Noname manuscript No.
(will be inserted by the editor)

Strategic Arrivals to Queues Offering Priority Service

Rajat Talak · D. Manjunath · Alexandre Proutiere

the date of receipt and acceptance should be inserted later

Abstract We consider strategic arrivals to a FCFS service system that starts service at a fixed time and has to serve a fixed number of customers, e.g., an airplane boarding system. Arriving early induces a higher waiting cost (waiting before service begins) while arriving late induces a cost because earlier arrivals take the better seats. We first consider arrivals of heterogenous customers that choose arrival times to minimize the weighted sum of waiting cost and and cost due to expected number of predecessors. We characterize the unique Nash equilibria for this system.

Next, we consider a system offering L levels of priority service with a FCFS queue for each priority level. Higher priorties are charged higher admission prices. Customers make two choices—time of arrival and priority of service. We show that the Nash equilibrium corresponds to the customer types being divided into L intervals and customers belonging to each interval choosing the same priority level. We further analyze the net revenue to the server and consider revenue maximising strategies—number of priority levels and pricing. Numerical results show that with only a small number of queues (two or three) the server can obtain nearly the maximum revenue.

1 Introduction

Consider an airline that starts boarding a plane at time 0 and the N customers that are booked to fly on the plane all arrive before boarding starts and wait in a FCFS queue. A customer that has a better rank in the queue gets a better choice of seats and luggage space. However, a better queue rank is achieved by arriving early and hence incurring a cost due to the waiting time before service (boarding) begins. Thus, with each customer we can associate a total inconvenience cost that has two components—(1) the *waiting cost* from the inconvenience due to arriving early (before boarding time), and

Address(es) of author(s) should be given

(2) the *boarding cost* from the inconvenience due to customers that are served before the focal customer, i.e., inconvenience due to its rank or position in the boarding queue. An immediate model would be to assume that these costs are separable, and that the total cost is a weighted sum of the waiting time and the rank in the boarding sequence. (Note that the second component accounts for possibly non negligible service time once boarding begins.) Each customer then has to choose the time at which to arrive into the queue, i.e., the waiting time before service begins. Of course, for a given customer the queue rank that it obtains depends on the choice of arriving (and hence waiting) times of the other customers. In this paper we assume that the customers are strategic and that they choose their arrival time to minimize their individual cost. We also assume that none of the customers renege and that all N take service. Thus we have the rudiments of the *airplane boarding game* which we will define more generally in the following.

Now assume that the customers are from a heterogeneous population and that the different types of customers calculate their total costs differently. Specifically, for the linear cost function described in the preceding, different types of customers have different weights for the queue rank and for the waiting time before boarding. In this case, rather than have a single queue and hence provide egalitarian service, the airline could reduce the social cost, and possibly increase its revenue, by providing service differentiation as follows. Maintain two FCFS queues with queue 2 having strict boarding priority over queue 1, i.e., board customers from queue 2 before boarding customers from queue 1. As before, assume that all the customers arrive before boarding begins and that there is no reneging by any of the N customers. The airline could now charge a fixed premium to the customers who join the higher priority queue. To obtain better seats, i.e., to make their boarding costs lower, the customers can now choose to join the high priority queue and trade off some waiting time (before boarding begins) for payment of the premium in the higher priority queue. In such a system the customers make two choices—the arrival time and the queue that they plan to join and hence the premium that they will be paying. Once again, the actual total cost for each customer depends on the choice of these parameters by the other customers. We thus have the outline of the basic *airplane boarding game* in which customers strategically choose two parameters to minimize their individual total cost.

