Personalized Recommendation for Balancing Content Generation and Usage on Two-Sided Entertainment Platforms

Hao Arthur Zhang
Zhejiang University, haozhang.2021@phdis.smu.edu.sg

Zhiling Guo
Singapore Management University, ZHILINGGUO@smu.edu.sg

Mingzheng Wang
Zhejiang University, wangmzh@zju.edu.cn

Personalized Recommendation for Balancing Content Generation and Usage on Two-Sided Entertainment Platforms

Completed Research Paper

Hao Zhang
Zhejiang University
Singapore Management University
Singapore
haozhang.2021@phdis.smu.edu.sg

Zhiling Guo
Singapore Management University
Singapore
zhilingguo@smu.edu.sg

Mingzheng Wang
Zhejiang University
Hangzhou, China
wangmzh@zju.edu.cn

Abstract

Online entertainment platforms such as Youtube host a vast amount of user-generated content (UGC). The unique feature of two-sided UGC entertainment platforms is that creators' content generation and users' content usage can influence each other. However, traditional recommender systems often emphasize content usage but ignore content generation, leading to a misalignment between these two goals. To address the challenge, this paper proposes a prescriptive uplift framework to balance content generation and usage through personalized recommendations. Specifically, we first predict the heterogeneous treatment effects (HTEs) of recommended contents on creators' content generation and users' content usage, then consider these two predicted HTEs simultaneously in an optimization model to determine the recommended contents for each user. Using a large-scale real-world dataset, we demonstrate that the proposed recommendation method better balances content generation and usage and brings a 42% increase in participants' activity compared to existing benchmark methods.

Keywords: Recommender systems, prescriptive analytics, user-generated content, content generation and usage, heterogeneous treatment effects

Introduction

The advancement of mobile technology and the rising entertainment demands of consumers have contributed to the unprecedented growth of the digital entertainment industry (Foutz 2017). For example, the global video streaming market is estimated to grow from \$7.5 billion in 2021 to \$17.5 billion in 2026 at a compound annual growth rate of 18.5% (MarketsandMarkets 2021). Meanwhile, the global music streaming market size is valued at \$25.1 billion in 2021 and is expected to grow to \$37.8 billion in 2026 (Statista 2021). The prosperity of the entertainment industry has led to the rapid development of many two-sided entertainment platforms such as YouTube, TikTok, and Triller for video streaming services and Spotify, Pandora, and Tidal for music streaming services. To illustrate, Spotify, the world's largest music streaming service provider, has 406 million monthly active users (including 180 million premium subscribers) with a year-over-year growth rate of 18% respectively as of February 2022.[1]

The rapid growth of the two-sided entertainment platforms is largely attributed to its rich collection of user-generated content (henceforth UGC) that continuously attracts users. On these two-sided UGC entertainment platforms, creators' generated contents can influence users' content usage behaviors (e.g., Likes and Shares), and users' content usage behaviors subsequently provide creators with performance feedback (e.g., the number of likes and shares) which further influences creators' content generation.[2] However, maximizing creators' content generation and users' content usage simultaneously can be an incompatible goal for platforms (Jannach and Adomavicius 2016). For example, if platforms were to always recommend the most relevant contents to users, some new or less popular contents would not be able to receive enough exposure. Such a lack of feedback from users would reduce creators' incentives for new content generation, potentially leading to content underproduction and eventually negatively affecting platforms' revenue (e.g., advertising revenue) and user retention (Huang et al. 2019). Van Mierlo (2014) find that only 1% of participants actively produce contents. It is thus very important to design recommender systems to balance content generation and usage.

However, most existing recommendation methods mainly focus on maximizing users' content usage, which essentially overlooks content generation (Jannach and Adomavicius 2016). In addition, current recommendation methods make recommendations based on the predicted user response to recommended contents (Besbes et al. 2016, Lian et al. 2018, Wu et al. 2019), which cannot identify the heterogeneous treatment effects (HTEs) of recommendations on content usage and content generation. Without considering HTEs, platforms may not optimize content generation and usage effectively (McFowland III et al. 2021).

This research aims to balance content generation and usage through personalized recommendations on two-sided UGC entertainment platforms. We answer the following research questions: How to identify the HTE of recommendation on creators' content generation and users' content usage? Based on the identified HTEs, how to design an optimal content recommendation method to balance content generation and usage? We integrate the causal inference into the optimization model and propose a prescriptive uplift framework to make the recommendation decisions. Specifically, we first develop an Engagement-based Click-interact Input-output hidden Markov model to predict user activity which measures the content usage level of a user. Based on the predicted activity, we use the linear double machine learning model (LDML) to predict incremental user activity and incremental creator activity, which represent the HTEs of recommendation on content usage and content generation, respectively. Finally, we propose an optimization model to take into account these two HTEs simultaneously to balance content generation and usage and make personalized recommendation decisions.

This paper makes several methodological contributions to the literature. First, we propose a new method to effectively balance creators' content generation and users' content usage, which is more suitable to make recommendations on two-sided UGC entertainment platforms. Second, we identify the HTE of recommendation on creators' content generation and users' content usage, which is an improvement over the traditional correlation-based models. Third, the proposed method is generalizable to other two-sided platforms that need to balance content generation and usage.

We validate our proposed framework using real-world instances from NetEase Cloud Music. The results show that our proposed model makes content generation and usage more balanced and brings a 42% increase in participants' activity compared to existing benchmark methods. The significant improvement in participants' activity means that our method not only encourages users' consumption of existing contents but also incentivizes creators' generation of new contents. Our model identifies the HTEs of various factors (e.g., novelty and popularity) on users' content usage behavior and creators' content generation behavior. It further helps identify different user types (e.g., novelty-driven and popularity-

---

[1] https://newsroom.spotify.com/2022-02-02/spotify-reports-fourth-quarter-2021-earnings/.

[2] Although some two-sided entertainment platforms also include professionally-generated content (PGC), we mainly focus on the UGC on these platforms in this paper. Thus, we refer to the UGC part of these platforms as two-sided UGC entertainment platforms.

driven) and creator types (e.g., popularity-driven). These findings provide important managerial insights into the design of personalized recommendation strategies on two-sided UGC entertainment platforms.

The remainder of this paper is organized as follows. Section 2 reviews the relevant literature. Section 3 introduces the proposed prescriptive uplift framework. Section 4 presents the experimental results using real-world data. Section 5 discusses the managerial implications on user and creator behaviors. Section 6 concludes the paper with directions for future research.

Literature Review

Our research is broadly related to two streams of literature: content generation and usage, and product recommendation. In this section, we briefly review related work in each category.

Content Generation and Usage

Current studies on UGC have focused on online reviews from e-commerce platforms and review websites and posts from social media (Archak et al. 2011, Deng et al. 2021, Yang et al. 2019). It has been documented that UGC is an effective tool for supporting consumers' decisions and improving firms' decisions (Chen et al. 2011, Abbasi et al. 2019). However, most of the existing works only focus on either generation or usage of UGC but not both. In contrast, we study UGC on entertainment platforms (e.g., music and video) and consider both creators' generation and users' usage for UGC simultaneously.

Prior studies have found that creators' content generation behaviors are influenced by such factors as performance feedback, financial incentives, virtual rewards, and cumulative past contributions (Moon and Sproull 2008, Jabr et al. 2014, Aaltonen and Seiler 2016, Huang et al. 2019, Burtch et al. 2021, Liu and Feng 2021). Especially, the lack of performance feedback can prevent creators from generating new contents, which leads to the underprovision of UGC (Zhang and Zhu 2011). Because the underprovision problem can affect the sustainability of platforms, motivating creators to contribute content is vital for UGC platforms. In this research, we consider both the direct effect of users' performance feedback information and the indirect effect of platforms' recommendation decisions on creators' content generation.

