How Do You Earn Money on Live Streaming Platforms? — A Study of Donation-Based Markets

Ming Tang, *Member, IEEE,* and Jianwei Huang, *Fellow, IEEE*

***Abstract*—Donation-based markets have been implemented by many online platforms, such as live streaming platforms. In these markets, producers provide services without mandatory charges, and customers enjoy the services and voluntarily donate money to the producers. The donation is split between the producers and platform with a pre-agreed fraction. To gain insights into the market operation, we use a two-stage game to capture the sequential decision process between the platform and producers. In Stage I, the platform decides a donation-split-fraction (DSF), i.e., the fraction of donation kept by the producers. In Stage II, producers decide whether to participate in the platform and (if yes) how to choose their service attributes considering the DSF as well as the producers' and customers' preferences. We prove that the Stage II game is a potential game with a counter-intuitive equilibrium result: although a larger DSF leads to more producer participation and a better match between the producers' choices and the customers' preferences, it does not necessarily lead to more total donation. The Stage I problem, nevertheless, is challenging to solve analytically due to its non-convexity. To gain insights regarding the optimal DSF that maximizes the platform's payoff, we characterize both its upper-bound and lower-bound. We show numerically that the platform's optimal payoff always decreases with the mismatch between the producers' and customers' preferences. Finally, we conduct a case study with the dataset from Twitch and demonstrate the approach of computing the platform's optimal DSF without the producers' inherent preferences.**

***Index Terms*—Live streaming platform, donation-based market, potential game, non-atomic game.**

I. Introduction

A. Background and Motivation

Recently, many online platforms have chosen to implement a donation-based two-sided market between two groups of users: producers who provide services without mandatory charges, and customers who enjoy the services and voluntarily donate to the producers. The customers donate mainly due to their desires of being acknowledged on the platforms (e.g., to gain community presence) and supporting the producers for future high-quality service provisions [2]. The donation is split between the producers and the platform with a fixed *donation-split-fraction* (DSF), which corresponds to the fraction of donation kept by the producers.

Many live streaming platforms are good examples of the donation-based markets. One of them is Twitch, which is the largest live streaming platform in the US [3]. It has 1.44 million concurrent viewers and 56,000 concurrent streamers on average, as of March 2020 [4]. On live streaming platforms, individuals can become streamers to live stream their videos of game play for viewers.[1] The viewers watch live streaming videos for free. These viewers can donate to the streamers in order to receive personal recognition and obtain the permission of doing extra actions (e.g., using diverse chat icons in the stream chat window) [2].[2] A fixed fraction of the donation will be kept by the streamers (e.g., at this moment, $1/1.4 \approx 0.71$ on Twitch), while the rest will be kept by the platform. The implementation of the donation-based markets helps the platforms gamify the interaction between streamers and viewers [2] and hence attract more streamers and viewers to participate in the platforms. Meanwhile, the direct monetary transfer from viewers to streamers motivates the streamers (who may not be professional video producers) to provide high quality live streaming services. The total donation volume on live streaming platforms is huge. In 2018, a total of \$141 million dollars of donation were received by top live streaming platforms including Twitch, YouTube Live, Mixer, and Facebook Live [3]. Other donation-based market examples are blogging platforms (e.g., WeChat Subscription) and online music platforms (e.g., Songtradr).[3]

The donation-based markets lead to distinctive monetary ecosystems for online platforms. For example, for traditional video streaming services (e.g., video on demand services on YouTube), the majority income of a streamer comes from advertisement or sponsorship [5]. In comparison, on live streaming platforms (e.g., Twitch), the majority income of a streamer comes from viewers directly [5]. Hence, on live streaming platforms, streamers have more incentive to provide services that match the viewers' preferences. On the other hand, as the platform keeps part of the donation from viewers

Ming Tang is with the Department of Electrical and Computer Engineering, the University of British Columbia, Vancouver, BC, Canada (e-mail: mingt@ece.ubc.ca).

Jianwei Huang is with the School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, Shenzhen 518172, China, and the Shenzhen Institute of Artificial Intelligence and Robotics for Society, Shenzhen 518129, China (corresponding author, e-mail: jianweihuang@cuhk.edu.cn).

This work is supported by the Shenzhen Institute of Artificial Intelligence and Robotics for Society, and the Presidential Fund from the Chinese University of Hong Kong, Shenzhen.

Part of this work has been published at IEEE INFOCOM [1].

Manuscript received Oct. 15, 2019; revised Oct. 2, 2020.

[1]Since streamers and viewers are the ones who provide and enjoy services on live streaming platforms, respectively, they correspond to the producers and customers in donation-based markets, respectively.

[2]There are other types of monetization methods on live streaming platforms. For example, on Twitch, streamers can earn money through not only donation but also viewers' subscription, advertisement, and sponsorship [5]. We will not discuss those monetization methods other than donation in this paper.

[3]WeChat Subscription (https://mp.weixin.qq.com/?lang=en_US) is a blogging platform for individual article publishing. Songtradr (https://www.songtradr.com/) is a music platform for independent musicians.

to streamers, it has the incentive to optimize the DSF in order to motivate streamers to provide the services matching the viewers' preferences and to optimize its payoff.

The donation-based feature of these markets brings two unique questions as follows:

First, from the producers' point of view, *how should they decide their service attributes (e.g., on a live streaming platform, what game to stream at what time of day) given a fixed DSF?* The producers and customers may have different preferences over the service attributes. The producers' choices (which can be different from their preferences) will affect the competition among producers and the satisfaction of the customers.

As an example, Figure 1 illustrates the mismatch of the number of concurrent streamers and viewers on Twitch.[4] Figure 1 (a) shows the average number of concurrent streamers and viewers (over a period of two weeks) of different games. Some games with a large number of streamers have a small number of viewers (e.g., {3} Fortnite). This implies that those streamers may improve their payoffs by switching to stream the other games with less competitors. Figure 1 (b) shows the average number of concurrent streamers and viewers of game *League of Legends* at different time of day. Some streamers may increase their payoffs by changing the time of day of their stream (e.g., from 3am to 11am) to gain more attention and potentially more donation from viewers.

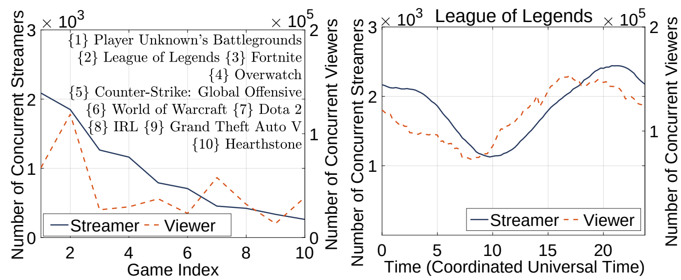
Second, from the platform's point of view, *how should it set the DSF to maximize its payoff?* A higher DSF reduces per-donation revenue to the platform. On the other hand, it can increase the incentive for the producers to participate in the platform and better match the customers' preferences, which may induce more donation.

Despite the fact that donation-based markets have been embraced by top online platforms (e.g., Twitch, YouTube Live, and Facebook Live) and attract millions of producers and customers, there does not exist a good understanding regarding the above two key questions theoretically. In this work, we aim to understand the two key questions and to investigate the interaction between the platform and producers.[5]

B. Solution Approach and Contribution

For the sake of concreteness, we focus on the live streaming example in this paper. The modeling approach and analysis techniques are applicable to other donation-based markets as well. We will use "streamer" and "producer" interchangeably and use "viewer" and "customer" interchangeably.

We use a two-stage game to capture the interaction between the platform and producers. In Stage I, the platform announces the DSF. In Stage II, each producer decides whether to participate and what service attribute to choose (for example, at what time of day to stream). This two-stage game captures



(a) (b)

Fig. 1. Mismatch of the number of concurrent streamers and viewers in Twitch: (a) game attribute; (b) time of day attribute.

the major feature of the donation-based markets and the main interactions between the platform and producers. In practical systems, there may also exist other behaviors of the platform and producers. For example, on a live streaming platform, there may exist social interactions between streamers (e.g., community formation) and other donation methods (e.g., donation through a third party transaction without the involvement of the platform). Detailed studies of those behaviors will be part of our future work. Moreover, we assume that both the platform and producers are rational, i.e., they aim at optimizing their payoffs. It will be an interesting future direction to further consider bounded rational decision makers, considering the impact of emotional and psychological factors.

First, regarding the Stage II problem where producers make their participation and service attribute selection decisions, this problem is an extended version of the Hotelling model [6] with a large (finite) number of producers. This is still an open problem [7]. To resolve this issue, we consider a large population approximation where each producer is non-atomic, i.e., a single producer's strategy choice does not affect the entire market. This approximation is reasonable given the large number of producers (and customers) on those platforms in practice. The remaining difficulty is to compute the asymmetric equilibrium, where producers of the same preference may choose different strategies at the equilibrium. This is significant more difficult than focusing on only the symmetric equilibrium as in many previous work discussed in the survey paper [8]. Despite these difficulties, we are able to prove that the Stage II game is a potential game [9], [10], based on which we derive the game equilibrium and corresponding equilibrium features.

Next, regarding the Stage I problem where the platform optimizes the value of DSF, this problem is non-convex and hence is challenging to solve. By exploiting the structure of the problem, we derive both the upper-bound and lower-bound of the optimal DSF. We also construct empirical examples to show how the optimal DSF and the optimal platform's payoff change with system parameters.

Our key contributions are listed as follows:

- *Donation-Based Market Formulation:* To the best of our knowledge, this is the first paper that presents a two-stage model of a donation-based market. We characterize how the platform optimizes DSF, and how the producers decide their participation and service attributes.

[4]Figure 1 is based on the stream data that we collected from Twitch. The data is collected every 15 minutes from Nov. 05 to Nov. 20, 2017.

[5]Note that although our analytical results are able to provide practical insights, there is still some significant work needed to help online platforms to make strategic marketing decisions. Many other factors need to be considered to reach a comprehensive marketing decision, including brand reputation, entire monetization ecosystem (e.g., advertisement, sponsorship), as well as user culture and psychology (e.g., sense of fairness).

- *Stage II Equilibrium:* Under the non-atomic producer approximation, we prove that the Stage II problem is a potential game and derive the asymmetric equilibrium. At the equilibrium, a larger DSF induces more producer participation and a better match to the customers' preferences. However, a larger DSF may not necessarily induce a larger total customers' donation.
- *Stage I Equilibrium:* Under the non-convex Stage I problem, we derive the upper-bound and lower-bound of the optimal DSF. We show that if it is harder to motivate producer participation, the lower-bound is larger. That is, the platform sets a higher DSF to motivate producers' participation. If the customer donation is more sensitive to the number of producers, the upper-bound is larger. That is, the platform sets a higher DSF to motivate the producers to fulfill the customers' preferences.
- *Numerical Results:* We further construct numerical examples, and show that as the mismatch between the producers' and customers' preferences increases, the optimal DSF either monotonically increases or decreases, while the platform's optimal payoff always decreases due to its increasing efforts on balancing the mismatch. As the producers' opportunity and deviation costs increase, the optimal DSF changes non-monotonically, while the platform's optimal payoff always decreases due to its increasing compensation on the costs.
- *A Case Study based on Empirical Twitch Data:* We collect two weeks' data about the streamer and viewer behaviors from Twitch. Based on the data, we demonstrate how to compute the platform's optimal DSF without knowing the producers' inherent preferences.