Several budget airlines in fact follow a simple version of the system described above and motivates this paper. In these airlines, seat numbers are not issued at check-in and typically two FCFS queues are maintained at the boarding gate. The premium queue allows service differentiation to a heterogeneous population where different types of customers have different relative values for their time and money and provides an opportunity for the airline to make additional revenue. If there is significant heterogeneity in the customer population then increased differentiation can provide increased revenue. Thus, rather than provide just two priorities, the airline could provide L priority queues with customers in priority l having strict priority over those of strictly less than l. In this paper we consider this general case of the following *airplane boarding game.*

N customers are to be served by a single server starting service at time 0. Different types of customers will be served through L FCFS priority queues with higher priorities charging higher premiums. Associated with each type of customer is a total cost function that depends on the waiting time, the rank in the boarding sequence, and the premium paid to receive service from the queue of choice. The distribution of the types in the customer population is assumed known to the server and to the other customers. Each customer strategically chooses two parameters, the arrival time and the priority level of the queue from which it wants to receive service, to minimize its expected total cost. None of the N customers renege and all of them obtain service and this is known to all the customers.

1.1 Preview of the Results

In this paper our objectives are twofold: (1) studying customers' strategic behavior in their choice of the priority and of the arrival time, and (2) analyzing how the server may tune the service parameters, the number of priority levels and the charge for each of these priority levels, to maximize its revenue. The following is the summary of the key results in the paper.

1. In Section 3, we consider a system where customers choose a single priority parameter that we call the grade of service. The service provider offers a continuum of grades of service and defines a pricing function that maps the service grade to a price. Higher service grades have strictly higher priority in boarding. Customers are strategic and choose the service grade and hence the price. For this case we show that the revenue is independent of the pricing function.
   Interestingly, we will see that the model considered in this section is very general in that it is applicable to 'finite duration' games like in the airline boarding game where a finite number of customers are to be allocated a finite resource. For this case our results should be interpreted in the spirit of [6]. The results are also applicable to 'infinite duration' games where the customer arrivals and service processes are in a stationary regime; in this case the results are to be interpreted in the spirit of, e.g., [4, 5, 15].
2. In Section 4, we consider the system where the server maintains several FCFS queues at different priority levels. Customers pay a higher price to join a queue with higher priority. The number of priority levels and the price of service from these priority levels is assumed known to all the customers. In this case the customers have to select two parameters—the arrival time before boarding starts, and the priority of the queue from which they will receive service. We show that when the prices are fixed, the game admits a unique Nash equilibrium that we characterize exactly. We also numerically illustrate the strategic choices of the priority parameters at equilibrium.

3. Finally, in Section 5, we investigate how the service provider may set the prices to join the various queues to maximize its revenue. When the number of priority queues are fixed, we show that these prices can be computed by solving a simple dynamic program. Numerically, we observe that the revenue increases with the number of available queues. However, the marginal gain of operating additional queues vanishes rapidly with increasing number of priorities. In fact we see that for several examples, maintaining a small number number of queues (two or three) may actually yield a revenue that is almost optimal, i.e., two or three levels priority levels is enough to extract a 'close to maximum' revenue.

All proofs are carried in the appendix.

In the next section we provide a brief overview of the relevant literature and delineate them from the results of this paper.

2 Related Work

The *meeting game* and the concert queueing game have similarities to the airplane boarding game . We first describe these briefly.

A meeting is scheduled to start at time t but the participants arrive at random times and it usually starts at a random time T when quorum is achieved. Knowing that the meeting does not necessarily start at time t, participants i chooses to arrive at τ_i to minimize her costs due to waiting (a function of $(T - \tau_i)^+$), due to inconvenience (a function of $(\tau_i - T)^+$), and due to loss of reputation (a function of $(\tau_i - t)^+$). The knowledge of the structure and distribution of these costs among the participants can be strategically used by participant i to choose τ_i. This meeting game has been introduced and studied in [6]. A variant of the meeting game is the 'concert queueing game'. Here the concert hall opens at time t and the attendees queue up for FCFS service; each attendee requires a random service time. Arriving late has the drawback that the better seats will be taken by those who came earlier, whereas arriving early induces a cost of waiting. Attendees strategically choose their arrival time distribution to mitigate these opposing costs. Several versions of this game have been studied in [8–10].

