# BEHAVIORAL RESPONSES TO THE SPOTIFY PLATFORM'S TRANSFORMATION: CONSUMER SENTIMENT ANALYSIS

by

Yehor Kraievskyi

A thesis submitted in partial fulfillment of the requirements for the degree of

MA in Business and Financial Economics

Kyiv School of Economics

2025

Thesis Supervisor: Professor Pavlo Prokopovych

Approved by

Head of the KSE Defense Committee, Professor [Type surname, name]

Date: November 1, 2025

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

**EDA** Exploratory Data Analysis

**VADER** Valence Aware Dictionary and sEntiment Reasoner

**BE** Behavioral Economics

**DiD** Difference-in-difference

**ML** Machine Learning

**KDE** Kernel Density Estimation

**UI/UX** User Interface / User Experience

# Chapter 1. Introduction

Spotify pioneered music streaming and established itself as the global leader in the market. Founded on the 23rd of April 2006 by Daniel Ek and Martin Lorentzon in Stockholm, Spotify has since dominated the music streaming market it once helped create. The Swedish media giant is now increasingly positioning itself as a "one-stop shop" for all things audio, expanding into podcasts in 2015 and audiobooks in 2022 (Spotify, 2025). This expansion shows the direction of Spotify's transformation. Spotify is aiming to reach new audiences, expand into new markets, and, most importantly, fulfil its promises to investors.

Spotify serves as an indicator of modern digital market. A digital streaming platform that acquired a large user base through investments, but is now struggling to compete. Spotify is not the only business to face this, and not the only platform to diversify. But how are its users reacting to changes? What are the mechanisms behind consumers' reactions to changes? This research tries to answer these questions.

This Master's thesis focuses on exploring consumer sentiment through Behavioral Economics lens. The core research hypotheses are formulated on well-researched BE mechanisms: loss aversion, status quo bias, psychological reactance. The core analysis studies how user feedback evolved before and after the rollout of new features and price changes. The thesis also explores the broader context of Spotify's business model. The findings are applicable to a broad spectrum of digital products and provides valuable insight into consumer sentiment.

The motivation for this research is in the author's interest in Behavioral Economics, digital platforms and Spotify in particular. Additionally, the topic allows to obtain rich dataset that span years of consumer reviews. Such datasets allow for a robust temporal analysis. From a business perspective, Spotify's transformation is also interesting, as many companies find themselves in a similar circumstance.

A broader business context is particularly significant for this research. To understand why Spotify decided to change its model this thesis explores Spotify's financial reports, academic papers and various business literature. Spotify's transformation shows whether it is possible to maintain long-term sustainability and profitability in a highly competitive market. The company has not been able to turn a profit until 2024 (Spotify, 2024) due to various reasons explored in the following chapters. The shift to podcasts and audiobooks might help Spotify reduce its dependency on costly music licensing fees and differentiate its product from the competition. For example, Spotify's own financial reports (Spotify, 2024) show that podcasts broke even in 2024.

Spotify's transformations increase listening times, and introduce Spotify to new audiences, according to Spotify's own Q3 2023 shareholder letter (Spotify, 2023). However, the shift can also have its downsides, such as diluting Spotify's core proposition, increasing premium subscription prices, and alienating the core audience by shifting resources from music streaming (Smith, 2023; Patel, 2022). These user reactions are particularly relevant from a behavioral economics perspective, as price increases or major product changes often trigger responses tied to loss aversion, status quo bias, and fairness considerations. This thesis' research hypotheses focus on these exact behavioral phenomena.

The findings reveal valuable insights. They demonstrate that consumer sentiment might not behave as expected and product managers should consider this when planning any changes to the platform. Consumer reactions to strategic pivots are measurable in real time through app review analysis, and BE theory provides useful frameworks for understanding these reactions. The results suggest that managers should anticipate negative sentiment effects from any major platform changes, and should carefully manage price increases and, new feature additions and major redesigns.

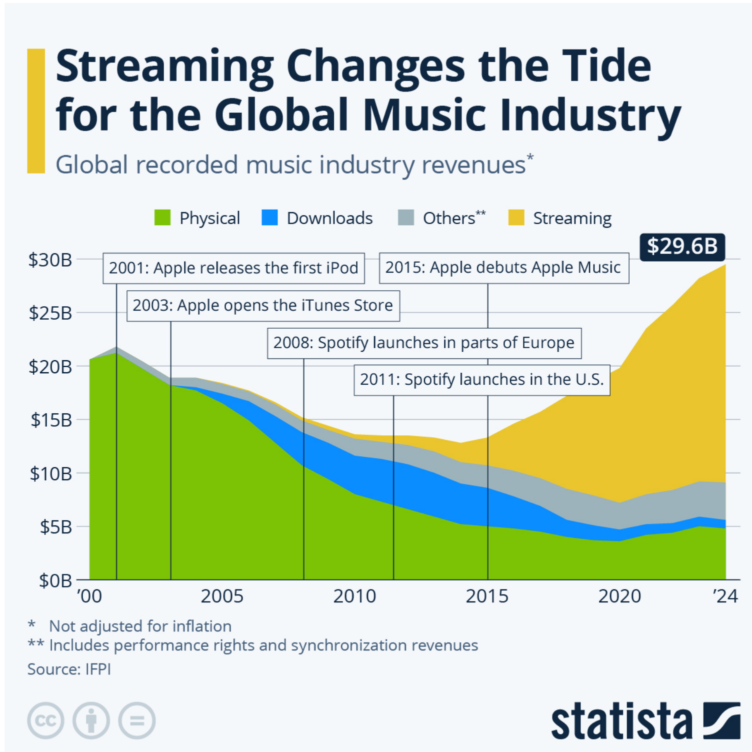
By combining sentiment analysis with behavioral economics theory, the study demonstrates how consumers react to strategic pivots in real time. Product managers, marketing managers and researchers should find this research useful. The findings show that behavioral economics theory is a useful technique to understand the underlying mechanisms behind consumer sentiment. The analysis should also be interesting to a general audience of readers interested in business topics, tech-savvy consumers, frequent Spotify users, and content creators.

# CHAPTER 2. INDUSTRY OVERVIEW AND RELATED STUDIES

## 2.1 Industry overview

Digital audio and the music industry have undergone several transformations over time. The invention of the phonograph in 1877 allowed the first commercial recordings and established the foundation for mass music distribution (Yellowbrick, 2024). The next pillar in the evolution was radio's invention and distribution in the early 20th century. In the late 20th century, vinyl records became the dominant form of music distribution, and CDs subsequently replaced them. 1999 saw the first sign of disruption in the music industry with the launch of Napster, a peer-to-peer music sharing platform. In 2001, Apple introduced legal digital music sales through the iTunes store, which popularized the digital ownership model (Clausius Press, 2023). Music streaming was the latest transformation of the industry and has since experienced explosive growth, as shown in Figure 1 below.

Figure 1. Global recorded music industry revenues

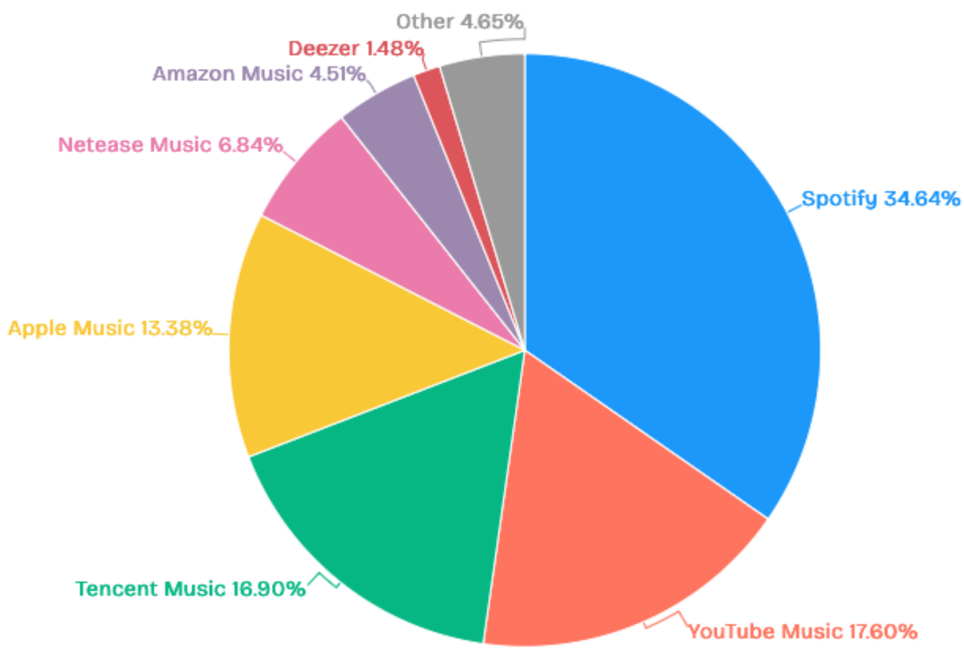



Note. Adapted from Streaming Changes the Tide for the Global Music Industry, by IFPI, 2024, [statista.com]. Copyright 2024 by IFPI. Reprinted under fair use for educational purposes.

The global music streaming market is highly competitive, with key players like Spotify, Apple Music, YouTube Music, and Amazon Music. Spotify currently maintains the leadership position with the highest market share among competitors. According to its financials (Spotify, 2025), in Q1 2025, Spotify had 678 million monthly active users (MAUs), 268 million of whom were subscribed to Spotify Premium. Its closest competitor, Apple Music, has 93 million MAUs (Kumar, 2025), all of whom are Premium users, since Apple Music does not offer an ad-supported tier.

As shown in Figure 2, Spotify continues to lead the global music streaming market by a significant margin (Business of Apps, 2024).

Figure 2. Global music streaming market share.




Note. Adapted from Music Streaming Market Share, by Business of Apps, 2024, https://www.businessofapps.com/data/music-streaming-market/. Copyright 2024 by Business of Apps. Adapted under fair use for educational purposes.

Thus, the four-firm concentration ratio for the music streaming industry is:

$$C_4 = 0.346 + 0.176 + 0.169 + 0.134 = 0.825$$

In recent years, podcasts have emerged as a rapidly growing market segment of the broader digital audio industry. In 2025 the global podcasting market is valued at approximately USD 32.54 billion. It is projected to grow at a compound annual growth rate (CAGR) of 27.0%, reaching USD 173.49 billion by 2032 (Coherent Market Insights, 2025) The increasing popularity of podcasts can be attributed to the general increase in demand, as people search for more convenient forms of entertainment, education and other various topics. Podcast monetization has also become more popular, as advertisers realised the growing potential of reaching new audiences with interest-targeted ads. Streaming platforms dominate the podcast market, which accounts for over USD 13 billion in revenue for platforms in 2025, according to Coherent Market Insights (2025).

The audiobook market is experiencing robust growth as well. This growth is driven by a broad adoption of smartphones and AI, more publishers getting into the audiobook industry, and the general growth of the broader book market. In 2025, the market revenue is projected to reach approximately USD 9.8 billion, with a steady compound annual growth rate (CAGR) of around 6.2% through 2030, reaching over USD 13 billion (Statista, 2025). North America dominates the audiobook market with over 44% of the overall global revenue, while the Asia-Pacific region is experiencing the most significant year-over-year growth according to Grand View Research (2024).

The following two tables provide industry analysis with Porter's Five Forces and PEEST framework.

Table 1. Porter's five forces of Spotify in the Music Streaming, Podcasts, and Audiobooks markets.

| Force | Music Streaming | Podcasts | Audiobooks |
|---|---|---|---|
| **Threat of new entrants** | **Low.** The barriers are: high licensing costs, extensive intellectual property agreements, and the need for large upfront investment into infrastructure and algorithms are needed. | **Medium.** Podcasting has lower production costs and freely available distribution platforms. However, reaching a significant audience and monetizing the podcast requires significant investments in quality content and marketing. Creating a distribution platform for podcasts also requires a significant investment, and the business model for it is unclear since consumers are not used to paying for this form of content. | **Medium.** The audiobook market has moderate entry barriers. Acquiring audiobook rights and owning audiobook production can involve high costs. However, independent and small publishing platforms can create opportunities for creating niche platforms at lower costs. |
| **Bargaining power of suppliers** | **Very high.** Music labels are consolidated into a couple of key players that control most of the market. Their power is underscored by high licensing fees that constitute music streaming platforms' most significant operational expenses. | **Low.** As the production costs for a podcast are minimal, anyone can create and upload a podcast. As the entry barriers for new creators are minimal, the bargaining power for individual creators is low. However, more established and popular creators have more leverage. | **Medium.** A relatively small number of publishers control audiobook rights. However, these publishers are less consolidated than music labels, and there are many independent ones. |
| **Bargaining power of buyers** | **Medium.** The switching cost for a listener is low. However, platforms create platform-specific algorithms, features, and personalized playlists that increase the perceived cost of switching to a new platform. | **High.** Most podcasts are freely available on the most popular platforms, so there are no switching costs for consumers. Algorithms do not play a significant role here since podcasts are produced periodically, so an average listener typically listens to their favorite shows. | **Low-medium.** Typically, consumers gather a non-transferrable library of audiobooks, so switching platforms is not an option. Recently, switching costs have been lower with a wider adoption of the subscription-based model. |
| **Threat of substitutes** | **High.** Consumers can switch to other forms of digital audio, such as YouTube, podcasts, and short-form content, which are free. | **Medium.** Podcasts are a unique form of audio content and cannot easily be replaced by other content. YouTube videos, music, and radio are some formats that can substitute podcasts, but the distinctive format of podcasts mitigates this threat. | **High.** Audiobook listeners have the option to switch to physical or e-copies of the book they are listening to at the moment at any time. |
| **Industry rivalry** | **Very high.** Large tech corporations (Google, Apple, Amazon, Tencent) entered the market and are currently engaged in price wars, continuous innovation in user experience, and aggressive content licensing strategies | **High.** Multiple competing platforms exist (Apple, Spotify, Google, and Amazon). These companies invest in exclusive deals and original content to attract listeners. However, podcasts do not offer a clear path to monetization for platforms, so the competition is lower than in the music streaming market. | **Medium-high.** The audiobook market is increasingly competitive, as it is marked by competition among established platforms like Audible (Amazon), Apple Books, and Spotify. While still expanding, the rivalry is accelerating due to significant investments in original and exclusive audiobook content to attract listeners. |

Table 2. STEEP analysis of Spotify in the Music Streaming, Podcasts, and Audiobooks markets.

| Factors | Music Streaming | Podcasts | Audiobooks |
|---|---|---|---|
| **Social** | Consumers increasingly prefer personalized, convenient digital access to music over the traditional legacy formats. Social media amplifies the trend for socially influenced listening habits. | The popularity of podcasts increases as people look for engaging and informative content for commuting, leisure, and everyday chores as "background entertainment." | Modern consumers' Fast-paced lifestyles decrease the available time for reading books, so readers look for convenient, hands-free media consumption while performing other activities like commuting. |
| **Technological** | Advancements in streaming algorithms and widespread AI adoption significantly increase the user experience, allowing streaming platforms to provide tailored recommendations. | The widespread availability of hardware and software tools allows anyone to start their podcast. Automated transcriptions and algorithms and improved internet accessibility enable users to engage with podcasts anytime. | Similarly, technological improvements in audio production enable more publishers to create audiobooks. The broad adoption of various playback devices allows listeners to listen to their favorite audiobooks on the go. |
| **Economic** | Subscription-based models brought the price of legally listening to music way down. However, pressures from music labels and the wider economic instability drive some listeners back to ad-supported plans or legacy media. Artists are unsatisfied with streaming earnings, discouraging talent from entering the industry. | Podcasts are less sensitive to economic fluctuations due to low production costs. However, podcast monetization can be economically volatile, as advertisers cut costs during economic hardships. | Audiobooks are sensitive to economic shifts since people cut costs on entertainment during economic disruptions. |
| **Ecological** | Music streaming significantly reduced the environmental footprint of the music industry compared to traditional media. Still, data centers are consuming much electricity, raising concerns. | Podcasts also benefit ecologically from the broad adoption of digital distribution platforms, significantly reducing the footprint compared with physical media. | Audiobooks have a significantly lower environmental impact compared to printed books. |
| **Political and Legal** | Complex copyright laws, licensing agreements, and royalty frameworks influence the business model of streaming platforms. There are multiple ongoing lawsuits from artists and labels regarding copyright infringement. There are also concerns regarding the unregulated use of AI-generated music on streaming platforms. | Podcasts face fewer legal issues than music. However, in some jurisdictions, concerns about spreading misinformation on podcasts are raised. | Audiobooks face issues similar to music streaming, as copyright infringement is common in the industry. |

2.2 Future Viability

Spotify's long-term viability depends on whether the company can address structural weaknesses in its business model while building on the advantages of its global scale. The reporting of its first annual profit in 2024 marked a significant turning point. However, the achievement should not be overstated. Nearly 70% of Spotify's revenues continue to flow directly to rights holders, putting Spotify's management in a tough spot (Strategic Management Insight, 2025). This heavy cost burden means that even with a growing user base, profitability remains fragile. Thus, to survive Spotify needs to diversify its revenue streams as well as lower its share of revenue paid to rights holders.

