Three Papers in Applied Microeconomic Theory

A Thesis in the Field of Applied Microeconomic Theory
for the Degree of Doctor of Philosophy in Economics

Xiaochang Lei
University of Glasgow
February 2025

Abstract

This thesis comprises three papers in Applied Microeconomic Theory. Chapter 1 is my job market paper titled *Repeated Signaling and Reputation*.

Chapters 2 and 3 include two papers published during my Ph.D. studies:

- **Chapter 2**: Lei, X. (2023). *Pro-rata vs User-centric in the Music Streaming Industry. Economics Letters*, 226, 111111.
- **Chapter 3**: Lei, X. (2023). *Optimal Queue to Minimize Waste. Mathematical Social Sciences*, 123, 87-94.

Declaration

I declare that, except where explicit reference is made to the contribution of others, this dissertation is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution.

Printed Name: Xiaochang Lei
Signature: Xiaochang Lei

Acknowledgments

I would like to express my deepest gratitude to my supervisors, Hervé Moulin and John Levy, for their invaluable guidance, encouragement, and support throughout the course of my Ph.D. Their expertise and insight have greatly shaped my research and intellectual growth.

Table of Contents

Introduction

This thesis presents three essays in Applied Microeconomic Theory, each exploring distinct yet interconnected themes within applied game theory, mechanism design and information design. These studies aim to deepen the understanding of (dynamic) interactions and decision-making processes in economically relevant environments, such as influencer marketing, the music streaming industry, and queue management for waste minimization. Collectively, the three essays contribute to Applied Microeconomic Theory by modeling complex behaviors and equilibria under varying informational asymmetries and incentive structures, which are prevalent in modern digital and service-based economies.

The first chapter investigates repeated signaling and reputation, where an influencer repeatedly interacts with myopic customers, a setting common in social media marketing. Reputation models, particularly those that integrate behavioral types, help explain how reputation evolves and influences customer actions. Traditional literature on reputation models, such as Kreps and Wilson (1982) and Milgrom and Roberts (1982), highlights how reputation can deter entry through long-term strategic incentives. However, existing models often assume customers can fully observe history, which is unrealistic in real-world influencer marketing. This chapter extends previous work by assuming that customers are limited to a single-period memory. Consequently, the optimal signaling strategy becomes unique, with the influencer using a mixed strategy to manage reputation while inducing purchases, contrasting with models that allow for pure strategy equilibria in repeated games.

The second chapter addresses remuneration rules in the music streaming industry, comparing the efficiency and fairness of pro-rata (P rule) versus user-centric (U rule) payment models. Music streaming platforms like Spotify and Deezer face ongoing debates regarding equitable revenue distribution, as demonstrated by empirical studies (Muikku, 2017; Hesmondhalgh, 2021). While previous studies often treat the streaming matrix as exogenous, this chapter contributes by developing an endogenous model where artists strategically adjust song quality to influence streaming frequency. Using this model, the study shows that the P rule outperforms the U rule in both efficiency and fairness when superstar artists' marginal costs are lower. This finding challenges current policy discussions that favor the U rule as inherently more equitable and highlights the strategic responses of artists, which can shift the fairness and efficiency of remuneration outcomes.

The third chapter explores queue management and waste minimization in service systems, drawing on theories from queueing theory and social choice. Optimizing queues, especially in contexts with significant waste potential (e.g., healthcare and food services), requires balancing fairness of queuing service and resource allocation to minimize unused resources. This essay contributes to queueing literature by proposing an optimal policy, which combines multiple disciplines in allocation, that reduces waste while preserving service fairness. Building on classic works in queue theory and social fairness (Naor, 1969), this optimization model offers insights for both private and public service providers seeking to minimize allocation waste.

Methodology

The methodologies employed across these chapters draw from game theory and microeconomic modeling. In Chapter 1, a repeated signaling model is developed under a Markov Decision Process framework to capture the influencer's strategic behavior in a limited-memory environment. The model is solved using elimination of dominated strategy to identify the unique mixed-strategy equilibrium, supported by analytical derivations of customer belief updating and posterior reputation calculations. Chapter 2 employs an endogenous modeling approach to analyze artist behavior under different revenue-sharing rules, leveraging a comparative static analysis to quantify equilibrium qualities and royalty distributions across remuneration models. In Chapter 3, a queueing optimization framework is utilized to identify policies that prioritize service order based on fairness and mathematically prove the theoretical model's capacity for waste minimization.

This research builds upon foundational theories in reputation modeling, revenue-sharing rules, and queue management while addressing gaps in the literature by incorporating real-world complexities such as memory constraints in reputation models and strategic quality adjustments by agents. The theoretical advancements contribute to a nuanced understanding of strategic decision-making and optimal policy design in scenarios marked by information asymmetry and competitive incentives.

In conclusion, this thesis offers a comprehensive examination of strategic behaviors in microeconomic settings with practical implications. By integrating behavioral insights with rigorous modeling, the three essays extend Applied Microeconomic Theory's scope, providing both theoretical advancements and policy-relevant recommendations for sectors grappling with asymmetric information, reputation dynamics, and resource allocation challenges.

Chapter 1

Repeated Signaling and Reputation

Preface

The first chapter lays the theoretical foundation of this thesis by exploring the dynamics of reputation in repeated information design interactions. Reputation, particularly in markets with asymmetric information, is a critical mechanism through which individuals and firms build trust and influence decisions. This chapter examines how a long-lived influencer's reputation evolves when interacting with a series of myopic customers who observe only limited historical information. By introducing a bounded-memory assumption, this model reflects the practical limitations of real-world settings like influencer marketing, where customers often lack perfect recall of past interactions.

This analysis establishes key insights that resonate throughout real-world decision making process: namely, that agents make intertemporal trade-offs between immediate rewards and the potential for reputation-based gains in the future. The chapter's findings highlight how the influencer's signaling strategy, shaped by these trade-offs, becomes strictly mixed in equilibrium, contrasting with pure equilibria found in standard reputation models. The themes of information asymmetry, bounded rationality, and strategic adaptation presented here provide a conceptual bridge to the subsequent chapters, where different forms of asymmetric information and competitive strategies are explored in applied settings.

Abstract

The Markov decision process (MDP) framework is widely applied in the behavioral-type-based reputation literature. A pure strategy stationary equilibrium can exist, often resulting from the linear transition dynamics linked to players' actions. In this paper, we examine a reputation model with an information design stage game, where the transition to the receiver's action is non-linear. A relevant example is influencer marketing on social media, where information asymmetry exists regarding both the product's quality and the influencer's type. The customer's decision-making process depends on the influencer's signal and reputation. We demonstrate that, with myopic customers, neither a separating nor a pooling equilibrium exists in the steady state. Instead, the optimal stationary Markov persuasion strategy is unique and strictly mixed. Finally, we capture the stochastic reputation dynamics in the steady state.

1 Introduction

A reputation model with behavioral types is a form of a repeated game with incomplete information, where Player 1's type is private, and the history of past play reveals information about his type, thereby influencing Player 2's beliefs. Typically, Player 1 faces a trade-off between choosing an action that yields a higher current payoff but harms future reputation—leading to a lower future value—and choosing an action that results in a lower current payoff but improves reputation, potentially generating higher future value.

We consider a new application of the reputation model in which the stage game is an information design game. For instance, in influencer marketing, customers often struggle to distinguish between low-quality and high-quality products and rely on influencers for information regarding the product's quality. The influencer is rewarded when a product is purchased, creating an incentive to mislead customers if misreporting the product's quality is profitable. However, the influencer's credibility diminishes if customers discover prior instances of dishonesty. Thus, the influencer faces a trade-off between achieving a higher current payoff through less accurate information, which lowers future value, and accepting a lower current payoff with more accurate reporting, which increases future value. Reputation effects also arise in other contexts involving information asymmetry, such as stock recommendations (Morgan & Stocken 2003), consulting services (Glückler & Armbrüster 2003), and political signaling (Honryo 2018).

We use the influencer marketing context as an example, where the information design literature, starting from (Kamenica & Gentzkow 2011), applies widely. The key difference between an information design game and a signaling game is that the influencer needs to commit to a signaling plan before the realization of states rather than choosing a signaling strategy after observing a realized state. We choose the information design game because it focuses on how the sender maximizes utility through information manipulation, which is central to marketing literature, and places less emphasis on the sender's incentive compatibility constraints compared to a signaling game by giving the sender commitment power.

Our aim is to capture the long-term interaction between a long-lived influencer and short-lived, myopic customers with limited capability to observe the history of past interactions. We are interested in addressing the following questions:

1. In the steady state, how do different histories affect the influencer's reputation, and what are the dynamics of reputation evolution?

2. Can a separating or pooling equilibrium exist in the steady state? If yes, what are the conditions for such an equilibrium? If not, what is the stationary Markov perfect equilibrium, and is it unique?

The first result is that a matched history of high quality and high signal does not improve the influencer's reputation. This may seem counterintuitive, as one might expect that a matched history would boost the influencer's reputation. However, because customers know that the influencer must send a high signal when the quality is high, this action is fully anticipated and, therefore, does not contribute to increasing reputation. Reputation improves only when an outcome is somewhat unexpected. For the same reason, a matched history of low quality and low signal can actually benefit the influencer's reputation, as the influencer has an incentive to send a high signal to encourage purchasing.

The second result is that there is a unique stationary Markov perfect equilibrium in a mixed strategy. It is characterized by a partially separating signaling strategy, where the influencer sends a high signal when the quality is high and a strictly mixed signal when the quality is low. The mixed strategy is essential for achieving the optimal outcome due to the non-linear transition in the customer's action. The inefficiency of a pure signaling strategy is evident: when the quality is low, any pure signaling strategy cannot induce purchasing, resulting in a zero instant reward. In contrast, a mixed strategy can induce purchasing and generate a positive expected payoff for the influencer, which is not a simple linear combination of the two zero rewards under pure strategies.

The uniqueness of the equilibrium arises from the assumption that customers are myopic and have limited memory of past interactions. The one-period memory of the customers guarantees that cheating has a very limited impact on future value, making cheating always beneficial for the influencer. This situation may change when customers have greater capabilities for observing longer histories or even perfect recall, which could allow cheating to have a longer-lasting negative effect on the influencer's reputation, making a separating equilibrium feasible. Nevertheless, the myopic assumption is suitable for the real world and shows an important result: all exogenous variables, such as the discount factor, the influencer's ex-post reward, the prior distributions of goods, and the influencer's prior reputation, do not affect the uniqueness of the stationary Markov perfect equilibrium. This is different from other reputation models where multiple equilibria exist and whose attainment depends on exogenous parameters.

The structure of the paper is as follows. In Section 2, we review the most relevant research in the literature on reputation and influencer marketing. In Section 3, we explain the setup of the stage signaling game and demonstrate the effect of a fixed reputation. In Section 4, we present the setup for the repeated version of the stage signaling game and derive the unique optimal stationary Markov signaling strategy for the influencer. In Section 5, we conclude with a brief discussion on the game with perfect recall.

2 Literature Review

Reputation literature: Mailath & Samuelson (2006) provides a comprehensive collection of models related to reputation in repeated games. Here, we focus on the models, whose solution concept is Markov perfect equilibrium (Maskin & Tirole 2001), where agents are rational and Bayesian. In general, papers are discussing the existence of three types of reputation equilibria:

1. Reputation-building equilibrium, in which a sufficiently patient firm always playing a strategy to build up its reputation (Höner 2002, Kreps et al. 1982, Kreps & Wilson 1982, Mailath & Samuelson 2001). Although these papers usually have different model settings, they usually assume that the customers have perfect recall of history play, which means that a bad behavior can persistently harm the long-run player's reputation, even with imperfect monitoring. In contrast, in our model, the customers are myopic and have only a one-period memory of history, meaning that cheating in the remote past is hidden. This assumption is more realistic in real-world business cases.

2. Cyclic equilibrium, in which the firm may choose to build reputation for a while and than exploit its reputation. Liu (2011) develops a model similar to ours. They also assume that the short-lived players have limited memory, but they also consider the customer's costly information acquisition strategy, which determines how many periods of past outcomes to observe. They prove the existence of a reputation cycle in the steady state, where the long-run player will build reputation when their reputation is low and cheat when their reputation is sufficiently high. Our result shares the same insight, and we also show that the influencer has a cyclic reputation in the steady state. The difference is that we do not consider the short-run player's information acquisition by imposing a strictly one-period memory restriction. This allows us to derive the explicit Markov signaling strategy and show the probability of reaching each reputation transition processes.

3. Levine (2021) discusses a new type of equilibrium called trap equilibrium, in which the firm's actions depend on the realization of a public history that is imperfectly monitored. A positive history signal induces high effort from the firm, while a negative history induces no effort. Thus, the firm's payoff depends on its luck in experiencing beneficial events. In our model, the influencer has more control over the history realization, since they can commit to a signaling plan before the quality is realized. Additionally, past actions of the firm are perfectly monitored, so the influencer does not need to worry about the possibility that their honest reporting will be ignored. In their model, the long-run player also has an exogenous type replacement process, which is not very useful in our setting, since the customers usually knows the influencer's identity so that the replacement will always be known.

Dynamic Information Design: There is a vast literature on information design, starting with Kamenica & Gentzkow (2011), and we focus on dynamic games here. Renault et al. (2017) and Ely (2017) both study a dynamic Bayesian persuasion game without cost. In their model, the

Markov transition of the state is exogenously given and publicly known. The information designer maximizes payoff by altering the receiver's autonomous belief updating. In our setting, the product's quality is independently and identically distributed (i.i.d.) across all periods. The Markov transition of history is endogenously determined by the long-run player's signaling strategy. There is also research on costly dynamic persuasion that restricts the sender's commitment power (Henry & Ottaviani 2019, Che et al. 2023). To the best of our knowledge, this paper is the first to combine dynamic information design with the reputation model.

Influencer Marketing: Here, we focus exclusively on economic theory papers. Fainmesser & Galeotti (2021) use an extensive-form one-shot game to discuss the equilibrium of the interaction between firms, influencers, and customers. They focus on the market as a whole and neglect the dynamics of interactions, which is the main focus of our work. Additionally, the influencer marketing discussion is more concerned with the influencer's trade-off between providing organic content, which generates more utility for customers but no income, and sponsored content, which generates less utility for customers but provides income. However, we aim to capture the repeated signaling and dynamic reputation evolution inherent in influencer marketing.

Mostagir & Siderius (2023) use a repeated game to model the dynamic interaction between firms, influencers, and customers. However, they focus more on market equilibrium, aiming to find market equilibria while neglecting the dynamic evolution of the influencer's reputation, which is one of the main goals of our work.

3 Stage Game

We characterize the influencer marketing story as a simple binary information design problem, where the influencer has state-independent preferences. There are two players. Player 1 is an influencer who tries to persuade Player 2 (the customer) to buy a good. The good has binary quality $q \in Q = \{H, L\}$. The common prior distribution on quality is $\mu = \{\mu(H), \mu(L)\}$, with $\sum_q \mu(q) = 1$ and $\mu(H) < \frac{1}{2}$. The customer has a binary action $a \in A = \{0, 1\}$, where action 1 means to buy. The customer's utility function is given by

$$u_2(0, H) = u_2(0, L) = 0, \quad u_2(1, H) = 1, \quad \text{and} \quad u_2(1, L) = -1,$$

which depends only on the action and the quality of the good. Let $u_1(a)$ denote the reward of the influencer, where $u_1(0) = 0$ and $u_1(1) = r > 0$. Here, we assume that the influencer's payoff depends only on the customer's action. The customer chooses an action a after observing a signal $s \in S = \{h, l\}$ sent by the influencer. We assume the customer chooses to buy when his expected payoff is non-negative.

Let $\alpha_1(s|q)$ denote the influencer's signaling strategy, which is the probability of sending signal s when quality q is realized. The key difference between an information design model and a signaling game is that, in the information design game, the influencer commits to a signaling strategy α_1 before quality q is realized.

The timing of the stage game is as follows:

1. The influencer commits to a signaling strategy α_1.
2. Nature selects the quality of the good q according to prior μ.
3. Signal s is realized based on q and α_1.
4. The customer observes the signal s, updates his quality belief, and chooses an action a.
5. Payoffs are realized.

