing the distribution of features, propensities, and outcomes between the training period and recent serving periods using discrepancy measures. Overlap degradation occurs when the deployed policy becomes more deterministic and concentrates treatment on a narrow region of the feature space, reducing the support available for learning about alternative policies [52]. To counter this, exploration is maintained by randomizing treatment for a small subset of traffic, logging propensities, and integrating exploration data into future uplift estimation.

Calibration diagnostics compare predicted uplift to realized differences in held-out or recent experimental cohorts. Features are binned according to predicted uplift, and within each bin, the empirical treatment-control difference is compared to the average prediction. Systematic deviations suggest miscalibration or concept drift [53]. When drift is detected, retraining and policy reevaluation are triggered under predefined governance rules.

Governance structures define how models are documented, approved, and audited. Documentation includes a description of the data sources, inclusion and exclusion criteria, exposure definitions, outcome windows, identification assumptions, estimation methods, hyperparameters, and limitations. Logging infrastructure records each decision with hashed identifiers for features, the predicted uplift, the chosen action, and the propensity of that action under the active policy and any exploration scheme [54]. These logs support post-hoc analysis of unexpected outcomes, reconstruction of exposure histories, and tracing of aggregate patterns to specific model updates.

Fairness and compliance considerations form part of governance. Group-level analyses examine whether treatments, benefits, and burdens are distributed unevenly across protected or sensitive segments. While uplift optimization inherently targets individuals with higher estimated incremental response, monitoring is used to determine whether this optimization inadvertently produces systematic exclusion or concentration patterns that war-

Evaluation Method	Estimator	Requirement	Purpose
Inverse Propensity	$\hat{V}$	Known p_{ij}	Offline counterfactual eval
Doubly Robust	$\hat{V}_{DR}$	Either p or μ_t correct	Lower variance, consistency
Uplift Curve	Qini index	Treatment-control split	Ranking performance
Bootstrap CI	Resampling	Time-aware folds	Uncertainty quantification

Table 11. Evaluation and validation metrics for uplift models.

Governance Aspect	Focus	Implementation	Outcome
Logging	Decision traceability	Hashed IDs, propensities	Auditability
Monitoring	Drift detection	Feature and outcome shift	Stability alerts
Fairness	Group-level analysis	Parity in uplift and exposure	Compliance tracking
Documentation	Transparency	Model/data registry	Reproducible records

Table 12. Governance and monitoring practices for causal uplift systems.

rant reassessment [55]. Corrective actions can include imposing constraints on policy parameters or revising feature sets.

8. Extensions: Dynamic Effects, Interference, and Robustness

The static uplift framework can be extended to address dynamic treatment effects, interactions among customers, and robustness to misspecification. Dynamic effects arise when the impact of an intervention persists or evolves over time, influencing subsequent purchases, engagement, and price sensitivity. To formalize this, define potential outcome trajectories $\{Y_{i,t}(\bar{a}_t)\}$, where $\bar{a}_t$ denotes a sequence of actions up to time t. A dynamic regime specifies actions as functions of evolving histories, and the evaluation objective becomes an expectation over trajectories under such regimes.

In practical terms, retailers often consider limited-horizon effects where interventions are not expected to substantially alter behavior beyond a finite window [56]. Nevertheless, discount habituation or overexposure to promotions can shift baselines. Sequential causal estimators such as marginal structural models reweight observed sequences to emulate alternative regimes under assumptions about treatment assignment and confounding. The stabilized weights depend on probabilities of observed actions conditional on past histories; truncated weights are used to manage variance [57].

Interference is inherent in settings with shared capacity, social contagion, or recommendation feedback [58]. One abstraction partitions customers into clusters where interference is assumed to operate locally. Potential outcomes $Y_{ij}(t,z)$ may depend on own treatment t and a summary z of treatments within the cluster. Identification then requires assumptions on how cluster-level treatment is assigned and how it affects outcomes. In promotion scenarios with inventory constraints, the availability of an item may be a function of total treated demand; policies must consider that encouraging additional demand could induce stock-outs that affect untreated customers.