Previous studies have also found that users' content usage behaviors are influenced by many factors including user engagement, user characteristics (e.g., preference, preference diversity, demographics), content characteristics (e.g., popularity, novelty), and time factors such as the hour of the day and the day of the week (Shim et al. 2008, Singh et al. 2014, Dewan et al. 2017, Liu and Toubia 2018, Song et al. 2019, Ghose et al. 2019a, Burtch et al. 2021). User engagement denotes a state of being involved, occupied, retained, and intrinsically interested in something (Pagani and Mirabello 2011). Besbes et al. (2016) find that user engagement state can influence users' browsing path on media sites. Zhang et al. (2019) consider four stages (i.e., aware, exploring, active, addicted) of user engagement and find that different stages can affect users' mobile content usage behaviors (e.g., giving up reading, reading by pay-per-content, and reading by subscription) differently. Cohen (2009) and Dietrich and Livingston (2012) model consumer behaviors as a conversion funnel that consists of awareness, consideration, conversion, loyalty, and advocacy stage. Following this prior literature, we adopt the conversion funnel framework to measure user engagement during content usage, which is more granular than other existing studies about engagement. Moreover, we consider users' sequential content usage process (i.e., first click the content on the summary page and then interact with the content on the detail page) to develop an Engagement-based Click-interact Input-output hidden Markov model to predict user activity. Our comprehensive model helps generate rich insights into users' heterogeneous usage behaviors.

Product Recommendation

Product recommendation has been widely studied in the existing literature (Adomavicius and Tuzhilin 2005, Ricci et al. 2011). In the context of digital products such as news, music, and video, advanced methods make the content recommendation based on consumer behaviors. For example, by considering the visit path of readers, Besbes et al. (2016) design a path-focused heuristic algorithm to make dynamic content recommendations to maximize the value generated by clicks. Song et al. (2019) propose a multicategory utility model that captures consumers' preference diversity dynamically in making the personalized content recommendation. However, current recommender systems typically focus on the usage-oriented goal but generally overlook the creation-oriented goal. Ghose and Han (2011) and Jannach and Adomavicius (2016) show that there is an inherent conflict between content generation and usage. This makes existing methods unsuitable for making recommendations on the two-sided UGC platforms. We address this research gap by proposing a new method to balance content generation and usage on two-sided UGC platforms.

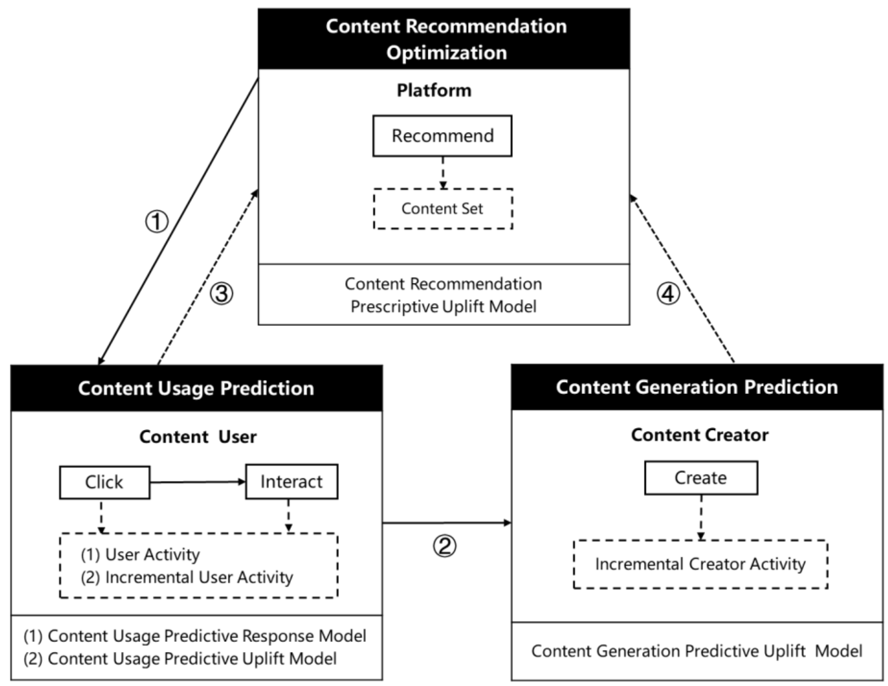
Compared with utilitarian products (e.g., computers, cars), entertainment products (e.g., music, video) are experiential in nature and typically have a short lifecycle (Foutz 2017). There are several unique aspects that need to be considered in making product recommendation decisions. First, the performance feedback is important for users to infer the entertainment products' quality before consumption because entertainment products usually have many intangible attributes (e.g., tempo, danceability) that are hard to characterize. It also affects platforms' recommendation decisions. Second, the demand curves of entertainment products such as songs usually exhibit a rapid exponential decay and their demand is highly uncertain (Moe and Fader 2003, Foutz 2017). Thus, the recommendation must be accurate enough to maximize the value of entertainment products during such a short lifecycle. We address the above challenges by considering the performance feedback of users and identifying the HTEs of recommendations on content generation and usage. We propose a systematic framework to estimate the HTEs and integrate it into the recommendation model to balance content generation and usage.

Current recommendation methods are mainly based on the correlation relationship, that is, estimating the user response to products with the choice models, neural networks, probabilistic graphical models, latent factor models, and collaborative filtering methods (Besbes et al. 2016, Lian et al. 2018, Peake and Wang 2018, Ghose et al. 2019b, He et al. 2019, Wu et al. 2019). Different from this stream of research, we use the uplift model to estimate the HTEs of recommendation on user activity and creator activity. Uplift model is a causal inference technique that estimates the HTE of a treatment on an individual's behavior. An overview of the uplift model is provided by Kane et al. (2014), Gutierrez et al. (2017), and Devriendt et al. (2018). Compared with the field experiment that only identifies the group-level treatment effect, uplift model can estimate the individual-level treatment effect (McFowland et al. 2021). Besides, in the context of content recommendation, each recommended content is regarded as a treatment, the number of treatments is too large to implement in the field experiment and to use the traditional causal inference methods (e.g., propensity score matching, difference in differences). Most uplift models in the existing literature mainly belong to the predictive uplift model category, which only predicts the HTE of a treatment. In contrast, we develop a prescriptive uplift model that not only predicts the HTE of each treatment but also optimizes it to make the optimal decision. Only a few recent studies have adopted the prescriptive uplift model. Lo and Pachamanova (2015) study how to assign multiple marketing treatments to maximize the total number of incremental responders in email marketing campaigns. Alley et al. (2022) first estimate consumers' heterogenous price sensitivity and then price ticket reselling to maximize revenue. McFowland et al. (2021) consider the treatment heterogeneity to target individuals in blood donations and referral marketing campaigns to maximize the total profit. To the best of our knowledge, we are the first to apply the prescriptive uplift model to personalized content recommendation on two-sided UGC platforms.

Model

As shown in Figure 1, the platform's recommendation decisions can influence users' content usage (①), which subsequently influences creators' content generation (②). The platform should consider both effects (③ and ④) when making the recommendation decisions. Thus, we propose a prescriptive uplift framework to combine user and creator sides to optimize the platform's content recommendation decisions. This framework includes three interconnected modules: content usage prediction, content generation prediction, and content recommendation optimization.