3.1 Equilibrium

Given the influencer's strategy, the posterior belief of the customer is:

$$\mu(q|s) = \frac{\mu(q)\alpha_1(s|q)}{\sum\limits_{q'} \mu(q')\alpha_1(s|q')}.$$

The customer's strategy $\alpha_2(s)$ assigns a probability of choosing action 1 (to buy) after observing signal s. Using backward induction, in stage 2, the customer obtains non-negative expected utility when:

$$\mu(H|s) \geq \frac{1}{2}.$$

Therefore, the customer's best response is determined by his posterior belief under signal s, which is:

$$\alpha_2^*(s) = \begin{cases} 1, & \text{if } \mu(H|s) \geq \frac{1}{2}, \\ 0, & \text{if } \mu(H|s) < \frac{1}{2}. \end{cases}$$

Given the customer's best response, in stage 1, the influencer receives a positive payoff r only when the customer chooses $a = 1$ under signal s. Since the prior belief $\mu(H) < \frac{1}{2}$, the customer cannot always buy the product under both signals h and l. Therefore, to induce purchasing, the influencer must design information such that the customer would take different actions under different signals. This results in two conditions:

$$\mu(H|h) \geq \frac{1}{2}, \quad \mu(H|l) < \frac{1}{2}.$$

(Here, we discuss the case where the customer only chooses to buy under signal h. The other case is symmetric.)

As a result, the influencer's utility maximization problem can be expressed as:

$$\max_{\alpha_1} \sum_{q \in \{H,L\}} \mu(q)\alpha_1(h|q)r$$

subject to:

$$\mu(H|h) \geq \frac{1}{2}, \quad \mu(H|l) < \frac{1}{2}.$$

The resulting equilibrium signaling strategy is:

$$\alpha_1^*(h|H) = 1,$$

$$\alpha_1^*(h|L) = \frac{\mu(H)}{\mu(L)}.$$

This generates $2\mu(H)r$ expected payoff for the influencer. From here, we can see that the influencer maximizes his utility by maximizing the realization probability of signal h while ensuring that signal h is informative for high quality H.

3.2 Reputation Effect

Assuming that the influencer's private type can be either honest or rational, an honest influencer always truthfully reports the product's quality, while a rational influencer maximizes expected payoff and, hence, the realization probability of signal h. The customer has a prior belief γ about the influencer being honest, which we refer to as the influencer's reputation. We are interested in the case where the influencer is rational; thus, when we refer to the influencer without mentioning their type, we mean the rational influencer.

Let $\alpha_1(s|q,\gamma)$ denote the signaling strategy of the influencer. After observing signal $s \in \{h, l\}$, the customer's posterior belief about high quality is given by:

$$\mu_{\alpha_1}(H|h,\gamma) = \frac{\mu(H)(\gamma + (1-\gamma)\alpha_1(h|H,\gamma))}{\mu(H)(\gamma + (1-\gamma)\alpha_1(h|H,\gamma)) + \mu(L)(1-\gamma)\alpha_1(h|L,\gamma)}$$

$$\mu_{\alpha_1}(H|l,\gamma) = \frac{\mu(H)(1-\gamma)\alpha_1(l|H,\gamma)}{\mu(H)(1-\gamma)\alpha_1(l|H,\gamma) + \mu(L)(\gamma + (1-\gamma)\alpha_1(l|L,\gamma))}$$

For a similar reason as explained before, the influencer must induce different actions under different signals. This results in the following conditions:

$$\mu(H|h,\gamma) \geq \frac{1}{2}, \quad \mu(H|l,\gamma) < \frac{1}{2}.$$

As a result, the influencer's utility maximization problem is:

$$\max_{\alpha_1} \sum_{q \in Q} \mu(q)\alpha_1(h|q,\gamma)r$$

subject to:

$$\mu_{\alpha_1}(H|h,\gamma) \geq \frac{1}{2}, \quad \mu_{\alpha_1}(H|l,\gamma) < \frac{1}{2}.$$

The optimal signaling strategy is given by:

$$\alpha_1(h|H,\gamma) = 1, \quad \alpha_1(l|H,\gamma) = \min\left\{\frac{\mu(H)}{\mu(L)}\frac{1}{1-\gamma}, 1\right\}.$$

Again, the influencer aims to maximize the probability of sending signal h, while ensuring that signal h remains informative.

Comparing the equilibrium signaling strategy with the default case (where there is no reputation), we find that the influencer's expected payoff $\mu(H)\left(1 + \frac{1}{1-\gamma}\right)r$ increases with reputation γ until a reputation threshold $\gamma = 1 - \frac{\mu(H)}{\mu(L)}$. When reputation γ exceeds this threshold, the optimal signaling strategy $\alpha_1(l|H,\gamma)$ remains at 1, since it is a probability and cannot exceed

1. The highest expected payoff achievable is r (see Figure 1).

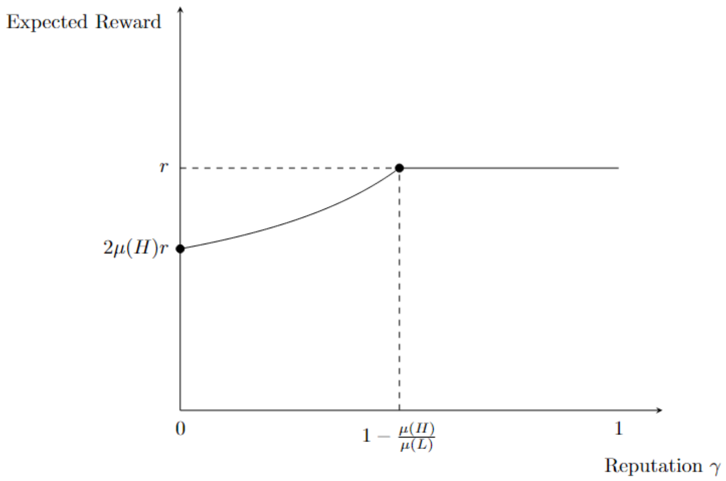

Figure 1: Expected Reward and Reputation

4 Repeated game

The static information design game with reputation is repeatedly played between a long-lived influencer and a short-lived customer. In each period, a new customer is born and leaves at the end of that period, while the influencer has an infinite lifetime and a persistent type. To allow past play to affect future reputation, we adopt a simple assumption that the short-lived customer is myopic and can only observe the last period's outcome. Let

$$Y = \{\{h, H\}, \{l, H\}, \{h, L\}, \{l, L\}\}$$

denote the set of all possible histories y. At period t, the customer t observes history $y_t = \{s_{t-1}, q_{t-1}\}$, where s_{t-1} and q_{t-1} denote the last period's signal and quality. All players share a common prior on the honest type γ_0. We assume $\gamma_0 < 1 - \frac{\mu(H)}{\mu(L)}$ so that the customers won't buy if the influencer sends the pure signal h. At period t, first, the public history y_t is realized. Then, the influencer commits to a signaling strategy α_{1t}. The customer observes signal s_t and history y_t, then chooses an action a_t.

Timing of the supergame:

1. In each period t, the public history y_t is realized.
2. A new customer t is born with prior belief γ_0 and μ.
3. After observing history y_t, both players update the influencer's reputation to γ_{y_t}.
4. The influencer commits to signaling strategy α_{1t}.
5. A new product is independently realized with quality q_t according to μ.
6. Signal s_t is realized based on α_{1t} and q_t.

7. After observing the signal, the customer chooses an action a_t.

8. Payoffs are realized, and the customer leaves.

9. The history then changes to $y_{t+1} = \{s_t, q_t\}$ for the next period.

In this game, history $y_t \in Y$ is the state. The state transition from y_t to y_{t+1} depends on the realization of q_t in period t and the signal s_t sent by the influencer in period t.

The Markov decision process (MDP) can be defined as a tuple (Y, S, A, P, u, δ):

1. A state space
$$Y = \{\{h, H\}, \{l, H\}, \{h, L\}, \{l, L\}\}$$
.

2. A finite space of signals $S = \{h, l\}$ and a finite space of actions $A = \{0, 1\}$.

3. A 4×4 transition matrix P, where each element represents a transition probability from state
$$y_t = \{s_{t-1}, q_{t-1}\} \text{ to state } y_{t+1} = \{s_t, q_t\},$$
given by
$$p(y_{t+1}|y_t) = \mu(q_t)\alpha_{1t}(s_t|q_t).$$

4. Reward functions:
$$u_1(a) = \begin{cases} r & \text{if } a = 1, \\ 0 & \text{otherwise} \end{cases}, \quad u_2(a, q) = \begin{cases} 1 & \text{if } q = H, a = 1, \\ -1 & \text{if } q = L, a = 1, \\ 0 & \text{otherwise} \end{cases}$$

5. A discount factor $\delta \in (0, 1)$ for the influencer to trade off between immediate and future rewards.

Table 1 collects notations and corresponding meanings. The key difference between the repeated information design MDP and other MDPs with simultaneous-move stage games (such as the chain-store entrance game or firm effort game) is that the information design MDP has nonlinear transitions into player 2's actions. This means that the repeated information design MDP is different from the standard MDP with linear transitions, as described in Blackwell (1965), where the optimal stationary Markov policy is pure. In this paper, we will show that, in the repeated information design MDP, the optimal strategy must be strictly mixed, which is necessary to induce purchasing when quality is low.

4.1 Markov Perfect Equilibrium

We are interested in the long-term interaction between the influencer and the customer. Thus, we do not need to specify the initial periods and can focus solely on the steady state of the

Variable	Meaning
S	Signal space, $S = \{h, l\}$
A	Action space, $A = \{0, 1\}$
Q	Quality space, $Q = \{H, L\}$
Y	History space, $Y = \{\{h, H\}, \{l, H\}, \{h, L\}, \{l, L\}\}$
s_t	Signal sent by the influencer at period t
a_t	Action chosen by the customer at period t
q_t	Current quality of the product at period t
y_t	History observed at period t, $y_t = \{s_{t-1}, q_{t-1}\}$
$p(y_{t+1}\|y_t)$	Transition probability from state y_t to y_{t+1}
r	Influencer's reward when the customer buys the product
μ	Prior distribution of product quality
δ	Discount factor for the influencer, $\delta \in (0, 1)$
γ_0	Prior reputation for the influencer, $\gamma_0 \in (0, 1)$
$u_1(a_t)$	Reward function for the influencer
$u_2(a_t, q_t)$	Reward function for the customer

Table 1: List of Variables and Their Meanings

Markov decision process. For any history y, a stationary strategy $\alpha_1(y)$ of the influencer, denoted as $\{\alpha_1(s|q, y)\}_{s\in S, q\in Q}$, is a plan for sending signals, where each element maps the current quality q and state y to a probability of sending signal s. A strategy of the customer, denoted as $\alpha_2(s, y)$, maps the history y and the observed signal s to a probability of choosing action 1.

The customer's best response: For any observed history y, the customer has the following two beliefs:

1. With probability γ_0, the influencer is honest and always reports the product's true quality. Thus, the distribution π_0 of states is given by:

$$\begin{aligned} \pi_0(\{h, H\}) = \mu(H) \qquad \pi_0(\{h, L\}) = 0 \\ \pi_0(\{l, H\}) = 0 \qquad \pi_0(\{l, L\}) = \mu(L) \end{aligned}$$

2. With probability $1 - \gamma_0$, the influencer is rational. Let π^{α_1} denote the stationary distribution of states induced by the rational influencer's stationary signaling strategy α_1.

From these two distributions over history, the customer updates the influencer's reputation to $\gamma_y^{\alpha_1}$ after observing history y:

$$\gamma_y^{\alpha_1} = \frac{\gamma_0 \pi_0(y)}{\gamma_0 \pi_0(y) + (1 - \gamma_0)\pi^{\alpha_1}(y)}$$

Then, after receiving a signal s, the customer chooses an action based on their updated quality belief:

$$\mu^{\alpha_1}(H|h,y) = \frac{\mu(H)(\gamma_y^{\alpha_1} + (1-\gamma_y^{\alpha_1})\alpha_1(h|H,y))}{\mu(H)(\gamma_y^{\alpha_1} + (1-\gamma_y^{\alpha_1})\alpha_1(h|H,y)) + \mu(L)(1-\gamma_y^{\alpha_1})\alpha_1(h|L,y)}$$

$$\mu^{\alpha_1}(H|l,y) = \frac{\mu(H)(1-\gamma_y^{\alpha_1})\alpha_1(l|H,y)}{\mu(H)(1-\gamma_y^{\alpha_1})\alpha_1(l|H,y) + \mu(L)(\gamma_y^{\alpha_1} + (1-\gamma_y^{\alpha_1})\alpha_1(l|L,y))}$$

The customer's strategy $\alpha_2(s,y)$ is a best response to α_1:

$$\alpha_2(s,y) = \begin{cases} 1 & \text{if } \mu^{\alpha_1}(H|s,y) \geq \frac{1}{2} \\ 0 & \text{otherwise} \end{cases}$$

The influencer's optimal signaling strategy: Let $\mathbf{P}^{\alpha_1}$ denote the transition matrix induced by α_1. The corresponding stationary distribution π^{α_1} of states under strategy α_1 can be computed from the system of equations:

$$\pi^{\alpha_1} = \pi^{\alpha_1}\mathbf{P}^{\alpha_1}$$
$$1 = \pi^{\alpha_1}\mathbf{e}$$

Given the current state $y \in Y$, let V_y denote the influencer's value function. Under signal s, the influencer's expected payoff at the current period is $\alpha_2(s,y)r$. Thus, the Bellman equation is: for all $y \in Y$,

$$V_y = \max_{\alpha_1} \sum_{q' \in Q} \sum_{s' \in S} \mu(q')\alpha_1(s'|q',y)\left[\alpha_2(s',y)r + \delta V(\{s',q'\})\right]$$

4.2 Stationary Equilibrium

We first eliminate dominated strategies to simplify the analysis of the equilibrium.

Lemma 1. *The influencer must send the signal h when the current quality is H, which implies that $\forall y \in Y$, $\alpha_1^*(h|H,y) = 1$.*

Proof. The reasoning is as follows: when high quality is realized,

1. Sending signal l decreases current expected payoff: For all y, the influencer's expected payoff in the current period is weakly increasing in the realization probability of signal h and thus $\alpha_1(h|H,y)$. In the current period, the customer either buys under signal h or does not buy under signal h. If the customer does not buy under signal h, then it has no impact. If the customer buys, the two constraints $\mu^{\alpha_1}(H|h,y) \geq \frac{1}{2}$ and $\mu^{\alpha_1}(H|l,y) < \frac{1}{2}$ only restrict the probability $\alpha_1(h|L,y)$ of sending signal h under low quality, since a very high cheating probability will make signal h uninformative. However, the two constraints do not restrict $\alpha_1(h|H,y)$, and the expected payoff in the current period is increasing in $\alpha_1(h|H,y)$.

2. Sending signal l decreases future value: If $\alpha_1(l|H,y) > 0$, there is a probability that the state y will transition to $\{l,H\}$ in the next period. The updated history $\{l,H\}$ contains a

mismatch between the past signal and the quality, which clearly indicates that the influencer is rational, generating a lower future value for the influencer.

Thus, any strategy with positive probability of sending signal l is dominated by the strategy of sending pure signal h, which implies that the optimal stationary Markov signaling strategy must be $\alpha_1^*(h|H,y) = 1$ for all history $y \in Y$. $\square$

Also, the strategy $\forall y, \alpha_1^*(h|H,y) = 1$ results in the following reputation updating process:

Proposition 1. *Under the optimal stationary Markov signaling strategy, the history $\{h,H\}$ does not affect the influencer's reputation. The history $\{h,L\}$ decreases the influencer's reputation to zero, while the history $\{l,L\}$ weakly increases the influencer's reputation.*

Proof. 1. For history $\{h,H\}$, its distribution is fixed as $\pi^{\alpha_1}(\{h,H\}) = \mu(H)$ under the optimal signaling strategy, since the rational influencer will honestly report the true quality. Therefore, in this case, the posterior reputation is:

$$\gamma^{\alpha_1}_{\{h,H\}} = \frac{\gamma_0\pi_0(\{h,H\})}{\gamma_0\pi_0(\{h,H\}) + (1-\gamma_0)\pi^{\alpha_1}(\{h,H\})} = \gamma_0.$$

2. The distribution of the history $\{h,L\}$ depends on the signaling strategy $\alpha_1(h|L,y)$. In this case, the posterior reputation is 0, due to the mismatch between the past signal and past quality:

$$\gamma^{\alpha_1}_{\{h,L\}} = \frac{\gamma_0\pi_0(\{h,L\})}{\gamma_0\pi_0(\{h,L\}) + (1-\gamma_0)\pi^{\alpha_1}(\{h,L\})} = 0.$$

3. No matter what the signaling strategy $\alpha_1(h|L,y)$ is, we have $\pi^{\alpha_1}(\{l,L\}) \leq \mu(L)$. So, in this case, the posterior reputation is:

$$\gamma^{\alpha_1}_{\{l,L\}} = \frac{\gamma_0\pi_0(\{l,L\})}{\gamma_0\pi_0(\{l,L\}) + (1-\gamma_0)\pi^{\alpha_1}(\{l,L\})} = \frac{\gamma_0}{\gamma_0 + (1-\gamma_0)\frac{\pi^{\alpha_1}(\{l,L\})}{\mu(L)}} \geq \gamma_0.$$

$\square$

The ineffectiveness of the history $\{h,H\}$ comes from the fact that it is anticipated by the customer, as the influencer always reports signal h when $q = H$. We can see that if the influencer always sends signal l when the quality is L, i.e., $\forall y$, $\alpha_1(l|L,y) = 1$, his reputation also stays at γ_0 with the history $\{l,L\}$. This demonstrates that if the influencer (rationally) honestly reports the product's quality and behaves like an honest influencer, it will not benefit his reputation. The result may seem counterintuitive, as one might expect that honest reporting would improve the influencer's reputation. However, reputation can only improve when a somewhat unexpected outcome is observed. For example, the history $\{l,L\}$ can improve the influencer's reputation only when the influencer has an incentive to send signal h with low quality L.