Robustness analysis addresses uncertainty in propensity estimates, outcome models, and structural assumptions [59]. Sensitivity to unobserved confounding can be expressed by bounding how much the odds of treatment might differ between units with equal observed features but different potential outcomes. Given such a bound, one can derive ranges for treatment effect estimates consistent with the observed data and the assumed degree of hidden bias. These ranges can inform whether conclusions used for targeting remain stable under plausible deviations.

Consider a simple bound on propensity distortion [60]. Let $\hat{e}(x)$ be the estimated propensity and assume the true propensity satisfies

$$\log \frac{e(x)}{1-e(x)} = \log \frac{\hat{e}(x)}{1-\hat{e}(x)} + \delta(x), \tag{17}$$

where $\delta(x)$ is unknown but constrained in magnitude. By varying $\delta(x)$ within a specified range, one obtains perturbed propensities and recomputed effect estimates, forming sensitivity intervals. This type of analysis can be implemented segment-wise to detect regions where conclusions are fragile.

Robust policy learning incorporates such uncertainties directly into optimization [61]. Instead of maximizing a point estimate of expected uplift, the decision-maker can optimize a conservative criterion, such as the worst-case value over a set of plausible models. Let $\mathscr{M}$ be a neighborhood of models for $\tau(x)$ and $e(x)$. A robust objective is

$$\max_{\pi} \min_{m \in \mathscr{M}} \mathbb{E}_m\{\pi(X)\tau_m(X)\}, \tag{18}$$

where τ_m denotes the treatment effect function under model m. While direct solution can be complex, approximate methods based on penalizing variance and incorporating safety margins in thresholds are tractable.

Another dimension of robustness concerns numerical and implementation stability [62]. Differences between training and serving feature computations, silent failures in identity resolution, and lagged ingestion of events can all induce discrepancies. Monitoring for such discrepancies includes cross-checking distributions of key features across pipelines, verifying that exposure definitions remain consistent over time, and validating that propensities implied by the live policy match expectations.

Finally, interpretability contributes to robustness by enabling practitioners to identify implausible patterns. Even when primary models are complex, simplified summaries of uplift as functions of a small number of features can reveal whether the learned structure aligns with domain understanding [63]. If treatment effects appear to depend strongly on proxies for attributes that are operationally or ethically sensitive, additional scrutiny is warranted, and feature sets or constraints can be adjusted accordingly.

9. Conclusion

This paper has presented a detailed and neutral account of how causal inference and uplift modeling can be integrated with customer 360 data to inform targeted personalization strategies in B2C digital retail platforms. The proposed perspective starts from explicit definitions of exposure units, temporal alignment of features and outcomes, and identity resolution constraints, recognizing that these infrastructural elements shape the validity of any downstream causal claim. Within this structured data environment, potential outcomes notation and associated identification assumptions provide a basis for interpreting uplift estimates as approximations to heterogeneous treatment effects rather than purely associative patterns. [64]

Estimation strategies built on outcome regression, propensity modeling, doubly robust scores, and orthogonalization enable flexible learners to be used while maintaining some protection against misspecification. These strategies produce individualized effect estimates that can be mapped into policies respecting budget, capacity, and frequency constraints. Simple ranking rules based on uplift per unit cost emerge naturally from constrained optimization formulations, with dynamic extensions allowing limited forms of sequential adaptation when warranted by data and operational requirements.

Evaluation and governance mechanisms are central to the framework [65]. Off-policy estimators grounded in logged propensities, uplift ranking diagnostics, calibration checks, drift detection, and continuous experimentation contribute to a controlled environment in which targeting rules can be updated with traceable justifications. Fairness, privacy, and interpretability considerations are incorporated not as external additions but as constraints and monitoring dimensions that interact with model design and deployment.