The rest of this paper is organized as follows. We review the existing works in Section II. We propose the system model in Section III. In Sections IV and V, we analyze the equilibrium of Stages II and I, respectively. We construct the numerical examples in Section VI, perform the case study with Twitch data in Section VII, and conclude in Section VIII.

II. Literature Review

A. Donation-Based Market

Most of the prior works on donation-based markets studied the customers' donation behaviors in these markets. Hu *et al.* [11] conducted an online survey to study why customers visit live streaming platforms. Hilvert-Bruce *et al.* [12] studied why customers donate. Zhu *et al.* [13] analyzed the data from Douyu (a live streaming platform in China) to investigate the customers' donation behaviors. Liu [14] investigated the monthly donation earned by streamers on live streaming platforms. Tang *et al.* [15] used an all-paid auction framework to understand the customers' donation behaviors.

Through data analysis, Jia *et al.* [16] discussed the producers' and customers' different preferences on live streaming platforms. The authors mentioned that the platform has to motivate producers to match the customers' preferences to increase the platform's revenue. However, as far as we know, there is no existing study analytically characterizing how the platform should motivate the producers' service attribute selection. As a first step, this work studies the platform's optimal DSF decision and analyzes the producers' service attribute selection in donation-based markets.

B. Hotelling Model

The Stage II of the two-stage game can be regarded as an extended version of the Hotelling model [6], [17] with a large finite number of producers. This still remains as an open problem in the literature when the number of producers is arbitrary [7]. Economides [18] studied a multi-producer model without discussing the producer equilibrium. Brenner [19] theoretically studied a three-producer case, and empirically studied four- to nine-producer cases. Behringer *et al.* [20] theoretically analyzed a four-producer case. However, the analyses in [19] and [20] cannot be easily generalized to the case of an arbitrary number of producers.

We overcome the difficulty by approximating the problem with non-atomic producers, where a single producer's strategy choice does not affect the market. We believe that such an approximation is reasonable given the very large number of streamers on practical live streaming platforms. Schmeidler [21] first analyzed a game with non-atomic players, and proved the existence of Nash equilibrium (without deriving the equilibrium). However, we are not aware of papers that explicitly characterizing the equilibrium of a general Hotelling model with non-atomic producers. Based on the reformulated model, we derive the asymmetric equilibrium with non-atomic players, which is a challenging problem according to [8].

III. System Model

We first introduce the system setting. Then, we define the two-stage game and the platform's and producers' payoffs.

A. System Setting

We first introduce the platform model and the service attribute. Then, we introduce the producer and customer models.

1) Platform: We consider a platform with a set of non-atomic producers and a set of non-atomic customers. Such a mass (non-atomic) setting corresponds to the case where the entire producer or customer population is large, under which a producer's or a customer's behavior does not affect the aggregated feature of the entire population. In practical systems, for example, Twitch often has thousands of streamers and millions of viewers on average (see Figure 1).

The producers provide services without mandatory charge. The customers enjoy the services and voluntarily donate to the producers. The donation will be shared between the producers and the platform with a fixed fraction.

2) Service Attribute: For simplicity, we consider one service attribute in this paper.[6] For example, on a live streaming platform, the service attribute can be the time of day of a stream or the type of the game to be streamed.

[6]Our model can be extended to the case of multi-attribute, under which the potential game in Section IV is still applicable. The extension can be done by extending the one-dimensional location set (to be discussed in the next paragraph) to a multi-dimensional one.

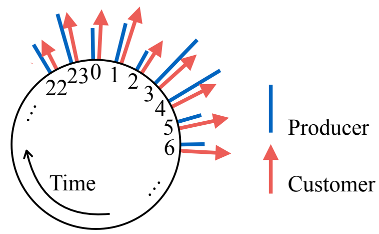


Fig. 2. An example of a circular service attribute model with time of day as the service attribute.

Similar as in Salop's circle Model (a circular extension of the Hotelling model) [22], we represent the service attribute using a unit length circle. Such a circular model has no extreme points (as in an interval model). Thus, it can better characterize realistic attributes, such as time of day and game type. We label the possible values of the service attribute by a set $\mathcal{L} = \{0, 1, 2, ..., L-1\}$, where we refer to each of the values as a location.[7] For the rest of this paper, we will use "location" and "value of the service attribute" interchangeably. Let $r_l \in [0, 1)$ denote the position along the circle that location l is located at. An example is shown in Figure 2. It shows a location set $\mathcal{L} = \{0, 1, ..., 23\}$, each representing an hour of the day, and location l is located at $r_l = l/24$ along the circle. Producers and customers are distributed along the circle based on their preferences over the service attribute, where the preferences are inherent and given. The heights of the bars and arrows denote the number of producers and customers preferring the corresponding locations, respectively.

3) Non-Atomic Producers: There are a total of N producers, where these N producers are regarded as a mass comprising non-atomic producers. Each producer has a preferred value of the service attribute. Let $n^l \in [0, 1]$ denote the fraction of the producers preferring location $l \in \mathcal{L}$, and $\sum_{l\in\mathcal{L}} n^l = 1$. Hence, the number of producers preferring location l is $N^l = n^l \times N$. For the rest of this paper, we will use either the fraction n^l or the number N^l depending on presentation simplicity. A producer can choose a location that is different from its preference, so as to avoid intense competition with other producers or to encounter more customers. For example, in Figure 2, a producer at hour 3 may want to deviate to hour 6 to encounter less producers and more customers.

4) Non-Atomic Customers: There are a total of M customers, where these M customers are regarded as a mass comprising non-atomic customers. Each customer has a preferred value of the service attribute. Let $m^l \in [0, 1]$ denote the fraction of customers preferring location $l \in \mathcal{L}$, and $\sum_{l\in\mathcal{L}} m^l = 1$. Hence, the number of customers preferring location l is $M^l = m^l \times M$. For simplification, we assume that these customers will stick to their preferences over the service attribute.

A customer will donate to the producers who provide services at the customer's preferred location. Instead of characterizing the donation behavior of each customer, we consider an aggregate donation function at each particular location. Let $D_l(M^\dagger, N^\dagger)$ denote the donation function at location $l \in \mathcal{L}$, given the number of producers $N^\dagger$ (according to the producers' choices) and the number of customers $M^\dagger$ (according to the customers' preferences) at that location. Note that the donation function $D_l(M^\dagger, N^\dagger)$ at different locations $l \in \mathcal{L}$ can be different. For example, considering the game attribute on live streaming platforms, viewers may be willing to donate more when they are watching a live streaming video on a more difficult game. In practice, this function can be obtained through data analysis [13]. We assume that $D_l(M^\dagger, N^\dagger)$ satisfies the following assumption.

Assumption 1 (Donation Function)**.** *For any location $l \in \mathcal{L}$, customer donation function $D_l(M^\dagger, N^\dagger)$ (i) is strictly increasing in $M^\dagger$, (ii) is strictly increasing and concave in $N^\dagger$, and (iii) has an elasticity that is smaller than one, i.e.,*

$$\eta_{l,M^\dagger}(N^\dagger) = \frac{[D_l(M^\dagger, N^\dagger)]_{N^\dagger} N^\dagger}{D_l(M^\dagger, N^\dagger)} \leq 1, \quad \forall M^\dagger, N^\dagger \in \mathbb{R}_+, \quad (1)$$

where $[D_l(M^\dagger, N^\dagger)]_{N^\dagger}$ denotes the partial derivative of $D_l(M^\dagger, N^\dagger)$ with respect to $N^\dagger$.

Assumption 1 is general and reasonable in practice. Specifically, Point (i) implies that as the number of customers increases, the total donation strictly increases. Point (ii) implies that as the number of producers increases, the total donation strictly increases but the marginal change decreases. On live streaming platforms, for example, more streamers imply a higher probability that a viewer can find his satisfactory streams so that he will donate more, while the probability of finding a satisfactory stream is concave in the number of streamers. Point (iii) on elasticity can be written as follows:

$$1 \geq \frac{[D_l(M^\dagger, N^\dagger)]_{N^\dagger} \cdot N^\dagger}{D_l(M^\dagger, N^\dagger)} \approx \frac{\%\Delta D_l(M^\dagger, N^\dagger)}{\%\Delta N^\dagger}. \quad (2)$$

This implies that a unit percentage increase in the number of producers leads to a percentage donation increase that is less than one. Because of this, the producers tend to avoid competition (e.g., a streamer would prefer to stream at a time when there are more viewers and less streamers). If Point (iii) is not satisfied, then all producers would prefer to stream at a few common time, which is trivial to analyze and does not fit into the practical data that we collect.

We now present an example of the donation function based on [13]. Paper [13] showed that the donation to a streamer increases with the number of viewers in the following manner:

$$\begin{aligned}&[\text{received donation per streamer}]\\ &\qquad = e^{b_0}([\text{viewers per streamer}])^{b_1}, \quad (3)\end{aligned}$$

where $b_0 = -1.17$ and $b_1 = 0.6$ based on empirical data. This leads to the following donation function:

$$D_l(M^\dagger, N^\dagger) = e^{b_0} \left(M^\dagger/N^\dagger\right)^{b_1} N^\dagger, \; l \in \mathcal{L}, \quad (4)$$

which is the per-producer donation $e^{b_0}(M^\dagger/N^\dagger)^{b_1}$ multiplied by the number of producers. We can show that this donation function $D_l(M^\dagger, N^\dagger)$ satisfies Assumption 1.

[7] The discrete locations characterize discrete attributes, such as game types. As the number of locations increases, the discrete set of locations can approximate a continuous set of locations.

TABLE I
KEY NOTATION.

Notation	Description
N, M	Total number of producers/customers
n^l, m^l	The fraction of producers/customers preferring location $l \in \mathcal{L}$
N^l, M^l	The number of producers/customers preferring location $l \in \mathcal{L}$
V	A producer's opportunity cost for participation
W	A producer's deviation cost per unit quadratic distance
$\mathcal{L}_s$	The set of locations served by the producers choosing location $s \in \mathcal{L}$
$\mathcal{S}_l$	$\mathcal{S}_l \triangleq \{s \in \mathcal{L} : l \in \mathcal{L}_s\}$, the set of producers' location choices that can serve location $l \in \mathcal{L}$
α	Platform's DSF decision
x_s^p	The fraction of the producers (over the entire producer mass) preferring location $p \in \mathcal{L}$ and choosing location $s \in \mathcal{L}$
$\boldsymbol{x}^p$	$\boldsymbol{x}^p = (x_s^p, \forall s \in \mathcal{L})$, the strategies of the producers preferring $p \in \mathcal{L}$
$\boldsymbol{x}$	$\boldsymbol{x} = (x_s^p, \forall p, s \in \mathcal{L})$, the strategies of all producers
$\widehat{N}_l(\boldsymbol{x})$	$\widehat{N}_l(\boldsymbol{x}) \triangleq \sum_{p\in\mathcal{L}} \sum_{s\in\mathcal{S}_l} x_s^p \times N$, the cumulative number of producers serving location $l \in \mathcal{L}$ under strategy $\boldsymbol{x}$

B. Two-Stage Game

We use a two-stage game to capture the sequential decision process among the platform and the producers. Note that this two-stage game is of complete information, i.e., the platform and producers are aware of the preference distributions of the customers and producers over the service attribute.[8] The two-stage game is follows.