The dependence on premium subscriptions is one of the major issues with Spotify's business model. Approximately nine out of every ten euros Spotify earns are derived from paid accounts (Investing.com, 2025). The Wall Street Journal (2025) called subscription-based revenue streams as "recession-resistant". However, the author finds it hard to agree with, as the latest research shows that during economic downturns subscriptions are the first to be cut from consumers' budgets (Barclays, 2025). Thus, Spotify's main revenue stream is under tremendous threat during the economic uncertainty and Trump-era isolationism.

For years, Spotify has promoted its data-driven approach as a competitive advantage. Spotify gathers millions of data points and tracks nearly every user interaction imaginable. However, critics point out that Spotify is not able to turn their data-driven model, compared, for example, with Meta (Business Insider, 2025). The push into podcasts highlights this issue. While exclusive content deals help Spotify to differentiate, the financial returns do not meet the expectations (Alexander, 2023). Spotify even tried to experiement with messaging features, short-form videos to try . Audiobooks represent a more focused commitment and a general pivot in strategy. Reuters (2025) points out that without clear evidence that these ventures will pay off, they risk complicating the business model without a positive impact to the bottom line.

The most problematic factor shaping Spotify's viability is its relationship with artists and music labels. The company has faced extensive criticism for low royalty payments to artists. "Discovery Mode" algorithm offers algorithmic visibility in exchange for lower payouts and this drew a lot of criticism in the industry (Music Business Worldwide, 2025). In addition, Pelly (2025) argues that Spotify increasingly positions music as a background utility, or a "mood machine," not a dedicated platform to connect music lovers and creators. This might work for casual listeners, but it risks alienating the musicians.

All in all, Spotify's future depends on more than growth in user numbers. To secure long-term stability, the company will need to diversify revenues beyond subscriptions, reduce dependency on costly licensing agreements, and repair problematic relationships with artists. Without demonstrable progress in these areas, the shift toward becoming a super app might not be enough to sustain the profitability of the platform.

## 2.3 Literature review

Studies related to the subject can be broadly organized into two categories: papers evaluating Spotify's transformation from a business perspective and Behavioral Economics studies that examine some aspects of Spotify's business model. Foundational Behavioral Economics studies have been utilized as well to lay the groundwork for the main contribution of this thesis.

### Business literature

The research literature on Spotify's change in strategy shows that the subject has been researched from many diverse perspectives. Researchers have been looking into its changing business model, how it's managing to retain users, where it's generating revenues, and how it's differentiating itself from competitors.

Kiberg and Spilker (2023) provide a conceptual perspective of how Spotify has evolved from a music streaming company to a comprehensive audio platform. They posit that the industry is evolving rapidly, and these days, what is most important is exclusive content and a sophisticated set of features to differentiate.

Wang (2023) looks into the algorithms, in particular Discover Weekly. He he refers to Discover Weekly as one of the most influential personalization innovations. The article states that personalization is one of the most important factors of long-term user retention.

A number of articles examine the impact of such algorithms on real data. Holtz et al. (2020) conducted randomized experiments with Spotify Premium users. He discovered that personalized podcast recommendations led to short-term user engagement increase. However, there was an adverse effect present too. The same algorithms that grab the attention of users also limit their exposure.  Users end up listening to a less diverse array of content with time. Anderson et al. (2020) come to the same conclusion.

Raj (2021)  conducts a quantitative examination of Spotify's in-house data. His argument is that algorithmic complicates the competition among Spotify's content creators. When one show or song is aggressively promoted, it artificially changes consumers' listening patters and also affects the shows and music in a related genre.

Spotify's expansion into podcasting has also been a subject of academic papers. Pérez-Alaejos, Terol-Bolinches, and Barrios-Rubio (2022) describe how Spotify uses exclusive podcast agreements and acquisitions to attract new consumers and inspire loyalty. Mei (2024) and Li (2024) examine these maneuvers via SWOT analysis and conclude that content diversification is an intelligent means for Spotify to lower their dependance on music labels. These reports also indicate that Spotify's new model favors additional revenue-generating methods, such as variable pricing and targeted advertising.

YEs, Knox, and Bronnenberg (2017) illustrate that listeners still crave variety. This is partly why Spotify tries to offer more content but relies on personalization to help people find what they like. Hracs and Webster (2020) refer to competitive advantage through data-driven curation, exclusive content, and sophisticated user experiences.

Behavioral economics literature

Behavioral economics has established that consumer decisions often deviate systematically from rational choice models. Prospect theory shows that individuals are loss-averse. They react more strongly to perceived losses compared to equivalent gains (Kahneman & Tversky, 1979). Prospect theory shows explains why Spotify users react negatively to subscription price increases and feature removals. Status quo bias (Samuelson & Zeckhauser, 1988) and psychological reactance (Brehm, 1966; Steindl et al., 2015) show that consumers are somewhat resistant to any imposed changes, even when those changes might offer more opportunities or expand functionality. Egelman et al. discuss that in digital products, users often see modifications to autonomy rather than enhancements.

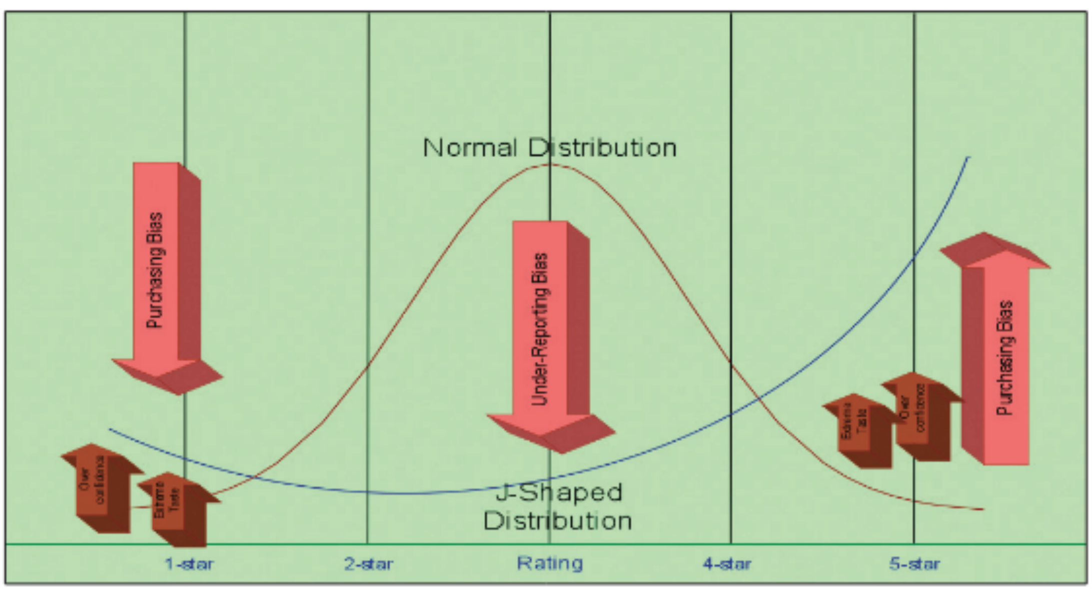
The choice overload phenomenon shows that expanding options can reduce satisfaction and hinder decision-making. When faced with an increasing number of options, users might reject most of them and use only the basic functions they are accustomed to (Iyengar & Lepper, 2000). Algorithm aversion desctibed by Dietvorst et al. (2015) is another phenomenon that is the most relevant for the audio supper app like Spotify. It reveals that the trust in the automated recommendation algorithm, can decline rapidly if there are visible errors in the algorithm. For example, multiple users complain that Spotify's curated playlists offer the same selection of music.

Multiple studies effects highlight how platform-selected options (autoplay, pinned content) shape user behavior (Johnson & Goldstein, 2003). Mental accounting (Thaler, 1999) and the zero-price effect (Shampanier, Mazar, & Ariely, 2007) further demonstrate how bundling "free" features can alter perceived value. Finally, fairness norms (Kahneman, Knetsch, & Thaler, 1986) and negativity bias (Baumeister et al., 2001) show why digital consumers frequently frame platform changes as unfair or manipulative, amplifying negative feedback in public reviews.

### Literature on consumer sentiment and app-store reviews

Digital feedback in the form of app reviews provides a necessary framework for the main analysis of this thesis. As outlined by Hu, Pavlou, & Zhang, product reviews most frequently show a J-shaped distribution. The study highlights an under-reported bias for 2-4 star reviews. Thus, textual analysis should offer richer insights than just star ratings (Guzman & Maalej, 2014).

Figure 3. Product review distribution




Note. From “Overcoming the J-shaped distribution of product reviews,” by N. Hu, P. A. Pavlou, and J. Zhang, 2009, *Communications of the ACM, 52*(10), p. 144. Copyright 2009 by the Association for Computing Machinery. Reprinted under fair use for educational purposes.

Empirical studies demonstrate that updates and redesigns provoke significant shifts in sentiment. Martin et al. (2018) show that “bad updates” in Google Play apps often trigger negative spikes as a result of feature removals or navigation changes. Saidani et al. (2022) confirm that review sentiment correlates strongly with disruptive update shocks. These findings are consistent with loss aversion and reactance mechanisms. Furthermore, Muchnik, Aral, and Taylor (2013) highlight the role of herding effects in online ratings, showing how early extreme reviews bias subsequent user evaluations.

# Chapter 3. Methodology

## 3.1 Research hypotheses

Based on the behavioral economics theory and the literature on digital platforms, this thesis tests four core hypotheses related to Spotify's consumer sentiment.

This study employs a quasi-experimental event study design using difference-in-differences (DiD) analysis to examine how platform changes affect user sentiment and behavioral responses. The research analyzes user-generated reviews from the iOS App Store to test four hypotheses related to psychological reactance (H1), status quo bias (H2), loss aversion and fairness concerns (H3), and positive value framing (H4).

The event study approach is well-suited for this research because it allows for causal inference by comparing user responses in treated markets (those experiencing platform changes) to control markets (those without changes) across pre-event and post-event periods.

To test the four research hypotheses, this study analyzes three distinct platform events that occurred at different points in time and affected different markets. This approach strengthens external validity and allows for hypothesis-specific identification strategies.

### Event 1: Audiobook and Podcast Rollout (H1 and H4)

Event Description: On November 8, 2023, Spotify launched audiobooks as part of its Premium subscription in the United States, providing subscribers with 15 hours of audiobook listening per month at no additional cost. This event represents a significant feature addition to the platform, corresponding to Spotify's commitment to transformation and expansion.

Treatment and Control Groups:

- Treatment: United States
- Control: Canada and India

Canada and India did not receive the audiobook feature during the study period, making them appropriate counterfactuals for isolating the effect of the audiobook rollout on US user sentiment.

Data Collection Window: September 15, 2023 to December 31, 2023 (approximately ±6 weeks around the event date)

Hypotheses Tested:

H1 (Psychological Reactance): Predicts that users may express complaints about being forced into unwanted changes or feeling restricted in their choices

H4 (Positive Value Framing): Predicts that users may appreciate the added value and express positive sentiment about the bundled audiobook feature

Event 2: Mobile Application Redesign (H2)

Event Description: On March 8, 2023, Spotify announced a major redesign of its mobile application interface at the "Stream On" event. The redesign introduced a new dynamic home page with visual previews (Canvas clips), separate feeds for Music, Podcasts & Shows, and Audiobooks, and an enhanced Search feed with scrollable visual content. This represented the most significant user interface change in Spotify's recent history.

Because this was a global rollout with gradual implementation rather than country-specific, this part of the study uses the date of the redesign launch as the proxy for treatment. At first, the app version was considered as an ideal indicator of treatment (whether the user used the old or the redesigned app). However, due to Spotify not releasing any public release note along with the updates, the author failed to identify the exact or even an approximate app version that utilizes this redesign. Thus, this hypothesis is tested within the US user base with the redesign release date acting as a proxy for treatment. A large window of observations is chosen to allow as many users as possible to update their apps.

Data Collection Window: February 1, 2023 to August 31, 2023 (approximately ±3 months around the event date to capture gradual rollout).

Hypothesis Tested: H2 (Status Quo Bias): Predicts that following platform updates or redesigns, users demonstrate change aversion by preferring prior versions and explicitly resisting changes.

Event 3: US Premium Price Increase (H3)

Event Description: On July 24, 2024, Spotify increased the price of its individual Premium subscription in the United States from $9.99 to $10.99 per month. This price increase was not accompanied by any new features or service improvements.

- Treatment: United States
- Control: India.

Data Collection Window: June 15, 2024 to August 31, 2024 (approximately ±6 weeks around the event date)

Hypothesis Tested: H3 (Loss Aversion and Fairness Concerns): Predicts that subscription price increases or removal of valued features disproportionately increase complaints about fairness and perceived exploitation relative to the benefits of added content

## 3.2 Data Collection

Review data were collected from the iOS App Store using AppFollow, a professional app store intelligence and review management platform. The iOS App Store was selected as the primary data source because it provides the option to filter country-specific reviews, while the Google Play Store reserves this option for developers only.

## 3.3 Sentiment analysis

The datasets are enriched through VADER (Valence Aware Dictionary and sEntiment Reasoner) sentiment analysis tool. VADER is specifically optimized for social media and informal text. It works by first breaking down text into individual words and assigning each a valence score from a pre-compiled lexicon indicating the word's emotional intensity and polarity. These scores are then combined into an overall sentiment score for the text, called the compound score, which ranges from -1 (most negative) to +1 (most positive). VADER provides sentiment polarity subscores (positive, negative, neutral) along with the compound score, which was used as the main sentiment measure in this research. This approach enables precise and interpretable sentiment quantification suitable for large-scale text analysis in social science contexts.

3.4 Analysis tools

The primary empirical approach employs difference-in-differences estimation, which compares the change in outcomes for treated units (countries or users experiencing the platform change) to the change in outcomes for control units (countries or users not experiencing the change) over the same time period.

The baseline DiD specification for datasets with control markets (H1/H4, H2):

$$Y_{it} = \alpha + \beta_1 Treatment_i + \beta_2 Post_t + \beta_3 (Treatment_i \times Post_t) + \gamma_k Country_k + \varepsilon_{it}$$

H2 hypothesis, where only US market data is analyzed has a similar setup:

$$Y_{it} = \alpha + \beta_1 Treatment_i + \beta_2 Post_t + \beta_3 (Treatment_i \times Post_t) + \varepsilon_{it}$$

Where:

- $Y_{it}$ is the outcome (sentiment score) for review $i$ at time $t$.
- $Treatment_i$ is an indicator for the treatment group;
- $Post_t$ is an indicator for post-event period (1 = after event 0 = before);
- $Treatment_i \times Post_t$ is the interaction term – the main DiD estimator;
- $\gamma_k Country_k$ – country fixed effects included to control for unobserved, time-invariant differences between countries in the dataset;
- $\varepsilon_{it}$ – error term;
- $\alpha, \beta_1 \ldots$ - coefficients.

3.5 Behavioral Mechanism Flagging.

To test specific psychological mechanisms, binary flags were created using keyword-based text analysis of review content. Full keyword dictionaries are provided in Appendix A. Each review in the dataset is enriched with hypothesis-specific flags, based on these keywords. Each flag equals 1 if the review contains one or more keywords associated with the mechanism, and 0 otherwise.