Specifically, we first propose a *content usage predictive response model* for predicting *user activity*, which represents users' content usage level. We then develop a *content usage predictive uplift model* for predicting *incremental user activity*, which represents the HTEs of the platform's recommendation decisions on users' content usage. Because users' content usage provides performance feedback for creators' content generation, building upon the predicted user activity, we further develop a *content* *generation predictive uplift model* for predicting *incremental creator activity*, which represents the HTEs of the platform's recommendation decisions on creators' content generation. Finally, based on predicted incremental user and creator activity, we build a *content recommendation prescriptive uplift model* to optimize content recommendation decisions and balance content generation and usage. Figure 1 shows the general structure of the proposed prescriptive uplift framework as well as the decision process of the platform, the content user, and the content creator. We next present the prescriptive uplift framework in detail.



Notes. The figure presents three inter-connected modules. For each module, the top box represents the predictions or decisions of the platform, the solid rectangle in the middle box represents the decisions of each agent, the dotted rectangle in the middle box represents the results of each decision, and the bottom box represents the models for making predictions or decisions.

Figure 1. Proposed Prescriptive Uplift Framework

Content Usage Prediction

Content Usage Predictive Response Model

We describe users' content usage process as an engagement-based, click-interact decision process. During content usage, user engagement is dynamically evolved and characterized as a four-stage conversion funnel including *awareness-consideration*, *conversion*, *loyalty*, and *advocacy*. When the content is recommended to a user, she immediately becomes aware of the content and might engage in information-seeking activities like examining the number of Likes that the content receives (i.e., *awareness-consideration* stage). After that, the user decides whether to click (*CL*) the content, visit the personal homepage (*VPH*) of the content creators, and view comments (*VC*) on the content. We refer to these activities as users' behaviors in the *conversion stage*. Afterward, she might spend a lot of time viewing the content (we denote the viewing time as *VT*) and may follow (*F*) the creator of the content. We characterize this stage as the *loyalty stage*. Finally, she may further progress into the *advocacy stage* by giving likes (*L*) and comments (*C*), and even sharing (*S*) her experience with friends and family.

On the platform, users first access limited content information (e.g., number of likes) from the summary page, and then obtain additional content information (e.g., creator profile) from the detail page after clicking on the content of interest. Therefore, we model users' content usage as a click-interact sequential decision process. Click (CL) belongs to the click stage and the rest of the actions (i.e., VC, VPH, VT, F, L, C, S) belong to the interaction stage.[3]

Because engagement states are latent variables, we employ an input-output hidden Markov model (IO-HMM) (Yin et al. 2017) to predict user activity. We hereafter refer to the proposed Engagement-based Click-interact IO-HMM as the ECI model for brevity. User activity UA_{ijt} is used for measuring the content usage level of user i for content j at time t, which is defined as the weighted sum of the actions from the conversion, loyalty, and advocacy stages in the conversion funnel since there are no actions in the awareness-consideration stage. We denote w_1, w_2, w_3 as the weights of the actions from the above three engagement stages and assume $w_1 < w_2 < w_3$ because actions from the higher engagement stage represent higher user activity.[4] Alternatively, we can express user activity as the sum of her click score $w_1 CL_{ijt}$ (i.e., multiplication of click behavior and its weight) and her interaction score IS_{ijt} (i.e., the weighted sum of various interaction behaviors) as follows:

$$\begin{aligned} UA_{ijt} &= w_1\left(CL_{ijt} + VC_{ijt} + VPH_{ijt}\right) + w_2\left(VT_{ijt} + F_{ijt}\right) + w_3\left(L_{ijt} + C_{ijt} + S_{ijt}\right) \\ &= w_1 CL_{ijt} + IS_{ijt}, \end{aligned} \tag{1}$$

where $IS_{ijt} = w_1\left(VC_{ijt} + VPH_{ijt}\right) + w_2\left(VT_{ijt} + F_{ijt}\right) + w_3\left(L_{ijt} + C_{ijt} + S_{ijt}\right)$. To be consistent in terms of variable measurement, we define VT_{ijt} as a binary variable, which equals 1 when user i's viewing time for the content is higher than average users' time length of viewing contents, and 0 otherwise.[5] All other variables are also binary variables, which equals 1 if user i takes the corresponding action on content j at time t.

The factors that influence user activity can be defined as a six-dimensional vector $\boldsymbol{X_{ijt}} = \left(PR_{ij}, PD_i, PL_{jt}, NV_{jt}, DE_i, TI_t\right)$, which measures the preference (PR) of user i for content j, the preference diversity (PD) of user i, the popularity (PL) of content j at time t, the novelty (NV) of content j at time t, the demographics (DE) of user i and the time (TI) of content usage, respectively. With the proposed ECI model, we first use these factors to predict the click score and interaction score respectively, and then we aggregate them to get the predicted user activity $\widehat{UA}_{ijt}$.

Content Usage Predictive Uplift Model

Recommending content j to user i at time t is defined as an impression $\boldsymbol{\mathcal{T}_{ijt}}$. The impression $\boldsymbol{\mathcal{T}_{ijt}}$ is regarded as the treatment of the recommender system on user i at time t. It can be decomposed into a number of attributes as follows:

$$\boldsymbol{\mathcal{T}_{ijt}} = \left(PR_{ij}, PL_{jt}, NV_{jt}\right). \tag{2}$$

Because PD_i, DE_i, TI_t are not the treatments of recommender systems on user activity, we don't include them in $\boldsymbol{\mathcal{T}_{ijt}}$. We implement the linear double machine learning model (LDML) (Chernozhukov et al. 2017) to identify the heterogeneous treatment effect (HTE) of impression $\boldsymbol{\mathcal{T}_{ijt}}$ on user activity, that is, the predicted incremental user activity $\widehat{\Delta UA}_{ijt}$:

$$\widehat{\Delta UA}_{ijt} = \boldsymbol{\hat{\theta}_{UA_{ijt}}}^{\boldsymbol{T}} \boldsymbol{\mathcal{T}_{ijt}}, \tag{3}$$

where $\boldsymbol{\hat{\theta}_{UA_{ijt}}}$ is the heterogeneous marginal treatment effect (HMTE) of $\boldsymbol{\mathcal{T}_{ijt}}$.

3 The proposed model is easily adaptable to other content usage behaviors by mapping them to the corresponding stages of the conversion funnel. For example, behaviors like Save/Add to Favorites, Download, Subscribe can be mapped to the loyalty stage.

4 Experiments show that our results are robust against different choices of w_1, w_2, w_3.

5 When a user spends more than the average time to view a specific content, she is considered as having a high engagement level (i.e., loyalty stage). Thus, we use the binary variable to measure the viewing time.

Content Generation Prediction

Creator activity CA_{ct} describes the content generation level of creator c on day t, which is defined as the number of contents published by this creator on the day.[6] Creators can observe the various performance feedback information of contents created by them. The performance feedback information can help them judge their creator popularity, which further influences their motivation to create contents. We define the creator popularity CP_{ct} of creator c on day t as the weighted sum of the number of various performance feedback information that she receives:

$$CP_{ct} = w_1(NCL_{ct} + NVC_{ct} + NVPH_{ct}) + w_2(NVT_{ct} + NF_{ct}) + w_3(NL_{ct} + NC_{ct} + NS_{ct}), \tag{4}$$

where the variables in Equation (4) represent the number of CL, VC, VPH, VT, F, L, C, and S that creator c receives on day t, respectively.[7] For example, $NCL_{ct} = \sum_{i\in I_c}\sum_{j\in J_c} CL_{ijt}$, where I_c is the set of users who consume the contents created by the creator c and J_c is the set of contents created by the creator c. Other variables are defined in a similar manner. Using the LDML model, we identify the HTE of creator popularity on creator activity, that is, incremental creator activity $\widehat{\Delta CA}_{ct}$:

$$\widehat{\Delta CA}_{ct} = \hat{\theta}_{CP_{ct}} CP_{ct}, \tag{5}$$

where $\hat{\theta}_{CP_{ct}}$ represents HMTE of creator popularity.