- Stage I: The platform decides the DSF $\alpha \in [0, 1]$. This is the fraction of donation kept by producers.
- Stage II: Based on the announced DSF α, each producer decides whether to participate and what location to choose (if participating) concerning the service attribute.

Next, we elaborate on the producers' decisions in Stage II. Note that when a producer decides its service attribute, it needs to take into account the preferences of not only the customers but also other producers. We consider the aggregated decisions of the producers. Let x_s^p denote the fraction of producers preferring location $p \in \mathcal{L}$ and choosing location $s \in \mathcal{L}$.[9] Note that this fraction $x_s^p \in [0, 1]$ is the fraction over the entire producer mass. The strategies of the producers preferring location $p \in \mathcal{L}$ is characterized by $\boldsymbol{x}^p = (x_s^p, \forall s \in \mathcal{L})$. Let $\boldsymbol{x} = (x_s^p, \forall p, s \in \mathcal{L})$ denote the strategies of all the producers. In addition, any feasible strategy $\boldsymbol{x}$ should satisfy producer population constraint:

$$\sum_{s\in\mathcal{L}} x_s^p \leq n^p, \forall p \in \mathcal{L}. \tag{5}$$

Specifically, for producers preferring any location $p \in \mathcal{L}$, the sum of the producers choosing all locations should be no larger than the total number of producers of this type.

For a producer selecting a location s, we assume that it can serve the customers preferring the following $K \leq L$ locations (including location s). The corresponding set of locations is denoted by $\mathcal{L}_s \triangleq (\text{mod}(s + k - 1, L), k = 1, ..., K)$, where $\text{mod}(x, y)$ is remainder of the division of x by y. For example, if the producers stream for two hours ($K = 2$), choosing hour 3 (in Figure 2) means that the producers can serve customers preferring hours 3 and 4, i.e., $\mathcal{L}_3 = \{3, 4\}$.

For presentation convenience, let $\mathcal{S}_l \triangleq \{s \in \mathcal{L} : l \in \mathcal{L}_s\}$ denote the set of producers' location choices that can serve location $l \in \mathcal{L}$; hence, $\mathcal{S}_l = (\text{mod}(l + 1 - k, L), k = 1, ..., K)$. For example, in Figure 2, if the producers stream for two hours ($K = 2$), customers in location 3 can be served by the producers choosing any location in set $S_3 = \{2, 3\}$. Let $\widehat{N}_l(\boldsymbol{x}) \triangleq \sum_{p\in\mathcal{L}} \sum_{s\in\mathcal{S}_l} x_s^p \times N$ denote the cumulative number of producers serving location l under strategy $\boldsymbol{x}$.

C. Payoff Functions

Given the platform strategy α and the producer strategy $\boldsymbol{x}$, we define their payoffs as follows.

1) Platform's Payoff: The platform's payoff is equal to $1-\alpha$ fraction of the total donation from all customers:

$$G(\alpha, \boldsymbol{x}) = (1 - \alpha) \sum_{l\in\mathcal{L}} D_l(M^l, \widehat{N}_l(\boldsymbol{x})). \tag{6}$$

2) A Producer's Payoff: If a producer does not participate in the platform, it gains a zero payoff.[10]

If a producer preferring location $p \in \mathcal{L}$ participates and chooses a location $s \in \mathcal{L}$, its payoff is equal to the difference between the donation gain and its cost, i.e.,

$$F_s^p(\alpha, \boldsymbol{x}) = \alpha \times \sum_{l\in\mathcal{L}_s} U_l(M^l, \widehat{N}_l(\boldsymbol{x})) - C_s^p(V, W), \ \forall p, s \in \mathcal{L}. \tag{7}$$

Specifically, the donation gain is the DSF α multiplied by the average donation[11] that a producer can gain at the locations in set $\mathcal{L}_s$. The average donation at a location $l \in \mathcal{L}_s$ is

$$U_l(M^l, \widehat{N}_l(\boldsymbol{x})) = \frac{D_l(M^l, \widehat{N}_l(\boldsymbol{x}))}{\widehat{N}_l(\boldsymbol{x})}. \tag{8}$$

The cost contains a fixed opportunity cost V and a distance-associated deviation cost with a coefficient W. Formally,

$$C_s^p(V, W) = V + W \times (\min\{|r_p - r_s|, 1 - |r_p - r_s|\})^2, \tag{9}$$

where $\min\{|r_p - r_s|, 1 - |r_p - r_s|\}$ is the producer's deviation distance along the circle. The quadratic form of the deviation cost is used to characterize the producers' increasing marginal costs on the deviation, similar as in the Hotelling model [17].

The key notation is summarized in Table I. The two-stage game is analyzed using backward induction, where Stage II and Stage I is analyzed in Sections IV and V, respectively.

[8] In practice, the preference distributions of the customers and producers can be estimated through real-world survey and historical data analysis. To address the potential incomplete information scenario, in the case study in Section VII, we demonstrate the approach of computing the platform's optimal DSF with only the producers' actual behaviors (instead of their inherent preferences).

[9] In this paper, we use superscripts to denote preferences and subscripts to denote decisions.

[10] If the non-participation induces a positive payoff, we can normalize it to zero by adjusting the value of the opportunity cost V to be defined in (9).

[11] We assume that the producers choosing the same location will equally share the donation from the customers at this location. The non-equal sharing case will be a future work.

IV. Stage II: Producer Location Equilibrium

In Stage II, given any DSF α, we analyze the producer location game as follows.

Definition 1 (Stage II Producer Location Game)**.**

- *Players: all producers;*
- *Strategies: each producer preferring any location $p \in \mathcal{L}$ selects a location $s \in \mathcal{L}$. The aggregate strategy is represented by $\boldsymbol{x} = (x_s^p, \forall p, s \in \mathcal{L})$;*
- *Payoffs: $F_s^p(\alpha, \boldsymbol{x})$ for each producer preferring a location $p \in \mathcal{L}$ and choosing a location $s \in \mathcal{L}$, as defined in* (7).

Next, we first define the producer location equilibrium, and then derive the equilibrium and its corresponding features.

A. Equilibrium Definition

We first define the support correspondence. Then, we define a producer's best response and the producer location equilibrium based on such a correspondence.

1) Support Correspondence: We define a correspondence that outputs a vector's positive elements.

Definition 2 (Support Correspondence)**.** *For a vector $z \in \mathbb{R}_+^{1\times L}$, the support correspondence $S(z) = \{s \in \mathcal{L} : z_s > 0\}$ is the set of indices corresponding to positive elements in z.*

For example, if $z = (0.3, 0, 0.2, 0)$, then $S(z) = \{1, 3\}$. For any strategy $\boldsymbol{x}^p$, the correspondence $S(\boldsymbol{x}^p) = \{s \in \mathcal{L} : x_s^p > 0\}$ indicates the set of locations which are chosen by the producers preferring location p under the strategy $\boldsymbol{x}^p$.

2) Best Response: We now define a producer's best response. For a producer preferring location $p \in \mathcal{L}$, its best response locations, denoted by a set $BR^p(\alpha, \boldsymbol{x})$, are the locations that induce the maximum producer payoff, i.e.,

$$BR^p(\alpha, \boldsymbol{x}) = \arg\max_{s\in\mathcal{L}} F_s^p(\alpha, \boldsymbol{x}). \tag{10}$$

Normally, the best response is defined as a correspondence of all other producers' strategies excluding the producer's own strategy. However, due to the non-atomic producer assumption, the change of one producer's strategy does not affect the aggregate strategies of all the producers. This allows us to directly write the best response as a correspondence of $\boldsymbol{x}$.

Based on a single producer's best response, we define a correspondence representing the aggregate best response of all the producers preferring location $p \in \mathcal{L}$:

$$ABR^p(\alpha, \boldsymbol{x}) = \{z \in \mathbb{R}_+^{1\times L} : \\ S(z) \subset BR^p(\alpha, \boldsymbol{x}), \textstyle\sum_{s\in\mathcal{L}} z_s \le n^p\}. \tag{11}$$

Specifically, the aggregate best response for the producers preferring location p is any vector z such that (i) all its elements (locations) with positive producer fractions belong to $BR^p(\alpha, \boldsymbol{x})$, and (ii) it satisfies the producer population constraint im (5). Note that this aggregate best response $ABR^p(\alpha, \boldsymbol{x})$ may not be unique.

3) Producer Location Equilibrium: Producer location equilibrium is defined as the fixed point of the best response.

Definition 3 (Producer Location Equilibrium)**.** *Given any α, producer location strategy $\boldsymbol{x}$ is an equilibrium if and only if the aggregate strategy of the producers preferring any location $p \in \mathcal{L}$ belongs to their aggregate best response under $\boldsymbol{x}$, i,e.,*

$$\boldsymbol{x}^p \in ABR^p(\alpha, \boldsymbol{x}), \forall p \in \mathcal{L}. \tag{12}$$

An interpretation of this equilibrium is that a strategy $\boldsymbol{x}$ is an equilibrium if and only if the producers' aggregate best responses under the strategy $\boldsymbol{x}$ can recover this strategy $\boldsymbol{x}$.

B. Deriving the Producer Location Equilibrium

Directly computing the equilibrium based on the best response is challenging, due to the challenge of computing the fixed point of the multi-dimensional best response mapping of an $L \times L$-dimensional vector $\boldsymbol{x} = (x_s^p, \forall p, s \in \mathcal{L})$. Instead of directly deriving the equilibrium distribution, we first prove that the Stage II game is a potential game. Under this, all the producers' payoffs can be related to a common potential function, which allows us to characterize the equilibrium by solving a problem that maximizes the potential function. Then, we can derive the producer location equilibrium.

We first define the potential game according to [9].

Definition 4 (Potential Game)**.** *The Stage II game with non-atomic players is a potential game if there exists a continuously differentiable potential function $f : \boldsymbol{x} \to \mathbb{R}$ such that*

$$\frac{\partial f(\alpha, \boldsymbol{x})}{\partial x_s^p} = F_s^p(\alpha, \boldsymbol{x}), \forall s, p \in \mathcal{L}. \tag{13}$$

We now show that the Stage II game is a potential game. The key proof is to identify the potential function $f(\alpha, \boldsymbol{x})$ that satisfies equation (13). Note that there does not exist a general methodology for doing this. We have to identify the potential function by exploiting the specific structure of the problem.

