Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Operational Identification of Purchase Intention in Retail Media Networks with Theory-Coherent Hidden Semi-Markov Dynamic

Tsung-Yin Ou

National Kaohsiung University of Science and Technology

Yenming J. Chen

yjjchen@nkust.du.tw

National Kaohsiung University of Science and Technology

Research Article

Keywords:

Posted Date: November 13th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10028465/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Operational Identification of Purchase Intention in Retail Media Networks with Theory-Coherent Hidden Semi-Markov Dynamics

Tsung-Yin Ou
Department of Marketing and Distribution Management, National Kaohsiung University of Science and Technology, Taiwan, R.O.C.,
outy@nkust.edu.tw

Yenming J. Chen
Department of Information Management, National Kaohsiung University of Science and Technology, Taiwan, R.O.C.,
yjjchen@nkust.edu.tw

Retail media networks need ad decisions that react to temporally shifting shopping behavior rather than only to static transaction summaries. This paper tries to infer purchase intention from ever-changing clickstreams as an operational identification problem for a deployed advertising system. The proposed method fits time-dependent data to a covariate-dependent hidden semi-Markov model whose latent states are anchored by the Theory of Planned Behavior, the Consumer Decision Journey, and the Persuasion Knowledge Model. A retrieval-augmented language model supplies a theory-coherent reference posterior γ_t^{T} to stabilize the categorization. Semi-Markov inference uses this reference as a Kullback–Leibler regularizer in the expectation step while mapping retrieved context into the initial, transition, emission, and duration kernels. Closed-loop directional validation and revalidation are conducted through a 3-campaign experiment on a live platform. For illustrative purposes, the experiment, conducted with an anonymized omni-channel retailer covering 5.2 million active members, confirms an AUC of 0.717 for our theory-identified model, compared to 0.635 for an action-only covariate baseline on n=1,008 paired test consumers. The results show that segment-level conversion lift, opt-out reduction, and 60-day retention lift can be retained in future deployments under a drifting environment.

Key words: retail media networks, operational identification, hidden semi-Markov models, retrieval-augmented language models, personalized advertising

1. Introduction

Retail media networks (RMNs) have become a central advertising infrastructure for retailers because they combine first-party transaction data, owned digital channels, and closed-loop purchase telemetry. Global retailers such as Carrefour and Walmart have invested heavily in this channel, and industry forecasts report continued growth in retail media spending (Statista, 2024; Agius, 2024; Cramer-Flood, 2024). The commercial promise is clear. The targeting problem is less clear. A retailer observes clicks, searches, views, basket additions, and checkouts rather than survey-measured attitudes, perceptions, or psychological intentions.

This distinction defines the paper's object of inference. The paper models clickstreams as evidence for an operational latent-state representation that improves ad-response prediction and is tested through deployment feedback. The relevant object is a model state whose semantic label is declared

in advance, whose dynamics are fitted from observed events, and whose value is tested by subsequent ad response.

Conventional RMN targeting often summarizes behavior through aggregate features such as click-through rates, conversion rates, recency, frequency, monetary value, dwell time, and basket activity. These summaries are useful, but they compress away temporal structure. Two shoppers can share the same recent event counts while differing in transition pressure, state duration, and response to an informational or emotional creative. A policy based only on action counts cannot distinguish these cases. Fig. 1 contrasts this action-only representation with the theory-coherent semi-Markov representation studied here.

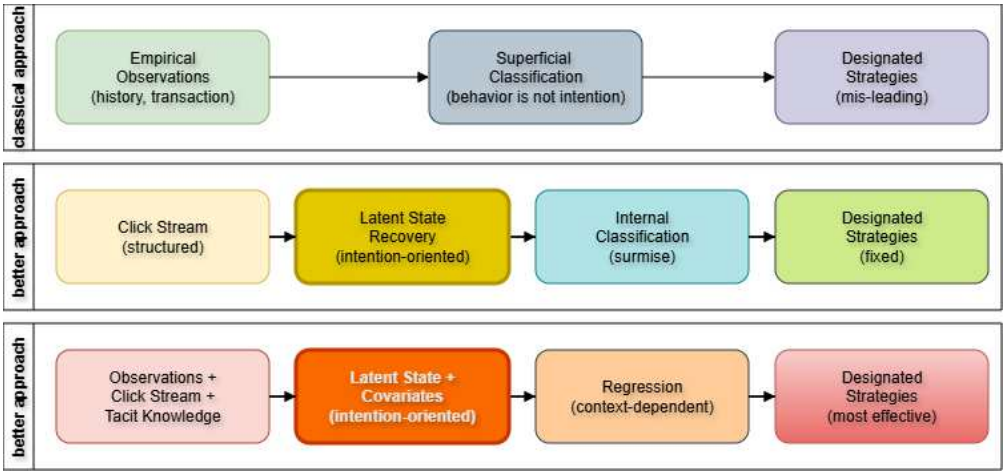


Figure 1 Action-only targeting compresses clickstreams into observed event summaries. The proposed system instead infers posteriors over a theory-coherent latent-state space, where a covariate-dependent hidden semi-Markov model is regularized by a retrieval-augmented reference distribution and then evaluated through closed-loop deployment.

The empirical motivation is also operational. Industry reports attribute large budget losses to mistargeted or generic advertising (Bishop, 2025; Accenture Interactive, 2017; Bureau, 2020). They motivate an engineering question. Can a retailer construct a state-aware predictor that uses the event sequence, product involvement, ad type, shopper trait, and retrieved contextual knowledge to recommend a next ad whose predicted response direction is later verified in the same RMN environment.

The paper answers this question through a single contribution, conceptualized as $\mathcal{S}$. Purchase-intention inference for RMN advertising can be formulated as operational identification of theory-coherent semi-Markov dynamics. The conceptual content of $\mathcal{S}$ is the identifiability reframing. The algorithmic one is a covariate-dependent HSMM whose E-step is regularized toward a retrieval-augmented reference posterior. The systematic one is the deployment audit that tests the directional implications of the inferred states. Without the theory-coherent reference distribution, the latent states have no declared semantic identity. Without the semi-Markov sequence and duration structure, the reference distribution is not calibrated to click trajectories. Without deployment feedback, the semantic regularizer remains unaudited.

The state semantics are anchored by three theory coordinates. Theory of Planned Behavior supplies an intention-strength coordinate. Consumer Decision Journey supplies a stage coordinate. The Persuasion Knowledge Model supplies a persuasion-resistance and trust-boundary coordinate. These coordinates define the reference distribution against which a latent-state representation is selected. The empirical question is whether that selected representation predicts subsequent ad response better than an action-only representation under the same data split and whether its directional recommendations are stable in later campaigns.

The deployed system is the anonymized retailer's omni-channel RMN platform. The lifecycle phase is online policy deployment with feedback monitoring. The reported engineering metrics are one-step AUC, cell-level directional agreement, verified checkout-direction accuracy, and temporal before-after checkout counts. The headline evidence uses metrics supported by the released export. Model-derived expected hitting time remains a diagnostic when the needed transition matrices are available. Segment-level conversion lift, opt-out reduction, and 60-day retention lift are reserved for deployments with a simultaneous comparison arm or the required telemetry fields.

The empirical audit through a live platform supports $\mathcal{S}$ in three ways. The one-step leakage-free comparison across 1,008 paired test consumers gives Full AUC $= 0.717$ versus action-only baseline AUC $= 0.635$ with DeLong $Z = 5.64$ and $p = 1.7 \times 10^{-8}$. Campaign 2 preserves 13 of 15 cell-level response directions, and its verified checkout table gives 80.65% directional agreement. Campaign 3 increases total checkouts from 358 to 512 relative to Campaign 1, which serves as a temporal before-after deployment audit of the selected policy. The ablations further show that non-$\mathcal{S}$ systems do not reproduce the full-model gain.

2. Literature

The literature relevant to this paper is not a list of antecedent findings. It discovers knowledge type of research gaps and supplies the primitives needed by the deployed system. Advertising-content and involvement studies motivate the observable covariates x_t. Dynamic purchase-readiness studies motivate a sequence model with transitions and durations. Consumer-behavior theory supplies semantic coordinates for the reference posterior. This section therefore maps prior work to model components rather than claiming a new psychological theory of intention.

2.1. Advertising content, product involvement, and delayed response

Advertising content affects consumer response through information, affect, salience, and product involvement. Prior work documents that advertising can have delayed effects and that short conversion windows may miss carryover from brand exposure to later search or purchase behavior (Filippou et al., 2025). Creative features also matter. Visual salience, psychological benefits, humorous tone, informational content, and emotional content have all been linked to engagement or sales under different settings (Hartnett et al., 2016; Bruce et al., 2020; Guitart and Stremersch, 2021; Gu et al., 2023; Tsai and Honka, 2021; Hamby and Jones, 2022). Product involvement changes the interpretation of these cues because high-involvement decisions usually require more diagnostic information, while low-involvement decisions often respond to simpler prompts and convenience cues (Quester and Lin Lim, 2003; Chen and Guo, 2022).

In the model, ad type and product involvement are decision covariates that change the emission and transition behavior of a clickstream process. We treat ad type and involvement as observed inputs in x_t, while the latent state remains an operational representation inferred from event dynamics.

2.2. Dynamic purchase readiness and intervention timing

Purchase readiness is temporally unstable. Trend estimates can change when structural breaks are present, and retail interventions often overlap in time (Moriyama et al., 2024). Funnel and customer-journey research similarly emphasizes that consumers move through stages rather than remaining in a fixed segment (Qin et al., 2024). Static segmentation and customer lifetime value models are useful for reporting, but they often remain backward-looking and do not explicitly encode migration pressure across states (Kanchanapoom and Chongwatpol, 2023).

This stream motivates the semi-Markov component. A hidden Markov model can represent transitions, but it treats state persistence implicitly through self-transition probabilities. An HSMM makes dwell time explicit. That distinction matters in RMN decisioning because the timing of an

ad depends not only on which state is likely but also on how long the consumer is expected to remain there. The model therefore conditions the initial, transition, emission, and duration kernels on (x_t, r_t) rather than fitting a covariate-stationary chain.

2.3. Theory-to-model mapping for latent-state semantics

The paper semantically anchors on Theory of Planned Behavior, Consumer Decision Journey, and the Persuasion Knowledge Model. TPB motivates an intention-strength coordinate because intention depends on attitudes, subjective norms, and perceived behavioral control (Ajzen, 2020). CDJ motivates a stage coordinate because consumers move among awareness, consideration, intent, and purchase stages rather than following a fixed linear funnel (Qin et al., 2024). The Persuasion Knowledge Model motivates a persuasion-resistance and trust-boundary coordinate because consumers interpret and cope with persuasion attempts, especially when targeting feels intrusive (Friestad and Wright, 1994).

This study uses an operational theory mapping. The model declares a finite state map $z = (\kappa, \sigma, \tau)$ and tests whether a retrieval-augmented theory reference distribution selects an HSMM parameterization with stronger deployment-tested predictive content. In this sense, TPB, CDJ, and PKM supply semantics for identification, while survey measurement remains outside the released telemetry.

A related stream on AI-mediated information search and personal knowledge management helps explain why retrieved context may influence consumer journeys (Batat, 2022; Forte, 2022; Kaya, 2024; Vorderer and Klimmt, 2023; North, 2023). We keep that stream conceptually separate from the Persuasion Knowledge Model. In the notation and empirical model below, PKM denotes persuasion knowledge and persuasion resistance, not personal knowledge management.

3. Theory-Identified Hidden Semi-Markov Dynamics for Purchase Intention

3.1. Problem Description

E-commerce clickstream data comprises timestamped user interactions, including page visits, search queries, product detail views, add-to-cart actions, and checkout events. These observed behaviors are compatible with many latent-state explanations. The binding methodological problem is therefore identification rather than raw modeling capacity.

A hidden semi-Markov chain learned from clickstream alone is generically under-identified. Different parameter sets (π, A, B, G) can produce indistinguishable distributions over click sequences while inducing different latent trajectories. The recovered states are therefore statistical clusters unless an additional semantic restriction is imposed. Layering an emission encoder on top of this chain does not resolve the identification problem. It re-parameterizes the same latent-state ambiguity.

Our construction intervenes at the identification step itself. A retrieval-augmented language model produces a theory-coherent reference distribution γ_t^{T} over operational intention states, conditioned on TPB, CDJ, and Persuasion Knowledge Model coordinates. Semi-Markov inference uses this reference distribution as a regularizer in the expectation step. The chain supplies the duration and transition structure in which γ_t^{T} can select among observationally compatible latent trajectories. The language model supplies the theory-conditioned reference distribution at clickstream scale, using textual theory, retrieved contextual knowledge, and the running session context.

Empirically, this construction supports what we call *state-response directional agreement*. The quantity is the agreement between an ad-response direction predicted from the inferred state distribution and the response direction realized after deployment. Its empirical role is deployment-tested prediction rather than survey measurement of inner attitudes. Sections 3 and 4 develop the construction, and Campaigns 1–3 close the loop in three sequential deployments with a major omni-channel retailer. Fig. 2 shows the end-to-end pipeline schematically; Table 1 fixes the notation used throughout the rest of this section.

The dynamic model is fitted on an event-level clickstream export with one row per timestamped event, including anonymized member identifier, session identifier, event time, event symbol y_t, shopper trait, product-involvement tag, ad-type tag, and campaign identifier. The tables in the Section 4 are aggregate audit summaries of that event-level export.

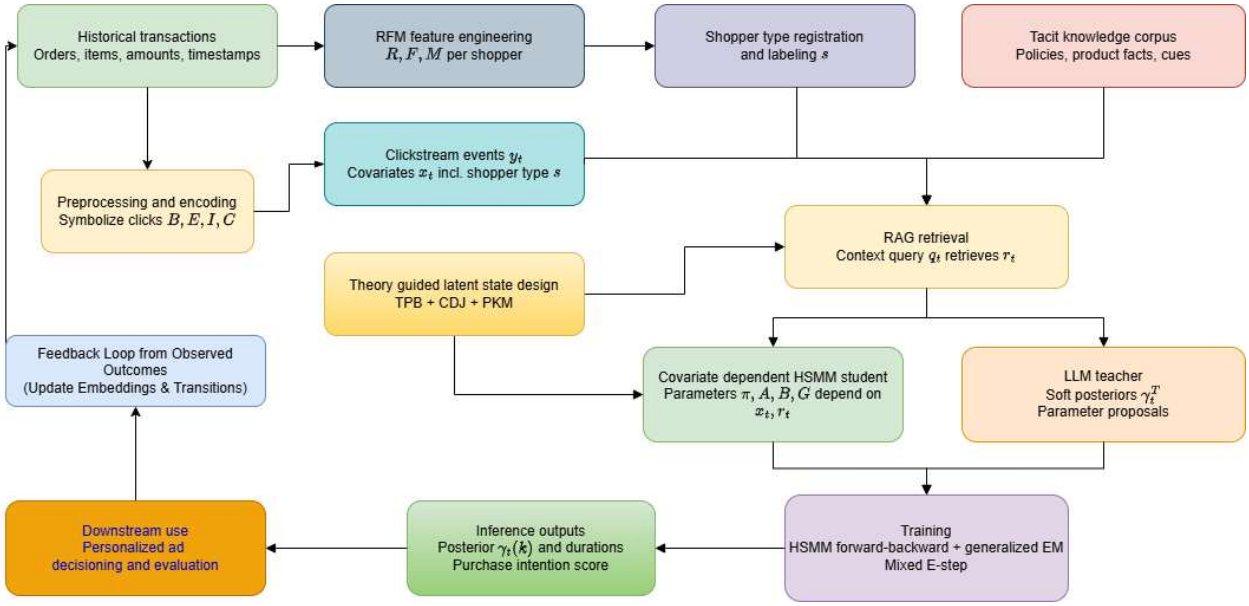


Figure 2 The research framework for suggesting effective ad strategies. The covariate-dependent semi-Markov inference is regularized by a retrieval-augmented theory-coherent reference distribution; the resulting posterior over latent states drives the ad-response policy whose directional predictions Campaigns 1–3 deploy and verify.

Let $\{1, \ldots, T\}$ index the clickstream timeline over N experiment sessions. Let $y_t \in \mathcal{Y} = \{B, E, I, C\}$ denote the observed click alphabet, where B is Browse, E is Evaluate, I is Intent, and C is Checkout. At each time t, we observe a structured covariate vector $x_t \in \mathcal{X}$ encoding shopper traits (Thrifty, Chill, Disciplined, Social, Opportunistic, Health-conscious), merchandise involvement (high- or low-involvement), and ad type (informational or emotional).

Additionally, we permit a retrieval-augmented context $r_t \in \mathcal{R}$ (text snippets describing local consumer traits, local events, seasonal patterns, etc.) obtained by RAG using a query constructed from the current session context and covariates. The pair (x_t, r_t) is treated as exogenous input.

Let $z_t \in \{1, \ldots, K\}$ denote the latent operational intention state at time t. Because operational states can persist for variable dwell times, we model $\{z_t\}$ as a hidden semi-Markov process (HSMM) with parameters $\pi(\cdot)$, $A(\cdot)$, $B(\cdot)$, and $G(\cdot)$. For notational ease, we denote $(x_{1:T}, y_{1:T})$ as the full observed sequence.

