# Estimating Position and Social Influence Effects in Online Search

Ata Jameei Osgouei [1]
Fairfield University

Andrew T. Ching [2]
Johns Hopkins University

Brian Ratchford [3]
The University of Texas at Dallas

Shervin Shahrokhi Tehrani [4]
The University of Texas at Dallas

December 8, 2024

[1]Dolan School of Business, Fairfield University, 1073 North Benson Road, Fairfield, CT 06824. Phone: (323) 768-4821. E-mail: ajameeiosgouei@fairfield.edu.

[2]Carey Business School, Johns Hopkins University, 100 International Drive, Baltimore, MD 21202. Phone: (410) 234-4504. E-mail: andrew.ching@jhu.edu.

[3]Naveen Jindal School of Management, The University of Texas at Dallas, 800 West Campbell Rd. SM32 Richardson, TX 75080. Phone: (972) 883-5975. E-mail: Brian.Ratchford@utdallas.edu.

[4]Naveen Jindal School of Management, The University of Texas at Dallas, 800 West Campbell Rd. SM32 Richardson, TX 75080. Phone: (972) 883-5921. E-mail: Shervin.Tehrani@utdallas.edu.

**Abstract**

This paper contributes to the literature on the structural modeling of online consumer search by incorporating both product position and social influence effects (e.g., product popularity). Empirically disentangling these two effects is challenging because product position and popularity are usually highly correlated in field data. To overcome this challenge, we make use of a publicly available dataset from a field experiment conducted on an online music platform. The experimental design motivates us to introduce a two-step estimation procedure to address our identification challenge. In addition, we introduce a new empirical structural search model that allows consumers to choose multiple items in each search session and adaptively learn about their match with the platform. Our counterfactual experiments show that (i) by revealing product popularity and sorting items accordingly, platforms can help consumers select products more efficiently, resulting in less search effort and more downloads; (ii) by sorting items randomly and disclosing their popularity, it leads to more consumer search activity and could result in higher advertising revenue.

Keywords: Consumer Search, Product Position, Social Influence, Information Disclosure, Music Platforms, Content-based Filtering

# 1 Introduction

Consumers often need to choose among alternatives with which they have little or no prior experience, e.g., books, songs, movies, TV shows, vacation destinations, hotels, etc. When making their decisions, consumers may rely on information supplied by others, e.g., sales ranking, product ratings, reviews, etc. This reliance on others' behaviors and opinions is known as social influence.

In addition to social influence information, a product's position may also play an important role in consumer choice. For instance, consumers may tend to read from top to bottom, and hence they may search for products listed near the top first. Different combinations of social information disclosure and product positioning strategies could have profound implications on which products are more likely to be chosen by consumers. However, previous research on online search provides little guidance about how platforms should act. This is because the literature has considered only one of these two factors so far, not both.[1]

To address this major shortcoming in the literature, this paper estimates the effects of social influence and product position on consumer search behavior. More specifically, we study product popularity as one form of social influence. By examining how position and social influence impact search behavior, we aim to provide a more comprehensive understanding of the consumer decision-making process and offer insights that can help platforms optimize their strategies for presenting products and information. However, disentangling these two effects is challenging because product position and popularity are usually highly correlated in field data, as platforms often sort products according to their popularity.

To overcome this challenge, we utilize data from a field experimental study conducted by Salganik (2008). We argue that the experiments, together with our structural search model, allow us to separately identify the effects of product position and popularity on consumer search behavior.

In these experiments, participants can sample from a total of 48 songs and download as many as they want in a single search session. Participants are randomly assigned to one of two conditions: (i) the Independent group and (ii) the Social Influence group. In the Independent group, no popularity information (i.e., number of downloads by previous participants) is provided. In the Social Influence

[1]For instance, Amazon and Expedia arrange products according to consumer queries, encompassing factors like product ratings, popularity (best-selling), and additional attributes. Conversely, Spotify does not reveal a song's popularity, whereas SoundCloud shows the number of times a song has been listened to. The choice of disclosing social influence information significantly impacts consumer responses and platform revenue, depending on whether the revenue model is based on product sales (e.g., Expedia and 7digital), subscriptions (e.g., Spotify Premium), or advertising (e.g., Pandora). These variations highlight the importance of considering both positions and social influence in platform strategy. See Figures A.12, A.13, A.14, A.15, and A.16 in Web Appendix F.

group, popularity information is disclosed to each participant. Under column display format, songs are sorted randomly in the Independent group for each participant (see Figure A.8); however, in the Social Influence group, songs are sorted by their popularity (see Figure A.9).

These two conditions allow us to implement a two-step estimation procedure to separately identify the effects of position and popularity on consumer search behavior. In the first step, we use data from the Independent group to isolate the position effect on search behavior, because the Independent condition ensures the absence of social influence. In the second step, we fix the position effect identified from the Independent group and use data from the Social Influence group to identify the effect of product popularity on consumer search behavior.

In addition to identifying the effects of position and social influence on search behavior, our paper makes a contribution by introducing a new sequential search model. This model addresses limitations in existing empirical studies, which often rely on the model proposed by Weitzman (1979), where agents face an optimal stopping problem and can only select one item per search session. However, this model is too restrictive for many real-world applications, including our experimental setting, where participants can select multiple items in a single search session. In our experiments, participants are allowed to download as many songs as they wish during each session. To more accurately capture this behavior, we propose a new structural empirical sequential search model based on the "*Generalized Weitzman Model*" developed by Olszewski and Weber (2015a,b). This model accommodates multiple selections per session and better reflects the conditions of our experiments. We extend it further by allowing consumers to adaptively learn about their match with the platform. We estimate the structural parameters of this model using simulated maximum likelihood (SML) methods, enabling us to quantify the effects of both position and social influence in our search model.[2]

Our third contribution is to incorporate a machine-learning method known as *content-based filtering* into our search model to capture participants' download decisions in a more contextually relevant manner. Traditional search models rely on unobserved random match values from a one-dimensional fixed distribution to capture consumers' conversion (i.e., downloads in our case). However, this approach does not account for the rich and unstructured content attributes of songs' content which

[2]Although some previous research in search behavior has quantified the effect of product position in the e-commerce setting (Chen and Yao 2017, Ursu 2018), the effect of social influence has not been studied in this literature. On the other hand, the social influence literature has mainly investigated the impact of others' choices on an individual's decision-making while ignoring the consumer search context (e.g., Cai et al. 2009, Ching 2010, Tucker and Zhang 2011, Chen et al. 2011, Zhang and Liu 2012). As far as we know, our study is the first to incorporate both position and social influence effects in a structural search model.

is revealed after sampling an item, nor does it accommodate the significant heterogeneity in user preferences.

In our study, participants initially observe limited information about songs, such as titles and band names (pre-search). Subsequently, they gain access to additional multi-dimensional attributes of the songs after sampling and listening to them (post-search), including genre, tempo, key, tune, and instruments used. Our content-based filtering approach uses multiple layers of neural networks to capture the interactions between songs' multi-dimensional attributes, popularity, and participants' characteristics. This advanced machine-learning method allows us to group participants with similar tastes and forecast their download probabilities for all songs, including those not sampled in the experiment. By capturing these interactions between songs' multi-dimensional attributes and participants' heterogeneous preferences, we can more accurately predict download probabilities, which is crucial for conducting reliable counterfactual simulations.

Our fourth contribution is to use counterfactual experiments to examine the impact of three methods of disclosing social influence information on a platform under two different display formats (matrix vs. column): (i) random assortment without popularity information, (ii) random assortment with popularity information, and (iii) assortment based on popularity with popularity information. Our results suggest that platforms primarily generating revenue through sales should favor a column format that sorts songs by displayed popularity, as this format may be able to increase consumer purchases. However, platforms relying on advertising revenue may prefer the random assortment with popularity information displayed in the column format. This is because it would generate higher consumer engagement, as measured by the number of searches captured by their sampling behavior, leading consumers to stay longer when visiting a platform. Consequently, the platform gains additional opportunities to display advertisements.

The rest of the paper is organized as follows. Section 2 reviews related work on consumer search and observational learning. Section 3 describes the data used for analysis and provides reduced-form evidence. In Section 4, we introduce our empirical sequential search model, which relies on the generalized Weitzman model. Section 5 discusses the estimation approaches and identification. Section 6 discusses the estimation results on the empirical data. Section 7 investigates the effect of different counterfactual designs on participants' search behavior and provides managerial insights. Finally, Section 8 concludes the paper.

## 2 Literature Review

Our paper is related to the literature on estimating product position effects in online platforms. Various methodological approaches have been employed to isolate the position effect, including latent instrumental variables (Rutz et al. 2012), simultaneous equations (Ghose et al. 2014), regression discontinuity (Narayanan and Kalyanam 2015), control function (De los Santos and Koulayev 2017), and randomized ranking (Ursu 2018). However, none of these studies take into account the effect of social influence on the search process. Social influence could be a significant factor affecting search behaviors and product rankings on online platforms. In this paper, we address this gap by examining both position effects and social influence in the context of online search.

Our paper is also related to the literature on social influence. Banerjee (1992) and Bikhchandani et al. (1992) show that herding can occur due to observational learning. Anderson and Holt (1997) and Çelen and Kariv (2004) use laboratory experiments to identify the time at which subjects join a herd. Using a field experiment at a restaurant in Beijing, Cai et al. (2009) show that displaying a small set of popular dishes in addition to providing the standard menu, which has about 60 dishes, the sales of this set of popular dishes generates substantially more sales compared with the control setting where customers are only shown the standard menu. In another field experiment, Tucker and Zhang (2011) find that displaying click count information attracts more visitors for narrow-appeal vendors compared with broad-appeal vendors. Zhang (2010) investigates observational learning and the adverse effects of herding in the U.S. kidney market. Ching (2010) develops and estimates a structural learning model where patients pool their experiences together to update their prior beliefs about the quality of generic drugs. Zhang and Liu (2012) find evidence of rational herding among lenders on Prosper.com. Using data from an online music platform to estimate a structural learning model, Newberry (2016) finds that observational learning benefits high-quality music providers, consumers, and the platform. Krumme et al. (2012) show that observational learning leads to greater unpredictability and inequality in cultural markets. Carrera and Villas-Boas (2023) finds that disclosing information on other consumers' purchase behavior increases generic purchase shares significantly. However, unlike our study, these works did not incorporate the role of social influence in consumer search behavior.

Our paper also contributes to the literature on structural search models based on Weitzman (1979). Early examples are Ratchford (1982), who provides an economic framework for studying consumer search and choice behavior, and Moorthy et al. (1997) who study the effect of consumers'

prior beliefs on their search strategies. More recently, a number of empirical studies have also applied the framework of Weitzman (1979) to various search contexts. Prominent examples in this stream are Kim et al. (2010), who apply the model to click-stream data; Honka and Chintagunta (2017), who extend the model to study the purchase decisions of consumers; and Kim et al. (2017), who use the probit model of sequential search to implement the Weitzman (1979) model.[3]

Despite its merits and widespread application, Weitzman (1979) only applies to situations where consumers only choose one item in each search session, which renders it unsuitable for our data and some other real-world situations. Recently, Olszewski and Weber (2015a,b) extended the Weitzman model by allowing consumers to choose multiple products in a search session. We extend their framework to an estimable structural model to capture participants' behavior in our empirical context. Since this model does not have a closed-form solution, we estimate it using a logit-smoothed AR simulator (Accept-Reject simulator), which has been used in the estimation of related search models (e.g., Honka 2014, Honka and Chintagunta 2017, Ursu 2018).[4]

As mentioned above, our research makes use of the data from Salganik (2008), who designed and ran a field experiment to study the social influence in entertainment markets. This experiment provides unique detailed data containing the individual-level sampling and downloading behavior dynamics under control/treatment conditions. However, Salganik et al. (2006), Salganik and Watts (2008), and Salganik and Watts (2009) only used the data to measure the impact of social influence and different assortments of products on the average market shares of songs. They use the Gini Index to illustrate that displaying products' popularity and sorting them based on it leads to greater inequality of success. Using the data from Salganik et al. (2006), Hendricks et al. (2012) model the effect of observational learning on the probability of herding for high and low-quality songs and use the aggregated songs' download shares to test their model predictions in the long run. However, they do not account for how search behavior influences song selection, thereby missing the interplay between product position and social influence on participants' search behavior.

From a managerial perspective, platforms must decide whether and how to disclose social information. Research has focused on ranking and recommendation systems, showing that utility-based

[3]Some recent works have also modeled consumer search in an equilibrium framework. For instance, Moraga-González et al. (2017) presents a theoretical model to characterize how consumers' heterogeneous search costs affect firms' pricing policies in an equilibrium.

[4]The alternative approach to estimate these types of models is GHK-estimator (named after Geweke (1989), Hajivassiliou and McFadden (1998), Keane (1994)). Note that we follow a similar approach as in Choi and Mela (2019) and Chung (2019) allow for a stochastic element that enters search costs. Therefore, we use the logit-smoothed AR simulator for convenience. However, we explain in Web Appendix G, that using an alternative modeling approach, the GHK-estimator can be applied in estimating a search model as in Jiang et al. (2021).

rankings, such as those examined by Ghose et al. (2014), De los Santos and Koulayev (2017), and Chen and Yao (2017), outperform traditional methods in revenue and engagement. Additionally, Ursu (2018) and Greminger (2022) provide insights into how utility-based rankings and position effects influence consumer behavior. Our study extends this line of research by taking social influence into account and conducting counterfactual analyses to assess the impact of different product popularity disclosure strategies by measuring their effect on consumer sampling, downloading, and search efficiency.

# 3 Data Description

Salganik (2008) created a website called MusicLab, where participants can sample and download any of the 48 songs available there. Our analysis makes use of three experiments conducted there. These three experiments have around 10,000 participants. The data collected from MusicLab is particularly well-suited for our research objectives. It tracks participants' sampling and downloading decisions at the individual level on a platform that mimics commercial platforms such as Spotify, iTunes, SoundCloud, Pandora, etc. (see figures in Web Appendix F).

## 3.1 Experimental Design

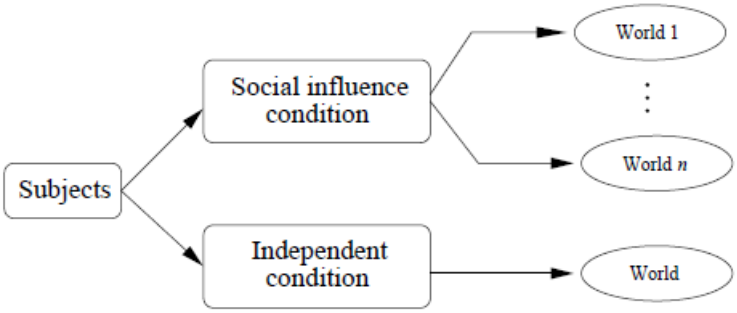
Three experiments were run at different times. Experiment 1 uses a matrix display format, where songs were arranged on the screen in a $16 \times 3$ matrix. Experiments 2 and 3 use a column display format, where songs were shown in a single column. In each experiment, participants arrived sequentially and were randomly assigned to one of the two groups, "Independent" or "Social Influence," as shown in Figure 1.[5] The Independent group participants only observed the names of the songs and the corresponding bands. The Social Influence group differs from the Independent group by disclosing past participants' choices, i.e., the number of times each song has been downloaded, which we refer to as "Popularity." Here, it captures the idea of social influence by providing an informative signal about the desirability of songs by others.[6]

It is important to note that in Experiment 1 (matrix format), songs are *always* sorted randomly

[5] Figure 1 is extracted from Salganik (2008).

[6] Sun et al. (2019) investigates the normative effect of social influence on people's sampling behavior. For example, a person might listen to a well-known song because her friends also listen. In this scenario, social influence affects someone's behavior by conforming to the crowd's choice without getting new information about the chosen item. Since all 48 songs are unknown to our participants, social influence has an informative role in our case in contrast to the normative effect. In sum, we study social influence from the informative viewpoint in conformity theory. https://www.jove.com/science-education/11055/conformity

Figure 1: Experiment Design




in both the Independent and Social Influence groups for each participant (see figures A.6 and A.7). In Experiments 2 and 3 (column format), songs are also sorted randomly in the Independent group (see Figure A.8); however, in the Social Influence group, songs are sorted by their popularity (see Figure A.9).[7] The setup in Experiments 2 and 3 closely mirrors real-world scenarios where product position and popularity are often confounded, creating challenges for researchers to disentangle these effects when analyzing search behavior data.

One advantage of the experimental design is that it had multiple parallel worlds of the Social Influence group, where each world had its own initial condition and evolved independently over time. This created multiple realizations of the same process and alleviated the sampling error problem when we estimated our model. Experiments 1 and 2 each have one Independent world and eight Social Influence worlds. Experiment 3 has one Independent world and two Social Influence worlds.

### 3.2 Experimental Procedure

Upon joining the experiment, each participant needed to complete a survey, which provided us with their demographic information, music expertise, and internet connection type. After the survey, participants were randomly assigned to one of the above groups and were directed to the songs screen page where they could sample the songs. After sampling a song, the participants decided whether to download it or not. Then, they choose to either return to the main menu to continue their search or exit the experiment (see Web Appendix E). It is important to highlight that participants cannot download a song directly without sampling it first.

The 48 songs in the experiment were chosen from bands unknown to participants to minimize

[7] If songs have identical popularity (i.e., the same number of downloads), their positions are randomly determined among these similarly popular songs. For instance, see the second and third-ranked songs in Figure A.9.

any external influence. Furthermore, participants were removed if: (1) they had heard about the experiment from one of the bands or their friends had informed them about a specific song or a band, or (2) they had logged off and re-entered the experiment again. The exclusion ensures that the songs are unknown to the participants, that nobody participates more than once, and that the participants do not hear about the experimental design from other people. This allows us to address two common problems in the field data: (i) consumers' initial prior information is often influenced by word of mouth (WOM) or other external factors; (ii) consumers may try to conform to what their friends do. More details about the experimental design are provided in Web Appendix E.

### 3.3 Summary Statistics

Table 1 presents the descriptive statistics of the data used for our analysis. We first focus on the matrix display format, i.e., experiment 1 (the upper panel). In the Independent group, there are 773 participants; the average age is 22.05, and 63% percent of participants are male. Twenty-one percent of participants joined the experiment with a dial-up connection. Participants sampled 6.39 songs and downloaded 2.02 songs on average. In the Social Influence group, there are 2,960 participants, the average age is 22.46, 65% percent of participants are male, and 23% use a dial-up connection. On average, participants sampled 6.79 songs and downloaded 2.23 songs.

We now turn to the column display format, i.e., experiments 2 and 3 (the lower panel). In the Independent group, there are 1,803 participants, the average age is 26.1, 46% of participants are male, and 16% use a dial-up connection. On average, participants sampled 9.2 songs and downloaded 2.15 songs. In the Social Influence group, there are 4,004 participants, the average age is 22.3, 35% of participants are male, and 24% use a dial-up connection. On average, participants sampled 7.1 songs and downloaded 2.5 songs.[8]

### 3.4 Preliminary Evidence

On average, participants sampled and downloaded 12.5% and 4% of songs, respectively. This indicates that participants face a significant search cost. In addition, Figure 2 displays the distribution of the maximum number of songs sampled by participants in each condition. Notably, there is significant heterogeneity in participants' search behavior. More specifically, 48.67% sampled less than four songs, and 1.59% sampled all 48 songs. This motivates us to allow for heterogeneous search

[8]These three experiments were conducted at different times and involved participants from different populations. We account for these variations by modeling individual behaviors and considering participants' characteristics.