In the airplane boarding game all customers arrive before the start of service which is unlike in the meeting game and in the concert queueing game. Further, all analyses on the concert queueing game consider a single queue while we consider multiple priority queues providing differentiated service to a heterogeneous customer population that values time, money and queue rank (boarding costs) differently. A finite number of customers being served in a finite time is also considered in [14]. They consider a discrete time system in which the customers incur a congestion cost due to other customers arriving in the same slot. This congestion cost could also be time-dependent. The airline boarding game that we describe here does not consider such a congestion cost.

Queueing systems with customers who value delays differently have been investigated in the queueing and economics literature. Specifically, providing lower delays for customers with higher delay costs and extracting a commensurate payment from them has been of interest for a while now. In an early work on pricing based on differential service, [13] considered a queue in which customers bribe the server and the server provides a 'highest bribe first' service. Strategic customers in such a queue has been considered in [2, 5, 15]. In fact, our Theorem 1 in Section 3 is a generalization of the equilibrium analysis in [5, 15]. Providing differentiated service to a heterogeneous population through priority queues and pricing the priorities has also been considered in, among others, [16, 19]. More recent work on priority pricing is in [1, 12] where only the expected delay contributes to the cost and the weights have a specific form. Priorities may also be auctioned like in [5, 15].

The models in all of the preceding studies have considered what can be termed a 'simple' system where there is only one parameter that the customer chooses. It is the arrival time in [6], the bid value in [2,5,13,15] and the priority value in [1, 12, 16, 19]. In contrast, in our model, the customer has to choose two parameters—the waiting time *and* the priority level. The closest system to the one that we consider is that in [7] where the attendees choose both the arrival time and the queue to join. The key difference with our model is that in [7] the queues start their service at different times and could be serving in parallel. Thus in this case there is 'isolation' between the customers that choose different queues.

A more recent work is that of [4] which considers a priority queueing system with a pricing menu. Here the customers have to declare their delay valuation and the focus is on deriving incentive compatible revenue maximizing pricing functions. In our model, type is private information and does not have to be revealed and the choice of the grade of service is made by the customer and not by the service provider.

3 Single Queue: Choosing the Arrival Time

A plane is to board N customers starting at time 0. The N customers arrive at different times before the boarding starts and form a FCFS queue. The population from which the N customers are drawn is heterogeneous in that different customers value their waiting time differently. Let $\mathcal{V} = [A, B] \subset \mathbb{R}^+$ be the set of all customer types, and let it be endowed with a probability measure that has a cumulative distribution function (CDF) given by $\mathcal{G}$. We assume $\mathcal{G}$ to be a continuous distribution.

A customer has to decide how much before the boarding time she has to arrive; we will denote this choice of waiting time (before boarding begins) by $T(v)$, for every $v \in \mathcal{V}$. Therefore, waiting time of $T(v)$ implies that the customer arrives at time $-T(v)$. We assume that all customers of the same type make the same decision. Further, we assume that none of the N customers renege and all of them join the queue before time $t = 0$. If a customer of class $v \in \mathcal{V}$ arrives at time $-t$ then her cost is given by

$$c_v(t) = NF(t) + n_v g(t) \tag{1}$$

where $F(t)$ denotes the fraction of customers that choose to arrive before $-t$ on average, $g(t)$ is the cost of waiting for t time units, and n_v is the weight that a type v customer assigns to it's waiting time cost. Notice that $F(t)$ is a monotonically decreasing function. Further, $g(t)$, being the cost of waiting for time t, is a monotonically strictly increasing function in t. We also assume that g is continuous and differentiable function with $g(0) = 0$. Note that $g(\cdot)$ is independent of the customer type v. Without loss of generality we let n_v to be a monotonically strictly increasing function in v. Thus $NF(t)$ represents boarding cost, or the cost corresponding to the customer's position in the queue. This cost includes the inconvenience due to the position in the boarding sequence and the service time of customers ahead of the focal customer after boarding begins. $n_v g(t)$ is the total waiting cost for a customer of type v. Now, $F(t)$ is given by

$$F(t) = \int \mathbb{1}_{T(v) \geq t} \, d\mathcal{G}(v). \tag{2}$$

An arriving customer of type v chooses $T(v)$ to minimize its expected cost where the expectation is taken over the customer type distribution $\mathcal{G}$. We seek an arrival profile $v \mapsto T(v)$ that minimizes the individual expected cost for all customers. This is also the Nash equilibrium (NE) policy as defined below.