The discussion acknowledges the limitations imposed by unobserved confounding, interference, dynamic feedback, and finite logging support. Rather than asserting universal solutions, it outlines how sensitivity analysis and robust optimization concepts can be applied to characterize the range of plausible effects and to construct policies that are conservative with respect to modeling uncertainty. The resulting view is one where causal uplift modeling is treated as an integrated systems problem, connecting data architecture, identification, estimation, policy learning, and governance, and where conclusions are conditioned on transparent and testable assumptions aligned with the realities of large-scale digital retail platforms. [66]

References

[1] M.-H. Huang and R. T. Rust, “Technology-driven service strategy,” Journal of the Academy of Marketing Science, vol. 45, pp. 906–924, 6 2017.

[2] I. V. Pestun and Z. M. Mnushko, “Methodology of modern marketing management pharmaceutical organizations,” Upravlìnnâ, ekonomìka ta zabezpečennâ âkostì v farmacìï, vol. 0, pp. 60–66, 11 2017.

[3] M. A. Rahman, A. Islam, B. H. Esha, N. Sultana, and S. Chakravorty, “Consumer buying behavior towards online shopping: An empirical study on dhaka city, bangladesh,” Cogent Business & Management, vol. 5, pp. 1514940–, 1 2018.

[4] J. G. Sidak and H. J. Singer, “Evaluating market power with two-sided demand and preemptive offers to dissipate monopoly rent: Lessons for high-technology industries from the antitrust division’s approval of the xm–sirius satellite radio merger,” Journal of Competition Law and Economics, vol. 4, pp. 697–751, 9 2008.

[5] S. Gu, B. Ślusarczyk, S. Hajizada, I. Kovalyova, and A. Sakhbieva, “Impact of the covid-19 pandemic on online consumer purchasing behavior,” Journal of Theoretical and Applied Electronic Commerce Research, vol. 16, pp. 2263–2281, 9 2021.

[6] Y. Şahin and D. M. Sara, “E-shopping sites preference analysis with multi-criteria decision-making methods,” Bilişim Teknolojileri Dergisi, vol. 12, pp. 265–275, 10 2019.

[7] M. Cotarelo, T. Fayos, H. Calderón, and A. Mollá, “Omni-channel intensity and shopping value as key drivers of customer satisfaction and loyalty,” Sustainability, vol. 13, pp. 5961–, 5 2021.

[8] P. A. Broome, “Conceptualizing the foundations of a regional e-commerce strategy: Open networks or closed regimes? the case of caricom,” Cogent Business & Management, vol. 3, pp. 1139441–1139441, 2 2016.

[9] H. S. M. Lipiäinen, "Crm in the digital age: implementation of crm in three contemporary b2b firms," Journal of Systems and Information Technology, vol. 17, pp. 2–19, 3 2015.

[10] I. Zennaro, S. Finco, M. Calzavara, and A. Persona, "Implementing e-commerce from logistic perspective: Literature review and methodological framework," Sustainability, vol. 14, pp. 911–911, 1 2022.

[11] S. H. Kukkuhalli, "Increasing digital sales revenue through 1:1 hyper-personalization with the use of machine learning for b2c enterprises," Journal of Artificial Intelligence, Machine Learning and Data Science, vol. 1, no. 1, 2023.

[12] J. Zhu, M. A. S. Goraya, and Y. Cai, "Retailer–consumer sustainable business environment: How consumers' perceived benefits are translated by the addition of new retail channels," Sustainability, vol. 10, pp. 2959–, 8 2018.

[13] B. Alkhalil and M. A. Siddiqui, "Factors affecting saudi consumer trust in e-commerce: A quantitative analysis," International Journal of Computer Applications, vol. 182, pp. 41–45, 11 2018.