Table 1: Summary Statistics of the Data Used

**Matrix Design**

Independent group (N=773)

| | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|
| songs listened | 6.39 | 8.74 | 1 | 48 |
| songs downloaded | 2.02 | 4.62 | 0 | 47 |
| participant age | 22.05 | 9.35 | 9 | 77 |
| participant male | 0.63 | 0.48 | 0 | 1 |
| internet connection (dial-up) | 0.21 | 0.41 | 0 | 1 |

Social Influence group (N=2960)

| | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|
| songs listened | 6.79 | 9.20 | 1 | 48 |
| songs downloaded | 2.23 | 4.51 | 0 | 48 |
| participant age | 22.46 | 9.84 | 9 | 80 |
| participant male | 0.65 | 0.48 | 0 | 1 |
| internet connection (dial-up) | 0.23 | 0.42 | 0 | 1 |

**Column Design**

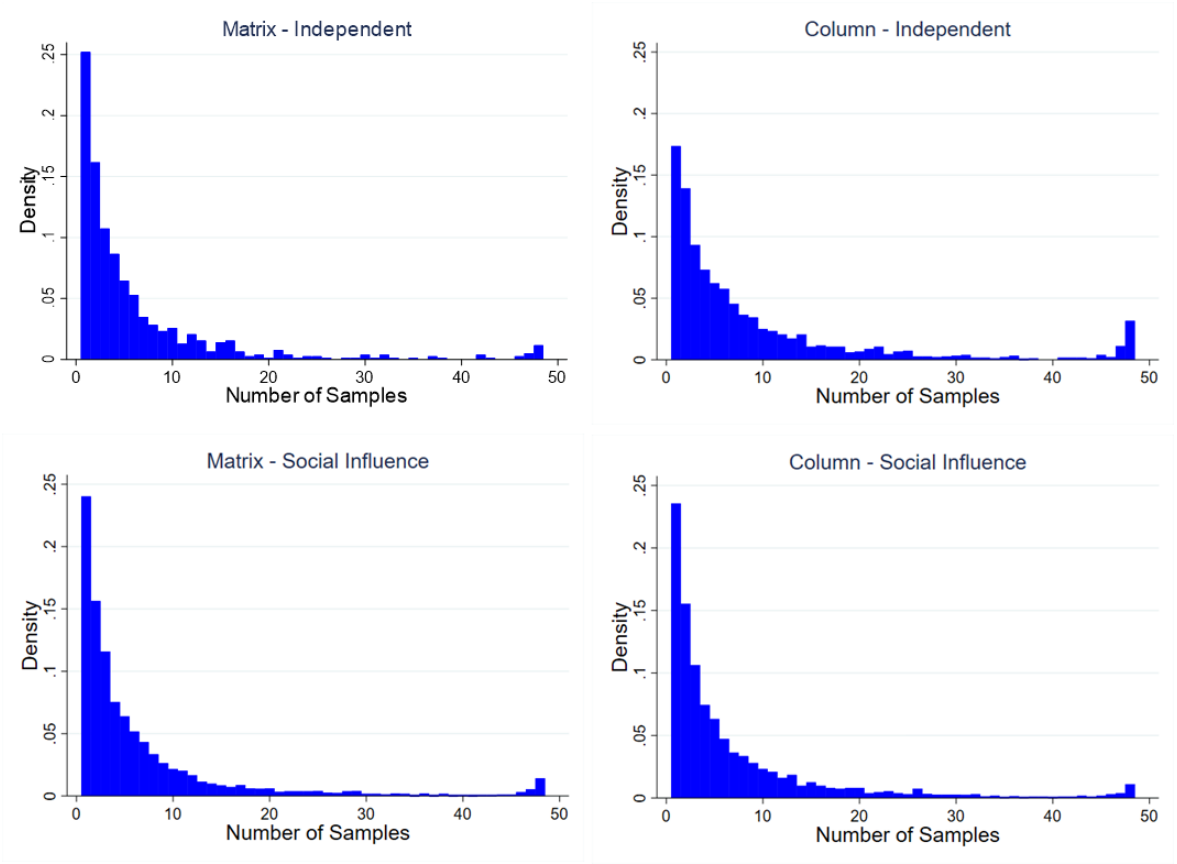
Independent group (N=1803)

| | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|
| songs listened | 9.18 | 11.44 | 1 | 48 |
| songs downloaded | 2.15 | 5.37 | 0 | 48 |
| participant age | 26.10 | 11.35 | 10 | 82 |
| participant male | 0.47 | 0.50 | 0 | 1 |
| internet connection (dial-up) | 0.16 | 0.36 | 0 | 1 |

Social Influence group (N=4004)

| | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|
| songs listened | 7.10 | 9.20 | 1 | 48 |
| songs downloaded | 2.53 | 5.54 | 0 | 48 |
| participant age | 22.28 | 10.07 | 11 | 81 |
| participant male | 0.35 | 0.48 | 0 | 1 |
| internet connection (dial-up) | 0.24 | 0.43 | 0 | 1 |

Figure 2: Distribution for the Maximum Number of Songs Sampled by Participants



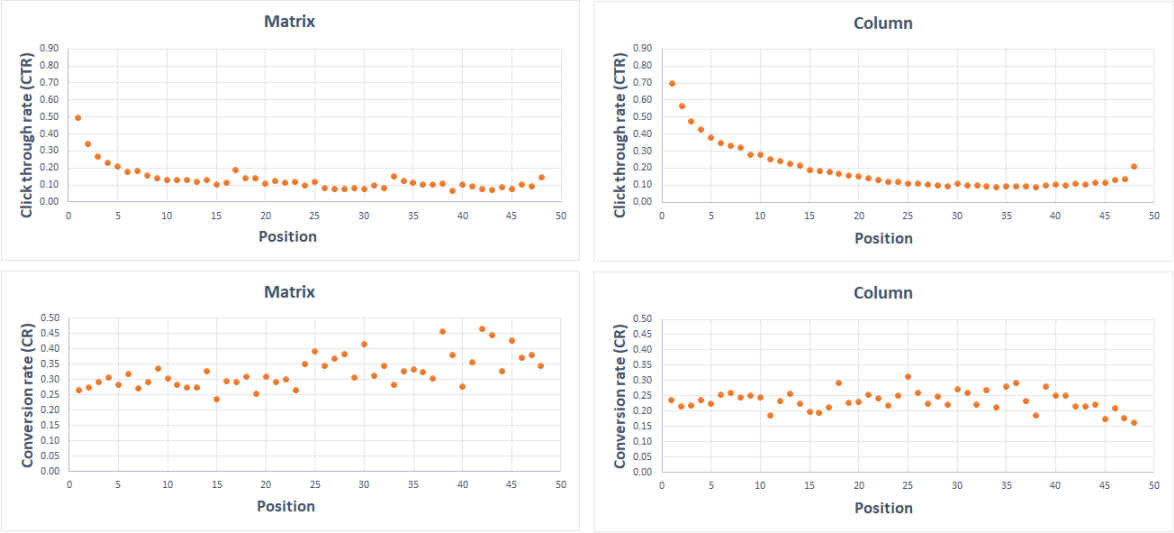

costs in our model.

By examining participants sampling and downloading decisions in the Independent groups where no social influence is present, Figure 3 provides reduced form evidence for the position effect. The top panel of Figure 3 illustrates the click-through rate (CTR) for different positions, representing the proportion of participants who sampled a song at each specific position. In the matrix format, it is assumed that participants read from the first column on the left, top to bottom, before moving to the next column.[9] It is clear from the top panel that songs in higher positions are sampled more frequently. The decreasing trend in CTR with increasing song positions indicates that a song's position significantly influences participants' sampling behavior. This pattern suggests a strong positional effect, where songs placed higher in the list receive more attention and sampling from participants.

In the bottom panel of Figure 3, we display the conversion rate (i.e., downloading rate) of a

[9]We also tested an alternative reading pattern where participants read each row from left to right before moving to the next row, yielding qualitatively similar results.

Figure 3: Click-Through Rate and Conversion Rate versus Product Position: Independent Group



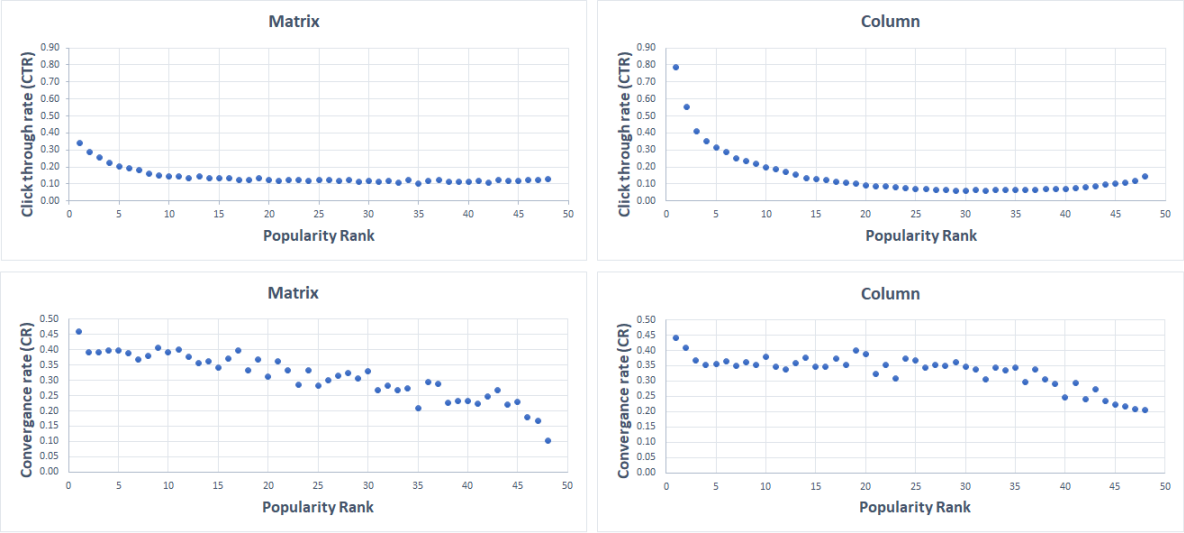

position, which is the download ratio of the sampled songs by participants in that position. Recall that the experimental design requires participants to sample a song before downloading it. The conversion/download rate of the sampled songs is quite stable across positions. This suggests that a song's position has very little impact on download decisions conditional on sampling, and it only affects the participants' search cost (Ursu 2018).[10]

Figure 4 shows the relationship between click-through/sampling rate and conversion/download rate relative to the popularity ranking within the Social Influence group. This figure offers two insights. Firstly, the top panel reveals that popular songs tend to be sampled more frequently, consistent with the idea that a more popular product is usually perceived to have higher quality (Cai et al. 2009, Ching 2010, Zhang 2010, Hendricks et al. 2012, Newberry 2016). This is the case even if we just focus on the matrix format, i.e., experiment 1, where songs are randomly sorted in the Social Influence group (see the left panel). Moreover, such a perception is consistent with the bottom panel of Figure 4, which shows that a song's conversion rate is positively correlated with its popularity; this indicates that participants' preferences are positively correlated with aggregate preferences.

Secondly, a comparison between the top panel of figures 3 and 4 suggests that individuals prefer sampling songs with higher positions and popularity. This is the case even if we just focus on the matrix display format (experiment 1), where songs are always randomly sorted even in the Social Influence groups, such that product position and popularity are independent. This implies that when

[10]Here, we follow the literature interpretation that the search cost is the clicking and scrolling costs for the list of items (Kim et al. 2010, Ursu 2018).

Figure 4: Click-Through Rate and Conversion Rate versus Popularity: Social Influence Group




the positions of songs are arranged based on popularity, as in column display format (experiments 2 and 3), disentangling the impacts of position and social influence becomes difficult. Thus, the confounding effects of position and social influence inherent in the column display format within Social Influence groups create an identification challenge.

Finally, we present some reduced-form evidence on how previous downloading outcomes may influence participants' continued sampling behavior on the MusicLab platform. Participants may adjust their sampling behavior based on their success rate in discovering high-quality songs. For instance, if a participant samples the first song and likes and downloads it as a high-quality song, she may become more optimistic about the remaining songs. Conversely, if the initial songs do not meet her expectations, she may become pessimistic about the quality of the remaining songs and decide to exit the platform. This behavior significantly impacts their search behavior, particularly their decision to continue or exit sampling.

Table 2 presents the results of a logit regression analysis on the likelihood of exiting the MusicLab platform. The key independent variable, $Dls_{ratio,t}$, represents the ratio of the number of downloads to the number of samples until the $t$-th search step at the individual level. This ratio normalizes downloads, as a higher number of samples naturally leads to more downloads. Additionally, the regression controls for participant characteristics such as age, gender, and internet connection type, and includes dummy variables for which experiment the participant was in.

This reduced-form evidence indicates that participants who experience lower download success (i.e., $Dls_{ratio,t} \downarrow$) in previous search steps are more likely to exit the platform. To capture this

behavior, we incorporate a feedback variable into our structural model, acknowledging the impact of past outcomes on future sampling decisions.

Table 2: Logit Regression Analysis of Exit Decision Based on Download Ratio

| Variable | Independent Group | Social Influence Group |
|---|---|---|
| $Dls_{ratio,t}$ | $-0.395^{***}$ | $-0.241^{***}$ |
| | (0.069) | (0.037) |
| $age_i$ | $-0.016^{***}$ | $-0.017^{***}$ |
| | (0.002) | (0.001) |
| $dial-up_i$ | $0.962^{***}$ | $0.985^{***}$ |
| | (0.062) | (0.033) |
| $male_i$ | $0.100^{**}$ | 0.033 |
| | (0.046) | (0.029) |
| $intercept$ | $-1.411^{***}$ | $-1.454^{***}$ |
| | (0.076) | (0.046) |
| Experiment Dummy | ✓ | ✓ |
| Number of Observations | 21488 | 48511 |

Note: Robust standard errors are shown in parentheses. * $p < .1$, ** $p < .05$, *** $p < .01$

# 4 Model

## 4.1 A Case Study of Generalized Weitzman Model

In the well-known sequential search model developed by Weitzman (1979), an agent called Pandora is presented with $n$ boxes. For $j = 1, ..., n$, box $j$ contains a prize $x_j$ that comes from an independent probability reward distribution $F_j$ known to Pandora. Pandora can pay a cost $c_j$ to open box $j$ and observe its prize, denoted as $x_j^o$.[11] Weitzman (1979) shows the optimal solution for the Pandora problem is to open the boxes in the descending order of their reservation values (selection rule) until the maximum realized prize found is greater than the reservation value of any unopened box (stopping rule). As noted in the literature review, this model has been commonly applied to estimate consumers' sequential search behavior in economics and marketing (e.g., Kim et al. 2010, 2017, Ursu 2018, Ursu et al. 2023).

Despite its merits, one main limitation of the classic Weitzman model is that it does not apply to situations where agents select multiple items in a search session. On average, participants sampled 7 songs and downloaded 2 songs in our context. To allow a decision maker to choose multiple items in

[11] The superscript "$o$" is provided as a mnemonic for "opened" or "observed."

a search session, we apply one specification of *Generalized Weitzman Model* developed by Olszewski and Weber (2015a,b). We first describe the general model below and then specify its elements to capture our MusicLab environment. Then we explain how to extend it to an estimable structural model in Section 4.2.

As in Olszewski and Weber (2015a,b), let $S$ be the set of opened boxes, $\vec{x}_S^o = (x_l^o; l \in S)$ be the vector of observed prize values from boxes in $S$, $u(\vec{x}_S^o)$ be the utility function which is increasing in $\vec{x}_S^o$. Similar to the Weitzman model, the agent's goal is to maximize the following expected net utility,

$$R(\vec{x}_S^o) = u(\vec{x}_S^o) - \sum_{j \in S} c_j. \tag{1}$$

The main difference is the agent gains utility from a combination of the observed prizes rather than solely the maximum of the observed prizes as in Weitzman (1979).[12] Olszewski and Weber (2015a,b) show sufficient conditions for $u(.)$ such that the following strategy leads to the optimal sequential search behavior.[13]

Consider $j \notin S$, where $S$ is the set of previously opened boxes. The "reservation value" for box $j$, $x_j^*$, is defined as the smallest prize such that Pandora is indifferent to opening box $j$ or stopping the search:

$$x_j^* = min\left\{ z : u(\vec{x}_S^o, z) \geq -c_j + \mathbb{E}\Big[u(\vec{x}_S^o, x_j^o, z)\Big], z \geq 0 \right\}, \tag{2}$$

where the expectation is taken over $x_j^o$ with the distribution $F_j$. Here, $z$ denotes any non-negative real number that can be added to the set of rewards in $S$. Unlike the reservation value in Weitzman (1979), $x_j^*$ is a function of not only $c_j$ and $F_j$, but also $\vec{x}_S^o$, i.e., the vector of all observed prizes before opening box $j$. Therefore, after opening a box, the reservation values of all the unopened boxes must be recomputed from Eq(2). This shows that the agent receives a combined utility of all opened boxes in the search process. Olszewski and Weber (2015a,b) show that the optimal strategy is to open the boxes in descending order of reservation values until there is no unopened box with reservation value greater than 0.

In what follows, we consider a specification of the above model outlined in Olszewski and Weber (2015b) and apply it to our MusicLab context. Let $F_{ij}$ be consumer $i$'s pre-search belief distribution about the chance of finding a high-quality song in box $j$. To simplify notations, we will drop the $i$

[12] In other words, Weitzman (1979) finds the selection and stopping rules under the assumption $u(\vec{x}_s^o) = \max(\vec{x}_s^o)$.

[13] Note that the utility function $u(.)$ maps vectors of any length to real values. We state these sufficient conditions of $u(.)$ in Web Appendix A.

subscript below. We assume that $F_j$ is a two-point distribution such that $x_j^o$ is either 0 or 1 with probabilities $q_j$ and $p_j = 1 - q_j$, respectively. In this context, $x_j^o = 1$ indicates that the participant finds a high-quality song after sampling it (i.e., she clicked and listened to it; she liked the song and downloaded it), while $x_j^o = 0$ indicates a low-quality song (i.e., she clicked and listened to it; she disliked the song and did not download it). We assume that if the participant discovers a high-quality song, she will download it. After the download decision, she may either continue searching on the main page or log off from the MusicLab platform.

We further assume that $u(.)$ is an increasing concave function of the total number of prizes found. Consider the objective function,

$$R(\vec{x}_S^o) = u\left(\sum_{j\in S} x_j^o\right) - \sum_{j\in S} c_j. \tag{3}$$

Eq(3) shows that the agent receives a utility of the sum of the observed prizes $\sum_{j\in S} x_j^o$, i.e., the combined utility from all downloaded items in a search session, minus the total search costs, $\sum_{j\in \vec{x}_S} c_j$. In contrast to the Weitzman model, the participant received a utility of a total number of high-quality songs in Eq(3). Here, the concavity of $u(.)$ captures the concept of diminishing marginal utility in downloading another song. In other words, if the participant has already downloaded several high-quality songs, they will gain less utility from finding one more. The assumption that payoffs of search outcomes are linearly additive and $u(\cdot)$ is concave, implies that rewards for songs are substitutes in the sense that they are exchangeable.

Olszewski and Weber (2015b) shows that in this particular specification of the model, the optimal solution takes the following form: for any binary reward distribution $F_j$ and any concave utility $u(.)$, the reservation value becomes $x_j^* = \max\left\{0, \frac{w_{k+1}-c_j/p_j}{w_{k+1}-w_{k+2}}\right\}$, where $w_i = u(i)-u(i-1)$ and $k$ denotes the total number of prizes have already been found.[14] Suppose that we have $n$ unopened boxes, and we sort these boxes in descending order according to their reservation values: $x_1^* \geq x_2^* \geq \cdots \geq x_{n-1}^* \geq x_n^*$. Without loss of generality, this ordering is mathematically equivalent to $\frac{p_1}{c_1} \geq \frac{p_2}{c_2} \geq \cdots \geq \frac{p_n}{c_n}$. For simplicity in notation, we use $x_j^* = \frac{p_j}{c_j}$ as the reservation value of box $j$.