- Reactance Flag (H1): Captures language indicating perceived threats to freedom or forced changes. Keywords include: forced, have to, must, need to, made me, cannot, won't let, mandatory, restriction, limited, no choice, blocked, dictate, compelled, intrusive, manipulative, etc.
- Status Quo Bias Flag (H2): Captures language indicating preference for previous versions or resistance to change. Keywords include: used to, old version, previous, before, revert, legacy, familiar, classic, tradition, accustomed, habit, change, different now, bring back, don't like change, etc.
- Fairness/Loss Aversion Flag (H3): Captures language indicating perceptions of unfairness, excessive pricing, or loss. Keywords include: unfair, expensive, overpriced, ripoff, not worth, robbery, steal, cheating, greedy, price hike, lost feature, took away, hidden fee, exploit, unjust, etc.
- Positive Value Framing Flag (H4): Captures language indicating appreciation for added value. Keywords include: worth it, good deal, love it, best, awesome, recommend, excellent, satisfied, value for money, happy, pleased, favorite, improvement, etc.

  To analyze those behavioral flags, several tests are conducted:

- Chi-Square test - evaluates whether the distribution of flags differs significantly across two time periods (before and after the event);
- Two-proportion z-tests analyze flag prevalence between pre-event and post-event periods

- Logistic regression for flag outcomes models flag presence as a binary outcome conditional on treatment assignment and event timing:

$$\log\left(\frac{P(Flag_i = 1}{1 - P(Flag_i = 1)}\right) = \gamma_0 + \gamma_1\, Treatment_i + \gamma_2 Post_t + \gamma_3(Treatment_i \times Post_t)$$

Logistic regression was chosen because the outcome is binary, and it provides interpretable effect sizes.

Finally, the flags are added to the main DiD model:

$$\begin{aligned} Y_{it} = \alpha + \beta_1 Treatment_i + \beta_2 Post_t + \beta_3(Treatment_i \times Post_t) + \beta_4 Flag_i \\ + \beta_5(Treatment_i \times Flag_i) + \beta_6(Post_t \times Flag_i) \\ + \beta_7(Treatment_i \times Post_t \times Flag_i) + \varepsilon_{it} \end{aligned}$$

The main DiD term $(\beta_3)$captures the average treatment effect of the event in the treatment group relative to controls. The flag main effect $(\beta_4)$ measures baseline sentiment differences between flagged and non-flagged reviews. The two-way interactions capture how treatment and timing effects differ as a function of flag status, while the three-way interaction $(\beta_7)$ tests whether the DiD effect is significantly different for flagged versus non-flagged reviews.

### 3.6 Robustness Checks

The main advantage in this research design is that it avoids single-method bias. Behavioral flagging is used in addition to VADER sentiment scores to check precisely for the behavioral mechanisms they represent, while using only sentiment scores could result in omitted variable bias where sentiment is influenced by unobserved factors. Multiple countries are analyzed where appropriate to minimize the selection bias. All regressions use heteroscedasticity-robust standard errors (HC1) to account for potential violations of homoscedasticity in review data.

### 3.7 Ethical considerations

This study analyzes publicly available user reviews posted voluntarily on the iOS App Store. No personally identifiable information is collected beyond what users chose to share publicly.

# CHAPTER 4. DATA

The data for this thesis consists of three datasets, each collected to test one of the research hypotheses. As outlined in the Methodology section, each research hypothesis has its own event date - a date where a major change on Spotify's platform and/or its pricing occurred. Each dataset consists of user-generated app reviews collected from the iOS platform, capturing user sentiment across the United States, India, and Canada. This section gives an overview of these datasets.

The data is gathered using the AppFollow.io (Appfollow, 2025) software package due to its flexibility in filtering. A custom review scraper has proven to be difficult to build due to restrictions in the google-play-scraper Python library. Spotify reviews were downloaded from Apple AppStore exclusively, because it allows to filter country-specific reviews, which is crucial for the analysis.

Each dataset has the following core variables used in the analysis or data preparation and filtering:

Table 3. Description of variables

| Variable | Explanation |
|---|---|
| Submission date | Review submission date |
| Country | Reviewer's country |
| Review language | Only English reviews were considered |
| Version | Spotify app version. Used to determine the presence of a treatment effect. |
| Rating | User's app rating (from 1 to 5) |
| Review | The text of the review itself |
| vader_neg, | The proportion of negative sentiment detected in the review text, as scored by the VADER sentiment analysis algorithm. Ranges from 0 (no negative sentiment) to 1 (entirely negative). |
| vader_neu, | The proportion of neutral sentiment detected in the review text by VADER. Indicates the degree to which the review is emotionally neutral, on a scale from 0 to 1. |
| vader_pos, | The proportion of positive sentiment identified in the review text by VADER. Scores range from 0 (no positive sentiment) to 1 (entirely positive). |
| vader_compound | The proportion of positive sentiment identified in the review text by VADER. Scores range from 0 (no positive sentiment) to 1 (entirely positive). |

## 4.1 Exploratory data analysis

H1: Psychological Reactance, H4: Positive Value Framing. Event: US Audiobooks introduction on November 8, 2023.

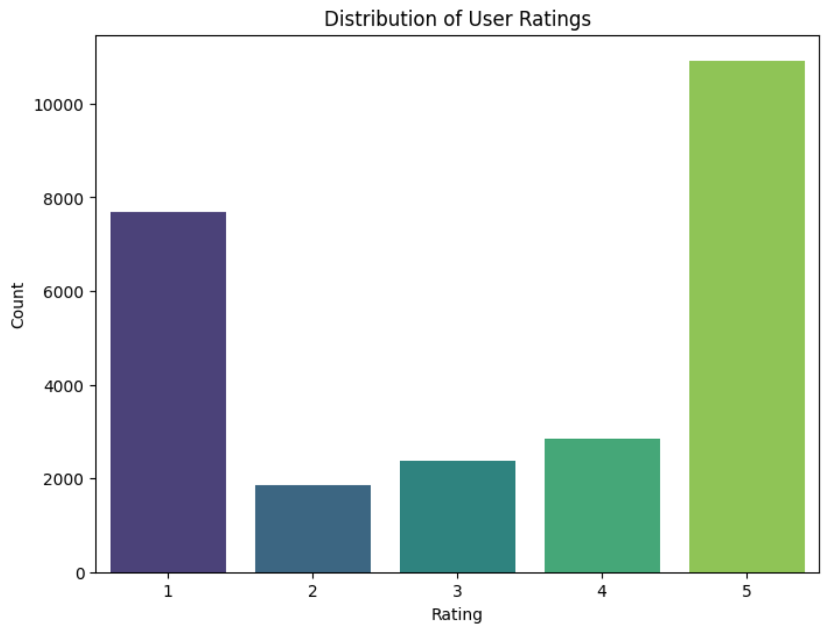
The first dataset contains 25688 observations: 18784 from the US (treatment group), 5077 from India (first control group), and 1827 from Canada (second control group). Observation dates range from 15th of September, 2023 to 31st of December, 2023 with the event date being the 8th of November, 2023. The dates were chosen to have a robust group of pre-event and post-event reviews to analyze the treatment effect. The range of dates is similar in the treatment and control groups. The detailed output is provided in Appendix A (Table 1).

The table above demonstrates that across countries, US users assign the highest average ratings and show the least negative sentiment, while Indian users give the lowest ratings and display the most negativity. Canadian reviews are generally moderate in both sentiment and ratings. Overall, review scores and sentiment metrics confirm substantial differences in user experience and satisfaction across markets.

The summary statistics table 2 in Appendix A shows that most users give high ratings (median 4, mean 3.29), but there is a wide spread across the full range from 1 to 5. Sentiment analysis indicates that the majority of review text is neutral (mean vader_neu 0.73), with positive sentiment (mean vader_pos 0.20) outweighing negative sentiment (mean vader_neg 0.07). The overall compound sentiment is moderately positive (mean 0.31), but both the ratings and sentiment scores have considerable variability, highlighting diverse user experiences in the dataset.

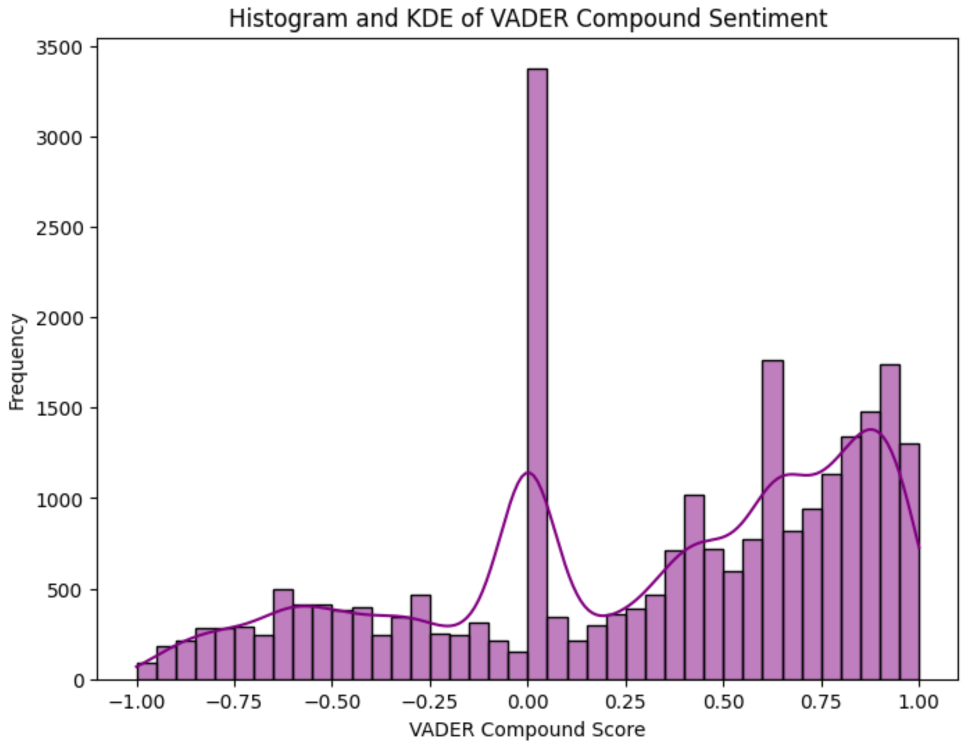
The distribution of user review scores is consistent with the most common J-shaped distribution for product reviews (Hu et al., 2009).

Figure 4. Distribution of user rating in dataset 1.



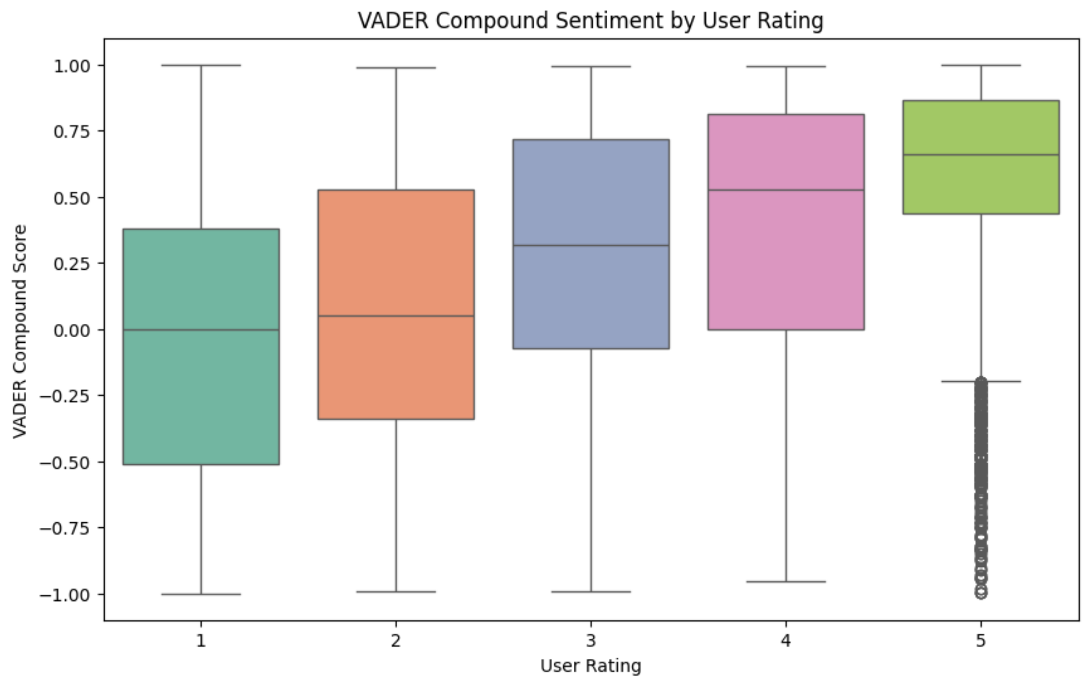

The histogram and KDE curve below reveal a multimodal pattern, with a strong peak at zero (neutral sentiment) and secondary peaks at both the highly positive (close to +1) and highly negative (close to -1) ends of the scale. The most frequent sentiment is neutral, indicating that many users write emotionally balanced or mixed reviews. However, there is also substantial representation of strongly positive sentiment, and to a lesser extent, strongly negative sentiment, highlighting different user opinions.

Figure 5. Histogram and KDE of VADER compound sentiment in Dataset 1



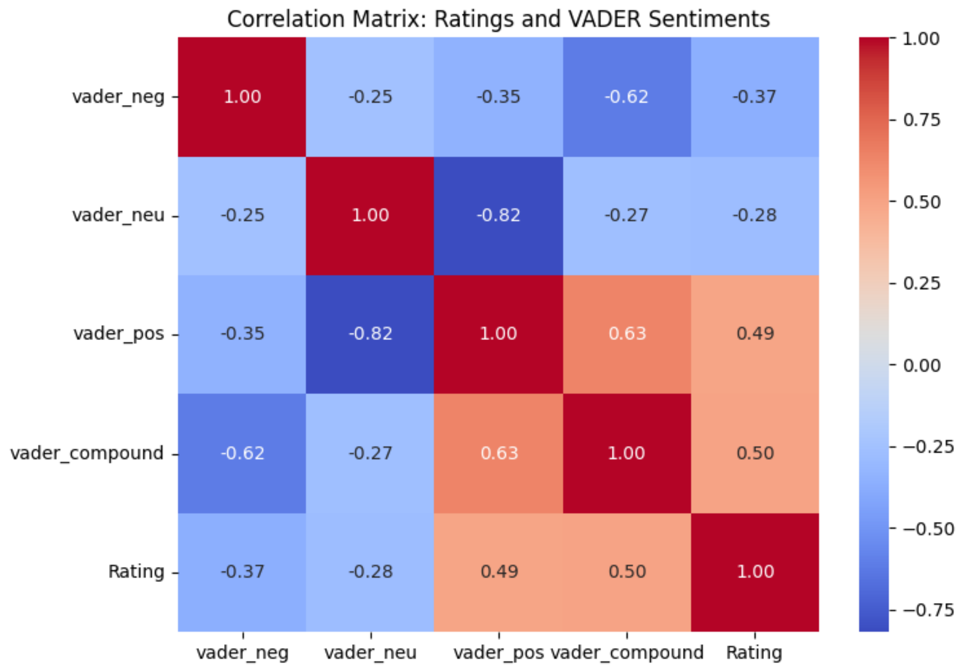

The boxplot below shows the relationship between the VADER sentiment analysis and user ratings. As user ratings increase from 1 to 5, the median compound VADER sentiment score increases as well. There are some outliers in 5-star and 1-star user ratings, indicating a mismatch between the sentiment score and the user score, possibly due to user mistakes, sarcasm, or other atypical behavior.

Figure 6. Boxplot of VADER compound sentiment by user rating in Dataset 1




The correlation matrix confirms that both user ratings and textual sentiment metrics are aligned: more positive numeric ratings tend to be accompanied by more positive and fewer negative review sentiments.

