Information Overload, Cognitive Capacity, and Customer Engagement: Evidence from Retail Messaging Data

1 Introduction and Major Research Question

Digital retail platforms rely heavily on repeated messaging—emails, push notifications, and in-app messages—to stimulate customer engagement. Yet these messages arrive in an environment where customers face limited attention and must make repeated decisions under uncertainty: whether opening the next message will be worth the cognitive cost.

This paper studies how different sources of information pressure—including a firm's own messaging frequency, competitor-side market pressure, contextual timing (weekday, hour, holidays), and major news events—affect customers' cognitive capacity, fatigue, and ultimately engagement decisions.

Central Research Question:

> How do multiple information-environment factors shape customer cognitive capacity and thereby influence message-opening behavior in retail settings?

A motivating idea is that attention is scarce. When customers experience high information pressure—whether from the firm, from competitors, or from the external environment—their ability to process additional commercial messages declines. This results in message fatigue, lower engagement, and potentially market failure where firms collectively oversupply commercial information.

We also ask a second, practical question:

> How do these effects differ across customer segments (loyal, occasional, dormant), and how should firms adjust messaging strategies for each group?

These questions are directly relevant for retail campaign planning. Ignoring information pressure risks inefficient campaigns, reduced ROI, and long-run habituation effects.

2 Conceptual Framework: Decision Under Uncertainty

We model customer message engagement as a decision under uncertainty.

Customer i at time t chooses action $a_{it} \in \{0, 1\}$ where $1 = \text{open}$, $0 = \text{ignore}$. Utility from opening:

$$U_{it}(1) = E_{it}[V_{it}] - C_{it},$$

$$U_{it}(0) = 0.$$

Decision rule:

$$a_{it} = 1 \iff E_{it}[V_{it}] - C_{it} \geq 0.$$

Expected Value

$$E_{it}[V_{it}] = \theta_i \cdot q_{it}$$

- θ_i: latent preference / inherent propensity to value offers.
- q_{it}: message-specific relevance signal.

Cognitive Cost and Fatigue

$$C_{it} = \alpha_i + \beta F_{it}$$

- α_i: baseline cognitive cost.
- F_{it}: fatigue state.
- β: conversion factor from fatigue to cost.

Fatigue dynamics:

$$F_{i,t} = \rho F_{i,t-1} + \gamma \cdot \mathrm{msgs}_{i,t-1} + \delta M_t + \varepsilon_{it}.$$

- $\mathrm{msgs}_{i,t-1}$: recent firm messages.
- M_t: market-wide or external information pressure.
- $\rho \in [0, 1]$: persistence.

Customers with high cognitive capacity behave almost risk-neutral; fatigued customers avoid messages even if potentially valuable.

This framework predicts heterogeneous responses across customer segments.

3 Testable Predictions

We derive three main predictions:

1. **Information pressure reduces engagement.** Higher messaging frequency (own or market-wide) lowers open probability by increasing cognitive cost.
2. **Heterogeneity across segments.**
    - Loyal customers exhibit slower fatigue accumulation.
    - Dormant customers show strong negative responses even at moderate frequency.
3. **Context matters.** Timing (weekday, hour, holiday periods) shifts cognitive cost and open rates.

4 Data Plan

We use fully engineered internal datasets.

Outcome Variable

- Message opened (binary)

Key Explanatory Variables

Own-firm information pressure

- Messages in last 1/7/30 days
- Recency since last message
- Frequency volatility
- Fatigue ratios (short-term vs long-term)

Market-wide information pressure

- Competitor or simulated market messaging averages
- External news cycles (if available)

Contextual cognitive-load shifters

- Day of week, time of day
- Holiday indicator
- Special events

Expected-value proxies

- Customer historical open rates (1-week, 1-month, lifetime)
- Subject line features, personalization, discount mentions
- Campaign historical open/click rates

- Customer value (visits, purchases)

Customer segmentation (three groups)

1. Loyal / high-engagement
2. Occasional
3. Dormant / low-engagement

5 Empirical Strategy

We implement three components:

5.1 1. Descriptive Exploration

- Open rate vs. messaging frequency
- Fatigue curves across segments
- Contextual timing patterns

5.2 2. Reduced-Form Econometric Model

We estimate:

$$\Pr(\text{open}_{it} = 1) = \Lambda\Big(\beta_1 \text{Fatigue}_{it} + \beta_2 \text{ExpectedValue}_{it} + \beta_3 (\text{Fatigue} \times \text{Segment}) + X_{it}\gamma\Big),$$

where X_{it} includes timing controls and campaign fixed effects.

5.3 3. Heterogeneous Effects & Policy Simulations

- Estimate models separately by segment.
- Simulate open probabilities under alternative messaging strategies.
- Provide segment-specific recommendations.

6 Expected Outputs

Figures

- Fatigue curves for each segment
- Heatmap: predicted open probability by (fatigue $\times$ expected value)
- Timing-effect plots

Tables

- Model coefficients and marginal effects
- Heterogeneous effects by segment
- Simulation results for optimal messaging frequencies

Contribution

- Demonstrate how multiple information pressures jointly determine cognitive capacity.
- Highlight strong heterogeneity across customer segments.

- Provide actionable guidance for messaging under limited attention.

7 Feasibility

We already have:

- Fatigue features, expected-value proxies, and market pressure features
- Segmentation options
- Clean panel data structure
- Experience with logit/probit modeling and interactions

The analysis is fully feasible within the available timeline.

Conclusion

This project combines behavioral economics, decision-making under uncertainty, and high-dimensional retail data to understand how information environments shape customer engagement—and how firms can design more effective, less intrusive communication strategies.