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
0165-4896/

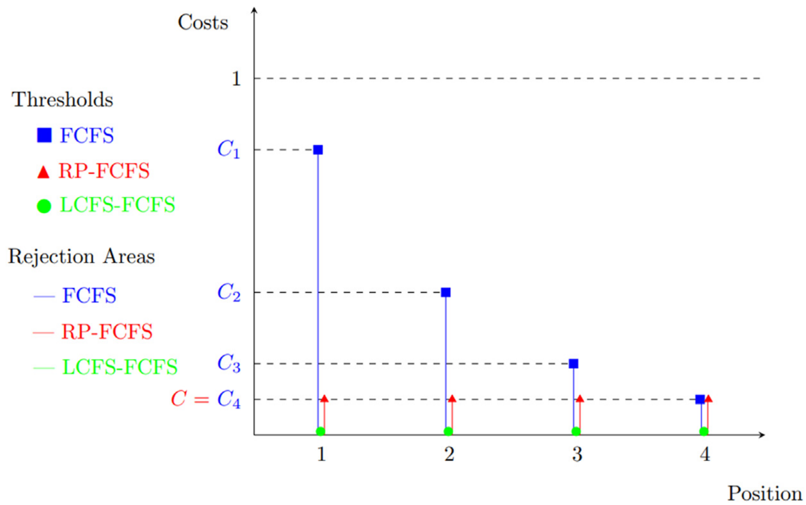

Fig. 1. Thresholds and rejection areas under different disciplines with fixed-length 4.

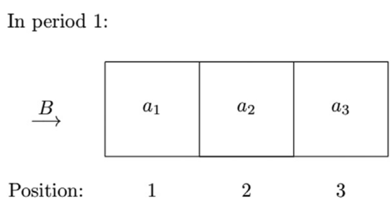

Fig. 2. Waste depends on the actions of all agents in the queue.

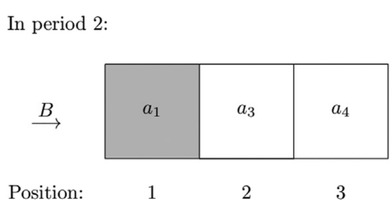

Fig. 3. Waste only depends on the actions of the last two agents, since an rejection is observed.

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

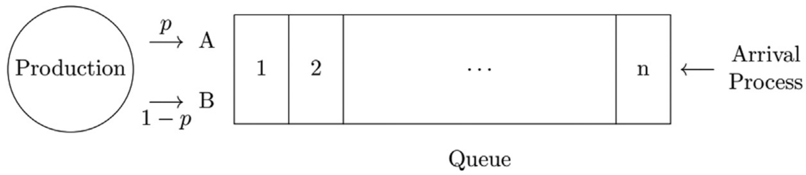

Fig. 4. The simple FCFS queue.

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

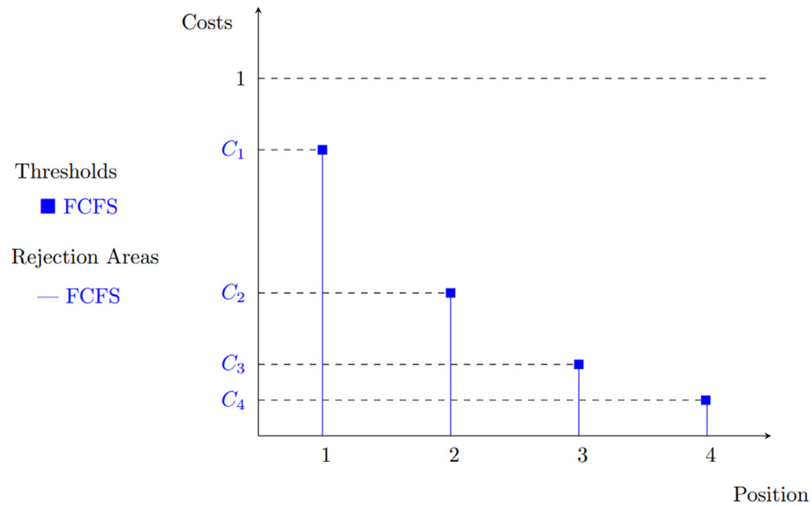

Fig. 5. Thresholds and rejection areas under simple FCFS queue with fixed-length 4.

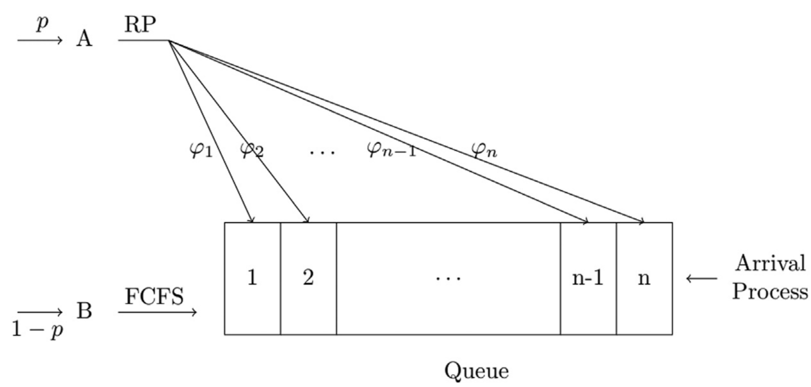

Fig. 6. The general complex queue discipline.

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

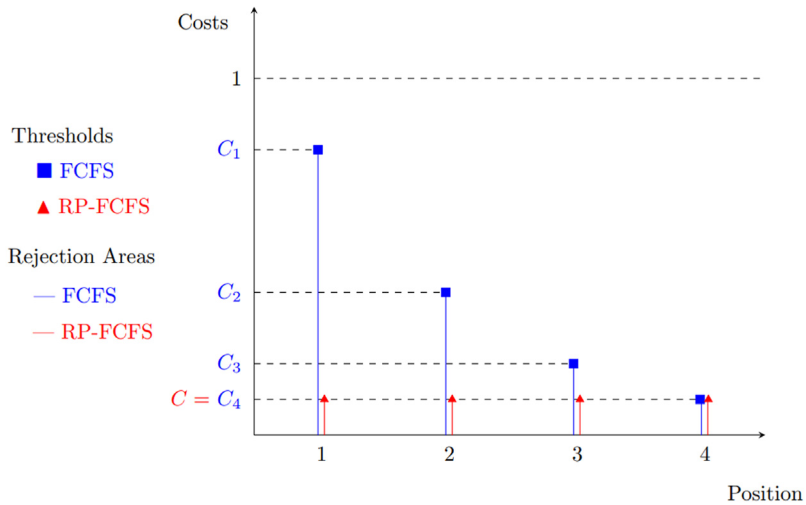

Fig. 7. Thresholds and rejection areas in different queue disciplines with fixed-length 4.

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:

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

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:

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

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:

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

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

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

...

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

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

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:

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

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.