Figure 7. Correlation matrix - VADER and ratings in Dataset 1



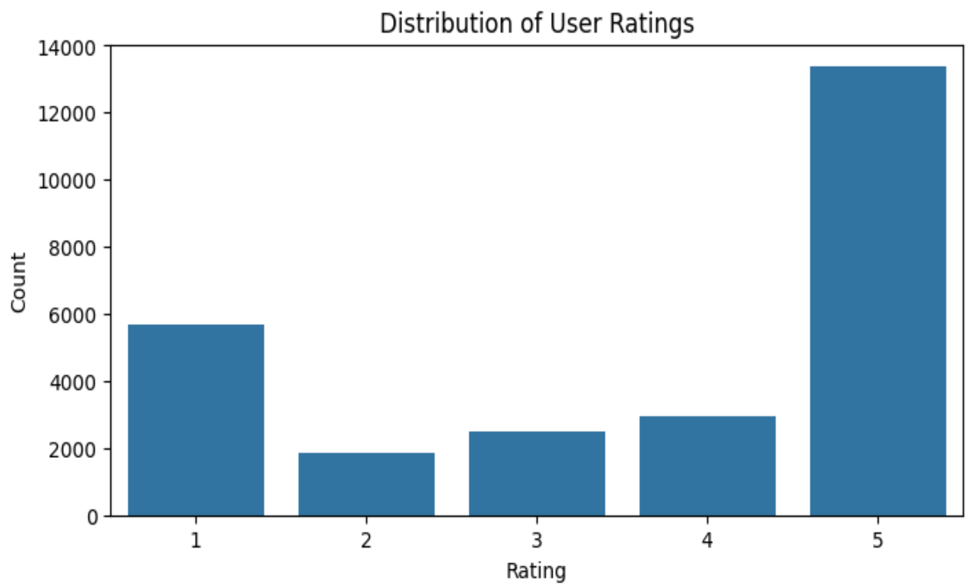

H2: Status Quo Bias. Event: Stream On UI/UX redesign, March 8, 2023.

The second dataset consists of 26361 app store reviews submitted by US users, specifically centered around the period of Spotify's Stream On redesign. Review submissions in this dataset span from January 1, 2023 to May 31, 2023, bracketing the date of the major interface update, which was introduced on March 8, 2023.

The summary statistics table (Appendix A, Table 3) below shows that, across 26361 reviews, the mean app rating was 3.63 out of 5, with notable dispersion (std = 1.64). Sentiment analysis with VADER shows that the average compound score was moderately positive (mean = 0.37), while the positive, neutral, and negative sentiment components demonstrate that most reviews were skewed toward neutral or positive language.

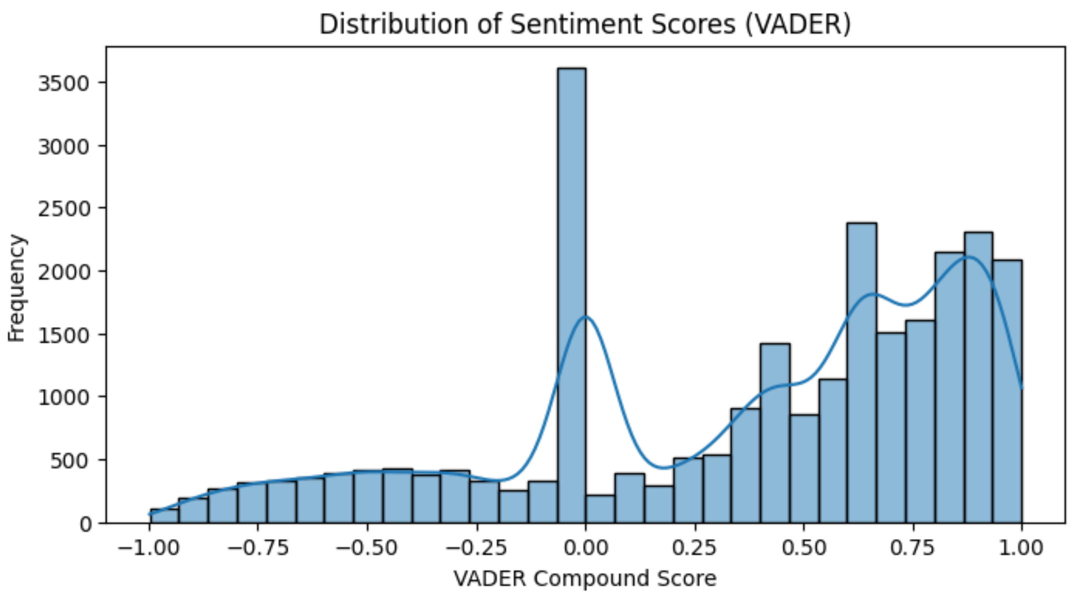
Similarly to the first dataset, the rating distribution follows the J-shaped curve:

Figure 8. Distribution of user ratings in Dataset 2



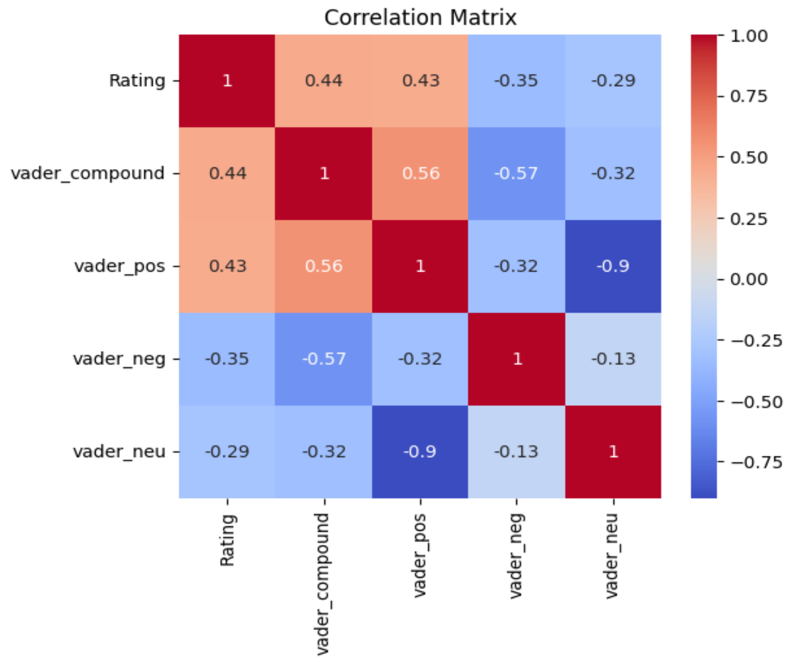

The distribution of VADER compound sentiment scores for US Spotify app reviews between January and May 2023 is bimodal, with clear peaks around 0 (neutral) and 0.9 (strongly positive). Most reviews show either neutral or positive sentiment, with a pronounced spike at exactly zero reflecting the many reviews containing neutral language. The right tail is notably heavier, indicating that positive experiences are more frequently or more strongly expressed than negative ones. Low and negative sentiment reviews are present but much less common, as shown on the figure below:

Figure 9. Distribution of VADER sentiment - Dataset 2



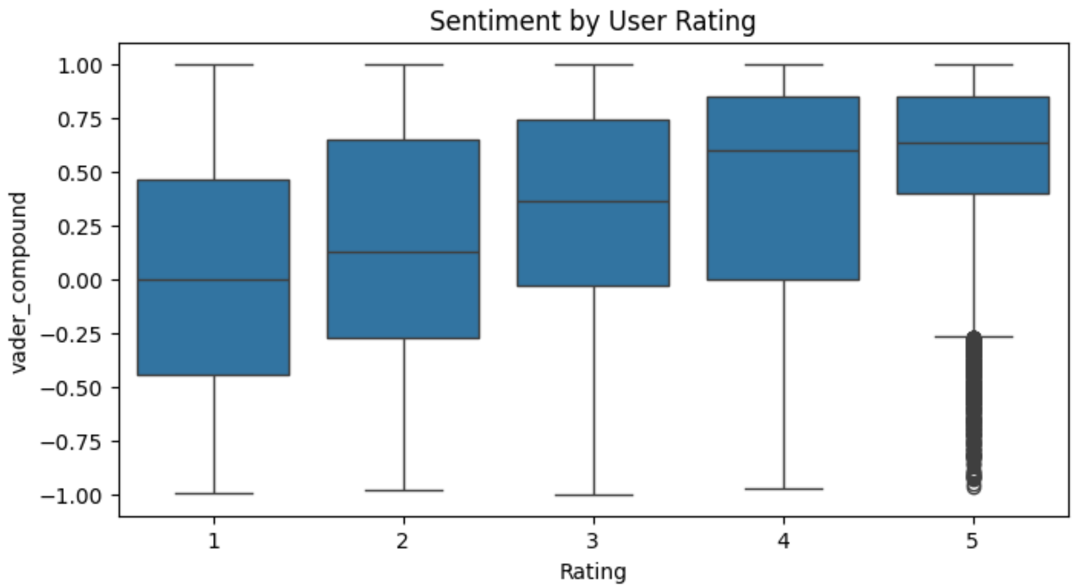

The correlation matrix shows that user ratings are moderately to strongly aligned with VADER sentiment scores: higher ratings correspond closely with higher VADER compound and positive scores, and show a negative relationship with negative sentiment.

Figure 10. Correlation matrix - VADER vs user ratings in Dataset 2




The boxplot confirms that sentiment rises stepwise with rating level, with more positive sentiment observed in higher-star reviews. This validates VADER as a reliable proxy for subjective user satisfaction in this dataset. Notably, there are some outliers, similar to the first dataset in 5-star rating group.

Figure 11. Boxplot of Sentiment scores by user rating in Dataset 2




H3: (Loss Aversion and Fairness Concerns). Event: US Premium Plan price increase on July 24, 2024

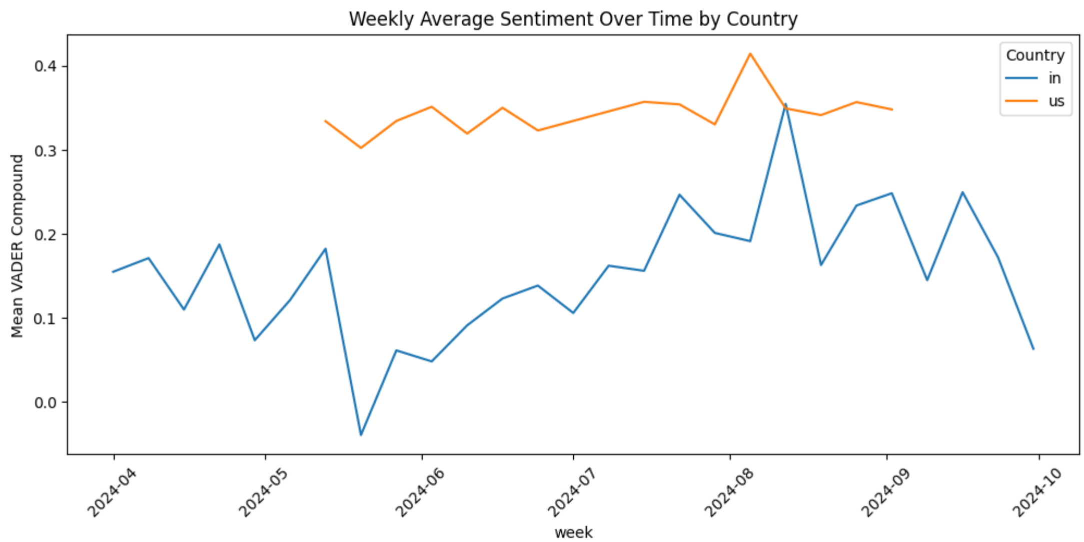
The Hypothesis 3 dataset comprises 22,003 user reviews of Spotify submitted across the United States (treatment), India (control group) from April 1 to September 30, 2024, spanning the July 24, 2024 US price increase. There are 19674 US reviews and 2329 reviews from India (Appendix A, Table 4).

The dataset is structured similarly to the previous two, with only exception in VADER sentiment variable names: they are named *neu*, *pos*, *neg*, and *compound*, corresponding to neutral, positive, negative, and overall compound sentiment respectively.

Pre-event sentiment in the US treatment market averaged 0.33 to 0.36, declining slightly to around 0.31 following the price hike in late July and early August before recovering in subsequent weeks. Notably, a sentiment spike occurred in early August across both US and India, likely driven by Billie Eilish's Olympics performance on August

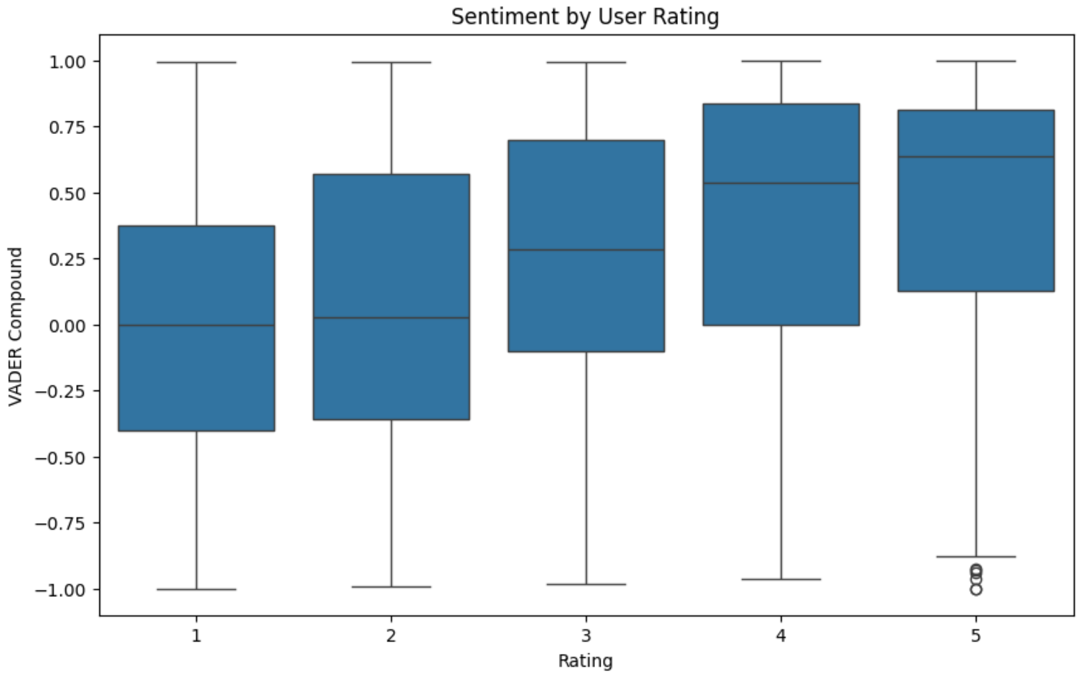
11, 2024, highlighting the importance of controlling for external cultural shocks. The control market (India) exhibited lower overall sentiment levels with greater volatility, suggesting that sentiment drivers differ across geographic markets independent of the US pricing event.

Figure 12. Weekly average sentiment over time - Dataset 3



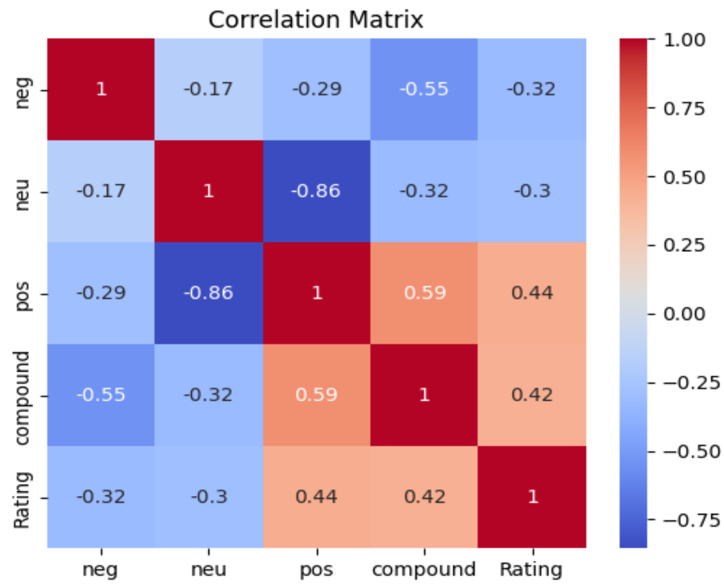

The exploratory data analysis reveals consistent, interpretable relationships between sentiment, ratings, and country in the context of the July 2024 Spotify price hike. As shown in the boxplot, there is a clear, nearly monotonic increase in VADER compound sentiment as user ratings increase from 1 to 5 stars. Reviews with lower ratings are notably more likely to feature negative or neutral sentiment, while those with higher ratings exhibit much more positive sentiment, as seen by the upward trend in the medians and overall spread.