Lemma 1 also results in a simplified expression for the customer's posterior quality belief:

$$\mu^{\alpha_1}(H|h,y) = \frac{\mu(H)}{\mu(H) + \mu(L)(1-\gamma^{\alpha_1}_y)\alpha_1(h|L,y)}$$

$$\mu^{\alpha_1}(H|l,y) = 0$$

which means that the customer always chooses not to buy under signal l. Also, If the influencer would like to induce purchasing in the current period, then there is a restriction on signal h: $\alpha_1(h|L,y) \leq \frac{\mu(H)}{\mu(L)(1-\gamma_y^{\alpha_1})}$, derived from $\mu^{\alpha_1}(H|h,y) \geq \frac{1}{2}$. From here, we show the influencer's optimal stationary signaling strategy must be mixed:

Proposition 2. *Neither separating signaling nor pooling signaling can induce a stationary Markov perfect equilibrium.*

Proof. A simple intuition is that the optimal signaling strategy cannot exceed the restriction value. The reason is as follows:

1. If $\alpha_1(h|L,y) > \frac{\mu(H)}{\mu(L)(1-\gamma_y^{\alpha_1})}$, there will be no purchase in the current period, resulting in a zero reward for the influencer.

2. A high $\alpha_1(h|L,y)$ increases the probability of transitioning to states $\{h,L\}$, where the history clearly indicates that the influencer's type is rational, leading to a lower future discounted payoff.

Therefore, the optimal signaling strategy must lie within the range where the customer buys in the current period: $\forall y \in Y, \alpha_1(h|L,y) \leq \frac{\mu(H)}{\mu(L)(1-\gamma_y^{\alpha_1})}$. Note that the analysis here has proved that a pooling equilibrium, where the influencer always sends signal h, does not exist.

Also, from the reputation updating process in **Proposition 1**, we know that $\gamma_{\{l,L\}}^{\alpha_1} \geq \gamma_{\{h,H\}}^{\alpha_1} = \gamma_0 > \gamma_{\{h,L\}}^{\alpha_1} = 0$. So, we have the the following ascending ranking of the restriction values of $\alpha(h|L,y)$ for different states:

$$\frac{\mu(H)}{\mu(L)} = \frac{\mu(H)}{\mu(L)(1-\gamma_{\{h,L\}}^{\alpha_1})} < \frac{\mu(H)}{\mu(L)(1-\gamma_0)} = \frac{\mu(H)}{\mu(L)(1-\gamma_{\{h,H\}}^{\alpha_1})} \leq \frac{\mu(H)}{\mu(L)(1-\gamma_{\{l,L\}}^{\alpha_1})}.$$

Firstly, we can eliminate the case where $\forall y \in Y, 0 \leq \alpha_1(h|L,y) \leq \frac{\mu(H)}{\mu(L)}$. In this case, a marginal increase in $\alpha_1(h|L,y)$ results in a marginal increase of $\mu(L)r$ in the current expected reward and a decrease of

$$\mu(L)\delta\left(V_{\{l,L\}} - V_{\{h,L\}}\right)$$

in the future value. This marginal analysis implies that if there is an internal solution, then $\alpha_1(h|L,\{h,H\}) = \alpha_1(h|L,\{h,L\}) = \alpha_1(h|L,\{l,L\})$. However, this results in the decrease in future value being zero since $V_{\{l,L\}} = V_{\{h,L\}}$ when the three strategies are equal. Therefore, the marginal increase $\mu(L)r$ is always higher than the marginal decrease of zero, and the influencer will increase the probability $\alpha(h|L,y)$ until it hits the lowest restriction value $\frac{\mu(H)}{\mu(L)(1-\gamma_{\{h,L\}}^{\alpha_1})} = \frac{\mu(H)}{\mu(L)}$.

Thus, we have derived the optimal signaling strategy in state $\{h,L\}$: $\alpha_1^*(h|L,\{h,L\}) = \frac{\mu(H)}{\mu(L)}$. Note that the analysis here has proved that a separating equilibrium, where the influencer sends signal h under high quality and sends signal l under low quality, does not exist. □

The analysis implies that a stationary Markov signaling strategy must be a strictly mixed strategy. Additionally, the optimal signaling strategies $\alpha_1^*(h|L, \{h, H\})$ and $\alpha_1^*(h|\{l, L\})$ must be higher than $\frac{\mu(H)}{\mu(L)}$.

To find the optimal strategies that maximize the value function, we give the value function system of the influencer:

$$
\begin{aligned}
V_{\{h,H\}} &= \mu(H)(r + \delta V_{\{h,H\}}) + \mu(L)\alpha_1(h|L, \{h, H\})(r + \delta V_{\{h,L\}}) + \mu(L)\alpha_1(l|L, \{h, H\})\delta V_{\{l,L\}}) \\
V_{\{h,L\}} &= \mu(H)(r + \delta V_{\{h,H\}}) + \mu(L)\frac{\mu(H)}{\mu(L)}(r + \delta V_{\{h,L\}}) + \mu(L)(1 - \frac{\mu(H)}{\mu(L)}))\delta V_{\{l,L\}}) \\
V_{\{l,L\}} &= \mu(H)(r + \delta V_{\{h,H\}}) + \mu(L)\alpha_1(h|L, \{l, L\})(r + \delta V_{\{h,L\}}) + \mu(L)\alpha_1(l|L, \{l, L\})\delta V_{\{l,L\}})
\end{aligned}
$$

A marginal increase in $\alpha_1(h|L, \{h, H\})$ or $\alpha_1(h|L, \{l, L\})$ results in a marginal increase of $\mu(L)r$ in the current reward, while inducing a marginal decrease in the future value of $\delta\mu(L)(V_{\{l,L\}} - V_{\{h,L\}})$, which equals:

$$
\delta\mu(L)(V_{\{l,L\}} - V_{\{h,L\}}) = \mu(L)r\frac{\delta(\mu(L)\alpha_1(h|L, \{l, L\}) - \mu(H))}{1 + \delta(\mu(L)\alpha_1(h|L, \{l, L\}) - \mu(H))}.
$$

Given that $\alpha_1(h|L, \{l, L\}) > \frac{\mu(H)}{\mu(L)}$, the expression is positive. We can easily check that this expression increases with $\alpha_1(h|L, \{l, L\})$, but it is always less than $\mu(L)r$, which represents the marginal increase in the current reward. Therefore, internal solutions are not possible, and the optimal value of $\alpha_1(h|L, \{h, H\})$ or $\alpha_1(h|L, \{l, L\})$ is achieved at their upper bound. We have already derived that the posterior reputation in state $L, \{h, H\}$ is γ_0. Thus, the optimal signaling strategy under state $\{h, H\}$ is:

$$
\alpha_1^*(h|L, \{h, H\}) = \frac{\mu(H)}{\mu(L)(1 - \gamma_0)}.
$$

The remaining task is to find the upperbound for $\alpha_1(h|L, \{l, L\})$. We need to firstly derive the steady-state distributions of states. Then, we can derive the endogenous upper bound of $\alpha_1(h|L, \{l, L\})$, which needs to be weakly higher than $\alpha_1(h|L, \{l, L\})$. Afterward, the optimal signaling strategy can be identified. The details are provided in the appendix. Here, we present the main result:

Theorem 1. *The unique optimal stationary Markov signaling strategy for the influencer is as follows:*

1. *Sending a pure signal h when the current product's quality is H.*
2. *Sending strictly mixed signals when the current product's quality is L, with the following specific strategies:*

   i. $\alpha_1^*(h|L, \{h, H\}) = \frac{\mu(H)}{\mu(L)(1-\gamma_0)}$.

   ii. $\alpha_1^*(h|L, \{h, L\}) = \frac{\mu(H)}{\mu(L)}$.

iii. $\alpha_1^*(h|L,\{l,L\}) = \frac{\mu(H)(\mu(L)-\mu(H))}{\mu(L)((1-\gamma_0)\mu(L)-\mu(H))}$.

The theorem demonstrates that the optimal signaling strategy is unique for all exogenous variables because the decrease in future value is always less than the increase in the current reward. With myopic customers, one-period cheating behavior does not result in long-lasting reputation damage for the influencer. Consequently, in each period with a low-quality product, the influencer maximizes the probability of cheating while still maintaining the informativeness of the high signal. Finally, we can derive the stochastic reputation dynamics under the unique stationary equilibrium.

Proposition 3. *Under the unique stationary equilibrium, the reputation dynamics is a stochastic transition process among three increasing reputation values* 0, γ_0 *and* $\gamma_{\{l,L\}}^{\alpha_1^*}$, *in which a higher reputation status corresponds to a higher probability of cheating.*

The reputation evolution process is directly computed from the optimal signaling strategy α_1^* and the prior quality distribution μ. It gives a explicit capture of how influencer's reputation flows dynamically in steady state. We show the reputation dynamics in Figure 2.

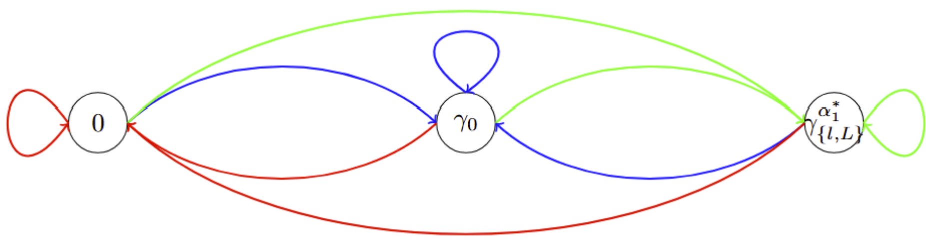

Figure 2: Stationary Reputation Dynamics

The blue arrows show the reputation updating to γ_0 when a high quality good is realized currently and the influencer optimally sends high signal. The red arrows show the reputation updating to 0 when a low quality good is realized and the influencer strategically sends signal h. The green arrows show the reputation updating to $\gamma_{\{l,L\}}^{\alpha_1^*}$ when a low quality good is realized and the influencer strategically sends signal l.

5 Conclusion

In this paper, we model the interaction between a long-lived influencer and a sequence of myopic customers using a repeated game with incomplete information. We define the Markov decision process of the influencer and illustrate how the current persuasion strategy can affect future reputation, thereby influencing the expected future payoff. Our first result addresses the reputation effect of each history, particularly for histories with matched signals and quality. We show that only when the history reflects that the influencer has sacrificed benefits does it positively impact the influencer's reputation. In contrast, histories that align with the influencer's incentives can only weakly decrease their reputation. Second, we derive the unique optimal stationary Markov signaling strategy for the influencer across all exogenous parameters and give the corresponding stochastic reputation evolution process.

The limitation of our model is that the main theorem relies on the assumption that the customer's recall of history is limited. If customers had longer observation histories, a single mismatched history would have a more prolonged negative impact on the influencer's reputation, potentially altering the optimal stationary Markov signaling strategy. For instance, with perfect recall—where the customer can observe all past signal and quality pairs—any mismatched signal would reveal the influencer's type to all future customers. In this case, for a sufficiently patient influencer, the optimal strategy would be honest reporting. However, perfect recall is unrealistic due to limitations in customer attention, platform data storage, and information decay over extended periods.

A Proof of Theorem 1

We know that the optimal signaling strategy at state $L, \{h, L\}$ falls within the range $\left(\frac{\mu(H)}{\mu(L)(1-\gamma_0)}, \frac{\mu(H)}{\mu(L)(1-\gamma^{\alpha_1}_{\{l,L\}})}\right]$. Since the range is endogenously determined, we need to find the steady-state distribution of states to capture the domain of $\alpha_1(h|L, \{l, L\})$. The system of equations is:

$$
\begin{aligned}
\pi^{\alpha_1}_{\{h,H\}} &= \mu(H)(\pi^{\alpha_1}_{\{h,H\}} + \pi^{\alpha_1}_{\{h,L\}} + \pi^{\alpha_1}_{\{l,L\}}) \\
\pi^{\alpha_1}_{\{l,H\}} &= 0 \\
\pi^{\alpha_1}_{\{h,L\}} &= \mu(L)\left(\pi^{\alpha_1}_{\{h,H\}}\frac{\mu(H)}{\mu(L)(1-\gamma_0)} + \pi^{\alpha_1}_{\{l,L\}}\alpha_1(h|L,\{l,L\}) + \pi^{\alpha_1}_{\{h,L\}}\frac{\mu(H)}{\mu(L)}\right), \\
\pi^{\alpha_1}_{\{l,L\}} &= \mu(L)\left(\pi^{\alpha_1}_{\{h,H\}}\left(1-\frac{\mu(H)}{\mu(L)(1-\gamma_0)}\right) + \pi^{\alpha_1}_{\{l,L\}}(1-\alpha_1(h|L,\{l,L\})) + \pi^{\alpha_1}_{\{h,L\}}\left(1-\frac{\mu(H)}{\mu(L)}\right)\right).
\end{aligned}
$$

The normalization condition is:

$$
1 = \pi^{\alpha_1}_{\{h,H\}} + \pi^{\alpha_1}_{\{h,L\}} + \pi^{\alpha_1}_{\{l,L\}}
$$

Thus, the steady-state distribution is:

$$
\begin{aligned}
\pi^{\alpha_1}_{\{h,H\}} &= \mu(H) \\
\pi^{\alpha_1}_{\{l,H\}} &= 0 \\
\pi^{\alpha_1}_{\{h,L\}} &= \frac{\mu^2(H) + \mu^2(L)(1-\gamma_0)\alpha_1(h|L,\{l,L\})}{(1-\gamma_0)(1+\alpha_1(h|L,\{l,L\}))} \\
\pi^{\alpha_1}_{\{l,L\}} &= \frac{\mu^2(L)(1-\gamma_0) - \mu^2(H)}{\mu(L)(1-\gamma_0)(1+\alpha_1(h|L,\{l,L\}))}
\end{aligned}
$$

Note that we have assumed $\gamma_0 < 1 - \frac{\mu(H)}{\mu(L)}$, ensuring that the distribution probabilities are positive. The updated reputation $\gamma^{\alpha_1}_{\{l,L\}}$ after observing history $\{l, L\}$ is:

$$
\gamma^{\alpha_1}_{\{l,L\}} = \frac{\gamma_0\mu(L)}{\gamma_0\mu(L) + (1-\gamma_0)\pi^{\alpha_1}_{\{l,L\}}}
$$

Substituting $\pi^{\alpha_1}_{\{l,L\}}$ into this expression, we get:

$$\gamma^{\alpha_1}_{\{l,L\}} = \frac{\gamma_0\mu^2(L)(1+\alpha_1(h|L,\{l,L\}))}{\gamma_0\mu^2(L)(1+\alpha_1(h|L,\{l,L\}))+\mu^2(L)(1-\gamma_0)-\mu^2(H)}$$

Next, the endogenous upper bound for $\alpha_1(h|L,\{l,L\})$ is derived from:

$$\frac{\mu(H)}{\mu(L)(1-\gamma^{\alpha_1}_{\{l,L\}})} = \frac{\mu(H)\left(\gamma_0\mu(L)(1+\alpha_1(h|L,\{l,L\}))+\mu(L)(1-\gamma_0)-\frac{\mu^2(H)}{\mu(L)}\right)}{\mu^2(L)(1-\gamma_0)-\mu^2(H)}$$

This upper bound is weakly higher than $\alpha_1(h|L,\{l,L\})$ if:

$$\alpha_1(h|L,\{l,L\}) \le \frac{\mu(H)(\mu(L)-\mu(H))}{\mu(L)\left((1-\gamma_0)\mu(L)-\mu(H)\right)}$$

Therefore, the domain of $\alpha_1(h|L,\{l,L\})$ is:

$$\frac{\mu(H)}{\mu(L)(1-\gamma_0)} < \alpha_1(h|L,\{l,L\}) \le \frac{\mu(H)(\mu(L)-\mu(H))}{\mu(L)\left((1-\gamma_0)\mu(L)-\mu(H)\right)}$$

Thus, we have derived the optimal signaling strategy in state $L,\{h,L\}$, which is just the upper bound of $\alpha_1(h|L,\{l,L\})$.