[14] R. Mącik, "Cross-border online shopping of polish consumers and cross- border online sales of polish companies as examples of the processes of the markets internationalisation," Annales Universitatis Mariae Curie-Skłodowska, sectio H, Oeconomia, vol. 51, pp. 185–194, 8 2017.

[15] A. Moued and S. Abdullah, "The development of reference model for social commerce e-marketplaces based on value co-creation framework," International Journal of Computer Applications, vol. 181, pp. 43–50, 12 2018.

[16] C. Yeo, C. Hur, and S. Ji, "The customer orientation of salesperson for performance in korean market case: A relationship between customer orientation and adaptive selling," Sustainability, vol. 11, pp. 6115–, 11 2019.

[17] B. Chakraborty and S. Das, "Introducing a new supply chain management concept by hybridizing topsis, iot and cloud computing," Journal of The Institution of Engineers (India): Series C, vol. 102, pp. 109–119, 9 2020.

[18] J. Strauss and A. I. El-Ansary, "Integrating the e in e-marketing," Journal of Business & Economics Research (JBER), vol. 2, 8 2004.

[19] B. Burroughs and W. J. Burroughs, "Digital logistics: Enchantment in distribution channels.," Technology in society, vol. 62, pp. 101277–101277, 5 2020.

[20] R. Dolan, J. Conduit, J. Fahy, and S. Goodman, "Social media engagement behaviour: a uses and gratifications perspective," Journal of Strategic Marketing, vol. 24, pp. 261–277, 12 2015.

[21] S. Suhartini, N. A. Mahbubah, and M. Basjir, "Development of sme's business cooperation information technology system design," Eastern-European Journal of Enterprise Technologies, vol. 6, pp. 78–86, 12 2022.

[22] U. Golob, M. A. P. Davies, J. Kernstock, and S. M. Powell, "Trending topics plus future challenges and opportunities in brand management," Journal of Brand Management, vol. 27, pp. 123–129, 12 2019.

[23] M. van der Westhuizen, "Recasting hypertext markup language in extended text markup language: is it the foundation for the future of e-commerce?," SA Journal of Information Management, vol. 4, 12 2002.

[24] A. Agerdal-Hjermind, "The enterprise social media relations strategy: The case of maersk line," Communication & Language at Work, vol. 3, pp. 3–17, 12 2014.

[25] K. Siau, E.-P. Lim, and Z. Shen, "Mobile commerce: Promises, challenges and research agenda," Journal of Database Management, vol. 12, pp. 4–13, 7 2001.

[26] J. Li, Y. Pan, Y. Yang, and C. H. Tse, "Digital platform attention and international sales: An attention-based view," Journal of International Business Studies, vol. 53, pp. 1817–1835, 6 2022.

[27] H. K. H. A.-S. Udayanan, "Factors influencing online shopping intention: A study among online shoppers in oman," International Journal of Academic Research in Business and Social Sciences, vol. 9, pp. 691–709, 4 2019.

[28] F. Ibrahim, I. Ramzi, and F. A. Razak, "E business in practice: Evidence from brunei darussalam," Journal of Entrepreneurship and Business Innovation, vol. 5, pp. 46–, 3 2019.

[29] "Book notes 2009/no. 1," Journal of Consumer Policy, vol. 32, pp. 75–90, 1 2009.

[30] K. Storbacka and T. Moser, "The changing role of marketing: transformed propositions, processes and partnerships," AMS Review, vol. 10, pp. 299–310, 10 2020.

[31] P. Das, "A prospective venue for marketing is social media marketing," RESEARCH REVIEW International Journal of Multidisciplinary, vol. 7, pp. 7–23, 10 2022.

[32] "Empowerment and poverty reduction: a sourcebook," Choice Reviews Online, vol. 40, pp. 40–2275–40–2275, 12 2002.

[33] S. Kretschmer and S. Dehm, "Sustainability transitions in university food service—a living lab ap-

proach of locavore meal planning and procurement," Sustainability, vol. 13, pp. 7305–7305, 6 2021.