Figure 13. Boxplot of sentiment by user rating - Dataset 3



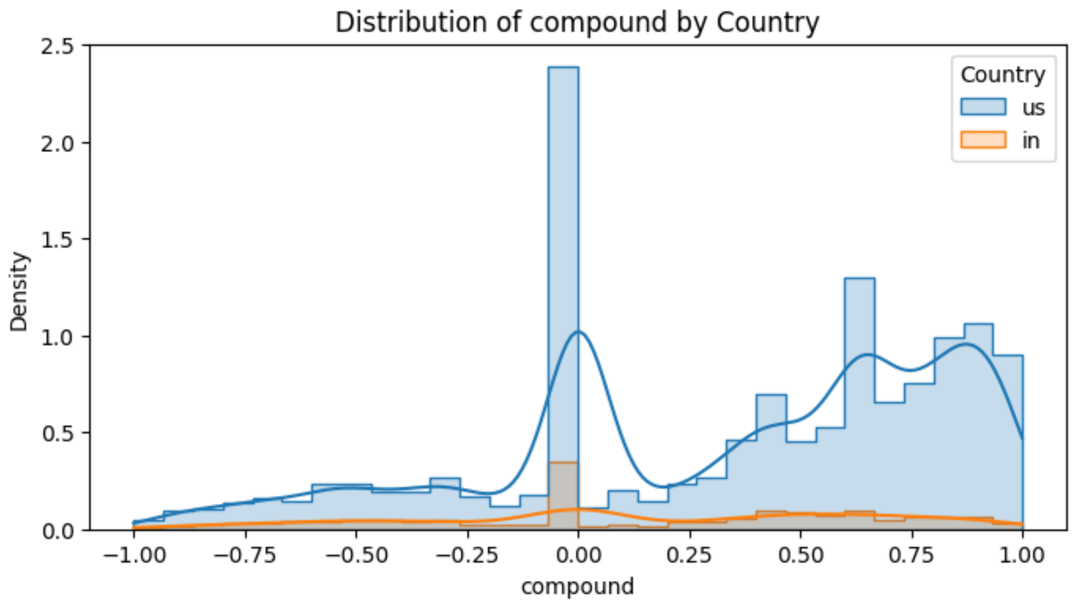

The correlation matrix quantifies these patterns: VADER compound and positive sentiment are moderately to strongly positively correlated with user ratings (0.42 and 0.44, respectively), while negative sentiment is inversely related to both compound sentiment (−0.55) and ratings (−0.32).

Figure 14. Sentiment Scores vs User Reviews correlation matrix - Dataset 3.




Further, a comparison of the distributions of compound sentiment by country shows that the US market (treatment group) exhibits greater density in positive sentiment and generally higher VADER compound scores, while India (control group) presents a more spread distribution, with a larger fraction of neutral or slightly negative reviews. These visualizations collectively highlight the strong alignment between star ratings and VADER sentiment, validate the consistency of the sentiment pipeline, and further underscore the meaningful differences in sentiment structure between treatment and control markets in the event window.

Figure 15. Distribution of sentiment by country in Dataset 3



5.1 H1: Psychological Reactance, H4: Positive Value Framing. Event: US Audiobooks introduction on November 8, 2023.

This subsection presents the empirical findings on how US user sentiment changed following the launch of audiobooks for Spotify Premium on November 8th, 2023, using India and Canada as control markets. As outlined in the methodology section, all sentiment measures are based on the VADER compound score, which captures the overall tone of user-written app reviews.

First, the original DiD model is estimated, as defined in the Methodology section:

Table 4. DiD estimation (sentiment) - H1/H4

OLS Regression Results

| | | | |
|---|---|---|---|
| Dep. Variable: | vader_compound | R-squared: | 0.091 |
| Model: | OLS | Adj. R-squared: | 0.091 |
| Method: | Least Squares | F-statistic: | 581.0 |
| Date: | Thu, 30 Oct 2025 | Prob (F-statistic): | 0.00 |
| Time: | 13:23:06 | Log-Likelihood: | -19119. |
| No. Observations: | 25688 | AIC: | 3.825e+04 |
| Df Residuals: | 25683 | BIC: | 3.829e+04 |
| Df Model: | 4 | | |
| Covariance Type: | HC1 | | |

| | coef | std err | z | P>\|z\| | [0.025 | 0.975] |
|---|---|---|---|---|---|---|
| Intercept | 0.3143 | 0.014 | 23.101 | 0.000 | 0.288 | 0.341 |
| C(Country)[T.in] | -0.3418 | 0.014 | -24.167 | 0.000 | -0.369 | -0.314 |
| C(Country)[T.us] | 0.0397 | 0.007 | 5.579 | 0.000 | 0.026 | 0.054 |
| treatment | 0.0397 | 0.007 | 5.579 | 0.000 | 0.026 | 0.054 |
| post | 0.0326 | 0.013 | 2.452 | 0.014 | 0.007 | 0.059 |
| treatment:post | -0.0399 | 0.016 | -2.497 | 0.013 | -0.071 | -0.009 |

| | | | |
|---|---|---|---|
| Omnibus: | 1835.519 | Durbin-Watson: | 1.975 |
| Prob(Omnibus): | 0.000 | Jarque-Bera (JB): | 1726.848 |
| Skew: | -0.575 | Prob(JB): | 0.00 |
| Kurtosis: | 2.462 | Cond. No. | 6.86e+15 |

The coefficient on the DiD interaction term, $\beta_3$, is -0.0399 (standard error 0.016, $p = 0.013$). This means that after Spotify introduced audiobooks in the US, the average user sentiment, as measured by VADER, decreased by approximately 0.04 points, compared to the change observed in India and Canada. This is a modest but statistically robust decline. In other words, US reviews became more negative compared to trends in countries that were not exposed to the new feature during the same period.

Other parts of the model behave as expected. The overall post-event period is associated with a slight increase in sentiment (coefficient 0.0326, $p = 0.014$) across all markets, but this is more than offset by the negative US-specific effect. Sentiment scores in India remain considerably lower than in Canada or the US, as reflected by the negative country coefficient for India (-0.3418, $p < 0.001$).

The model's R-squared is relatively low (0.091). However, this is typical for the analysis of the user-generated reviews, as explained by Peterson and Ozili (2023). The models in this dataset remain statistically significant with robust standard error estimates.

The first model shows that including the audiobooks in US Premium Plan led to a measurable and statistically significant drop in user sentiment. The control markets didn't observe such a change. Preliminary, we can conclude that major rollouts can be perceived negatively by platforms' users. The further behavioral flagging analysis tests whether we can truly attribute the sentiment decline to BE mechanisms.

The results of the flagging analysis (Appendix A, Table 5) show a disconnect between the overall sentiment decline and the presumed behavioral mechanisms that caused it. The behavioral flags were added into the OLS regression model with interaction terms for treatment and the post-event period. The DiD interaction coefficients for both H1 reactance and H4 positive framing were not statistically significant (treatment:post:flag_h1_reactance, $p = 0.585$; treatment:post:flag_h4_positive, $p = 0.354$). This indicates that the observed sentiment decline in US reviews was not disproportionately driven by reviews flagged with explicit reactance language. The sentiment decline was offset by increased positive framing language in the post-event period either.

At the aggregate level, chi-square tests (Appendix A, Tables 6,7) revealed that the proportion of US reviews flagged for H1 reactance decreased from 18.05% pre-event to 15.13% post-event ($p < 0.001$. The proportion of US reviews showing positive value (Hypothesis H4) framing increased slightly from 34.24% before the event to 35.96% after.

A logit regression (Appendix A, Table 8) test confirmed that the US reactance flag rate decreased by 1.19 percentage points more than control countries ($p < 0.001$), while the US positive framing flag rate increased 0.82 percentage points more than controls ($p = 0.041$).

5.2 H2: Status Quo Bias. Event: Stream On UI/UX redesign, March 8, 2023.

An ordinary least squares regression with heteroscedasticity-robust standard errors was estimated to assess whether the Stream On redesign on March 8, 2023 affected average review sentiment in the US market. The pre-event average VADER compound sentiment was 0.3839, whereas the post-event average declined to 0.3672, representing a statistically significant decrease of 0.0168 points (approximately 1.68 percentage points, $p = 0.007$). The post-event coefficient in the regression was $\beta = -0.0168$ ($SE = 0.006$, $z = -2.694$, 95% CI [-0.029, -0.005]). Although the magnitude of sentiment decline is modest and the overall R-squared is very small, the effect demonstrates that user reviews became measurably more negative following the redesign rollout. This pattern is consistent with prior research on user reactions to platform interface changes.

Table 5. DiD estimation (sentiment) - H2

```
                            OLS Regression Results
==============================================================================
Dep. Variable:         vader_compound   R-squared:                       0.000
Model:                            OLS   Adj. R-squared:                  0.000
Method:                 Least Squares   F-statistic:                     7.255
Date:                Sat, 01 Nov 2025   Prob (F-statistic):            0.00707
Time:                        09:18:03   Log-Likelihood:                -19246.
No. Observations:               26361   AIC:                         3.850e+04
Df Residuals:                   26359   BIC:                         3.851e+04
Df Model:                           1
Covariance Type:                  HC1
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept      0.3839      0.005     82.739      0.000       0.375       0.393
post          -0.0168      0.006     -2.694      0.007      -0.029      -0.005
==============================================================================
Omnibus:                     2366.685   Durbin-Watson:                   1.982
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             2820.161
Skew:                          -0.777   Prob(JB):                         0.00
Kurtosis:                       2.608   Cond. No.                         2.80
==============================================================================
```

To test for status quo bias, or the behavioral manifestation of user preference for the pre-redesign interface, reviews were flagged for language indicating resistance (Appendix B) to change or expressed preference for the original design. The proportion of reviews exhibiting status quo bias language increased significantly from 5.88% pre-event to 9.43% post-event, a difference of 3.55 percentage points. This pre-post change was confirmed to be statistically significant by multiple statistical tests. A chi-square test yielded $\chi^2 = 111.95$, $p < 0.001$, and a two-proportion z-test yielded $z = 10.604$, $p < 0.001$(Appendix A, Table 9). Additionally, a logistic regression (Appendix A, Table 10) estimating the odds of observing status quo bias language post-event versus pre-event shows a post-event coefficient of $\beta = 0.5108$ ($SE = 0.049$, $z = 10.512$, $p < 0.001$), corresponding to an odds ratio of approximately 1.67. This result indicates that the odds of a review displaying status quo bias language increased by 67% following the redesign launch.

Reviews were also flagged for expressions of enthusiasm toward or comfort with the redesign – "openness to change" language (Appendix B). The proportion of reviews containing such language remained almost stable, with 5.85% before and and 5.73% post-event. This change was not statistically significant ($\chi^2 = 0.1398$, $p = 0.7085$; $z = -0.401$, $p = 0.6888$).

A logistic regression (Appendix A, Table 11) showed a post-event coefficient of $\beta = -0.0213$ (SE = 0.053, $z = -0.401$, $p = 0.689$), indicating no meaningful shift in the positive framing.

A secondary event-study analysis aggregated review-level data into weekly bins to examine the redesign's effect in a more aggregate way. The aim was to reduce noise and reveal underlying trends more clearly. This weekly difference-in-differences specification (Appendix A, Table 12) produced a post-event coefficient of $\beta = -0.0061$ (SE = 0.008, $p = 0.438$). The overall model was highly significant ($F = 1206.2$, $p < 0.001$), but didn't show any meaningful shift.

The analysis provides evidence that the Stream On redesign triggered two primary behavioral and psychological responses in US user reviews. First, there was a statistically significant, though modest, absolute decline in review sentiment. Second, there was a robust and substantial increase in the prevalence of status quo bias language expressing preference for and resistance to the change. In contrast, explicit expressions of openness to or enthusiasm for the redesign did not increase post-event.

### 5.3 H3: (Loss Aversion and Fairness Concerns). Event: US Premium Plan price increase on July 24, 2024

The regression results show that the Spotify price increase on July 24, 2024, led to a statistically significant decline in user sentiment in the US market compared to Indian market (control). The key interaction term from the difference-in-differences model is negative and significant (coefficient = $-0.0756$, $p = 0.001$). This interaction term shows that after the price hike, sentiment in the US dropped by an additional 0.076 points compared to the control market.

These findings provide evidence that users responded negatively to the price increase. Although the model's R-squared is modes (0.015), this is typical for consumer review analysis Ozili (2023). But the model's statistical significance and direction of the effect are clear and consistent with expectations for user sentiment trends.

Table 6. DiD estimation (sentiment) - H3

```
                            OLS Regression Results
==============================================================================
Dep. Variable:               compound   R-squared:                       0.015
Model:                            OLS   Adj. R-squared:                  0.015
Method:                 Least Squares   F-statistic:                     108.9
Date:                Sat, 01 Nov 2025   Prob (F-statistic):           5.86e-70
Time:                        10:20:19   Log-Likelihood:                -15698.
No. Observations:               22003   AIC:                         3.140e+04
Df Residuals:                   21999   BIC:                         3.144e+04
Df Model:                           3
Covariance Type:                  HC1
==================================================================================
                     coef    std err          z      P>|z|      [0.025      0.975]
----------------------------------------------------------------------------------
Intercept          0.1183      0.013      9.004      0.000       0.093       0.144
treatment          0.2163      0.014     15.500      0.000       0.189       0.244
post               0.0954      0.021      4.483      0.000       0.054       0.137
treatment:post    -0.0756      0.022     -3.371      0.001      -0.120      -0.032
==============================================================================
Omnibus:                     1791.332   Durbin-Watson:                   1.995
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             1614.764
Skew:                          -0.595   Prob(JB):                         0.00
Kurtosis:                       2.411   Cond. No.                         14.9
==============================================================================
```

Chi-square tests and two-proportion z-tests were conducted to examine changes in loss aversion and fairness concerns language between pre-event and post-event periods within each geographic market. For the US treatment group, the prevalence of loss aversion language declined from 13.67% pre-event to 11.25% post-event (chi-square = 25.165, $p < 0.001$; z-test: $z = -5.04$, $p < 0.001$). Similarly, in India, loss aversion language decreased from 11.32% to 8.55, also statistically significant (chi-square = 4.121, $p = 0.042$; z-test: $z = -2.10$, $p = 0.036$).

Fairness concerns language showed a comparable pattern. In the US, fairness language decreased from 3.10% pre-event to 2.52% post-event (chi-square = 5.527, $p = 0.019$; z-test: $z = -2.39$, $p = 0.017$). In India, fairness language also declined from 2.98% to 1.22% (chi-square = 6.384, $p = 0.012$; z-test: $z = -2.67$, $p = 0.008$). In contrast, positive acceptance language showed no meaningful change in either market. The US showed a decline from 5.57% to 5.28% (chi-square = 0.768, $p = 0.381$), and India showed almost no change (from 1.32% to 1.34%, chi-square < 0.001, $p = 1.000$), not statistically significant (Appendix A, Table 13).

A logistic regression (Appendix A, Table 14) examined the probability of loss aversion language appearing in a review as a function of treatment status, post-event period, and their interaction. The treatment main effect was positive and significant (coefficient = 0.2149, $p = 0.012$), indicating that US reviews were more likely to contain loss aversion language than India reviews in baseline. The post-event main effect was negative and significant (coefficient = -0.3122, $p = 0.036$), reflecting the overall decline in loss aversion language across both markets after the price increase. Critically, the treatment-by-post interaction term was not statistically significant (coefficient = 0.0901, $p = 0.562$), meaning that the decline in loss aversion language was not significantly different between the US treatment market and the India control market. This suggests that loss aversion language receded in both markets comparably after the event.

An OLS regression with heteroscedasticity-robust standard errors examined sentiment as a function of treatment, post-event period, loss aversion flag, and all interaction terms (Appendix A, Table 15). The main DiD effect (treatment-by-post) remained negative and significant (coefficient = -0.1009, $p < 0.001$), consistent with prior analysis showing that the price hike reduced sentiment in the US relative to India. Reviews flagged for loss aversion language exhibited substantially lower baseline sentiment (coefficient = -0.1725, $p < 0.001$), validating the semantic alignment of the flag with negative sentiment. Importantly, the three-way interaction (treatment-by-post-by-loss aversion) was positive and statistically significant (coefficient = 0.2553, $p = 0.003$), revealing that the negative sentiment effect of the price hike was significantly mitigated for reviews expressing loss aversion language in the US. This three-way interaction suggests that among US users, those most explicitly voicing loss-aversion concerns experienced a smaller overall sentiment decline relative to non-flagged reviews.