Definition 1 $T^{NE} : \mathcal{V} \to [0, +\infty)$ is a Nash equilibrium policy if

$$T^{\mathrm{NE}}(v) = \operatorname*{argmin}_{t \geq 0} c_v(t), \tag{3}$$

for every $v \in \mathcal{V}$. [1]

Since customers with larger v value their time more than those with smaller v, we would expect that customers with larger v should arrive later in a NE policy. We prove this in Lemma 1. Further, in Theorem 1 we show that there is a unique NE and also characterize this policy.

Lemma 1 *Assume that a NE policy T^{NE} exists. Then $T^{NE}(v)$ is a non-increasing function in v.*

Proof See Appendix A. □

Theorem 1 *If $g(\cdot)$ is such that there exists a NE policy then it is unique and is given by*

$$T^{NE}(v) = g^{-1}\left(\int_v^B N \frac{d\mathcal{G}(x)}{n_x} \right). \tag{4}$$

Proof See Appendix B. □

[1] This definition of NE is consistent with that found in the literature on non-atomic games [11].

Remark 1 In this paper we will focus on the characterisation of the NE when it exists. The question of the existence of the NE is answered in [11, 17, 18] where it is shown that in non-atomic games that have a continuum of players, there exists a pure strategy NE and it is unique. For our model, i.e. considered assumptions on g and n_v, the NE policy exists. More general conditions on $g(\cdot)$ and n_v that yield a NE can be derived from [18].

Consider an example system where $g(t) = t^r$ for any $r > 0$. For this system, the NE policy exists and, as proved, is unique. Thus we see that there is a non empty set of $g(\cdot)$ for which Theorem 1 holds.

3.1 Extensions and Generalizations

Lemma 1 and Theorem 1 can be extended as follows.

1. The preceding results will follow even for the following total cost function for a customer of type v. Let $h(x)$ be an increasing and continuously differentiable function in x and let the cost for a customer of type v, when she arrives t units of time before boarding, be

$$c_v(t) = h\left(F(t)\right) + n_v g(t). \tag{5}$$

   Note that $h(\cdot)$ is independent of the customer type v. Even for this case Lemma 1 and Theorem 1 hold except that the unique NE arrival profile would be given by

$$T^{\mathrm{NE}}(v) = g^{-1}\left(\int_v^B \frac{h'(\mathcal{G}(x))}{n_x}\, d\mathcal{G}(x)\right), \tag{6}$$

   where $h'(x)$ denotes derivative of h.
2. In (5), note that $h(F(t))$ is a specific decreasing function of t; as $h(\cdot)$ and $F(\cdot)$ are increasing and decreasing functions, respectively. Instead, we could also replace t by any decreasing function of t. This is true because if g_1 and g_2 are two decreasing functions over the same domain $\mathcal{D}$, then we can find an increasing function h such that $g_1(x) = h(g_2(x))$ for all $x \in \mathcal{D}$.
3. In the model we assumed that n_v was an increasing function. The results follow if n_v is a decreasing function, except that $T^{\mathrm{NE}}(v)$ now would be a non-decreasing function in v. This suits our intuition as now the cost of waiting is more for customers with smaller v.

3.2 A Server with a Continuum of Service Grades

In the preceding, the customer had to choose the time of arrival and that time determined its total cost. If we could treat time as just a parameter we could have different interpretations to the parameter and hence apply it to different systems. The following is an example of such a generalization which results in a relatively different server model but yields an interesting NE arrival profile.