B Proof of Proposition 3

From **Theorem 1**, we get the expression for optimal signaling strategy $\alpha_1^*(h|L,\{l,L\})$, thus we can derive the reputation when history $\{l,L\}$ occurs:

$$\gamma^{\alpha_1^*}_{\{l,L\}} = \frac{\gamma_0\mu(H)\mu(L)(\mu(L)-\mu(H))}{\gamma_0\mu(H)\mu(L)(\mu(L)-\mu(H))+((1-\gamma_0)\mu(L)-\mu(H))((1-\gamma_0)\mu^2(L)-\mu^2(H))}$$

The transition probabilities are multiplications of signaling strategy α_1^* and quality distribution μ, we show each transition probability from reputation A to reputation B in the following table:

Transition from $A \to B$	Probability
$0 \to 0$	$\mu(H)$
$0 \to \gamma_0$	$\mu(H)$
$0 \to \gamma^{\alpha_1^*}_{\{l,L\}}$	$\mu(L)-\mu(H)$
$\gamma_0 \to 0$	$\frac{\mu(H)}{1-\gamma_0}$
$\gamma_0 \to \gamma_0$	$\mu(H)$
$\gamma_0 \to \gamma^{\alpha_1^*}_{\{l,L\}}$	$\mu(L)-\frac{\mu(H)}{1-\gamma_0}$
$\gamma^{\alpha_1^*}_{\{l,L\}} \to 0$	$\frac{\mu(H)(\mu(L)-\mu(H))}{((1-\gamma_0)\mu(L)-\mu(H))}$
$\gamma^{\alpha_1^*}_{\{l,L\}} \to \gamma_0$	$\mu(H)$
$\gamma^{\alpha_1^*}_{\{l,L\}} \to \gamma^{\alpha_1^*}_{\{l,L\}}$	$\mu(L)-\frac{\mu(H)(\mu(L)-\mu(H))}{((1-\gamma_0)\mu(L)-\mu(H))}$

Table 2: Reputation Transition Probabilities

References

Blackwell, D. (1965), 'Discounted dynamic programming', *The Annals of Mathematical Statistics* **36**(1), 226–235.

Che, Y.-K., Kim, K. & Mierendorff, K. (2023), 'Keeping the listener engaged: a dynamic model of bayesian persuasion', *Journal of Political Economy* **131**(7), 1797–1844.

Ely, J. C. (2017), 'Beeps', *American Economic Review* **107**(1), 31–53.

Fainmesser, I. P. & Galeotti, A. (2021), 'The market for online influence', *American Economic Journal: Microeconomics* **13**(4), 332–372.

Glückler, J. & Armbrüster, T. (2003), 'Bridging uncertainty in management consulting: The mechanisms of trust and networked reputation', *Organization studies* **24**(2), 269–297.

Henry, E. & Ottaviani, M. (2019), 'Research and the approval process: The organization of persuasion', *American Economic Review* **109**(3), 911–955.

Höner, J. (2002), 'Reputation and competition', *American economic review* **92**(3), 644–663.

Honryo, T. (2018), 'Risky shifts as multi-sender signaling', *Journal of Economic Theory* **174**, 273–287.

Kamenica, E. & Gentzkow, M. (2011), 'Bayesian persuasion', *American Economic Review* **101**(6), 2590–2615.

Kreps, D. M., Milgrom, P., Roberts, J. & Wilson, R. (1982), 'Rational cooperation in the finitely repeated prisoners' dilemma', *Journal of Economic theory* **27**(2), 245–252.

Kreps, D. M. & Wilson, R. (1982), 'Reputation and imperfect information', *Journal of economic theory* **27**(2), 253–279.

Levine, D. K. (2021), 'The reputation trap', *Econometrica* **89**(6), 2659–2678.

Liu, Q. (2011), 'Information acquisition and reputation dynamics', *The Review of Economic Studies* **78**(4), 1400–1425.

Mailath, G. J. & Samuelson, L. (2001), 'Who wants a good reputation?', *The Review of Economic Studies* **68**(2), 415–441.

Mailath, G. J. & Samuelson, L. (2006), *Repeated games and reputations: long-run relationships*, Oxford university press.

Maskin, E. & Tirole, J. (2001), 'Markov perfect equilibrium: I. observable actions', *Journal of Economic Theory* **100**(2), 191–219.

Morgan, J. & Stocken, P. C. (2003), 'An analysis of stock recommendations', *RAND Journal of economics* pp. 183–203.

Mostagir, M. & Siderius, J. (2023), 'Strategic reviews', *Management Science* **69**(2), 904–921.

Renault, J., Solan, E. & Vieille, N. (2017), 'Optimal dynamic information provision', *Games and Economic Behavior* **104**, 329–349.

Chapter 2

Pro-rata vs User-Centric in the music streaming industry

Preface

Building on the foundational ideas of signaling and strategic behavior in Chapter 1, Chapter 2 turns to the digital economy, focusing on the music streaming industry. Here, the strategic interaction shifts to a setting where artists compete for royalties through adjustments in content quality. This chapter explores two dominant remuneration models—pro-rata and user-centric—and models artist responses to these schemes within an endogenous framework. By situating artists within a competitive, information-limited market, this chapter expands on the strategic adaptation themes established in Chapter 1, illustrating how platform rules influence agents' incentives and ultimately the fairness and efficiency of outcomes.

This chapter's findings on the role of endogenous responses in remuneration outcomes underscore the broader importance of designing incentive-compatible rules in platform-based economies, where agent behavior is shaped by the regulatory structure. The examination of artists' strategic quality adjustments under different payment schemes parallels the influencer's signaling strategy from Chapter 1, emphasizing that competitive environments and remuneration schemes fundamentally influence agent behavior. These insights set the stage for Chapter 3, where resource allocation and waste minimization are examined as part of a broader discussion on minimizing allocation waste within a information constrained environment.

Economics Letters 226 (2023) 111111

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

Pro-rata vs User-centric in the music streaming industry

Xiaochang Lei

Adam Smith Business School, University of Glasgow, Glasgow, United Kingdom

A R T I C L E I N F O

Article history:
Received 9 January 2023
Received in revised form 31 March 2023
Accepted 4 April 2023
Available online 6 April 2023

Keywords:
Pro-rata and user-centric
Digital streaming platform
Endogenous modeling

A B S T R A C T

An endogenous model, which allows the artists to determine the streaming times strategically, is used to compare two remuneration rules in the music streaming industry: Pro-rata (P rule) and User-centric (U rule). The two judgement criteria are 1. efficiency, in terms of no dominance on quality profile. 2. egalitarian fairness, in terms of the lowest royalty among all artists. Our main result is that P rule always outperforms U rule in efficiency and fairness when the superstar's marginal cost is the lowest. This means that the transition from P rule to U rule can not only enlarge the existing royalty gap but also decrease the efficiency of the music streaming industry.

1. Introduction

In October 2020, the House of Commons Digital, Culture, Media and Sport Committee launched an inquiry into the music streaming industry to deal with the long-lasting complaints on unfair creator remuneration (The Digital, Culture, Media and Sport Committee, 2021). While the methods based on Shapley value from cooperative game theory provide a theoretical foundation for revenue sharing, it requests stand-alone costs, which are not observable or not simple to compute (Shiller and Waldfogel, 2013). There are two widely used payment methods in the music streaming industry. Most industry heavyweights (e.g. Spotify and Apple Music) use the pro-rata rule (P rule). Under P rule, the subscription fees of all subscribers are aggregated at first as a royalty pot and then proportionally divided. In contrast, under the user-centric rule (U rule), which Deezer uses, the subscription fee of each subscriber is first proportionally split and then aggregated (See Example 1).

Example 1. P rule vs. U rule in 2 × 2 case

There are two songs A, B and two subscribers 1,2. The streaming matrix is in Table 1. In each small cell, the number is the streaming times for a song of a subscriber.

Suppose the subscription fee is £10, which is the same for all subscribers. So, the total royalty pot is £20. Under P rule, the royalty R for each artist are $R_A^P = \frac{20+10}{20+10+10+60} \times £20 = £6$ and $R_B^P = \frac{10+60}{20+10+10+60} \times £20 = £14$. Under U rule, the royalty for each artist are $R_A^U = \frac{20}{20+10} \times £10 + \frac{10}{10+60} \times £10 = £8.1$ and $R_B^U = \frac{10}{20+10} \times £10 + \frac{60}{10+60} \times £10 = £11.9$.

There is a hot debate on how the royalty pot should be allocated and what the properties of different rules are (Page and Safir, 2018a,b). There is an argument from empirical research (Muikku, 2017; Pedersen, 2014; Hesmondhalgh, 2021), Law (Dimont, 2018) and industry news (Dredge, 2021) that U rule can benefit the specialists (unpopular artists with low total streaming times) more than P rule by giving them more royalty (as Example 1 shows). This may be true when the streaming matrix is exogenously given. However, this relation reverses when the streaming matrix is endogenously determined.

In this paper, we construct an endogenous model that allows the artists to strategically change their songs' quality to influence the streaming matrix and maximize their payoff. Preferences and quality profiles jointly determine the consumers' streaming time on each song, and the streaming matrix is endogenously determined in equilibrium. By comparing equilibria, we aim to answer two widely debated questions: 1. which is more efficient? 2. which is fairer? No dominance of quality profile captures efficiency. Egalitarianism captures fairness since, in reality, the specialists complain about their low royalty. Our main theorem is that in the two-artist model, P rule outperforms U rule in efficiency and fairness. We find the theorem still holds when there are one superstar and two identical specialists. Although the cases are special, our endogenous model does give a unique result that contradicts the empirical research, which neglects the artists' strategic behavior.

The structure of the paper is as follows: we discuss other related theoretical papers, and then, Section 2 shows the general setting. Section 3 gives equilibrium and discusses the 2-artist model and 3-artist model with identical specialists. Section 4 concludes the paper with limitations and contributions.

1.1. Literature review

To the best of our knowledge, the only theoretical paper comparing the two rules is Alaei et al. (2020). They use an

E-mail address: Xiaochang.Lei@glasgow.ac.uk.

https://doi.org/10.1016/j.econlet.2023.111111
0165-1765/

Table 1
Streaming matrix in 2 × 2 case.

		Songs	
		A	*B*
Subscribers	1	20	10
	2	10	60

extensive-form game to capture the interactions between the platform, the artists and the consumers. They focus on the platform's pricing strategy and how it can sustain a set of artists (participation constraint). We do not consider these two since, in reality, the artists can work part-time. Also, the subscription fee has been fixed for decades due to intensive competition among platforms with highly homogeneous content. We focus on the incentive compatibility constraint that makes the streaming matrix endogenously determined.

For other related stories, Ginsburgh and Zang (2003) considers a problem of sharing the revenue from selling museum passes. Their model can be extended to many real-life cases (e.g., travel cards). Under their setting, the complex Shapley value has a very simple form. However, their model cannot be used in the music streaming industry. In the museum pass game, the consumers' consumption for each museum is binary: 0 for non-visit and 1 for a visit. Here, consumers' consumption of each song in music streaming has intensity. Flores-Szwagrzak and Treibich (2020) considers revenue sharing in teamwork. They use an individual's productivity in his stand-alone project as a proxy for his contribution to teamwork. However, millions of songs are on the platform in the music streaming industry, and it is impossible to find every song's stand-alone price. The last paper (Bergantinos and Moreno-Ternero, 2020) considers the problem of sharing the revenue from selling tickets between football clubs. Again, the story differs from the music streaming industry due to the lack of consumers' consumption intensity.

2. Setting

There are $m+1$ artists on the platform, and each artist publishes one song. The set of the artists is $M=\{0,1,2,\dots,m\}$. The artist 0 is a superstar, and all subscribers like him. The artists $1\sim m$ are specialists and have their fan base. There are n subscribers with private preferences who have paid a subscription fee of p. Let $\pi_0\in(0,1)$ denote the probability that a subscriber only likes the song 0. let $\pi_k\in(0,1)$ denote the probability that a subscriber likes song $k\in\{1,2,\dots,m\}$ as well as song 0. The distribution of subscribers' types $\pi=\{\pi_0,\pi_1,\pi_2,\dots,\pi_m\}$ is common knowledge. Our assumption that consumers are interested in only one specialist is realistic. For example, we can imagine that artists $1\sim m$ are the ones who create songs in different styles: classic, jazz, folk, etc. Each of them attracts a small group of consumers with special tastes. While as a superstar, the artist 0 publishes popular music, which is widely accepted. Another explanation is that the superstar publishes songs in English, while the specialists publish songs in their languages like Dutch, Japanese, German, etc.

Given the quality profile $q=\{q_0,q_1,q_2,\dots,q_m\}$. Each consumer determines his streaming times for each song. If a user likes the song k then his utility for streaming song k is $u_k(t_k)=q_kt_k-\frac{1}{2}t_k^2$. Else, his utility for streaming song k is 0. The quadratic and additive utility function allows the simplest linear marginal utility function: $MU(q_k)=q_k-t$. So the streaming time for song k is $t_k=q_k$ for its fans and 0 for the other consumers. So, given any quality profile q, the streaming matrix is given in Table 2. In each cell, there is a corresponding streaming time (equal to quality) for a song of a subscriber.

Table 2
The general streaming matrix.

		Songs/Artists				
		0	1	2		m
Subscribers	π_0	q_0	0	0	$\cdots$	0
	π_1	q_0	q_1	0	$\cdots$	0
	π_2	q_0	0	q_2	$\cdots$	0
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	π_m	q_0	0	0	$\cdots$	q_m

The artist k can only control quality q_k while taking other quality q_{-k} as given. So, under P rule, his royalty is:

$$R_k^P(q_k,q_{-k})=\begin{cases}\frac{q_0}{q_0+\sum_{j=1}^m\pi_jq_j}np & k=0\\ \frac{\pi_kq_k}{q_0+\sum_{j=1}^m\pi_jq_j}np & k>0\end{cases}$$

Under U rule, the artist k's royalty is:

$$R_k^U(q_k,q_{-k})=\begin{cases}\pi_0np+\sum_{j=1}^m\frac{q_0}{q_0+q_j}\pi_jnp & k=0\\ \frac{q_k}{q_0+q_k}\pi_knp & k>0\end{cases}$$

Taking the first derivative on q_k, we can find first-order condition $MR_k^P(q_k,q_{-k})=c_k$ under the P rule or $MR_k^U(q_k,q_{-k})=c_k$ under the U rule. Equilibrium can be found by combining all artists' best response functions. The formal definition of this static game is (M,π,c,n,p). A royalty distribution rule r is a partition of total royalty pot np among the artists in M.

Definition (*Efficiency*)**.** A royalty distribution rule r_1 is more efficient than rule r_2 if r_1 can induce a weakly higher quality profile for all songs and at least one song's quality should be strictly higher:

$$\forall k\in M,\ q_k^1\ge q_k^2$$
$$\exists k\in M,\ q_k^1>q_k^2$$

Definition (*Fairness*)**.** Fixed the royalty pot, a royalty distribution rule r_1 is egalitarian fairer than rule r_2 if r_1 can induce a higher lowest royalty: let $R^{1*}=min\{R_0^1,R_1^1,\dots,R_m^1\}$ and $R^{2*}=min\{R_0^2,R_1^2,\dots,R_m^2\}$. Rule r_1 is fairer than r_2 if $R^{1*}>R^{2*}$.

3. Equilibrium

To express the equilibrium under P rule, we first define a $m+1\times m+1$ matrix with marginal costs and consumer's type distribution:

$$A=\begin{bmatrix}1 & 0 & 0 & 0 & \cdots & 0\\ c_0 & -\frac{c_1}{\pi_1} & c_0-\frac{c_1}{\pi_1} & c_0-\frac{c_1}{\pi_1} & \cdots & c_0-\frac{c_1}{\pi_1}\\ c_0 & c_0-\frac{c_2}{\pi_2} & -\frac{c_2}{\pi_2} & c_0-\frac{c_2}{\pi_2} & \cdots & c_0-\frac{c_2}{\pi_2}\\ c_0 & c_0-\frac{c_3}{\pi_3} & c_0-\frac{c_3}{\pi_3} & -\frac{c_3}{\pi_3} & \cdots & c_0-\frac{c_3}{\pi_3}\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ c_0 & c_0-\frac{c_m}{\pi_m} & c_0-\frac{c_m}{\pi_m} & c_0-\frac{c_m}{\pi_m} & \cdots & -\frac{c_m}{\pi_m}\end{bmatrix}$$

The matrix A is derived from the linear system of the FOCs under P rule. The first row is special due to the superstar's popularity in 1. Also we define a group of matrix: A_k, which replace the $(k+1)$'th column of A with a $m+1$-dimension vector: $v=(1,0,\dots,0)$. Then by using Cramer's rule, the equilibrium can be found.