A hidden semi-Markov model (HSMM) $\mathcal{H}_\theta = \big(\pi_\theta(x,r), A_\theta(x,r), B_\theta(x,r), G_\theta(x,r)\big)$ is defined by the initial distribution $\pi(\cdot)$ on $\{1, \ldots, K\}$, transition matrix $A(\cdot)$ with entries $a_{ij}(\cdot)$ for transitions between latent states at session boundaries, emission distributions $B(\cdot)$ mapping latent states to observed clicks

$$b_i(y \mid \cdot) = \mathbb{P}(y_t = y \mid z_t = i, \cdot), \quad y \in \{B, E, I, C\},$$

and duration distributions $G(\cdot)$ for dwell times

$$g_i(d \mid \cdot) = \mathbb{P}(\text{duration} = d \mid z \text{ is } i, \cdot), \quad d \in \{1, \ldots, D_{\max}\}.$$

Table 1 Notation for the covariate-dependent hidden semi-Markov model and the retrieval-augmented teacher posterior.

Symbol	Type/Domain	Meaning
HSMM		
$\mathcal{H}_\theta$	HSMM	$\mathcal{H}_\theta = \big(\pi_\theta(x,r), A_\theta(x,r), B_\theta(x,r), G_\theta(x,r)\big)$
t	set	index $t \in \{1,\dots,T\}$ for the clickstream timeline over N experiment sessions
z_t	set	$z_t \in \{1,\dots,K\}$ the latent operational intention state, implemented as a Cartesian product of CDJ=$\{A,C,I,P\}$, TPB=High,Medium,Low, and PKM=High,Low
$\mathcal{Y}$	set	$y_t \in \mathcal{Y} = \{B,E,I,C\}$ the observed click alphabet (B=Browse, E=Evaluate, I=Intent, and C=Checkout)
$\pi(\cdot)$	vector	initial distribution
$A(\cdot)$	matrix	$a_{ij}(\cdot)$ for transitions between latent states
$B(\cdot)$	matrix	$b_{ij}(\cdot)$ emission distributions mapping latent states to observed clicks
$G(\cdot)$	vector	duration distributions for dwell times of z_t at state k
$(x_{1:T}, y_{1:T})$	timeseries	covariates and clickstream
$\gamma_t(i)$	matrix	the posterior $\gamma_t(i) \equiv (z_t = i \mid x_{1:T}, r_{1:T}, y_{1:T})$
LLM		
c_t	timeseries	context $(x_t, y_{1:t})$ the observable information at time t.
$\mathcal{K} = \{d_m\}$	database	tacit knowledge-base and accessed via a retrieval distribution $p(D_t = m \mid c_t)$.
r_t	string	tacit knowledge representation = the expectation of $\psi(d_m)$ under $p(D_t \mid c_t)$.
$p(z_t \mid c_t)$	soft distr.	teacher distribution $\approx \sum_{m=1}^{M} p(z_t \mid c_t, d_m)\, p(D_t = m \mid c_t)$,.
γ_t	distr.	EM posterior for student.
$\tilde{\gamma}_t$	distr.	mixed posterior for student.

A covariate-dependent HSMM conditions on the exogenous inputs $(x_{1:T}, r_{1:T})$ with the joint probability

$$\begin{aligned}&\mathbb{P}_\theta\big(y_{1:T}, q_{1:N}, d_{1:N} \mid x_{1:T}, r_{1:T}\big)\\ &= \pi_\theta(q_1 \mid x_1, r_1) \prod_{n=1}^{N} g_{\theta,q_n}\big(d_n \mid x_{s_n}, r_{s_n}\big) \prod_{n=1}^{N-1} a_{\theta,q_n q_{n+1}}\big(x_{s_{n+1}}, r_{s_{n+1}}\big) \prod_{t=1}^{T} b_{\theta,z_t}\big(y_t \mid x_t, r_t\big), \qquad (1)\end{aligned}$$

where the $b_{\theta,z_t}(y_t \mid x_t, r_t)$ is the output distribution over $\{B,E,I,C\}$.

The marginal likelihood from observed data sums the probability (1) over all sessions and latent states

$$\mathcal{L}(\theta) = \mathbb{P}_\theta(y_{1:T} \mid x_{1:T}, r_{1:T}) = \sum_{N} \sum_{q_{1:N}} \sum_{d_{1:N}} \mathbb{P}_\theta\big(y_{1:T}, q_{1:N}, d_{1:N} \mid x_{1:T}, r_{1:T}\big).$$

Direct summation is intractable, but dynamic programming yields exact inference for fixed K and $D_{\max}$. After HSMM inference (forward-backward for HSMM), we aim to get the posterior

$$\gamma_t(i) \equiv (z_t = i \mid x_{1:T}, r_{1:T}, y_{1:T}), \qquad i = 1,\dots,K.$$

The conventional method for creating personalized advertising was often clustering. Multi-category data help differentiate customers and create operationally usable covariate labels. In multi-run advertising experiments, consumers assigned to targeted ads are treated as operationally comparable within the same cluster. Because this equivalence relation is defined by observed attributes,

it supplies the covariate layer rather than the latent dynamic state. The HSMM then adds the event-level transition and duration structure that static segmentation cannot provide.

To build a transaction table assembly (Table 2) keyed by shopper (user) identifier $\{(u, t_i, \text{amount}_i, \text{order_id}_i)\}_{i=1}^{N}$, we start by defining an analysis window and anchor time. Choose a fixed lookback window W (180 days) and an anchor time T (session start scoring date). Compute raw RFM, for each shopper u within $[T-W, T]$,

$$R_u = T - \max\{t_i : i \in \mathcal{I}(u)\}, \quad F_u = \#\{\text{orders of } u\}, \quad M_u = \sum \text{amount}_i.$$

Normalize and stabilize the number by applying transforms such as $\log(1+F_u)$, $\log(1+M_u)$, winsorization for heavy tails, and robust scaling or quantile binning. Derive discrete shopper types by clustering (R, F, M) (k-means after scaling) and mapping clusters to interpretable labels to obtain s_u. Register types in a profile store by persisting into a profile record $(u, s_u, R_u, F_u, M_u, T)$. At scoring time, attach shopper type s_u as a covariate:

$$x_t \leftarrow [x_t;\ \text{onehot}(s_u)]$$

and incorporate s_u into retrieval queries q_t to obtain more type consistent context r_t.

The hidden state space is designed to be interpretable by construction, and then learned statistically. Using consumer-behavior theory, we decompose the state space into operational coordinates whose semantics are declared before estimation.

Based on the theory of planned behavior (TPB), we embed an intention-strength coordinate $\sigma \in \mathcal{S}$, e.g., low, medium, and high. Based also on consumer decision journey (CDJ), we embed a stage-aware intention coordinate $\kappa \in \mathcal{K}$, e.g., awareness, consideration, intent, and purchase. Furthermore, based on the persuasion knowledge model (PKM), we embed a bounded persuasion and trust preservation coordinate $\tau \in \mathcal{T}$, e.g., low-trust/high-resistance and high-trust/low-resistance, capturing resistance and trust preservation.

The Cartesian product can construct the composite state space.

$$z_t \equiv (\kappa_t, \sigma_t, \tau_t) \in \mathcal{K} \times \mathcal{S} \times \mathcal{T}, \qquad K = |\mathcal{K}|\,|\mathcal{S}|\,|\mathcal{T}|. \tag{2}$$

This construction provides interpretability for shopper-state discovery and sufficient flexibility for personalized ad decisioning. Because complete enumeration is inefficient and some combinations are rare in practice, the implementation selects the representative states with the highest empirical relevance in the observed clickstream.

Conceptually, each latent state k corresponds to a triple

$$\big(S(k), J(k), P(k)\big),$$

where $S(k)$ encodes intention strength (TPB), $J(k)$ encodes consumer decision journey stage (CDJ), and $P(k)$ encodes persuasion knowledge and trust stance (PKM).

Define three mapping functions

$$\phi_{\text{TPB}} : \mathcal{Z} \to \mathcal{S}, \quad \phi_{\text{CDJ}} : \mathcal{Z} \to \mathcal{J}, \quad \phi_{\text{PKM}} : \mathcal{Z} \to \mathcal{P},$$

where $\mathcal{S}$, $\mathcal{J}$, and $\mathcal{P}$ are finite sets of labels for intention strength, CDJ stage, and persuasion stance, respectively. These mappings, constructed using domain theory, map K latent states to interpretable states and can be implemented as simple lookup tables, where the K states serve as indices, with the mappings used for interpretation, reporting, and policy simulation.

The current framework employs a covariate-dependent HSMM whose latent states represent theory-grounded intention states combining TPB, CDJ, and PKM dimensions. The LLM supports

this framework at three levels. 1) State-space construction: proposing and labeling K interpretable intention states and their TPB/CDJ/PKM mappings. 2) Prior elicitation: translating qualitative theory into Dirichlet and Gaussian priors for emissions, transitions, and covariate effects. 3) Posterior guidance: providing soft, theory-consistent state assignments for observed sequences to regularize HSMM inference. This framework yields a Markov chain that estimates operational purchase-intention trajectories from clickstream data, conditioned on known covariates such as shopper characteristics, merchandise involvement, and ad type, while keeping the latent-state semantics tied to declared consumer-behavior theory.

We use γ_t to compute 1) dominant state trajectories, 2) dwell-time profiles per state, and 3) covariate-conditioned transition patterns, for shopper-state discovery and personalized ads. Given the current posterior over latent states and covariates, the next ad-type covariate value $a \in \{\text{informational}, \text{emotional}\}$—defined operationally by the retailer's creative-tagging codebook in the Analysis section, not as a psychological appeal type—is selected to maximize expected downstream conversion while respecting the PKM-driven trust constraint $\mathcal{R}(\theta)$ defined later in this section.

3.2. Retrieval-Augmented Teacher for Tacit Knowledge r_t

The identification argument requires a reference distribution γ_t^{T} that depends on textual theory and on tacit, session-level context that no fixed-format spreadsheet can encode. This subsection constructs that distribution by routing the LLM teacher through a retrieval-augmented pipeline. The operational distinction is that the retrieved text is mapped into all four HSMM kernels before inference, rather than being appended after inference to a downstream classifier.

Let the RAG system return a knowledge r_t given a query $q_t = q(c_t)$ with context $c_t = (x_{1:t}, y_{1:t-1})$. The context influences the teacher's inference parameters, that is, $c_t \to q_t \to r_t \to p_t \to h_t \to \mathcal{H}_\theta$. The context also influences the teacher's posterior distribution $c_t \to d_m \to z_t$.

Let $\mathrm{Enc}(x_t)$ be a deterministic serialization of the fixed-format covariates into tokens. Define the LLM input at time t such that

$$p_t = \mathrm{Concat}\big(\mathrm{Enc}(x_t),\ r_t\big).$$

Let $h_t = f_\theta(p_t) \in \mathbb{R}^H$ denote a learned representation produced by the pretrained LLM with parameters θ, possibly adapted by lightweight fine-tuning.

The teacher run used a fixed prompt template, fixed retrieval corpus, temperature 0, top-K_R retrieval, and a fixed random seed for all deterministic preprocessing. The API model identifier, prompt hash, retrieval-corpus hash, and parse-status log are recorded with each run. Reported teacher posteriors are accepted only when the parsed vector is non-negative and sums to one within numerical tolerance. These settings make the reference distribution auditable even though the LLM itself is not a generative clickstream model.

To guarantee valid probabilities, define the following softmax heads,

$$\pi_\theta(i \mid x_1, r_1) = \frac{\exp(u_i^\pi(h_1))}{\sum_{k=1}^K \exp(u_k^\pi(h_1))}.$$

For each source state i, define a row-wise softmax,

$$a_{\theta,ij}(x_t, r_t) = \frac{\exp(u_{ij}^A(h_t))}{\sum_{k=1}^K \exp(u_{ik}^A(h_t))}.$$

For each latent state i,

$$b_{\theta,i}(y \mid x_t, r_t) = \frac{\exp(u_{i,y}^B(h_t))}{\sum_{y' \in \mathcal{Y}} \exp(u_{i,y'}^B(h_t))}.$$

For each latent state i,

$$g_{\theta,i}(d \mid x_t, r_t) = \frac{\exp(u^G_{i,d}(h_t))}{\sum_{d'=1}^{D_{\max}} \exp(u^G_{i,d'}(h_t))}.$$

The functions u^π, u^A, u^B, u^G are affine maps of h_t, which will be calibrated empirically.

Assumption 1 (Boundedness and positivity) *For all (x, r) and all states, emissions, transitions, and durations satisfy:*

$$\pi_\theta(i \mid x, r) \geq \epsilon, \quad a_{\theta,ij}(x, r) \geq \epsilon, \quad b_{\theta,i}(y \mid x, r) \geq \epsilon, \quad g_{\theta,i}(d \mid x, r) \geq \epsilon,$$

for some fixed $\epsilon > 0$.

Theorem 1 (Properness and stability of the LLM-parameterized HSMM) *Under Assumption 1, the distributions produced by the LLM heads define a valid HSMM for every input sequence $(x_{1:T}, r_{1:T})$, and the marginal likelihood $\mathcal{L}(\theta)$ is finite and differentiable in θ. Proof in Appendix A.4.*

For convenience, define the observable context at time t as

$$c_t := \big(\mathbf{x}_i, y_{1:t}\big),$$

that is, covariates plus click history up to time t.

The reference distribution γ_t^{T} is constructed from a knowledge base $\mathcal{K} = \{d_m : m = 1, \ldots, M\}$ of theory-derived snippets and historical patterns. Marginalizing the retrieval-augmented teacher's posterior over which document is consulted, the latent-state posterior factorizes as

$$p(z_t \mid c_t) = \sum_{m=1}^{M} p\big(z_t \mid c_t, d_m\big)\, p(D_t = m \mid c_t), \tag{3}$$

where $p(D_t = m \mid c_t)$ is a softmax retrieval distribution over the knowledge base. Equation (3) is what makes the reference distribution operationally cheap (a small top-K_R retrieval suffices) and theoretically transparent (the dependence on textual theory enters in a single tractable place). The full derivation—knowledge-base embedding $\mathbf{k}_m$, discrete retrieval variable D_t, query encoder $\mathbf{q}_t$, similarity score $s_t(m)$, softmax distribution with temperature τ, and the deterministic mapping $r_t = \psi(d_{D_t})$—is in Appendix A.1.

Assumption 2 (Approximation capacity) *There exists a target measurable mapping $(x, r) \mapsto (\pi^\star, A^\star, B^\star, G^\star)$ and for any $\delta > 0$ there exists θ such that the induced distributions satisfy*

$$\sup_{(x,r)} \Big(\|\pi_\theta - \pi^\star\|_1 + \|A_\theta - A^\star\|_{1,\mathrm{row}} + \|B_\theta - B^\star\|_{1,\mathrm{row}} + \|G_\theta - G^\star\|_{1,\mathrm{row}} \Big) \leq \delta,$$

where $\|\cdot\|_{1,\mathrm{row}}$ denotes the maximum row-wise ℓ_1 distance.

Theorem 2 (Expressiveness: conditional HSMM parameter approximation) *Under Assumption 2, the LLM-parameterized family can approximate arbitrarily well any covariate- and context-dependent HSMM with finite K and $D_{\max}$. Consequently, tacit information encoded in r_t can be represented as systematic changes in the conditional kernels $\pi(x, r), A(x, r), B(x, r), G(x, r)$. Proof in Appendix A.4.*

The LLM uses tacit knowledge in r_t to adjust the structured covariate semantics by producing context-dependent $\mathcal{H}_\theta = \big(\pi_\theta(x,r), A_\theta(x,r), B_\theta(x,r), G_\theta(x,r)\big)$. As mentioned, the pre-modeled covariates collected in a fixed-format spreadsheet miss subtle and tacit information for each consumer and each transaction. We operationalize the challenges in an LLM plus RAG with versatile consumer reports.

Assumption 3 (Tacit factor and informative retrieval) *There exists an unobserved random variable U (tacit local context) such that the actual click model satisfies*

$$\mathbb{P}(y_{1:T} \mid x_{1:T}) = \int \mathbb{P}(y_{1:T} \mid x_{1:T}, u)\, d\mu(u),$$

and the retrieved text R is informative about U in the sense that $I(U;R \mid X) > 0$, where $I(\cdot;\cdot \mid \cdot)$ is conditional mutual information.

Theorem 3 (Value of retrieval: achievable log-loss strictly improves with informative retrieval) *Consider a one-step click prediction under the log-loss metric, and let $\ell^\star(\mathcal{F})$ denote the Bayes-optimal expected log-loss when conditioning on a sigma-field $\mathcal{F}$. Then*

$$\ell^\star(\sigma(X,R)) \le \ell^\star(\sigma(X)),$$

with strict inequality whenever $I(Y;R \mid X) > 0$. Consequently, a retrieval-augmented predictor that can represent $\mathbb{P}(Y \mid X, R)$ achieves strictly lower expected log-loss than any model restricted to $\mathbb{P}(Y \mid X)$. Proof in Appendix A.4.

The theorems formalize why the LLM-assisted HSMM is well-posed, expressive, and statistically beneficial under explicit assumptions.

3.3. Forward-Backward EM for the Covariate-Dependent HSMM

The identification argument requires a tractable E-step that produces γ_t^{S}, the student-side posterior that the next subsection's KL constraint will pull toward γ_t^{T}. This subsection delivers that E-step by extending the standard HSMM forward-backward to the covariate- and retrieval-conditional setting; the M-step's gradient form, the duration-aware recursion, and the PKM-consistent regularizer $\mathcal{R}(\theta)$ are deferred to Appendix A.2 because they are mechanical once the kernels are conditioned on (x_t, r_t).