The agent should open boxes in the order of $1, 2, \cdots, n$, and stop according to the following

[14] See Olszewski and Weber (2015b) for the proof. Note that this reservation value $x_j^* = \max\left\{0, \frac{w_{k+1}-c_j/p_j}{w_{k+1}-w_{k+2}}\right\}$ (or equivalently $p_j/c_j$) is only valid in the above setting. We discuss some possible extensions in Web Appendix G.

stopping rule,

$$u(k) \geq \mathbb{E}[u(\vec{x}_S^o \cup \{j\})] \iff u(k) \geq -c_j + p_j \cdot u(k+1) + (1-p_j) \cdot u(k), \tag{4}$$

or equivalently,

$$\frac{1}{x_j^*} = c_j/p_j \geq u(k+1) - u(k) = w_{k+1}. \tag{5}$$

Here, $\mathbb{E}[u(\vec{x}_S^o \cup \{j\})]$ denotes the expected search utility by adding box $j$ to $\vec{x}_S^o$ (which is the vector of the prizes that have been observed), and $u(k+1) - u(k)$ represents the expected marginal gain conditional on box $j$ having a reward (i.e., $x_j^o = 1$). Note that the stopping rule depends on the total number of high-quality songs that the agent has found so far.

The reservation value $x_j^* = p_j/c_j$ has an intuitive interpretation: a participant seeks to maximize her expected utility by opening these boxes and receiving rewards (i.e., sampling and downloading songs). She endures a cost $c_j$ for opening box $j$, where she will find a reward (i.e., $x_j^o = 1$) with probability $p_j$. These boxes differ in the probability of providing a reward to a participant (i.e., a high-quality song). Thus, boxes with higher $p_j/c_j$ will be more beneficial for the participant to search first. Therefore, the participant will start by opening the box $j$ with the greatest reservation value $x_j^* = p_j/c_j$ and continue searching until the expected benefit of opening the next box is lower than the cost of opening that box, i.e., $\frac{1}{x_j^*} < u(k+1) - u(k)$ if $p_j$ remains unchanged in the search process. In the next subsection, we will extend the above model to a structural empirical setting by specifying $p_j$ and $c_j$ to capture participants' search behavior on MusicLab.

## 4.2 Empirical Model Specification

We first specify the search costs. Figure 3 demonstrates that participants are more likely to sample songs at higher positions. Moreover, it suggests that, conditional on being sampled, the song's position does not affect its downloading probability. Thus, we model the agent's search cost as a function of the song's position. We follow Kim et al. (2010) and Ursu (2018) to specify the search cost of song $j$ for participant $i$ as,

$$c_{ij} = exp(\alpha_i + \beta \cdot Position_{ij}), \tag{6}$$

where the mean search cost component $\alpha_i$ captures the opportunity cost of time listening to any song, and $\beta$ captures the position effect of song $j$. Intuitively, songs positioned higher in the list

may benefit from increased visibility and accessibility.

Recall that Figure 2 provides evidence that there is significant heterogeneity in participants' search costs. This could be attributed to participants' individual characteristics, internet connections, or unobserved heterogeneity. Hence, we model $\alpha_i$ as,

$$\alpha_i = \alpha_{i0} + \alpha_1 \cdot age_i + \alpha_2 \cdot male_i + \alpha_3 \cdot dial{-}up_i, \quad \alpha_{i0} \sim N(\alpha_0, \sigma^2_{\alpha_0}), \tag{7}$$

where $\alpha_{i0}$ captures the unobserved heterogeneity in participants' search costs; $age_i$ is the participant's age; $male_i$ is a dummy variable which is one if the participant is male; and $dial{-}up_i$ is a dummy variable for dial-up connection. Because a dial-up connection has a lower speed, it can increase one's search cost.

A participant can sample a song to see whether she likes it or not. If participant $i$ likes song $j$, she will download it and receive the reward $x^o_{ij} = 1$ (i.e., it is high-quality). On the contrary, if she dislikes it, she will not download it and receive the reward $x^o_{ij} = 0$ (i.e., it is low-quality). Let $t$ index the sampling sequence. We set $u(.) = \sqrt{.}$ in the estimation.[15] It follows from Eq(3) and the realizations of $x^o_{ij}$ that $u\big(\sum_{j\in S} x^o_{ij}\big) = \sqrt{\sum_{j\in S} x^o_{ij}} = \sqrt{k}$, where $k$ denotes the number of liked songs discovered up to the current search step $t$, i.e., $k$-songs has been downloaded.

Before sampling song $j$, participant $i$ believes that she will like it with probability $p_{ijt}$, which we refer to as her *pre-search belief probability*. We model $p_{ijt}$ as a logit function of observable information about song $j$ before sampling, and her past experiences with previously sampled songs:

$$p_{ijt} = \frac{exp(\gamma_0 + \gamma_1.feedback_{it} + \gamma_2.appeal_j + \gamma_3.popularity_{ij})}{1 + exp(\gamma_0 + \gamma_1.feedback_{it} + \gamma_2.appeal_j + \gamma_3.popularity_{ij})}, \tag{8}$$

where $\gamma_0$ captures the participant's initial belief about the average likelihood of the platform's songs being high-quality; $feedback_{it}$ is the participant $i$'s experience from her previously sampled songs; $appeal_j$ is the average attractiveness of a song due to its name and band; and $popularity_{ij}$ is song $j$'s number of downloads by previous participants who arrived before participant $i$.

We should make two remarks here. First, we allow participant's preferences to be heterogeneous. This is why $p_{ijt}$ and $x^o_{ij}$ are participant $i$ specific. More specifically, a song can be liked by some participants but disliked by others. Second, $popularity_{ij}$ does not depend on $t$ in Eq(8) because it does not change during participant $i$'s search session; moreover, it only applies if participant $i$ is in

[15] We consider different functional forms of increasing and concave functions for $u(.)$ such as $\log(.)$, $\sqrt[3]{.}$, etc. All results are consistent and qualitatively the same.

a Social Influence group.

The parameter $\gamma_3$ measures the effect of $popularity_{ij}$. This is our main parameter of interest because it captures the social influence effect on search behavior. Intuitively, a song's popularity can be an "informative signal" for its quality. As shown in Figure 4, the higher the popularity, the more likely participants would sample it. This indicates that a participant's belief about her chance of liking the song may increase with its popularity.

A participant may change her pre-search beliefs about the remaining songs' according to the rewards discovered so far. As shown in Table 2, if the participant samples the first song and likes it, she may become more optimistic about the remaining songs in the list, and vice versa. Thus, we model $feeback_{it} = \frac{TP_{it}-TN_{it}}{TP_{it}+TN_{it}}$, where $TP_{it}$ (resp. $TN_{it}$) denote the total number of songs liked (resp. disliked) by participant $i$ up to the $t$-th step of the search. Our approach to modeling the updating of beliefs from past outcomes aligns closely with adaptive learning methodologies, as illustrated in Doraszelski et al. (2018) and Li and Ching (2024).[16]

We also control for the intrinsic appeal of a song to accurately assess the impact of other factors, such as position or social influence. We model *appeal* as follows:

$$appeal_j = l_j / \sum_k l_k, \tag{9}$$

where $l_j$ denotes the total number of times that song $j$ has been sampled in the Independent group. Intuitively, $appeal_j$ captures the average sampling rate of song $j$ when there is no social influence, and songs are randomly positioned. Thus, if song $j$ has a higher appeal compared with song $j'$ (i.e., $appeal_j > appeal_{j'}$), then on average, song $j$ is more attractive for all participants due to features other than its position. In this case, a song's appeal should only depend on its name and the band because these are the only observable characteristics before sampling songs in the Independent group. The parameter $\gamma_2$ captures the effect of the overall appeal of song $j$.[17]

[16] While we recognize the importance of allowing participants to update their pre-search beliefs based on outcomes, we restrict them from considering the exploration value. In other words, we assume that participants are myopic in a correlated learning problem as in Erdem (1998), Ching and Lim (2020). Nonetheless, we assume participants are forward-looking in determining the optimal search and stopping rules, aligning with previous sequential search literature. Note that, in the context of sequential search with correlated learning over boxes (as in this paper), fully rational exploration-exploitation tradeoff learning is NP-hard, meaning it is computationally infeasible for participants to perfectly balance exploration and exploitation (Chawla et al. 2020). Even the optimal decision cannot be approximated in polynomial time (Chawla et al. 2023). Therefore, people can exhibit bounded rationality when facing an NP-hard learning process (Tehrani and Ching 2024).

[17] Both appeal and popularity variables are standardized by subtracting them from their mean and then dividing them by their standard deviation. This puts them on the same scale, making comparisons and analyses more accurate. Importantly, the results remain robust and consistent even when using non-standardized data.

Note that one key distinction between our empirical structural model here and the Olszewski and Weber (2015a,b)'s theoretical model in Section 4.2 is that $p_{ijt}$ depends on the search step $t$ in our model, whereas it remains static in Olszewski and Weber (2015a,b). It is straightforward to show that the optimality remains valid using induction on $t$ based on $x^*_{ijt} = p_{ijt}/c_{ij}$ because $x^*_{ijt}$ is the corresponding reservation values of boxes at step $t$. In other words, the reservation values need to be recomputed after a search in our model, and the optimality goes through if this is done at each search step $t$.[18]

# 5 Estimation and Identification

## 5.1 Identification

Many platforms sort products based on their popularity, as in the column display format/Social Influence group of our experiment. If individuals prefer sampling from top positions and more popular products (see figures 3 and 4), these confounding effects pose a challenge to separately identify the position and social influence effects.

### 5.1.1 Two-Step Estimation for Column Display Experiments

Typically, experiments are designed so that one can identify the treatment effect in a straightforward manner using regression analysis or some other technique. Consistent with the goal of providing evidence to show that social influence can impact products' market shares, the experiments of Salganik et al. (2006) employed this standard approach. The Social Influence group is their treatment group, and the Independent group is their control group. However, they made no attempt to model and disentangle the product position and social influence effects in the consumer search process.[19]

Our key insight is that their experimental setup for column format provides a unique opportunity to address the above identification problem. Because songs in the Independent group are randomly sorted without displaying any popularity information, the data under this condition can be used to estimate the product position effect in the search model (via the search costs). The estimated position effect from the Independent group allows us to control its impact on search behavior under

[18] Alternatively, we can demonstrate the optimality using either the argument in Adam (2001) or by applying the Gittins index, as Olszewski and Weber (2015a,b) shows that the generalized Weitzman problem can be mapped to a multiple-armed bandit problem.

[19] More precisely, Salganik et al. (2006) show evidence that popularity information only matters in explaining the Gini coefficient of a song's success measured by the number of downloads, and they did not investigate the role of participants' search behavior and song's position and social influence on it.

Social Influence groups.[20] Hence, by holding the position effect and search cost parameters fixed when we estimate the model using only the social influence group data, we can identify the social influence effect from the search behavior not explained by the position effect.

This motivates us to use a two-step approach to estimate our structural search model. In step 1, we estimate the position effect ($\beta$) along with other search cost parameters ($\alpha$'s) and pre-search belief parameters ($\gamma$'s) by *only* using the Independent group data. This determines the search cost function on the platform, especially the position effect $\beta$. In step 2, we take the parameter estimates of the search costs obtained in step 1 as given and then re-estimate the pre-search belief parameters ($\gamma$'s), which include the social influence effect, (i.e., $\gamma_3$), by *only* using the Social Influence group data. This two-step estimation procedure allows us to disentangle the confounding effects of position and social influence.[21]

#### 5.1.2 One-step Estimation for Matrix Display Experiment

Another way to solve the identification problem is to randomly sort songs independent of their popularity in the Social Influence group. Salganik's Experiment 1 did just that. Although this experiment only considers the matrix display format, in this situation, we can estimate both product position and social influence effects simultaneously in one step. However, using a one-step estimation approach instead of the two-step method described above can lead to misleading estimates in the column format/Social Influence group (i.e., experiments 2 and 3) because it doesn't handle the confounding effects of position and popularity. We will revisit this point when we discuss the estimation results.

To check the robustness of our results, we compare the one-step and two-step estimation results under each display-social information disclosure condition. The comparison also provides evidence of how well the two-step method can address the aforementioned identification issue. Table 3 presents all estimation results under one-step and two-step approaches. As expected, the results for the matrix display design reveal that the effect of position and social influence in the one-step and two-step estimation results are similar because there is no confounding issue due to the random sorting

[20] Our identification argument is similar to Ursu (2018) who also uses the random assortment to identify the cost function and position effect. Furthermore, this paper takes into account the effect of social influence on the search process.

[21] There are several prominent examples of the two-step estimation procedures, e.g., Heckman selection models (Heckman 1979), the control-function approach for the robust estimation of nonlinear models with endogenous variables (Terza et al. 2008, Wooldridge 2014), dynamic stockpiling structural model (Hendel and Nevo 2006), etc. Our approach also takes advantage of the existence of the Independent vs. Social Influence groups to estimate the position and social influence in two distinctive stages to find unbiased causal effects.

of songs in both Independent and Social Influence groups.

However, we observe significant differences in the position and social influence effects between one-step and two-step estimation results in the column design. Specifically, the one-step approach yields a negative and insignificant effect of position, which is counter-intuitive and contradicts the reduced-form evidence indicating that participants sample top positions more frequently (see Figure 3), whereas the position effect is positive and significant in the two-step approach. Moreover, the social influence effect is significantly larger in the one-step estimation, with a value of 0.031 compared with 0.018 obtained in the two-step estimation.[22] We take this as evidence that the two-step approach is able to address the endogeneity issue.

Lastly, our two-step identification strategy relies on the availability of both Independent and Social Influence groups to separately estimate the effects of position and social influence. However, in practice, platforms may not always have access to an Independent group where social influence is not disclosed. We acknowledge this limitation in certain practical scenarios. However, if platforms can randomly sort items regardless of their popularity, our structural model can be estimated using a one-step approach without the need for an Independent group.

### 5.1.3 Search Costs and Pre-search Belief Intercepts

Besides the product position and social influence effects, the other parameters can be identified straightforwardly due to the presence of exogenous variation in the data. Note that $\gamma_0$ in pre-search belief has the same effect as $\alpha_0$ in search costs in terms of participants' search behavior. Intuitively, both lower base search cost (i.e., $\alpha_0 \downarrow$) and a higher base prior belief for a song being high-quality (i.e., $\gamma_0 \uparrow$) lead to more searches. Mathematically, $\alpha_0$ and $\gamma_0$ move the reservation values $x^*_{ijt}$ in the same direction through the numerator or denominator. Thus, we need to normalize one of them and estimate the other. Here, we fix $\gamma_0 = 0$ and estimate the rest of the parameters. Finally, it is worth noting that the identification of the appeal and popularity effects relies on the cross-sectional variation in the data. However, the feedback effect is identified from both cross-sectional and within-subject variation over time. Each participant has her own experience of sampling songs and whether

[22] Beyond comparing point estimates, we also calculate the marginal effects to address the non-linearity of the model. We consider a woman of average age with a dial-up connection from our data. In the matrix display, the marginal effects of the 1st position in the one-step versus two-step models are similar, with values of 0.00010 versus 0.00012, respectively. However, these marginal effects are significantly different in the column format, with values of -7.28E-06 versus 0.0002. Furthermore, in the matrix display, the marginal effects of social influence in the one-step versus two-step models are also similar, with values of 0.00234 versus 0.00235, respectively. In contrast, these marginal effects are significantly different in the column format, with values of 0.0076 versus 0.0046. These results are robust across different individual characteristics.

to download them or not.

### 5.2 Estimation

There are three sets of parameters to be estimated in the model: (i) participant-specific search cost parameters, denoted as $\vec{\alpha} = (\alpha_0, \sigma^2_{\alpha_0}, \alpha_1, \alpha_2, \alpha_3)$, (ii) the song-specific search cost with respect to the position effect $\beta$, and (iii) the participant's pre-search belief probability parameters $\vec{\gamma} = (\gamma_0, \gamma_1, \gamma_2, \gamma_3)$. Suppose participant $i$ sampled $s$ songs of the total $J = 48$ songs displayed on the platform.

Let $R_i = \{R_i(1), \cdots, R_i(s)\}$ denote the set of the sampled songs' positions and the order in which they were sampled. For example, if participant $i$ sampled three songs sequentially (i.e., $s = 3$) at positions 3rd, 20th and 8th, then $R_i(1) = 3$, $R_i(2) = 20$ and $R_i(3) = 8$, respectively. Note that we observe $R_i$ for each participant in our data. Given $F_{ijt} \sim Bern(p_{ijt})$ and utility function $u(k) = \sqrt{k}$ described in Section 4.2, participant $i$'s optimal sequential search strategy translates into the following three restrictions.

First, the selection rule requires that participant $i$ samples songs in the descending order of their reservation values $x^*_{ijt} = p_{ijt}/c_{ij}$. Therefore, if song $j$ is sampled at the $t$-th search step, its reservation value must be greater than the reservation values of the unsampled songs. Let $S'_{it}$ denote the set of songs not sampled by participant $i$ up to $t$-th search. Formally, we have

$$C1 := x^*_{ijt} \geq \max_{\eta \in S'_{it}} \{x^*_{i\eta t}\}, \ \ \forall t \in \{1, \cdots, s\}. \tag{10}$$

Second, if song $j$ is sampled at the $t$-th search, its expected benefit must be greater than the cost of sampling it; otherwise, the participant should have stopped searching. Therefore, we have

$$C2 := 1/x^*_{ijt} \leq w_{k+1}, \ \ \forall t \in \{1, \cdots, s\}, \tag{11}$$

where $w_{k+1} = u(k+1) - u(k)$, and $k$ is the number of downloaded songs until the $t$-th search step. Again, note that $w_{k+1}$ is the marginal gain conditional that song $j$ is a high-quality song to be downloaded by participant $i$, i.e., participant $i$ has $k+1$ high-quality songs in her library by adding song $j$.[23]

Finally, if participant $i$ has downloaded $K$ songs in total before stopping the search session (i.e.,

[23]Note that $w_{k+1} = u(k+1) - u(k)$ can be rewritten as $\mathbb{E}[u(\vec{x}^o_S \cup \{j\} \mid x_j = 1)] - u(k)$, which is the expected marginal utility of song $j$ being high-quality.

leaving the platform at the $s$-th search step when downloaded $K$ songs), the stopping rule will imply that the expected benefit of sampling those songs not searched by participant $i$ (marginal utility) must be lower than the search cost she would have had. This leads to the following condition:

$$C3 := 1/x^*_{ijt} > u(K+1) - u(K) = w_{K+1}, \ \forall j \notin R_i. \tag{12}$$

Let $P_{iR_i}$ denote the probability that participant $i$ follows the search sequence $R_i$ observed in the data, given the set of parameters. Then, we have

$$P_{iR_i} = Pr\Big(C1 \cap C2 \cap C3\Big) = \int I(cond)\phi(\alpha_{0i})d_{\alpha_{0i}}, \tag{13}$$

where $I(cond)$ is an indicator for whether conditions (10), (11), and (12) hold jointly. Thus, the log-likelihood over all participants can be written as:

$$LL = \sum_i \sum_{R_i} d_{iR_i} log(P_{iR_i}), \tag{14}$$

where $d_{iR_i} = 1$ if participant $i$ has the search sequence $R_i$; otherwise $d_{iR_i} = 0$. Since the integral in Eq(14) does not have a closed-form solution, we should use the simulated log-likelihood approach commonly used in the search literature by replacing $P_{iR_i}$ with the simulated choice probability $\hat{P}_{iR_id}$ with respect to the $d$-th draw from the unobserved heterogeneity distribution of base cost $N(\alpha_0, \sigma^2_{\alpha_0})$. This leads to the following expression:

$$SLL = \sum_i \sum_{R_i} \sum_d d_{iR_id} log(\hat{P}_{iR_id}) \tag{15}$$

Multiple approaches are available to determine the above probability $\hat{P}_{iR_id}$. We follow the empirical search models in economics and marketing (Honka 2014, Honka and Chintagunta 2017, Ursu 2018), and use the *logit-smoothed AR simulator*, in which the indicator function $I(cond)$ is replaced with a logit function that has first and second derivatives (McFadden 1989).[24] The logit-smooth AR simulator is implemented as follows:

1. Draw $d = \{1, ..., D\}$ samples of $\alpha_i^d$ for each participant to capture unobserved heterogeneity

[24]Using a direct approximation of $SLL$ by finding $P_{iR_id}$ and the step-wise indicator function leads to two well-known issues in the estimating parameters: (a) a low finite number of draws can result in $\hat{P}_{iR_id} = 0$; (b) the simulated probabilities may lack smoothness due to the non-twice differentiable nature of the indicator function. Consequently, employing the first and second derivatives to optimize the SLL will not be feasible.