Proposition 1. *Under P rule, the equilibrium quality profile and the royalty profile are:*

$$\forall k > 0,\ q_k^{*P} = \frac{|A_k|}{\pi_k}\frac{np}{c_0}\frac{\sum_{j=1}^{m}|A_j|}{(\sum_{j=0}^{m}|A_j|)^2}$$
$$q_0^{*P} = |A_0|\frac{np}{c_0}\frac{\sum_{j=1}^{m}|A_j|}{(\sum_{j=0}^{m}|A_j|)^2}$$
$$\forall k,\ R_k^{*P} = \frac{|A_k|}{\sum_{k=0}^{m}|A_k|}np$$

Proposition 2. *Under U rule, the equilibrium quality profile and the royalty profile are:*

$$\forall k > 0,\ q_k^{*U} = np\frac{\sum_{j=1}^{m}\sqrt{\pi_j c_j}}{\sum_{j=0}^{m}c_j}\left(\sqrt{\frac{\pi_k}{c_k}} - \frac{\sum_{j=1}^{m}\sqrt{\pi_j c_j}}{\sum_{j=0}^{m}c_j}\right)$$
$$q_0^{*U} = \frac{np(\sum_{j=1}^{m}\sqrt{\pi_j c_j})^2}{(\sum_{j=0}^{m}c_j)^2}$$
$$\forall k > 0,\ R_k^{*U} = np\left(\pi_k - \frac{\sum_{j=1}^{m}\sqrt{\pi_j c_j}}{\sum_{j=0}^{m}c_j}\sqrt{\pi_k c_k}\right)$$
$$R_0^{*U} = \pi_0 np + np\frac{(\sum_{j=1}^{m}\sqrt{\pi_j c_j})^2}{\sum_{j=0}^{m}c_j}$$

There is no easy way to mathematically compare the quality profile and the royalty gap. We discuss two special cases: the 2-artist and 3-artist models with identical specialists.

Theorem 1. *In the two-artist model, P rule is better than U rule in efficiency and fairness.*

The intuition for better performance of efficiency is from the competition. Under P rule, the artists must compete on the streaming of all subscribers. Under U rule, the artists only compete on the streaming of their fans. The intuition for the dominance of fairness is that the P rule can significantly increase the incentive for the specialists to increase their songs' quality. In contrast, the superstar's incentive increase is not that significant.

Theorem 1 can be easily checked. From Propositions 1 and 2, in this 2-artist model, under P rule, the equilibrium quality profile is $q_0^{*P} = \frac{c_1}{c_0}\frac{\pi_1 np}{c_0(\frac{c_1}{c_0}+\pi_1)^2}$ and $q_1^{*P} = \frac{\pi_1 np}{c_0(\frac{c_1}{c_0}+\pi_1)^2}$. The royalty profile is $R_0^{*P} = \frac{\frac{c_1}{c_0}}{\frac{c_1}{c_0}+\pi_1}np$ and $R_1^{*P} = \frac{\pi_1}{\frac{c_1}{c_0}+\pi_1}np$. Under U rule, the equilibrium quality profile is $q_0^{*U} = \frac{c_1}{c_0}\frac{\pi_1 np}{c_0(\frac{c_1}{c_0}+1)^2}$ and $q_1^{*U} = \frac{\pi_1 np}{c_0(\frac{c_1}{c_0}+1)^2}$. The royalty profile is $R_0^{*U} = \pi_0 np + \pi_1 np\frac{\frac{c_1}{c_0}}{\frac{c_1}{c_0}+1}$ and $R_1^{*U} = \pi_1 np\frac{1}{\frac{c_1}{c_0}+1}$. Given $\pi_1 < 1$, both equilibrium qualities under P rule are higher than that under U rule ($q_0^{*P} > q_0^{*U}$ and $q_1^{*P} > q_1^{*U}$). Also, $R_1^{*P} > R_1^{*U}$, the specialist gets more under P rule.

3.1. Extension: Model with one superstar and two identical specialists

To give a mathematically tractable extension, we assume that there are only two types of artists: one superstar and 2 homogeneous specialists. The superstar 0 has low marginal cost c_L, and the specialists have uniform high marginal cost c_H. The marginal costs satisfy $c_H \geq c_L > 0$. Also, all the specialists share the same popularity. To make the notation coherent, we assume the specialists share uniform popularity $\pi_H(\pi_H \leq \frac{1}{m})$. We find that:

Proposition 3. *In the 3-artist with identical specialists model, Theorem 1 still holds.*

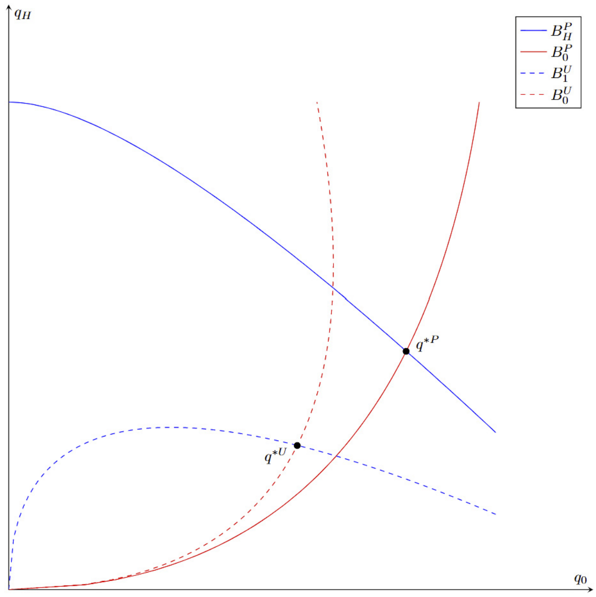

Fig. 1. Best Response functions and Equilibrium Quality Profile.

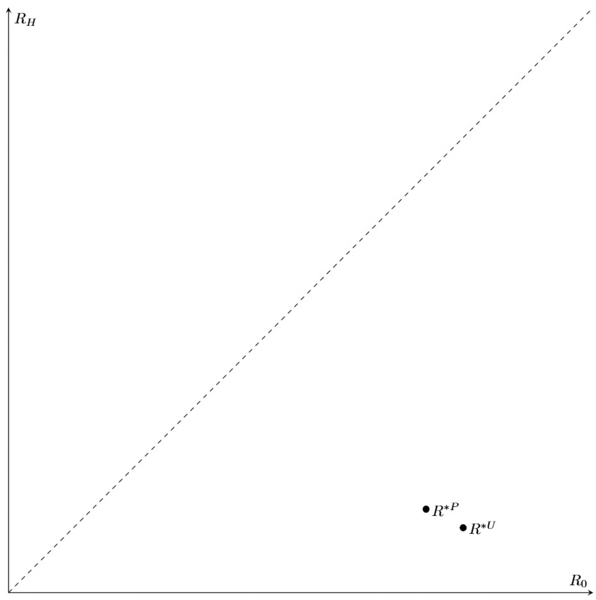

Fig. 2. Royalty profile and egalitarian fairness.

We give a numerical example. Considering the case $\pi_H = \frac{1}{3}$, $c_H = c_L = 10$, $n = 100$, $p = 8$. Let B_0 denote the best response function of the superstar and B_H denote the best response function of the homogeneous specialists. The result is shown in Figs. 1 and 2. Obviously, P rule induces a higher equilibrium quality profile. Also, the royalty profile closer to the 45° line is more egalitarian.

4. Discussion and conclusion

This paper proposes an endogenous model, which shows P rule outperforms U rule in efficiency and fairness. It contradicts the results from other exogenous models. Our result indicates that the current transition from P to U rule can reduce efficiency and fairness in the music streaming industry. So, instead of using U rule, the platforms should increase the exogenous variables (e.g., expanding the market and attracting more subscribers). Although the theorems are compelling, there are some limitations to the model. First, we assume the consumers have uniform marginal utility functions on streaming, which can be heterogeneous. For example, some people are more addicted to music. Second, our analysis can only be limited in some special cases for mathematical tractability.

Data availability

No data was used for the research described in the article.

Appendix A. Proof of propositions

Proof of Proposition 1. The marginal royalty under P rule is:

$$MR_k^P(q_k, q_{-k}) = \begin{cases} \frac{\sum_{j=1}^m \pi_j q_j}{(q_0+\sum_{j=1}^m \pi_j q_j)^2} np & k=0 \\ \frac{\pi_k(q_0+\sum_{j\neq k|j>0}^m \pi_j q_j)}{(q_0+\sum_{j=1}^m \pi_j q_j)^2} np & k>0 \end{cases}$$

To solve the FOC system, let $\forall k > 0,\ \pi_k q_k = x_k q_0$ and $x_0 = 1$. The linear system is:

$$\forall k>0,\ q_0 = \frac{\pi_k np}{c_k}\frac{\sum_{j\neq k}^m x_j}{(\sum_{j=0}^m x_j)^2}$$
$$q_0 = \frac{np}{c_0}\frac{\sum_{j=1}^m x_j}{(\sum_{j=0}^m x_j)^2}$$
$$x_0 = 1$$

First, we solve the linear system for $\boldsymbol{x}$ and then the equilibrium quality profile can be found. The linear system for $\boldsymbol{x}$ is:

$$\frac{\sum_{j\neq 0}^m x_j}{c_0} = \frac{\pi_1\sum_{j\neq 1}^m x_j}{c_1} = \cdots = \frac{\pi_m\sum_{j\neq m}^m x_j}{c_m}$$
$$x_0 = 1$$

In the matrix form:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ c_0 & -\frac{c_1}{\pi_1} & c_0-\frac{c_1}{\pi_1} & c_0-\frac{c_1}{\pi_1} & \cdots & c_0-\frac{c_1}{\pi_1} \\ c_0 & c_0-\frac{c_2}{\pi_2} & -\frac{c_2}{\pi_2} & c_0-\frac{c_2}{\pi_2} & \cdots & c_0-\frac{c_2}{\pi_2} \\ c_0 & c_0-\frac{c_3}{\pi_3} & c_0-\frac{c_3}{\pi_3} & -\frac{c_3}{\pi_3} & \cdots & c_0-\frac{c_3}{\pi_3} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ c_0 & c_0-\frac{c_m}{\pi_m} & c_0-\frac{c_m}{\pi_m} & c_0-\frac{c_m}{\pi_m} & \cdots & -\frac{c_m}{\pi_m} \end{bmatrix} \times \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

The size of the parameter matrix is $m+1 \times m+1$. The solution $\boldsymbol{x}$ can be found using Cramer's rule. □

Proof of Proposition 2. The marginal royalty under U rule is:

$$MR_k^U(q_k, q_{-k}) = \begin{cases} \sum_{j=1}^m \frac{q_j}{(q_0+q_j)^2}\pi_j np & k=0 \\ \frac{q_0}{(q_0+q_k)^2}\pi_k np & k>0 \end{cases}$$

The FOC linear system is:

$$\sum_{j=1}^m \frac{\pi_j q_j}{(q_0+q_j)^2} np = c_0$$
$$\frac{\pi_1 q_1}{(q_0+q_1)^2} np = c_1$$
$$\vdots$$
$$\frac{\pi_m q_m}{(q_0+q_m)^2} np = c_m$$

From the equation $1 \sim m$, there is a relation between $q_k (k > 0)$ and q_0:

$$q_k = \sqrt{\frac{\pi_2}{c_2} npq_0} - q_0$$

Substituting the relation into the equation 0, q_0^{*U} can be found. Then, all the other equilibrium quality q_k^{*U} can be found. □

Proof of Proposition 3. From Propositions 1 and 2, under U rule, the equilibrium quality profile and royalty profile are:

$$\forall k>0,\ q_k^{*U} = np\frac{2\pi_H}{c_L+2c_H}(1-\frac{2c_H}{c_L+2c_H})$$
$$q_0^{*U} = np\frac{4\pi_H c_H}{(c_L+2c_H)^2}$$
$$\forall k>0,\ R_k^{*U} = np\pi_H(1-\frac{2c_H}{c_L+2c_H})$$
$$R_0^{*U} = np((1-2\pi_H)+\pi_H\frac{4c_H}{c_L+2c_H})$$

In equilibrium, under the P rule, the quality profile and royalty profile are:

$$\forall k>0,\ q_k^{*P} = \frac{c_L^2}{(\frac{2c_Lc_H}{\pi_H}+c_L^2)^2}\frac{2npc_L}{\pi_H}$$
$$q_0^{*P} = \frac{\frac{2c_Lc_H}{\pi_H}-c_L^2}{(\frac{2c_Lc_H}{\pi_H}+c_L^2)^2}2npc_L$$
$$\forall k>0,\ R_k^{*P} = \frac{c_L}{\frac{2c_H}{\pi_H}+c_L}np$$
$$R_0^{*P} = \frac{\frac{2c_H}{\pi_H}-c_L}{\frac{2c_H}{\pi_H}+c_L}np$$

To check the efficiency property, for the superstar, we have the following:

$$q_0^{*P} - q_0^{*U} = 2npc_L^3\frac{(\frac{8}{\pi_H}-12)c_H^2 - c_L^2 + (\frac{1}{\pi_H}-2-\pi_H)2c_Hc_L}{(c_L+2c_H)^2(\frac{2c_Lc_H}{\pi_H}+c_L^2)^2}$$

The check the sign, the denominator is:

$$(\frac{8}{\pi_H}-12)c_H^2 - c_L^2 + (\frac{1}{\pi_H}-2-\pi_H)2c_Hc_L$$
$$\geq 4c_H^2 - c_L^2 - c_Hc_L\ \ (Since\ \pi_H \leq \frac{1}{2})$$
$$\geq 4c_H^2 - c_H^2 - c_H^2\ \ (Since\ c_L \leq c_H)$$
$$= 2c_H^2 \geq 0$$

For the specialists:

$$\forall k>0,\ q_k^{*P} - q_k^{*U} = 2npc_L^3\frac{4(\frac{1}{\pi_H}-1)c_Lc_H + (\frac{1}{\pi_H}-\pi_H)c_L^2}{(c_L+2c_H)^2(\frac{2c_Lc_H}{\pi_H}+c_L^2)^2} > 0$$

To check the fairness property, we compare the specialist's royalty under two rules:

$$\forall k > 0, R_k^{*P} - R_k^{*U} = np\frac{(1-\pi_H)c_L^2}{(c_L + 2\frac{c_H}{\pi_H})(c_L + 2c_H)}$$

Since $\pi_H \leq \frac{1}{2}$ and $c_H \geq c_L > 0$, the gap is always positive. □

References

Alaei, S., Makhdoumi, A., Malekian, A., Pekeč, S., 2020. Revenue-Sharing Allocation Strategies for Two-Sided Media Platforms: Pro-Rata Versus User-Centric. Rotman School of Management Working Paper 3645521.

Bergantinos, G., Moreno-Ternero, J.D., 2020. Sharing the revenues from broadcasting sport events. Manage. Sci. 66 (6), 2417–2431.

Dimont, J., 2018. Royalty inequity: Why music streaming services should switch to a per-subscriber model. Hast. Law J. 69 (2), 675–700.

Dredge, S., 2021. Deezer steps up its efforts to introduce user- centric payments. https://musically.com/2019/09/11/deezer-steps-up-its-efforts-to-introduce-user-centric-payments/.

Flores-Szwagrzak, K., Treibich, R., 2020. Teamwork and individual productivity. Manage. Sci. 66 (6), 2523–2544.

Ginsburgh, V., Zang, I., 2003. The museum pass game and its value. Games Econom. Behav. 43 (2), 322–325.

Hesmondhalgh, D., 2021. Is music streaming bad for musicians? Problems of evidence and argument. New Media Soc. 23 (12), 3593–3615.

Muikku, J., 2017. Pro rata and user centric distribution models: a comparative study. http://www.muusikkojenliitto.fi/wp-content/uploads/2018/02/UC_report_FINAL-2018.pdf.

Page, W., Safir, D., 2018a. Money in, money out: Lessons from CMOs in allocating and distributing licensing revenue. http://www.serci.org/congress_documents/2018/money_in_money_out.pdf.

Page, W., Safir, D., 2018b. 'User-centric' revisited: The unintended consequences of royalty distribution. http://www.serci.org/congress_documents/2019/user_centric_revisited.pdf.

Pedersen, R.R., 2014. A meta study of user-centric distribution for music streaming. https://www.koda.dk/media/224782/meta-study-of-user-centric-distribution-model-for-music-streaming.pdf.

Shiller, B., Waldfogel, J., 2013. The challenge of revenue sharing with bundled pricing: An application to music. Econ. Inq. 51 (2), 1155–1165.

The Digital, Culture, Media and Sport Committee, 2021. Economics of music streaming. https://committees.parliament.uk/publications/6739/documents/72525/default/.