Lemma 1 (Stage II Game as Potential Game)**.** *Given any α, the Stage II game is a potential game with non-atomic players, which has a potential function*

$$f(\alpha, \boldsymbol{x}) = \alpha \times \sum_{l\in\mathcal{L}} \int_0^{\widehat{N}_l(\boldsymbol{x})} U_l(M^l, z)dz - V \times \sum_{s\in\mathcal{L}} \sum_{p\in\mathcal{L}} x_s^p \\ - W \times \sum_{s\in\mathcal{L}} \sum_{p\in\mathcal{L}} x_s^p \left(\min\{|r_p - r_s|, 1 - |r_p - r_s|\}\right)^2. \tag{14}$$

Lemma 1 is proven by showing that the potential function (14) satisfies equation (13).

Showing that the game is a potential game allows us to characterize the Stage II producer location equilibrium by solving an optimization problem, which is easier than finding the fixed point of the producers' best responses. Formally,

Theorem 1 (Producer Location Equilibrium)**.** *The set of producer location equilibria of the Stage II game is the set of global optimal solutions to the following problem:*

$$\boldsymbol{x}^*(\alpha) \triangleq \mathop{\textit{arg maximize}}_{\boldsymbol{x}\ge 0} \quad f(\alpha, \boldsymbol{x}) \tag{15a}$$

$$\textit{subject to} \quad \textstyle\sum_{s\in\mathcal{L}} x_s^p \le n^p, \ \forall p \in \mathcal{L}, \tag{15b}$$

(STAGE II-NE)

which maximizes the potential function $f(\alpha, \boldsymbol{x})$ subjecting to producer population constraint in (5).

More specifically, a vector $\boldsymbol{x}$ is an equilibrium, i.e., $\boldsymbol{x} \in \boldsymbol{x}^*(\alpha)$, if and only if there exists a pair of $\boldsymbol{\mu} \in \mathbb{R}^{1\times L}$ and $\boldsymbol{\lambda} \in \mathbb{R}^{L\times L}$ such that the following constraints are satisfied:

$$F_s^p(\alpha, \boldsymbol{x}) = \mu^p - \lambda_s^p, \quad \forall p, s \in \mathcal{L}, \tag{16a}$$
$$\lambda_s^p x_s^p = 0,\ \lambda_s^p \geq 0,\ x_s^p \geq 0, \quad \forall p, s \in \mathcal{L}, \tag{16b}$$
$$(\textstyle\sum_{s\in\mathcal{L}} x_s^p - n^p)\mu^p = 0,\ \mu^p \geq 0, \quad \forall p \in \mathcal{L}, \tag{16c}$$
$$\textstyle\sum_{s\in\mathcal{L}} x_s^p \leq n^p, \quad \forall p \in \mathcal{L}. \tag{16d}$$

(STAGE II-NE-CONDITION)

Proof. We have proven in Lemma 1 that the Stage II game is a potential game, so its equilibria are the solutions to Problem (STAGE II-NE) [9]. On the other hand, the conditions (STAGE II-NE-CONDITION) are the KKT conditions of Problem (STAGE II-NE). To show that (STAGE II-NE-CONDITION) are the conditions for the equilibria, we have to show that the KKT conditions of Problem (STAGE II-NE) is necessary and sufficient conditions to its global optimal solutions. This is true because the $f(\alpha, \boldsymbol{x})$ in (15a) is concave (by checking its Hessian Matrix), and the constraint (15b) fulfills the Slater's condition. □

The producer location equilibrium may not be unique. However, we can show that any of the equilibria leads to the same set of producers' payoffs, the same set of producer distribution, and the same platform's payoff,

Corollary 1 (Unique Equilibrium Outputs)**.** *Under any given α, any equilibrium of the Stage II game induces*

- *the same set of producers' payoffs, i.e.,*
$$F_s^p(\alpha, \boldsymbol{x}) = \widetilde{F}_s^p(\alpha), \forall p, s \in \mathcal{L}, \boldsymbol{x} \in \boldsymbol{x}^*(\alpha), \tag{17}$$
where $\widetilde{F}_s^p(\alpha)$ denotes the identical payoff of the producer preferring location p and selecting location s in any of the producer location equilibria given α.
- *the same producer distribution, i.e.,*
$$\widehat{N}_l(\boldsymbol{x}) = \widetilde{N}_l(\alpha), \forall l \in \mathcal{L}, \boldsymbol{x} \in \boldsymbol{x}^*(\alpha), \tag{18}$$
where $\widetilde{N}_l(\alpha)$ denotes the identical number of producers serving the customers preferring location l in any of the producer location equilibria given α.
- *the same platform's payoff, i.e.,*
$$G(\alpha, \boldsymbol{x}) = G(\alpha), \forall \boldsymbol{x} \in \boldsymbol{x}^*(\alpha), \tag{19}$$
where $G(\alpha)$ denotes the identical platform's payoff under any of the producer location equilibria given α.

Specifically, the dual variables of Problem (STAGE II-NE), i.e., $\boldsymbol{\mu}$ and $\boldsymbol{\lambda}$, are unique, because the constraints (15b) are linearly independent [23]. Hence, according to (16a), under a fixed α, any of the equilibria induces the same set of producers' payoffs, i.e., $F_s^p(\alpha, \boldsymbol{x}) = \widetilde{F}_s^p(\alpha), \forall p, s \in \mathcal{L}, \boldsymbol{x} \in \boldsymbol{x}^*(\alpha)$. Based on this, we can show that under a fixed α, any of the equilibria induces the same set of cumulative number of producers at all locations, i.e., $\widehat{N}_l(\boldsymbol{x}) = \widetilde{N}_l(\alpha), \forall l \in \mathcal{L}, \boldsymbol{x} \in \boldsymbol{x}^*(\alpha)$.

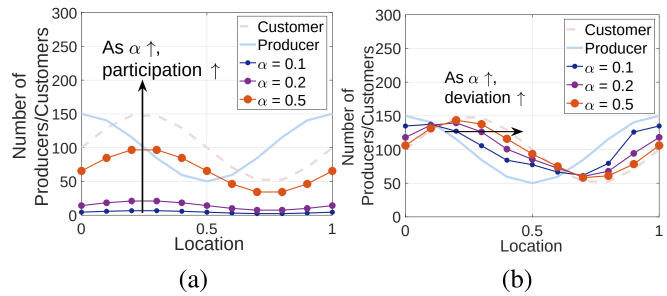


Fig. 3. The impact of α on producer location equilibrium: (a) $W = 0$ and $V = 0.1$, under which producers always perfectly match the customers' preferences; (b) $W = 0.4$ and $V = 0$, under which producers always fully participate.

This is because the mapping from the set of producers' payoffs (defined in (7)) to the set of cumulative number of producers at the locations is a one-to-one correspondence, due to the strictly increasing donation function in the number of producers as in Assumption 1. Hence, from the platform's point of view, given any α, it achieves the same payoff under any of the producer location equilibria in Stage II, as its payoff (defined in (6)) depends on only the cumulative number of producers at those locations, i.e., $\widehat{N}_l(\boldsymbol{x})$, $l \in \mathcal{L}$.

C. Impact of α on Producer Location Equilibrium

Based on the conditions in Theorem 1, we show how the producer location equilibrium changes with DSF α. A key insight is that a larger α leads to more producer participation and a better match to the customers' preferences.

Given any opportunity cost, deviation cost, and producers' and customers' preferences, the producer location equilibrium changes with α as follows.

Proposition 1 (Impact of α on Producer Location Equilibrium)**.** *As α increases, for any $\boldsymbol{x} \in \boldsymbol{x}^*(\alpha)$,*

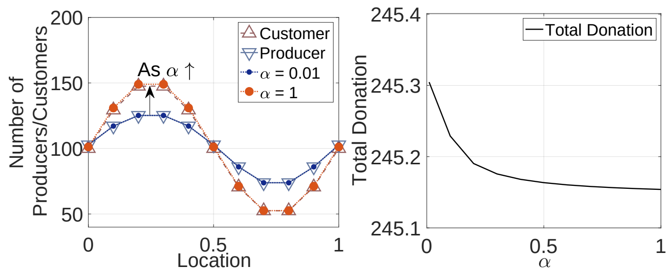
(1) *the total participation (i.e, $\sum_{s\in\mathcal{L}} \sum_{p\in\mathcal{L}} x_s^p$) increases;*
(2) *the total deviation (i.e, $\sum_{s\in\mathcal{L}} \sum_{p\in\mathcal{L}} x_s^p (\min\{|r_p - r_s|, 1 - |r_p - r_s|\})^2$) increases.*

Specifically, the potential function (14) is a weighted sum of three functions with the corresponding weights α, V, and W, respectively. As these weights change, the producer location equilibrium changes accordingly as in Proposition 1. The proof is given in Appendix A.

Figure 3 shows the impact of α on the producer location equilibrium. The donation function is the one defined in (4). We set $N = M = 1000$ and $K = 1$. The x-axis represents the location. The y-axis shows the number of producers or customers. The "producer" and "customer" curves correspond to the producers' and customers' location preferences, respectively. The curves labeled with $\alpha = 0.1$, 0.2, and 0.5 are the producer location equilibrium under the particular α.

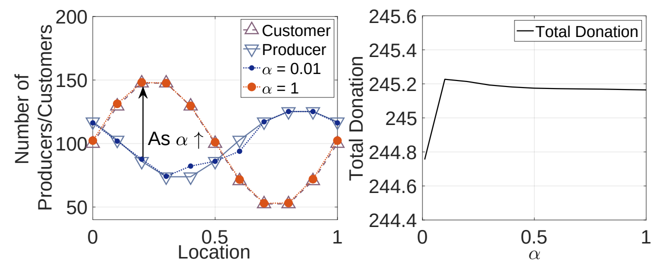
Figure 3 (a) shows the results with $W = 0$ and $V = 0.1$, under which producers always perfectly match[12] the customers' preferences due to the zero deviation cost W. In

[12]The term "perfectly match" means that all the locations induce the same average donation for the producers such that no producer can increase its payoff by further deviation even if there is no deviation cost.



(a) (b)

Fig. 4. The impact of α under the producer preference distribution that maximizes the total donation (with $W = 0.4$, $V = 0$): (a) producer and customer distributions; (b) total donation.



(a) (b)

Fig. 5. The impact of α under the producer preference distribution that does not maximize the total donation (with $W = 0.4$, $V = 0$): (a) producer and customer distributions; (b) total donation.

this case, under any α, producers are distributed in a shape that is similar as the customers' preferences. As α increases, the producer participation increases, i.e., the total number of participating producers increases. Figure 3 (b) shows the result with $W = 0.4$ and $V = 0$, under which producers always fully participate due to the zero opportunity cost V. In this case, as α increases, the producer matching increases, i.e., producers' location choices deviate from producers' preferences (i.e., the blue solid line) to match customers' preferences (i.e., the red dash line). To sum up, a larger α leads to more producer participation and a better match to the customers' preferences.