2. Use $\alpha_i^d$ to form cost $c_{ij}^d$ with respect to $d$-th draw

3. Compute probability of rewards $p_{ijt}$ and the reservation values $x_{ijt}^{*d} = p_{ijt}/c_{ij}^d$

4. Define the following expressions for $d$-th draw:

   (a) $v_1^d = x_{ijt}^{*d} - \max_{\eta \in S'_{it}}\{x_{i\eta t}^{*d}\}$,

   (b) $v_2^d = w_{k+1} - 1/x_{ijt}^{*d}$, and

   (c) $v_3^d = 1/x_{ijt}^{*d} - w_{K+1}$.

5. Calculate $S^d$ w.r.t $d$-th draw using the logit expression:

$$S^d = \frac{1}{1 + \sum_{n=1}^{3} e^{-v_n^d/\lambda}}. \tag{16}$$

6. Compute the simulated probability $P_{iR_id}$ by taking the average of the $S^d$ over $D$ draws:

$$\hat{P}_{iR_id} = \frac{1}{D}\sum_{d} S^d. \tag{17}$$

Note that $v_i^d$ (for any $i = 1, 2, 3$) will be a larger positive value when the condition $Ci$ is satisfied much more strongly given parameters' values; otherwise, $v_i^d$ has a negative value. To put it differently, if participant $i$ follows $R_i$ sequence search as in Section 4, the $S^d$ and $SLL$ will approach 1 and 0, respectively. This effectively means that we are smoothing the log-likelihood function by using the logit-smooth AR simulator. Consequently, by maximizing the simulated log-likelihood function $SLL$, we are able to estimate the set of parameters such that the participants follow the generalized Weitzman sequential model proposed in Section 4.2. Here, $\lambda$ is called the smoothing scaling parameter. As $\lambda \to 0$, the simulator approaches the AR simulator and is thus unbiased. Hence, the researcher should use a small enough $\lambda$ that prevents the numerical problems of optimizing with a non-smooth function. In our case, we chose $\lambda = 1/50$, a value commonly utilized in the search literature, and that also works best in recovering the true parameters in the Monte Carlo simulation (see Web Appendix B for details).[25]

In Web Appendix B, we generate a simulated data set of 2,000 participants, each searching among 24 songs. For estimation, we use 100 draws to capture the heterogeneity in participants'

[25] Our results are consistent and qualitatively the same across other smoothing scaling parameters we tested, such as $\lambda = 1/5, 1/10, 1/25, 1/50, 1/100$. An alternative approach could involve using the GHK-estimator (Jiang et al. 2021), which does not rely on the choice of $\lambda$. Moraga-González et al. (2023) propose another way to estimate a structural search model by formulating the search problem as a discrete choice model.

search cost and use the inverse scaling factor $1/\lambda = 50$. Table A.1 summarizes the simulation results where the first column shows the true value of the parameters and the second column exhibits the estimated parameters. The results demonstrate that we can accurately recover the true values of the parameters.

# 6 Estimation Results

In this section, we provide the estimation results of the structural model, along with a discussion of the key findings. Recall that our identification argument leads to the conclusion that we should apply a two-step method to estimate the model from the column display format/Social Influence group (see Section 5.1.1). However, we can also apply a one-step method to estimate the model from the matrix display format/Social Influence group (see Section 5.1.2). For comparison purposes, we apply both one-step and two-step estimation methods to the social influence condition, regardless of the display format. Table 3 summarizes the estimation results.

The two-step estimation for the Social Influence groups deserves more explanation. As explained in Murphy and Topel (2002) and Wooldridge (2014), the standard errors for the parameters obtained in step 1 (i.e., search costs parameters here) are valid for two-step estimation procedures. However, when calculating the standard errors of the parameter estimates obtained in step 2, i.e., pre-search belief parameters in the Social Influence group, we need to take into account the standard error of the parameter estimates obtained in step 1.

Following Kasahara and Shimotsu (2008), we use the bootstrap method to address this problem. More precisely, we draw 100 bootstrap samples for the search cost parameters from the normal distributions with the mean set at the point estimate and the standard deviation set at the standard error, estimated in the Independent group. For each set of simulated search cost parameters from the bootstrap samples, we fix these parameters and then estimate the pre-search belief parameters using the Social Influence group data. The point estimate and standard errors of the parameter estimates in step 2 are the mean and standard deviation of the estimates obtained from the above bootstrap procedure.

We first discuss the two-step estimation results from the column display/Social Influence condition (the right panel in Table 3). The estimates of the position effect ($\beta$) and the popularity effect ($\gamma_3$) are 0.001 and 0.018, respectively, and they are both positive and significant. However, in the one-step estimation results, the position effect becomes negative at -3.4e-05 and it is insignificant;

Table 3: Estimation Results

| | Matrix | | | Column | | |
|---|---|---|---|---|---|---|
| | Independent | Social Influence | | Independent | Social Influence | |
| | | Two-Step | One-Step | | Two-Step | One-Step |
| Search Cost | | | | | | |
| $\alpha_0$ | $-1.55^{***}$ | $-1.55^{***}$ | $-1.53^{***}$ | $-1.62^{***}$ | $-1.62^{***}$ | $-1.49^{***}$ |
| | (0.02) | (0.02) | (0.01) | (0.01) | (0.01) | (0.01) |
| $\sigma_{\alpha_0}$ | $0.37^{***}$ | $0.37^{***}$ | $0.41^{***}$ | $0.41^{***}$ | $0.41^{***}$ | $0.43^{***}$ |
| | (0.02) | (0.02) | (0.01) | (0.01) | (0.01) | (0.01) |
| $\alpha_1(age)$ | $-0.0018^{*}$ | $-0.0018^{*}$ | $-0.0018^{***}$ | $-0.0033^{***}$ | $-0.0033^{***}$ | $-0.0019^{***}$ |
| | (0.0010) | (0.0010) | (0.0005) | (0.0005) | (0.0005) | (0.0004) |
| $\alpha_2(male)$ | $-0.064^{***}$ | $-0.064^{***}$ | $-0.035^{***}$ | 0.016 | 0.016 | $0.021^{***}$ |
| | (0.022) | (0.022) | (0.011) | (0.014) | (0.014) | (0.010) |
| $\alpha_3(dial-up)$ | $0.27^{***}$ | $0.27^{***}$ | $0.18^{***}$ | $0.26^{***}$ | $0.26^{***}$ | $0.22^{***}$ |
| | (0.03) | (0.03) | (0.02) | (0.02) | (0.02) | (0.01) |
| $\beta(position)$ | $0.0006^{***}$ | $0.0006^{***}$ | $0.0005^{***}$ | $0.0010^{***}$ | $0.0010^{***}$ | $-3.4E-05$ |
| | (9.2E-05) | (9.2E-05) | (5.2E-05) | (5.3E-05) | (5.3E-05) | (5.9E-05) |
| Pre-search Belief | | | | | | |
| $\gamma_0$ | 0 | 0 | 0 | 0 | 0 | 0 |
| $\gamma_1(feedback)$ | $1.25^{***}$ | $1.23^{***}$ | $1.16^{***}$ | $1.27^{***}$ | $1.25^{***}$ | $1.04^{***}$ |
| | (0.02) | (0.04) | (0.01) | (0.02) | (0.02) | (0.01) |
| $\gamma_2$ (appeal) | $0.008^{***}$ | $0.007^{***}$ | $0.005^{***}$ | $0.005^{***}$ | $0.005^{***}$ | $0.004^{***}$ |
| | (0.002) | (0.002) | (0.001) | (0.002) | (0.002) | (0.001) |
| $\gamma_3(popularity)$ | | $0.009^{***}$ | $0.009^{***}$ | | $0.018^{***}$ | $0.031^{***}$ |
| | | (0.0002) | (0.001) | | (0.0005) | (0.001) |
| Log-Likelihood | -5100 | -19700 | -19685 | -12636 | -25629 | -25285 |
| Observations | 37104 | 142080 | 142080 | 86544 | 192192 | 192192 |

Note: Robust standard errors are shown in parentheses. * $p < .1$, ** $p < .05$, *** $p < .01$

moreover, the popularity effect becomes much larger at 0.031. The counter-intuitive results from the one-step estimation of column display/Social Influence illustrate the identification problem that we discussed earlier in Section 5.1.1.

We now turn to discuss the estimation results from the matrix display/Social Influence condition (the left panel in Table 3). We start with the one-step estimation results. The estimates of the position effect ($\beta$) and the popularity effect ($\gamma_3$) are 0.0005 and 0.009, respectively; like the results from the column display format, they are both positive and significant. As a robustness check, we also apply the two-step approach to estimate our model using the data from this condition. In principle, because a song's position and popularity do not confound with each other here, we should still obtain similar results. In fact, this is the case; using the two-step approach, the estimates of the position effect ($\beta$) and the popularity effect ($\gamma_3$) are 0.0006 and 0.009, respectively, and they are very close to the estimates from the one-step approach. As discussed earlier (see footnote 22), to address the non-linearity of our model, we further compare the calculated marginal effects instead of point estimates. Our findings remain robust and consistent.

The positive estimate of the position parameter ($\beta$) indicates that the search cost is higher for songs at lower positions (i.e., located closer to the bottom of the list). Also, the higher value of $\beta$ in the column format vs. matrix format (i.e., 0.001 vs. 0.0006) suggests that the position effect is higher in the column format compared with the matrix format. In the column format, participants must scroll down to access lower-positioned songs, whereas in the matrix design, all songs are displayed simultaneously on the screen. Again, the results are robust and consistent if we consider the marginal effects rather than point estimates.

The positive estimate of $\gamma_3$ suggests that popular songs are more likely to be sampled by participants. Furthermore, the greater estimate of $\gamma_3$ in the column design shows that popularity has a much more pronounced effect on participants' pre-search beliefs under column display, compared with matrix display. This could be because, under the matrix format, all songs are randomly sorted; as a result, it would likely be harder for participants to figure out the ranking of songs by their popularity. For intuitive interpretation, we compare the percentage change in pre-search belief between column and matrix formats when popularity, measured by the number of downloads, changes by $x$ units. Assume a song with average appeal and popularity has been sampled by identical participants in both display formats (i.e., all cost covariates, positions, and feedback are identical at average levels). If the number of downloads increases by 1, 10, or 100 units, there is a 0.45%, 4.3%, and 20.12% higher chance, respectively, that the song will be considered high-quality in the

column format compared with the matrix format. This demonstrates the significant impact of social influence on perceived quality in search behavior.

We now turn to discuss other parameter estimates. All the search cost and pre-search belief parameter estimates are statistically significant and economically meaningful. The negative estimate of $\alpha_1$ indicates that older participants have slightly lower search costs and have continued to sample more songs than younger participants. The negative estimate of $\alpha_2$ in the matrix format indicates that male participants have lower search costs and have sampled more songs on average than their female counterparts; however, we should note that it is insignificant in the column format. The positive estimate of $\alpha_3$ shows that participants joining the experiment with a dial-up connection have greater search costs, and consequently, they also sampled fewer songs than others. Finally, $\sigma^2_{\alpha_0}$, which captures the participants' unobserved search cost heterogeneity, is significantly positive and large.

For the pre-search belief parameters, the positive and statistically significant estimate of feedback ($\gamma_1$) captures participants' adaptive learning behavior from sampling songs. This implies that when a participant samples a high-quality song, she will revise her pre-search belief that the remaining songs are more likely to be of high-quality and vice versa. Importantly, the effect of feedback is about 3.3% and 4.1% higher in Independent groups compared with Social Influence groups in matrix and column format, respectively.

Additionally, we observe a significant positive impact of appeal on pre-search belief across all estimation results. This indicates that songs with appealing names are more likely to be sampled during a search session. However, the appeal effect is also much smaller in the Social Influence groups compared with the Independent groups, regardless of the display format. The appeal has 37.5% and 40% less weight in the search behavior conditional on disclosing songs' popularity in matrix and column displays, respectively. Our result is consistent with the consumer behavior literature wherein decision-makers may re-weight the importance of heuristic clues in a decision process if there is a new information clue in the environment (Payne et al. 1993).

To illustrate the importance of social influence, we re-estimate the model from the Social Influence group by setting $\gamma_3 = 0$, and refer to it as the "No Popularity Model." Table 4 shows our Full Model performance compared with the No Popularity Model. Full Model yields substantial AIC enhancements of 117 and 1041 for the matrix and column formats, respectively. These results highlight the importance of taking into account the impact of popularity in explaining search behavior data.

We also consider a Random Search Model, where the parameters $\beta$, $\gamma_1$, $\gamma_2$, and $\gamma_3$, which capture the effects of position, feedback, songs' appeal, and popularity, are all set to zero. This model assumes that all songs have identical search costs and pre-search belief probabilities of a reward; this implies that participants randomly sample songs with equal probability regardless of their position, appeal, or popularity. Comparing the Full and Random Search models' estimation results in the Independent groups, the Full Model improves the AIC by 1358 and 2715 for matrix and column experiments, respectively. Furthermore, comparing the Full and Random Search models' estimation results in the Social Influence groups yields substantial AIC enhancements of 4029 and 7485 for the matrix and column formats, respectively. These results underscore the critical role of social influence and popularity in shaping search behavior and highlight the robustness and explanatory power of our model.

Table 4: Models Performance Comparison

| | Matrix | | Column | |
|---|---|---|---|---|
| | Independent | Social Influence | Independent | Social Influence |
| Log-Likelihood | | | | |
| Full Model | -5100 | -19700 | -12636 | -25629 |
| No Popularity Model | — | -19760 | — | -26150 |
| Random Search Model | -5782 | -21719 | -13997 | -29375 |
| AIC | | | | |
| Full Model | 10218 | 39421 | 25290 | 51278 |
| No Popularity Model | — | 39538 | — | 52319 |
| Random Search Model | 11576 | 43450 | 28005 | 58763 |

# 7 Counterfactual Analyses

Our structural estimation provides insights into how product popularity and position influence consumer search decisions. By running counterfactuals to evaluate different product assortments and disclosing social influence, platforms can strategically present products to enhance their click-through rate, conversion rate, and/or overall user experience.

## 7.1 Popularity Disclosure and Platforms' Policies

For each display format (column and matrix), we consider three counterfactual experiments to understand the impact of disclosing popularity information under different situations. In the counter-

factual experiment CF1, songs are sorted randomly, and participants do not observe their popularity. In the counterfactual experiment CF2, song popularity is revealed, but the sorting remains random. Finally, in the counterfactual experiment CF3, songs are sorted based on their displayed popularity, from the most popular to the least. In summary, we quantify the impact of all six scenarios on three measures of platform performance: (a) the average number of songs being sampled (click-through rate), (b) the average number of downloads (conversion rate), and (c) the search efficiency (user experience), the average of $\frac{\#\text{ downloaded songs}}{\#\text{ sampled songs}}$ for participants.

### 7.2 Advantages of Structural Counterfactual Analyses

We implement our counterfactuals based on the estimated structural model. It has the advantage of predicting outcomes of scenarios which Salganik (2008) did not consider in the experiments. Note that two cases were not included in the experimental design: (i) CF2 in the column display format and (ii) CF3 in the matrix display format. Without conducting new experiments, our structural approach allows us to evaluate all counterfactual scenarios, including these two. Additionally, we only observe the behavior of individual participants just under one of the above scenarios. In our counterfactual experiments, we will be able to simulate participants' behavior under scenarios to which they were not assigned.

Note further that the evolution of feedback and popularity are endogenous. When we change the way the platform sorts the songs or discloses popularity information, it will affect a participant's decision on which songs to sample and download. This will change the feedback and popularity variables and affect the next participant's search behavior. Our counterfactual experiments explicitly capture such endogenous dynamics.

### 7.3 Simulating Counterfactual Environments

We use all the participants and their actual demographic for the counterfactual simulations. Although each participant was only assigned to one condition, we simulate their behavior under all counterfactual scenarios. To account for the unobserved search cost for each participant, denoted by $\alpha_{i0}$, we make 100 draws of $\alpha_{i0}$ for each participant from the normal distribution with parameters estimated from the empirical results in Table 3. Then, we run simulations for each draw and report the average results.

To address possible new dynamics and market outcomes, we use a dynamically updated popularity measure for each participant based on the simulated download behavior of previous participants

in that world under the existing counterfactual design. We keep the order in which participants join a counterfactual experiment to be the same as what we observe in the data. Detailed explanations of the simulation procedures and calculations of the results are provided in Web Appendix D.

### 7.4 Enhancing Counterfactual Simulations Using Machine Learning

In our study, participants can sample any of the 48 songs in a counterfactual scenario. Therefore, we need an accurate prediction of the probability that a song will be downloaded if sampled in the counterfactual scenario. To precisely predict this likelihood, we must consider two critical factors: (a) participants' exposure to the songs' attributes, which are unstructured data representing the content of the songs are revealed after sampling, and (b) the significant heterogeneity in preferences among participants.

Given that participants, on average, sampled seven out of 48 songs, a standard random utility model (e.g., logit) may struggle to effectively capture the effect of songs' unstructured attributes and their interactions with heterogeneous consumer preferences observed in the data. To improve the accuracy of our predictions of participants' download behaviors, we employ a machine-learning method known as "content-based filtering." This well-known machine learning approach, offers several advantages over traditional models like logit. It can handle unstructured data more effectively by considering potential interactions through hidden layers of neural networks. Additionally, it adapts to individual user preferences without assuming explicit functional forms, allowing for greater flexibility in modeling user behavior. Moreover, content-based filtering can identify nuanced patterns in user interactions, leading to more accurate and personalized predictions.

To apply content-based filtering, we used the participant characteristics from survey data (Table A.3) alongside the songs' popularity and extracted unstructured features like genre, tempo, key, and used instruments (Table A.4). We extracted the features of these songs using the Librosa package, a widely recognized Python tool for analyzing and processing audio signals. By analyzing these data, we can group participants with similar tastes and forecast their download probabilities for all songs under different counterfactual scenarios. This method employs neural networks to create user and item vectors, essentially numerical representations of participants and songs based on their preferences and features. Then, these vectors are merged by using a sigmoid activation function, giving us a more accurate prediction of a participant's likelihood to download songs. When compared with a standard logit, content-based filtering proves superior in predicting download behaviors for all songs. We discuss the detailed implementation and comparison in Web Appendix C.

### 7.5 Counterfactual Results

We evaluate our counterfactual results in three dimensions: (1) $l$ denotes the average number of songs sampled, (2) $k$ denotes the average number of songs downloaded, and (3) $w$ denotes the search efficiency, which is the average ratio of the number of downloaded songs to the total number of sampled songs at the individual level. We interpret $w$ as the platform's efficiency in assisting consumers to find more high-quality songs in a search session.