Chapter 3

Optimal Queue to Minimize Waste

Preface

Chapter 3 shifts the focus to a new application area—queue management—while maintaining the overarching theme of strategic decision-making in environments with competing incentives and informational constraints. This chapter applies the lessons from Chapters 1 and 2 on optimizing outcomes in strategic settings to the problem of waste minimization in service systems. Here, service providers must balance the competing objectives of maintaining fairness and reducing waste.

By modeling an optimal policy for waste minimization in queuing allocation, this chapter extends the thesis's examination of competitive environments to include public and private sectors where resource allocation is critical. The framework developed here connects to the incentive design issues highlighted in Chapter 2, as the queueing policies and prioritization rules influence agent actions and system-wide expected waste of resource. The theme of managing trade-offs for optimal outcomes—whether for maximizing marketing income, remuneration, or resource efficiency—ties together the chapters, underscoring a consistent narrative that explores how economic agents navigate complex incentive structures in diverse settings. This final application reinforces the thesis's contribution to understanding strategic adaptation in environments shaped by information and incentive constraints.

Mathematical Social Sciences 123 (2023) 87–94

Contents lists available at ScienceDirect

Mathematical Social Sciences

journal homepage: www.elsevier.com/locate/mss

Optimal queue to minimize waste

Xiaochang Lei

Adam Smith Business School, University of Glasgow, Glasgow, United Kingdom

ARTICLE INFO

Article history:
Received 10 September 2022
Received in revised form 3 March 2023
Accepted 3 March 2023
Available online 11 March 2023

Keywords:
Dynamic matching
Waste minimization
Fair queue

ABSTRACT

We study an application of stochastic games in the dynamic allocation of two types of goods when agents have deferral rights. If all individuals strictly prefer one good to the other, the worse good can be wasted by successive rejections. We allow different goods to be allocated in different ways and study the combinations of three popular disciplines in an overloaded waiting list: FCFS (first-come-first-serve), LCFS(last-come-first-serve) and RP(random-priority). The first result is that the LCFS–FCFS queue (the better good allocated under LCFS and the worse good allocated under FCFS) does result in zero waste, but it is unfair. To restore fairness, the agent's age matters and the older agent has a weakly higher probability of receiving goods. Our second result is that RP–FCFS is fair and induces less expected waste than FCFS when the waiting cost is uniformly distributed.

1. Introduction

Using a queue is very popular in dynamic matching, especially when there is a supply shortage and a lack of a price mechanism (e.g. allocating donated food, organ, public house, etc.). In this paper, we consider how the social planner can reduce the waste of free goods. For example, the planner allocates two types of vaccines (A and B). Suppose scientists have announced that A has milder side effects and is more powerful than B. Then, B is wasted when all individuals in the queue reject it and prefer to wait for A. Then, it comes to the question: What is the optimal queue discipline to minimize the expected waste? Is the optimal queue feasible in reality?

The intuition for finding the optimal queue to minimize waste is simple. There are two reasons for agents' rejections. 1. A offers them a strictly higher utility than B. 2. Although waiting is costly, their expected waiting times for A are short enough to make rejections more attractive. Since goods' utilities are fixed, the only way to reduce the expected waste is to increase the agents' expected waiting times, thus reducing the probability of rejection and letting them be less selective. As a result, the optimal queue discipline should induce the longest expected waiting time. Besides that, it also needs to match some social norms to be feasible in reality. A pilot study on queue fairness is Larson (1987), which shows from a psychological perspective that an agent is more willing to join the queue if the front agents have a relatively smaller waiting time. So, in this paper, we assume that a longer waiting time in the queue must correspond to a weakly higher probability of receiving goods. We will define it in the model.

Our main contribution to the literature on dynamic allocation in queuing systems is introducing the complex queue disciplines, which allow different goods to be allocated differently. Otherwise, if both goods are allocated similarly, we call it a simple queue discipline. For example, a simple FCFS queue means both A and B are allocated under FCFS, while a complex RP–FCFS queue means that A is allocated under RP, but B is allocated under FCFS. The complex queue is allowed since we assume that the good's type is common knowledge after its realization. We show that LCFS–FCFS can result in zero expected waste when there are at least two agents since the first agent has an infinite waiting time for A. However, this queue is not feasible since LCFS will cause reneging. To restore fairness, we establish a criterion that the probability of receiving A can only (weakly) decrease on positions in the queue. We will show the intuition that with fairness guaranteed, the best the planner can do is RP–FCFS. Also, we prove that RP–FCFS is better than FCFS when the waiting cost is uniformly distributed.

We first establish how agents act in the simple FCFS overloaded queue. We show there is a rejection threshold of the agent's private waiting cost at each position. An agent at a specific position will reject B if his private waiting cost is below the corresponding threshold. Both LCFS–FCFS and RP–FCFS queues dramatically increase the expected waiting times for all agents, thus reducing the thresholds and rejection areas at all positions (See Fig. 1).

Besides the rejection probability, we also need to find the steady-state expected waste. This is not straightforward since agents' past rejections can change the probability of waste in the future. Under an overloaded queue, if an agent rejects B, his expected waiting time for A can only weakly decrease in the future. So, whenever the planner observes a rejection, he knows the corresponding agent will still reject it whenever he gets an offer B (See Example 1).

E-mail address: Xiaochang.Lei@glasgow.ac.uk.

https://doi.org/10.1016/j.mathsocsci.2023.03.001

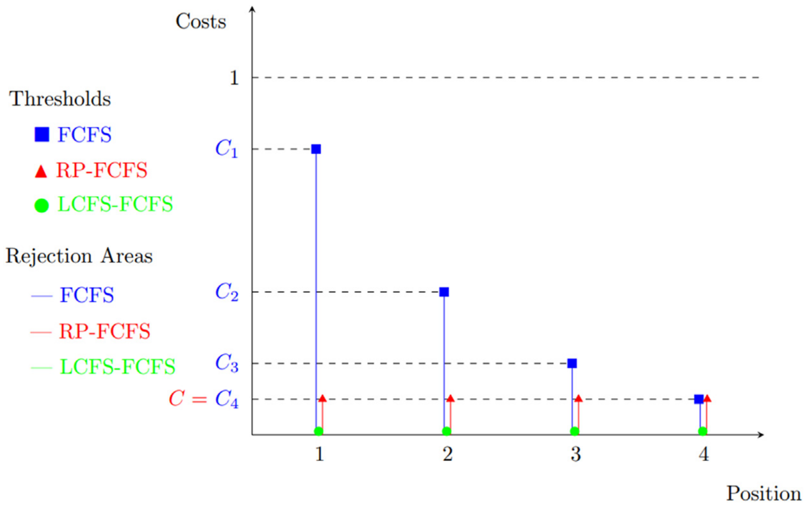

Fig. 1. Thresholds and rejection areas under different disciplines with fixed-length 4.

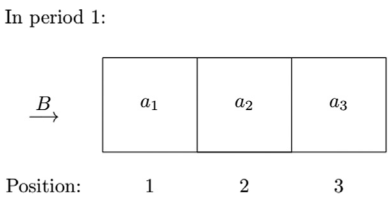

Fig. 2. Waste depends on the actions of all agents in the queue.

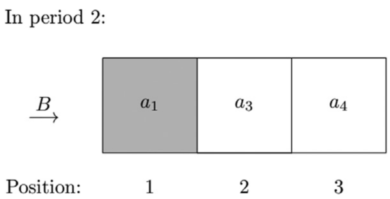

Fig. 3. Waste only depends on the actions of the last two agents, since an rejection is observed.

Example 1. Given a simple FCFS queue with 3 agents: $\{a_1, a_2, a_3\}$, we assume the production realization in period 1 and 2 is $\{B, B\}$. Suppose when offered with B, a_1 will reject it, but a_2 will accept it. The planner does not know the agents' actions at the beginning. Before the allocation in period 1, waste probability depends on the actions of all three agents. The allocation finishes after a_1 rejects B; a_2 accepts B and leaves the queue; a_3 moves up one position; and a new agent a_4 is born at position 3. At the beginning of period 2, the new queue is $\{a_1, a_3, a_4\}$. Now, the planner knows a_1 will still reject B. So, waste probability only depends on the decisions of a_3 and a_4. Here, the probability of waste in period 2 is different from that in period 1 due to an observed rejection of agent a_1 (See Figs. 2 and 3). □

The structure of the paper is as follows. We first list some relevant papers. In Section 2, we construct the benchmark model when the allocation of B is fixed under FCFS. We first capture the model in a simple FCFS queue and then show how the planner can control the discipline and explain why LCFS–FCFS can result in zero waste. After that, we introduce the fairness criterion and show the intuition that RP–FCFS is fair and waste-minimizing. In Section 3, we model the evolution of observed rejections in a one-step transition matrix and find expected waste. We show that RP–FCFS can induce lower expected waste than FCFS when the waiting cost is uniformly distributed. In the last section, we conclude with the limitations and contributions.

1.1. Literature review

There is a huge amount of literature from operations research on dynamic matching in queuing systems. A detailed review is Ashlagi and Roth (2021). Here, we just list recent papers that incorporate agents' dynamic tradeoffs. The most relevant research to our paper is Bloch and Cantala (2017). They analyze the welfare and waste in a constant size overloaded probabilistic queue when agents' have heterogeneous or homogeneous valuations. They show that FCFS is Pareto-superior to the lottery but can generate more expected waste. The main difference in settings is that goods in different periods are independent in their setting, while we assume the goods are the same if they belong to the same type. The difference results in a much more difficult ex-ante waste expression in our model.

Su and Zenios (2004, 2005) analyze the effects of offer rejection in $M/M/1$ dynamic kidney transplant. They compare FCFS and LCFS queues and show that FCFS makes agents more selective and induces a higher organ discard rate. By contrast, LCFS can maximize the expected life years, but it is practically infeasible. Our model also has the same intuition, and we innovatively combine different queue disciplines and make one step forward to find the optimal discipline when fairness is restored.

Leshno (2022) investigates dynamic allocation in minimizing misallocation under thresholds strategies. The way to reach that is similar to ours: to let the agents be less selective. He introduces a Loaded Independent Expected Waits (LIEW) queue, which can balance the expected waiting time for all agents. Compared with FCFS, the LIEW queue sacrifices the front agents and benefits the agents in the end. In his model, the agents have heterogeneous preferences but a homogeneous waiting cost. Instead, we model the agents with homogeneous preferences and heterogeneous waiting costs. We aim to minimize the expected waste and show that under the RP–FCFS queue, all agents' expected waiting times will increase.

Baccara et al. (2020) studies bilateral dynamic matching in general queuing systems. They aim to capture the utilitarian welfare maximizing mechanism by the number of remaining agents. Arnosti and Shi (2020) compares matching welfare and quality under different versions of lotteries and waiting lists in dynamic matching. Schummer (2021) discusses whether the social planner should give deferral rights to the agents on the

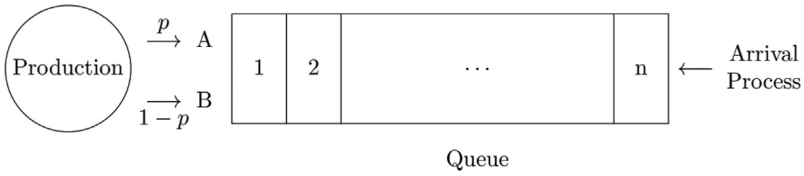

Fig. 4. The simple FCFS queue.

FCFS waiting list. He finds that the agents' welfare depends on the type of agents' preferences. Instead, we focus on how the planner should design the optimal mechanism when the deferral rights have been given to the agents. Thakral (2016, 2019) analyzes queuing systems from an axiomatic view. He discusses strategyproofness, efficiency and envy-freeness in different queues.

2. Benchmark model

One good is produced in each discrete period and has to be allocated within that period. Otherwise, it is wasted. Before realization, it can be one of two types: A with probability p or B with probability $1-p$. After realization, the type is known by all agents. A batch of (at least 2) agents forms the initial queue with length n. Each agent waits to get one good and then leaves. The agents have a homogeneous preference: $A \succ B$ (e.g., vaccine A has higher efficacy than B). They all get instantaneous utility 1 from accepting A or instantaneous utility u $(0 < u < 1)$ from accepting B. Agents have deferral rights, which means that when an agent is offered the worse object B, he can reject the offer, keep his position in the queue, and wait for the better good A. However, waiting is costly. A cost is subtracted from his utility if an agent stays in the queue for one more period. Each agent i has his private waiting cost c_i. Costs are i.i.d. distributed on $(0, 1)$ with CDF $F(\cdot)$ and will linearly decrease agents' utility. We assume that all agents' reservation values are sufficiently low (e.g. $-\infty$) so that no one will opt out. For example, no one will leave the queue for a vaccine and put himself at high risk of death. This means that when an agent's utility is zero or even negative, he will still stay until he gets a good.

We follow the tradition of calling that the position i is higher than the position j if $i < j$. When an agent accepts an offer, he leaves the queue, and all agents positioned behind will move one step forward. Also, a new agent is born at the last position. If no one accepts the realized item, the good is wasted, and no new agent is born in this period. This arrival process guarantees that the length of the queue is always n. The only private information is the waiting cost of each agent. Initially, the planner only knows the distribution of waiting costs $F(\cdot)$ and aims to find a discipline to minimize the steady-state expected waste.

2.1. Agents' strategies under the simple FCFS queue

We first show the story under the benchmark simple FCFS queue (See Fig. 4). An agent offered A will accept it since waiting is costly, and there is no better offer in the future. So, in the simple FCFS queue, when A is realized, it can only be offered to and accepted by the agent at position 1. However, an agent faces a binary choice when offered B: accept or reject it. Given the fixed instantaneous utilities of the two goods, the decision depends on his expected waiting time and private waiting cost. The intuition is that under simple FCFS, an agent positioned ahead faces a shorter expected waiting time, so he is more likely to be selective. Also, a more patient agent is more likely to reject B since waiting is not a big deal for him. The formal expression is that an agent i at position k faces an optimal stopping problem when offered B: If he accepts B now, he gets utility u. If he rejects B, he gets expected utility $1 - c_i w_k$. w_k is the expected waiting time for A at position k.

The first agent's expected waiting time for A follows a geometric distribution with parameter p. So, the expected waiting time of the first agent is $1/p$. Considering the agent at position k, when offered with B, he knows all front agents have already rejected B. Otherwise, B must have been accepted by one of them. He can also infer that they will reject B whenever B is realized since their expected waiting time for A can only weakly decrease. As a result, if he rejects B now, he can only get A after all front agents have been served with A. So, the expected waiting time of an agent at position k is k/p. We know rejection happens only when $1 - c_i w_k \geq u$. Given the expression $w_k = \frac{k}{p}$, we can find a rejection range for agent i's private waiting cost at position k: $c_i \leq (1-u)p/k$. Let $C_k = \frac{(1-u)p}{k}$ denote the threshold at position k. For any agent at position k, he will reject B if his private waiting cost is below the threshold C_k. Although the waiting cost is private, the threshold is common knowledge. C_k is a decreasing function on position k, which means that a more patient agent is required to reject B at a lower position (see Fig. 5).

After finding the rejection thresholds, the probability of waste can be easily found. Let i denote the number of agents who have rejected B, the probability of waste is:

$$PW_i^{FCFS} = \Pi_{k=i+1}^{n} F(C_k)$$

Next, we show how LCFS–FCFS and RP–FCFS can reduce the rejection thresholds, thus reducing the expected waste.

2.2. Waste minimization

We assume that the planner can arbitrarily control the probability φ_k that the good A is allocated to an agent at position k when it is realized. However, for simplicity, the way of B's allocation still follows the FCFS (See Fig. 6). For example: FCFS queue of A is captured by: $\varphi_1 = 1$ and $\forall 2 \leq k \leq n$, $\varphi_k = 0$. LCFS queue of A is captured by: $\varphi_n = 1$ and $\forall 1 \leq k \leq n-1$, $\varphi_k = 0$. RP of A is captured by $\forall 1 \leq k \leq n$, $\varphi_k = \frac{1}{n}$. Let $\boldsymbol{\varphi} = \{\varphi_1, \varphi_2, \ldots, \varphi_n\}$ denote the set of probabilities, we have: $\sum_{k=1}^{n} \varphi_k = 1$.

We fix the allocation of B in FCFS since it is difficult to capture how B can be wasted once we allow randomization in the allocation of B. The planner has to record which agent has rejected it. Under randomization, these agents can be separately positioned, which results in a complicated expression for the probability of waste. Also, adding randomization of A has already changed the expected waiting time of all agents. Under this setting, the planner aims to find an optimal $\boldsymbol{\varphi}$ to minimize the expected waste.