Inference proceeds by a covariate- and retrieval-conditional forward-backward EM, extending the standard HSMM recursion of Yu (2010) to kernels of the form $\big(\pi(x,r), A(x,r), B(x,r), G(x,r)\big)$. The E-step computes posteriors over (state, duration, segment-boundary) triples; the M-step updates the LLM heads u^π, u^A, u^B, u^G by gradient ascent on the EM auxiliary $Q(\theta;\theta^{old})$ defined in equation (5) below. A PKM-consistent regularizer $\mathcal{R}(\theta)$ penalizes aggressive Checkout emissions in low-trust states. Full recursion (including the forward variable $\alpha_t(i)$ (13), backward variable $\beta_t(i)$ (14), segment-duration posterior $\eta_t(i,d)$ (15), smoothed occupancy $\gamma_s(i)$ (16), transition-count posterior $\xi_t(i,j)$ (17), one-step predictive $p^{(t)}(y)$ (19), gradient form for the softmax heads, and the trust-preserving regularizer) is given in Appendix A.2. The recursion is standard; what is non-standard, and what carries the identification argument, is that all four kernels are conditioned on (x_t, r_t) rather than on x_t alone.

Algorithm 1 Theory-Identified HSMM EM

Require: Event sequences $y_{1:T}$, covariates $x_{1:T}$, retrieval corpus $\mathcal{K}$, state map $z = (\kappa, \sigma, \tau)$, duration cap $D_{\max}$, teacher weight λ_{T}, tolerance ε.
Ensure: Conditional kernels $\mathcal{H}_\theta = (\pi_\theta, A_\theta, B_\theta, G_\theta)$ and posteriors $\gamma^{\mathrm{S}}_{1:T}$.

Initialize $\theta^{(0)}$ and retrieve r_t from $q(c_t)$ for each covariate key or event context.
for $m = 0, 1, 2, \ldots$ **do**
  Run the covariate-conditional HSMM forward-backward recursion with $\mathcal{H}_{\theta^{(m)}}$ to obtain $\gamma_t^{\mathrm{S},(m)}$, $\eta_t^{(m)}$, and $\xi_t^{(m)}$.
  Query the deterministic retrieval-augmented teacher, or load cached teacher outputs, to obtain γ_t^{T} from the TPB/CDJ/PKM state definitions.
  Form the identification-regularized occupancy counts

$$\widetilde{\gamma}_t^{(m)} = (1 - \lambda_{\mathrm{T}})\gamma_t^{\mathrm{S},(m)} + \lambda_{\mathrm{T}}\gamma_t^{\mathrm{T}}.$$

  Update $\theta^{(m+1)}$ by generalized M-step ascent on $Q(\theta; \theta^{(m)}) - \mathcal{R}(\theta)$ using $\widetilde{\gamma}_t^{(m)}$, $\eta_t^{(m)}$, and $\xi_t^{(m)}$ as soft counts.
  if $|\mathcal{J}(\theta^{(m+1)}) - \mathcal{J}(\theta^{(m)})|/(1 + |\mathcal{J}(\theta^{(m)})|) < \varepsilon$ **then**
    break
  end if
end for
Return $\mathcal{H}_{\theta^{(m+1)}}$ and the final smoothed posterior $\gamma^{\mathrm{S}}_{1:T}$.

3.4. Theory-Coherent Reference Distribution as the Identifying Constraint

This subsection formalizes the identifying mechanism that distinguishes our construction from any modular combination of a language model and a probabilistic sequence model. The reference distribution γ_t^{T} is not an "LLM module" that we plug into a sequence model; it is the constraint that resolves a specific structural pathology of the semi-Markov chain. We use the language "teacher" and "student" for continuity with prior literature, but the relationship is not assembly. The teacher quantity γ_t^{T} exists only because TPB, CDJ, and PKM are textual theories and clickstreams are textual events; the student inference exists only because γ_t^{T}, by itself, is non-generative, non-Markovian, and uncalibrated against the empirical click distribution. Neither side is interpretable as a deliverable on its own. The remainder of the subsection makes explicit what each side cannot do without the other. Throughout, let

$$\gamma_t^{\mathrm{T}}(k) = q_{\mathrm{T}}(z_t = k \mid c_t) \quad \text{and} \quad \gamma_t^{\mathrm{S}}(k) = p_{\mathrm{HSMM}}(z_t = k \mid \mathbf{y}_i, \mathbf{x}_i, \Theta),$$

denote the teacher and student posteriors over latent states z_t, and let

$$\tilde{\gamma}_t(k) = (1 - \lambda_{\mathrm{T}})\, \gamma_t^{\mathrm{S}}(k) + \lambda_{\mathrm{T}}\, \gamma_t^{\mathrm{T}}(k)$$

be the mixed posterior used in the HSMM E-step.

The construction confronts seven coupled identification issues, each of which a non-coupled pipeline would expose. The most consequential is the under-identification of HSMMs over clickstreams: distinct parameter sets Θ, Θ' produce indistinguishable click distributions $p_{\mathrm{HSMM}}(\mathbf{y} \mid \mathbf{x}, \Theta) \approx p_{\mathrm{HSMM}}(\mathbf{y} \mid \mathbf{x}, \Theta')$ while inducing different latent posteriors $p_{\mathrm{HSMM}}(z_t \mid \cdot, \Theta) \neq p_{\mathrm{HSMM}}(z_t \mid \cdot, \Theta')$. Mixing γ_t^{T} into $\tilde{\gamma}_t$ implicitly adds a regularizer

$$-\lambda_{\mathrm{T}} \sum_{i,t} \mathrm{KL}\big(\gamma_t^{\mathrm{T}}(\cdot) \,\big\|\, \gamma_t^{\mathrm{S}}(\cdot)\big)$$

to the EM objective, selecting, among observationally equivalent parameter sets, the one whose posteriors are theory-coherent. The remaining six issues—LLM non-generativity, textual theory that resists hard-coding, noisy student posteriors on short clickstreams, the auditability of LLM influence through $\lambda_{\mathrm{T}} \in [0,1]$, duration estimation under short sequences, and the cost asymmetry between offline elicitation and online inference—are coupled to the first one and to each other. Appendix A.3 treats each in detail; what matters for the main argument is that none of these issues can be addressed by either side acting alone.

From (1), the complete-data log-likelihood is

$$\begin{aligned}\log \mathbb{P}_\theta(y,q,d \mid x,r) = \log \pi_\theta(q_1 \mid x_1,r_1) + \sum_{n=1}^{N-1} \log a_{\theta,q_n q_{n+1}}(x_{s_{n+1}}, r_{s_{n+1}}) \\ + \sum_{n=1}^{N} \log g_{\theta,q_n}(d_n \mid x_{s_n}, r_{s_n}) + \sum_{t=1}^{T} \log b_{\theta,z_t}(y_t \mid x_t, r_t). \end{aligned} \tag{4}$$

Given current parameters θ^{old}, the EM auxiliary function is

$$Q(\theta;\theta^{old}) = \mathbb{E}_{\theta^{old}}\big[\log \mathbb{P}_\theta(y,q,d \mid x,r) \;\big|\; y,x,r\big].$$

Using posteriors from (15)–(17)–(16), one obtains an explicit form

$$\begin{aligned} Q(\theta;\theta^{old}) = &\sum_{i=1}^{K} \Gamma_1(i)\ \log \pi_\theta(i \mid x_1,r_1) \\ &+ \sum_{t=1}^{T-1}\sum_{i=1}^{K}\sum_{j=1}^{K} \Xi_t(i,j)\ \log a_{\theta,ij}(x_{t+1},r_{t+1}) \\ &+ \sum_{t=1}^{T}\sum_{i=1}^{K}\sum_{d=1}^{D_{\max}} \mathcal{H}_t(i,d)\ \log g_{\theta,i}(d \mid x_{t-d+1}, r_{t-d+1}) \\ &+ \sum_{s=1}^{T}\sum_{i=1}^{K} \gamma_s(i)\ \log b_{\theta,i}(y_s \mid x_s, r_s), \end{aligned} \tag{5}$$

where $\Gamma_1(i) = \mathbb{P}(z_1 = i \mid \cdot)$, $\Xi_t(i,j) = \xi_t(i,j)$, and $\mathcal{H}_t(i,d) = \eta_t(i,d)$.

Theorem 4 (EM monotonicity for the unregularized likelihood) *Let θ^{old} be the current parameters and define $Q(\theta;\theta^{old})$ as in (5) using the HSMM posterior. If an update produces θ^{new} with $Q(\theta^{new};\theta^{old}) \geq Q(\theta^{old};\theta^{old})$, then $\log \mathcal{L}(\theta^{new}) \geq \log \mathcal{L}(\theta^{old})$. Proof in Appendix A.4.*

When $\lambda_{\mathrm{T}} > 0$, the algorithm is best read as generalized regularized EM for the objective below. The monotonicity statement in Theorem 4 applies to the unregularized HSMM likelihood and should not be interpreted as a guarantee that $\log \mathcal{L}$ increases after adding the teacher and trust regularizers.

Writing the identification-regularized objective explicitly, the algorithm optimizes

$$J(\theta) \;=\; \log \mathcal{L}(\theta) \;-\; \lambda_{\mathrm{T}} \sum_{i,t} \mathrm{KL}\big(\gamma_t^{\mathrm{T}}(\cdot) \,\big\|\, \gamma_t^{\mathrm{S}}(\cdot;\theta)\big) \;-\; \mathcal{R}(\theta), \tag{6}$$

where $\log \mathcal{L}(\theta)$ is the marginal data log-likelihood from (1), the KL term is the identifying constraint that breaks observational-equivalence ties among parameterizations producing the same click distribution, and $\mathcal{R}(\theta)$ is the PKM-consistent trust-preserving regularizer that penalizes aggressive Checkout emissions in low-trust states. The mixing weight $\lambda_{\mathrm{T}} \in [0,1]$ in the mixed posterior (8) is

the empirical control on the strength of the identifying constraint. At $\lambda_{\mathrm{T}} = 0$, the objective degenerates to a pure HSMM likelihood. At $\lambda_{\mathrm{T}} = 1$, it collapses to teacher-only assignment. The useful regime is the intermediate region where the teacher reference selects among likelihood-compatible state labelings while the HSMM still fits the click distribution.

Proposition 1 (Regularized selection within an observational equivalence class) *Let $\mathcal{E}(\ell) = \{\theta : \log \mathcal{L}(\theta) = \ell\}$ be a non-empty level set of the HSMM marginal log-likelihood. Fix $\lambda_{\mathrm{T}} > 0$ and ignore $\mathcal{R}$ or hold it constant on $\mathcal{E}(\ell)$. If the map $\theta \mapsto \sum_t \mathrm{KL}(\gamma_t^{\mathrm{T}} \| \gamma_t^{\mathrm{S}}(\theta))$ has a unique minimizer $\theta^\star$ on $\mathcal{E}(\ell)$, then $\theta^\star$ is the unique maximizer of $J(\theta)$ on $\mathcal{E}(\ell)$. Thus the KL term does not identify the HSMM from observations alone. It selects the theory-coherent representative inside a likelihood-equivalent set under the declared reference distribution.*

O n $\mathcal{E}(\ell)$, the likelihood term in (6) is constant. Maximizing J is therefore equivalent to minimizing the summed KL discrepancy to γ_t^{T} when $\lambda_{\mathrm{T}} > 0$. Uniqueness of that minimizer gives the result.

The reference distribution $\gamma_t^{\mathrm{T}}(k) = q_{\mathrm{T}}(z_t = k \mid c_t)$ is parsed from the LLM's output under retrieval-augmented prompting. Using the mixture representation (3),

$$\gamma_t^{\mathrm{T}}(k) \approx \sum_{m \in \mathcal{N}_t} q_{\mathrm{T}}\big(z_t = k \mid c_t, d_m\big)\, p(D_t = m \mid c_t), \tag{7}$$

where $\mathcal{N}_t$ is a top-K_R retrieval set; full operational pipeline in Appendix A.1. The semi-Markov inference defines its own posterior $\gamma_t^{\mathrm{S}}(k) = p_{\mathrm{HSMM}}(z_t = k \mid \mathbf{y}_i, \mathbf{x}_i, \Theta)$. The two posteriors enter EM through a single mixed posterior

$$\tilde{\gamma}_t(k) = (1 - \lambda_{\mathrm{T}})\, \gamma_t^{\mathrm{S}}(k) + \lambda_{\mathrm{T}}\, \gamma_t^{\mathrm{T}}(k), \tag{8}$$

with $\lambda_{\mathrm{T}} \in [0, 1]$ controlling the strength of the identification regularizer. In the E-step, forward-backward computes γ_t^{S} and the mixture (8) is formed; in the M-step, the LLM heads are updated by gradient ascent on $Q(\theta; \theta^{old})$ with $\tilde{\gamma}_t$ used as soft state-occupancy counts.

The seven points above describe one mutual dependency rather than seven independent advantages. Identification (point 1) requires a textual theory expressible only as γ_t^{T} (point 3); γ_t^{T} acquires generative force, duration semantics, and forecastability only inside the semi-Markov chain (points 2, 6); the mixing weight λ_{T} is what makes the identification empirically auditable rather than dogmatic (point 5); sample efficiency and operational deployment (points 4, 7) follow because the resulting inference is well-posed rather than because we composed two efficient parts. Removing any one of these dependencies breaks the others. The point is not that we have built a flexible pipeline whose stages can be swapped; the point is that the dynamical system, the reference distribution, and the identification argument are co-defined and stand or fall together.

The teacher-student paradigm can be extended in the future. We can perform domain adaptation and model compression based on the current result. When $\mathcal{D}^{(s)}$ (source domain) differs from $\mathcal{D}^{(t)}$ (target domain), the teacher can label target sequences based on their posteriors γ_t^{T} and ξ_t^{T}. The student can learn a target-specialized model by simultaneously imitating the teacher's latent structure (high β), and fitting the target distribution ($\mathcal{L}_{\mathrm{NLL}}$), and optimizing downstream metrics ($\mathcal{L}_{\mathrm{sup}}$ if labels exist). When the student is a reduced-complexity model (with fewer states and lower emissions), distillation transfers the teacher's richer latent representation.

To assess computational complexity and deployment cost, we know that the forward-backward recursion is $\mathcal{O}(TK^2 D_{\max})$ per session per EM iteration, where T is the session length, $K = 24$ is the composite state cardinality, and $D_{\max}$ is the truncation of the duration distribution. Across 24,000 sessions in a Campaign-1 cohort with average session length 64 events and $D_{\max} = 8$, one EM iteration completes in ≈ 7 minutes on a single 16-core workstation; the run converged to a $< 10^{-3}$

relative log-likelihood change within six iterations (Fig. 5(a)). The teacher pass is per *covariate key* rather than per session: with 24 keys, the LLM is called once per key per EM iteration regardless of cohort size, so the per-session LLM cost amortizes to zero as the cohort grows. The retrieval index size is dominated by the TPB/CDJ/PKM-aligned corpus, which is <2 MB on disk; retrieval is in-memory at ad-decisioning time and adds <2 ms per query at the 5.2-million-member scale. The cost profile is therefore asymmetric: offline elicitation is bounded by $\mathcal{O}$(keys) LLM calls per training pass, and online inference per ad-decisioning event is the standard HSMM forward step. This fits a realistic RMN ad-decisioning budget.

4. Analysis

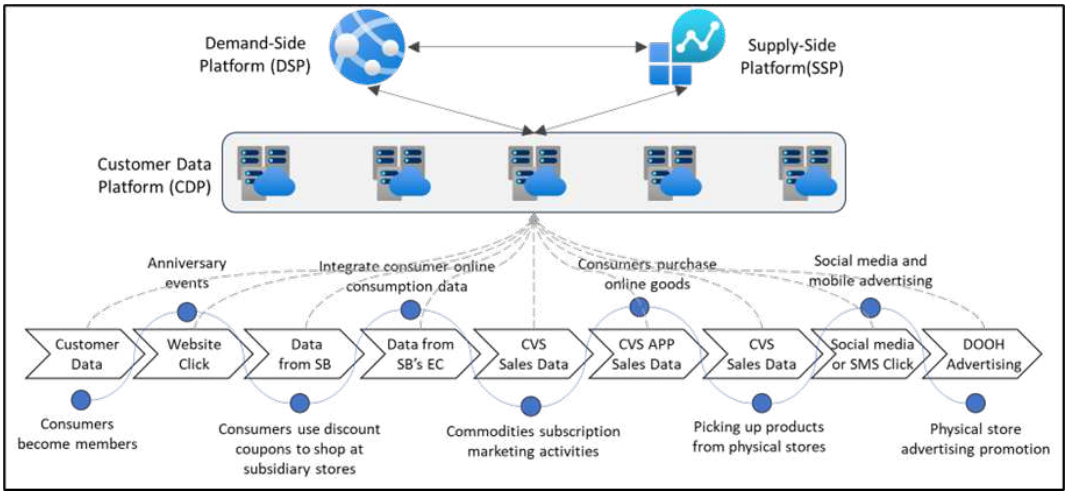


Figure 3 Omni-channel RMN setting for method illustration. The platform delivers targeted advertisements through first-party channels and records browse, evaluate, add-to-cart, and checkout events used by the model and by the closed-loop validation.

To demonstrate our method, we conduct experiments with associated implementation steps in a real-world e-commerce platform. The case company is an anonymized large omni-channel convenience-store retailer in East Asia with a mature loyalty program and a first-party retail-media infrastructure. The identifying commercial details are suppressed to preserve the data-use agreement. The analysis uses 5.2 million active loyalty members after excluding utility-only and negligible-spend accounts, and all reported campaign cohorts are drawn from this active-member universe. As shown in Fig. 3, the omnichannel RMN platform connects demand-side, supply-side, and customer-side components through first-party telemetry.

The accurate, personalized ads and consolidated clickstream data enable marketers to identify and reach individuals with a higher purchase propensity. The platform provides accurate feedback on the effectiveness of advertising strategies by monitoring the full funnel, including which shoppers viewed an ad, who engaged, who added items to the basket, who abandoned, and who completed a purchase. This closed-loop capability enables marketers to audit campaign response across the observed funnel and compute promotional ROI from first-party telemetry, subject to the comparison design actually deployed.