Table 5: Counterfactual Results

| | **Matrix** | | | **Column** | | |
|---|---|---|---|---|---|---|
| | Samples ($l$) | Downloads ($k$) | Efficiency ($w$) | Samples ($l$) | Downloads ($k$) | Efficiency ($w$) |
| **CF1:** | 7.74 | 1.30 | 0.302 | 8.69 | 1.50 | 0.293 |
| **Randomly Sorted without Popularity** | (0.114) | (0.02) | (0.003) | (0.11) | (0.02) | (0.002) |
| **CF2:** | 8.22 | 1.38 | 0.333 | 9.18 | 1.63 | 0.333 |
| **Randomly Sorted with Popularity** | (0.11) | (0.02) | (0.004) | (0.09) | (0.02) | (0.004) |
| **CF3:** | 7.85 | 1.45 | 0.351 | 8.80 | 1.72 | 0.354 |
| **Sorted based on Popularity** | (0.11) | (0.02) | (0.006) | (0.11) | (0.03) | (0.006) |

Note: Please refer to Web Appendix D for an explanation of how $l$, $k$, $w$ are calculated. Robust standard errors are shown in parentheses. The t-test reveals significant differences between the relevant variables.

Our findings indicate that in the column format/CF1 — random assortment without popularity — consumers, on average, sampled 8.69 songs and downloaded 1.5 songs with a search efficiency of 0.293. Displaying the popularity of the songs in CF2 increased the average number of songs sampled and downloaded to 9.18 and 1.63, respectively, and raised the search efficiency to 0.333. Thus, disclosing social information increases the platform's search efficiency by 13.7%. In CF3, sorting songs based on popularity improves the impact of social influence on the search session. On average, participants sampled fewer songs (8.80) but downloaded more frequently (1.72), leading to the highest search efficiency of 0.354. Thus, sorting items based on their popularity further boosted the platform's search efficiency by nearly 6.3%.

Similarly, under the matrix format, participants sampled on average of 7.74, 8.22, and 7.85 songs in CF1, CF2, and CF3, respectively. The average number of songs downloaded by participants in CF1, CF2, and CF3 are 1.30, 1.38, and 1.45 songs, respectively, and the corresponding efficiencies are 0.302, 0.333, and 0.351, respectively. Thus, under the matrix format, only disclosing popularity information increases the platform's search efficiency by 10.3%, and sorting items based on their popularity further increases the search efficiency by 5.4%.

Interestingly, ranking songs based on their popularity leads to the following *asymmetric results* in both column and matrix display formats: (a) reducing the average number of songs sampled and (b) increasing the average number of songs downloaded. This leads to an increase in the search efficiency.

Also, conditional on disclosing the social information, all metrics will increase more significantly in the column format compared with the matrix format; in particular, for the case of sorting songs by their popularity, the number of songs sampled and the number of songs downloaded increases by 12.1% and 18.6%, respectively. This demonstrates that the column format not only enhances user sampling rate but also significantly drives higher conversion rates.

Interestingly, this is consistent with the concept of "Choice Overload Bias" in consumer behavior. Participants see more alternatives on the screen in the matrix display format compared with the column display format. Intuitively, having too many options can make a decision overwhelming due to many potential outcomes and risks involved with making a wrong choice, leading the decision maker to walk away (Shah and Wolford 2007).

### 7.6 Managerial Implications

Our counterfactuals have implications about how online platforms can increase their profits by implementing more suitable strategies to assort and disclose social influence depending on their revenue generation models. Note that our counterfactual analysis indicates that keeping everything else the same, the column format has a greater impact compared with the matrix format, as shown in Table 5. Thus, we will discuss the implications in terms of column format.

- **Product Sales Platforms:** Platforms like 7Digital, which derive revenue from product sales, would be better off adopting a column format where products are ranked according to popularity, as this approach is likely to result in a higher number of downloads and purchases, thereby enhancing platforms' revenue.

- **Subscription-Based Platforms:** Platforms like premium Spotify that charge a monthly subscription fee are also advised to use a column format where products are sorted based on their popularity. This design can improve utility and efficiency for consumers, discouraging them from switching to competing platforms. If the platform's goal is to maximize the search efficiency for consumers, our results suggest that, under the column format, it is preferable to sort products based on their popularity because higher search efficiency should lead to higher consumer satisfaction.

- **Freemium Platforms:** Platforms like the free-to-use version of Spotify or Pandora, which rely on advertising revenue, should consider a column design with random assortment while displaying popularity information. With CF2's 4.3% higher sample rate than CF3, consumers

are encouraged to explore more items and spend more time on the platform. This should generate increased ad display opportunities.

# 8 Conclusion

Leveraging a publicly available dataset from an online music platform field experiment, this paper contributes to the existing research on online consumer search by disentangling the impacts of product position and social influence on consumer search behavior. Our findings highlight the significant impact that social information can have on decision-making processes, reinforcing the need for careful consideration of these elements in digital marketplaces.

Furthermore, by employing a sequential search model that does not restrict to (i) one choice per search session and (ii) known reward distribution, we address two common limitations of the previous research based on the Weitzman Model. More specifically, our empirical search model captures consumers choosing multiple items per search session and adaptively learning about their matches with the platform by extending the theoretical model as in Olszewski and Weber (2015a,b).

In platforms with restricted pre-search information about products' attributes, such as music or short video platforms, users often make initial decisions based on limited cues like brief titles or the popularity level of items. However, after sampling—such as listening to a portion of a song—consumers encounter a richer set of features that can significantly influence their behavior, including decisions to download or engage further with the content. Previous research has abstracted away from these institutional details. To address this, we use content-based filtering to model the interactions between these detailed contextual attributes and user heterogeneous preferences in their conversion decisions. This approach allows us to conduct more reliable counterfactual analyses, offering valuable insights for platform managers. Specifically, platforms generating revenue through sales should use a column format sorting products by popularity to increase purchases, while those relying on advertising should opt for a random assortment with popularity info to increase consumer engagement and ad opportunities.

Our study significantly diverges from previous sequential search models, which are summarized in Ursu et al. (2023). These papers model the difference between pre-search and post-search utility using a random match value from a fixed one-dimensional distribution (see Web Appendix G). We highlight that when modeling platforms for music streaming (e.g., Spotify), short video streaming (TikTok), etc., users decide whether to sample items with limited initial information, and the multi-

dimensional unstructured attributes of these products are revealed only after sampling. This implies that pre-search utility and post-search utility may involve different sets of product attributes. This is why we have decided to deviate from the previous literature by modeling the sampling and downloading decisions separately.

We conclude by acknowledging some limitations and suggesting directions for future research. Our study assumes that participants use adaptive learning based on previous outcomes without accounting for the exploration value. We leave it for future research to investigate whether consumers are fully forward-looking when making exploration-exploitation trade-offs in search problems, especially when they have the option to choose multiple items.

Moreover, our findings are based on a controlled experiment, which may differ from real-world scenarios. The fact that participants were only exposed to one session and did not engage in monetary transactions for the downloaded music raises questions about the internal versus external validity of our results. Addressing these issues would require future studies to conduct field experiments on actual platforms to better capture new insights.

## Funding and Competing Interests

All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

## References


Adam K (2001) Learning while searching for the best alternative. *Journal of Economic Theory* 101(1):252–280.

Anderson LR, Holt CA (1997) Information cascades in the laboratory. *The American Economic Review* 847–862.

Banerjee AV (1992) A simple model of herd behavior. *The Quarterly Journal of Economics* 107(3):797–817.

Bikhchandani S, Hirshleifer D, Welch I (1992) A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy* 100(5):992–1026.

Cai H, Chen Y, Fang H (2009) Observational learning: Evidence from a randomized natural field experiment. *American Economic Review* 99(3):864–82.

Carrera M, Villas-Boas S (2023) Generic aversion and observational learning in the over-the-counter drug market. *American Economic Journal: Applied Economics* 15(3):380–410.

Çelen B, Kariv S (2004) Distinguishing informational cascades from herd behavior in the laboratory. *American Economic Review* 94(3):484–498.

Chawla S, Gergatsouli E, McMahan J, Tzamos C (2023) Leibniz international proceedings in informatics (lipics): Approximation, randomization, and combinatorial optimization. algorithms and techniques (approx/random 2023). *APPROX RANDOM* (Schloss Dagstuhl–Leibniz-Zentrum für Informatik).

Chawla S, Gergatsouli E, Teng Y, Tzamos C, Zhang R (2020) Pandora's box with correlations: Learning and approximation. *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*, 1214–1225 (IEEE).

Chen Y, Wang Q, Xie J (2011) Online social interactions: A natural experiment on word of mouth versus observational learning. *Journal of Marketing Research* 48(2):238–254.

Chen Y, Yao S (2017) Sequential search with refinement: Model and application with click-stream data. *Management Science* 63(12):4345–4365.

Ching AT (2010) Consumer learning and heterogeneity: Dynamics of demand for prescription drugs after patent expiration. *International Journal of Industrial Organization* 28(6):619–638.

Ching AT, Lim H (2020) A structural model of correlated learning and late-mover advantages: The case of statins. *Management Science* 66(3):1095–1123.

Choi H, Mela CF (2019) Monetizing online marketplaces. *Marketing Science* 38(6):948–972.

Chung JH (2019) *Estimation of sequential search models* (The University of Chicago).

De los Santos B, Koulayev S (2017) Optimizing click-through in online rankings with endogenous search refinement. *Marketing Science* 36(4):542–564.

Doraszelski U, Lewis G, Pakes A (2018) Just starting out: Learning and equilibrium in a new market. *American Economic Review* 108(3):565–615.

Erdem T (1998) An empirical analysis of umbrella branding. *Journal of Marketing Research* 35(3):339–351.

Geweke J (1989) Bayesian inference in econometric models using monte carlo integration. *Econometrica: Journal of the Econometric Society* 1317–1339.

Ghose A, Ipeirotis PG, Li B (2014) Examining the impact of ranking on consumer behavior and search engine revenue. *Management Science* 60(7):1632–1654.

Greminger RP (2022) Heterogeneous position effects and the power of rankings. *arXiv preprint arXiv:2210.16408* .

Hajivassiliou VA, McFadden DL (1998) The method of simulated scores for the estimation of ldv models. *Econometrica* 863–896.

Heckman JJ (1979) Sample selection bias as a specification error. *Econometrica: Journal of the econometric society* 153–161.

Hendel I, Nevo A (2006) Measuring the implications of sales and consumer inventory behavior. *Econometrica* 74(6):1637–1673.

Hendricks K, Sorensen A, Wiseman T (2012) Observational learning and demand for search goods. *American Economic Journal: Microeconomics* 4(1):1–31.

Honka E (2014) Quantifying search and switching costs in the us auto insurance industry. *The RAND Journal of Economics* 45(4):847–884.

Honka E, Chintagunta P (2017) Simultaneous or sequential? search strategies in the us auto insurance industry. *Marketing Science* 36(1):21–42.

Jiang Z, Chan T, Che H, Wang Y (2021) Consumer search and purchase: An empirical investigation of retargeting based on consumer online behaviors. *Marketing Science* 40(2):219–240.

Kasahara H, Shimotsu K (2008) Pseudo-likelihood estimation and bootstrap inference for structural discrete markov decision models. *Journal of Econometrics* 146(1):92–106.

Keane MP (1994) A computationally practical simulation estimator for panel data. *Econometrica: Journal of the Econometric Society* 95–116.

Kim JB, Albuquerque P, Bronnenberg BJ (2010) Online demand under limited consumer search. *Marketing Science* 29(6):1001–1023.

Kim JB, Albuquerque P, Bronnenberg BJ (2017) The probit choice model under sequential search with an application to online retailing. *Management Science* 63(11):3911–3929.

Krumme C, Cebrian M, Pickard G, Pentland S (2012) Quantifying social influence in an online cultural market. *PloS one* 7(5):e33785.

Li X, Ching AT (2024) How does a firm adapt in a changing world? the case of prosper marketplace. *Marketing Science* 43(3):673–693.

McFadden D (1989) A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica: Journal of the Econometric Society* 995–1026.

Moorthy S, Ratchford BT, Talukdar D (1997) Consumer information search revisited: Theory and empirical analysis. *Journal of Consumer Research* 23(4):263–277.

Moraga-González JL, Sándor Z, Wildenbeest MR (2017) Prices and heterogeneous search costs. *The RAND Journal of Economics* 48(1):125–146.

Moraga-González JL, Sándor Z, Wildenbeest MR (2023) Consumer search and prices in the automobile market. *The Review of Economic Studies* 90(3):1394–1440.

Murphy KM, Topel RH (2002) Estimation and inference in two-step econometric models. *Journal of Business & Economic Statistics* 20(1):88–97.

Narayanan S, Kalyanam K (2015) Position effects in search advertising and their moderators: A regression discontinuity approach. *Marketing Science* 34(3):388–407.

Newberry PW (2016) An empirical study of observational learning. *The RAND Journal of Economics* 47(2):394–432.

Olszewski W, Weber R (2015a) A more general pandora rule? *Journal of Economic Theory* 160:429–437.

Olszewski W, Weber R (2015b) A more general pandora rule? *Working Paper: arXiv:1509.09095 [math.OC]* .

Payne JW, Bettman JR, Johnson EJ (1993) *The adaptive decision maker* (Cambridge University Press).

Ratchford BT (1982) Cost-benefit models for explaining consumer choice and information seeking behavior. *Management Science* 28(2):197–212.

Rutz OJ, Bucklin RE, Sonnier GP (2012) A latent instrumental variables approach to modeling keyword conversion in paid search advertising. *Journal of Marketing Research* 49(3):306–319.

Salganik MJ (2008) Data from success and failure in cultural markets. *Supplement to doctoral dissertation. Department of Sociology, Columbia University. http://opr. princeton. edu/archive/cm/data_ release. pdf* .

Salganik MJ, Dodds PS, Watts DJ (2006) Experimental study of inequality and unpredictability in an artificial cultural market. *science* 311(5762):854–856.

Salganik MJ, Watts DJ (2008) Leading the herd astray: An experimental study of self-fulfilling prophecies in an artificial cultural market. *Social Psychology Quarterly* 71(4):338–355.

Salganik MJ, Watts DJ (2009) Web-based experiments for the study of collective social dynamics in cultural markets. *Topics in cognitive science* 1(3):439–468.

Shah AM, Wolford G (2007) Buying behavior as a function of parametric variation of number of choices. *PSYCHOLOGICAL SCIENCE-CAMBRIDGE-* 18(5):369.

Sun M, Zhang X, Zhu F (2019) U-shaped conformity in online social networks. *Marketing Science* 38(3):461–480.

Tehrani SS, Ching AT (2024) A heuristic approach to explore: The value of perfect information. *Management Science* 70(5):3200–3224.

Terza JV, Basu A, Rathouz PJ (2008) Two-stage residual inclusion estimation: addressing endogeneity in health econometric modeling. *Journal of Health Economics* 27(3):531–543.

Tucker C, Zhang J (2011) How does popularity information affect choices? a field experiment. *Management Science* 57(5):828–842.

Ursu R, Seiler S, Honka E (2023) The sequential search model: A framework for empirical research. *Working paper* .

Ursu RM (2018) The power of rankings: Quantifying the effect of rankings on online consumer search and purchase decisions. *Marketing Science* 37(4):530–552.

Weitzman ML (1979) Optimal search for the best alternative. *Econometrica: Journal of the Econometric Society* 641–654.

Wooldridge JM (2014) Quasi-maximum likelihood estimation and testing for nonlinear models with endogenous explanatory variables. *Journal of Econometrics* 182(1):226–234.

Zhang J (2010) The sound of silence: Observational learning in the us kidney market. *Marketing Science* 29(2):315–335.

Zhang J, Liu P (2012) Rational herding in microloan markets. *Management science* 58(5):892–912.

# WEB APPENDIX

# A The Sufficient Conditions in Olszewski and Weber (2015a,b)

Our model in Section 4 achieves the optimality of sequential search following Assumptions 1, 2, and 3 and using Theorem 3 in Olszewski and Weber (2015a,b). Consider the following three assumptions, which have been explained in more detail in Olszewski and Weber (2015a,b)

- **Assumption 1:** Let the utility function $u(.)$ has the following properties:
  - $u(\emptyset) = 0$ and $u(0, x_2, ..., x_k) = u(x_2, ..., x_k)$.
  - $u(.)$ is continuous, nonnegative, symmetric, nondecreasing, and submodular in its arguments.
- **Assumption 2:** The benefit of increasing a component of $x$ from 0 to some positive value $x_i^o$ is independent of the values of other components of $x$ which are greater than $x_i^o$. That is, for $x_S^o$ and any $x_i \leq \underline{x}_j < \overline{x}_j$, with $i, j \notin S$,

$$u(x_S^o, \overline{x}_j, x_i) - u(x_S^o, \overline{x}_j, 0) = u(x_S^o, \underline{x}_j, x_i) - u(x_S^o, \underline{x}_j, 0)$$

- **Assumption 3:** Olszewski and Weber (2015a,b) called this assumption the "History-Independence of the Ordering of Reservation Values." The ordering of reservation values $x_k^*$ of an unopened box is independent of both the number of boxes that have been already opened and their realizations. That is, for any $S$, $x_S^o$, and $k, j \notin S$,

$$x_k^*(x_S^o) \geq x_j^*(x_S^o) \Longleftrightarrow x_k^*(\emptyset) \geq x_j^*(\emptyset)$$

Now according to Olszewski and Weber (2015a,b), we have

**Theorem 1** *If Assumptions 1, 2, and 3 are satisfied, then the generalized Pandora's rule maximizes the expected utility.*

# B The Monte Carlo Simulation

We initially conducted simulation experiments to ensure that all parameters are identifiable. Here, we present the simulation results, demonstrating that the simulated maximum likelihood using the logit AR simulator effectively recovers the search cost and pre-search belief parameters. We generate a data set of 2000 participants, each searching among 24 songs. For estimation, we use 100 draws to account for the heterogeneity in participants' search cost and use the inverse scaling factor $1/\lambda = 50$. To overcome the simultaneous identification problem of $\alpha_0$ and $\gamma_0$, we set the $\gamma_0$ as its true value in the model. Table A.1 summarizes the simulation results where the first column shows the true values of the parameters and the second column shows the estimated parameters. Our results demonstrate that our model's estimations successfully recover the true values of all parameters. In Section 6, we present the estimated results for the field data.

Table A.1: Monte Carlo Simulation Results

| | True values | Estimated values |
|---|---|---|
| **Search Cost** | | |
| $\alpha$ | -1.5 | $-1.39^{***}$ |
| | | (.03) |
| $\sigma_\alpha$ | .5 | $.51^{***}$ |
| | | (.01) |
| $\alpha_1(age)$ | -.02 | $-.021^{***}$ |
| | | (.0006) |
| $\alpha_2(male)$ | .3 | $.311^{***}$ |
| | | (.02) |
| $\alpha_3(dial-up)$ | -0.3 | $-.299^{***}$ |
| | | (.02) |
| $\beta(position)$ | .001 | $.0012^{***}$ |
| | | (.0003) |
| Pre-search Belief (p) | | |
| $\gamma_0$ | 0 | 0 |
| $\gamma_1(feedback)$ | 1 | $0.88^{***}$ |
| | | (.04) |
| $\gamma_2(appeal)$ | .5 | $0.60^{***}$ |
| | | (.02) |

Note: Robust standard errors are shown in parentheses. * $p < .1$, ** $p < .05$, *** $p < .01$

## C Content-Based Filtering

There are two major types of recommender systems: collaborative filtering and content-based filtering methods. The collaborative filtering method utilizes past interactions between users and items to suggest new recommendations. The leading idea in the collaborative method is to use the interaction from all users and items to detect similar users and similar items. The method creates a vector for each user that captures their preferences for different attributes in the items and a corresponding vector for each item that incorporates these attributes. Collaborative filtering recommends items for you that you have not selected yet but have been liked by people similar to you. *Collaborative filtering doesn't use any information from users and items to create the user and item vectors; it merely relies on their interactions, such as ratings given to a product or consumers' purchase decisions.* This makes the collaborative method suffer from the *"cold start problem,"* making it challenging to handle new users or new items and inefficient for users or items with only a few interactions.