D. Impact of α on Total Customer Donation

Based on Theorem 1, we further show how the total customer donation (i.e., $\sum_{l\in\mathcal{L}} D_l(M^l, \widetilde{N}_l(\alpha))$) changes with DSF α. It is intuitive to expect that as α increases, the total donation increases, because the producers increasingly deviate to match the customers' preferences. However, we show that this is not always true for the following reason. The increasing deviation of the producers to match the customers' preferences implies a decreasing number of the producers selecting the locations preferred by fewer customers. It also implies an increasing number of producers selecting the locations preferred by more customers. Due to the concavity of the donation function, the value of the donation decreasing at the locations with fewer customers (resulting from the decreasing number of producers selecting those locations) may be greater than the value of the donation increasing at the locations with more customers (resulting from the increasing number of producers selecting those locations). Hence, as α increases, the total donation may not increase. We also show that under some particular donation functions, as α increases, the total donation increases.

We first show two examples of how total donation changes with α in Figures 4 and 5. We consider a donation function $D_l(M^\dagger, N^\dagger) = M^\dagger \times N^\dagger/(4N^\dagger + 8)$ (which satisfies Assumption 1), under which the producer distribution maximizing the total donation is not the producer distribution perfectly matching the customers' preferences. The Figures 4 (a) and 5 (a) show the customer preference (dash triangle line), producer preference (solid reverse triangle line), and the producer equilibrium distributions when $\alpha = \{0.01, 1\}$ at different locations. The Figures 4 (b) and 5 (b) show how the total donation changes with α. In Figure 4, we set the producer's inherent preference distribution to be the one that maximizes the total donation, under which any deviation will lead to total donation decreasing. Consequently, although imposing a larger α may lead to a better match to customers' preferences (triangle dash line) in Figure 4 (a), it leads to a lower total donation in Figure 4 (b). In Figure 5, the producer preference distribution is not the one that maximizes the total donation, under which deviation may lead to total donation increasing. Hence, a larger α may lead to a higher donation in Figure 5 (b), because of the resulting better match between the producers and the customers in Figure 5 (a).

Remark 1 (Impact of α on Total Donation)**.** *As α increases, the total donation is not necessarily monotonically increasing or decreasing, which depends on the customers' and producers' preferences over the service attribute.*

However, under some specific donation function, the total donation increases in α.

Lemma 2 (Impact of α on Total Donation under a Special Case)**.** *If the donation function is in the form of $D_l(M^\dagger, N^\dagger) = a_0(M^\dagger/N^\dagger)^{a_1} N^\dagger$ with $a_1 < 1$ for $l \in \mathcal{L}$, the total donation $\sum_{l\in\mathcal{L}} D_l(M^l, \widetilde{N}_l(\alpha))$ increases in α.*

The detailed proof is given in Appendix B. Specifically, under the donation function $D_l(M^\dagger, N^\dagger) = a_0(M^\dagger/N^\dagger)^{a_1} N^\dagger$ with $a_1 < 1$ for $l \in \mathcal{L}$, the following equality holds:

$$\sum_{l\in\mathcal{L}} \int_0^{\widehat{N}_l(\boldsymbol{x})} U_l(M^l, z)dz = \frac{1}{1-a_1} \sum_{l\in\mathcal{L}} D_l(M^l, \widehat{N}_l(\boldsymbol{x})). \quad (20)$$

By substituting (20) into the potential function (14), according to Problem (STAGE II-NE), we can prove that as α increases, the total donation $\sum_{l\in\mathcal{L}} D_l(M^l, \widetilde{N}_l(\alpha))$ increases.

V. STAGE I: PLATFORM DSF DECISION

In Stage I, the platform chooses the DSF α to maximize its payoff. We first present the platform's payoff optimization problem. As the problem is non-convex and cannot be solved in closed-form, we derive the lower-bound and upper-bound of the optimal solution. Finally, we demonstrate how system parameters affect the optimal DSF.

A. Platform's Payoff Maximization Problem

In Stage I, the platform selects the optimal fraction α^* that maximizes its payoff. Formally,

$$\alpha^* \triangleq \arg \max_{\alpha \in [0,1]} \quad G(\alpha). \qquad \text{(STAGE I-NE)}$$

Here $G(\alpha)$ is the platform's payoff under the producer location equilibrium given an α, as defined in (19).

Problem (STAGE I-NE) is a non-convex optimization problem due to the non-convex objective function $G(\alpha)$. Specifically, $G(\alpha)$ is a piece-wise function that is not always differentiable. In addition, this piece-wise function may not be quasi-concave, so we cannot use an effective bisection algorithm [24] to solve the problem. Hence, it is difficult to derive the closed-form optimal solution to Problem (STAGE I-NE).

B. Optimal Solution Bound and Approximate Solution

Despite the non-convexity of Problem (STAGE I-NE), we can characterize the lower-bound and upper-bound of the optimal solution. The bounds reveal some insights over the optimal DSF. Meanwhile, those bounds shrink the range of optimal DSF, within which we would use an exhausted searching method to obtain the optimal DSF approximately.

1) Lower-Bound of α^*: The lower-bound of the optimal DSF α^* is given as follows:

Proposition 2 (Lower-Bound of α^*)**.** *The optimal α^* is lower-bounded by $\underline{\alpha}$ as follows:*

$$\underline{\alpha} = \min\left\{\underline{\alpha}_1, \underline{\alpha}_2\right\}, \tag{21}$$

where $\underline{\alpha}_1$ is the optimal solution to the following problem:

$$\begin{aligned}
&\text{maximize} && \alpha && \text{(22a)}\\
&\text{subject to} && \alpha \times \sum_{l \in \mathcal{L}_p} U_l(M^l, x_l^l \times N) = V, \forall p \in \mathcal{L}, && \text{(22b)}\\
& && x_l^l \le n^l, \forall l \in \mathcal{L}, && \text{(22c)}\\
&\text{variables} && 0 \le \alpha \le 1, x_l^l \ge 0, \forall l \in \mathcal{L}, && \text{(22d)}
\end{aligned}$$

(LOWER BOUND - 1)

and $\underline{\alpha}_2$ is the optimal solution to the following problem:

$$\begin{aligned}
&\text{maximize} && (1-\alpha) \times \sum_{l \in \mathcal{L}} D_l(M^l, x_l^l \times N) && \text{(23a)}\\
&\text{subject to} && \alpha \times \sum_{l \in \mathcal{L}_p} U_l(M^l, x_l^l \times N) = V, \forall p \in \mathcal{L}, && \text{(23b)}\\
&\text{variables} && 0 \le \alpha \le 1, x_l^l \ge 0, \forall l \in \mathcal{L}. && \text{(23c)}
\end{aligned}$$

(LOWER BOUND - 2)

The detailed proof is shown in Appendix C. For the definition of $\underline{\alpha}_1$ and $\underline{\alpha}_2$ in Problems (LOWER BOUND - 1) and (LOWER BOUND - 2), we consider the scenario that at the producer location equilibrium, each producer participating in the platform chooses its preferred location and gains a zero payoff. In this case, each participating producer is already gaining his maximum payoff (i.e., the best response to other producers' strategy choices), and deviating to another location (other than his current preferred location) will only lead to a zero or negative payoff gain. We refer to this scenario as *non-deviation producer participation*. Under this case, the conditions (16a), (16b), and (16c) in (STAGE II-NE-CONDITION) can be transformed into one constraint, as in (22b) and (23b). In Problem (LOWER BOUND - 1), $\underline{\alpha}_1$ is the maximum value of the DSF α in Stage I that leads to the non-deviation producer participation scenario at the producer location equilibrium, given the producer population constraint (23c). In Problem (LOWER BOUND - 2), $\underline{\alpha}_2$ is the value of the DSF α in Stage I that maximizes the platform's payoff, under the relaxation of the producer population constraint. That is, $\underline{\alpha}_2$ corresponds to the optimal DSF if there is no producer population constraint. Consequently, if $\alpha \le \underline{\alpha}_1 \le \underline{\alpha}_2$, then the platform's payoff increases in α until $\alpha = \underline{\alpha}_1$, above which the non-deviation producer participation scenario may no longer hold at the producer location equilibrium. If $\alpha \le \underline{\alpha}_2 \le \underline{\alpha}_1$, the platform's payoff increases in α until $\alpha = \underline{\alpha}_2$, above which the platform's payoff decreases with α. Hence, the platform's payoff always increases in α when $\alpha \le \min\left\{\underline{\alpha}_1, \underline{\alpha}_2\right\}$, so $\min\left\{\underline{\alpha}_1, \underline{\alpha}_2\right\}$ is the lower-bound of the optimal α^*.

Intuitively, $\underline{\alpha}_1$ reveals how hard it is to motivate producers to deviate from their preferred locations (in order to better match the customers' preferences). Meanwhile, $\underline{\alpha}_2$ reveals how beneficial it is for the platform to increase α under the scenario of non-deviation producer participation. If either it is more difficult to motivate (i.e., larger $\underline{\alpha}_1$) or it is more beneficial (i.e., larger $\underline{\alpha}_2$), the lower bound of α^* is larger such that the platform provides more incentives to the producers.

2) Upper-Bound of α^*: It is difficult to characterize the upper-bound of the optimal DSF α^* for any K (where K is the number of locations a producer can serve), as the expression of the producer location equilibrium changes in α. Next, we first show the upper-bound of α^* under $K = 1$. Then we numerically demonstrate that for some particular donation functions, the same upper-bound applies for any $K \ge 1$.

The upper-bound of α^* under $K = 1$ is as follows:

Proposition 3 (Upper-Bound of α^* under $K = 1$)**.** *The optimal α^* is upper-bounded by $\overline{\alpha}$ that satisfies*

$$\overline{\alpha} = \max\{\eta_{l,M^l}(\tilde{N}_l^\dagger), \forall l \in \mathcal{L}\}, \tag{24}$$

where $\tilde{N}_l^\dagger$ is the value that satisfies $[D_l(M^l, \tilde{N}_l^\dagger)]_{\tilde{N}_l^\dagger} = V$ for $l \in \mathcal{L}$, and function $\eta_{l,M^\dagger}(N^\dagger)$ is defined in (1).

The proof is given in Appendix D. Specifically, the upper-bound $\overline{\alpha}$ is characterized by the elasticity of the customers' donation. When the elasticity is larger, the customers' donation is more sensitive to the number of producers serving their preferred locations. Thus, the platform should increase DSF to incentivize producers to participate and to match customers' preferences. Such an upper-bound exists when the following conditions hold: $\lim_{N^\dagger \to 0}[D_l(M^\dagger, N^\dagger)]_{N^\dagger} \ge V$ and $\lim_{N^\dagger \to \infty}[D_l(M^\dagger, N^\dagger)]_{N^\dagger} \le V$ for all $M^\dagger \in \mathbb{R}_+$, $l \in \mathcal{L}$. Intuitively, these conditions imply that when there is no producer choosing a location, the marginal donation increase (by adding a producer to the location) is sufficiently large. When there are infinite number of producers choosing a location, the marginal donation increase is sufficiently small. Under these conditions, $\tilde{N}_l^\dagger$ always exists and is unique for $l \in \mathcal{L}$ due to the strictly concavity of $D_l(M^\dagger, N^\dagger)$ in $N^\dagger$.

