ChatGPT

Message Frequency vs Customer Fatigue in E-commerce Communications

Paper Title: *"Balancing Information and Overload: Optimizing E-commerce Messaging Frequency"*
Main Research Question: *How does messaging frequency affect customer engagement in a multi-channel retail setting, and how can firms optimize communication strategies to reduce information asymmetry without inducing customer fatigue?*

Literature Review

Research on marketing communications highlights a trade-off between **information provision** and **consumer overload**. Firms send messages to reduce information asymmetry (e.g. informing customers of deals), but consumers have limited attention and processing capacity. In the "attention economy," a *wealth of information creates a poverty of attention* [1], implying many messages can overwhelm customers. This bounded rationality perspective [2] suggests consumers satisfice rather than fully optimize, processing only a fraction of received messages. Empirical marketing studies confirm *excessive messaging causes fatigue*: for example, 70% of consumers recently reported unsubscribing from brands in a three-month period due to **overwhelming message volume** [3]. High-frequency messaging often leads to diminished open rates and increased churn [4] [5].

Behavioral decision models indicate that **fatigue** lowers consumer engagement. Theoretical work on search behavior finds that as fatigue increases, consumers *search fewer options* [6] and exert less effort in decision-making. In our context, a "fatigue curve" could capture how recent exposure to messages reduces current openness: as customers receive more messages in the short term, their probability of opening further messages declines. Conversely, **long-term engagement and expectations** matter: a history of positive interactions can raise a consumer's propensity to engage (since they expect useful information). Past studies highlight that personalization and relevance mitigate fatigue (e.g. tailored content is 67–81% more likely to be opened) [7].

Competition among senders further complicates this. Game-theoretic models show that when multiple firms vie for limited attention, each has an incentive to send more messages, leading to **over-communication with no equilibrium** [8]. For example, a model of "celebrities" racing to post on social networks finds no pure-strategy equilibrium in message frequency: each sender increases rate to attract attention, but collectively they swamp users [8]. By analogy, retailers in a competitive market may end up over-messaging to avoid losing share, even though this creates negative externalities (customer fatigue) for everyone. This is a kind of market failure where information needs (in theory) justify messaging, but bounded rationality means too many messages become counterproductive.

Key references include works on information overload and advertising wear-out (diminishing returns to message frequency) and on boundedly rational information consumption. For example, ad-frequency studies note "wear-out" effects, where response declines after repeated exposure. The classic attention-economy insight by Herbert Simon warns that when information is abundant, attention is the bottleneck

[1]. Behavioral decision theory (e.g. Simon's bounded rationality [2]) and recent consumer studies on fatigue [3] [6] will guide our modeling of how uncertainty and messaging costs affect open rates. We will also draw on marketing science results on message saturation and response dynamics to frame our empirical hypotheses.

Data and Empirical Plan

Dataset and Variables

We will use the **REES46 CDP multichannel messaging dataset** (Apr 2021–Apr 2023), which logs every message (email, web push, mobile push, SMS) sent by a mid-sized retailer. Key fields likely include: *customer ID*, *timestamp*, *channel*, *campaign type*, and binary outcomes (*opened, clicked, converted*). The dataset has ~10 million records, enough for detailed analysis. We will preprocess it as follows: - **Clean & filter data:** Remove system/test messages, deduplicate, handle missing fields, and define a customer's active period. - **Create temporal features:** For each message, calculate **recent frequency**: e.g. number of messages the customer received in the prior 7 days (overall and per channel). Also track **engagement history**: e.g. past 30-day open-rate or past purchases from messages. - **Customer-level features:** Define a *long-term engagement trend*, such as cumulative open-rate or recency/frequency of past clicks. This proxies expectation or interest. Possibly cluster customers by engagement profile (high vs low responders). - **Competitor/simulation variables:** The dataset has only this retailer's sends. To capture competitive pressure, we may *simulate an external "market message volume"* variable. For instance, assume a uniform additional messaging rate or use a proxy from industry benchmarks (e.g., average ads per week in e-commerce). Even a constructed index (like competitor message count = fraction of this firm's count) can allow a sensitivity analysis on competition effects.

Metrics and Model Specification

We will empirically model the probability that a given message is opened (and possibly clicked or converted), focusing on **open rate** as the primary engagement metric. The key independent variables will be: - **Recent Fatigue Index:** e.g. count of messages in past 1–2 weeks (higher = more fatigue). - **Historical Engagement:** e.g. customer's past open/click rate or time since last click (higher = stronger interest/loyalty). - **Message and Campaign Controls:** channel type, campaign category, day of week/time, subject-line length, etc. - **Customer Controls:** account age, total lifetime spends (if available). - **Competitive Pressure (simulated):** a hypothetical variable for overall market messaging rate.

Our core hypothesis: *higher short-term messaging frequency reduces open probability (negative fatigue effect) after controlling for baseline interest; while higher long-term engagement increases it (loyal customers remain more receptive).*

Regression Analysis

We plan logistic regression (or multilevel logistic) for `P(open=1)`. Formally:

$$\mathrm{logit}(P(\mathrm{open})) = \beta_0 + \beta_1 \times \mathrm{Frequency}_{\mathrm{recent}} + \beta_2 \times \mathrm{PastOpenRate} + \mathbf{X}\beta + \epsilon.$$

Here **Frequency_recent** is expected to have negative coefficient (fatigue), while **PastOpenRate** positive. Controls (X) include channel fixed effects, day/time dummies, etc. Customer random effects could capture unobserved heterogeneity. We might use generalized estimating equations to account for repeated messages per customer or cluster standard errors.