Theorem 1. *LCFS–FCFS induces zero expected waste, but it is unfair.*

Intuitively, the LCFS–FCFS queue does minimize the expected waste since the expected waiting times for A of the agents are infinite (except the last one). Suppose the agent at position 1 rejects B. Since A is allocated under LCFS, it will be allocated to the agent at position n whenever it is realized. After getting A, the agent at position n will leave the queue, and a new agent will join the queue at position n. So, the agent 1 will never get A. As

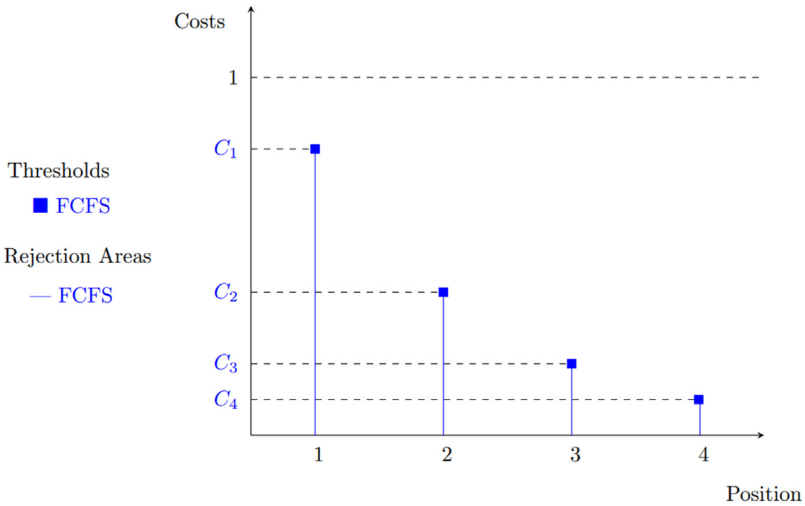

Fig. 5. Thresholds and rejection areas under simple FCFS queue with fixed-length 4.

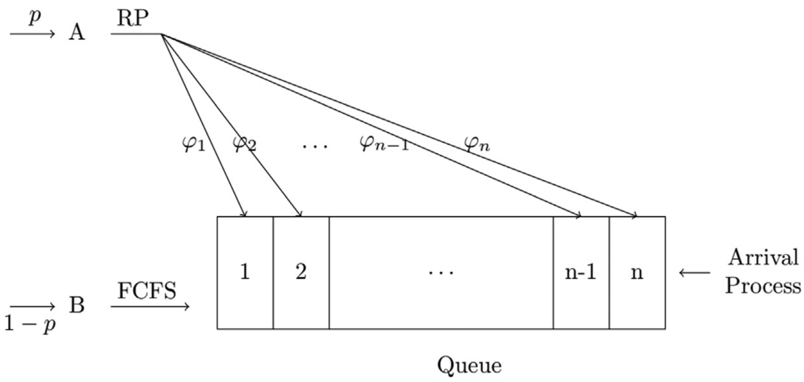

Fig. 6. The general complex queue discipline.

a result, the first agent will always accept B, even if he is very patient. This implies there is no waste.

Since the expected waiting times for the front agents are infinite under LCFS–FCFS, the corresponding rejection thresholds are zero for the front agents. Also, the last agent will never get an offer B, so he has no rejection area (See Fig. 1). However, LCFS is unfair to the front agents, who should be rewarded for their long waiting time in the queue. Despite that, it gives us an intuition that waste can be reduced if the planner makes φ_k of the agents at the lower positions (larger k) as high as possible to increase the expected waiting time of the agents positioned ahead. To restore fairness, we assume that probability φ_k must satisfy: $\varphi_1 \geq \varphi_2 \geq \cdots \geq \varphi_n$. So, every agent is not treated worse than anyone positioned behind him. This gives a range of φ_n: $0 \leq \varphi_n \leq \frac{1}{n}$. From the LCFS–FCFS queue, we know that to minimize the waste, φ_n should be as large as possible ($\varphi_n = \frac{1}{n}$). So, RP–FCFS is both fair and waste-minimizing.

2.3. Agents' strategies under RP–FCFS

Since the allocation of B still follows FCFS, once B is offered to an agent i at position k, he still knows that all agents positioned ahead have rejected B and will reject B in the future. Good A's probability of realization is p, and the probability of getting A after realization is $1/n$, which does not depend on his position. So, the probability that an agent is offered with A is p/n, and his expected waiting time for A is n/p. The agent i will reject B if $1 - c_i\frac{n}{p} \geq u$, else, he will accept B. Since the expected waiting time is independent of position, there is a uniform threshold under RP–FCFS for all positions: $C = \frac{(1-u)p}{n}$. Comparing the uniform threshold under RP–FCFS with the thresholds under simple FCFS, we find that RP–FCFS weakly reduces the thresholds for all positions (See Fig. 7).

After finding the rejection thresholds, the probability of waste can be easily found. Given the number of observed rejections i, the probability of waste is:

$$PW_i^{RP-FCFS} = \Pi_{k=i+1}^{n} F(C) = F(C)^{n-i}$$

We know that the threshold under RP–FCFS is always below the thresholds under simple FCFS: $\forall k, C \leq C_k$. So, for any realization of the number of observed rejections i, the probability of waste under RP–FCFS is less than the probability of waste under simple FCFS: $\forall i \leq n, PW_i^{RP-FCFS} \leq PW_i^{FCFS}$.

Proposition 1 (*Ex-post Improvement*)**.** *For any number of observed rejections, RP–FCFS induces less probability of waste than FCFS.*

3. Steady state

3.1. Steady state under simple FCFS

As mentioned above, the agent at position k will reject B if his waiting cost is below C_k. Else, he will choose to accept B immediately. Although we assume that initially, the planner only knows the distribution of waiting costs, he can infer its range from observing the agent's decision. When an acceptance is observed, the planner knows that the agent's waiting cost exceeds

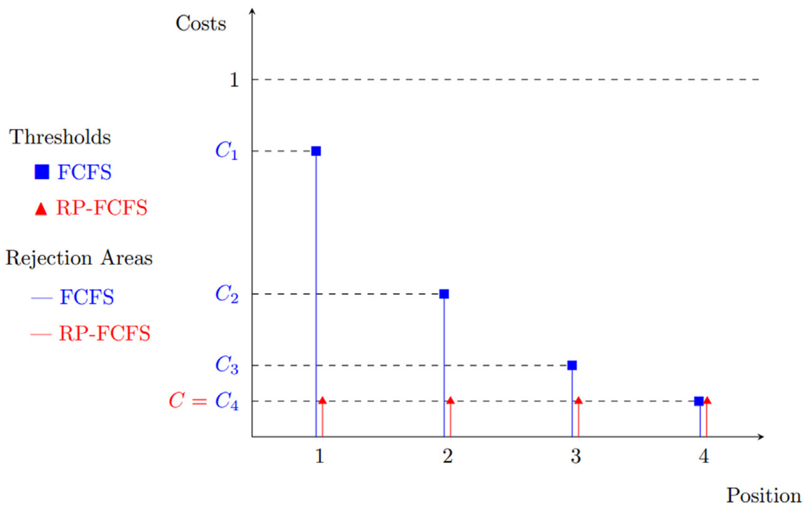

Fig. 7. Thresholds and rejection areas in different queue disciplines with fixed-length 4.

his position's threshold. When the planner observes rejection, the agent's waiting cost is below his position's threshold. This information disclosure process is vital in calculating the probability of waste. For example, when B is realized if the planner observes rejection at position k ($1 \leq k \leq n-1$), he knows the agent (at position) k's private waiting cost is below C_k. Also, he knows all front agents' ($1 \sim k-1$) private waiting costs are below their corresponding threshold since they must have already rejected B under FCFS. So, the probability of waste is the probability of successive rejection for the agents behind: $\Pi_{i=k+1}^{n} F(C_i)$. If $k = n$, the expected waste is 1. Since when all agents reject B, B is automatically wasted after realization.

Now, we use a Markov chain to capture this information disclosure process. In period t, the number of agents who have rejected B, is a stochastic process $\{X^{(t)}, t = 1, 2, 3, \ldots\}$ with a finite state space $M = \{0, 1, 2, \ldots, n\}$. We define the one-step transition probability as $P_{ij} = P(X^{(t+1)} = j | X^{(t)} = i)$ with $0 \leq P_{ij} \leq 1$ and $\sum_{j \in M} P_{ij} = 1, i = 1, 2, 3, \ldots$, since any state in period t must transit into a state in period $t+1$. Let π^t denote the distribution of $X^{(t)}$ and $\boldsymbol{P}$ denote the one-step transition matrix, we have:

$$\boldsymbol{P}\boldsymbol{\pi}^t = \boldsymbol{\pi}^{t+1}$$

The aim is to find the stationary distribution $\boldsymbol{\pi} = \{\pi_0, \pi_1, \pi_2, \ldots, \pi_n\}$ of the number of information-disclosed positions. The linear system is:

$$\boldsymbol{P}\boldsymbol{\pi} = \boldsymbol{\pi}$$
$$\boldsymbol{e}\boldsymbol{\pi} = 1$$

$\boldsymbol{e}$ is a vector of $\{1, 1, 1 \cdots, 1\}$. By solving the system of equations, the steady state can be found.

Proposition 2. *In the steady state of the simple FCFS queue, $\boldsymbol{\pi}$ is:*

$$\pi_k^{FCFS} = \frac{Q_k}{\sum_{i=0}^{n} Q_i}, k = 0, 1, 2, \ldots, n$$

where $\forall k \geq 1$, $Q_k = \frac{1-p}{p^k} \Pi_{i=1}^{k} F(\frac{(1-u)p}{i})$, and $Q_0 = 1$.

Q_k is the coefficient of π_0 in the equation $\pi_k = Q_k \pi_0$. Obviously, Q_k is decreasing on k, so π_k^{FCFS} is also decreasing. The intuition is simple: one more rejection is less possible. The probability that The expected waste under FCFS queue is:

$$EW^{FCFS} = \sum_{k=0}^{n} \pi_k^{FCFS} * PW_k^{FCFS}$$

$$= \frac{\Pi_{i=1}^{n} F(\frac{(1-u)p}{i})(1 + \frac{1-p}{p} + \frac{1-p}{p^2} + \cdots + \frac{1-p}{p^n})}{1 + \frac{1-p}{p} F((1-u)p) + \frac{1-p}{p^2} \Pi_{i=1}^{2} F(\frac{(1-u)p}{i}) + \cdots + \frac{1-p}{p^n} \Pi_{i=1}^{n} F(\frac{(1-u)p}{i})}$$

3.2. Steady state under RP–FCFS

Under RP–FCFS, the information disclosure process is similar to the simple FCFS queue above. To find the steady-state expected waste, we only need to know the distribution of the disclosed information. Again, we can solve it by using a Markov chain.

Proposition 3. *In the steady state of the RP–FCFS queue, the number of observed rejections is distributed as:*

$$\pi_k^{RP-FCFS} = \frac{R_k}{\sum_{i=0}^{n} R_i}, \forall k = 0, 1, 2, \ldots, n$$

where $\forall k \geq 2$, $R_k = \Pi_{i=1}^{k-1}(n - (n-i)p) n \frac{1-p}{k! p^k} F(\frac{(1-u)p}{n})^k$, and $R_1 = n \frac{1-p}{p} F(\frac{(1-u)p}{n})$, and $R_0 = 1$.

R_k is the coefficient of π_0 in the equation $\pi_k = R_k \pi_0$. Obviously, R_k is decreasing on k, so $\pi_k^{RP-FCFS}$ is also decreasing. Under uniform distribution $Q_1 = R_1$ and $\forall k \geq 2$, $Q_k \geq R_k$. This means π_k can decrease faster under the RP–FCFS queue. The expected waste is:

$$EW^{RP-FCFS} = \sum_{k=0}^{n} \pi_k^{RP-FCFS} * PW_k^{RP-FCFS}$$

$$= \frac{F(\frac{(1-u)p}{n})^n (1 + n\frac{1-p}{p} + \sum_{k=2}^{n} \Pi_{i=1}^{k-1} \frac{(n-(n-i)p)n}{k!} \frac{1-p}{p^k})}{1 + n\frac{1-p}{p} F(\frac{(1-u)p}{n}) + \sum_{k=2}^{n} \Pi_{i=1}^{k-1}(n - (n-i)p) n \frac{1-p}{k! p^k} F(\frac{(1-u)p}{n})^k}$$

The direct comparison between the two expected wastes is intractable, let alone proof of RP–FCFS's optimality. We only find a tractable comparison when the waiting cost is uniformly distributed.

Theorem 2 (*Ex-ante Improvement*). *When $F(x) = x$, RP–FCFS induces less expected waste than FCFS.*

Theorem 2 is much stronger than Proposition 1. The ex-post improvement does not necessarily implies ex-ante improvement since the expected waste also depends on the stationary distribution of the number of observed rejections.

4. Conclusion

This paper investigates the waste minimization problem in dynamic queuing allocation. We find that the expected waiting time for the better good should be maximized to decrease the probability of rejection, thus reducing the steady-state expected waste. Our main theoretical contribution is to allow different goods to be allocated under different queue disciplines. This can be achieved when the goods' type is public information after its realization. Our first result is that LCFS–FCFS generates zero waste by inducing infinite expected waiting times for the front agents. However, it is unfair and impossible to be used in reality. A basic fairness criterion is that the agent must be prioritized (weakly) higher than anyone behind him. In other words, he should have a higher probability of receiving goods than the agents arriving later. Our second result is that RP–FCFS is both ex-post and ex-ante better than FCFS when fairness is restored.

While the paper only discusses the fixed-length deterministic queue, our results can be easily extended into M/M/1 environment. Since, given any queue's length, the rejection thresholds under RP–FCFS are always below the thresholds under simple FCFS. The only difference under M/M/1 environment is that the queue length can change. So, there are two elements in the one-step transition matrix under M/M/1: 1. The number of observed rejections. 2. The queue's length. Also, the expected waiting time for A of the agents under RP–FCFS will change when the queue's length changes. Despite the differences, the intuition on waste minimization is the same as in the fixed-length deterministic environment. Also, the waiting cost in our model is different from the discount factor. The discount factor exponentially decreases the utility, while the waiting cost linearly decreases the utility. We adopt the waiting cost since it is more mathematically tractable. There is not much difference in the main results when using the discount factor.

The limitation of this paper is that while the intuition of RP–FCFS's optimality is easy, the proof is intractable. Also, we mainly discuss the complex queues under which $B's$ allocation is fixed under FCFS. There are other possible combinations (e.g., LCFS–RP and FCFS–RP). Although this direction of extension is interesting, the information disclosure process is hard to capture when the discipline of B moves away from the FCFS. The main reason is that the agents need to know which agent in front of them has rejected B before calculating their expected waiting times. So, adding randomization in the allocation of B will complicate the story.

Data availability

No data was used for the research described in the article.

Appendix. Proof of theorems and propositions

Proof of Theorem 1. Under LCFS–FCFS, for any $n > 2$, suppose at period t, B is realized. If the agent at position 1 rejects it, his expected waiting time is $-\infty$. $\forall c_1 \in (0, 1)$, his expected utility under rejection is $1 - \infty * c_1 = -\infty$. His expected utility under acceptance is $u \in (0, 1)$. So, he will accept B. This means no rejection will happen, and there is no waste. □

Proof of Theorem 2. If $F(x) = x$, the expected wastes are:

$$EW^{FCFS} = \frac{\frac{(1-u)^n p^n}{n!}(1 + \frac{1-p}{p} + \frac{1-p}{p^2} + \cdots + \frac{1-p}{p^n})}{1 + (1-p)(1-u) + \frac{(1-p)(1-u)^2}{2!} + \cdots + \frac{(1-p)(1-u)^n}{n!}}$$

$$EW^{RP-FCFS} = \frac{\frac{(1-u)^n p^n}{n^n}(1 + n\frac{1-p}{p} + \sum_{k=2}^n \Pi_{i=1}^{k-1} \frac{(n-(n-i)p)n}{k!} \frac{1-p}{p^k})}{1 + (1-p)(1-u) + \sum_{k=2}^n \frac{\Pi_{i=1}^{k-1}(n-(n-i)p)n}{n^k} \frac{(1-p)(1-u)^k}{k!}}$$

When $n = 2$, the above expressions are reduced to:

$$EW^{FCFS} = \frac{(1-u)^2 p^2 (\frac{1}{2} + \frac{1}{2}\frac{1-p}{p} + \frac{1}{2}\frac{1-p}{p^2})}{1 + (1-p)(1-u) + \frac{(1-p)(1-u)^2}{2!}}$$

$$EW^{RP-FCFS} = \frac{(1-u)^2 p^2 (\frac{1}{4} + \frac{1}{2}\frac{1-p}{p} + \frac{(2-p)}{2}\frac{1}{2}\frac{1-p}{p^2})}{1 + (1-p)(1-u) + \frac{(2-p)}{2}\frac{(1-p)(1-u)^2}{2!}}$$

Since $p > 0$, $\frac{2-p}{2} < 1$, $EW^{FCFS} > EW^{RP-FCFS}$.