A population of heterogeneous customers, characterized by a type $\mathcal{V} = [A, B] \subset \mathbb{R}^+$, arrive for service. They have to buy a priority w from the server which is priced by the server. Without loss of generality we assume the class of all priorities to be $\mathcal{W} = [0, 1]$ and $P : \mathcal{W} \to \mathbb{R}^+$ to be the pricing function such that $P(w)$ denotes the price for priority w. We assume P to be any continuous, differentiable, monotonically increasing function with $P(0) = 0$ and $P(1) = P_{\max}$. Then the cost incurred by a customer of type v when she buys priority w is given by

$$c_v(w) = NF(w) + n_v P(w), \tag{7}$$

where n_v is an increasing function in v and $F(w)$ is the fraction of customers who choose a priority higher than w. Note that $n_v P(w)$ is the cost incurred by the type v customer in buying priority w.

Since, this is only a different interpretation of our original system model, the results of Lemma 1 and Theorem 1 hold in this scenario too. Thus, the unique NE is given by

$$w^{\mathrm{NE}}(v) = P^{-1}\left(\int_v^B N \frac{d\mathcal{G}(x)}{n_x}\right). \tag{8}$$

Let us now compute the expected revenue earned by the service provider, and see what price function P will maximize it. For a given price function $P(\cdot)$, the revenue of the service provider is

$$R(P) = \int_A^B P\left(w^{\mathrm{NE}}(v)\right) d\mathcal{G}(v).$$

Substituting (8), we obtain

$$R(P) = \int_A^B \left(\int_v^B N \frac{d\mathcal{G}(x)}{n_x}\right) d\mathcal{G}(v).$$

Observe that $R(P)$ defined above is independent of $P(\cdot)$. Thus, when the customers are strategic and $\mathcal{G}(v)$ distribution is known to the customers, the revenue to the service provider is invariant to the pricing function that is increasing in the priorities and has range with range $[0, P_{\max}]$! This leads us to argue there is only so much willingness in the market to pay, and that any pricing function will fully extract it. We argue that this is because with a continuous pricing function, it is possible to achieve perfect discrimination which in turn leads to revenue maximisation for the service provider [3]. In contrast, we will see in the next two sections that a finite number of service levels does not extract the maximum revenue. We argue that this is because of the 'quantisation' effect which in turn does not admit perfect discrimination because the customers will have to choose the 'nearest' grade of service.

4 Multiple Queues: Customers Choose Arrival Time and Priority

Now we consider the system of the previous section but with L FCFS queues. Queue $l+1$ has strict priority over all queues of priority level l or less, i.e., it is served before queue l for all $l = 0$ to $L-1$. Customers who join queue l have to pay an admission price of P_l. Since queue $l+1$ has priority over queue l the server enforces that $P_{l+1} > P_l$. Queue 0, however, has no admission price, i.e., $P_0 = 0$.

Let $c_v(l,t)$ denote the cost of joining queue l at time t before boarding begins. Then they are defined as follows:

$$c_v(l,t) = m_v P_l + N \sum_{j=l+1}^{L-1} F_j(0) + N F_l(t) + n_v g(t), \tag{9}$$

where $F_l(t)$, for $0 \le l \le L-1$, denote the fraction of customers that wait for t or longer in queue l. Further, $m_v P$ is how much a customer v values price P in relation to his rank in the queue.

Each customer has to determine which queue she wants to join and when to arrive. First, consider a customer of type v who wants to join Queue l. Clearly, her optimal arrival time $T_l(v)$ is

$$T_l(v) = \operatorname*{argmin}_{t \ge 0} c_v(l,t). \tag{10}$$

For this choice of $T_l(v)$, the cost of joining Queue l will be $c_l(v) \triangleq c_v(l, T_l(v))$. Thus the optimal randomized choice of queue is obtained from

$$q(v) = \operatorname*{argmin}_{q_j} \sum_{l \in \mathcal{L}} q_l c_v(l, T_l(v)), \tag{11}$$

where $q(v) = (q_0(v), q_1(v), \ldots, q_{L-1}(v))$ and $q_l(v)$ denotes the probability that a customer of type v would join Queue l. An equilibrium strategy is defined as follows.