For the initial labeling, 5.2 million currently active members from the retailer's loyalty universe were used for the experiment, excluding accounts that only pay utility bills or spend $<$\$5 in 180 days. The clustering was conducted in 12 dimensions in terms of economic (discounted/coupons/redeem points), product (food/beverages/tobacco/alcohol/household/snacks/health), and consumption (expense/frequency/time slot), resulting in the nominal six labels as in Table 2, accounting

for 31%, 22%, 18%, 12%, 11%, and 6% among the total population, respectively, reflecting a even distribution of dataset and clustering method.

Table 2 Shopper-trait labels derived from 180-day purchase histories. The labels are operational covariates used for targeting and reporting, not psychological measurements.

Shopper Type	**Characteristics**	**Proportion**
Thrifty	Prioritize cost efficiency, with over 80% of their purchases utilizing discount codes, coupons, or accumulated loyalty points to earn free rewards. While their shopping frequency falls below average, their purchase volumes and spending exceed the mean. Their purchases predominantly consist of food and beverages, with no distinct geographic concentration	31%
Chill	Purchase based on immediate needs, exhibiting low price sensitivity. Their consumption peaks during breakfast hours, morning commutes, lunch breaks, and post-work commutes, with higher transaction density observed in transit-oriented stores (e.g., near stations). Their purchases are similarly skewed toward food and beverages, with minimal acquisition of household or beauty products	22%
Disciplined	Demonstrate indifference to discounts, adhering to fixed schedules for purchasing specific or homogeneous product categories. Though their per-transaction spending is modest, their purchase frequency significantly surpasses the average. For instance, they procure breakfast in the morning, beverages/coffee/cigarettes during work hours, and rice balls/boxed meals at noon. This behavioral pattern is markedly prevalent in metropolitan areas	18%
Social	Stockpile beverages, particularly coffee or tea, leveraging volume discounts. For example, their daily coffee consumption averages 5 times that of the general consumer base. Characterized by frequent redistribution and sharing of products, their purchase volumes far exceed individual daily consumption needs. While predominantly concentrated in urban areas, a smaller yet notable proportion is found in non-metropolitan areas. These consumers exhibit extensive mobility across multiple retail locations	12%
Opportunistic	Strategically await promotional periods, exhibiting acute sensitivity to brand-specific campaigns (e.g., corporate-sponsored team championships, milestone store openings). They prioritize optimal timing to bulk-purchase non-perishables and high-value items (e.g., gold), capitalizing on arbitrage opportunities through ultra-high bonus point incentives	11%
Health-Conscious	Less sensitive to discounts and promotions. They purchase at least one item that carries a health label and contains fewer than 500 kilocalories, including leafy green salads, fruit, soup, rice balls, sandwiches, bread, sweet potatoes, chicken breast, and soy milk. Their spending is not high and is only slightly above average, but their purchasing habits are pretty stable, and they are also willing to try newly launched health-oriented products. 6%	

To illustrate our method, this study launched three campaigns, as shown in Fig. 4. Campaign 1 ran from August 4 to August 17, 2023. Campaign 2 ran from November 24 to December 7, 2023. Campaign 3 ran from February 10 to February 23, 2025. Each launch had different purposes and outcomes. The first campaign aims to establish the initial framework for model training and delivery strategies. The second campaign is to measure the accuracy of our designed strategies by

collecting consumer purchase feedback. The third campaign evaluates the operational performance of the selected ad-delivery policy in a temporal deployment re-validation.

The experimental assignment logic differs across the three campaigns. *Campaign 1* draws 24,000 consumers uniformly at random from the 5.2M active members and allocates them uniformly across the cells of the 24-key covariate grid (six shopper traits × two involvement levels × two ad types), so each cell receives the same expected exposure regardless of cell base rate. The purpose is data generation for the initial fit, not policy comparison. *Campaign 2* draws a fresh 24,000 consumers (no overlap with Campaign 1's cohort, enforced by member-id de-duplication) and assigns them only to the 15 covariate cells flagged by the Campaign 1 fit as either high-conversion or low-conversion, in a 8-vs-7 high/low split. The remaining nine cells are intentionally not exposed in Campaign 2; they appear as "–" in Table 5. The purpose is to verify whether the Campaign 1-predicted ranking of cells survives in fresh consumers under a fresh deployment. *Campaign 3* again draws a fresh 24,000 consumers and exposes them only to the policy recommended by Table 4, the post-Campaign-2 strategy table; the purpose is operational validation under the deployed policy. This closed-loop deployment audit complements the paired holdout AUC benchmark reported later in the section, where the action-only covariate-cell model (baseline) serves as the reproducible within-test-set comparator.

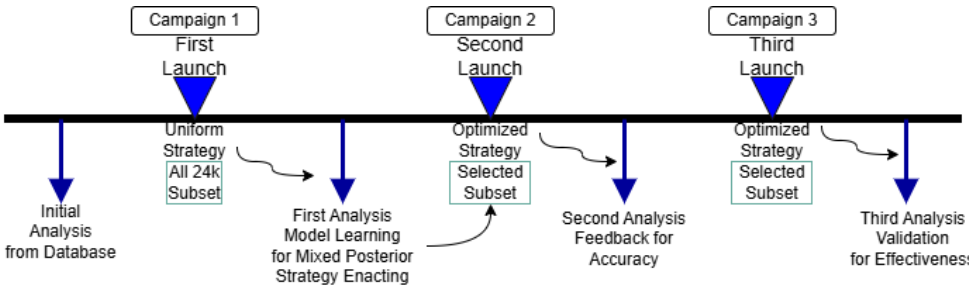


Figure 4 Timeline and purpose of the three deployment waves to demonstrate our method. Campaign 1 fits the initial model, Campaign 2 checks directional stability in fresh consumers, and Campaign 3 re-validates the recommended policy over a later temporal window.

4.1. A First Campaign for Model Building and Strategy Development

The first advertising deployment targeted 24,000 individuals, who were randomly selected from the 5.2M active consumers. The experimental design was based on a uniform allocation of covariates across combinations of consumer types, product involvement levels, and advertisement types. As shown in Table 3, this campaign generated 3,362 clicks (browse), 1,830 views (evaluate), 638 add-to-cart (intent) actions, and 358 purchases (checkout).

The retailer's standard creative production workflow was used, and creatives were tagged at production time along two binary dimensions that we use as covariates downstream: (a) *ad-type* $\in \{\text{informational}, \text{emotional}\}$, where informational creatives lead with factual product information (price, ingredients/specifications, calorie or composition labels, delivery and return terms) and emotional creatives lead with affective imagery, brand-belonging cues, or social-proof framing; (b) *involvement level* $\in \{\text{high}, \text{low}\}$, where high-involvement creatives feature considered-purchase categories (premium beverages, durable household goods, health-labeled food) and low-involvement creatives feature impulse or routine-purchase categories (snacks, ready-meal items, mainstream beverages). Each consumer in the targeted cohort was exposed to one creative per delivery slot, drawn uniformly within the assigned $(x_{\text{shopper}-\text{trait}}, x_{\text{involvement}}, x_{\text{ad}-\text{type}})$ cell of the experimental allocation. The creative-production workflow, the placement channel (the retailer's first-party app and partner properties), and the impression-frequency cap were held constant across the three campaigns;

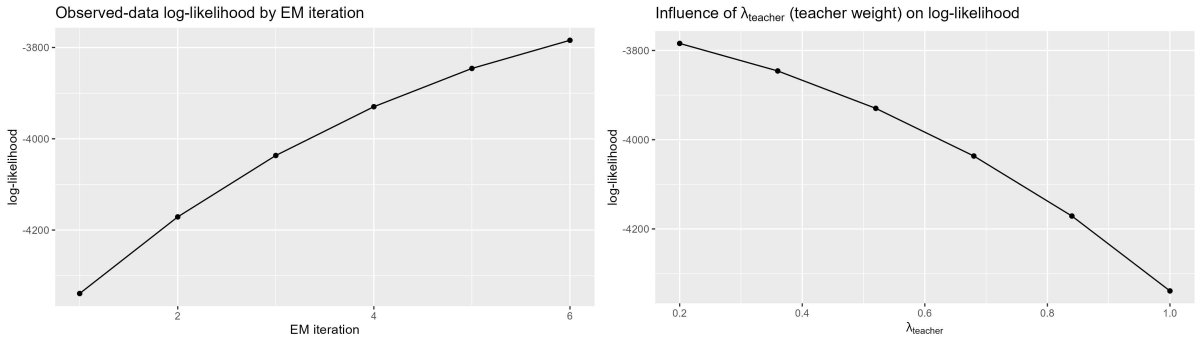
only the targeting policy varied across deployments. The full creative-tagging codebook is available from the case retailer under data-use agreement; we do not reproduce individual creatives here for confidentiality reasons.

The click-to-purchase ratio C/B represents the number of completed checkouts divided by browse events in a covariate cell. The cart-to-purchase ratio C/I represents the number of completed checkouts divided by add-to-cart events. Tables report these as ratios rather than percentages.

To screen consumers for the next phase of the campaign, we select the top head and tail effective checkout consumers for evaluation. The effective screening threshold was $Q_2 = 0.08$ for C/B and $Q_2 = 0.54$ for C/I. Conversion rates above Q2 were categorized as high conversion (denoted"+"), and those below Q2 as low conversion (denoted"−"). Based on these thresholds, two ad delivery configurations were selected in the last column of Table 3, one exhibiting high conversion on both metrics and the other exhibiting low conversion on both, for validation in the next phase. For example, presenting high-involvement and emotionally charged ads to the Thrifty Shopper resulted in above-average conversion rates for both C/B and C/I ratios. Conversely, presenting high-involvement and emotional ads to the Chill Shopper resulted in below-average conversion rates for both paths.

Table 3 Campaign 1 event counts by shopper trait, product involvement, and ad type. B, E, I, and C denote browse, evaluate, add-to-cart intent, and checkout. The final column marks cells selected for Campaign 2 directional re-validation.

Shopper	Involv.	Ad type	#	B	E	I	C	C/B	C/I	Next
Thrifty	High	Informational	1860	234	176	49	26	0.11 (+)	0.53 (-)	
Thrifty	High	Emotional	1860	243	185	32	25	0.10 (+)	0.78 (+)	Yes
Thrifty	Low	Informational	1860	214	114	31	18	0.08 (+)	0.58 (+)	Yes
Thrifty	Low	Emotional	1860	209	98	44	22	0.11 (+)	0.50 (-)	
Chill	High	Informational	1320	99	34	10	4	0.04 (-)	0.40 (-)	Yes
Chill	High	Emotional	1320	87	28	12	6	0.07 (-)	0.50 (-)	Yes
Chill	Low	Informational	1320	86	51	34	11	0.13 (+)	0.32 (-)	
Chill	Low	Emotional	1320	102	22	11	9	0.09 (+)	0.82 (+)	Yes
Disciplined	High	Informational	1080	102	11	2	1	0.01 (-)	0.50 (-)	
Disciplined	High	Emotional	1080	98	22	11	1	0.01 (-)	0.09 (-)	Yes
Disciplined	Low	Informational	1080	304	272	155	89	0.29 (+)	0.57 (+)	Yes
Disciplined	Low	Emotional	1080	244	143	76	33	0.14 (+)	0.43 (-)	Yes
Social	High	Informational	720	122	43	12	7	0.06 (-)	0.58 (+)	Yes
Social	High	Emotional	720	101	31	9	8	0.08 (+)	0.89 (+)	Yes
Social	Low	Informational	720	98	43	6	4	0.04 (-)	0.67 (+)	
Social	Low	Emotional	720	256	112	32	29	0.11 (+)	0.91 (+)	
Opportunistic	High	Informational	660	104	66	22	12	0.12 (+)	0.55 (+)	Yes
Opportunistic	High	Emotional	660	111	79	29	15	0.14 (+)	0.52 (-)	
Opportunistic	Low	Informational	660	156	103	32	24	0.15 (+)	0.75 (+)	Yes
Opportunistic	Low	Emotional	660	99	54	12	5	0.05 (-)	0.42 (-)	Yes
Health	High	Informational	360	98	51	4	2	0.02 (-)	0.50 (-)	
Health	High	Emotional	360	76	29	4	1	0.01 (-)	0.25 (-)	Yes
Health	Low	Informational	360	66	31	4	3	0.05 (-)	0.75 (+)	
Health	Low	Emotional	360	53	32	5	3	0.06 (-)	0.60 (+)	Yes
(total)	—	—	24000	3362	1830	638	358	Q2=0.08	Q2=0.54	



(a) iteration (b) weight

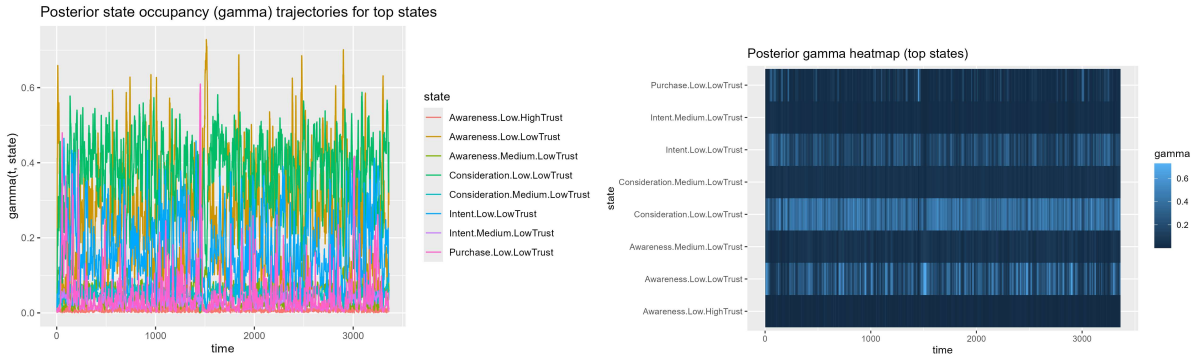
Figure 5 The log-likelihood values to iteration and teacher–student weighting. This diagram highlights calibrated soft guidance, where the teacher regularizes the posterior without dominating the data-fitted HSMM. The practical utility of the trained model is evaluated through follow-on campaigns in the retailer's operating environment.

4.2. Posterior Training and Policy Design

Based on the click stream data from the first campaign, Fig. 5 demonstrates that the HSMM provides a good fit for the observed click streams under teacher–student transfer, using the observed log-likelihood on the vertical axis (a higher value, less negative, indicates a better fit). When λ_{T} in $\tilde{\gamma}_t$ is higher, more weight is given to the teacher model.

The observed log-likelihood by EM iteration for a fixed λ_{T} indicates the log-likelihood increases monotonically across iterations (Fig. 5 (a)). The monotone increase is the expected behavior when the M step correctly maximizes the expected complete-data objective for the chosen model and λ_{T}. The flattening suggests near convergence. A stopping rule based on relative improvement (for example, a change in log-likelihood below a small threshold) would likely terminate soon after iterations.

In Fig. 5(b), the teacher information behaves as a soft constraint that regularizes the parameters toward theory-induced transition and emission structure. The decline at large λ_{T} shows why the LLM component is most effective as a calibrated prior. The empirical clickstream distribution keeps control of the fit, and the trained model is evaluated through follow-on campaigns in the retailer's operating environment.



(a) trajectory (b) heatmap

Figure 6 (a) The multi-line trajectories showing posterior state occupancy, $\gamma_t(s)$, for the top latent states. Several states fluctuate substantially over time with frequent spikes, while others remain near zero for most of the horizon, indicating that only a subset of states carries most of the posterior mass at any given time. (b) The same $\gamma_t(s)$ values are shown as a state-by-time matrix. Brighter horizontal bands mark states that are repeatedly assigned higher occupancy across time, whereas darker bands indicate persistently low occupancy, making state dominance patterns and temporal intermittency easier to see.

Fig. 6 visualizes the posterior state occupancy probabilities from the fitted HSMM. For each time index t and latent state s, the trajectory shows $\gamma_t(s) = P(S_t = s \mid y_{1:T})$. Large $\gamma_t(s)$ means the model believes the user is in state s at time t, given the entire observed sequence. A small set of states repeatedly dominate, especially "Consideration.Low.LowTrust" and "Awareness.Low.LowTrust", with "Intent.Low.LowTrust" also frequently substantial. "Purchase.Low.LowTrust" appears mostly near zero with occasional bursts, consistent with purchase being episodic rather than persistent.

The dominant "Consideration" with Low trust indicates prolonged evaluation without strong confidence, which is a natural target for interventions that reduce uncertainty. Recurrent movement between "Awareness" and "Consideration" with Low trust indicates repeated re-entry into early funnel cognition, suggesting friction or insufficient reassurance. Sparse "Purchase" spikes suggest that conversion is event-like, so triggering and timing of checkout-oriented stimuli should be tied to those brief windows rather than applied uniformly.