Unlike collaborative methods that merely rely on user-item interactions, content-based filtering incorporates additional information and features both from the users and items to make predictions. This additional information for users can be their demographic information such as age, gender, or any other personal information. For items such as songs, these features can be their singer, class type, etc. The content-based filtering model uses these user and item features to explain the observed user-item interactions. For example, the model will track which demographics like what types of songs, and for consumers with few interactions, it will suggest similar songs liked by users sharing similar demographic attributes.

The main challenge to implementing our counterfactual analyses is that we do not observe the download decisions of those songs that were not sampled by the participants in the actual experiment. Due to the small size of our data and the fact that many participants only sampled a few songs, using the simple logit/probit regression or even collaborative methods to predict the download decisions for the unsampled songs for these participants would not be precise and practical. We need to take into account several songs' unstructured content features and possible hidden layers to obtain a reliable prediction of the download probabilities for unsampled songs. We face a prediction task whose precision will lead to more accurate counterfactual outcomes.

We use content-based filtering on our data to predict the download probability for all the songs for each participant. Figure A.1 displays the model used for our predictions. The model consists of two neural networks to generate the user and item vectors whose dot product will be given to a

sigmoid activation node to predict the download probability. The model input data consists of the user and item content vectors. The user content vector contains the features of each participant. We used the survey's demographic information and engineered features from participants' ratings in the user content vector. The features used in the user content are participants' age, gender, expertise, country, internet connection (Dial-up or Broad-band), ability to use the internet, frequency of music sites visited, number of songs downloaded in the past 30 days, and the average rating given by the participant for each song genre. Table A.3 displays the variables used in the user content vector.

The item content vector contains the features of the songs. To generate the features for the songs, we use the Librosa[1] package in Phyton to extract features such as Mel-Frequency Cepstral Coefficients (MFCC)(20 in number), spectral centroid, zero crossing rate, Chroma frequencies, and spectral roll-off. Further, we use a neural network to identify the genre of the songs in our data set. We use the GTZAN genre collection data set to train this model. The GTZAN data set consists of 1000 songs with ten genres (Blues, Classical, Country, Disco, Hip-hop, Jazz, Metal, Pop, Reggae, and Rock), each genre having 100 songs. We applied the trained model to the songs in the experiment and used the softmax output of the model, which is a vector of size 10 predicting the probability of a song belonging to each genre. We use the songs' extracted spectrogram features, the softmax output of the genre classifier model, and the song's download rate, quality, and popularity in our item content vector. Table A.4 displays the features used in the item content vector.

Table A.2 displays the accuracy of our content-based filtering trained model along with a benchmark simple logit model in predicting the participants' download decisions for the sampled songs. In detail, we consider all variables in tables A.3 and A.4 as predictors in a linear additive form in the logit regression. To evaluate the predictive power, we train both models on a random 80% sample of the data (training data) and test the models on the remaining holdout sample. Now, Table A.2 provides evidence that our content-based filtering approach is superior in predicting download probability compared with the logit model.

As displayed, our trained model with F1-score=.71, precision=.64, and recall=.79 significantly outperforms the simple logit model with F1-score=.55, precision=.42, and recall=.79.[2]

[1] https://librosa.org/doc/latest/index.html

[2] Precision focuses on the accuracy of positive predictions (true downloads out of all predicted downloads); recall, also known as sensitivity, measures the true positive rate (proportion of correctly identified downloads out of all downloaded songs), and the F1-score provides a combined measure of precision and recall, offering a balanced assessment of a model's performance in binary classification tasks. A higher F1-score indicates a better performance and balance between precision and recall.

Figure A.1: Content-Based Filtering Neural Network

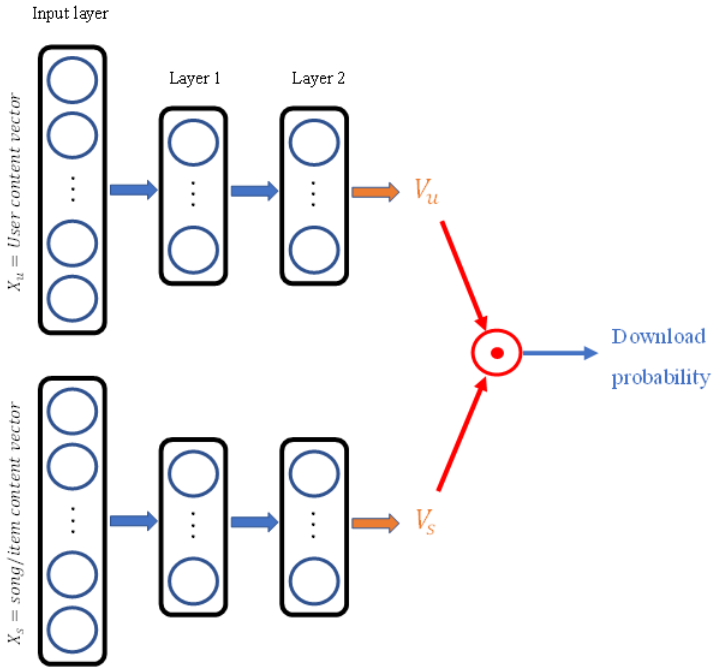



Table A.2: Model Performance

| | **Content-Based Filtering** | | | **Logit Model** | | |
|---|---|---|---|---|---|---|
| Threshold | F1-score | Precision | Recall | F1-score | Precision | Recall |
| 0.1 | 0.66 | 0.52 | 0.92 | 0.50 | 0.33 | 0.97 |
| 0.2 | 0.70 | 0.59 | 0.86 | 0.51 | 0.36 | 0.89 |
| 0.3 | 0.71 | 0.64 | 0.79 | 0.55 | 0.42 | 0.79 |
| 0.4 | 0.70 | 0.68 | 0.73 | 0.49 | 0.52 | 0.47 |
| 0.5 | 0.70 | 0.73 | 0.67 | 0.19 | 0.57 | 0.11 |
| 0.6 | 0.67 | 0.77 | 0.59 | 0.05 | 0.59 | 0.02 |
| 0.7 | 0.62 | 0.81 | 0.50 | 0.01 | 0.51 | 0.00 |
| 0.8 | 0.55 | 0.86 | 0.40 | 0.00 | 1.00 | 0.00 |
| 0.9 | 0.42 | 0.92 | 0.27 | 0.00 | 0.00 | 0.00 |

Note: Models were trained on a random 80% sample and tested on the remaining 20%. The results show that the content-based filtering approach outperforms the logit model in predicting download probability based on the holdout-sample performance.

Table A.3: Variables Used in User Content Vector

| | |
|---|---|
| Expert | Binary variable equal 1 if the self-reported expertise(asked for advice) of the participant is greater than 2 |
| Age | Age of the participant |
| Male | Binary variable equal 1 if the participant is male |
| Rate_avg | Average of the ratings given by participant |
| Asked advice | compared with your circle of friends, how likely are you to be asked for advice about music?<br>(1=much less likely, 2=less likely, 3=more likely, 4=much more likely) |
| Int_connection | Participant's type of Internet connection<br>(0=broadband , 1=dial-up) |
| Ability_www | How would you rate your ability to use the World Wide Web?<br>(1=excelent, 2=good, 3=fair, 4=poor, 5=very poor) |
| Visited | In the past 30 days, how often have you visited a site for information about music or concerts?<br>(0=never , 1=1-2 times , 2=3-5 times , 3=more than 5 times) |
| Purchased | Ever purchased a record as a result of hearing it on the web?<br>(0=no , 1=yes) |
| Songs_dl | Approximately, how many songs have you downloaded in the past 30 days?<br>(self-reported value) |
| Email_given | Email address provided by participant (optional);<br>0=no email given , 1=email given |
| Country_live | Hot 1 vector of size 5 for participant's country<br>(1=US , 2=UK, 3=Canada, 4=Brazil, 5=Other) |
| Place | Hot 1 vector of size 4 for participant's place while participating<br>(1=home, 2=office, 3=puplic computer at school, 4=other) |
| Avg_rate_class1 | Average rating given by the participant to songs of genre Classical |
| Avg_rate_class2 | Average rating given by the participant to songs of genre Country |
| Avg_rate_class3 | Average rating given by the participant to songs of genre Disco |
| Avg_rate_class4 | Average rating given by the participant to songs of genre Hiphop |
| Avg_rate_class5 | Average rating given by the participant to songs of genre Jazz |
| Avg_rate_class6 | Average rating given by the participant to songs of genre Metal |
| Avg_rate_class7 | Average rating given by the participant to songs of genre Pop |
| Avg_rate_class8 | Average rating given by the participant to songs of genre Reggae |
| Avg_rate_class9 | Average rating given by the participant to songs of genre Rock |
| Social | Binary variable equal 1 if the participant is in the social influence group |

Table A.4: Variables Used in Song Content Vector

| | |
|---|---|
| Dl_rate | Download rate of the song by the participants in the experiment |
| Q_song | Quality of the song, the average rating for the song by participants in the independent (control) group |
| Dls_obs | Observed popularity for the song by the participant (0 for control group participants) |
| Chroma_stft | Chroma features represent the distribution of energy in the chromatic scale, Short-Time Fourier Transform (STFT): The power spectrum of each frame is computed using the Fast Fourier Transform (FFT) |
| RMSE | Root Mean Square Energy of audio signals |
| Spectral_centroid | Spectral centroid of the song, information about the frequency content of a sound signal |
| Spectral_bandwidth | Spectral bandwidth of the song, information about distribution of energy in the signal |
| Rolloff | Rollof, information about the spectral shape or distribution of energy in the signal |
| Zero_crossing_rate | Zero-Crossing Rate (ZCR) is a measure of the rate at which the audio signal changes its sign |
| MFCC1 to MFCC20 (20 features) | Mel-Frequency Cepstral Coefficients (MFCCs) are a feature widely used in speech and audio processing. They represent the short-term power spectrum of a sound |
| Blues | Classifier's predicted probability of song belonging to genre Blues |
| Classical | Classifier's predicted probability of song belonging to genre Classical |
| Country | Classifier's predicted probability of song belonging to genre Country |
| Disco | Classifier's predicted probability of song belonging to genre Disco |
| Hiphop | Classifier's predicted probability of song belonging to genre Hiphop |
| Jazz | Classifier's predicted probability of song belonging to genre Jazz |
| Metal | Classifier's predicted probability of song belonging to genre Metal |
| Pop | Classifier's predicted probability of song belonging to genre Pop |
| Reggae | Classifier's predicted probability of song belonging to genre Reggae |
| Rock | Classifier's predicted probability of song belonging to genre Rock |

## D Counterfactual Setup

In this section, we provide a detailed explanation of how we set up our counterfactual analyses. Our counterfactual scenarios are conducted under the following conditions:

- We use all the participants and their actual demographic as our data for the counterfactual simulations. We observe a participant's behavior only under one condition. In other words, we consider an identical counterpart of each participant in the data who samples and downloads under each counterfactual environment.

- Note that there is an unobserved search cost for each participant denoted by $\alpha_{i0}$. For each participant and her corresponding counterparts, we randomly draw a $\alpha_{i0}$ from the normal distribution with parameters estimated from our empirical estimations in Table 3. We repeat each simulation 100 times and report the average results.

- For each participant and her counterparts, we use the dynamically updated popularity based on the download behavior of previous participants in that world under the existing counterfactual design. The analyses will run separately at each world based on the true ordering of the people as it is in the data in each experimental world. We run the counterfactual for the first person in the world and save his/her new download record. Then, we use this new record to show the following participant in the list, i.e., the next participant will see the download behavior of the first person under the new counterfactual. We will continue the above algorithm for all participants. Therefore, this approach provides us with comparable counterfactual environments to quantify the participants' search behavior under new design formats and social information disclosure following the actual dynamics under a different counterfactual setting.

- The download decision for song $j$ sampled by participant $i$ is drawn from a Bernoulli distribution with success probability $Bern(p_{ij}^{post})$, where $p_{ij}^{post}$ is the download probability of song $j$ by participant $i$. To precisely run the counterfactuals, we predict the Bernoulli distribution $Bern(p_{ij}^{post})$ by using an advanced machine-learning approach known as the "*content-based filtering*." The model consists of two neural networks to generate the user and item vectors whose dot product will be given to a sigmoid activation nod to predict the download probability. In simpler terms, if we can identify participants with similar music preferences and personal characteristics, we will be able to predict the download probability of each song within this group

of similar participants. Thus, we predict the download probability of song $j$ by participant $i$ by using neural networks based on features such as genre, tempo, key, instruments used, and displayed popularity (conditional on being in CF2 and CF3), along with the participant's characteristics provided in the survey data, as well as their rating and downloading behavior. A thorough elucidation of this approach and its accuracy can be found in Web Appendix C.

The three metrics, (1) $l$ the average number of songs sampled, (2) $k$ the average number of songs downloaded, and (3) $w$ the search efficiency, in counterfactual results (Table 5) are measured as follow:

Under each counterfactual, we run $D = 100$ simulations. In each simulation $d$, for each of $N = 9540$ participants, we randomly draw a $\alpha_{i0}$ from the normal distribution with parameters estimated from our empirical estimation in Table 3 to capture the unobserved based cost heterogeneity. With respect to each draw, we run the counterfactual setting in which the download decision for the song $j$ sampled by participant $i$ is drawn from a Bernoulli distribution with success probability $Bern(p_{ij}^{post})$, where $p_{ij}^{post}$ is the predicted download probability using content-based filtering. The popularity of songs for each participant is updated based on the download behavior of the previous participant in each simulation. For each participant $i$, we have the number of samples, the number of downloads, and search efficiency denoted by $l_{id}$, $k_{id}$, and $\zeta_{id}$, respectively. Thus, for simulation $d$, we have the average performance over all $N$-participants as $l_d = \frac{1}{N}\sum_{n=1}^{N} l_{id}$, and $k_d = \frac{1}{N}\sum_{n=1}^{N} k_{id}$, and $\zeta_d = \frac{1}{N}\sum_{n=1}^{N} \zeta_{id}$. Table 5 summarizes the aggregated results, with the numbers in parentheses indicating the standard deviations for outcomes across the 100 repeated simulations where $l = \frac{1}{d}\sum_{d=1}^{D} l_d$, and $k = \frac{1}{d}\sum_{d=1}^{D} k_d$, $\zeta = \frac{1}{d}\sum_{d=1}^{D} \zeta_d$.

# E The Experiment Details

## The Experimental Design

Our description here closely follows Salganik (2008), which explains the experimental design and data collection in detail. The MusicLab experiment was conducted at the Department of Sociology at Columbia University by Mathew J.Salganik and Duncan J Watts between 2004 and 2007. The experiment was designed to study the impact of social influence on the outcome of cultural markets. It aimed to disentangle factors that could contribute to the success of popular products like best-selling books, top hit songs, and blockbuster movies. Questions such as the extent to which their success is attributed to intrinsic quality versus social influence, and how social influence shapes market outcomes and inequality, were central to the study. To answer these questions, they empirically constructed a website where participants can sample and listen to new music songs, rate them, and download them upon liking. They use two different experimental conditions: the Independent group and the Social Influence group, where they control the information available to participants about the behavior of others and the popularity of the songs.

In the Independent group, there is no social information, and participants only observe the names of the songs and their band names. In the Social Influence group, in addition to the information available in the independent group, each participant can observe the number of times each song has been downloaded by previous participants. This is an online field experiment where the majority of the participants were recruited through the banner advertisement of the experiment on popular blogs such as www.bolt.com and www.kottke.org. When participants enter the experiment website, they are presented with a welcome screen informing them that this is a study on musical taste where they would be offered free songs by the new up-and-coming artists to download. After the welcome screen, participants fill out a survey that gathers their demographic information. Then, they are shown a page of instructions. Finally, a menu design of 48 songs is shown to participants to sample from. Figures A.2 to A.5 present the screenshots of the experiment's steps.