Definition 2 $(T_l(v), q_l(v))_{l=0}^{L-1}$ is a NE policy if (10) and (11) are satisfied, where for all $l \in \mathcal{L}$,

$$F_l(t) = \int_{v \in \mathcal{V}} \mathbb{1}_{T_l(v) \ge t} \; q_l(v) \; d\mathcal{G}(v) \tag{12}$$

is the fraction of customers that join Queue l and arrive at least t units of time before 0.

Let $\left(T_l^{\mathrm{NE}}(v), q_l^{\mathrm{NE}}(v)\right)_{l=0}^{L-1}$ denote an equilibrium strategy. We first analyze the system with a single queue, i.e., $L = 1$, and then proceed to the analysis when there are L queues.

Lemma 1 is first extended to L queues to show that at NE, the optimal joining times at each queue, $T_l^{\mathrm{NE}}(v)$, is non increasing in v.

Lemma 2 *$T_l^{NE}(v)$ is a non increasing function in v for each $l \in \mathcal{L}$.*

Proof See Appendix C. □

Recall that A and B are, respectively, the minimum and the maximum values of the support of v. Thus an implication of this lemma is that at NE, type A customers will be the first to arrive while type B customers will be the last to arrive and this is independently of the queue that they join. This leads us to conclude that[2]

$$T_l^{\mathrm{NE}}(B) = 0, \tag{13}$$

and

$$F_l\left(T_l^{\mathrm{NE}}(A)\right) = 0, \tag{14}$$

for every $l \in \mathcal{L}$.

We now see how $T_l^{\mathrm{NE}}(v)$ would vary from the case when there is only a single queue to that when there are two or more queues. Intuitively, when there are two or more queues, customers would want to come later than they would if there was only a single queue. This is confirmed by the following lemma.

Lemma 3 *For all l, we have $T_l^{NE}(v) \leq g^{-1}\left(N \int_v^B \frac{d\mathcal{G}(x)}{n_x}\right)$.*

Proof See Appendix D. □

From Theorem 1, $g^{-1}\left(N \int_v^B \frac{d\mathcal{G}(x)}{n_x}\right)$ is the arrival time at NE in a single queue system. Using the previous results, we next show that there is a unique NE and at this NE, the strategies are pure. Toward that proof the following assumptions will be made on the system parameters.

(A1) We assume that m_v and n_v are such that $n_v' \int_v^B \frac{d\mathcal{G}(x)}{n_x}$ does not grow faster than $-m_v'$. Specifically,

$$y(v) \triangleq \frac{n_v'}{(-m_v')} \int_v^B \frac{d\mathcal{G}(x)}{n_x},$$

is a bounded function over $v \in \mathcal{V}$. Also, $\int_A^B \frac{d\mathcal{G}(x)}{n_x} < \infty$.

(A2) We also make an assumption on the minimum difference between the prices to join 'adjacent' queues. Specifically, for every l and $l+1$ in $\mathcal{L}$,

$$P_{l+1} - P_l > N \max\left\{\sup_{v\in\mathcal{V}} y(v), \frac{2}{m_A}\right\}. \tag{15}$$

(A3) Finally, we assume that m_B is small enough so that all queues are occupied, and also that $P_{l+1} - P_l < \frac{NF_l(0)}{m_B}$.

The following example illustrates that (A1) and (A2) are not very restrictive. Let v be uniformly distributed over $[A, B]$, $n_v = v$, and

$$m_v = \frac{N}{\epsilon(B-A)}\left(B\log\left(\frac{B}{v}\right) - (B-v)\right), \tag{16}$$

[2] The proofs of (13) and (14) are trivial.

where $\epsilon > 0$. In this case,

$$y(v) = \frac{\epsilon}{N} \frac{\log\left(\frac{B}{v}\right)}{\left(\frac{B}{v} - 1\right)} \leq \frac{\epsilon}{N}.$$