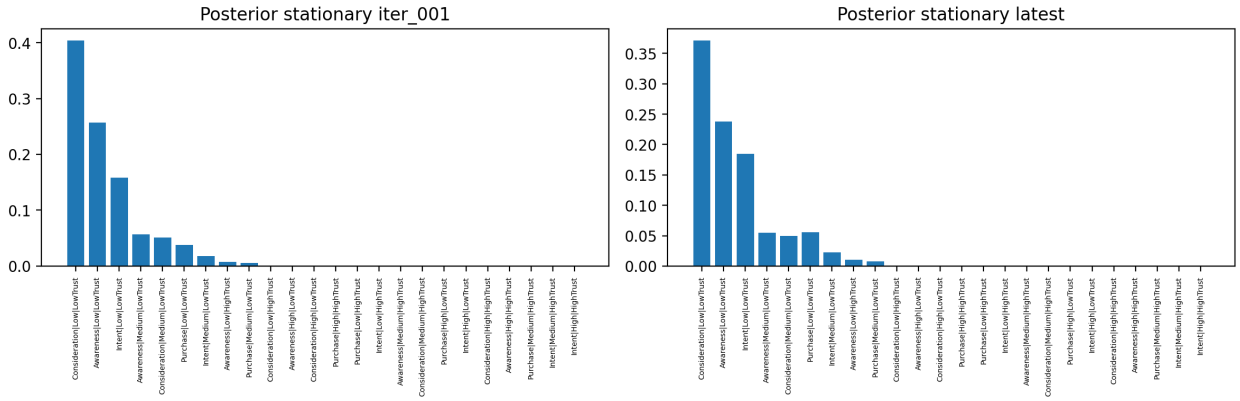


Figure 7 The comparison of the time-averaged HSMM posterior occupancy, i.e., the empirical limiting distribution of $\gamma_t(s)$ over time. The observed exponential decay in occupancy time is consistent with a sparse fitted latent-state representation in the HSMM. From iteration 1 to 6, the posterior stationary distribution over states remained sparse but showed a modest redistribution among the top states. The tail states remain negligible, indicating the chain still effectively operates on a small subset of important states.

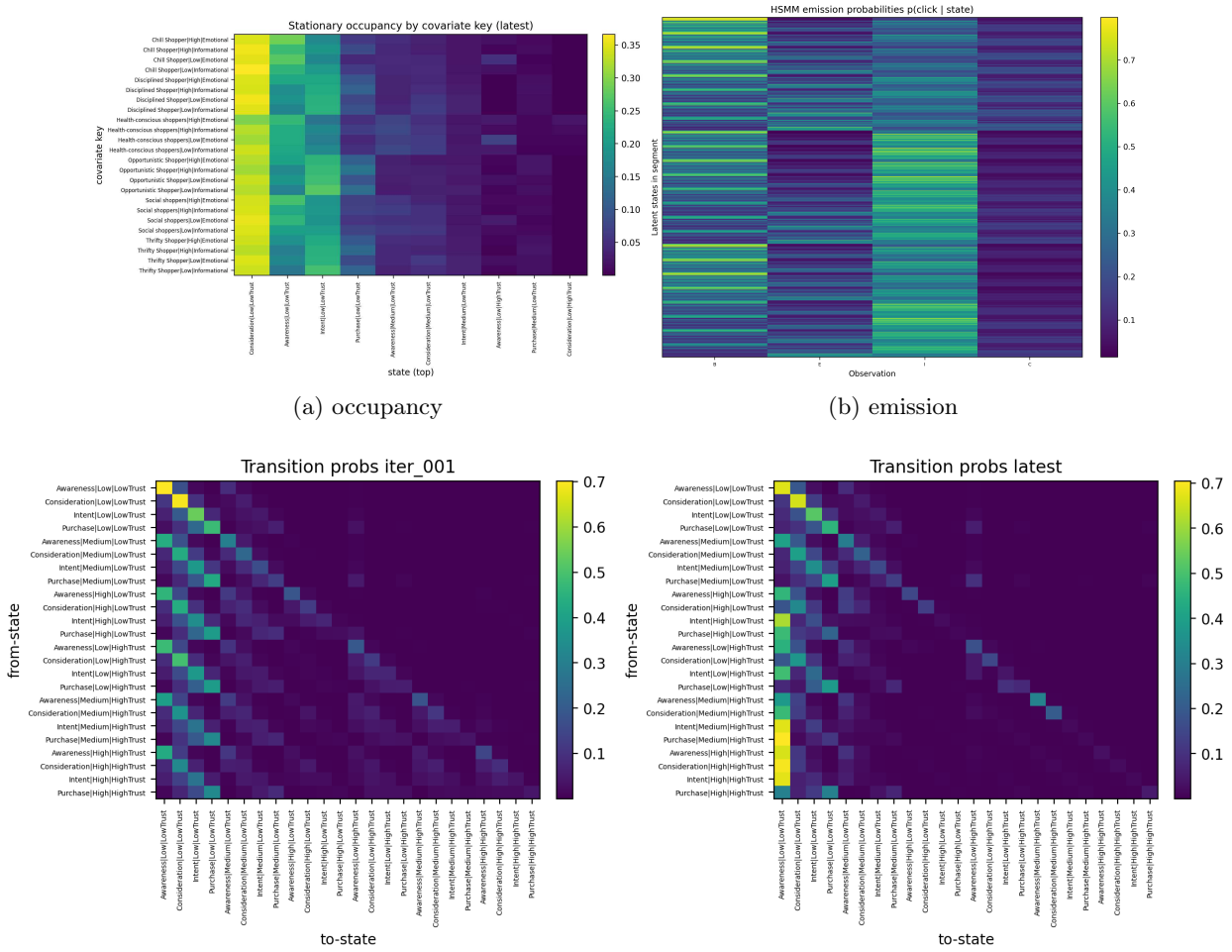
Fig. 7 shows the learning iteration of the empirical limiting distribution of γ_t over time and represents the time-averaged HSMM posterior occupancy. The occupancy values are sorted from largest to smallest. Therefore, the leftmost value implies that the states dominate the clickstream in terms of posterior time mass. For each latent state i, the HSMM inference produces $\gamma_t(i) = P(z_t = i \mid y_{1:T}, x_{1:T}, r_{1:T})$. The values in the chart are the time average $\bar{\gamma}(i) = \frac{1}{T}\sum_{t=1}^{T}\gamma_t(i)$, which is the expected fraction of time steps spent in state i under the posterior. The subgraphs (a) and (b) represent the posteriors at iterations 1 and 6, respectively. The evolution observed is stable.

The states are labeled in terms of theory-guided composition "CDJ stage|TPB intention strength|PKM trust level" Therefore "Consideration|Low|LowTrust" means the consumer is in the deliberation phase (comparing options and evaluating information), has weak purchase intention (unlikely to move toward Intent or Checkout), and is skeptical or resistant to persuasion (trust preservation). In Fig. 7 (a) and (b), most time steps are spent in early or mid-funnel states with low intention strength and low trust, requiring trust-building and uncertainty reduction (often consistent with informational evidence, transparent pricing, delivery, and returns) before they can stably transition toward Checkout.

The first few columns (especially "Consideration|Low|LowTrust") dominate across almost all covariate types, indicating a broad bottleneck, that is, consumers are frequently inferred to be evaluating with low intention and low trust. For "Awareness|Low|LowTrust", consumers are still early in the journey, with weak intentions and low trust. The next states, "Intent|Low|LowTrust" and

"Purchase|Low|LowTrust", imply that consumers approach intent or purchase but still exhibit low trust, which can manifest as hesitation, repeated evaluation, and abandonment. States with higher intention or higher trust receive substantially smaller posterior mass. This pattern is a fitted-model summary of the deployment data and provides the basis for the policy interpretation below.

Operational finding 1. Within the campaign study period, we observed that the fitted posterior concentrates on a small set of early and mid-funnel LowTrust states across covariate keys, which motivates trust-building and uncertainty-reducing creative policies. This observation highly matches common wisdom.



(a) occupancy (b) emission

(c) transition comparison

Figure 8 (a) The posterior time mass to the top latent states (columns) for each covariate type (row); (b) The emission probabilities from states to clicks for all covariates; (c) The evolution of the HSMM embedded-chain transition matrix over the 24 composite states from iteration 1 to 6.

Through the full operational policy table (24 covariate keys $\times$ 24 latent states $=$ 576 cells, of which 288 distinct ad-eligible *(covariate-state, ad-type)* pairs are exercised at deployment time), the advertising types can be detailed to the covariate level. At the covariate level, for personalized advertising strategies, we consider the expected one-step click probabilities $\widehat{P}(C \mid a) = \sum_i \gamma_t(i), b_i^{(covariate,a)}(C)$ for ad types $a = \{$ Informational, Emotional$\}$, plus a suggested budget mix derived from the predicted Checkout uplift. Given the predicted next-click distribution $(pred_B, pred_E, pred_I, pred_C)$ and state condition, our strategies can encourage consumers to take the Checkout action. The numeric

KPI targets are computed directly from the predicted emissions for that covariate-state under the recommended action.

Managers should treat LowTrust posterior mass as an actionable bottleneck rather than as a descriptive segment label. Budget should first fund evidence oriented creatives, transparent price and delivery claims, return assurances, and credible third party signals that can move consumers from Awareness or Consideration toward Intent. Campaign dashboards should track the posterior mass of the Consideration Low LowTrust state jointly with C/I, because a falling low trust state share without a stronger cart to purchase ratio would signal engagement without conversion quality.

To retrieve the operational manageability of 24 covariate keys, we take the two facts addressing operational manageability of the 24 covariate keys at deployment time. First, the 24 keys are not independent decision objects; the policy table collapses to a much smaller *ad-strategy alphabet* when one looks at which $(a, \text{involvement})$ combination is recommended per shopper trait. In our deployment, the 24 keys mapped to five distinct ad-strategy recommendations (Table 4), a cardinality the retailer's campaign-management workflow handles natively. Second, the framework is invariant to coarsening: collapsing two adjacent shopper traits (e.g., Disciplined and Health-Conscious) into a single key reduces the table to 20 keys without re-fitting the model, because the HSMM kernels are conditioned on x_t rather than on a partition of x_t. The framework also supplies a model-derived prediction of the merger cost even absent a controlled re-deployment: under the reading that policy loss scales with the posterior-distance between the two merged segments, the cost is the difference between the segments' mean posterior state distributions. For Disciplined and Health-Conscious shoppers, the realized posteriors (Fig. 8) give a mean per-row γ that differs by total variation $\text{TV} = 0.051$ ($L_1 = 0.101$, Jensen-Shannon 0.0035); both segments concentrate roughly three-quarters of their mass on LowTrust states, and the per-key policy in Table 4 maps both to the same Informational recommendation at Low involvement and the same Emotional recommendation at High involvement. The predicted merger cost is therefore the small but non-zero $\text{TV} = 0.051$: deployments whose directional response is only weakly sensitive to fine-grained state allocation can absorb the merger, while deployments whose response is sharply state-sensitive will see a proportional loss. A measured-in-deployment cost still requires a controlled re-deployment of the merged audience, which we do not run here.

Fig. 8 (a) reports, for each covariate type (row), how much posterior time mass the HSMM assigns to the top latent states (columns). The covariates are observable and serve as the factors for strategies. Each covariate is indexed by the covariate label "ShopperTrait | Involvement | AdType". The values of the posterior occupancy $\bar{\gamma}_g(i) := \frac{1}{|T_g|} \sum_{t \in T_g} \gamma_t(i)$ are represented by color intensity, where T_g is the set of time indices whose covariates map to key g, and $\gamma_t(i) = P(z_t = i \mid \text{data})$.

For each covariate row, brighter cells indicate the latent states in which consumers from that type spend more posterior time. The posterior occupancy conditioned on covariate type supplies practical decision rules for the next advertisement. Because most types concentrate on "Consideration|Low|LowTrust", the fitted policy favors creatives that supply information, transparency, and friction reduction. Retargeting sequences should follow the inferred progression from Awareness to Consideration to Intent and then to Purchase, while the effect of such sequences on transition probabilities must be tested in deployment rather than assumed from the fitted model alone.

In Fig. 8 (b), for each covariate key (shopper trait, involvement, ad type) and each latent state, (a) gives the emission probabilities from the 24 composite states based on $\{A, C, I, P\}$ to the observed clicks $y \in B, E, I, C$.

Fig. 8 (c) demonstrates the HSMM embedded-chain transition matrix among states, recording $P(S_{t+1} = j \mid S_t = i), i, j \in \{A, C, I, P\}$ from iteration 1 to 6. The bright diagonal blocks indicate persistence (the process tends to stay in the same state across jumps). The off-diagonal bright regions indicate the dominant progression routes and the dominant "reset" routes, that is, returning to the earliest awareness states.

From the Fig. 8, the embedded stationary distribution is dominated by low-trust early and mid-funnel states. The top mass is on $\{A, C, I\}$ with Low Intention and Low Trust states. This concentration is consistent with a trust and perceived-control bottleneck; therefore, the strategies should prioritize evidence, transparency, and friction reduction over aggressive emotional persuasion. The Health-conscious high-involvement covariate retains a larger share under emotional ads, but it still does not outperform informational ads on expected Checkout. Emotional creative is therefore treated as an upper-funnel test cell with strict evidence anchors.

Operational finding 2. Within the campaign study period, we observed that the fitted policy table favors informational creatives in most covariate-state cells, with only narrow emotional-creative exceptions for selected high-involvement states. This observation also matches our expectation about high-involvement merchandise needing information support.

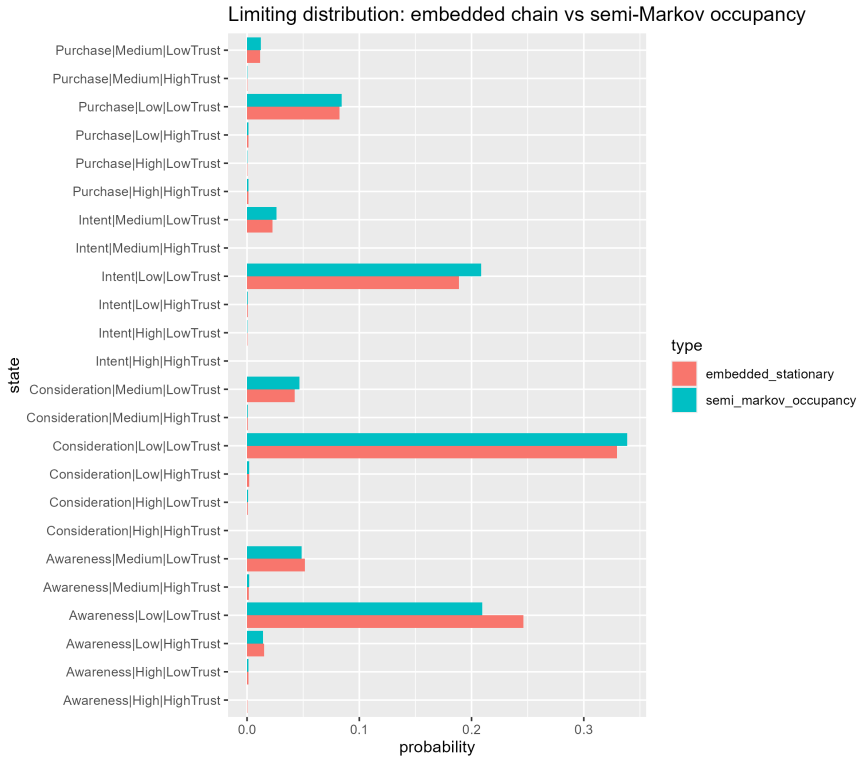


Figure 9 Two limiting distribution representations over the same set of states. For the embedded chain stationary distribution (red) and the semi-Markov long-run occupancy (teal), most probability mass concentrates on a small set of LowTrust states, especially "Consideration|Low|LowTrust" and "Intent|Low|LowTrust", with "Awareness|Low|LowTrust" and "Purchase|Low|LowTrust" as secondary contributors. The systematic gaps between red and teal indicate that accounting for state dwell times (semi-Markov occupancy) shifts the long-run weight relative to the embedded transition-based stationary distribution.

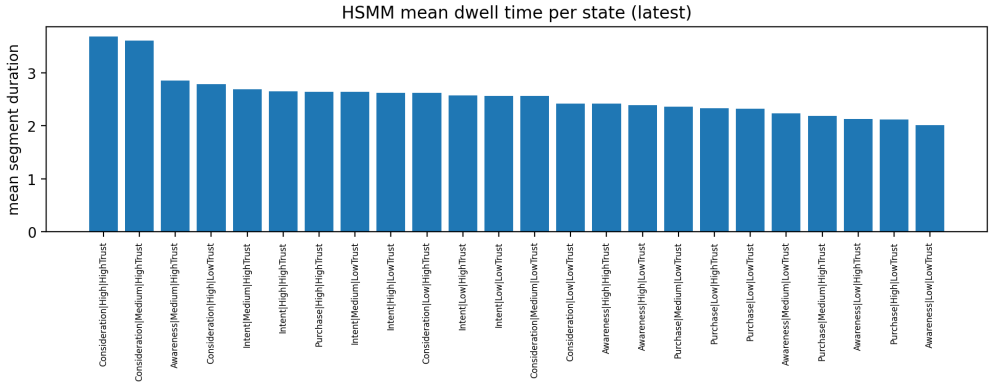
Fig. 9 compares two different notions of a long-run distribution for the same fitted HSMM type model. The embedded stationary (red) represents the stationary distribution of the embedded jump chain, namely the discrete-time Markov chain observed only at transition instants (state entry times), where each state visit contributes 1 count, regardless of the process's duration in that state. Formally, if P is the transition matrix among states at jump times, the embedded stationary vector satisfies $\pi^{\text{emb}} = \pi^{\text{emb}} P, \quad \sum_i \pi_i^{\text{emb}} = 1$. This limiting distribution can be interpreted as the long-run fraction of state entries that occur in each state.