Suppose when $n = m - 1$, $EW^{FCFS} > EW^{RP-FCFS}$. Then, when $n = m$, we have equation given in Box I

Proof of Proposition 2. The one-step transition probability needs to be presented separately for different situations: If $0 < i < n$, conditional on $X^{(t)} = i$, $X^{(t+1)}$ can be $i - 1, i, i + 1, \ldots, n$:

$$P(X^{(t+1)} = j | X^{(t)} = i) = \begin{cases} p & \text{if } j = i - 1 \\ (1-p)(1 - F(C_{j+1})) & \text{if } j = i \\ (1-p)(1 - F(C_{j+1}))\Pi_{k=i+1}^{j} F(C_k) & \text{if } i + 1 \le j \le n - 1 \\ (1-p)\Pi_{k=i+1}^{j} F(C_k) & \text{if } j = n \end{cases}$$

If $i = n$,

$$P(X^{(t+1)} = j | X^{(t)} = n) = \begin{cases} p & \text{if } j = n - 1 \\ (1-p) & \text{if } j = n \end{cases}$$

If $i = 0$, conditional on $X^{(t)} = i$, $X^{(t+1)}$ can be $i, i + 1, \ldots, n$:

$$P(X^{(t+1)} = j | X^{(t)} = 0) = \begin{cases} p + (1-p)(1 - F(C_1)) & \text{if } j = i = 0 \\ (1-p)(1 - F(C_{j+1}))\Pi_{k=i+1}^{j} F(C_k) & \text{if } 0 < j < n \\ (1-p)\Pi_{k=i+1}^{j} F(C_k) & \text{if } j = n \end{cases}$$

Then, we have a system of equations:

$$\pi_0 = \pi_0(p + (1-p)(1 - F(C_1))) + \pi_1 p$$

$$\pi_k = \sum_{i=0}^{k-1} \pi_i (1-p)(1 - F(C_{k+1}))\Pi_{j=i+1}^{k} F(C_j) + \pi_k(1-p)(1 - F(C_{k+1})) + \pi_{k+1} p, \quad \forall 1 \le k \le n - 1$$

$$\pi_n = \sum_{i=0}^{n-1} \pi_i (1-p)\Pi_{j=i+1}^{n} F(C_j) + \pi_n(1-p)$$

We can derive:

$$\pi_1 = \frac{1-p}{p} F(C_1)\pi_0$$

$$\pi_2 = \frac{1-p}{p^2} F(C_1)F(C_2)\pi_0$$

$$\pi_3 = \frac{1-p}{p^3} F(C_1)F(C_2)F(C_3)\pi_0$$

...

$$\pi_n = \frac{1-p}{p^n} \Pi_{i=1}^{n} F(C_i)\pi_0$$

$$
\begin{aligned}
EW^{FCFS} &= \frac{\frac{(1-u)^m p^m}{m!}\left(1+\frac{1-p}{p}+\frac{1-p}{p^2}+\cdots+\frac{1-p}{p^m}\right)}{1+(1-p)(1-u)+\frac{(1-p)(1-u)^2}{2!}+\cdots+\frac{(1-p)(1-u)^m}{m!}} \\
&= \frac{\frac{(1-u)p}{m}\left(\frac{(1-u)^{m-1}p^{m-1}}{(m-1)!}\left(1+\frac{1-p}{p}+\cdots+\frac{1-p}{p^{m-1}}\right)\right)+\frac{(1-p)(1-u)^m}{m!}}{1+(1-p)(1-u)+\frac{(1-p)(1-u)^2}{2!}+\cdots+\frac{(1-p)(1-u)^{m-1}}{(m-1)!}+\frac{(1-p)(1-u)^m}{m!}} \\
&> \frac{\left(\frac{(1-u)^m p^m}{m(m-1)^{m-1}}\left(1+\frac{(m-1)(1-p)}{p}+\sum_{k=2}^{m-1}\frac{\Pi_{i=1}^{k-1}((m-1)-(m-1-i)p)(m-1)}{k!}\frac{1-p}{p^k}\right)\right)+\frac{(1-p)(1-u)^m}{m!}}{1+(1-p)(1-u)+\sum_{k=2}^{m-1}\frac{\Pi_{i=1}^{k-1}((m-1)-((m-1)-i)p)(m-1)}{(m-1)^k}\frac{(1-p)(1-u)^k}{k!}+\frac{(1-p)(1-u)^m}{m!}} \\
&> \frac{\frac{(1-u)^m p^m}{m^m}\left(1+m\frac{1-p}{p}+\sum_{k=2}^{m-2}\Pi_{i=1}^{k-1}\frac{(m-(m-i)p)m}{k!}\frac{1-p}{p^k}\right)+\frac{(1-p)(1-u)^m}{m!}}{1+(1-p)(1-u)+\sum_{k=2}^{m-1}\frac{\Pi_{i=1}^{k-1}((m-1)-((m-1)-i)p)(m-1)}{(m-1)^k}\frac{(1-p)(1-u)^k}{k!}+\frac{(1-p)(1-u)^m}{m!}} \\
&= \frac{\frac{(1-u)^m p^m}{m^m}\left(1+m\frac{1-p}{p}+\sum_{k=2}^{m}\Pi_{i=1}^{k-1}\frac{(m-(m-i)p)m}{k!}\frac{1-p}{p^k}\right)+\frac{m^m}{m!}\frac{(1-p)}{p^m}}{1+(1-p)(1-u)+\sum_{k=2}^{m-1}\frac{\Pi_{i=1}^{k-1}((m-1)-((m-1)-i)p)(m-1)}{(m-1)^k}\frac{(1-p)(1-u)^k}{k!}+\frac{(1-p)(1-u)^m}{m!}} \\
&> \frac{\frac{(1-u)^m p^m}{m^m}\left(1+m\frac{1-p}{p}+\sum_{k=2}^{m}\Pi_{i=1}^{k-1}\frac{(m-(m-i)p)m}{k!}\frac{1-p}{p^k}\right)}{1+(1-p)(1-u)+\sum_{k=2}^{m}\frac{\Pi_{i=1}^{k-1}(m-(m-i)p)m}{m^k}\frac{(1-p)(1-u)^k}{k!}} = EW^{RP-FCFS} \quad \square
\end{aligned}
$$

Box I.

Given that $F(C_i) = F(\frac{(1-u)p}{i})$, so $\forall k > 0$:

$$
\begin{aligned}
\pi_k &= \frac{1-p}{p^k}\Pi_{i=1}^{k}F(C_i)\pi_0 \\
&= \frac{1-p}{p^k}\Pi_{i=1}^{k}F(\frac{(1-u)p}{i})\pi_0
\end{aligned}
$$

$\forall k \geq 1$, let $Q_k = \frac{1-p}{p^k}\Pi_{i=1}^{k}F(\frac{(1-u)p}{i})$, and $Q_0 = 1$. Since π is a probability distribution, then $\sum_{i=0}^{n} Q_i\pi_0 = 1$. So, we have the stationary distribution π:

$$\pi_k = \frac{Q_k}{\sum_{i=0}^{n} Q_i}, k = 0, 1, 2, \cdot, n \quad \square$$

Proof of Proposition 3. If $0 < i < n$, conditional on $X^{(t)} = i$, $X^{(t+1)}$ can be $i-1, i, i+1, \ldots, n$:

$$
\begin{aligned}
&P(X^{(t+1)} = j | X^{(t)} = i) \\
&= \begin{cases} p\frac{i}{n} & \text{if } j = i-1 \\ p\frac{n-i}{n} + (1-p)(1-F(C)) & \text{if } j = i \\ (1-p)(1-F(C))\Pi_{k=i+1}^{j}F(C) & \text{if } i+1 \leq j \leq n-1 \\ (1-p)\Pi_{k=i+1}^{n}F(C) & \text{if } j = n \end{cases}
\end{aligned}
$$

If $i = n$,

$$P(X^{(t+1)} = j | X^{(t)} = n) = \begin{cases} p & \text{if } j = n-1 \\ (1-p) & \text{if } j = n \end{cases}$$

If $i = 0$, conditional on $X^{(t)} = i$, $X^{(t+1)}$ can be $i, i+1, \ldots, n$:

$$
\begin{aligned}
&P(X^{(t+1)} = j | X^{(t)} = 0) \\
&= \begin{cases} p + (1-p)(1-F(C)) & \text{if } j = i = 0 \\ (1-p)(1-F(C))\Pi_{k=i+1}^{j}F(C) & \text{if } 0 < j < n \\ (1-p)\Pi_{k=i+1}^{j}F(C) & \text{if } j = n \end{cases}
\end{aligned}
$$

Then, we have a system of equations:

$$
\begin{aligned}
\pi_0 &= \pi_0(p + (1-p)(1-F(C))) + \pi_1\frac{p}{n} \\
\pi_k &= \sum_{i=0}^{k-1}\pi_i(1-p)(1-F(C))\Pi_{j=i+1}^{k}F(C) \\
&\quad + \pi_k(p\frac{n-k}{n} + (1-p)(1-F(C))) + \pi_{k+1}p\frac{k+1}{n}, \\
&\quad \forall 1 \leq k \leq n-1 \\
\pi_n &= \sum_{i=0}^{n-1}\pi_i(1-p)\Pi_{j=i+1}^{n}F(C) + \pi_n(1-p)
\end{aligned}
$$

We can derive:

$$
\begin{aligned}
\pi_1 &= n\frac{1-p}{p}F(C)\pi_0 \\
\pi_2 &= (n-(n-1)p)n\frac{1-p}{2!p^2}F(C)^2\pi_0 \\
\pi_3 &= (n-(n-2)p)(n-(n-1)p)n\frac{1-p}{3!p^3}F(C)^3\pi_0 \\
&\ldots \\
\pi_n &= \Pi_{i=1}^{n-1}(n-(n-i)p)n\frac{1-p}{n!p^n}F(C)^n\pi_0
\end{aligned}
$$

$\forall k \geq 2$, let $R_k = \Pi_{i=1}^{k-1}(n-(n-i)p)n\frac{1-p}{k!p^k}F(C)^k$, $R_1 = n\frac{1-p}{p}F(C)$ and $R_0 = 1$. We have the steady state distribution of X^t in RP–FCFS queue:

$$\pi_k = \frac{R_k}{\sum_{i=0}^{n} R_i}, k = 0, 1, 2\cdots, n \quad \square$$

References

Arnosti, N., Shi, P., 2020. Design of lotteries and wait-lists for affordable housing allocation. Manage. Sci. 66 (6), 2291–2307.

Ashlagi, I., Roth, A.E., 2021. Kidney exchange: an operations perspective. Manage. Sci. 67 (9), 5455–5478.

Baccara, M., Lee, S., Yariv, L., 2020. Optimal dynamic matching. Theor. Econ. 15 (3), 1221–1278.
Bloch, F., Cantala, D., 2017. Dynamic assignment of objects to queuing agents. Am. Econ. J. Microecon. 9 (1), 88–122.
Larson, R.C., 1987. OR forum—perspectives on queues: Social justice and the psychology of queueing. Oper. Res. 35 (6), 895–905.
Leshno, J.D., 2022. Dynamic matching in overloaded waiting lists. Amer. Econ. Rev. 112 (12), 3876–3910.
Schummer, J., 2021. Influencing waiting lists. J. Econom. Theory 195, 105263.
Su, X., Zenios, S., 2004. Patient choice in kidney allocation: The role of the queueing discipline. Manufacturing & Service Operations Management 6 (4), 280–301.
Su, X., Zenios, S., 2005. Patient choice in kidney allocation: A sequential stochastic assignment model. Oper. Res. 53 (3), 443–455.
Thakral, N., 2016. The public-housing allocation problem. Technical report, Harvard University.
Thakral, N., 2019. Matching with stochastic arrival. In: AEA Papers and Proceedings. vol. 109, pp. 209–212.

Conclusion

This thesis has explored Applied Microeconomic Theory across three interrelated areas: reputation and signaling dynamics, remuneration schemes in digital platforms, and efficient queue management. Each chapter addressed a distinct application, yet together, they advance our understanding of strategic interactions in environments characterized by information asymmetry, competing incentives, and decision-making under constraints. This concluding synthesis will highlight the theoretical contributions of the thesis as a unified project, acknowledge its limitations, and suggest avenues for future research.

A common theme across the chapters is the investigation of strategic behaviors in environments with asymmetric information and (intertemporal) trade-offs. Chapter 1's model of repeated signaling and reputation provides insight into how information asymmetry affects long-term interactions, particularly when one party (the customer) can only observe recent behavior. This assumption deviates from traditional reputation models by introducing bounded memory and signalling stage game, making it particularly applicable to real-world contexts where limited historical knowledge influences decisions, as seen in influencer marketing. This approach provides a more realistic account of reputation evolution, diverging from classical reputation models that assume perfect recall and thereby overestimate the potential for sustained cooperation or reputation enhancement.

In Chapter 2, the focus shifts to how strategic responses within an endogenous framework can shape equity and efficiency in digital platform markets. By modeling the music streaming industry as a competitive environment where artists adjust song quality to maximize their payoffs, the study challenges conventional beliefs about the equity of user-centric (U) versus pro-rata (P) remuneration models. This finding demonstrates the potential for strategic quality adjustments to shift the perceived fairness and efficiency of different payment schemes. These results inform ongoing policy debates, particularly those advocating for user-centric payment models as inherently fairer.

Chapter 3 further contributes to the literature on resource allocation and queue management, offering a novel framework to minimize waste in service systems by using complex queuing disciplines. The study's queue optimization model addresses a critical need in industries where Non-pricing allocation can result in significant waste of public goods, such as healthcare or food services. This model demonstrates that efficient resource allocation policies can substantially reduce waste while maintaining high service fairness. By integrating queueing theory with waste minimization principles, this chapter provides valuable insights for managers in both public and private sectors seeking to improve service delivery while minimizing costs.

The methodological contributions across the chapters reflect a versatile approach to modeling complex interactions and equilibria. Chapter 1's use of a Markov Decision Process (MDP) framework underpins a mixed-strategy equilibrium for influencer signaling that accounts for bounded memory. Chapter 2 employs comparative statics within an endogenous model, enabling nuanced assessments of quality and remuneration across different payment schemes. Chapter 3's queueing model blends social choice principles with optimization techniques to balance competing priorities of fairness and minimal waste. This diversity in methodologies allows for a comprehensive analysis across contexts, each benefiting from tailored approaches that address their unique structural characteristics. Together, these methodological choices underscore the thesis's contribution to Applied Microeconomic Theory, highlighting versatile frameworks that address dynamic behaviors in competitive environments.

While these contributions are significant, the thesis also has limitations that suggest potential directions for future research. In Chapter 1, the assumption of bounded memory in the reputation model captures realistic customer behavior but limits the analysis of long-term strategic equilibrium effects. Extending the model to include varying degrees of customer memory or even imperfect recall could yield insights into how different memory structures impact reputation dynamics and equilibria. Chapter 2's endogenous model assumes that artists adjust only song quality in response to remuneration rules, which simplifies the strategic space. Incorporating additional dimensions, such as marketing effort or song release frequency, could offer a more comprehensive view of how artists respond to competitive pressures in a platform economy. Also, the mathematical intractability limits our discussion with limited number of artists. Finally, Chapter 3's model for queue optimization assumes a single service type, which might limit applicability in settings where multiple services compete for prioritization. Expanding the model to a multi-queue or multi-service framework could enhance its relevance to complex service environments.

The synthesis of these chapters into a coherent whole is justified by their shared focus on strategic interactions within economically relevant contexts of asymmetric information, competitive incentives, and (intertemporal) decision-making. While each chapter addresses a specific application—signaling in reputation, remuneration in digital content, and queueing in service delivery—the common theoretical foundation unites them. Each chapter contributes to a broader understanding of how agents balance (immediate) payoffs with (long-term) incentives, adapting their strategies based on others' actions, beliefs, and the economic environment's structural characteristics. These insights are not only relevant within their respective fields but also underscore a broader theme of strategic adaptation in response to incentives, a fundamental concern in Applied Microeconomic Theory.

In conclusion, this thesis advances Applied Microeconomic Theory by developing models that capture (dynamic)incentive-driven behaviors across diverse yet interconnected contexts. By contributing theoretical insights relevant to real-world scenarios, the research highlights the importance of adapting traditional economic models to reflect the evolving realities of digital and service-based economies. Future research that builds on these findings will further enhance our understanding of complex economic interactions, helping policymakers and practitioners navigate the challenges of asymmetric information and strategic decision-making in increasingly interconnected markets.