Nonlinear/Behavioral Models

To capture potential nonlinear fatigue, we may define a “fatigue curve”: e.g. allow frequency effects to saturate (e.g. piecewise linear or quadratic). Decision-tree or random forest models can explore such nonlinearity and interactions (e.g. maybe email vs SMS fatigue differ). A tree could reveal thresholds: perhaps beyond 5 messages/week, open rate drops sharply.

Simulations for Competition

While data lacks actual competitor actions, we can perform scenario simulations. For example, assume an exogenous “market messaging intensity” that either is constant or varies seasonally. We could simulate how changes in this pressure (holding our own strategy fixed) would alter open rates, illustrating the game-theoretic incentive to message more when others do. The empirical plan could include fitting a separate model adding this dummy competition variable to see its qualitative effect.

Additional Analyses

- **Channel Differences:** Build separate models for email, SMS, etc., since channels have different norms (SMS open rates are high but customers also expect low volumes).
- **Time-Series Effects:** Use time-interval variables (e.g. days since last message to customer) to detect short-term fatigue vs recovery period.
- **Segment analysis:** Investigate if “fatigue curves” differ by customer segments (e.g. heavy purchasers vs occasional shoppers).
- **Robustness:** Check for reverse causality (e.g. if uninterested customers ignore emails and then receive more messages, it could look like frequency causing disengagement). Instrumental variables are tough; instead we may lag variables to ensure causal ordering.

Expected Findings and Implications

We expect to find that **message frequency negatively impacts immediate open rates**, consistent with a fatigue effect. In particular, a spike in communications in a short window should lower the open probability of subsequent messages [6] [3]. Conversely, customers with high historical engagement or loyalty likely show resilience: they may tolerate more messages without sharp drop in opens, reflecting *long-term expectation effects*. If so, this suggests a “fatigue–expectation trade-off”: emailing a generally interested customer often is okay, but even loyal customers have a tipping point beyond which fatigue kicks in. We can illustrate this with a hypothetical *fatigue function* (e.g. exponential decay of open probability with cumulative daily sends).

Managerial implications: Firms should optimize *dynamic messaging policies*. For instance, use our model to predict a “safe” frequency for each customer segment (e.g. 3 emails/week for casual shoppers vs 7 for highly engaged ones). Personalization tools (like adaptive send limits) can incorporate customer-specific fatigue thresholds. Over-messaging, even if driven by competitor pressure, harms everyone – so firms might coordinate on industry best practices (e.g. loyalty programs or preference centers where customers set their preferred contact frequency). The regression results will quantify the trade-offs: how many extra sales (conversions) are gained by an extra message versus how many opens or long-term engagements are lost.

Policy implications: From a policy perspective, this scenario resembles a negative externality of messaging. Regulators or industry bodies might consider setting guidelines: for example, mandatory opt-out mechanisms (already required by law in many countries) and stricter caps on message frequency (e.g. limiting promotional emails to a certain number per week). The literature suggests **regulatory interventions** can mitigate the "race to the bottom" in communications quality. Mandatory unsubscribe or do-not-disturb periods can protect consumers from fatigue. On the other hand, policies promoting *permission-based marketing* and personalization (ensuring messages match recipient interests) can achieve the informational benefits without the cost of fatigue [7].

If competition is intense, a "tragedy of the commons" can arise: each retailer individually tries to message more to grab attention, but collectively they degrade the medium. Our plan's brief game-theory overview (citing no-equilibrium results [8]) supports the idea that without coordination or limits, firms will likely end up over-saturating customers. Thus, besides technical personalization, broader strategies like industry self-regulation (best-practice charters) or consumer choices (preference dashboards) are important.

Project Timeline and Deliverables

The plan is executable over 4–5 weeks, yielding a 10–12 page paper and a 10-minute presentation. A possible timeline: 1. **Week 1:** Finalize research question, conduct detailed literature review, and acquire/understand the dataset. Define all variables, start data cleaning. 2. **Week 2:** Compute engagement metrics (fatigue index, past open rates), do exploratory analysis (e.g. plot open rate vs frequency). Draft empirical strategy section. 3. **Week 3:** Estimate baseline models (logistic regression with controls). Interpret key coefficients (frequency, past engagement) and refine features (try nonlinearity or interactions). Begin writing results. 4. **Week 4:** Extend analysis (segment models, tree-based checks, or simulation of competition). Draft discussion and implications, including policy recommendations. 5. **Week 5:** Finalize write-up (10–12 pages with citations and figures), prepare presentation slides. Rehearse speaking points on main findings and implications.

Expected outcomes: The final paper will clearly state the research question, situate it in the literature on information overload and bounded rationality, detail the empirical design, and discuss findings in depth. It will include citations like studies of consumer fatigue [3] [4] and game-theory of messaging competition [8]. The discussion will explicitly address how firms can adjust messaging strategy and what policies (opt-out rules, channel caps, personalization standards) could improve social welfare by balancing the information-communication trade-off.

---

[1] Paying Attention: The Attention Economy – Berkeley Economic Review
https://econreview.studentorg.berkeley.edu/paying-attention-the-attention-economy/

[2] Bounded rationality - Wikipedia
https://en.wikipedia.org/wiki/Bounded_rationality

[3] [5] [7] 2025 Marketing Fatigue Report | Optimove
https://www.optimove.com/blog/marketing-fatigue-insights

4 What Affects Email Open Rates? A Deep Dive
https://www.mailstand.com/blog/what-affects-email-open-rates

6 anderson.ucla.edu
https://www.anderson.ucla.edu/sites/default/files/document/2022-01/SearchGaps_UrsuZhangHonka_Dec2021.pdf

8 Title
https://www.contrib.andrew.cmu.edu/~ravi/waw10.pdf