[34] P. Wetzel and E. Hofmann, "Toward a multi-sided model of service quality for logistics service providers," Administrative Sciences, vol. 10, pp. 79–, 10 2020.

[35] K. A. Berger and M. T. Topol, "Report card on e-mail: An investigation of replies to consumers," Journal of Business & Economics Research (JBER), vol. 1, 2 2011.

[36] E. R. D. Oliveira, G. O. G. Cremonezi, J. L. Casagrande, and A. J. D. Souto, "Marketing strategies: innovation in retail during the restriction and social distance measures caused by the covid-19 pandemic / estratégias de marketing: inovação no varejo durante as medidas de restrição e distância social causadas pela pandemia de covid-19," Brazilian Journal of Development, vol. 7, pp. 67155–67173, 7 2021.

[37] K. Szczukiewicz and M. Makowiec, "Characteristics and specificities of local innovation accelerators: A case of poland," Sustainability, vol. 13, pp. 1689–, 2 2021.

[38] K. Gohain, S. Thambiah, and M. C. Changmai, "Behavioural components of online shopping among consumers of malaysia," The Journal of Social Sciences Research, vol. 5, pp. 436–443, 1 2019.

[39] V. Susanti, U. Sumarwan, M. Simanjuntak, and E. Z. Yusuf, "Rational antecedent framework of brand satisfaction in the industrial market: Assessing rational perceived quality and rational perceived value roles," International Review of Management and Marketing, vol. 10, pp. 19–26, 1 2020.

[40] A. D. Mitropoulou and G. T. Tsoulfas, "Using a modified servqual approach to assess the quality of supply chain services in greek online supermarkets," Logistics, vol. 5, pp. 69–, 10 2021.

[41] A. Ashley, M. V. Laric, P. M. Lynagh, and C. A. Vollmer, "A conceptual paradigm for internet and search engine marketing," Journal of Applied Business Research (JABR), vol. 24, 1 2011.

[42] R. Teixeira, R. Przyczynski, and J. Shockley, "B2b buyer operational capability for the effective delivery of technology-enabled services in a multi-facility context," Service Business, vol. 16, pp. 645–679, 7 2022.

[43] P. Bai, Q. Wu, Q. Li, L. Zhang, Y. Jiang, and B. Chen, "The value cocreation influence mechanism of network freight transport platform in iot-based environments: Under the service-dominant logic," Wireless Communications and Mobile Computing, vol. 2021, pp. 1–13, 5 2021.

[44] C. Loro and R. Mangiaracina, "The impact of e-marketplace on the b2b relationships," Industrial Management & Data Systems, vol. 122, pp. 37–54, 9 2021.

[45] N. Awadallah and S. G. Saad, "Studying the impact of egyptian hotels websites marketing on customers e-satisfaction," International Journal of Information Technology Convergence and Services, vol. 7, pp. 1–11, 6 2017.

[46] T. Reardon, A. Heiman, L. Lu, C. S. R. Nuthalapati, R. Vos, and D. Zilberman, ""pivoting" by food industry firms to cope with covid-19 in developing regions: E-commerce and "copivoting" delivery intermediaries," Agricultural economics (Amsterdam, Netherlands), vol. 52, pp. 459–475, 4 2021.

[47] I. S. Horn, T. Taros, S. Dirkes, L. Hüer, M. Rose, R. Tietmeyer, and E. Constantinides, "Business reputation and social media: A primer on threats and responses," Journal of Direct, Data and Digital Marketing Practice, vol. 16, pp. 193–208, 4 2015.

[48] A. Payne, P. Frow, and A. Eggert, "The customer value proposition: evolution, development, and application in marketing," Journal of the Academy of Marketing Science, vol. 45, pp. 467–489, 3 2017.