According to the definition of creator popularity CP_{ct}, performance feedback terms like NCL_{ct}, and user activity UA_{ijt}, we further find CP_{ct} equals the sum of all user activity from consuming contents created by creator c on day t, that is, $CP_{ct} = \sum_{i\in I_c}\sum_{j\in J_c} UA_{ijt}$. Thus, the predicted incremental creator activity of creator c on day t can be expressed as a linear function of user activity:

$$\widehat{\Delta CA}_{ct} = \hat{\theta}_{CP_{ct}} CP_{ct} = \hat{\theta}_{CP_{ct}}\left(\sum_{i\in I_c}\sum_{j\in J_c} UA_{ijt}\right) = \sum_{i\in I_c}\sum_{j\in J_c} \hat{\theta}_{CP_{c_j t}} UA_{ijt}, \tag{6}$$

where c_j is the creator of content j. This implies that recommender systems can influence incremental creator activity indirectly by directly influencing user activity. Hence, platforms can balance incremental user activity and incremental creator activity through content recommendation decisions. Because UA_{ijt} are unknown before making decisions, we first need to predict it with the proposed ECI model. With $\widehat{UA}_{ijt}$, we can get the predicted incremental creator activity of creator c on day t as follows:

$$\widehat{\Delta CA}_{ct} = \sum_{i\in I_c}\sum_{j\in J_c} \hat{\theta}_{CP_{c_j t}} \widehat{UA}_{ijt}. \tag{7}$$

Content Recommendation Optimization

Based on the identified HTEs, we build an integer programming model to make the personalized content recommendation decisions for balancing content generation and usage. d_{ijt} is the binary decision variable that equals 1 if recommending content j to user i at time t, otherwise 0. $\widehat{\Delta UA}_{ijt}(d_{ijt})$ is the incremental user activity incurred by recommendation decision d_{ijt}. $\widehat{\Delta CA}_{ct}\left(CP_{ct}(d_{ijt})\right)$ is the incremental creator activity brought by recommendation decision d_{ijt}. It is a function of creator popularity $CP_{ct}(d_{ijt})$ which is a function of recommendation decision d_{ijt}. β and $1-\beta$ are the weights of the total incremental user activity and the total incremental creator activity, respectively. The objective is to maximize the weighted sum of the total incremental user activity and the total incremental creator activity (total incremental activity for brevity) brought by recommendation decisions:

$$\max_{d_{ijt}\in\{0,1\}} \beta \sum_{i\in I}\sum_{j\in J}\sum_{t\in T} \widehat{\Delta UA}_{ijt}(d_{ijt}) + (1-\beta)\sum_{c\in C}\sum_{t\in T} \widehat{\Delta CA}_{ct}(CP_{ct}(d_{ijt})), \tag{8}$$

where $\widehat{\Delta UA}_{ijt}(d_{ijt}) = \widehat{\Delta UA}_{ijt} d_{ijt}$, $\widehat{UA}_{ijt}$ is the predicted incremental user activity that user i consumes content j at time t in the recommendation stage. According to Equation (7) and the fact that incremental creator activity is from each interaction between users and contents, we can obtain

[6] Creator activity and the variables that influence it are measured at daily level.

[7] We normalize these variables to the range from 0 to 1.

$\sum_{c\in C}\sum_{t\in T}\widehat{\Delta CA}_{ct}\left(CP_{ct}(d_{ijt})\right)=\sum_{c\in C}\sum_{t\in T}\sum_{i\in I_c}\sum_{j\in J_c}\hat{\theta}_{CP_{c_{j}t}}\widehat{UA}_{ijt}d_{ijt}=\sum_{i\in I}\sum_{j\in J}\sum_{t\in T}\hat{\theta}_{CP_{c_{j}t}}\widehat{UA}_{ijt}\,d_{ijt}$.[8] Thus, Equation (8) is equivalent to the following optimization problem:

$$\max_{d_{ijt}\in\{0,1\}}\sum_{i\in I}\sum_{j\in J}\sum_{t\in T}[\beta\widehat{\Delta UA}_{ijt}+(1-\beta)\,\hat{\theta}_{CP_{c_{j}t}}\widehat{UA}_{ijt}]d_{ijt}, \tag{9}$$

where $\beta\widehat{\Delta UA}_{ijt}+(1-\beta)\hat{\theta}_{CP_{c_{j}t}}\widehat{UA}_{ijt}$ represents the weighted incremental activity corresponding to each recommendation decision d_{ijt}. When recommending content j created by creator c_j to user i, user i will generate incremental user activity $\widehat{\Delta UA}_{ijt}$. Besides, her user activity $\widehat{UA}_{ijt}$ influences the creator popularity of creator c_j, which further affects the content generation of creator c_j, that is, incremental creator activity $\hat{\theta}_{CP_{c_{j}t}}\widehat{UA}_{ijt}$. By varying β, the platform can adjust the weighted incremental activity to change the recommended contents for each user, which further balances the content generation and usage on the platform by the individual-level recommendation. When $\beta\widehat{\Delta UA}_{ijt}+(1-\beta)\hat{\theta}_{CP_{c_{j}t}}\widehat{UA}_{ijt}>0$, $d_{ijt}=1$.[9] This means only the content that brings the positive weighted incremental activity is recommended. Finally, we get the recommendation content set S_{it} from available contents for each user i at time t as follows:

$$S_{it}=\{j|d_{ijt}=1, j\in J\}. \tag{10}$$

Experiment

In this section, we validate the proposed method using data from NetEase Cloud Music (henceforth NCM), which is one of the largest music streaming platforms in China, with approximately 800 million users and a valuation of approximately 9 billion dollars in 2019 (Zhang et al. 2022). Specifically, our sample data come from NCM's Cloud Village, which is an active online community where creators create music cards that are then consumed by users. The data set consists of more than 57 million impressions of music cards created by 90 thousand creators and recommended to 2 million users from November 1st, 2019, to November 30th, 2019. Information related to each impression includes the corresponding user, music card, time, and user's usage behaviors (e.g., clicks and shares). Each music card has its corresponding creator, type, category, topic, artist, publish time, and various performance feedback information (e.g., the number of clicks and shares). In addition, the data set includes information on each user (e.g., gender and age) and each creator (e.g., gender, genre, and the number of music cards created by her for a given date). A detailed description of the data can be found in Zhang et al. (2022).

Performance of Prediction Methods

We use 70% of data for model training, 10% of data as a validation set to choose the best parameters, and the remaining 20% as a test set for evaluating model performance. We compared the proposed ECI model with the hidden Markov model (HMM), IO-HMM, and other common machine learning models including linear regression, support vector regression and tree regression model (e.g., random forest, AdaBoost, and XGBoost).[10] Table 1 illustrates the out-of-sample prediction performance of the proposed ECI model and the benchmark models.

We find the proposed ECI performs best because it has the smallest mean square error (MSE) and mean absolute error (MAE). There are several reasons for its superior performance. First, it considers the effects of various covariates on transition probability and emission outcome, which is ignored by the standard

[8] We normalize $\widehat{\Delta UA}_{ijt}$ and $\hat{\theta}_{CP_{c_{j}t}}\widehat{UA}_{ijt}$ to prevent the case where one factor dominates the other. Specifically, each measure is normalized by the maximum of the absolute value respectively. After rescaling, the respective values of each measure belong to [-1,1].