The semi-Markov occupancy (teal) is the limiting time-average probability of being in each state for a continuous-time or duration-aware process, explicitly weighting each state by its expected

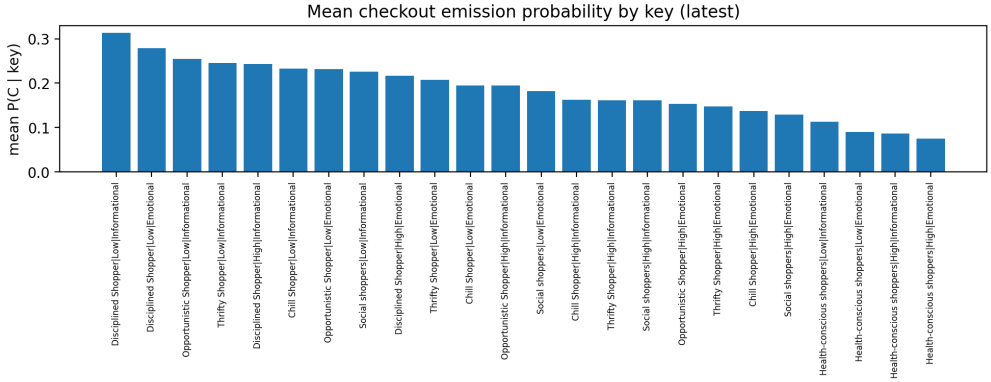
sojourn time (the expected duration of staying in that state). The occupancy distribution is $\pi_i^{\text{occ}} = \frac{\pi_i^{\text{emb}} m_i}{\sum_j \pi_j^{\text{emb}} m_j}$, where $m_i = \mathbb{E}[D_i]$ is the mean duration in state i. This limiting distribution can be interpreted as the long-run fraction of time spent in each state.

If all states had the same mean duration, the two distributions would match. When durations differ, states with longer mean durations have $\pi_i^{occ} > \pi_i^{emb}$, and states with shorter mean durations have $\pi_i^{occ} < \pi_i^{emb}$. The state "Awareness|Low|LowTrust" has a higher red bar than its teal bar. This suggests the process enters this state relatively often (high visit frequency at transition times) but does not stay long (short mean duration), so its time share is smaller. The state's "Intent|Low|LowTrust" and "Purchase|Low|LowTrust" show teal at least as strong as red, suggesting fewer entries relative to Awareness but longer stays when reached. The state "Consideration|Low|LowTrust" dominates both measures, and teal is slightly larger than red, indicating it is both frequently visited and relatively persistent.

In the retail media context, because we care more about event-driven opportunities, that is, how often users arrive at a state where you can trigger a message, we therefore prefer embedded stationary.



(a) duration



(b) emission

Figure 10 (a) The time-averaged posterior occupancy of each latent HSMM state, computed from the smoothed posterior $\gamma_t(s)$; (b) The mean predicted checkout probability conditioned on a key k, for the specific Checkout (C) event.

Fig. 10(a) reports the time-averaged posterior occupancy of each latent HSMM state, computed from the smoothed posterior $\gamma_t(s)$. The mass is concentrated in a few early and mid-funnel states, especially Awareness and Consideration states with low intention or low trust. This pattern is consistent with the posterior summaries in Figs. 7 and 8. In practice, $\bar{\gamma}(s)$ is best viewed as an

empirical summary of what the fitted model assigns to this dataset. In contrast, the stationary and semi-Markov quantities are model-level summaries of long-run behavior. These average quantities provide an interpretable funnel composition under the latent state taxonomy and enable before-after comparisons by tracking shifts in $\bar{\gamma}(s)$ toward higher trust or later funnel states.

Among the Purchase-stage variants, "Purchase|High|HighTrust" receives the largest fitted occupancy. This is consistent with the model's interpretation that trust and high intention jointly characterize checkout-adjacent states, but it remains a model-based interpretation rather than a psychological measurement.

Fig. 10 (b) summarizes the mean predicted checkout probability conditioned on a key k, for the specific Checkout (C) event. The decomposition

$$P(C \mid k) = \mathbb{E}[P(C \mid S, k)] \approx \sum_{s} \underbrace{\pi_s(k)}_{\substack{\text{state occupancy driven} \\ \text{by transitions and durations}}} \underbrace{P(C \mid S = s, k)}_{\text{checkout emission}}.$$

The values indicate, on average, how likely checkout is under a given key when the HSMM is generating predictive behavior.

If a key has a high bar in Fig. 10, it may be attributed to either emission advantage, i.e., $P(C \mid S = s, k)$ is high in late-stage states (Intent or Purchase) for that key, or trajectory advantage, i.e., the transitions in Fig. 8 imply that users associated with that key spend more probability mass in late-stage states (higher $\pi_s(k)$), so even moderate emissions yield a high average. Fig. 10 indicates the key configurations that monetize those routes best, in the specific sense of producing checkout and ranks, which covariate and message configurations convert within those journeys.

Figures 8 and 10 show that the emission probabilities based on the Informational ad type dominate those based on the Emotional type. Generally, for every covariate (shopper trait and involvement), informational ads increase the model-implied probability of Checkout (C) and also generally increase Intent (I), while reducing Browse (B). This pattern appears both globally, after averaging across states using the embedded stationary distribution, and state by state. Under the one-step Checkout objective, 286 of 288 covariate-state pairs recommend Informational. Only two health-conscious high-involvement purchase states were close enough that Emotional can be defended as a tie-breaker or a niche test. The differences are slight.

According to the covariate types, we suggest strategies in Table 4 to achieve a high conversion rate.

Table 4 Deployable ad-strategy recommendations selected from the fitted policy table. The table reports the covariate cells used in Campaign 3 rather than all 24 possible covariate keys.

Product Involvement	**Shopper Type**	**Ad Strategy**
Low	Chill	Emotional
Low	Thrifty	Informational
Low	Disciplined	Informational
High	Thrifty	Emotional
High	Disciplined	Emotional

The fitted policy table gives managers a default creative rule. Informational content should be the production baseline for most RMN cells, while emotional content should be held for bounded tests in high involvement cells where the model predicts a narrow advantage or an approximate tie. Creative governance should require each emotional cell to contain factual evidence anchors and a preset stop rule, so that brand imagery does not displace the uncertainty reduction needed for checkout.

4.3. A Second Campaign for Directional Re-Validation

The second deployment tests directional stability rather than causal lift. Campaign 1 selected eight high-conversion and seven low-conversion ad configurations for fresh exposure. Campaign 2 then asks whether those 15 selected cells preserve their predicted direction in a new cohort with no member overlap.

Among the 15 tested configurations, 13 maintain the same direction as in Campaign 1, giving cell-level directional agreement of 86.67%. As shown in Table 5, 24,000 individuals were targeted. CTR increased from 14.0% in Campaign 1 to 15.7% in Campaign 2, while the aggregate purchase rate changed from 1.49% to 1.40%. The high-conversion group records a purchase rate of 1.8% (276/15,280), while the low-conversion group records a purchase rate of 0.69% (60/8,720). This supports directional re-validation of the selected cells.

Operational finding 3. Within the campaign study period, we observe that our strategy preserves 13 of 15 previously directional signs in fresh consumers, which supports directional stability in a new deployment cohort.

Table 5 Campaign 2 directional re-validation. A dash denotes a cell not exposed in this wave. The signs in C/B and C/I compare each cell with the Campaign 2 median thresholds, and the consistency column records agreement with the Campaign 1 direction.

Shopper Type	Involv.	Ad type	#	B	E	I	C	C/B	C/I	consis.
Thrifty	High	Informational	-	-	-	-	-	-	-	-
Thrifty	High	Emotional	3720	521	368	71	52	0.10 (+)	0.73 (+)	Yes
Thrifty	Low	Informational	3720	498	209	81	44	0.09 (+)	0.54 (+)	Yes
Thrifty	Low	Emotional	-	-	-	-	-	-	-	-
Chill	High	Informational	1760	119	62	9	6	0.05 (-)	0.67 (+)	No
Chill	High	Emotional	1760	195	47	31	9	0.05 (-)	0.29 (-)	Yes
Chill	Low	Informational	-	-	-	-	-	-	-	-
Chill	Low	Emotional	1760	201	31	18	6	0.03 (-)	0.33 (-)	Yes
Disciplined	High	Informational	-	-	-	-	-	-	-	-
Disciplined	High	Emotional	1440	99	31	14	2	0.02 (-)	0.14 (-)	Yes
Disciplined	Low	Informational	1440	461	269	196	101	0.22 (+)	0.52 (+)	Yes
Disciplined	Low	Emotional	1440	392	176	92	33	0.08 (+)	0.36 (-)	No
Socials	High	Informational	1440	250	91	29	18	0.07 (+)	0.62 (+)	Yes
Socials	High	Emotional	1440	198	64	33	17	0.09 (+)	0.52 (+)	Yes
Socials	Low	Informational	-	-	-	-	-	-	-	-
Socials	Low	Emotional	-	-	-	-	-	-	-	-
Opportunistic	High	Informational	880	147	69	21	5	0.03 (-)	0.24 (-)	Yes
Opportunistic	High	Emotional	-	-	-	-	-	-	-	-
Opportunistic	Low	Informational	880	266	158	39	26	0.10 (-)	0.67 (+)	No
Opportunistic	Low	Emotional	880	142	66	21	12	0.08 (+)	0.57 (+)	Yes
Health	High	Informational	-	-	-	-	-	-	-	-
Health	High	Emotional	720	178	67	24	2	0.01 (-)	0.08 (-)	Yes
Health	Low	Informational	-	-	-	-	-	-	-	-
Health	Low	Emotional	720	96	34	22	3	0.03 (-)	0.14 (-)	Yes
(total)	—	—	24000	3763	1742	701	336	Q2=0.07	Q2=0.52	

In the advertisement experiments, Chill Shoppers who received high-involvement, informational ads raised their "add-to-cart"-to-"purchase" conversion from 0.40 in the first deployment to an

above-threshold 0.67 in the second. Upon analyzing this group's purchase behavior, 64% had previously purchased food-related items, 51% had purchased beverages, and 46% had purchased tobacco or alcohol products.

Table 6 The verification counts reflect final checkout outcomes after the second campaign. As summarized in Table 5, among 336 checkout customers, 244 exhibit increased checkouts and 27 exhibit decreased checkouts, consistent with the strategic expectation. Only 32 and 33 cases, respectively, contradict the expectation. The resulting prediction accuracy is estimated to be 80.65%.

	Expecting +	**Expecting -**
Verified +	244	32
Verified -	33	27

As shown in Table 6, the second campaign obtained accuracy of 80.65% with verified outcomes (precision 88.09%, recall 88.41%), where $(FPR, TPR) = \left(\frac{FP}{FP+TN}, \frac{TP}{TP+FN}\right) = (0.55, 0.8841)$. Therefore, the experiment received support from high directional agreement, with the false-positive rate defining the main operating boundary for future threshold tuning.

Managers can use the model as a ranking screen that moves cells into pilot, scale, or diagnose categories, rather than as a final causal proof. Cells that change sign, such as the inconsistent Chill and Opportunistic cases in Table 5, form an investigation list for assortment fit, creative execution, price, inventory, and frequency before allocating ad budget.

4.4. Ablation Study

We report the comparative evidence in two parts. First, the paired AUC benchmark compares the theory-identified posterior with an action-only covariate-cell representation under the same training horizon and outcome split (the baseline). Second, the ablation suite isolates which components generate the Full-model gain.

The action-only baseline used for the paired AUC comparison is a reproducible covariate-cell purchase predictor, $\widehat{P}(C \mid s, \ell, a)$, where s is shopper trait, ℓ is involvement level, and a is ad type. It uses the same training horizon, the same outcome variable, and the same 24 covariate keys as the intention-identified predictor, but it contains no latent-state posterior, no retrieval-augmented context, and no duration model. A richer lagged-event logistic baseline would require the full per-event timeline with stable session identifiers and dwell-time fields. It is not reported from the aggregate campaign export.

The intention-identified predictor uses the smoothed posterior $\gamma_t(i)$ over the identified latent states from Section 3. The paired comparison therefore tests whether replacing a 24-cell action-only representation by a posterior over theory-identified states improves one-step checkout prediction under the same outcome and evaluation split.

The reported $\Delta\text{AUC} = 0.082$ is a one-step, leakage-free, paired holdout comparison that establishes the predictive value of substituting theory-identified posteriors for an action-only covariate-cell representation. Causal lift from a deployed baseline policy is outside the Campaign 3 design because no simultaneous action-only control arm was run. Opt-out and retention outcomes are also outside the released export.

The Full-vs-baseline contrast establishes the predictive gap. The λ_{T} sweep in Fig. 5(b) is, by construction, an LLM-influence ablation. At $\lambda_{\mathrm{T}} = 0$ the inference ignores the language model entirely, and the curve traces empirical degradation as the LLM's influence grows beyond a small soft-prior weight. We complement that sweep with the five controlled ablations T1–T5 reported in Table 7.

T1–T4 each disable exactly one of the four Full-model mechanisms, namely the KL identifying constraint, the retrieved tacit context, the semi-Markov sequence structure, and the theory-coherent state coordinates. T5 substitutes a capacity-only $K = 64$ HMM as a contrasting model class. All five are re-evaluated on the same training horizon and outcome variable as the Full model.

The closest published sequence-modeling baseline in our ablation suite is the scaled hidden Markov language model (HSLM) of Chiu and Rush (2020), which provides a high-capacity sequential baseline trained without any theory-coherent regularization. The implied prediction of our identification argument is that increasing latent-state capacity alone (HSLM with comparable or larger K) does not close the AUC gap our framework opens over the action-only baseline, because the gap is generated by the identifying constraint, not by capacity. Chiu and Rush (2020) use HSLM as a high-capacity benchmark: blocked emissions (Eq. 3 of their paper, instantiated here with a funnel-progression partition {B,E} vs {I,C} because $|\mathcal{Y}| = 4$ makes Brown clustering degenerate), state dropout at $\lambda = 0.5$ (Eq. 5 of their paper, disabled at evaluation), and exact $O(Tk^2)$ inference with $k = K/M$ via the block constraint. The single adaptation is parameterization: we use the scalar parameterization explicitly considered in Chiu & Rush's Table 2 (row "− neural param", which they report reaches similar training perplexity to the neural variant). The realized HSLM AUC on the common $n = 1{,}008$ paired hold-out is 0.602 (Table 7), 3.4 percentage points below the action-only baseline ($Z = -0.81$, $p = 0.42$). The high-capacity HSLM thus does not close the AUC gap our framework opens; the realized number is consistent with prediction (T5) that latent-state capacity alone does not produce identification gains.

- (T1) Set $\lambda_{\mathrm{T}} = 0$ and keep everything else, including the theory state space, RAG, and semi-Markov structure. This test isolates the KL identifying constraint. The expected pattern is a material AUC loss relative to Full, with performance still above the action-only baseline because the other Full-model mechanisms remain active.
- (T2) Drop the retrieved context r_t from all four kernels. This test isolates retrieval as a context-shaping component. The expected pattern is lower AUC and policy-table movement in covariate keys whose posterior state distributions are retrieval-sensitive, especially high-reactance PKM keys.
- (T3) Feed the LLM's per-key posterior to a non-sequential classifier. This test removes transition and duration structure. The expected pattern is a clear failure to match Full on paired AUC and no support for hitting-time analysis, since the model has no transition or duration kernels.
- (T4) Fit an HSMM with $K = 24$ statistical states under the same kernels but without TPB/CD-J/PKM coordinates. This test isolates theory-coherent state semantics. The expected pattern is failure to reproduce the Full-model AUC gain and failure to recover the same deployment policy structure despite using the same state cardinality.
- (T5) Run the scaled HSLM of Chiu and Rush (2020) with $K \geq 64$ under the same training horizon. This test isolates latent-state capacity. The expected pattern is that higher capacity alone remains below Full because the empirical gain comes from the identifying constraint, not from a larger state space.

These mechanism tests are specified before reporting the realized outcomes. Table 7 reports realized results for all five ablations T1–T5. T3 is implemented as the LLM teacher's per-key marginal $\sum_i \pi_x(i)\, B_x(i, C)$, which removes the sequence and duration structure while keeping the LLM teacher's covariate-conditional state belief; the per-row predictor depends only on the covariate key and not on y_t, avoiding the leakage that an earlier $\pi_x \cdot B_x[\cdot, y_t]$ formulation produced.

T1 vs. Full removes the KL identifying constraint and leaves the theory state space, retrieved context, and duration-aware kernels in place. T1 remains above baseline, but its AUC is 4.4 pp below Full. The KL constraint therefore carries a material share of the realized gain. Because T1 still retains the other three mechanisms, the remaining positive AUC confirms that the Full model

Table 7 Realized ablation results, DeLong paired test on $n = 1{,}008$ common test consumers (the action-only baseline's stratified hold-out). ΔAUC is the ablation's AUC minus the baseline's. The p-value is the DeLong two-sided p against the baseline. T1, T2, T4, and T5 each evaluate the same one-step-ahead next-click predictor $\hat{P}(y_{t+1} = C \mid y_{1:t}, x_{1:t+1}) \propto \sum_i \gamma_{t-1}(i) \sum_j A_{ij} B_j(C \mid \text{key}_{t+1})$ to remove smoothed-posterior leakage (details in Appendix A.2). T3 is intentionally the non-sequential exception: its predictor is the LLM teacher's per-key marginal $\sum_i \pi_x(i) B_x(i, C)$, with no transition kernel and no time index, because the ablation's purpose is to remove the semi-Markov sequence structure entirely.

Model	K	AUC	ΔAUC	Z	p-value
Action-only baseline	—	0.635	—	—	—
Full (theory-coherent identification)	24	**0.717**	**+0.082**	**5.64**	$\mathbf{1.7 \times 10^{-8}}$
T1 ($\lambda_T = 0$; no KL regularizer)	24	0.673	+0.037	5.57	2.5×10^{-8}
T2 (no retrieved context; empty r_t)	24	0.700	+0.065	6.53	6.7×10^{-11}
T3 (flat LLM marginal; no sequence structure)	24	0.508	−0.128	−2.93	3.3×10^{-3}
T4 (K=24 data-driven HSMM; no theory)	24	0.633	−0.002	−0.56	0.58
T5 (Chiu & Rush 2020 HSLM; no theory, no RAG)	64	0.602	−0.034	−0.81	0.42

is a coupled architecture. The KL term is necessary for the best result, and the retained theory state space, retrieval context, and semi-Markov kernels preserve part of the predictive signal.