An OLS regression examined sentiment in relation to treatment, post-event period, fairness concerns flag, and all interaction terms (Appendix A, Table 16). The treatment and post-event main effects were positive and significant (coefficient = 0.2236, $p < 0.001$ and coefficient = 0.1006, $p < 0.001$, respectively). The main DiD effect (treatment-by-post) remained negative and statistically significant (coefficient = -0.0814, $p < 0.001$). Reviews flagged for fairness concerns language showed no significant baseline sentiment difference (coefficient = 0.0849, $p = 0.372$), but the two-way interaction between treatment and the fairness flag was negative and significant (coefficient = -0.2362, $p = 0.020$), indicating that fairness language was particularly associated with lower sentiment in the US market specifically. The three-way interaction (treatment-by-post-by-fairness concerns) was not statistically significant (coefficient = 0.2904, $p = 0.198$), suggesting that fairness language did not meaningfully modulate the post-event sentiment change.

An OLS regression assessed user ratings as a function of treatment, post-event period, fairness concerns flag, and interactions (Appendix A, Table 17). The treatment main effect was strong and positive (coefficient = 1.0386, $p < 0.001$), and the post-event effect was also positive (coefficient = 0.3703, $p < 0.001$), indicating higher baseline ratings in the US and in the post-event period. The main DiD effect was negative and significant (coefficient = -0.2384, $p = 0.004$), replicating the sentiment finding at the rating level. The fairness concerns flag itself showed no significant main effect on ratings (coefficient = -0.0642, $p = 0.786$), but the interaction between treatment and fairness flag was negative and highly significant (coefficient = -1.1794, $p < 0.001$), revealing that reviews in the US treatment market expressing fairness concerns assigned substantially lower ratings. The three-way interaction (treatment-by-post-by-fairness concerns) was positive and marginally significant (coefficient = 1.2030, $p = 0.035$), indicating that fairness-focused reviews in the post-event period in the US assigned modestly higher ratings relative to the pre-event baseline, a partial offset to the overall treatment effect.

# CHAPTER 6. CONCLUSIONS AND RECOMMENDATIONS

This thesis examined how Spotify users responded behaviorally to four major strategic transformations: the introduction of audiobooks to Premium plan in November 2023, the Stream On UI/UX redesign in March 2023, the premium subscription price increase in July 2024. These events indicate Spotify's larger business model transformation, examined in the first part of this thesis.

Using difference-in-differences regression analysis combined with behavioral tagging frameworks, this work tried to link consumer sentiment to behavioral economics theory. The findings reveal that some of those Behavioral Economics pillars hold, consumer sentiment cannot be explained by them alone.

## 6.1 Hypothesis 1: Psychological Reactance

The DiD analysis provided some support for H1. After the introduction of audiobooks to Spotify Premium in the US, November 8, user sentiment in the US declined by 0.0399 points relative to control markets India and Canada. The decline is relatively small, but statistically significant.

However, using the behavioral flagging framework, the author observed that the psychological reactance itself isn't observed. Counterintuitively, reactance language expressing forced change decreased from 18.05% to 15.13% in after the audiobooks are introduced ($p < 0.001$).

Since the DiD captured the decline in sentiment, but the behavioral flagging didn't confirm the psychological reactance is the cause of the decline, the underlying research hypothesis H1: Psychological Reactance is rejected.

6.2 Hypothesis 4: Positive Value Framing.

Hypothesis H4, which is a competing hypothesis to H1, receives only partial support. Positive value framing language, when present in reviews, was strongly associated with higher sentiment. However, using the behavioral flagging framework, such language increased only marginally from 34.24% to 35.96% after the introduction of audiobooks. The flags contained audiobook-specific keywords, which means that the addition of audiobooks was received positively, even though the overall consumer sentiment has decreased.

In conclusion, the addition of audiobooks in the US on November 8, 2023 cannot be linked to the overall sentiment decline around this period. The proportion of positive value framing language increased, suggesting an overall positive reception of this feature.

6.3 Hypothesis 2: Status Quo Bias

Hypothesis H2 receives strong empirical support. The DiD analysis revealed that following the March 8, 2023 Stream On user interface and user experience redesign, US sentiment declined only by 0.0168. However, the status quo bias language expressing preference for the old design or resistance to change increased from 5.88% to 9.43% of reviews post-event. A logistic regression analysis showed that such language was 67% more likely to occur after the redesign .

This demonstrates that major user interface changes trigger explicit, measurable user backlash. An attempt to measure the opposite trend was made by measuring the language with “openness to change” keywords. Such a trend didn’t manifest itself in the data with such language proportion remaining virtually flat over the course of the event.

The analysis confirms that status quo bias is a robust consumer response to digital product changes. Spotify users didn’t appreciate their app suddenly becoming unfamiliar to navigate with a sudden learning curve to understand how to operate the app lying ahead of them. This finding has implications product managers: interface changes should anticipate user resistance and may benefit from gradual rollouts or hybrid transitional periods rather than abrupt replacement of familiar navigation patterns.

## 6.4 Hypothesis 3: Loss Aversion and Fairness Concerns

Hypothesis H3 is supported through DiD, but similarly to H1 has important caveats regarding the behavioral language patterns (behavioral flags). DiD reveals that the July 24, 2024 price increase resulted in a clear negative sentiment effect: US sentiment declined by 0.0756 points relative to the control market of India.

When behavioral flagging framework was applied to this dataset, it revealed a surprising result. Loss aversion language declined from 13.67% before the event to 11.25% in US reviews after the price increase. Fairness concerns language decreased from 3.10% to 2.52% post-event. Positive acceptance language showed no meaningful change.

This hypothesis' analysis employed a three-way interaction model which revealed an important effect. Loss aversion language provided a protective effect on overall sentiment. When loss aversion language was present, the negative sentiment impact of the price increase was significantly mitigated (coefficient 0.2553, p 0.003). This suggests that among users who explicitly expressed loss aversion, their sentiment decline was lower compared to reviews without such text.

Overall, loss aversion and fairness concerns language became less common following the price increase in both Indian and US market. This hints at a possible omitted variable bias for this specific dataset. Behavioral language wasn't able to explain the decline in the sentiment. Thus, the third hypothesis H3 is supported partially by the DiD model.

## 6.5 Cross-hypotheses conclusion

Several important patterns emerge when considering the four hypotheses together. First, explicit behavioral language and overall sentiment often diverge. In H1, overall sentiment declined while explicit reactance language counterintuitively decreased. In H3, overall sentiment declined while loss aversion language declined. Such a diversion shows that consumer sentiment might be hard to capture through simple sentiment analysis tools. Users can express their dissatisfaction or satisfaction through language that is not typical for established behavioral science mechanisms.

Second, positive framing and enthusiasm for changes emerge weakly. In H1, positive value framing increased only marginally and failed to change the overall decline in sentiment. In H2, openness to change language remained essentially flat despite a major redesign. This shows that users are more prone to criticism during major platform changes, especially when their established usage patterns are being disrupted through app redesign.

## 6.6 Practical Implications

The results of this research are useful for digital platform product managers, business leaders considering major changes to digital platforms and academic researchers who are interested in similar area of research.

First, major platform changes will likely trigger negative sentiment regardless of the change.  The analysis consistency suggests that platform-level sentiment declines are a predictable consequence of substantial strategic changes. Product managers should always anticipate such reactions, even if prior A/B tests are showing otherwise. When considering major redesigns, decision makers should not only rely on performance metrics, such as click-rate, acceptance rate, etc., but also consider the underlying behavioral mechanisms, the most important of which this analysis reveals to be loss aversion. Sentiment declines alone shouldn't be interpreted as indicating failure when the change is introduced. Managers should therefore establish appropriate benchmarks for acceptable sentiment decline around major events and avoid reactive decisions based on short-term sentiment declines.

Second, positive framing of changes is not enough to counteract negative sentiment, especially when changes impose costs or restrictions on user autonomy. A marginal 1.73 percentage point increase in positive value framing following the audiobooks launch failed to offset the overall sentiment decline. Product managers might consider: providing transition periods that preserve prior functionality alongside new features, gradual rollouts that allow user adaptation. Transparent communication about changes is also extremely important.

Third, price increases trigger clear negative sentiment effects that cannot be offset by feature addition. Despite adding more audiobook hours, the 2024 price increase still triggered a decline in sentiment. Loss aversion proves to be the most important behavioral mechanism and platform managers should consider this factor when, inevitably, they have to reconsider the subscription increase. As analyzed in the first part of the thesis, Spotify's business model still relies heavily on investments, so future price increases are also imminent.

Fifth, sentiment analysis of user reviews provides important information that should be analyzed beyond the simple user rating. J-shaped product rating distribution holds true for Spotify, however important insights emerge with a more granular sentiment analysis. Research similar or adjacent to this thesis may contribute to developing new tools for sentiment analysis that can be implemented platform-wide or as a third-party review analysis.

## 6.7 Limitations and Future Research

There are several limitations in this research. Firstly, as discussed in the academic literature, user reviews are inherently biased. Users tend to leave reviews when they are extremely dissatisfied or satisfied with their experience, hence the J-Shaped distribution (Hu, et. al., 2018). Thus, when analyzing review datasets such bias should always be considered.

The behavioral flagging framework is limited. Due to short keyword dictionaries, presence of unexpected language and sarcasm it cannot fully capture the complexity of user responses to platform transformation. While behavioral flags successfully identified status quo bias language in H2 (increasing from 5.88% to 9.43%), the same framework failed to explain sentiment declines in H1 and H3. Future research might use this method with extended dictionaries, or use machine learning algorithms to perform a more robust textual analysis.

Another limiting factor that emerged during the research is the low ability to establish event-sentiment causality. Review datasets also suffer from the omitted variable bias, as the sample selection is broad. For example, external events like the Billie Eilish Olympics performance noted in the H3 analysis can influence sentiment independent of platform changes. Future research should incorporate additional control variables.

The four events studied represent only a subset of possible platform transformations and changes. Since the event sample selection is limited, the research results are most directly applicable to consumer-facing digital platforms engaging in feature expansion, interface redesign, or pricing changes.

Future research directions might include: extending or narrowing down the event selection; extending the observation window; adding platform metrics like user engagement and conversion to the dataset (might prove difficult due to data availability).

Additionally, researchers interested in a similar field might use surveys and questionaries to asses deeper user engagement patterns in addition to sentiment analysis.

REFERENCES


Alexander, J. (2023, October 24). Spotify's podcast bet is finally paying off. *The Verge*. https://www.theverge.com/2023/10/24/23929885/spotify-podcasts-profit-q3-2023

Anderson, A., Maystre, L., Anderson, I., Mehrotra, R., & Lalmas, M. (2020). Algorithmic effects on the diversity of consumption on Spotify. *Proceedings of The Web Conference 2020*, 2155–2165. https://doi.org/10.1145/3366423.3380281

Barclays**.** (2025). *January 2025: Key spend trends | UK unlocked*. https://home.barclays/insights/2025/02/January-2025-spend-trends/

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology, 5*(4), 323–370. https://doi.org/10.1037/1089-2680.5.4.323

Brehm, J. W. (1966). *A theory of psychological reactance*. Academic Press.

Business Insider. (2025). The streaming wars are over. The rich won. *Business Insider*. https://www.businessinsider.com/streaming-tv-movies-music-expensive-unequal-spotify-netflix-taylor-swift-2025-3

Business Insider. (2025a). Coverage critiquing Spotify's monetization and strategy. *Business Insider*. https://www.businessinsider.com (use the precise article URL you cited)

Business of Apps. (2024). Music streaming market share. https://www.businessofapps.com/data/music-streaming-market/

Clausius Press. (2023). *The evolution of the music industry in the digital age: From records to streaming*.

https://www.clausiuspress.com/assets/default/article/2023/11/27/article_1701102221.pdf

Coherent Market Insights. (2025). *Audiobooks market trends, share and forecast, 2025–2032*. https://www.coherentmarketinsights.com/industry-reports/audiobooks-market

Coherent Market Insights. (2025, April 9). Podcasting market size to worth USD 173.49 billion by 2032. *GlobeNewswire*. https://www.globenewswire.com/news-release/2025/04/09/3058669/0/en/Podcasting-Market-size-to-worth-USD-173-49-Billion-by-2032-says-Coherent-Market-Insights.html

Datta, H., Knox, G., & Bronnenberg, B. J. (2017). Changing their tune: How consumers' adoption of online streaming affects music consumption. *Marketing Science, 37*(1), 5–21. https://doi.org/10.1287/mksc.2017.1051

Deloitte. (2023). *Digital media trends survey, 17th edition*. Deloitte Insights. https://www.deloitte.com/global

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General, 144*(1), 114–126. https://doi.org/10.1037/xge0000033

Exploding Topics. (2024, December 10). Music streaming services stats (2025). https://explodingtopics.com/blog/music-streaming-stats

GMI Insights. (2025). *Audiobooks market size & share forecast report 2025–2034*. https://www.gminsights.com/industry-analysis/audiobooks-market

GlobeNewswire. (2025, March 14). Audiobooks market report 2025: Global & country-level trends and forecasts to 2030. *GlobeNewswire.* https://www.globenewswire.com/news-release/2025/03/14/3042894/28124/en/Audiobooks-Market-Report-2025-Global-Country-Level-Trends-and-Forecasts-to-2030-by-Content-Type-Language-Format-Technology-Pricing-Model-Target-Audience-Sales-Channel-and-End-User.html

Grand View Research. (2023). *Audiobooks market size, share & trends analysis report, 2023–2030*. https://www.grandviewresearch.com/industry-analysis/audiobooks-market

Grand View Research. (2024). *Audiobooks market size & share industry report, 2030*. https://www.grandviewresearch.com/industry-analysis/audiobooks-market

Guzman, E., & Maalej, W. (2014). How do users like this feature? A fine grained sentiment analysis of app reviews. *2014 IEEE 22nd International Requirements Engineering Conference*, 153–162. https://doi.org/10.1109/RE.2014.6912257

Harvard Business Review. See Johnson (2021).

Hern, A. (2021, November 11). Spotify acquires audiobook platform Findaway as it expands audio offerings. *The Guardian*. https://www.theguardian.com/technology/2021/nov/11/spotify-acquires-audiobook-platform-findaway

Holtz, D., Carterette, B., Chandar, P., Nazari, Z., Cramer, H., & Aral, S. (2020). The engagement–diversity connection: Evidence from a field experiment on Spotify. *Proceedings of the ACM Conference on Economics and Computation*. https://doi.org/10.1145/3391403.3399482

Hu, N., Pavlou, P. A., & Zhang, J. (2009). Overcoming the J-shaped distribution of product reviews. *Communications of the ACM, 52*(10), 144–147. https://doi.org/10.1145/1562764.1562800

Hracs, B. J., & Webster, J. (2020). The Spotify effect? Digital platforms and the transformation of the music industry. In *The Oxford Handbook of the Creative Industries* (pp. xxx–xxx). Oxford University Press. (add full chapter details if used)

Hutto, C. J., & Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the Eighth International Conference on Weblogs and Social Media* (pp. 216–225). https://ojs.aaai.org/index.php/ICWSM/article/view/14550

IFPI. (2025). *Global music report 2025: State of the industry*. International Federation of the Phonographic Industry. https://www.ifpi.org

Investing.com. (2025). Spotify revenue mix and subscription share. *Investing.com*. https://www.investing.com (use the precise article URL you cited)

Insider Intelligence. (2023). *Podcast listener forecast 2023*. https://www.insiderintelligence.com

Johnson, M. (2021, September). How Spotify's algorithm manages artists and listeners. *Harvard Business Review*. https://hbr.org/2021/09/how-spotifys-algorithm-manages-artists-and-listeners

Johnson, E. J., & Goldstein, D. (2003). Do defaults save lives? *Science, 302*(5649), 1338–1339. https://doi.org/10.1126/science.1091721

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263–291. https://doi.org/10.2307/1914185

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1986). Fairness and the assumptions of economics. *Journal of Business, 59*(4), S285–S300. https://doi.org/10.1086/296367

Kiberg, H., & Spilker, H. (2023). One more turn after the algorithmic turn? Spotify's colonization of the online audio space. *Popular Music and Society, 46*(5), 473–493. https://doi.org/10.1080/03007766.2023.2184160

Kumar, S. (2025). Apple Music monthly active users. *Tech market brief.* (replace with the exact source you cited)

Li, X. (2024). Strategic analysis of Spotify's content diversification: A SWOT perspective. *Journal/Conference name.* (add full details if used)

MacKinlay, A. C. (1997). Event studies in economics and finance. *Journal of Economic Literature, 35*(1), 13–39.