Also, note that, if m_v satisfies (A2) then so does $m_v + \delta$. Since $y(v) \leq \frac{\epsilon}{N}$, it suffices to have

$$P_{l+1} - P_l > \max\left\{\epsilon, \frac{N}{m_A}\right\},$$

to satisfy (A2). Further, if $A = 0$ then $m_A = \infty$, in which case we only require $P_{l+1} - P_l > \epsilon$. In general, this would be the case if m_A is sufficiently large. If (A3) is violated the empty queues will be those of higher priority. From an operational point of view there is no reason to have a queue if no customer is going to join it, especially a high priced queue.

Theorem 2 *The NE strategy is unique and is characterized by*

$$q_l^{NE}(v) = \mathbb{1}_{v_l < v \leq v_{l+1}}. \tag{17}$$

Here $A = v_0 < v_1 < v_2 < \cdots < v_{L-1} < v_L = B$ *are given by*

$$c_{l-1}(v_l) = c_l(v_l),$$

for all $l = 1$ *to* $L - 1$, *each of which has a unique solution.*

Proof The outline of the proof is as follows.

1. Lemma 4 first shows that for type A customers, the optimal queue joining cost, $c_l(A)$, increases in l, whereas for type B customers this cost, $c_l(B)$, decreases in l.
2. Lemma 5 shows that $\frac{dc_0(v)}{dv} > 0$, while for $l \geq 1$, $0 > \frac{dc_l(v)}{dv} > \frac{dc_{l+1}(v)}{dv}$.
3. Using this, in the third part of Lemma 5, we show that for every l the cost functions $c_{l-1}(v)$ and $c_l(v)$ intersect at a unique point v_l and these thresholds $\{v_l\}_{l=1}^{L-1}$ satisfy

$$A < v_1 < v_2 < \cdots < v_{L-1} < B.$$

4. Given the thresholds $\{v_l\}_{l=1}^{L-1}$, we show that the optimal joining cost for Queue l is greater than those of other queues for all $v \in (v_l, v_{l+1}]$, i.e., $c_l(v) \leq c_j(v)$ for all $v \in (v_l, v_{l+1}]$ and $j \in \mathcal{L}$.[3]

This proves that at NE we have (17). The details are in Appendix E. □

We now give an explicit characterization of the NE arrival times, $T_l^{\mathrm{NE}}(v)$, and the thresholds, $\{v_l\}_{l=0}^{L}$, in the following theorem.

[3] It does not matter whether the customer of type v_l joins either Queue l or Queue $l+1$. Set of all such customers form a negligible (measure 0 w.r.t. $\mathcal{G}$) set in $\mathcal{V}$.

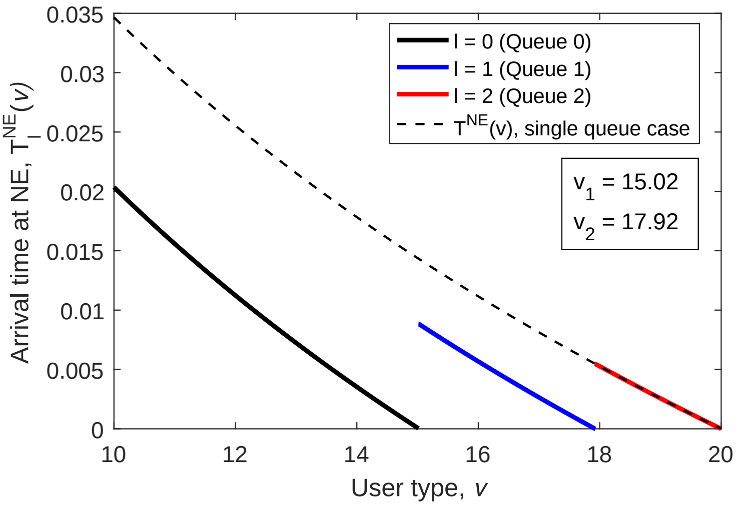

Fig. 1 Comparison of NE arrival times at different queues for a system with three queues ($v \sim \mathcal{U}[0, 20]$, $N = 10$, $P_1 = 8.75$, $P_2 = 11.45$ and $\delta = 0.05$).