[9] In practice, platforms can recommend top H contents with the largest $\beta\widehat{\Delta UA}_{ijt}+(1-\beta)\hat{\theta}_{CP_{c_{j}t}}\widehat{UA}_{ijt}$ to each user.

[10] Due to the page limit, we are not able to introduce benchmark models in detail. All details and specifications are available upon request.

HMM. Second, it incorporates user engagement and click-interact sequential decision process during content usage, which is not considered by IO-HMM and standard machine learning methods (e.g., SVR).

Model	MSE	MAE
ECI	0.0096	0.0231
HMM	0.0432	0.0378
IO-HMM	0.0668	0.0256
Linear Regression	0.0143	0.0356
Ridge Regression	0.0143	0.0356
Lasso Regression	0.0166	0.0373
Support Vector Regression (SVR)	0.0232	0.1119
Random Forest	0.0146	0.0277
AdaBoost	0.0461	0.0839
XGBoost	0.0137	0.0336

Table 1. Comparison of Predictive Model Performance

Performance of Recommendation Methods

To validate the performance of the proposed prescriptive uplift model, we compare it with the prescriptive response model. The prescriptive response model optimizes decisions based on the estimated user response (The details of this model are given in the Appendix). In contrast, our proposed model effectively identifies the HTE of recommendation on users and creators. As a result, it only recommends contents that bring positive incremental activity to users, which ensures that all recommended contents can increase activity. This leads to performance improvement of our proposed model over the prescriptive response model in terms of the total incremental activity. We define *performance improvement* as the absolute difference between the total incremental activity under our model and the benchmark model. Table 2 presents the improvement of our model under different β.

Weight of Content Usage (β)	0	0.2	0.4	0.6	0.8	1
Improvement	0.47	120.79	198.73	223.61	213.05	192.03

Table 2. Comparison of Prescriptive Model Performance

We observe that the advantage of our method is more obvious when the goals of platforms are more balanced (e.g., β is 0.4 or 0.6). This is because both its advantages in increasing incremental user activity and in increasing incremental creator activity are fully reflected. In contrast, when platforms are creator-oriented (β=0), the advantage brought by 'considering the HTE of recommendation on user activity' will vanish, which results in similar performance as the prescriptive response model. When platforms are user-oriented (β=1), the prescriptive response model performs much worse than our proposed prescriptive uplift model because its drawback of ignoring the HTEs becomes more obvious.

We can further calculate the *improvement ratio* as the absolute performance improvement divided by the total incremental activity of the benchmark model. The *average improvement ratio* is the average value of the improvement ratio under different β. Our model achieves 42% performance improvement over the existing benchmark on average, demonstrating superior performance in increasing user and creator activity on the platform by effectively balancing content generation and usage.

Note that the prescriptive response model can't balance content generation and usage because it assumes that the effect of user activity on creator activity is homogeneous rather than heterogeneous. The higher the user activity is, the higher the creator activity is. As a result, platforms only need to maximize user activity. By contrast, our model allows platforms to balance content generation and usage by choosing a preferred content usage weight. Specifically, creator-oriented platforms (β=0) only emphasize content generation, where incremental creator activity is highest but incremental user activity is lowest. As platforms emphasize content usage more (β increases), incremental user activity increases but incremental creator activity decreases. When platforms become fully user-oriented (β=1), the incremental user activity reaches the maximum level while incremental creator activity drops to the minimum. Different platforms or different development stages of the same platform can flexibly determine their preferred decision weight to balance content usage and content generation.

Implications on Personalized Recommendation

From the perspective of understanding heterogeneous user and creator behaviors, our model provides many useful managerial insights to platforms.

Dual Role of Content Creators and Users. Platform participants include those who serve one role (pure content creator or pure content user) and those who serve the dual role of a content creator and user. The number of served roles can influence content generation and usage, which is one important demographic feature in our previous prediction models. Compared with undertaking one role, our study shows that serving the dual role hinders content usage while contributing to content generation. Furthermore, the mechanisms are different for inactive users (via "stickiness") and active users (via "resource allocation").

To see the difference, we first classify users/creators into active and inactive user/creator groups. We classify users/creators whose activity intensity levels are above average as active users/creators, and inactive users/creators otherwise. The activity intensity of participants is calculated by their time and frequency of interacting with the NCM app and their values are given in the raw data. *Stickiness* is measured by the total number of content usage and content generation, which describes the strength of the connection between participants and platforms. *AllocatedResource is* the resource allocated to content usage, which is measured by the percentage of content usage in the total number of content usage and content generation. *Dual* is a binary variable that equals 1 if the participant serves the dual role, otherwise 0. *Controls* include a number of control variables like user features, content features, time features, and creator features.

We next estimate the effect of dual role on user activity and creator activity, respectively. Because user activity is a continuous variable, we use the linear regression model as follows:

$$UA_{ijt} = \alpha Stickness_i + \beta AllocatedResource_i + \gamma Dual_i \times Stickness_i$$
$$+\eta Dual_i \times AllocatedResource_i + \theta Controls_{ijt} + \varepsilon_{ijt} \quad (11)$$

Because creator activity is a count variable, we use the zero-inflated negative binomial regression model to estimate the effects of various factors on creator activity. We add $CreatorPopularity_{ct}$ to the model, which denotes the creator popularity and it has the same meaning as CP_{ct} in the previous section. The estimation model is as follows:

$$CA_{ct} = f(Stickness_c, AllocatedResource_c, Dual_c \times Stickness_c,$$
$$Dual_c \times AllocatedResource_c, CreatorPopularity_{ct}, Controls_{ct}) + \varepsilon_{ct} \quad (12)$$

The estimated results of the above models are given in Table 3.

Factors	User Activity		Creator Activity	
	Inactive Users	Active Users	Inactive Creators	Active Creators
$Stickiness$	0.0001*	0.0000***	0.0007***	0.0002***
$Dual \times Stickiness$	-0.0022***	0.0000	0.0051***	0.0043***
$AllocatedResource$	-0.4585***	-0.0490***	-6.9084***	-6.8517***

$Dual \times AllocatedResource$	0.0004	-0.0091***	2.1245***	1.5827***
$CreatorPopularity$			0.3069***	0.1968***
$Controls$	Yes	Yes	Yes	Yes
Note. *** p<0.01, ** p<0.05, * p<0.1.				

Table 3. Factors that Influence User Activity and Creator Activity

For inactive users, the "stickiness" between them and platforms is one of the most important factors that influence their content usage. Distraction brought by serving the dual role can hinder content usage by reducing the effect of their "stickiness" on content usage. For active users, how to allocate limited resources (e.g., time and energy) for content generation and usage is vital. Distraction caused by serving two roles can decrease content usage by reducing the effect of "the allocated resources" on content usage. In contrast, serving the dual role can contribute to the content generation of all creators by strengthening the effect of "stickiness" and "the allocated resources" on content generation. This could be because creators learn useful knowledge and skills about generating contents from content usage.

Heterogeneous Transition of User Engagement. According to the proposed ECI model, we can infer the transition probabilities of different engagement stages during each content usage. Table 4 gives the engagement transition probabilities of active users and inactive users.