Figure A.2: Splash Screen from the Website

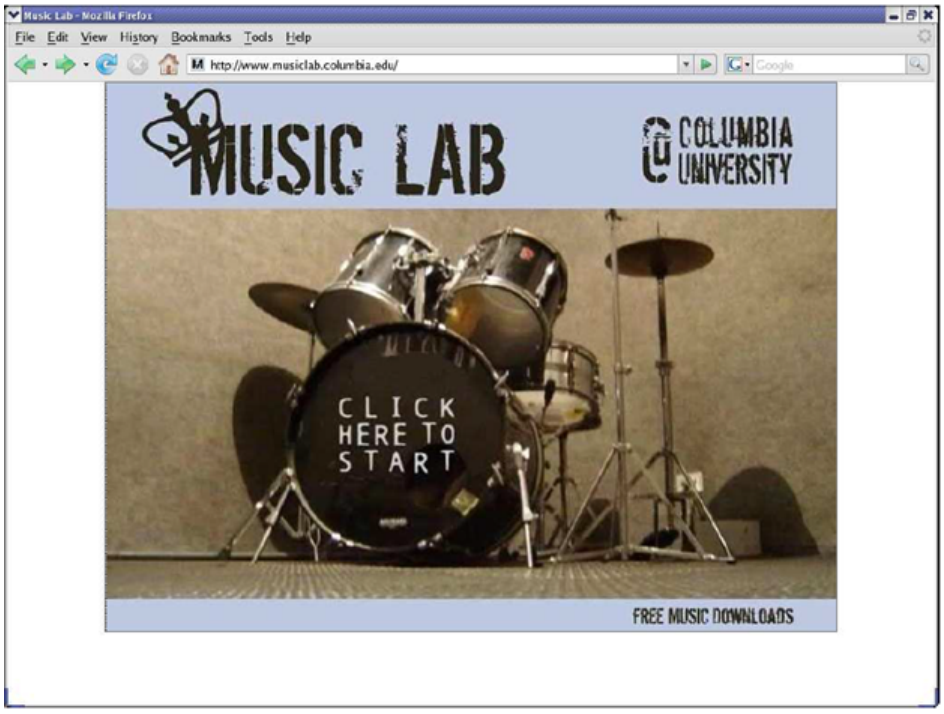



Figure A.3: Welcome Screen from the Website

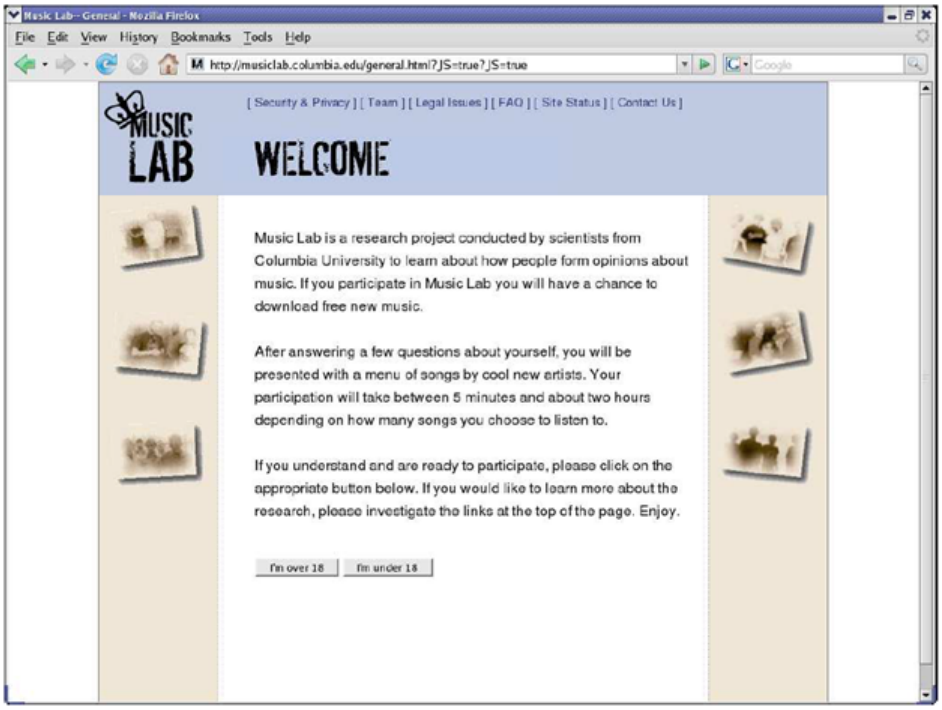


Figure A.4: Survey Screen from the Website

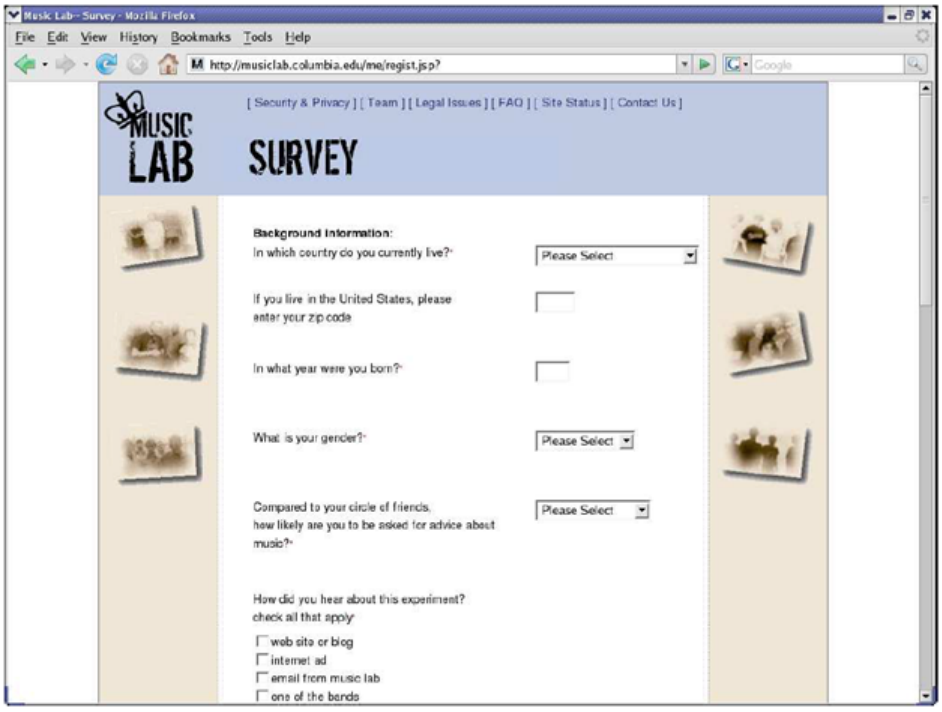
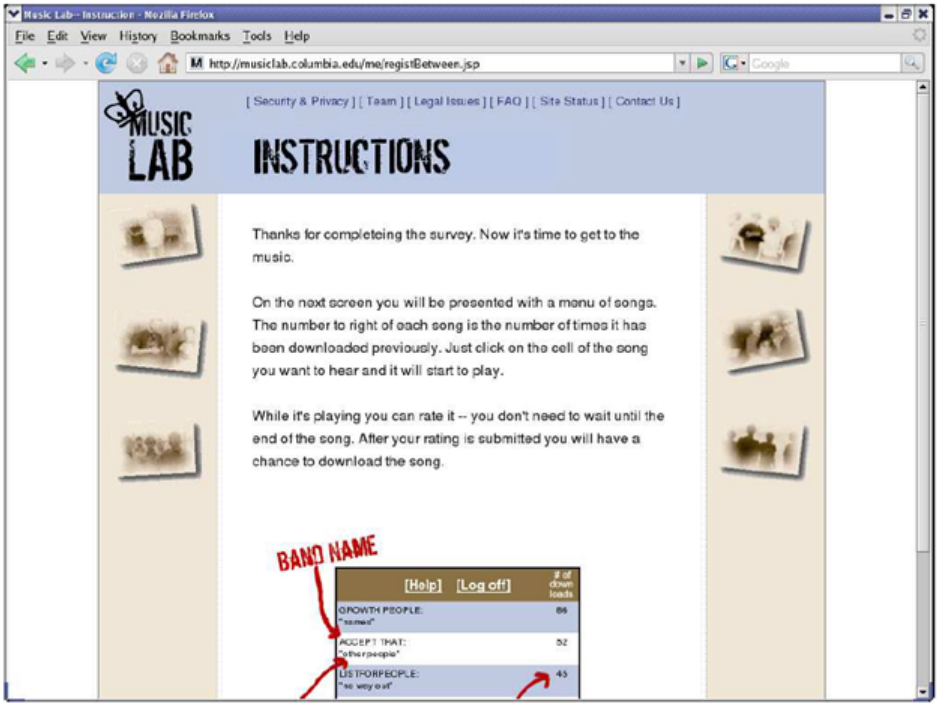



Figure A.5: Instructions Screen from the Website



Two designs are used to display the songs to participants: column and matrix designs.

We first discuss the column design. In this design, songs are displayed in a single column to the participants, where they need to scroll up and down to select a song. In the Social Influence group condition under column design, songs are sorted by their popularity-number of downloads by previous participants. In the Independent group, they are sorted randomly for each participant.

We now turn to the matrix design. In this design, songs are presented in a 16*3 grid that fits the computer screen, and all songs are visible to the participant in one screen. Songs are sorted randomly in both Independent and Social Influence groups.

Regardless of display design, after clicking on a song, the participant is directed to a new screen where the selected song automatically starts to play, and the participant is asked to rate it on a scale of 1 (hate it) to 5 (love it) while listening to it. Upon rating the song, the participant is provided with the download decision for the song, and then the participant is returned to the song menu, where they continue to sample another song or leave the experiment. Figures A.6 to A.11 show the different displays of each conditional group, format, and the rating and downloading steps.

Figure A.6: The Song Menu Screen from the Independent Group with Matrix Format

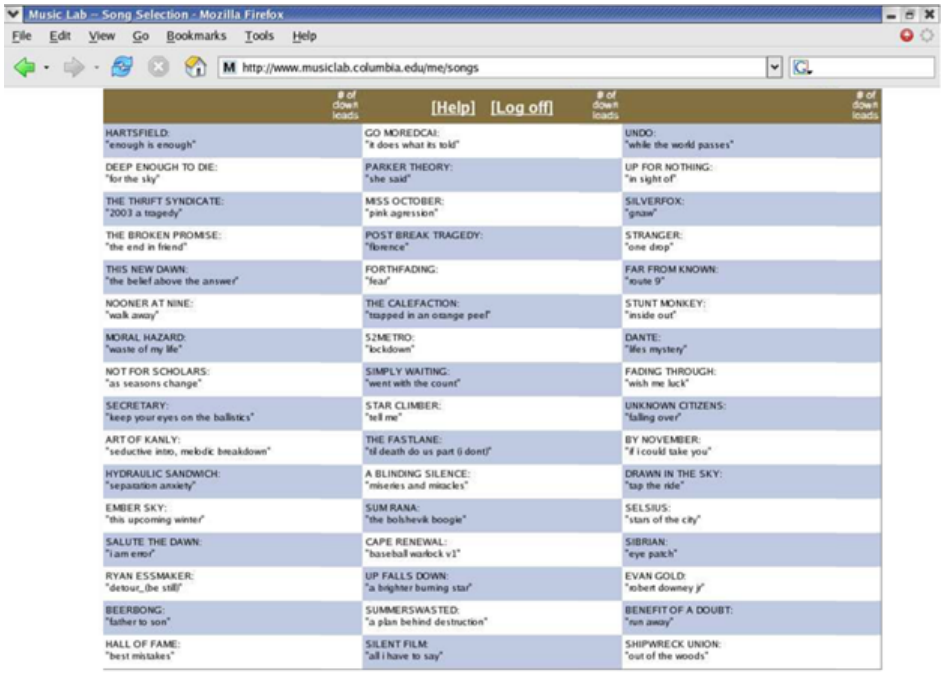



Figure A.7: The Song Menu Screen from the Social Influence Group with Matrix Format

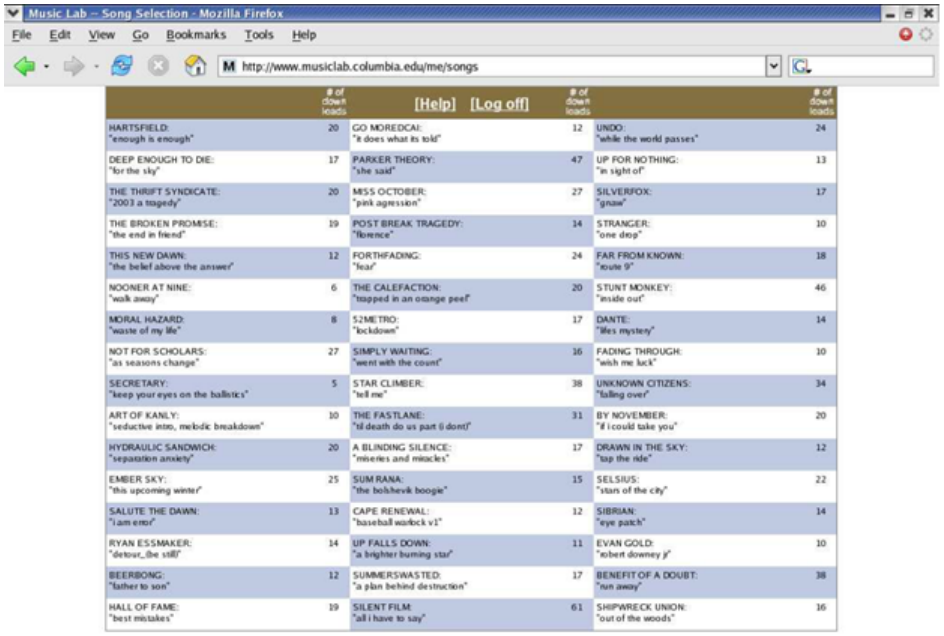


Figure A.8: The Song Menu Screen from the Independent Group with Column Format

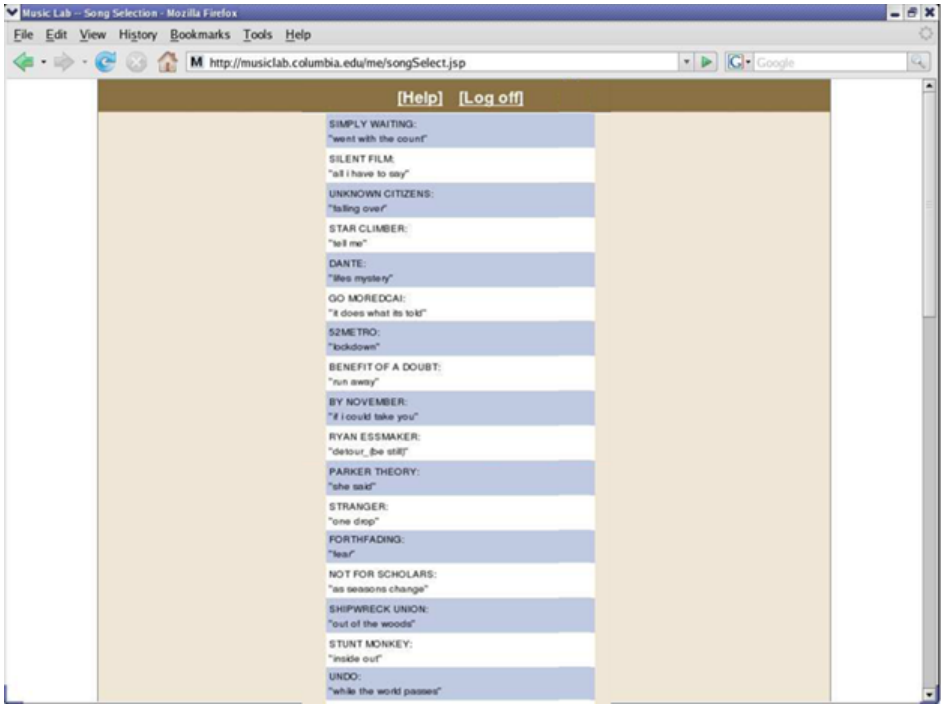



Figure A.9: The Song Menu Screen from the Social Influence Group with Column Format

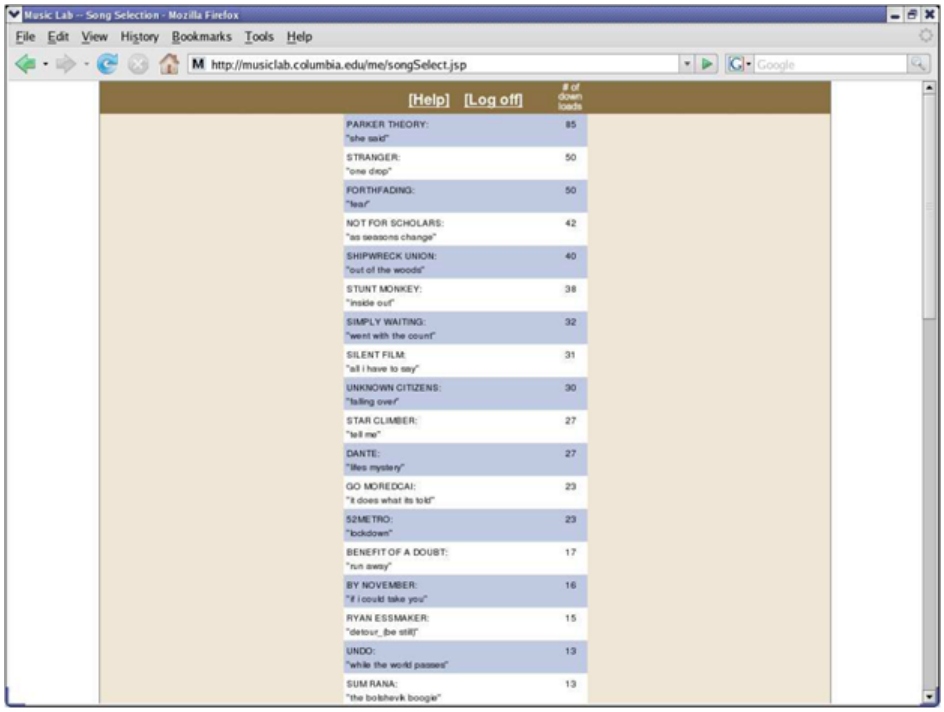


Figure A.10: Listening Screen. While a Song Was Playing, Subjects Were Required to Rate It. Rating Could Be Submitted Before the Song Was Finished Playing.

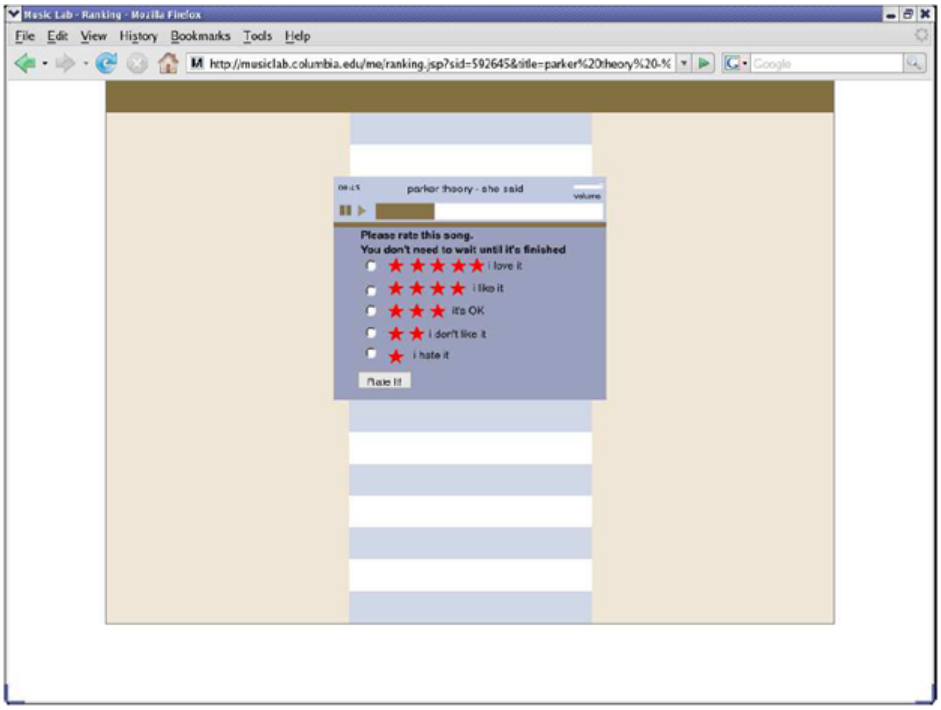
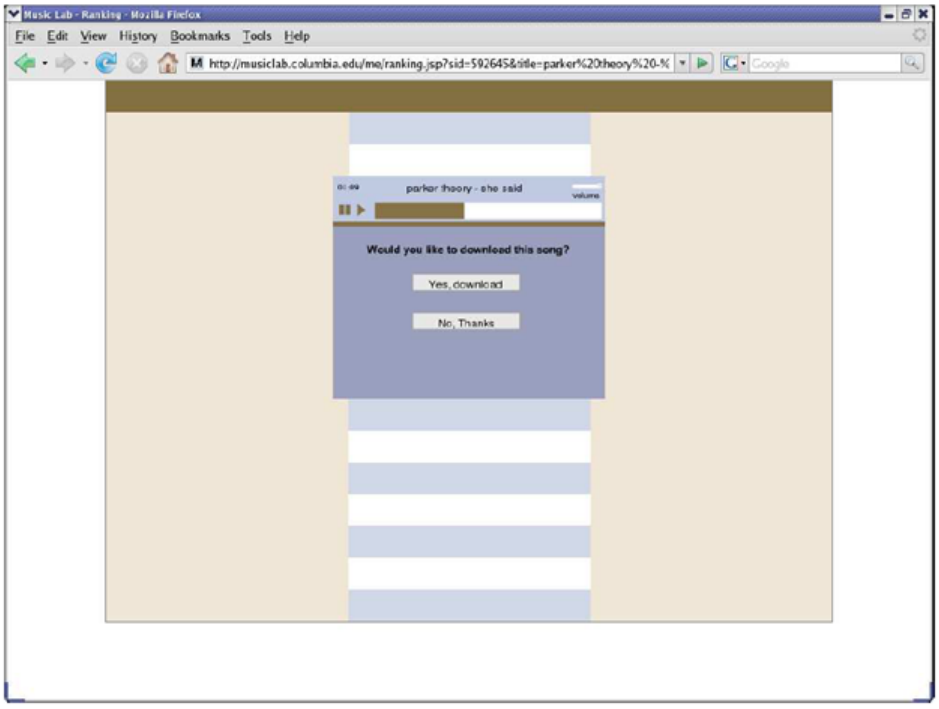



Figure A.11: Download Decision Screen. After Rating the Song, Subjects Had to Decide Whether to Download the Song or Not.



### The Songs

The music bands for the experiment are sampled from www.purevolume.com, a website where bands could create homepages and post their music for download. The sample was restricted to 1,000 premium member bands (those who paid $10 per month for additional features on their homepages) to ensure good audio quality. Initially, 200 bands were selected for consideration, but in order to use songs that were unknown to participants, any band that had played in more than 10 states, had played more than 15 concerts in the past 30 days, had appeared on the Warped Tour, or had 30,000 or more hits on their Purevolume page were screened out. Among the sampled bands, 51 agreed to participate in the study and provided a song of their choice for the experiment.