[49] S. K. Msosa, "Integrating technology and digital connectivity in service recovery strategies to achieve recovery satisfaction during the covid-19 pandemic: A hybrid approach," Baltic Journal of Economic Studies, vol. 7, pp. 1–10, 12 2021.

[50] G. Guihang, W. Yanqin, and G. Chuyao, "Research on logistics cost control of e-commerce enterprise from the perspective of value chain– a case study of pinduoduo," International Journal of Economics and Finance, vol. 13, pp. 42–, 6 2021.

[51] M. I. Ahmad and R. A. Abdelsalam, "E-commerce adoption in egypt insights from covid-19 and beyond," , vol. 12, pp. 37–55, 4 2021.

[52] S. H. Kukkuhalli, "Enabling customer 360 view and customer touchpoint tracking across digital and non-digital channels," Journal of Marketing & Supply Chain Management, vol. 1, no. 3, 2022.

[53] A. Aytekin and D. Tunalı, "An e-loyalty model proposal for online travel reservation websites," Journal of Tourism Theory and Research, vol. 4, pp. 98–110, 7 2018.

[54] W. O. Widyarto, M. J. Shofa, and N. D. Jamal, "Key performance indicators on supply chain performance measurement in an electronic commerce: A literature review," International Journal of Engineering and Advanced Technology, vol. 8, pp. 137–141, 9 2019.

[55] R. Subačienė and D. Tamulevičienė, “Accounting, audit, analysis: Transformation of science and business in new economic reality book of extended abstracts,” Vilnius University Proceedings, vol. 18, pp. 1–133, 11 2021.

[56] A. Fenwick and G. Molnar, “The importance of humanizing ai: using a behavioral lens to bridge the gaps between humans and machines,” Discover Artificial Intelligence, vol. 2, 8 2022.

[57] S. H. Kukkuhalli, “Improving digital sales through reducing friction points in the customer digital journey using data engineering and machine learning,” International Journal of Innovative Research in Engineering & Multidisciplinary Physical Sciences, vol. 10, no. 3, 2022.

[58] A. de la Calle, I. Freije, and A. Oyarbide, “Digital product–service innovation and sustainability: A multiple-case study in the capital goods industry,” Sustainability, vol. 13, pp. 6342–, 6 2021.

[59] A. Doulkaid, “The impact of covid-19 on the use of online shopping:technology acceptance model,” Zenodo (CERN European Organization for Nuclear Research), 5 2022.

[60] V. Sidenko, “New institutional formats in the development of global e-commerce,” Ekonomìka ì prognozuvannâ, vol. 2020, pp. 85–110, 4 2020.

[61] R. Öztürk and A. O. Dündar, “The effect of online shopping services on customer satisfaction:the mediation role of return management,” European Journal of Science and Technology, pp. 539–548, 8 2020.

[62] F. de Oliveira Santini, W. J. Ladeira, D. C. Pinto, M. M. Herter, C. H. Sampaio, and B. J. Babin, “Customer engagement in social media: a framework and meta-analysis,” Journal of the Academy of Marketing Science, vol. 48, pp. 1211–1228, 5 2020.

[63] J. Lee, “Corporate brand management of sk,” Asia Marketing Journal, vol. 20, 4 2018.

[64] S. Pai and S. Mayya, “A study on consumer preferences with reference to online food delivery amenities,” Zenodo (CERN European Organization for Nuclear Research), 8 2022.

[65] K. Balgomera, A. E. D. Cruz, J. E. G. Santiago, and R. R. Fernandez, “Consumer trust in mobile phone industry: Comparative study on traditional commerce & e-commerce,” Journal of Business and Management Studies, vol. 4, pp. 100–116, 3 2022.

[66] R. AlGhamdi, S. Drew, and T. Alhussain, “A conceptual framework for the promotion of trusted online retailing environment in saudi arabia,” International Journal of Business and Management, vol. 7, pp. 140–149, 2 2012.