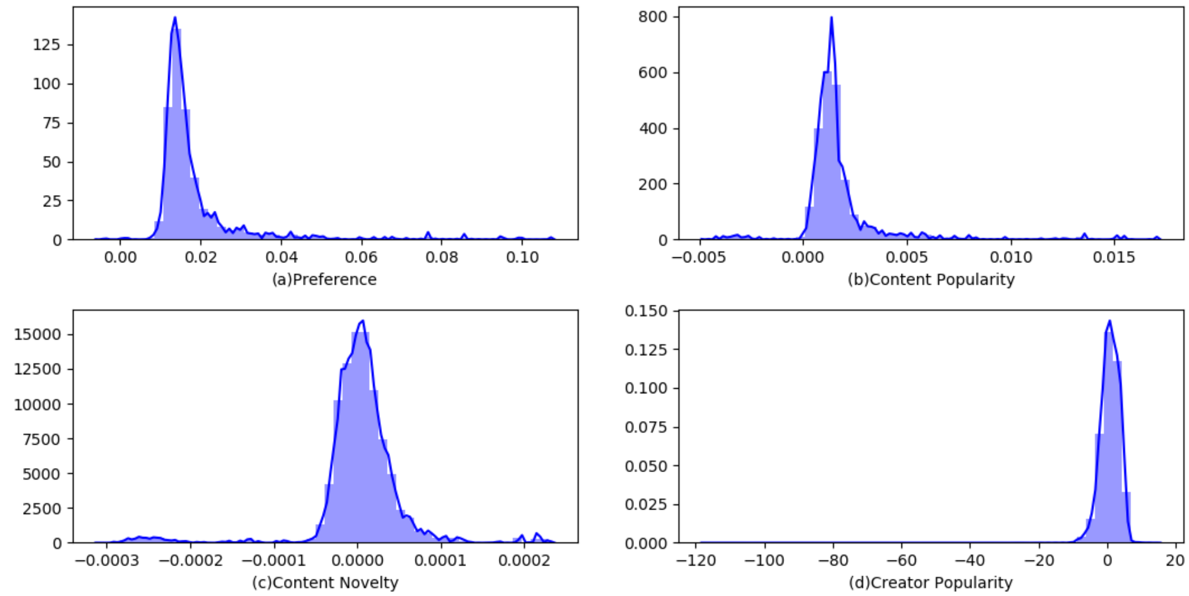
		e' =Awareness-consideration	e' = Conversion	e' = Loyalty	e' = Advocacy
Inactive Users	e = Awareness-consideration	0.979	0.001	0.018	0.000
	e = Conversion	0.408	0.205	0.194	0.191
	e = Loyalty	0.369	0.122	0.387	0.120
	e = Advocacy	0.291	0.218	0.218	0.271
Active Users	e = Awareness-consideration	0.980	0.002	0.016	0.000
	e = Conversion	0.464	0.198	0.170	0.166
	e = Loyalty	0.404	0.088	0.421	0.086
	e = Advocacy	0.297	0.192	0.198	0.311

Notes. e is the engagement state in the current stage, e' represents the engagement state in the next stage, each value in the table is the mean value of the probability that the engagement state transforms from e to e' for different groups of users.

Table 4. Transition Probability of User Engagement for Different Users

We find that improving user engagement can prevent users from transiting into a lower engagement stage in the future, especially for active users. Compared with inactive users, active users are more likely to transit into the lower-level stage (i.e., awareness-consideration stage) when they are at the conversion stage. Once they enter the loyalty stage, they prefer to stay in the stage rather than progress into the higher-level stage (i.e., advocacy stage). After that, if they enter the advocacy stage, they are more likely to maintain the highest engagement level and are less likely to transit into other lower-level stages. Thus, platforms can recommend the most tailored contents to active users to improve their current engagement, which helps maintain their high level of engagement in the future.

Heterogeneous Effect of Recommendation and User Types. According to the LDML model, we find that the treatment effect of recommendation on user activity is user-content heterogeneous in the terms of magnitude and direction (See Figure 2(a), 2(b), 2(c)).



Note. In each sub-figure, x-axis represents the HMTE of the corresponding factor, y-axis represents the density.

Figure 2. Heterogeneous Marginal Treatment Effect (HMTE)

According to the predicted HMTEs, we identify different types of users (e.g., popularity-driven users, preference-driven users, novelty-driven users). This will provide important management insights for two-sided UGC entertainment platforms to design personalized content recommendation strategies. For example, for popularity-driven users (i.e., $\hat{\theta}_{PL,i}$ is large), the proposed method will recommend popular contents to them, which is the best way to improve their user activity. [11] For preference-driven users (i.e., $\hat{\theta}_{PR,i}$ is large), the proposed model will recommend contents that are most tailored to their taste to them. Besides, it will be the most effective to recommend the latest contents to novelty-driven users (i.e., $\hat{\theta}_{NV,i}$ is large).

Heterogeneous Effect of Creator Popularity and Creator Types. Creator popularity can increase creators' content generation significantly, where the effect is stronger for inactive creators than active creators (See Table 3). The potential explanation is that the marginal utility of creator popularity is diminishing. Specifically, the average creator popularity of active creators and inactive creators are 0.0134 and 0.0107 respectively. The active creators are on average more popular, and thus, the effect of popularity on them will not be as strong as on inactive creators. Using the LDML model, we find that the treatment effect of creator popularity on creator activity is creator-heterogeneous in the terms of magnitude and direction (See Figure 2(d)). This finding is consistent with previous studies (Srivastava and Rangarajan 2008, Wozniak 2012, Jabr et al. 2014), that is, the effect of performance feedback on creators' content generation depends on creators' preferences for peer recognition and social exposure. According to the predicted HMTE, we identify different types of creators (e.g., popularity-driven creators).

[11] $\hat{\theta}_{PL,i}$ is the mean value of estimated heterogeneous marginal treatment effect of $\hat{\theta}_{PL,ijt}$ related to user i. $\hat{\theta}_{PR,i}$ and $\hat{\theta}_{NV,i}$ are calculated with the similar method.

For popularity-driven creators (i.e., $\hat{\theta}_{CP_c}$ is large), the proposed recommendation method incentivizes them to create more contents by improving their popularity. [12]

Conclusion

This research develops a prescriptive uplift framework to balance content generation and usage on two-sided UGC entertainment platforms through personalized recommendation decisions. Specifically, the proposed model first predicts users' content usage level. Then, it identifies HTEs of recommendation decisions on creators' content generation and users' content usage. Finally, the model balances both aspects by recommending contents based on identified HTEs. We make methodological contributions from the following perspectives. First, the proposed prescriptive uplift framework integrates causal inference into the optimization model, which brings significant performance improvement over existing correlation-based benchmark recommendation methods. Second, we contribute to the current literature in the areas of content management and recommender systems. To the best of our knowledge, this is the first study to balance the conflict between content generation and usage through the platform's recommendation decisions. Third, our proposed model can generalize to other two-sided platforms that need to balance content generation and usage such as online question & answer platforms (e.g., Quora, Zhihu), social media platforms (e.g., Twitter, Weibo), and news platforms (e.g., Dailyhunt, Toutiao).

We validate the proposed model performance using data from NetEase Cloud Music. Experimental results demonstrate that our method not only balances content generation and usage effectively but also increases the total incremental activity by 42% compared to benchmark models. The significant increase in participants' activity suggests that our method effectively improves both users' consumption of existing contents and creators' generation of new contents. In addition, we identify several managerial insights to understand heterogeneous user and creator behaviors on the platform. First, serving the dual role of content creators and content users can hinder content usage but contribute to content generation. Second, users' feedback affects creators' popularity, which significantly influences their content generation incentives. The effect is greater for inactive creators compared with active creators. Third, compared with inactive users, active users are more likely to become highly engaged in the future if their current engagement is high, and they are more likely to become less engaged if their current engagement is low.

We also provide practical implications for platforms to improve the current recommendation strategies. First, our model enables platforms to set the appropriate weight to balance content generation and usage according to their specific goals and development stages. Second, it helps platforms design personalized recommendation policies for different types of users (e.g., popularity-driven, preference-driven, and novelty-driven users) and creators (e.g., popularity-driven creators). As new data about user and creator behaviors are collected over time, we can easily and adaptively recalibrate the parameters of the proposed model. Overall, the proposed method increases user activity and creator activity, which are important for the platforms' revenue growth (e.g., ad revenue, subscription fees) and sustainable development.