T2 vs. Full loses 1.7 pp (T2 = +6.5 pp, Full = +8.2 pp; T2 still beats baseline at $Z = 6.53$, $p < 10^{-10}$). Removing the retrieved context costs a measurable but smaller share of the gain than removing the KL regularizer (T1) or the theory state space (T4 below), consistent with the kernel-shape interpretation in prediction (T2) that RAG enters as a structural shift on top of the identifying constraint. The aggregate harness reports the AUC effect. Segment-decomposed AUC is the appropriate next diagnostic for high-reactance PKM keys.

T3 vs. baseline is significantly *below* the action-only baseline (T3 AUC = 0.508, $\Delta = -0.128$; $Z = -2.93$, $p = 0.0033$). This is the strongest ablation result for the semi-Markov component. The flat LLM marginal $\sum_i \pi_x(i) B_x(i, C)$ uses the teacher's belief about $P(\text{checkout} \mid \text{covariate})$ with no time, no transitions, and no duration. It is poorly calibrated to this dataset, with a per-key range from 0.08 to 0.13, whereas the empirical baseline LR fits per-key rates from the training set directly. The result shows that the LLM teacher is not a standalone predictor. It functions as a prior that the EM/HSMM inference reshapes into an identifying distribution. Removing sequence and duration structure therefore eliminates the key advantage of the Full model and produces worse prediction than an empirical action-only baseline.

T4 vs. baseline is statistically indistinguishable from the action-only baseline (T4 AUC = 0.633, baseline = 0.635; $Z = -0.56$, $p = 0.58$). A data-driven HSMM at the *same* cardinality $K = 24$ as the full model, but without the theory-coherent state coordinates, fails to recover the identification gain. This directly confirms prediction (T4): the gain comes from the theory-coherent reference distribution, not from the cardinality of the state space.

T5 vs. baseline is below the action-only baseline and statistically indistinguishable from it (T5 AUC = 0.602, baseline = 0.635; $p = 0.42$). A high-capacity $K = 64$ HMM with blocked emissions and state dropout, trained without any theory-coherent reference distribution and without retrieved context, fails to close the gap that the full model opens. This directly confirms prediction (T5): *capacity alone is insufficient*; the identifying constraint is the source of the empirical gain.

Taken together, the Full model is the top-performing method in Table 7. The realized pattern is T3 $\ll$ baseline, T5 $<$ baseline, T4 $\approx$ baseline, T1 and T2 both less than the Full but both greater than the baseline. The ordering supports the integral identification argument across all four levers. Removing the theory-coherent state space collapses prediction to baseline (T4); removing the semi-Markov sequence and duration structure drops prediction below baseline (T3); pure latent-state capacity does not substitute for either (T5); and the KL constraint and the retrieved context each

carry quantifiable shares of the remaining gain (T1, T2). No single mechanism on its own reproduces the Full model, and the two most aggressive removals (T3, T5) actively degrade beneath the empirical action-only baseline. The paired holdout results support the one-step prediction improvement claimed by the model comparison.

Operational finding 4. Within the campaign study period, posteriors over the theory-identified latent states predict the next checkout event more accurately than the action-only covariate-cell baseline in the paired holdout comparison.

The expected hitting-time calculation is reported as a model diagnostic. Let "Purchase" be absorbing and let P_0 and P_1 denote baseline-policy and intention-policy transition matrices. When both matrices are available, the diagnostic reduction is

$$\frac{\mathbb{E}_{P_0}[\tau] - \mathbb{E}_{P_1}[\tau]}{\mathbb{E}_{P_0}[\tau]}.$$

The present export supports Campaign-2 directional agreement in Tables 5–6 and the Campaign-3 temporal before/after audit in Table 8 as headline deployment evidence. Segment-level conversion lift, opt-out reduction, and 60-day retention lift require baseline-policy transitions, opt-out flags, and retention indicators, so those metrics are reserved for deployments that collect the corresponding telemetry.

The predictive gain justifies using latent state posteriors as operational decision variables in bid allocation, frequency capping, trigger timing, and creative sequencing. Budget decisions can be conditioned on posterior shifts that forecast checkout. Because the result is predictive rather than causal, managers should still reserve a simultaneous holdout when claiming incremental lift, opt out reduction, or retention improvement.

4.5. A Third Campaign for Temporal Deployment Re-Validation

The third deployment evaluates the recommended policy over a later temporal window. The deployed policy produced higher clicks, page views, cart additions, and completed purchases than Campaign 1. Completed purchases rose from 358 to 512, an increase of 154 transactions and 43% temporal growth.

The segment patterns should be read as operational diagnostics. Thrifty shoppers retained strong cart-to-purchase ratios under high-involvement emotional and low-involvement informational exposure, although click-to-purchase conversion declined slightly. Chill shoppers improved from 19 to 30 cart-to-purchase conversions, while their overall purchase volume remained modest. Disciplined shoppers were targeted only with low-involvement informational ads and generated 134 more completed transactions than in Campaign 1. Social and Opportunistic shoppers show stable cart-to-purchase ratios in the exposed cells. Health-Conscious shoppers were exposed to health-aligned products such as chicken breast, yogurt drinks, and salad greens, and completed purchases increased by 29. These patterns support the operational usefulness of the policy table under the observed deployment.

Across the three ad testing launches, Health-Conscious Shoppers demonstrated subpar engagement, including ad clicks, browsing, and purchasing behavior. However, pure transactional analysis revealed that 85% had previously purchased food-related products, and 81% had purchased beverage-related products. In the third advertising campaign, according to our method, ads were specially curated to assess whether relevance to their identified preferences could improve ad engagement and effectiveness.

Retailers should manage RMN targeting as a state-aware allocation problem. First-party data can help match shopper type, product involvement, ad type, and inferred journey state. Managers should shift budget toward ad cells with stronger cart-to-purchase response instead of relying only on aggregate campaign metrics. This design converts targeting from an aggregate end-of-campaign

Table 8 Campaign 3 temporal deployment audit under the recommended policy. The table reports event counts and ratios for the exposed cells, which is a before-after audit relative to Campaign 1. Differing from other statistics methods, our method is effective in capturing shifting consumer purchase intension.

Shopper Type	Involv.	Ad type	#	B	E	I	C	C/B	C/I	improv.
Thrifty	High	Informational	-	-	-	-	-	-	-	-
Thrifty	High	Emotional	3720	644	425	68	48	0.07	0.71	Yes
Thrifty	Low	Informational	3720	506	199	65	41	0.08	0.63	Yes
Thrifty	Low	Emotional	-	-	-	-	-	-	-	-
Chill	High	Informational	1320	86	23	17	9	0.10	0.53	Yes
Chill	High	Emotional	1320	101	37	21	11	0.11	0.52	Yes
Chill	Low	Informational	1320	91	68	31	12	0.13	0.39	Yes
Chill	Low	Emotional	1320	178	65	29	17	0.10	0.59	Yes
Disciplined	High	Informational	-	-	-	-	-	-	-	-
Disciplined	High	Emotional	-	-	-	-	-	-	-	-
Disciplined	Low	Informational	4320	1201	683	412	262	0.22	0.64	Yes
Disciplined	Low	Emotional	-	-	-	-	-	-	-	-
Socials	High	Informational	1440	211	76	22	14	0.07	0.64	Yes
Socials	High	Emotional	1440	276	66	31	18	0.07	0.58	Yes
Socials	Low	Informational	-	-	-	-	-	-	-	-
Socials	Low	Emotional	-	-	-	-	-	-	-	-
Opportunistic	High	Informational	-	-	-	-	-	-	-	-
Opportunistic	High	Emotional	-	-	-	-	-	-	-	-
Opportunistic	Low	Informational	1320	243	56	44	26	0.11	0.59	Yes
Opportunistic	Low	Emotional	1320	118	41	25	16	0.14	0.64	-
Health	High	Informational	360	78	55	41	15	0.19	0.37	Yes
Health	High	Emotional	360	54	21	16	9	0.17	0.56	Yes
Health	Low	Informational	360	87	44	16	8	0.09	0.50	Yes
Health	Low	Emotional	360	71	54	31	6	0.08	0.19	Yes
(total)	—	—	24000	3945	1913	869	512			

evaluation problem into a deployment-auditable prediction problem, where each ad strategy has a directional response implication that can be checked in future campaigns.

The deployment evidence shows consistent response across covariate cells. High-involvement products should take emotional exposure for Thrifty and Social Shoppers, while low-involvement products should take informational exposure for Thrifty and Disciplined Shoppers.

Campaign 3 increased completed purchases from 358 to 512, a 43% temporal increase over Campaign 1. This result should be interpreted as a deployment evidence of the selected policy, not as causal lift against a simultaneous control arm. Overall, the evidence supports joint selection of ad type and product involvement according to the inferred CDJ stage and the PKM trust constraint. The cart-to-purchase ratio then functions as an operational diagnostic for policy selection and budget allocation within the observed RMN environment.

Operational finding 5. Within the campaign study period, the PKM coordinate functions as a policy constraint when it indicates high persuasion resistance, and later deployments can measure consumer-side trust and welfare outcomes directly.

This study emphasizes advertising content design and constructs a behavioral-state model from a ever-changing clickstream data. Marketers can use our method to facilitate closed-loop advertising strategies that track events from exposure to purchase.

For managers, the operational value is a repeatable decision cycle. Infer the state, choose the creative policy, deploy the selected cell, and update thresholds from observed checkout and trust diagnostics. The PKM coordinate should be implemented as a guardrail on persuasion intensity. When high resistance or low trust is inferred, the RMN should cap retargeting frequency, prefer transparent informational content, expose value and privacy cues, and route the cell to trust preserving measurement such as opt out, complaint, unsubscribe, and repeat purchase indicators.

5. Conclusions

This paper makes a unique contribution. It formulates purchase-intention inference from time-dependent RMN clickstreams as operational identification of theory-coherent semi-Markov dynamics. The resulting latent states obtain empirical content through closed-loop directional validation and provide a deployable state representation for ad-response prediction.

Three components are inseparable in this contribution. First, the retrieval-augmented language model produces a theory-coherent reference distribution under TPB, CDJ, and PKM coordinates. Second, the covariate-dependent HSMM turns that reference into a generative, duration-aware dynamics over click events. Third, the RMN deployment tests the directional implications of the inferred state distribution. Removing any one of these components changes the object being validated.

The strongest reproducible evidence is the one-step paired AUC comparison. The Full model achieves AUC 0.717 against an action-only covariate baseline of 0.635 on 1,008 paired test consumers, with DeLong $Z = 5.64$ and $p = 1.7 \times 10^{-8}$. Campaign 2 provides directional re-validation, preserving 13 of 15 predicted cell directions and yielding 80.65% verified checkout-direction agreement. Campaign 3 provides a temporal before/after audit of the selected policy, increasing checkouts from 358 to 512 relative to Campaign 1.

The ablations establish the component logic of the method. Full is the best-performing model. Removing the KL identifying constraint and removing retrieved context both reduce performance but remain above the action-only baseline. Removing the theory-coherent state space collapses performance to the baseline. Removing the semi-Markov sequence and duration structure drops performance below the baseline. A capacity-only $K = 64$ HMM also fails to close the gap. Thus the empirical pattern supports $\mathcal{S}$. Deployable predictive value comes from the coupled theory-identified semi-Markov dynamics, not from a detachable LLM teacher or from larger latent-state cardinality alone.

The results show that segment-level conversion lift, opt-out reduction, and 60-day retention lift can be retained in future deployments under a drifting environment. The same reproducibility standard can then be extended from the one-step AUC benchmark to hitting-time reduction, trust-preservation effects, and horizon-specific AUC.

References

Accenture Interactive (2017) U.S. Consumers Turn Off Personal Data Tap as Companies Struggle to Deliver the Experiences They Crave. https:/newsroom.accenture.com/news/2017/us-consumers-turn-off-personal-data-tap-as-companies-struggle.

Agius N (2024) Walmart connect opens up to small and non-endemic advertisers. https:/searchengineland.com/walmart-connect-small-non-endemic-advertisers-438913.

Ajzen I (2020) The theory of planned behavior: Frequently asked questions. *Human Behavior and Emerging Technologies* 2(4):314–324, URL http://dx.doi.org/10.1002/hbe2.211, a critical update from the founder of TPB addressing modern digital contexts.

Batat W (2022) The customer decision journey (cdj) revised in the digital era: The role of phygital experiences and social media. *Journal of Retailing and Consumer Services* 65:102824, URL http://dx.doi.org/10.1016/j.jretconser.2021.102824.

Bishop E (2025) Brands threw away $6 billion in wasted digital ad spend. https:/www.nextandco.com.au/the-next/brands-threw-away-6-billion-in-wasted-digital-ad-spend.

Bruce NI, Becker M, Reinartz W (2020) Communicating brands in television advertising. *Journal of Marketing Research* 57(2):236–256.

Bureau IA (2020) IAB predicts a 19% to 25% drop in digital ad revenues for the near-term. https:/yougov.com/articles/29085-iab-predicts-19-25-drop-digital-ad-revenues-near-t.

Chen J, Guo Z (2022) New-media advertising and retail platform openness. *MIS Quarterly* 46(1):431–456.

Chiu JT, Rush AM (2020) Scaling hidden Markov language models. *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1341–1349.

Cramer-Flood E (2024) Worldwide retail media ad spending forecast 2024. emarketer. URL https:/www.emarketer.com/content/worldwide-retail-media-ad-spending-forecast-2024.

Filippou G, Georgiadis AG, Jha AK (2025) Unveiling digital dynamics: Do digital media investments impact organic branded searches? *Journal of Marketing Analytics* 1–13.

Forte T (2022) *Building a second brain: A proven method to organize your digital life and unlock your creative potential* (Simon and Schuster).

Friestad M, Wright P (1994) The persuasion knowledge model: How people cope with persuasion attempts. *Journal of consumer research* 21(1):1–31.

Gu W, Chan KW, Kwon J, Dhaoui C, Septianto F (2023) Informational vs. emotional b2b firm-generated-content on social media engagement: Computerized visual and textual content analysis. *Industrial Marketing Management* 112:98–112.

Guitart IA, Stremersch S (2021) The impact of informational and emotional television ad content on online search and sales. *Journal of Marketing Research* 58(2):299–320.

Hamby A, Jones N (2022) The effect of affect: An appraisal theory perspective on emotional engagement in narrative persuasion. *Journal of Advertising* 51(1):116–131.

Hartnett N, Kennedy R, Sharp B, Greenacre L (2016) Creative that sells: How advertising execution affects sales. *Journal of Advertising* 45(1):102–112.

Kanchanapoom K, Chongwatpol J (2023) Integrated customer lifetime value (clv) and customer migration model to improve customer segmentation. *Journal of Marketing Analytics* 11(2):172–185.

Kaya T (2024) Personal knowledge management in the age of generative ai: A systematic review and future research agenda. *Computers in Human Behavior* 151:108012, the most current bridge between PKM and AI-driven decision making.

Moriyama T, Kuwano M, Nakayama M (2024) A statistical method for estimating piecewise linear sales trends. *Journal of Marketing Analytics* 12(2):436–444.

North M (2023) Personal knowledge management (pkm): A review of literature and a proposed framework for the digital age. *Information and Organization* 33(2).

Qin V, Pauwels K, Zhou B (2024) Data-driven budget allocation of retail media by ad product, funnel metric, and brand size. *Journal of Marketing Analytics* 12(2):235–249.

Quester P, Lin Lim A (2003) Product involvement/brand loyalty: is there a link? *Journal of product & brand management* 12(1):22–38.

Statista (2024) Carrefour retail media ad revenue worldwide 2021-2025, statista research department. retrieved nov 19. URL https:/www.statista.com/statistics/1358357/carrefour-retail-media-ad-revenue/#statisticContainer.

Tsai YL, Honka E (2021) Informational and noninformational advertising content. *Marketing Science* 40(6):1030–1058.

Vorderer P, Klimmt C (2023) Explaining the intentions to use ai-based personal assistants: An extension of the theory of planned behavior. *Journal of Media Psychology* Integrates TPB with modern AI-driven knowledge tools.

A. Appendix

A.1. Detailed RAG Factorization for the Theory-Coherent Reference Distribution

This appendix supplies the detailed derivation that supports equation (3) in the main text.

Knowledge base and embeddings

We posit that, beyond the explicit context $c_t = (\mathbf{x}_i, y_{1:t})$, there exists a set of tacit knowledge items that help infer the latent state z_t. These items are stored in the knowledge base $\mathcal{K} = \{d_m : m = 1, \ldots, M\}$, where each d_m is a document or snippet of expert or empirical knowledge—theory-based patterns, historical clickstream templates, behavioral rules, or consumer reports. Each item is embedded into a vector space via an encoder

$$\mathbf{k}_m = \zeta_\phi(d_m) \in \mathbb{R}^H, \qquad m = 1, \ldots, M,$$

with encoder parameters ϕ (e.g., a sentence-embedding model). We introduce a discrete retrieval variable $D_t \in \{1, \ldots, M\}$ indicating the knowledge item consulted at time t, and a continuous representation $r_t \in \mathbb{R}^R$ summarizing the retrieved tacit knowledge.

Marginalization over the retrieved item

The teacher posterior of interest is $p(z_t \mid c_t)$. We introduce r_t as an unobserved mediator:

$$p(z_t \mid c_t) = \int p(z_t \mid c_t, r_t)\, p(r_t \mid c_t)\, dr_t. \tag{9}$$

Tacit knowledge is stored in discrete items $d_m \in \mathcal{K}$, and r_t is induced by selecting one of these items, so

$$p(r_t \mid c_t) = \sum_{m=1}^{M} p(r_t \mid D_t = m, c_t)\, p(D_t = m \mid c_t). \tag{10}$$

Defining $r_t = \psi(d_{D_t})$ as a deterministic function of the chosen document, $p(r_t \mid D_t = m, c_t)$ becomes a point mass at $\psi(d_m)$, and substituting (10) into (9) yields equation (3) of the main text.