Preliminary testing revealed that 48 songs were the maximum number that could be legibly presented on a typical computer screen for the matrix design. Therefore, a sample of 48 out of 51 bands was used in the experiments, with one song from each band. To ensure that the screened music was not known to the participants, the list of bands and songs was presented to two experts in popular music, a DJ at the Barnard College student radio station and the music editor for www.bolt.com, neither of whom recognized any of the bands or songs. Additionally, on the registration survey, subjects were asked about their familiarity with five bands, including the three potential bands who agreed to participate but were ultimately not included, an imaginary band, and an extremely well-known band. Table A.5 summarizes the participant's familiarity with these bands. Participants reported having a similar level of familiarity with the fake band (Peter on Fire) as with three actual bands from the sample for the experiment. The survey results, along with the screening and queries to experts, indicate that the songs used in the experiment were unknown to the participants.

Table A.5: Comparing the Popularity of the Sample Bands to a Fake Band.

| | | How familiar are you with the following bands? | | |
|---|---|---|---|---|
| Band type | Name | Don't know it at all (% of participants) | Heard of it (% of participants) | Know it pretty well (% of participants) |
| Real | Guys on Couch | 91.0 | 8.1 | 1.0 |
| Real | Grover Dill | 91.2 | 7.8 | 0.9 |
| Fake | Peter on Fire | 88.1 | 10.5 | 1.4 |
| Real | U2 | 4.6 | 24.6 | 70.9 |
| Real | Remnant Soldier | 83.2 | 14.7 | 2.1 |

Furthermore, to ensure that the data is generated by the appropriate participants in situations that match the experimental design where the participants have no prior knowledge of the songs or any malicious intent, all potentially contaminated data has been excluded from the analysis. The

data in cases where the participant's behavior could have been influenced by information outside of the experimental design have been flagged and excluded. This flagging process began with the survey, in which participants were asked to report how they heard about the experiment. Any data generated by participants who reported hearing about specific songs or bands were flagged for exclusion. Additionally, data generated after the participant logged off or 2 hours passed since registration were also flagged in order to exclude any information that may have been influenced by discussions with friends. Further measures were taken to prevent potential manipulation by band members. Data from participants who reported on the survey that they heard about the experiment from "one of the bands" were flagged for exclusion.

### The Survey

Survey questions and available demographic information in the data:

- Field 1: In which country do you currently live? (numerical code)
  1 = United states, 2 = United kingdom, 3 = Canada, 4 = Brazil, 5 = Other
- Field 2: In which country do you currently live? (string)
- Field 3: If you live in the United States, please enter your zip code.
  00000 = missing value (non-US residents)
- Field 4: In what year were you born? (converted to age)
- Field 5: What is your gender? (numeric code)
  0 = female, 1 = male
- Field 6: compared with your circle of friends, how likely are you to be asked for advice about music? (Opinion leader score)
  1 = much less likely, 2 = less likely, 3 = more likely, 4 = much more likely
- Field 7: How did you hear about the experiment?
  - website or blog (1 = yes, 0 = no)
  - internet ad (1 = yes, 0 = no)
  - email from MusicLab (1 = yes, 0 = no)
  - one of the bands (1 = yes, 0 = no)
  - friend told me about a specific song (1 = yes, 0 = no)
  - friend told me about a specific band (1 = yes, 0 = no)
  - friend told me about the site in general (1 = yes, 0 = no)
  - search engine (1 = yes, 0 = no)
  - other (1 = yes, 0 = no)
- Field 8: What type of Internet connection are you currently using? (numerical code)
  1 = broadband, 2 = dial-up
- Field 9: Where are you while you are participating in this experiment? (numerical code)
  1 = home, 2 = office, 3 = public computer at school, 4 = other

- Field 10: How would you rate your ability to use the World Wide Web? (numerical code)
  1 = excellent, 2 = good, 3 = fair, 4 = poor, 5 = very poor

- Field 11: In the past 30 days, how often have you visited a site for information about music or concerts? (numerical code)
  0 = never
  1 = 1 - 2 times
  2 = 3 - 5 times
  3 = more than 5 times

- Field 12: Have you ever purchased a record as a result of hearing it on the web? (numerical code)
  0 = no, 1 = yes

- Field 13: Approximately, how many songs have you downloaded in the past 30 days?
  self-reported value

- How familiar are you with the following bands?
  Field 14: Guys on Couch (real)
  (1 = don't know it at all, 2 = heard of it, 3 = know it pretty well)
  Field 15: Grover Dill (real)
  (1 = don't know it at all, 2 = heard of it, 3 = know it pretty well)
  Field 16: Peter on Fire (fake)
  (1 = don't know it at all, 2 = heard of it, 3 = know it pretty well)
  Field 17: U2 (real)
  (1 = don't know it at all, 2 = heard of it, 3 = know it pretty well)
  Field 18: Remnant Soldier (real)
  (1 = don't know it at all, 2 = heard of it, 3 = know it pretty well)

- Field 19: Please provide your email address so that we can tell you about the results of Music Lab (optional)
  0 = no email address given
  1 = email address given

# F The Online Music Platforms' Designs in the Market

Figure A.12: Spotify Displays Songs in a Matrix Format With No Popularity

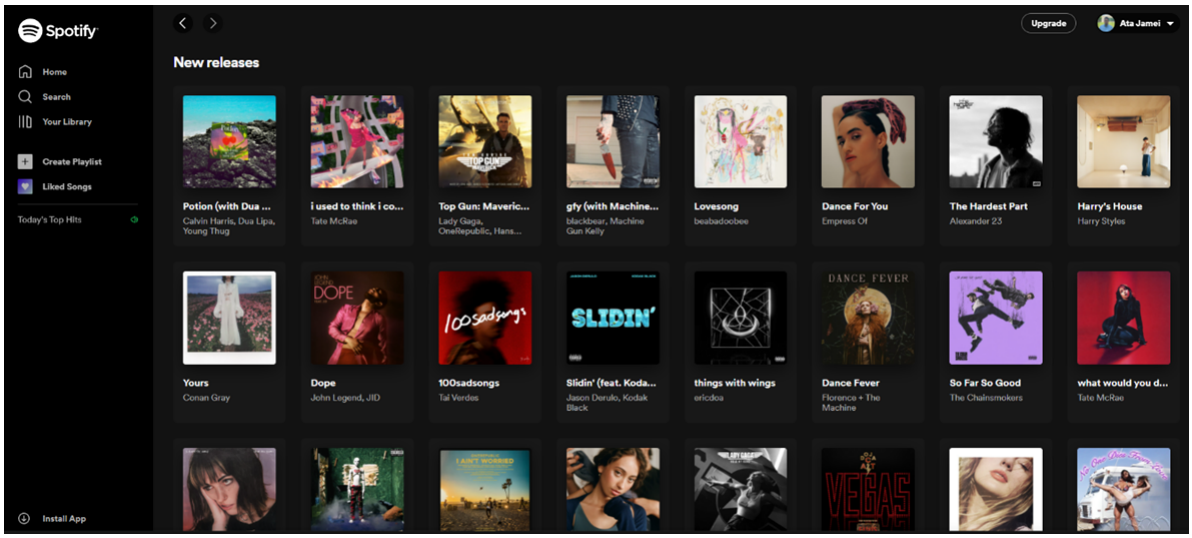



Figure A.13: Spotify Displays Songs in a Column Format With No Popularity

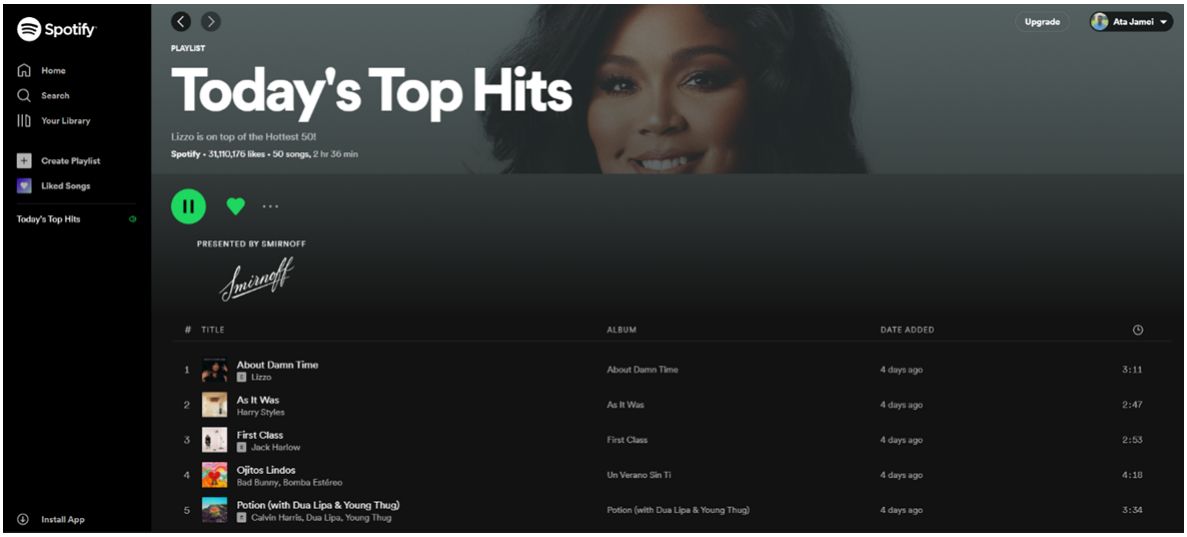


Figure A.14: SoundCloud Displays Songs in a Column Format With Popularity and Random Assortment

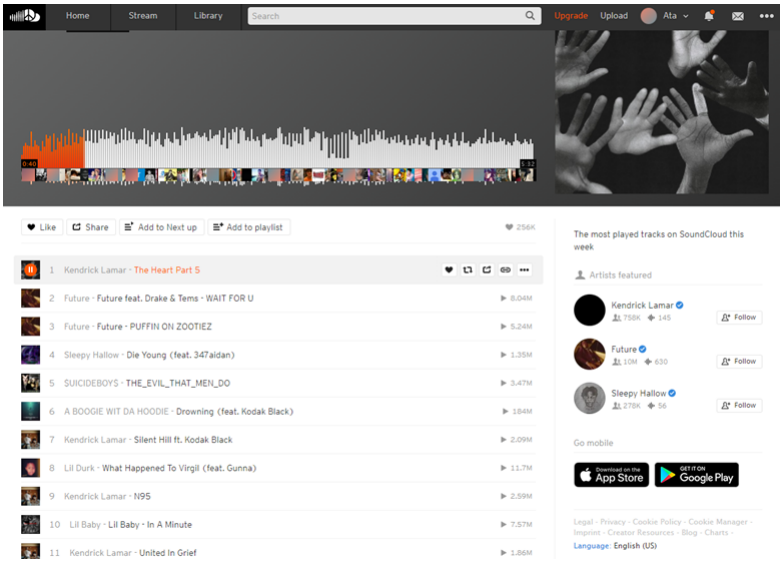



Figure A.15: 7Digital Displays Songs in a Column Format Without Popularity

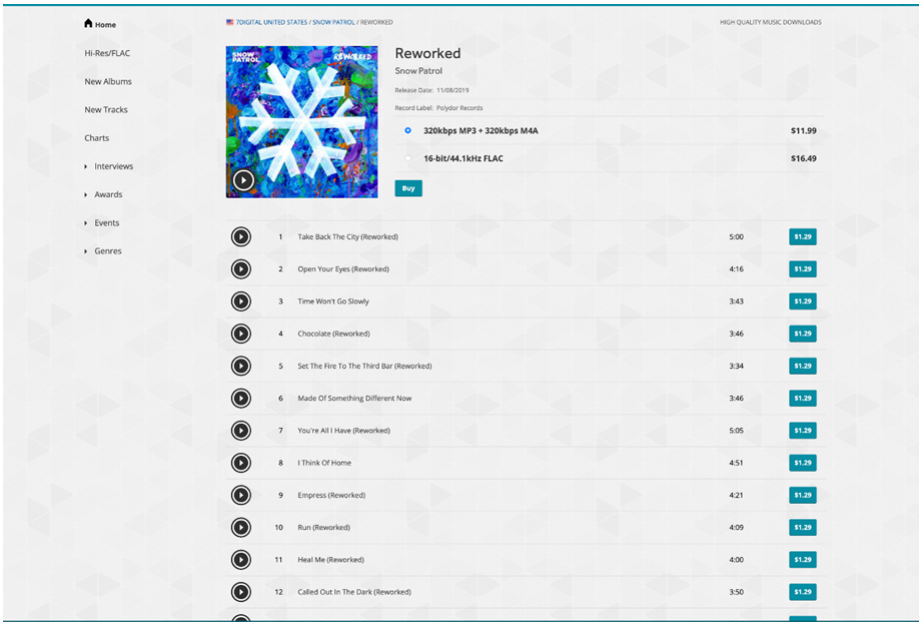
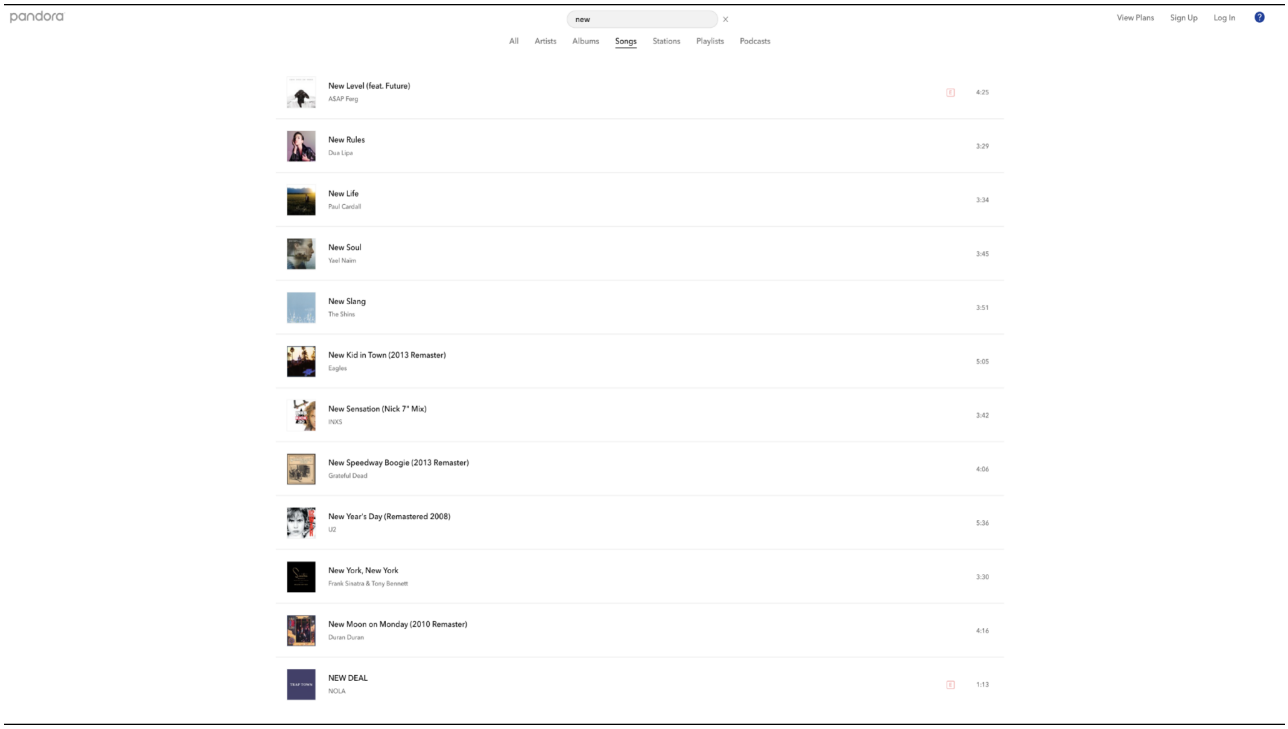


Figure A.16: Pandora Displays Songs in a Column Format Without Popularity



# G The Generalized Weitzman in Pre-search/Post-search Empirical Context

Our study distinguishes itself from previous research by focusing on platforms like Spotify, SoundCloud, and short video streaming (TikTok), where users rely on limited initial information about songs, necessitating a different modeling approach due to the absence of detailed product attributes impacting pre-search utility. Unlike traditional e-commerce platforms, where product attributes are fully observable before purchase, music and video streaming services involve unique challenges. Specifically, pre-search utility and post-search utility may involve different sets of product attributes because users only uncover many of a song's or movie's features after sampling.

This distinction is critical, as traditional search models (e.g., those used in online shopping contexts) typically assume that most relevant product attributes are known upfront, informing pre-search utility calculations (Kim et al. 2010, Ursu et al. 2023). In contrast, the hidden attributes revealed post-sampling play a significant role in streaming platforms, necessitating models that can dynamically adjust based on new information obtained during the search process.

Unlike studies in less heterogeneous markets where users observe almost all product features before sampling (Kim et al. 2010, Ursu et al. 2023), our empirical context poses unique challenges. While our method is tailored to our specific context, the question arises: can the generalized Weitzman approach be adapted to such contexts with multiple purchasing options? In this Web appendix, we explore an alternative proposed approach, using Olszewski and Weber (2015a,b) that may be applicable in these scenarios based on the pre-search/post-search framework.

We follow the notation in Ursu et al. (2023), which outlines the sequential empirical search models accounting for observed product features and pre-search utility considerations. For example, Kim et al. (2010) applied the sequential search model in purchasing camcorders on Amazon, and Ursu (2018) considered booking hotels on the Expedia platform. Therefore, researchers assume pre-search and post-search utilities for each item in the market. They assume consumer $i$'s utility from product $j$ has two additively separable components:

$$u_{ij} = X_j^{'}\beta_j + \epsilon_{ij} \tag{A.1}$$

where $X_j^{'}$ denote the observable product $j$'s attributes pre-search behavior and $\epsilon_{ij}$ denotes the part of utility that is only known by the consumer after search (called "post-search taste shock"). Therefore, consumer $i$ evaluates the Weitzman reservation value of product $j$ based on pre-search

utility obtained as $X_j^{'}\beta_j$. In this approach, researchers assume that $\epsilon_{ij}$ is normally distributed taste shock that is revealed after sampling. Thus, consumer $i$ decides the purchase behavior based on $u_{ij}$ at the purchase stage. These papers assume that consumer $i$ purchases only one item per search session and chooses the item that provides the maximum utility $u_{ij}$ among the sampled products. Therefore, researchers use the Weitzman model to formulate the optimal sequential search behavior.

However, if consumer $i$ is enabled to purchase multiple items per search session, the classic Weitzman model cannot be applied anymore. Here, it is possible to follow Olszewski and Weber (2015a,b) to expand our modeling approach in the above context. Let's assume that $x_{iSt}^o$ is the set of products that have been sampled until the search step $t$. Now, researcher can define a search utility $u(.)$ as follows:

$$u(x_{iSt}^o) = \sum_{j \in X_{iSt}^o} d_{ij} \times \psi(u_{ij}), \tag{A.2}$$

where $d_{ij}$ is an indicator which is 1 if consumer $i$ purchases product $j$; otherwise it is zero. Equation A.2 means consumer $i$ receives an accumulated reward of purchasing some items until search step $t$. This utility is based on reward $u_{ij}$ in box $j$, which can be parameterized in product attributes that have been observed in pre-search behavior. If $\psi$ is a concave function, we can show that the utility function $u(.)$ is a quasi-concave function. This is mathematically equivalent to consumers having convex preferences over products, which is a common assumption in choice theory. Then, researchers can follow Theorem 1 in Olszewski and Weber (2015a) to solve the optimal sequential search model.

The estimation can be done by following *logit-smoothed AR simulator* (Honka 2014, Honka and Chintagunta 2017, Ursu 2018) or using *GHK-type simulator* (Jiang et al. 2021). Note that the reservation values will be different than $x_j^*$ used in this paper. Since the above approach is not in the scope of our empirical context and our research questions in this paper, we leave the above extension for future research.