Our work has several limitations, which can be further explored in future research. First, for the sake of interpretability, we assume the treatment effects of recommendation on the user activity and creator activity are linear. Future research can explore the nonlinear heterogeneous treatment effects. Second, because users consume contents and creators create contents in multiple periods, sequential content recommendation model will be an interesting topic of future research. Third, this paper focuses on using explicit feedback data to recommend contents. Future studies may collect multi-modal data (e.g., videos, figures, texts) of contents and use deep neural networks to extract multi-modal features for recommending contents.

References

Aaltonen, A., and Seiler, S. 2016. "Cumulative Growth in User-Generated Content Production: Evidence from Wikipedia," *Management Science* (62:7), pp. 2054-2069.

[12] $\hat{\theta}_{CP_c}$ is the mean value of heterogeneous marginal treatment effect of creator popularity $\hat{\theta}_{CP_{ct}}$ related to creator c.

Abbasi, A., Li, J., Adjeroh, D., Abate, M., and Zheng, W. 2019. "Don't Mention It? Analyzing User-Generated Content Signals for Early Adverse Event Warnings," *Information Systems Research* (30:3), pp. 1007-1028.

Adomavicius, G., and Tuzhilin, A. 2005. "Toward the Next Generation of Recommender Systems: a Survey of the State-of-the-art and Possible Extensions," *IEEE Transactions on Knowledge and Data Engineering* (17:6), pp.734-749.

Alley, M., Biggs, M., Hariss, R., Herrmann, C., Li, M. L., and Perakis, G. 2022. "Pricing for Heterogeneous Products: Analytics for Ticket Reselling," *Manufacturing & Service Operations Management*, forthcoming.

Archak, N., Ghose, A., and Ipeirotis, P. G. 2011. "Deriving the Pricing Power of Product Features by Mining Consumer Reviews," *Management Science* (57:8), pp. 1485-1509.

Bengio, Y., and Frasconi, P. 1994. "An Input Output HMM Architecture," in *Proceedings of the 7th International Conference on Neural Information Processing Systems*, G. Tesauro, D.S. Touretzky, and T. K. Leen (eds.), Denver, CO, pp. 427–434.

Besbes, O., Gur, Y., and Zeevi, A. 2016. "Optimization in Online Content Recommendation Services: Beyond Click-through Rates," *Manufacturing & Service Operations Management* (18:1), pp. 15-33.

Burtch, G., He, Q., Hong, Y., and Lee, D. 2021. "How Do Peer Awards Motivate Creative Content? Experimental Evidence from Reddit," *Management Science*, forthcoming.

Chen, Y., Wang, Q., and Xie, J. 2011. "Online Social Interactions: A Natural Experiment on Word of Mouth Versus Observational Learning," *Journal of Marketing Research* (48:2), pp. 238-254.

Cheng, X., Zhang, J., and Yan, L. 2020. "Understanding the Impact of Individual Users' Rating Characteristics on the Predictive Accuracy of Recommender Systems," *INFORMS Journal on Computing* (32:2), pp. 303-320.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., and Newey, W. 2017. "Double/Debiased/Neyman Machine Learning of Treatment Effects," *American Economic Review* (107:5), pp. 261-265.

Cohen, A. 2009. "The New Marketing Funnel," August 13 (http://adamhcohen.com/the-new-marketing-funnel/).

Deng, Y., Zheng, J., Khern-am-nuai, W., and Kannan, K. 2021. "More Than the Quantity: The Value of Editorial Reviews for a User-Generated Content Platform," *Management Science*, forthcoming.

Devriendt, F., Moldovan, D., and Verbeke, W. 2018. "A Literature Survey and Experimental Evaluation of the State-of-the-Art in Uplift Modeling: A Stepping Stone toward the Development of Prescriptive Analytics," *Big Data* (6:1), pp. 13-41.

Dietrich, G., and Livingston, G. 2012. *Marketing in the Round: How to Develop an Integrated Marketing Campaign in the Digital Era,* Hoboken, NJ: Que Publishing.

Foutz, N. Z. 2017. "Entertainment Marketing," *Foundations and Trends® in Marketing* (10:4), pp. 215-333.

Ghose, A., and Han, S. P. 2011. "An Empirical Analysis of User Content Generation and Usage Behavior on the Mobile Internet," *Management Science* (57:9), pp. 1671-1691.

Ghose, A., Ipeirotis, P. G., and Li, B. 2019(a). "Modeling Consumer Footprints on Search Engines: An Interplay with Social Media," *Management Science* (65:3), pp. 1363-1385.

Ghose, A., Li, B., and Liu, S. 2019(b). "Mobile Targeting Using Customer Trajectory Patterns," *Management Science* (65:11), pp. 5027-5049.

Gutierrez, P., and Gérardy, J. Y. 2017. "Causal Inference and Uplift Modelling: A Review of the Literature," in *Proceedings of the 3rd International Conference on Predictive Applications and APIs*, C. Hardgrove, L. Dorard, K. Thompson, and F. Douetteau (eds.), Boston, MA, pp. 1-13.

He, J., Fang, X., Liu, H., and Li, X. 2019. "Mobile App Recommendation: An Involvement-Enhanced Approach," *MIS Quarterly* (43:3), pp. 827-849.

Huang, N., Burtch, G., Gu, B., Hong, Y., Liang, C., Wang, K., Fu, D., and Yang, B. 2019. "Motivating User-Generated Content with Performance Feedback: Evidence from Randomized Field Experiments," *Management Science* (65:1), pp. 327-345.

Jabr, W., Mookerjee, R., Tan, Y., and Mookerjee, V. S. 2014. "Leveraging Philanthropic Behavior for Customer Support: The Case of User Support Forums," *MIS Quarterly* (38:1), pp. 187-208.

Jannach, D., and Adomavicius, G. 2016. "Recommendations with a Purpose," in *Proceedings of the 10th ACM Conference on Recommender Systems*, S. Sen, W. Geyer, J. Freyne, and P. Castells (eds.), Boston, MA, pp. 7-10.

Kane, K., Lo, V. S. Y., and Zheng, J. 2014. “Mining for the Truly Responsive Customers and Prospects Using True-Lift Modeling: Comparison of New and Existing Methods,” *Journal of Marketing Analytics* (2:4), pp. 218-238.

Lian, J., Zhang, F., Xie, X., and Sun, G. 2018. “Towards Better Representation Learning for Personalized News Recommendation: a Multi-Channel Deep Fusion Approach,” in *Proceedings of the 27th International Joint Conference on Artificial Intelligence*, J. Lang (eds.), Stockholm, Sweden, pp. 2576-2584.

Liu, J., and Toubia, O. 2018. “A Semantic Approach for Estimating Consumer Content Preferences from Online Search Queries,” *Marketing Science* (37:6), pp. 930-952.

Liu, Y., and Feng, J. 2021. “Does Money Talk? The Impact of Monetary Incentives on User-Generated Content Contributions,” *Information Systems Research* (32:2), pp. 394-409.

Lo, V. S. Y., and Pachamanova, D. A. 2015. “From Predictive Uplift Modeling to Prescriptive Uplift Analytics: A Practical Approach to Treatment Optimization While Accounting for Estimation Risk,” *Journal of Marketing Analytics* (3:2), pp. 79-95.

MarketsandMarkets. 2021. “Video Streaming Software Market with COVID-19 Impact, by Component (So lutions (Video Distribution and Video Analytics) and Services), Streaming Type (Live Streaming and Video-on-Demand Streaming), Vertical, and Region - Global Forecast to 2026,” MarketsandMarkets, November (https://www.marketsandmarkets.com/Market-Reports/video-streaming-market-1811351 20.html).