Retrieval distribution

We interpret $p(D_t = m \mid c_t)$ as an approximate posterior over which tacit knowledge item is most relevant. A query encoder maps the context to an embedding $\mathbf{q}_t = f_\theta(c_t) \in \mathbb{R}^H$. With a similarity score $s_t(m) := \mathrm{sim}(\mathbf{q}_t, \mathbf{k}_m)$ and a temperature $\tau > 0$,

$$p(D_t = m \mid c_t) = \frac{\exp(\tau^{-1} s_t(m))}{\sum_{m'=1}^{M} \exp(\tau^{-1} s_t(m'))}. \tag{11}$$

The tacit knowledge vector r_t is then the expectation

$$r_t = \mathbb{E}_{D_t \sim p(\cdot \mid c_t)}[\psi(d_{D_t})] = \sum_{m=1}^{M} p(D_t = m \mid c_t)\, \psi(d_m). \tag{12}$$

In practice we approximate (12) using a top-K_R subset $\mathcal{N}_t \subset \{1, \ldots, M\}$:

$$p(D_t = m \mid c_t) \approx \frac{\exp(\tau^{-1} s_t(m))}{\sum_{m' \in \mathcal{N}_t} \exp(\tau^{-1} s_t(m'))}\, \mathbf{1}\{m \in \mathcal{N}_t\}.$$

This is the standard retrieval-augmented generation step: the context is encoded into $\mathbf{q}_t$, similarities to knowledge embeddings $\mathbf{k}_m$ are computed, and a probability distribution over documents is obtained via softmax.

Teacher posterior under retrieval

Under retrieval, the teacher conditions on r_t explicitly: $q_{\mathrm{T}}(z_t = k \mid c_t, r_t) \approx p(z_t = k \mid c_t, r_t)$, and the resulting reference distribution is the equation (7) of the main text. Operationally, at time t, we (a) compute $\mathbf{q}_t$, (b) retrieve $\mathcal{N}_t$ via (11), (c) construct r_t via (12), (d) prompt the LLM teacher with c_t, the latent-state definitions (TPB/CDJ/PKM semantics), and $\{d_m\}_{m\in\mathcal{N}_t}$, and (e) parse the teacher output as the probability vector $\gamma_t^{\mathrm{T}}(\cdot)$ approximating (7).

A.2. Forward-Backward Recursion for the Covariate-Dependent HSMM

This appendix presents the full forward-backward recursion for the covariate- and retrieval-conditioned HSMM, the EM auxiliary function $Q(\theta;\theta^{old})$, the gradient form for the softmax heads, and the PKM-consistent regularizer. The main text uses these results to claim that EM updates are well-posed and monotone (Theorem 4); the recursion is standard for HSMMs but the covariate- and retrieval-conditioning is what our identification argument requires.

Define the session emission block-likelihood for state $k \in \{1,\dots,K\}$ spanning times $t-d+1,\dots,t$:

$$\ell_i(t,d) = \prod_{s=t-d+1}^{t} b_{\theta,i}\big(y_s \mid x_s, r_s\big).$$

To define forward variables, let $\alpha_t(i)$ be the probability of the prefix $y_{1:t}$ with a session ending at time t in state i:

$$\alpha_t(i) = \mathbb{P}_\theta\big(y_{1:t},\ \text{a session ends at } t \text{ with state } i \mid x_{1:t}, r_{1:t}\big).$$

Initialization and recursion are:

$$\alpha_t(i) = \sum_{d=1}^{\min(D_{\max},t)} \left[\ell_i(t,d)\ g_{\theta,i}(d \mid x_{t-d+1}, r_{t-d+1}) \left(\mathbf{1}_{t=d}\,\pi_\theta(i \mid x_1, r_1) + \mathbf{1}_{t>d} \sum_{j=1}^{K} \alpha_{t-d}(j)\ a_{\theta,ji}(x_{t-d+1}, r_{t-d+1})\right)\right]. \tag{13}$$

The likelihood is $\mathcal{L}(\theta) = \sum_{i=1}^{K} \alpha_T(i)$.

To define backward variables, let $\beta_t(i)$ denote the probability of the suffix $y_{t+1:T}$ given that a session ended at t in state i:

$$\beta_t(i) = \mathbb{P}_\theta\big(y_{t+1:T} \mid \text{a session ends at } t \text{ with state } i,\ x_{1:T}, r_{1:T}\big),$$

with $\beta_T(i) = 1$. The recursion is:

$$\beta_t(i) = \sum_{j=1}^{K} a_{\theta,ij}(x_{t+1}, r_{t+1}) \sum_{d=1}^{\min(D_{\max},T-t)} g_{\theta,j}(d \mid x_{t+1}, r_{t+1})\ \ell_j(t+d,d)\ \beta_{t+d}(j). \tag{14}$$

Define the posterior probability that a session of state i and duration d ends at time t:

$$\begin{aligned}\eta_t(i,d) &= \mathbb{P}_\theta\big(\text{a session ends at } t \text{ with state } i \text{ and duration } d \mid y_{1:T}, x_{1:T}, r_{1:T}\big)\\ &\propto \ell_i(t,d)\ g_{\theta,i}(d \mid x_{t-d+1}, r_{t-d+1}) \left(\mathbf{1}_{t=d}\,\pi_\theta(i \mid x_1, r_1) + \mathbf{1}_{t>d} \sum_{j=1}^{K} \alpha_{t-d}(j)\ a_{\theta,ji}(x_{t-d+1}, r_{t-d+1})\right) \beta_t(i),\end{aligned} \tag{15}$$

normalized so that $\sum_i \sum_d \eta_t(i,d) = \mathbb{P}(\text{session ends at } t \mid \cdot) = 1$ at each t.

The posterior occupancy (state at time s) is recovered by summing over all sessions covering s:

$$\gamma_s(i) = \mathbb{P}_\theta(z_s = i \mid y_{1:T}, x_{1:T}, r_{1:T}) = \sum_{t=s}^{\min(T, s+D_{\max}-1)} \sum_{d=t-s+1}^{\min(D_{\max}, t)} \eta_t(i,d)\ \mathbf{1}\{s \in [t-d+1, t]\}. \tag{16}$$

Posterior expected transition counts at boundaries are

$$\begin{aligned}\xi_t(i,j) &= \mathbb{P}_\theta\big(\text{a session ends at } t \text{ in } i \text{ and next session starts at } t+1 \text{ in } j \mid y_{1:T}, x_{1:T}, r_{1:T}\big) \\ &\propto \alpha_t(i)\ a_{\theta,ij}(x_{t+1}, r_{t+1}) \sum_{d=1}^{\min(D_{\max}, T-t)} g_{\theta,j}(d \mid x_{t+1}, r_{t+1})\ \ell_j(t+d, d)\ \beta_{t+d}(j).\end{aligned} \tag{17}$$

To evaluate the HSMM one-step predictive distributions, let the emission probabilities be

$$b_i(y \mid X_{t+1}) \equiv (Y_{t+1} = y \mid Z_{t+1} = i, X_{t+1}), \qquad y \in \mathcal{Y},$$

and let the transition probabilities at session boundaries be

$$a_{ij}(X_{t+1}) \equiv (Z_{t+1} = j \mid \text{session ended at } t, Z_t = i, X_{t+1}).$$

Conditioning on $Z_t = i$, the next state Z_{t+1} equals i if the session continues, and transitions via $A(X_{t+1})$ if the session ends. The one-step predictive distribution of the next click is the mixture

$$p_i^{(t)}(y) \equiv (Y_{t+1} = y \mid Z_t = i, Y_{1:T}, X_{1:T}) = (1 - e_t(i))\, b_i(y \mid X_{t+1}) + e_t(i) \sum_{j=1}^{K} a_{ij}(X_{t+1})\, b_j(y \mid X_{t+1}), \tag{18}$$

where $e_t(i) \equiv (\text{segment ends at time } t \mid z_t = i, y_{1:T}, x_{1:T})$ is the conditional segment-termination probability, obtained from end-of-segment responsibilities aggregated over durations. Marginalizing over the posterior γ_t,

$$p^{(t)}(y) \equiv (Y_{t+1} = y \mid Y_{1:T}, X_{1:T}) = \sum_{i=1}^{K} \gamma_t(i)\, p_i^{(t)}(y). \tag{19}$$

In equation (5) of the main text, each term is a posterior-weighted log-probability under the LLM-parameterized kernels; maximizing Q is equivalent to fitting the LLM heads to the soft labels (posterior responsibilities). For emissions at time s, the contribution is the cross-entropy

$$\sum_{i=1}^{K} \gamma_s(i)\ \log b_{\theta,i}(y_s \mid x_s, r_s).$$

Fixing a time s and state i, with $\boldsymbol{p}_{i,s}(\cdot)$ the softmax emission distribution and $\boldsymbol{e}_{y_s}$ the one-hot vector of the observed symbol, the gradient of the emission term with respect to the logits is

$$\frac{\partial}{\partial \text{logit}} \Big(\gamma_s(i) \log p_{i,s}(y_s)\Big) = \gamma_s(i)\big(\boldsymbol{e}_{y_s} - \boldsymbol{p}_{i,s}\big),$$

and similarly for π, each row of A, and each row of G. Standard optimizers (SGD, Adam) suffice to increase Q.

A PKM-consistent regularizer protects trust by penalizing aggressive Checkout pushes when the latent trust component is low:

$$\mathcal{R}(\theta) = \lambda \sum_{t=1}^{T} \mathbb{E}\Big[\mathbf{1}\{\tau_t = \text{low trust}\}\ \log \frac{b_{\theta,z_t}(C \mid x_t, r_t)}{b_{\theta,z_t}(B \mid x_t, r_t)} \ \Big|\ y, x, r\Big]_+,$$

so the full objective is $\max_\theta\ Q(\theta; \theta^{old}) - \mathcal{R}(\theta)$.

A.3. Coupled Identification Issues Between the Reference Distribution and the Semi-Markov Inference

This appendix expands the seven-problem decomposition referenced in the main-text subsection on the theory-coherent reference distribution. The main text argues that the reference distribution and the semi-Markov inference are mutually defining; here we go through the seven specific structural problems that this mutual dependence addresses, in the order they appear in the inference pipeline.

(1) HSMM under-identification A pure HSMM estimated only from click sequences $\mathbf{y}_i = (y_1, \ldots, y_{T_i})$ and covariates $\mathbf{x}_i$ faces severe under-identification. Different configurations of initial intensities, transition probabilities, duration distributions, and emission probabilities can produce similar distributions over click sequences, and without additional structure the latent states z_t drift into purely statistical clusters with no clear interpretation in TPB/CDJ/PKM terms. Formally, let Θ and Θ' be two different parameter sets with

$$p_{\mathrm{HSMM}}(\mathbf{y}_i \mid \mathbf{x}_i, \Theta) \approx p_{\mathrm{HSMM}}(\mathbf{y}_i \mid \mathbf{x}_i, \Theta') \quad \forall i,$$

but

$$p_{\mathrm{HSMM}}(z_t \mid \mathbf{y}_i, \mathbf{x}_i, \Theta) \neq p_{\mathrm{HSMM}}(z_t \mid \mathbf{y}_i, \mathbf{x}_i, \Theta').$$

The data likelihood alone cannot distinguish Θ from Θ', so the latent states lose behavioral meaning. The reference distribution $\gamma_t^{\mathrm{T}}(k)$, mixed with γ_t^{S} in $\tilde{\gamma}_t$, implicitly adds the regularizer $-\lambda_{\mathrm{T}} \sum_{i,t} \mathrm{KL}(\gamma_t^{\mathrm{T}} \| \gamma_t^{\mathrm{S}})$ to the EM objective, which selects parameterizations whose posteriors are theory-coherent. Accordingly, the regularizer is an operational identifying constraint. It selects among observationally compatible parameterizations when $\lambda_{\mathrm{T}} > 0$, and the choice of λ_{T} is audited empirically in the Analysis section.

(2) LLM non-generativity A pure LLM approach (even with retrieval augmentation) can in principle infer $q_{\mathrm{T}}(z_t = k \mid c_t)$ for every (i, t), but it does not provide an explicit generative model $p(\mathbf{y}_i \mid \mathbf{x}_i)$ or duration distributions $p(d \mid z)$ required for forecasting and simulation. Each LLM call is a conditional judgment $c_t \mapsto \gamma_t^{\mathrm{T}}$ without a global likelihood. The semi-Markov inference supplies a generative model with explicit transition $p(z_t \mid z_{t-1})$, duration $p(d \mid z)$, and emission $p(y_t \mid z_t)$ kernels. Once fitted, the inference can be evaluated without further LLM queries to simulate click trajectories, compute long-run limiting distributions, and evaluate ad strategies.

(3) Textual theory and tacit knowledge resist hard-coding TPB, CDJ, PKM, and domain rules are textual; encoding them directly into transition matrices and duration distributions would be brittle and incomplete. Instead, the language model produces theory-coherent soft posteriors $\gamma_t^{\mathrm{T}}(k)$, and the semi-Markov inference adjusts its parameters to approximate them in a manner compatible with the observed data. Operationally, EM uses $\tilde{\gamma}_t$ as effective state weights, with expected sufficient statistics $\mathbb{E}[\mathbf{1}\{z_t = k\} \mid \mathbf{y}, \mathbf{x}] \approx \tilde{\gamma}_t(k)$. Theory and data are blended without committing to any single hard constraint.

(4) Noisy student posteriors on short clickstreams Clickstream sequences at the single-session level are often short—a handful of clicks leading to either abandonment or checkout. For such sequences, γ_t^{S} is noisy and sensitive to initialization, while $\gamma_t^{\mathrm{T}}(k)$ tends to be better behaved for extremely short or ambiguous histories. The mixed posterior $\tilde{\gamma}_t = (1 - \lambda_{\mathrm{T}})\, \gamma_t^{\mathrm{S}} + \lambda_{\mathrm{T}}\, \gamma_t^{\mathrm{T}}$ smooths sampling noise and improves sample efficiency in under-observed regions of the state space.

(5) Auditability of LLM influence Because the LLM teacher can be biased or miscalibrated, the control parameter $\lambda_{\mathrm{T}} \in [0, 1]$ makes LLM influence empirically tunable. At $\lambda_{\mathrm{T}} = 0$ the inference ignores the LLM entirely. At $\lambda_{\mathrm{T}} = 1$ it tracks the LLM's posterior as closely as it can. The choice of λ_{T} is reported empirically in the Analysis section, so LLM influence is an auditable design parameter.

(6) Duration estimation under short sequences Each latent state z_t has an explicit duration distribution $g(d \mid z)$. Estimating this from short, noisy sequences alone is difficult. The teacher posteriors $\gamma_t^{\mathrm{T}}(k)$ provide theory-coherent soft occupancy weights for the declared operational states. The student uses these weights to update duration parameters while remaining constrained by observed click timing.

(7) Cost asymmetry between offline elicitation and online inference The reference distribution γ_t^{T} is expensive to elicit and is computed offline on a sampled training horizon, while the identified semi-Markov inference $p(z_t \mid \mathbf{y}_i, \mathbf{x}_i)$ is cheap and runs at scale on every production request. This asymmetry is a property of the identification argument itself. The reference distribution is needed precisely where identification is at stake during learning. Once the dynamical system has been identified, every subsequent inference is a forward pass on the resulting kernels.

A.4. Proofs of Theorems Stated in the Main Text

This appendix collects the proofs of Theorems 1–4 stated in Section 3.

⌊ Proof of Theorem 1 (Properness and stability)⌉ Each LLM head $u^{\pi}, u^{A}, u^{B}, u^{G}$ feeds a softmax over a finite alphabet, so each produces non-negative entries that sum to one. Assumption 1 lower-bounds all probabilities away from zero, preventing $\log 0$ in (4). Since (1) is a finite sum/product of differentiable functions of θ (composition of LLM mapping and softmax), $\mathcal{L}(\theta)$ is finite and differentiable by standard calculus.

⌊ Proof of Theorem 2 (Expressiveness)⌉ Assumption 2 is a uniform approximation statement for the mapping from inputs (x, r) to the simplex-valued parameters. Given the approximation in logits (or directly in probabilities), the HSMM likelihood computed by the forward recursion (13) is continuous in these parameters, so the induced distribution over $y_{1:T}$ can be made arbitrarily close in total variation for any fixed T.

⌊ Proof of Theorem 3 (Value of retrieval)⌉ For any random variables, conditioning cannot increase conditional entropy:

$$H(Y \mid X, R) \leq H(Y \mid X),$$

with strict inequality iff $I(Y; R \mid X) > 0$. The Bayes-optimal expected log-loss equals the conditional entropy under the true distribution, so $\ell^{\star}(\sigma(X, R)) = H(Y \mid X, R)$ and $\ell^{\star}(\sigma(X)) = H(Y \mid X)$.

⌊ Proof of Theorem 4 (EM monotonicity)⌉ By Jensen's inequality and the standard EM decomposition,

$$\log \mathcal{L}(\theta) = Q(\theta; \theta^{old}) + \mathrm{KL}\big(\mathbb{P}_{\theta^{old}}(q, d \mid y, x, r) \,\|\, \mathbb{P}_{\theta}(q, d \mid y, x, r)\big),$$

up to an additive constant that does not depend on θ. The KL term is non-negative and equals zero at $\theta = \theta^{old}$. Therefore, increasing $Q(\theta; \theta^{old})$ cannot decrease $\log \mathcal{L}(\theta)$.