Martin, W., Sarro, F., Jia, Y., Zhang, Y., & Harman, M. (2018). A survey of app store analysis for software engineering. *IEEE Transactions on Software Engineering, 43*(9), 817–847. (If you used a specific "bad updates" paper, replace with that exact citation.)

Mei, Y. (2024). Spotify's platform strategy: SWOT analysis and implications. *Business and Management Review.* (add volume, issue, pages)

Muchnik, L., Aral, S., & Taylor, S. J. (2013). Social influence bias: A randomized experiment. *Science, 341*(6146), 647–651. https://doi.org/10.1126/science.1240466

Music Ally. (2025). Spotify ad revenues and Q2 2025 performance. *Music Ally.* https://musically.com (use the precise article URL you cited)

Music Business Worldwide. (2025). Discovery Mode and artist payout debates. *Music Business Worldwide.* https://www.musicbusinessworldwide.com (use the precise URL)

OpenAI. (2023). *ChatGPT* (Mar 14 version) [Large language model]. https://chat.openai.com/chat

Ozili, P. K. (2023). The acceptable R-square in Empirical Modelling for Social Science Research. *Advances in Knowledge Acquisition, Transfer, and Management*, 134–143. https://doi.org/10.4018/978-1-6684-6859-3.ch009

Pandey, P. (2019, November 8). *Simplifying sentiment analysis using Vader in Python (on social media text)*. Medium. https://medium.com/analytics-vidhya/simplifying-social-media-sentiment-analysis-using-vader-in-python-f9e6ec6fc52f

Patel, N. (2022). Risks of platform overextension in streaming. *Tech policy review*. (replace with the exact piece you cited)

Pelly, L. (2025). Spotify as a "mood machine" and the politics of playlists. *Publication name*. (add URL if you cited a 2025 article; the classic piece is 2018 in *The Baffler*)

Pérez-Alaejos, M., Terol-Bolinches, R., & Barrios-Rubio, A. (2022). Exclusive podcast agreements and platform strategy: The Spotify case. *Journal/Conference name*.

Perplexity AI. (2023). *Perplexity AI* [AI language model]. https://www.perplexity.ai

Raj, A. (2021). Algorithmic popularity and competition among creators on Spotify. *Platform Studies Journal*. (add full details)

Reuters. (2025). Spotify's product experiments and financial outlook. *Reuters*. https://www.reuters.com (use the precise article URL you cited)

Samuelson, W., & Zeckhauser, R. (1988). Status quo bias in decision making. *Journal of Risk and Uncertainty, 1*(1), 7–59. https://doi.org/10.1007/BF00055564

Saidani, M., et al. (2022). App updates, user sentiment, and disruption shocks: Evidence from Google Play. *Proceedings/Journal*. (add full details)

Shampanier, K., Mazar, N., & Ariely, D. (2007). Zero as a special price: The true value of free products. *Marketing Science, 26*(6), 742–757. https://doi.org/10.1287/mksc.1060.0268

Smith, A. (2023). Core user alienation risks in platform pivots. *Media strategy review*. (replace with the exact piece you cited)

Spotify. (2023). Q3 2023 shareholder letter. *Investor relations*. https://investors.spotify.com

Spotify. (2024, June 3). Adjusting Spotify Premium prices in the U.S. *Spotify Newsroom*. https://newsroom.spotify.com/2024-06-03/adjusting-spotify-premium-prices-in-the-us/

Spotify. (2024). Annual and quarterly financial results. *Investor relations*. https://investors.spotify.com

Statista. (2025). Audiobook market revenue worldwide. https://www.statista.com

Strategic Management Insight. (2025). Spotify cost structure and margins. https://www.strategicmanagementinsight.com (use the precise article URL you cited)

Thaler, R. H. (1999). Mental accounting matters. *Journal of Behavioral Decision Making, 12*(3), 183–206. https://doi.org/10.1002/(SICI)1099-0771(199909)12:3<183::AID-BDM318>3.0.CO;2-F

The Guardian. (2025). Artists remove catalogues over Daniel Ek's defense investments. *The Guardian*. https://www.theguardian.com (use the precise article URL you cited)

The Wall Street Journal. (2025, August 4). Spotify to raise subscription prices around the world amid chase for consistent profit. *The Wall Street Journal*.

https://www.wsj.com/business/media/spotify-to-raise-subscription-prices-around-the-world-amid-chase-for-consistent-profit-ff229b54

Wang, J. (2023). A critical research of Spotify's business model: The case of Discover Weekly. *BCP Social Sciences & Humanities, 21*, 790–795. https://doi.org/10.54691/bcpssh.v21i.3873

Warren, T. (2024, June 3). Spotify is increasing US prices again. *The Verge*. https://www.theverge.com/2024/6/3/24170301/spotify-us-price-increase-plans

Wikipedia contributors. (2025, September). Spotify. In *Wikipedia*. https://en.wikipedia.org/wiki/Spotify

Yellowbrick. (2024). The evolution of the music industry: A comprehensive overview. https://www.yellowbrick.co/blog/music/the-evolution-of-the-music-industry-a-comprehensive-overview-2

a

# APPENDIX A

## PYTHON OUTPUT

Table 1

```
Summary statistics by Country:
       vader_neg                                                  vader_neu  \
           count      mean       std  min  25%    50%    75%  max      count
Country
ca        1827.0  0.056544  0.102473  0.0  0.0  0.000  0.084  1.0     1827.0
in        5077.0  0.128297  0.158036  0.0  0.0  0.088  0.190  1.0     5077.0
us       18784.0  0.051997  0.094983  0.0  0.0  0.000  0.077  1.0    18784.0

                  ... vader_compound          Rating                     \
             mean ...            75%     max   count      mean       std
Country           ...
ca       0.732106 ...         0.7650  0.9993  1827.0  3.405583  1.669989
in       0.751581 ...         0.4571  0.9974  5077.0  1.651763  1.321484
us       0.722384 ...         0.8151  0.9999 18784.0  3.720826  1.563847


         min  25%  50%  75%  max
Country
ca       1.0  1.5  4.0  5.0  5.0
in       1.0  1.0  1.0  1.0  5.0
us       1.0  2.0  5.0  5.0  5.0
```

Table 2

```
Summary Statistics (Ratings and VADER Sentiments):
             Rating     vader_neg     vader_neu     vader_pos  vader_compound
count  25688.000000  25688.000000  25688.000000  25688.000000    25688.000000
mean       3.289474      0.067401      0.728846      0.203755        0.306908
std        1.731696      0.114863      0.187774      0.193789        0.534335
min        1.000000      0.000000      0.000000      0.000000       -0.999900
25%        1.000000      0.000000      0.627000      0.046000        0.000000
50%        4.000000      0.000000      0.751000      0.160000        0.440400
75%        5.000000      0.100000      0.851000      0.308000        0.777500
max        5.000000      1.000000      1.000000      1.000000        0.999900
```

Table 3

```
Summary Statistics (Ratings and VADER Sentiments):

                  count      mean       std     min    25%     50%     75%  \
Rating          26361.0  3.625394  1.637406  1.0000  2.000  5.0000  5.0000
vader_neg       26361.0  0.052895  0.097004  0.0000  0.000  0.0000  0.0790
vader_neu       26361.0  0.711152  0.211176  0.0000  0.607  0.7430  0.8480
vader_pos       26361.0  0.235951  0.220847  0.0000  0.069  0.1850  0.3430
vader_compound  26361.0  0.374465  0.502231 -0.9975  0.000  0.5284  0.7964

                   max
Rating          5.0000
vader_neg       1.0000
vader_neu       1.0000
vader_pos       1.0000
vader_compound  0.9999
```

| | count | mean | std | min | 25% | 50% | 75% | max |
|---|---|---|---|---|---|---|---|---|
| Rating | 26361 | 3.62539 | 1.63741 | 1 | 2 | 5 | 5 | 5 |
| vader_neg | 26361 | 0.0528948 | 0.0970045 | 0 | 0 | 0 | 0.079 | 1 |
| vader_neu | 26361 | 0.711152 | 0.211176 | 0 | 0.607 | 0.743 | 0.848 | 1 |
| vader_pos | 26361 | 0.235951 | 0.220847 | 0 | 0.069 | 0.185 | 0.343 | 1 |
| vader_compound | 26361 | 0.374465 | 0.502231 | -0.9975 | 0 | 0.5284 | 0.7964 | 0.9999 |

Table 4

| | count | mean | std | min | 25% | 50% | 75% | max |
|---|---|---|---|---|---|---|---|---|
| neg | 22003.0 | 0.058019 | 0.111625 | 0.0000 | 0.000 | 0.0000 | 0.0845 | 1.0 |
| neu | 22003.0 | 0.707579 | 0.233119 | 0.0000 | 0.589 | 0.7420 | 0.8670 | 1.0 |
| pos | 22003.0 | 0.230447 | 0.236086 | 0.0000 | 0.000 | 0.1680 | 0.3480 | 1.0 |
| compound | 22003.0 | 0.322810 | 0.497684 | -0.9999 | 0.000 | 0.4404 | 0.7482 | 1.0 |
| Rating | 22003.0 | 3.581375 | 1.642886 | 1.0000 | 2.000 | 4.0000 | 5.0000 | 5.0 |

Date range: 2024-04-01 05:26:52 to 2024-09-30 23:48:54

Table 5

OLS Regression Results

| | | | |
|---|---|---|---|
| Dep. Variable: | vader_compound | R-squared: | 0.222 |
| Model: | OLS | Adj. R-squared: | 0.222 |
| Method: | Least Squares | F-statistic: | 665.4 |
| Date: | Fri, 31 Oct 2025 | Prob (F-statistic): | 0.00 |
| Time: | 13:10:33 | Log-Likelihood: | -17122. |
| No. Observations: | 25688 | AIC: | 3.427e+04 |
| Df Residuals: | 25675 | BIC: | 3.438e+04 |
| Df Model: | 12 | | |
| Covariance Type: | HC1 | | |

| | coef | std err | z | P>\|z\| | [0.025 | 0.975] |
|---|---|---|---|---|---|---|
| Intercept | 0.2455 | 0.014 | 17.398 | 0.000 | 0.218 | 0.273 |
| C(Country)[T.in] | -0.2849 | 0.014 | -20.394 | 0.000 | -0.312 | -0.258 |
| C(Country)[T.us] | 0.0305 | 0.007 | 4.090 | 0.000 | 0.016 | 0.045 |
| treatment | 0.0305 | 0.007 | 4.090 | 0.000 | 0.016 | 0.045 |
| post | 0.0326 | 0.015 | 2.203 | 0.028 | 0.004 | 0.062 |
| treatment:post | -0.0480 | 0.018 | -2.642 | 0.008 | -0.084 | -0.012 |
| flag_h1_reactance | -0.1384 | 0.022 | -6.377 | 0.000 | -0.181 | -0.096 |
| flag_h4_positive | 0.3074 | 0.023 | 13.515 | 0.000 | 0.263 | 0.352 |
| treatment:flag_h1_reactance | -0.1415 | 0.025 | -5.662 | 0.000 | -0.191 | -0.093 |
| post:flag_h1_reactance | -0.0168 | 0.036 | -0.470 | 0.638 | -0.087 | 0.053 |
| treatment:post:flag_h1_reactance | 0.0247 | 0.045 | 0.546 | 0.585 | -0.064 | 0.113 |
| treatment:flag_h4_positive | 0.0946 | 0.024 | 3.947 | 0.000 | 0.048 | 0.142 |
| post:flag_h4_positive | 0.0125 | 0.035 | 0.362 | 0.717 | -0.055 | 0.080 |
| treatment:post:flag_h4_positive | -0.0355 | 0.038 | -0.926 | 0.354 | -0.111 | 0.040 |

| | | | |
|---|---|---|---|
| Omnibus: | 1038.759 | Durbin-Watson: | 1.993 |
| Prob(Omnibus): | 0.000 | Jarque-Bera (JB): | 1168.076 |
| Skew: | -0.521 | Prob(JB): | 2.27e-254 |
| Kurtosis: | 2.924 | Cond. No. | 1.42e+16 |

Table 6

```
========================================
Testing: flag_h1_reactance
========================================

US - flag_h1_reactance:
Pre-event proportion: 0.1805
Post-event proportion: 0.1513
Change: -0.0292 (-2.92%)
Chi-square: 18.4191, p-value: 0.0000
Significant at α=0.05: Yes

IN - flag_h1_reactance:
Pre-event proportion: 0.2430
Post-event proportion: 0.2069
Change: -0.0360 (-3.60%)
Chi-square: 8.0837, p-value: 0.0045
Significant at α=0.05: Yes

CA - flag_h1_reactance:
Pre-event proportion: 0.1785
Post-event proportion: 0.1532
Change: -0.0254 (-2.54%)
Chi-square: 1.9439, p-value: 0.1632
Significant at α=0.05: No
```

Table 7

```
=====================================
Testing: flag_h4_positive
=====================================

US - flag_h4_positive:
Pre-event proportion: 0.3424
Post-event proportion: 0.3596
Change: 0.0173 (1.73%)
Chi-square: 4.0839, p-value: 0.0433
Significant at α=0.05: Yes

IN - flag_h4_positive:
Pre-event proportion: 0.1483
Post-event proportion: 0.1372
Change: -0.0112 (-1.12%)
Chi-square: 1.0525, p-value: 0.3049
Significant at α=0.05: No

CA - flag_h4_positive:
Pre-event proportion: 0.3034
Post-event proportion: 0.2899
Change: -0.0135 (-1.35%)
Chi-square: 0.3348, p-value: 0.5629
Significant at α=0.05: No
```

Table 8

```
                           Logit Regression Results
==============================================================================
Dep. Variable:      flag_h1_reactance   No. Observations:                18784
Model:                          Logit   Df Residuals:                    18782
Method:                           MLE   Df Model:                            1
Date:                Fri, 31 Oct 2025   Pseudo R-squ.:                0.001101
Time:                        13:26:20   Log-Likelihood:                -8680.6
converged:                       True   LL-Null:                       -8690.1
Covariance Type:            nonrobust   LLR p-value:                 1.221e-05
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     -1.5129      0.021    -70.740      0.000      -1.555      -1.471
post          -0.2112      0.049     -4.309      0.000      -0.307      -0.115
==============================================================================
```

Table 9

```
Status Quo Bias:
  Pre-event proportion: 0.0588
  Post-event proportion: 0.0943
  Change: 0.0355 (3.55%)
  Chi-square: 111.9510, p-value: 0.0000
  Significant: Yes

Openness to Change:
  Pre-event proportion: 0.0585
  Post-event proportion: 0.0573
  Change: -0.0012 (-0.12%)
  Chi-square: 0.1398, p-value: 0.7085
  Significant: No

Status Quo Bias Proportion Z-test:
  z = 10.604, p = 0.0000
  Pre-post change: 0.0355 (3.55%)
  Significant: Yes

Openness to Change Proportion Z-test:
  z = -0.401, p = 0.6888
  Pre-post change: -0.0012 (-0.12%)
  Significant: No
```

Table 10

```
LOGIT REGRESSION (pre-post) for Status Quo Bias:
                           Logit Regression Results
==============================================================================
Dep. Variable:     flag_h2_status_quo_bias   No. Observations:            26361
Model:                               Logit   Df Residuals:                26359
Method:                                MLE   Df Model:                        1
Date:                     Sat, 01 Nov 2025   Pseudo R-squ.:            0.007938
Time:                             09:15:57   Log-Likelihood:            -7212.8
converged:                            True   LL-Null:                   -7270.5
Covariance Type:                 nonrobust   LLR p-value:             6.339e-27
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     -2.7736      0.040    -69.911      0.000      -2.851      -2.696
post           0.5108      0.049     10.512      0.000       0.416       0.606
==============================================================================
```