Mazurova, E. 2017. “Exploratory Analysis of the Factors Affecting Consumer Choice in E-Commerce: Conjoint Analysis,” *Journal of Information Systems Engineering & Management* (2:2), pp. 12.

McFowland Iii, E., Gangarapu, S., Bapna, R., and Sun, T. 2021. “A Prescriptive Analytics Framework for Optimal Policy Deployment Using Heterogeneous Treatment Effects,” *MIS Quarterly* (45:4), pp. 1807-1832.

Moe, W. W., and Fader, P. S. 2002. “Fast-Track: Article Using Advance Purchase Orders to Forecast New Product Sales,” *Marketing Science* (21:3), pp. 347-364.

Moon, J. Y., and Sproull, L. S. 2008. “The Role of Feedback in Managing the Internet-Based Volunteer Work Force,” *Information Systems Research* (19:4), pp. 494-515.

Peake, G., and Wang, J. 2018. “Explanation Mining: Post hoc interpretability of latent factor models for recommendation systems,” in *the Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, Y. Guo, and F. Farooq (eds.), London, pp. 2060-2069.

Ricci, F., Rokach, L., and Shapira, B. 2011. *Recommender systems handbook*, New York, NY: Springer.

Shim, J. P., Park, S., and Shim, J. M. 2008. “Mobile Tv Phone: Current Usage, Issues, and Strategic Implications,” *Industrial Management & Data Systems* (108:9), pp. 1269-1282.

Song, Y., Sahoo, N., and Ofek, E. 2019. “When and How to Diversify—a Multicategory Utility Model for Personalized Content Recommendation,” *Management Science* (65:8), pp. 3737-3757.

Srivastava, R., and Rangarajan, D. 2008. “Understanding the Salespeople's “Feedback ‐ Satisfaction” Linkage: What Role Does Job Perceptions Play,” *Journal of Business & Industrial Marketing* (23:3), pp. 151-160.

Statista. 2021. “Digital Media Report 2021 - Digital Music,” Statista, June (https://www.statista.com/stud y/39314/digital-music/).

Van Mierlo, T. 2014. “The 1% Rule in Four Digital Health Social Networks: An Observational Study,” *Journal of medical Internet research* (16:2), e33.

Wozniak, D. 2012. “Gender Differences in a Market with Relative Performance Feedback: Professional Tennis Players,” *Journal of Economic Behavior & Organization* (83:1), pp. 158-171.

Wu, C., Wu, F., An, M., Huang, J., Huang, Y., and Xie, X. 2019. “Npa: Neural News Recommendation with Personalized Attention,” in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, A. Teredesai, V. Kumar, Y. Li, R. Rosales, E. Terzi, and G. Karypis (eds.), Anchorage, AK, pp. 2576-2584.

Yang, M., Ren, Y., and Adomavicius, G. 2019. “Understanding User-Generated Content and Customer Engagement on Facebook Business Pages,” *Information Systems Research* (30:3), pp. 839-855.

Yin, M., Sheehan, M., Feygin, S., Paiement, J.-F., and Pozdnoukhov, A. 2018. “A Generative Model of Urban Activities from Cellular Data,” *IEEE Transactions on Intelligent Transportation Systems* (19:6), pp. 1682-1696.

Zhang, D. J., Hu, M., Liu, X., Wu, Y., and Li, Y. 2022. “Netease Cloud Music Data,” *Manufacturing & Service Operations Management* (24:1), pp. 275-284.

Zhang, X., and Zhu, F. 2011. "Group Size and Incentives to Contribute: A Natural Experiment at Chinese Wikipedia," *American Economic Review* (101:4), pp. 1601-1615.

Zhang, Y., Li, B., Luo, X., and Wang, X. 2019. "Personalized Mobile Targeting with User Engagement Stages: Combining a Structural Hidden Markov Model and Field Experiment," *Information Systems Research* (30:3), pp. 787-804.

Appendix

We present the details of the benchmark model implementation below. Prescriptive response model is to maximize the total activity (i.e., the weighted sum of total user activity and total creator activity) related to the content recommendation decisions rather than the total incremental activity brought by these decisions. d_{ijt} is the binary decision variable that equals 1 if recommending content j to user i at time t, otherwise 0. $\widehat{UA}_{ijt}(d_{ijt})$ is the user activity related to recommendation decision d_{ijt}. $\widehat{CA}_{ct}(CP_{ct}(d_{ijt}))$ is the creator activity related to recommendation decision d_{ijt}. It is the function of creator popularity $CP_{ct}(d_{ijt})$ which is a function of recommendation decision d_{ijt}. β and $1-\beta$ are the weights of the total user activity and the total creator activity, respectively. The objective is to maximize the total activity related to recommendation decisions.

$$\max_{d_{ijt}\in\{0,1\}} \beta \sum_{i\in I}\sum_{j\in J}\sum_{t\in T}\widehat{UA}_{ijt}\,(d_{ijt}) + (1-\beta)\sum_{c\in C}\sum_{t\in T}\widehat{CA}_{ct}(CP_{ct}(d_{ijt})) \tag{A1}$$

where $\widehat{UA}_{ijt}(d_{ijt}) = \widehat{UA}_{ijt}d_{ijt}$, $\widehat{UA}_{ijt}$ is the predicted user activity that user i consumes content j at time t. Similar to Equation (7) and based on the fact that creator activity is from each interaction between users and contents, we can get $\sum_{c\in C}\sum_{t\in T}\widehat{CA}_{ct}\left(CP_{ct}(d_{ijt})\right) = \sum_{c\in C}\sum_{t\in T}\widehat{CA}_{ct}\left(\sum_{i\in I_c}\sum_{j\in J_c}\widehat{UA}_{ijt}\,d_{ijt}\right) = \sum_{c\in C}\sum_{t\in T}\sum_{i\in I_c}\sum_{j\in J_c}\hat{\gamma}_{CP}\widehat{UA}_{ijt}\,d_{ijt} = \sum_{i\in I}\sum_{j\in J}\sum_{t\in T}\hat{\gamma}_{CP}\widehat{UA}_{ijt}\,d_{ijt}$, where $\hat{\gamma}_{CP}$ is the estimated parameter of creator popularity in the zero-inflated binomial regression model. It can also be estimated by other models. Our analysis shows that it does not affect the optimal recommendation decision.

Thus, the problem (A1) can be converted into the following formula equivalently.

$$\max_{d_{ijt}\in\{0,1\}} \sum_{i\in I}\sum_{j\in J}\sum_{t\in T}[\beta + (1-\beta)\,\hat{\gamma}_{CP}]\widehat{UA}_{ijt}d_{ijt} \tag{A2}$$

Because β and $\hat{\gamma}_{CP}$ are constants, we further simplify Equation (A2) to get the following optimization problem. The weight β doesn't influence the optimal recommendation decision, which means it can't balance content generation and usage.

$$\max_{d_{ijt}\in\{0,1\}} \sum_{i\in I}\sum_{j\in J}\sum_{t\in T}\widehat{UA}_{ijt}d_{ijt} \tag{A3}$$

When $\widehat{UA}_{ijt} > 0$, $d_{ijt} = 1$. This means the content that brings positive user activity will be recommended.[13] Thus, we can get the recommendation content set S_{it} from available contents for each user i at time t.

$$S_{it} = \{j|d_{ijt} = 1\} \tag{A4}$$

[13] In practice, platforms can recommend top H contents with the largest $\widehat{UA}_{ijt}$ to each user.