According to Proposition 3, we could obtain the upper-bound of the optimal DSF under the donation function in (4):

Corollary 2 (Upper-Bound of α^* under $K = 1$)**.** *Under donation function $D_l(M^\dagger, N^\dagger) = e^{b_0}\left(M^\dagger/N^\dagger\right)^{b_1} N^\dagger$ for $l \in \mathcal{L}$, the optimal α^* is upper-bounded by $\overline{\alpha}$ that satisfies*

$$\overline{\alpha} = 1 - b_1. \tag{25}$$

When the donation function is the one defined in (4), we perform a simulation to show that Corollary 2 holds for any $K \geq 1$. We evaluate the results under $b_1 = \{0.2, 0.4, 0.6, 0.8\}$. For each value of b_1, we randomly generate 1000 simulation rounds, each of which is generated based on the following distributions: K follows a discrete uniform distribution over $\{1, 2, ..., 10\}$; M^l and N^l follow a discrete uniform distribution over $\{1, 2, ..., 1000\}, \forall l \in \mathcal{L}$; V follows a uniform distribution within $[0, 0.4]$; W follows uniform distribution within $[0, 0.8]$. For each simulation, we perform an exhausted searching over discrete set $\alpha = \{0.01, 0.02, ..., 1\}$ to find the optimal α^* approximately. The key observation is as follows.

Remark 2 (Upper-Bound of α^* under $K \geq 1$)**.** *The simulation result suggests that under the donation function in* (4)*, the upper-bound of α^* is equal to* $1 - b_1$ *for general $K \geq 1$, the same as the upper-bound in Corollary 2 for $K = 1$.*

3) Computing the Optimal α: The upper-bound and lower-bound shrink the range of the optimal DSF α^*. To compute the optimal DSF, we can perform an exhausted searching within interval $[\underline{\alpha}, \overline{\alpha}]$. Specifically, we divide the internal into J segments. The approximate optimal solution is $\alpha_J^* = \arg\max\{G(\alpha) | \alpha \in \{\underline{\alpha} + (\overline{\alpha} - \underline{\alpha})j/(J-1), j = 0, 1, ..., J-1\}\}$. Let $\alpha^* = \arg\max_{\alpha \in [0,1]} G(\alpha)$ be the actual optimal solution. We can prove that the gap between $G(\alpha_J^*)$ and $G(\alpha^*)$ is bounded. Formally,

Lemma 3 (Optimal Solution Approximation)**.** *Given any ϵ, there always exists a threshold $\underline{J}$ such that $|G(\alpha_J^*) - G(\alpha^*)| \leq \epsilon$ for any $J \geq \underline{J}$.*

Proof. We first prove that the platform's payoff $G(\alpha)$ is continuous in α. Specifically, $G(\alpha)$ can be written as follows:

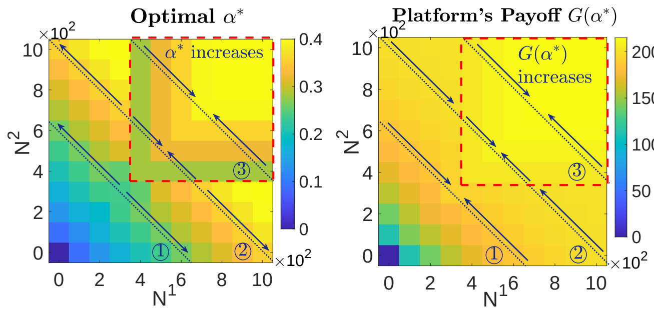
$$G(\alpha) = \underset{\alpha \in [0,1]}{\text{maximize}} \quad (1-\alpha) \sum_{l \in \mathcal{L}} D_l(M^l, \widehat{N}_l(\boldsymbol{x})) \tag{26a}$$

$$\text{subject to} \quad \boldsymbol{x} = \arg\max_{\boldsymbol{x} \geq 0} f(\alpha, \boldsymbol{x}). \tag{26b}$$

According to Maximum Theory [25],[13] $\boldsymbol{x}$ is continuous in α and further $G(\alpha)$ is continuous in α.

Then, we prove Lemma 3. Due to the continuity of $G(\alpha)$ in α, for any $\epsilon > 0$, there always exists a $\delta_\epsilon > 0$ such that $|G(\alpha) - G(\alpha^*)| \leq \epsilon$ for any $|\alpha - \alpha^*| \leq \delta_\epsilon$. We define $\underline{J} = \lceil 1/\delta_\epsilon \rceil + 1$. For any $J \geq \underline{J}$, there always exists an $\hat{\alpha}$ (which may not be equal to α_J) such that $|\hat{\alpha} - \alpha^*| \leq \delta_\epsilon$, so that $|G(\hat{\alpha}) - G(\alpha^*)| \leq \epsilon$. According to the definition of α_J, $|G(\alpha_J) - G(\alpha^*)| \leq |G(\hat{\alpha}) - G(\alpha^*)| \leq \epsilon$. □

[13]For a function $f^* : \Theta \to \mathbb{R}$ defined by $f^*(\theta) = \sup\{f(x, \theta) : x \in C(\theta)\}$, if C is continuous in θ, then f^* is continuous.



(a) (b)

Fig. 6. An example with different N^1 and N^2: (a) optimal α^*; (b) optimal platform's payoff $G(\alpha^*)$. The area within the red dash boundaries correspond to the case where the non-deviation producer participation scenario happens at the producer location equilibrium under the optimal α^* in Stage I.

C. Impact of V and W on the Optimal DSF α^*

We characterize the impact of costs V and W on α^*, as given in the following proposition.

Proposition 4 (Impact of V and W on α^*)**.** *There exists a V° such that for all $V \geq V^\circ$, the optimal DSF $\alpha^* = \underline{\alpha}_2$, under which producers do not fully participate at the producer location equilibrium (i.e., $\mu^p = 0$ for all $p \in \mathcal{L}$ in (STAGE II-NE-CONDITION)). In addition, there exists a W° such that for all $W \geq W^\circ$, the optimal $\alpha^* = \widetilde{\alpha}$, where the producers never deviate at the producer location equilibrium. Here $\widetilde{\alpha}$ is the optimal solution to the following problem:*

$$\text{maximize} \quad (1-\alpha) \times \sum_{l \in \mathcal{L}} D_l(M^l, x_l^l \times N), \tag{27a}$$

$$\text{subject to} \quad \alpha \times \sum_{l \in \mathcal{L}_p} U_l(M^l, x_l^l \times N) \geq V, \forall p \in \mathcal{L}, \tag{27b}$$

$$x_l^l \leq n^l, \forall l \in \mathcal{L}, \tag{27c}$$

$$\text{variables} \quad 0 \leq \alpha \leq 1, x_l^l \geq 0, \forall l \in \mathcal{L}. \tag{27d}$$

The proof is provided in Appendix E. The insights of Proposition 4 are as follows. When the opportunity cost is larger than the threshold V°, the platform should set the DSF to be $\underline{\alpha}_2$ (which is defined in (LOWER BOUND - 2)) and give up motivating producer participation. When the deviation cost W is larger than the threshold W°, the platform should set the DSF to be $\widetilde{\alpha}$ and give up motivating producers to match the customers' preferences. The $\widetilde{\alpha}$ is the DSF that maximizes the platform's payoff where no producer deviates.

VI. NUMERICAL RESULTS

In this section, we use simulations to understand how the optimal DSF and the platform's optimal payoff are affected by certain factors. These factors include the mismatch between the producers' and customers' preferences, the producers' opportunity cost, and their deviation cost.

Our numerical results show several insights. (i) As the mismatch between the producers' and customers' preferences increases, the optimal DSF either monotonically increases or decreases. However, the platform's optimal payoff always decreases due to its increasing efforts on balancing the mismatch. Hence, a lower degree of preference mismatch benefits the platform more. (ii) As the opportunity and deviation costs increase, the optimal DSF changes non-monotonically. However, the platform's optimal payoff always decreases due to its increasing compensation on the costs. Thus, lower opportunity and deviation costs benefit the platform more.

A. Impact of the Preference Mismatch

We aim to study how the mismatch between the producers' and customers' preferences affects the optimal DSF and the platform's optimal payoff. To visually demonstrate the impact, we consider an example with two locations $\mathcal{L} = \{0, 1\}$ and $K = 1$. We set $V = 0.1$ and $W = 0.1$. We assume that the number of customers is the same at both locations, i.e., $M^1 = M^2 = 500$.[14] We evaluate the system results under different N^1 and N^2 as in Figure 6. The results include the optimal α^* in Figure 6 (a) and the optimal platform's payoff $G(\alpha^*)$ in Figure 6 (b). The x-axis and y-axis correspond to N^1 and N^2, respectively. The colors represent the corresponding values.

In Figures 6 (a) and 6 (b), the areas within red dash boundaries correspond to the case $N^1 \geq N^\circ = 400$ and $N^2 \geq N^\circ = 400$, under which the non-deviation producer participation scenario happens at the producer location equilibrium under the optimal α^* in Stage I. The dot blue lines correspond to fixed total number of producers, i.e., 1 for $N^1 + N^2 = 600$, 2 for $N^1 + N^2 = 1000$, and 3 for $N^1 + N^2 = 1400$. Along these lines, moving towards the two endpoints implies a larger difference between N^1 and N^2, i.e., a larger $|N^1 - N^2|$. This further implies a worse match to the customers' preference (as $M^1 = M^2$). We use solid arrows to denote the directions that α^* and $G(\alpha^*)$ increase.

Based on Figures 6 (a) and 6 (b), we observe the following.

Remark 3 (Impact of Preference Mismatch on α^* and $G(\alpha^*)$)**.** *In Figures 6 (a) and 6 (b), if both producer populations N^1 and N^2 are large (i.e., inside the red dash boundary), a larger $|N^1 - N^2|$ (i.e., a worse match to the customers' preference) leads to a smaller α^* and a smaller $G(\alpha^*)$; otherwise, a larger $|N^1 - N^2|$ leads to a larger α^* and a smaller $G(\alpha^*)$.*

Specifically, when both producer populations are large (i.e., $N^1 \geq 400$ and $N^2 \geq 400$), producers never deviate. Under this case, the bottleneck of the platform's payoff is the minimum producer population, i.e., $\min\{N^1, N^2\}$. Hence, when the difference $|N^1 - N^2|$ is larger (i.e., the $\min\{N^1, N^2\}$ is smaller under the fixed values of $N^1 + N^2$ along the dot blue lines), the platform should set a smaller α^*, as imposing a high motivation is not beneficial due to the smaller $\min\{N^1, N^2\}$. Meanwhile, the platform gains a smaller optimal payoff $G(\alpha^*)$ due to the smaller $\min\{N^1, N^2\}$. On the other hand, when either of the producer populations is small (i.e., either $N^1 < 400$ or $N^2 < 400$), producers may deviate from the location preferred by more producers to the location preferred by less producers. Hence, when the difference $|N^1 - N^2|$ is larger (i.e., the mismatch between the producers' and customers' preferences is more severe), the platform should set a larger α^* to further motivate producer deviation (to match the customers' preferences). However, the platform's optimal payoff is smaller due to its profit share with the producers for motivating the deviation.