Table 11

```
LOGIT REGRESSION (pre-post) for Openness to Change:
                           Logit Regression Results
==============================================================================
Dep. Variable:     flag_h2_openness_to_change   No. Observations:         26361
Model:                                  Logit   Df Residuals:             26359
Method:                                   MLE   Df Model:                     1
Date:                        Sat, 01 Nov 2025   Pseudo R-squ.:        1.376e-05
Time:                                09:15:57   Log-Likelihood:         -5826.0
converged:                               True   LL-Null:                -5826.1
Covariance Type:                    nonrobust   LLR p-value:             0.6889
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     -2.7783      0.040    -69.884      0.000      -2.856      -2.700
post          -0.0213      0.053     -0.401      0.689      -0.125       0.083
==============================================================================
```

Table 12

```
==============================================================================
OPTION 4: Weekly Aggregated DiD
==============================================================================
                            OLS Regression Results
==============================================================================
Dep. Variable:         vader_compound   R-squared:                       0.027
Model:                            OLS   Adj. R-squared:                 -0.019
Method:                 Least Squares   F-statistic:                     1206.
Date:                Sat, 01 Nov 2025   Prob (F-statistic):           2.13e-22
Time:                        08:56:04   Log-Likelihood:                 44.246
No. Observations:                  23   AIC:                            -84.49
Df Residuals:                      21   BIC:                            -82.22
Df Model:                           1
Covariance Type:                  HC1
==================================================================================
                     coef    std err          z      P>|z|      [0.025      0.975]
----------------------------------------------------------------------------------
Intercept          0.1927      0.006     32.188      0.000       0.181       0.204
treatment          0.1927      0.006     32.188      0.000       0.181       0.204
post              -0.0061      0.008     -0.776      0.438      -0.021       0.009
treatment:post    -0.0061      0.008     -0.776      0.438      -0.021       0.009
==============================================================================
Omnibus:                        1.451   Durbin-Watson:                   2.039
Prob(Omnibus):                  0.484   Jarque-Bera (JB):                1.037
Skew:                           0.235   Prob(JB):                        0.595
Kurtosis:                       2.072   Cond. No.                     2.28e+17
==============================================================================
```

Table 13

```
=== Loss Aversion FLAG ANALYSIS ===
US: Pre 0.1367, Post 0.1125, Δ -0.0242, chi2=25.165, p=0.0000
    Z-test: z=-5.04, p=0.0000
IN: Pre 0.1132, Post 0.0855, Δ -0.0278, chi2=4.121, p=0.0423
    Z-test: z=-2.10, p=0.0356

=== Fairness Concerns FLAG ANALYSIS ===
US: Pre 0.0310, Post 0.0252, Δ -0.0058, chi2=5.527, p=0.0187
    Z-test: z=-2.39, p=0.0167
IN: Pre 0.0298, Post 0.0122, Δ -0.0176, chi2=6.384, p=0.0115
    Z-test: z=-2.67, p=0.0076

=== Positive Acceptance FLAG ANALYSIS ===
US: Pre 0.0557, Post 0.0528, Δ -0.0030, chi2=0.768, p=0.3807
    Z-test: z=-0.91, p=0.3637
IN: Pre 0.0132, Post 0.0134, Δ 0.0002, chi2=0.000, p=1.0000
    Z-test: z=0.04, p=0.9702

=== LOGIT DiD FOR LOSS AVERSION ===
```

Table 14

```
=== LOGIT DiD FOR LOSS AVERSION ===
                           Logit Regression Results
==============================================================================
Dep. Variable:     flag_h3_loss_aversion   No. Observations:            22003
Model:                             Logit   Df Residuals:                21999
Method:                              MLE   Df Model:                        3
Date:                   Sat, 01 Nov 2025   Pseudo R-squ.:            0.002471
Time:                           10:28:07   Log-Likelihood:            -8228.0
converged:                          True   LL-Null:                   -8248.4
Covariance Type:               nonrobust   LLR p-value:             7.356e-09
==================================================================================
                     coef    std err          z      P>|z|      [0.025      0.975]
----------------------------------------------------------------------------------
Intercept         -2.0580      0.081    -25.342      0.000      -2.217      -1.899
treatment          0.2149      0.086      2.508      0.012       0.047       0.383
post              -0.3122      0.149     -2.095      0.036      -0.604      -0.020
treatment:post     0.0901      0.155      0.579      0.562      -0.215       0.395
==================================================================================
```

```
=== LOGIT DiD FOR FAIRNESS CONCERNS ===
                              Logit Regression Results
==================================================================================
Dep. Variable:     flag_h3_fairness_concerns   No. Observations:            22003
Model:                                 Logit   Df Residuals:                21999
Method:                                  MLE   Df Model:                        3
Date:                       Sat, 01 Nov 2025   Pseudo R-squ.:            0.002787
Time:                               10:28:08   Log-Likelihood:            -2805.6
converged:                              True   LL-Null:                   -2813.5
Covariance Type:                   nonrobust   LLR p-value:              0.001319
==================================================================================
                     coef    std err          z      P>|z|      [0.025      0.975]
----------------------------------------------------------------------------------
Intercept         -3.4829      0.151    -23.014      0.000      -3.780      -3.186
treatment          0.0407      0.161      0.253      0.800      -0.274       0.356
post              -0.9103      0.352     -2.584      0.010      -1.601      -0.220
treatment:post     0.6989      0.363      1.924      0.054      -0.013       1.411
==================================================================================
```

```
=== LOGIT DiD FOR POSITIVE ACCEPTANCE ===
                               Logit Regression Results
=====================================================================================
Dep. Variable:     flag_h3_positive_acceptance   No. Observations:            22003
Model:                                   Logit   Df Residuals:                21999
Method:                                    MLE   Df Model:                        3
Date:                         Sat, 01 Nov 2025   Pseudo R-squ.:             0.01154
Time:                                 10:28:08   Log-Likelihood:            -4325.8
converged:                                True   LL-Null:                   -4376.3
Covariance Type:                     nonrobust   LLR p-value:             9.282e-22
==================================================================================
                     coef    std err          z      P>|z|      [0.025      0.975]
----------------------------------------------------------------------------------
Intercept         -4.3108      0.225    -19.150      0.000      -4.752      -3.870
treatment          1.4812      0.229      6.474      0.000       1.033       1.930
post               0.0141      0.378      0.037      0.970      -0.727       0.755
treatment:post    -0.0721      0.383     -0.188      0.851      -0.823       0.679
==================================================================================
```

Table 15

```
=== OLS: Compound ~ treatment * post * Loss Aversion ===
                            OLS Regression Results
==============================================================================
Dep. Variable:               compound   R-squared:                       0.053
Model:                            OLS   Adj. R-squared:                  0.053
Method:                 Least Squares   F-statistic:                     129.0
Date:                Sat, 01 Nov 2025   Prob (F-statistic):          8.17e-187
Time:                        10:28:08   Log-Likelihood:                -15267.
No. Observations:               22003   AIC:                         3.055e+04
Df Residuals:                   21995   BIC:                         3.061e+04
Df Model:                           7
Covariance Type:                  HC1
=========================================================================================================
                                            coef    std err          z      P>|z|      [0.025      0.975]
---------------------------------------------------------------------------------------------------------
Intercept                                 0.1379      0.014     10.191      0.000       0.111       0.164
treatment                                 0.2389      0.014     16.714      0.000       0.211       0.267
post                                      0.1099      0.022      5.104      0.000       0.068       0.152
treatment:post                           -0.1009      0.023     -4.457      0.000      -0.145      -0.057
flag_h3_loss_aversion                    -0.1725      0.048     -3.614      0.000      -0.266      -0.079
treatment:flag_h3_loss_aversion          -0.1355      0.051     -2.669      0.008      -0.235      -0.036
post:flag_h3_loss_aversion               -0.2258      0.082     -2.753      0.006      -0.387      -0.065
treatment:post:flag_h3_loss_aversion      0.2553      0.087      2.948      0.003       0.086       0.425
==============================================================================
Omnibus:                     1741.487   Durbin-Watson:                   1.992
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             1128.326
Skew:                          -0.434   Prob(JB):                    9.71e-246
Kurtosis:                       2.310   Cond. No.                         50.6
==============================================================================
```

Table 16

```
=== OLS: Compound ~ treatment * post * Fairness Concerns ===
                            OLS Regression Results
==============================================================================
Dep. Variable:               compound   R-squared:                       0.018
Model:                            OLS   Adj. R-squared:                  0.017
Method:                 Least Squares   F-statistic:                     52.83
Date:                Sat, 01 Nov 2025   Prob (F-statistic):           3.11e-75
Time:                        10:28:08   Log-Likelihood:                -15670.
No. Observations:               22003   AIC:                         3.136e+04
Df Residuals:                   21995   BIC:                         3.142e+04
Df Model:                           7
Covariance Type:                  HC1
==============================================================================================================
                                                coef    std err          z      P>|z|      [0.025      0.975]
--------------------------------------------------------------------------------------------------------------
Intercept                                     0.1158      0.013      8.753      0.000       0.090       0.142
treatment                                     0.2236      0.014     15.917      0.000       0.196       0.251
post                                          0.1006      0.021      4.712      0.000       0.059       0.142
treatment:post                               -0.0814      0.022     -3.617      0.000      -0.125      -0.037
flag_h3_fairness_concerns                     0.0849      0.095      0.893      0.372      -0.101       0.271
treatment:flag_h3_fairness_concerns          -0.2362      0.101     -2.334      0.020      -0.435      -0.038
post:flag_h3_fairness_concerns               -0.3019      0.218     -1.384      0.166      -0.729       0.126
treatment:post:flag_h3_fairness_concerns      0.2904      0.226      1.287      0.198      -0.152       0.732
==============================================================================
Omnibus:                     1797.471   Durbin-Watson:                   1.996
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             1562.709
Skew:                          -0.578   Prob(JB):                         0.00
Kurtosis:                       2.394   Cond. No.                         114.
==============================================================================
```

Table 17

```
=== OLS: Rating ~ treatment * post * Fairness Concerns ===
                            OLS Regression Results
==============================================================================
Dep. Variable:                 Rating   R-squared:                       0.047
Model:                            OLS   Adj. R-squared:                  0.046
Method:                 Least Squares   F-statistic:                     144.2
Date:                Sat, 01 Nov 2025   Prob (F-statistic):          7.66e-209
Time:                        10:28:09   Log-Likelihood:                -41619.
No. Observations:               22003   AIC:                         8.325e+04
Df Residuals:                   21995   BIC:                         8.332e+04
Df Model:                           7
Covariance Type:                  HC1
===========================================================================================================
                                              coef    std err          z      P>|z|      [0.025      0.975]
-----------------------------------------------------------------------------------------------------------
Intercept                                   2.6198      0.046     56.781      0.000       2.529       2.710
treatment                                   1.0386      0.049     21.376      0.000       0.943       1.134
post                                        0.3703      0.079      4.673      0.000       0.215       0.526
treatment:post                             -0.2384      0.083     -2.888      0.004      -0.400      -0.077
flag_h3_fairness_concerns                  -0.0642      0.237     -0.271      0.786      -0.528       0.400
treatment:flag_h3_fairness_concerns        -1.1794      0.250     -4.722      0.000      -1.669      -0.690
post:flag_h3_fairness_concerns             -1.0259      0.556     -1.845      0.065      -2.116       0.064
treatment:post:flag_h3_fairness_concerns    1.2030      0.571      2.106      0.035       0.083       2.323
==============================================================================
Omnibus:                    14285.035   Durbin-Watson:                   1.966
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             2495.912
Skew:                          -0.590   Prob(JB):                         0.00
Kurtosis:                       1.846   Cond. No.                         114.
==============================================================================

Notes:
```

APPENDIX B

BEHAVIORAL FLAGS DICTIONARIES

status_quo_bias_keywords = ["bring back the old version", "miss the old version", "the old [interface/feature] was better", "hate the new update", "the old way was easier", "wish it never changed", "don't like change", "still getting used to", "forced to use", "ruined by update", "why did they change", "go back to old", "the old look", "ruined it", "stop changing", "let us choose the old one", "previously was easier", "old was perfect", "I liked it better before", "new one is confusing", "do not want the redesign", "never should have changed", "I want the old", "feature removed", "unnecessary change"]

openness_to_change_keywords = [ "like the new look", "love this update", "happy with the new design", "the redesign is better", "finally updated", "getting better", "useful improvements", "looks much better now", "I'm glad they changed", "long overdue", "easier to use now", "prefer the new version", "works well after update", "the change was for the best", "needed fresh look"]

loss_aversion_keywords = ["too expensive", "too costly", "overpriced", "price increase", "price hike", "increased price", "raised price", "price went up", "prices went up", "cost more", "costs more", "paying more", "more expensive", "expensive now", "getting expensive", "price increase not worth", "not worth the price", "not worth the cost", "not worth paying", "canceling", "cancelling", "cancelled", "cancel", "unsubscribe", "unsubscribed", "switching", "switched", "leaving", "left the app", "leaving spotify", "will cancel", "going to cancel", "thinking of canceling", "considering canceling", "might cancel", "considering leaving", "may cancel", "about to cancel", "lost access", "lost feature", "lost features", "removed feature", "removed features", "no longer have", "no longer get", "took away", "took features", "downgrade", "reduced", "fewer features", "quality decreased", "service declining", "regret", "regrets", "wish i hadn't", "wish i didn't", "mistake", "wasted money", "waste of money", "ripoff", "rip off", "scam", "scamming", "disappointed", "disappointing", "dismayed", "frustrated", "frustrating", "angry at", "upset about", "upset with", "annoyed by", "annoying", "unfair", "unfairly", "unfairness", "unjust", "unjustly", "wrong decision", "wrong choice", "greed", "greedy", "money grab", "cash grab", "exploiting", "exploitation", "taking advantage", "take advantage of", "users get less", "less value", "worse deal", "bad deal", "poor decision", "terrible decision", "considering alternatives", "looking for alternatives", "switching services", "switching to", "switching platforms", "try apple music", "try youtube music", "try amazon music", "other services", "better alternatives", "cheaper alternative", "free service", "free app", "free option", "better deal elsewhere", "disgusted", "disgusting", "despise", "hate", "despise spotify", "never again", "no more", "done with", "fed up", "sick of"]

fairness_concerns_keywords = ["value for money", "not value", "poor value", "bad value", "little value", "overcharging", "overcharge", "charged too much", "charge too much", "pricing unfair", "unfair pricing", "unjustified price", "unjustified", "unreasonable price", "unreasonable pricing", "unreasonable increase", "excessive", "exorbitant", "absurd", "ridiculous", "laughable", "not reasonable", "without warning", "no warning", "no notice", "sudden increase", "surprised", "surprising increase", "unexpected price", "not informed", "weren't told", "didn't inform", "didn't tell", "poor communication", "no communication", "sneaky increase", "hidden increase", "surprise charge", "surprise fee", "more expensive than", "expensive compared to", "expensive vs", "vs apple music", "vs youtube music", "vs amazon music", "vs other services", "compared to", "competition", "competitors", "rival services", "other apps", "other platforms", "should charge less", "should be cheaper", "should cost", "others charge less", "competitors cheaper", "cheaper elsewhere", "best price", "too high compared", "should include", "should come with", "should provide", "why no", "why didn't", "expected", "expected to", "used to get", "used to have", "previously had", "we paid", "we deserve", "deserve better", "users deserve", "injustice", "unjust", "unjustly", "not right", "not fair to", "not fair that", "shouldn't charge", "shouldn't have increased", "shouldn't raise", "shouldn't raise prices", "not okay", "not acceptable", "not justified", "wrong to charge", "selective increase", "only us", "only america", "unfair to us users", "unfair to americans", "why only", "why us", "unfair to certain", "discrimination", "discriminatory", "targeting", "not worth it", "doesn't justify", "doesn't justify the price", "doesn't warrant", "doesn't deserve", "not enough features", "same features", "no new features", "no added value", "no improvements", "nothing new", "nothing different", "same as before", "what am i paying for", "why am i paying more"]

positive_acceptance_keywords = ["understand the increase", "understand the price increase", "understandable", "fair price", "fair pricing", "worth the price", "worth the cost", "worth the money", "value for money", "good value", "acceptable", "reasonable", "reasonable price", "reasonable increase", "justified", "justified price", "good service", "good product", "great service", "love spotify", "happy to pay", "willing to pay", "okay with the increase", "okay with increase", "understand why", "makes sense", "makes sense to", "quality increase", "improved service", "quality worth paying for", "worth the money"]