[14]When the customers are non-uniform distributed, the shape of the simulation figures will change, while the key insights will be the same.

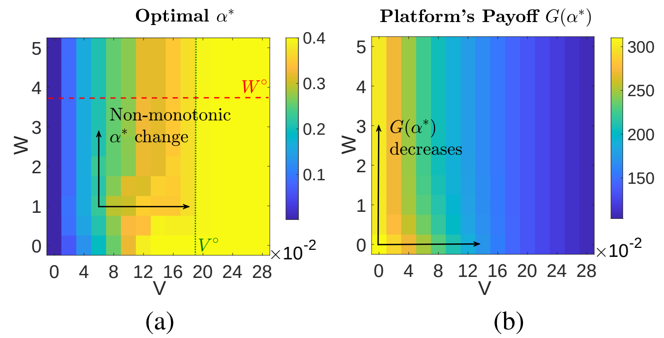


Fig. 7. An example with different V and W: (a) optimal α^* (with thresholds $W^\circ = 4$ and $V^\circ = 20$); (b) optimal platform's payoff $G(\alpha^*)$.

B. Impact of Opportunity and Deviation Costs

We aim to understand how the opportunity and deviation costs affect the optimal DSF and the platform's optimal payoff. We consider a location set $\mathcal{L} = \{0, 1, ..., 9\}$ with $r_l = 0.1 \times l, \forall l \in \mathcal{L}$, and $K = 1$. We set $N^l = 45 \times (2 + \cos(2\pi r_l))$ and $M^l = 45 \times (2 + \sin(2\pi r_l))$, where the producers' and customers' preferences are mismatched. We evaluate the results under different V and W in Figure 7, including the optimal α^* in Figure 7 (a) and the optimal platform's payoff $G(\alpha^*)$ in Figure 7 (b). The x-axis and y-axis correspond to V and W, respectively. The colors represent the corresponding values.

Based on Figures 7 (a) and 7 (b), we observe the following.

Remark 4 (Impact of Costs on α^* and $G(\alpha^*)$)**.** *In Figures 7 (a) and 7 (b), as W increases, α^* non-monotonically changes when $W < W^\circ$ and remains unchanged when $W \geq W^\circ$. As V increases, α^* non-monotonically changes when $V < V^\circ$ and remains unchanged when $V \geq V^\circ$. The platform's optimal payoff $G(\alpha^*)$ decreases in both W and V.*

In terms of α^*, when either the deviation cost or the opportunity cost is small, α^* non-monotonically changes with W or V, respectively. When either of them is large, α^* is independent of V and W, which is consistent with Proposition 4. In terms of $G(\alpha^*)$, as the costs V and W increase, the platform's optimal payoff decreases, since the platform imposes an increasing effort on compensating the costs of the producers, so as to provide them sufficient participation and deviation incentives.

VII. Case Study with Real Data from Twitch

In Section VI, we perform simulations to show the optimal DSF and the platform's payoff under different system settings. In practice, the system parameters (such as the producers' preference distribution) are unknown to the platform, which makes it challenging to determine the optimal DSF. In this section, we aim to demonstrate how to compute the platform's optimal DSF with only the producers' actual behaviors (instead of their inherent preferences) using data collected from Twitch. The study also suggests that under the collected data and our model

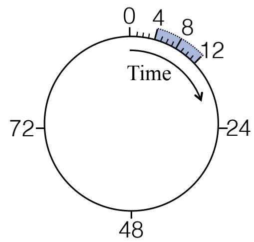


Fig. 8. A circular service attribute model with time of day as the service attribute. This model contains $24 \times 4 = 96$ locations, each corresponding to a period of 15 minutes in a day. Each streamer streams for an consecutive period of two hours (i.e., his or her service covers $2 \times 4 = 8$ locations).

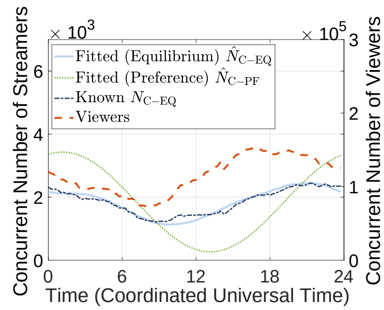


Fig. 9. A fitting example under the parameters (c_1, c_2, c_3) minimizing RMSE. The streamers deviate from $\hat{\boldsymbol{N}}_{\text{C-PF}}$ to $\hat{\boldsymbol{N}}_{\text{C-EQ}}$ to better match the viewers' preferences.

settings, Twitch may improve its payoff by reducing the DSF (comparing with its current practice).

This case study is based on the real-world dataset that we collected from Twitch every 15 minutes from Nov. 05 to Nov. 20, 2017. The information contains user_id, game_id, streamer_type, viewer_count, started_at, and language.

We first discuss the system setting. We then explain how to map from the producers' choices of the time of day (of the stream) to their preferences. Note that the producers' choices are observable from the collected data, while their preferences are not directly observable. Finally, we derive the optimal DSF.

A. System Setting

1) Service Attribute: We use the data of game *League of Legends*. We consider the service attribute of the time of day, where the service attribute setting is shown in Figure 8. There are $24 \times 4 = 96$ locations (which corresponds to the time of day at which the data are collected), denoted by $\mathcal{L} = \{0, 1, ..., 95\}$. We assume that each streamer streams for an consecutive period of 2 hours, i.e., $K = 2 \times 4 = 8$, which is the average stream time of the streamers in *League of Legends*. For example, as in Figure 8, when a streamer selects 1am (location 4), he will continue to serve until 3am (location 12), represented by the shaded area, i.e., $\mathcal{L}_4 = \{4, 5, ..., 11\}$.

2) Mass of Producers: Producers are distributed along the circle according to their preferences, denoted by $\boldsymbol{N}_{\text{PF}} = \{N^0_{\text{PF}}, N^1_{\text{PF}}, ..., N^{95}_{\text{PF}}\}$ (i.e., how many producers prefer to start at each location), which cannot be observed directly from the dataset. These producers make decisions on their actual start over the locations. Under the producers' decisions, let $\boldsymbol{N}_{\text{C-EQ}} = \{N^0_{\text{C-EQ}}, N^1_{\text{C-EQ}}, ..., N^{95}_{\text{C-EQ}}\}$ denote the cumulative number of producers serving customers at each location (considering the producers' two-hour streaming), which can be observed from the dataset. We are going to estimate the preference $\boldsymbol{N}_{\text{PF}}$ based on the observed $\boldsymbol{N}_{\text{C-EQ}}$, and then compute the optimal DSF.

3) Mass of Customers: Customers are distributed along the circle according to their preferences, where the number of customers at each location is derived from the dataset. These customers donate to the producers who select their preferring locations. The donation amount is given by the donation function defined in (4) from the dataset in [13].

4) Current DSF: On Twitch, viewers purchase 100 bits (i.e., a virtual currency on Twitch) with \$1.4, while streamers can exchange 100 bits with \$1. Hence, the current DSF $\alpha = 1/1.4 \approx 0.71$. We are going to derive the optimal DSF based on our model, and compare the result with this current practice of Twitch.

B. Mapping from Producer Distribution to Their Preference

Before deriving the optimal DSF, we need to first estimate the producers' inherent preference locations $\boldsymbol{N}_{\text{PF}}$ based on the observed cumulative producer distribution $\boldsymbol{N}_{\text{C-EQ}}$.

Let $\hat{\boldsymbol{N}}_{\text{PF}} = \{\hat{N}^0_{\text{PF}}, \hat{N}^1_{\text{PF}}, ..., \hat{N}^{95}_{\text{PF}}\}$ denote the estimated preferences. Under the estimated preference, let $\hat{\boldsymbol{N}}_{\text{C-EQ}} = \{\hat{N}^0_{\text{C-EQ}}, \hat{N}^1_{\text{C-EQ}}, ..., \hat{N}^{95}_{\text{C-EQ}}\}$ denote the fitted cumulative number of producers serving customers at each location if all the producers follow the producer location equilibrium. We aim to guess the most proper $\hat{\boldsymbol{N}}_{\text{PF}}$ such that the root-mean-square-error (RMSE) between the observed $\boldsymbol{N}_{\text{C-EQ}}$ and the fitted $\hat{\boldsymbol{N}}_{\text{C-EQ}}$ is minimized.

We assume that the producer actual preference $\boldsymbol{N}_{\text{PF}}$ follows a sine function, i.e., the number of producers at a location $l \in \mathcal{L}$ is $S(l) = c_1 \times \sin(2\pi l + c_2) + c_3$. Although we cannot observe the shape of $\boldsymbol{N}_{\text{PF}}$ from the dataset, we can verify through simulation that a sine function preference $\boldsymbol{N}_{\text{PF}}$ is likely to output a sine function equilibrium $\boldsymbol{N}_{\text{C-EQ}}$, where the $\boldsymbol{N}_{\text{C-EQ}}$ is observed to be a sine function from the dataset as its lag plot[15] follows a circular shape [26].

To estimate $\boldsymbol{N}_{\text{PF}}$, the key is to estimate the set of parameters (c_1, c_2, c_3). We choose the parameters that minimize the RMSE between the known cumulative distribution $\boldsymbol{N}_{\text{C-EQ}}$ and fitted cumulative distribution $\hat{\boldsymbol{N}}_{\text{C-EQ}}$:

$$\text{RMSE} = \sqrt{\sum_{l \in \mathcal{L}} \left(N^l_{\text{C-EQ}} - \hat{N}^l_{\text{C-EQ}}\right)^2 / L}. \tag{28}$$

Figure 9 shows the fitting result (under the parameters (c_1, c_2, c_3) that lead to the minimum RMSE).[16] The "Known $\boldsymbol{N}_{\text{C-EQ}}$" is the observed actual cumulative producer distribution. The "Fitted (Equilibrium) $\hat{\boldsymbol{N}}_{\text{C-EQ}}$" is the fitted cumulative producer distribution in equilibrium under the estimated producer preference $\hat{\boldsymbol{N}}_{\text{PF}}$, and the "Fitted (Preference) $\hat{\boldsymbol{N}}_{\text{C-PF}}$" is the fitted cumulative producer distribution when all the producers

[15]Lag plot shows the periodical feature of samples $\{X_1, X_2, \cdots, X_N\}$ by plotting points $(X_i - 1, X_i)$ for all $i = 1, 2, \cdots, N$.

[16]The probability density function of the fitting residuals roughly follows a normal distribution with zero mean. This implies that the residuals are random, which suggests that our fitting model works well [26].

TABLE II
OPTIMAL DSF.

$V \backslash W$	2	10	20	100	200	1000
0.2	0.08	0.09	0.08	0.07	0.07	0.07
1	0.32	0.31	0.30	0.28	0.28	0.28
2	0.39	**0.38**	0.38	0.40	0.40	0.40
4	0.40	0.40	0.40	0.40	0.40	0.40

TABLE III
PLATFORM'S PAYOFF INCREASE UNDER THE OPTIMAL DSF IN TABLE II, COMPARING WITH THAT UNDER 0.71 SET BY TWITCH.

$V \backslash W$	2	10	20	100	200	1000
0.2	215%	205%	206%	217%	217%	218%
1	132%	120%	120%	132%	132%	132%
2	55%	**54%**	53%	43%	43%	43%
4	41%	41%	41%	41%	41%	41%

choose their preferring locations under the estimated producer preference $\hat{\boldsymbol{N}}_{\text{PF}}$. We can see that the streamers deviate from "Fitted (Preference) $\hat{\boldsymbol{N}}_{\text{C-PF}}$" to "Fitted (Equilibrium) $\hat{\boldsymbol{N}}_{\text{C-EQ}}$" to better match the viewers' preferences (the red dash line).

C. Deriving Optimal Donation-Split-Fraction

Based on the estimated $\hat{\boldsymbol{N}}_{\text{PF}}$, we derive the platform's optimal DSF based on Theorem 1 and the searching method mentioned in Section V-B3. Table II shows the optimal DSF values under different values of V and W (as we do not know the actual values). Table III shows how much the platform's payoff increases under the optimal DSF in Table II, comparing with that under 0.71 currently implemented by Twitch.

In both tables, we use bold fonts to represent the values corresponding to $V = 2$ and $W = 10$, because this combination of parameters leads to the minimum RMSE (defined in (28)) over all the possible values (hence is most likely to be the one in reality). The key observations are as follows.

Remark 5 (Optimal α^*)**.** *In Table II, the bold text suggests that the optimal DSF should be* $\alpha^* = 0.38$*. Furthermore, all the optimal DSF values under various V and W values are significantly smaller than* 0.71 *currently adopted by Twitch.*

Remark 6 (Platform's Payoff Increase)**.** *In Table III, the bold text suggests that the platform's payoff can be increased by* 54% *under the optimal DSF choice of* $\alpha^* = 0.38$*, comparing with that under* 0.71 *currently adopted by Twitch.*

We would like to clarify that our case study results are derived based on our abstract model and the limited amount of data that we collected. On real-world live streaming platforms, however, many other factors have to be taken into account, such as the cultural and psychological factors of the streamers, the social interactions among them, as well as the reputation and future marketing plan of the platform.

VIII. CONCLUSION

In this paper, we focused on the study of donation-based markets, which have been embraced by various platforms in practice but have not been well-studied in the existing literature. In our study, we characterized the producers' behaviors in terms of their participation and service attribute decisions and the platform's behavior in terms of its DSF choice. Our main findings are summarized as follows.

- First, regarding the producer service attribute selection, if a platform aims at achieving a better match between the streamers' service attribute choices and the viewers' preferences, then it will increase DSF. However, if the platform aims at enhancing the total donation from the viewers, then increasing DSF may not be helpful.
- Second, regarding the platform's optimal DSF, it does not necessarily increase or decrease with the mismatch between the producers' and customers' preferences. Meanwhile, it changes non-monotonically with the producers' opportunity and deviation costs. Hence, there may not exist a particular rule for updating the optimal DSF in response to the changes of the producers' and customers' preferences as well as the costs of the producers.
- Third, the platform's optimal payoff decreases with the mismatch between the producers' and customers' preferences. It also decreases with the opportunity and deviation costs of the producers. Hence, in practice, if the platform can reduce the mismatch between the producers' and customers' preferences (or the opportunity and deviation costs of the producers) through particular governance methods, then the platform's payoff can be improved.
- Fourth, our case study on the dataset from Twitch demonstrates the approach of computing the platform's optimal DSF with only the producers' actual behaviors (instead of their inherent preferences) for practical systems.

This work serves as a first step towards understanding the operation and revenue management in donation-based markets. Our theoretical results were conducted based on the assumption that the platform and producers are rational. Meanwhile, as we focused on the donation-based markets, we did not capture other factors of more generic online platforms, such as the social interactions among producers and other monetization methods. There are several directions to extend this work. First, it is interesting to incorporate the idea of bounded human rationality (from behavioral economics) into the systems. For example, producers may regard the donation sharing with the platform as a loss. Those producers may be loss aversion so that they want to avoid the losses more than seeking the equivalent gains. Second, we can take into account the social interactions among the producers, e.g., understanding how the behavior of a producer affects those of others. Third, our analysis of the two-stage game (in Sections IV and V) is under complete information. To address the incomplete information scenario, it would be interesting to propose a distributed algorithm for Stage II game in order to enable the producers to achieve the producer location equilibrium in a decentralized manner. It is also beneficial to propose an online algorithm for Stage I game to enable the platform to determine the optimal DSF in an online fashion.

REFERENCES

[1] M. Tang and J. Huang, "How to earn money in live streaming platforms?—A study of donation-based markets," in *Proc. IEEE INFOCOM*, Paris, France, Apr. 2019.

[2] M. Siutila, “The gamification of gaming streams,” in *Proc. GamiFIN*, Pori, Finland, May 2018.

[3] Streamlabs, “Live streaming Q4’18 report — First decline for Fortnite, $141M in tips, YouTube and Facebook Gaming growing,” https://blog.streamlabs.com/live-streaming-q418-report-fortnite-sees-first-decline-in-mau-mobile-users-up-266-mixer-94f281e5b31f, Accessed 15 Sep. 2020.

[4] M. Iqbal, “Twitch revenue and usage statistics (2020),” https://www.businessofapps.com/data/twitch-statistics/, Accessed 15 Sep. 2020.

[5] M. R. Johnson and J. Woodcock, ““and today’s top donator is": How live streamers on Twitch.tv monetize and gamify their broadcasts,” *Social Media+ Society*, vol. 5, no. 4, Nov. 2019.

[6] C. D’Aspremont, J. J. Gabszewicz, and J.-F. Thisse, “On Hotelling’s “stability in competition",” *Econometrica*, vol. 47, no. 5, pp. 1145–1150, Sep. 1979.

[7] H. A. Eiselt and V. Marianov, *Foundations of location analysis*. Springer Science & Business Media, 2011.

[8] M. A. Khan and Y. Sun, “Non-cooperative games with many players,” *Handbook of game theory with economic applications*, vol. 3, pp. 1761–1808, 2002.

[9] W. H. Sandholm, “Potential games with continuous player sets,” *Journal of Economic theory*, vol. 97, no. 1, pp. 81–108, Mar. 2001.

[10] W. H. Sandholm, “Large population potential games,” *Journal of Economic Theory*, vol. 144, no. 4, pp. 1710–1725, Jul. 2009.

[11] M. Hu, M. Zhang, and Y. Wang, “Why do audiences choose to keep watching on live video streaming platforms? An explanation of dual identification framework,” *Comput. Human Behav.*, vol. 75, pp. 594–606, Oct. 2017.

[12] Z. Hilvert-Bruce, J. T. Neill, M. Sjöblom, and J. Hamari, “Social motivations of live-streaming viewer engagement on Twitch,” *Computers in Human Behavior*, vol. 84, pp. 58–67, Jul. 2018.

[13] Z. Zhu, Z. Yang, and Y. Dai, “Understanding the gift-sending interaction on live-streaming video websites,” in *Proc. Int’l Conf. on Social Computing and Social Media (SCSM)*, BC, Canada, May 2017.

[14] X. Liu, “Live streaming in china: Boom market, business model and risk regulation,” *Journal of Residuals Science & Technology*, vol. 13, no. 8, pp. 284.1–284.7, 2016.

[15] P. Tang, Y. Zeng, and S. Zuo, “Fans economy and all-pay auctions with proportional allocations.” in *Proc. AAAI Conf. on Artificial Intelligence*, San Francisco, CA, Feb 2017.

[16] A. L. Jia, S. Shen, D. H. Epema, and A. Iosup, “When game becomes life: The creators and spectators of online game replays and live streaming,” *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, vol. 12, no. 4, p. 47, Aug. 2016.

[17] H. Hotelling, “Stability in competition,” *The Economic Journal*, vol. 39, no. 153, pp. 41–57, Mar. 1929.

[18] N. Economides, “Minimal and maximal product differentiation in Hotelling’s duopoly,” *Economics Letters*, vol. 21, no. 1, pp. 67–71, 1986.

[19] S. Brenner, “Hotelling games with three, four, and more players,” *Journal of Regional Science*, vol. 45, no. 4, pp. 851–864, Oct. 2005.

[20] S. Behringer and L. Filistrucchi, “Hotelling competition and political differentiation with more than two newspapers,” *Information Economics and Policy*, vol. 30, pp. 36–49, Mar. 2015.

[21] D. Schmeidler, “Equilibrium points of nonatomic games,” *Journal of statistical Physics*, vol. 7, no. 4, pp. 295–300, Apr. 1973.

[22] S. C. Salop, “Monopolistic competition with outside goods,” *The Bell Journal of Economics*, vol. 10, no. 1, pp. 141–156, Spring 1979.

[23] J. Kyparisis, “On uniqueness of Kuhn-Tucker multipliers in nonlinear programming,” *Mathematical Programming*, vol. 32, no. 2, pp. 242–246, Jun. 1985.

[24] R. L. Burden and D. J. Faires, *Numerical analysis*. PWS Publishing Company, 1985.

[25] B. Beavis and I. Dobbs, *Optimisation and stability theory for economic analysis*. Cambridge university press, 1990.

[26] NIST/SEMATECH, “NIST/SEMATECH e-handbook of statistical methods,” https://www.itl.nist.gov/div898/handbook/, Accessed 15 Sep. 2020.

Ming Tang (S’16, M’18) is a postdoctoral research fellow in the Department of Electrical and Computer Engineering of the University of British Columbia, Vancouver, BC, Canada. She received her Ph.D. degree from the Department of Information Engineering, The Chinese University of Hong Kong, Hong Kong, China, in 2018. Her research interests include mobile networking and network economics.

Jianwei Huang (F’16) received the Ph.D. degree in ECE from Northwestern University in 2005, and worked as a Postdoc Research Associate in Princeton University during 2005-2007. From 2007 until 2018, he was on the faculty of Department of Information Engineering, The Chinese University of Hong Kong. Since 2019, he has been on the faulty at The Chinese University of Hong Kong, Shenzhen, where he is currently a Presidential Chair Professor and an Associate Dean of the School of Science and Engineering. He also serves as a Vice President of Shenzhen Institute of Artificial Intelligence and Robotics for Society. His research interests are in the area of network optimization, network economics, and network science, with applications in communication networks, energy networks, data markets, crowd intelligence, and related fields. He has co-authored 9 Best Paper Awards, including the 2011 IEEE Marconi Prize Paper Award in Wireless Communications. He has co-authored seven books, including the textbook on "Wireless Network Pricing." He is an IEEE Fellow, and was an IEEE ComSoc Distinguished Lecturer and a Clarivate Web of Science Highly Cited Researcher. He is the Editor-in-Chief of IEEE Transactions on Network Science and Engineering, and was the Associate Editor-in-Chief of IEEE Open Journal of the Communications Society.