Theorem 3 *At NE, the arrival time to Queue l is given by*

$$T_l^{NE}(v) = \begin{cases} 0, & \text{for } \; v > v_{l+1} \\ g^{-1}\left(N \int_v^{v_{l+1}} \frac{d\mathcal{G}(x)}{n_x}\right) & \text{for } \; v_l \le v \le v_{l+1}. \\ T_l^{NE}(v_l), & \text{for } \; v < v_l \end{cases}$$

Further, the thresholds $\{v_l\}_{l=0}^{L}$ can be computed by solving

$$\mathcal{G}(v_{l-1}) = \mathcal{G}(v_{l+1}) - \left(\frac{P_l - P_{l-1}}{N}\right) m_{v_l} - n_{v_l} \int_{v_l}^{v_{l+1}} \frac{d\mathcal{G}(x)}{n_x}, \tag{18}$$

for $1 \le l \le L - 1$.

Proof See Appendix F. □

We illustrate the preceding results with an example. Consider a boarding game with $\mathcal{V} = [0, 20]$ and the customer population distributed uniformly over $\mathcal{V}$. The service provider operates three queues with respective prices $P_0 = 0$, $P_1 = 8.75$, and $P_2 = 11.45$. Let $N = 10$, $n_v = v$, and m_v modified from (16) as follows.

$$m_v = \frac{N}{\epsilon(B - A)} \left(B \log\left(\frac{B}{v}\right) - (B - v)\right) + \frac{N\delta}{\epsilon}. \tag{19}$$

For this example, we choose $\delta = 0.05$.

Note that, we do not need the value of ϵ to determine the NE policy, as it can be subsumed in the prices P_l. Whenever we plot the price P_l and revenue, they are normalized by ϵ.

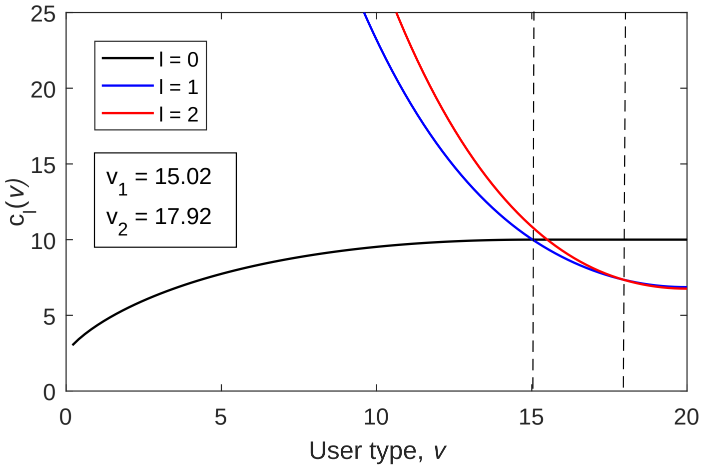

Fig. 2 Comparison of optimal costs to join Queue l and illustration of the thresholds ($v \sim \mathcal{U}[0, 20]$, $N = 10$, $P_1 = 8.75$, $P_2 = 11.45$ and $\delta = 0.05$).

Figure 1 plots the arrival times at the three queues. Observe that $T_l^{\mathrm{NE}}(v)$ is a decreasing function of v, which concurs with Lemma 2. In equilibrium, we observe that customers of a lower priority start arriving before the customers of higher priority. It is, however, not true that, in equilibrium, a customer which chooses a higher priority will arrive later than a customer that chooses a lower priority. We can observe this in Figure 1. For example, customer $v = 16$ chooses a higher priority than customer $v = 14$ but arrives earlier than the customer $v = 14$.

Also plotted in Figure 1 is the optimal arrival time as a function of the customer type when there is only a single queue. We see that this arrival time is greater than the arrival times at each queue for the three queue example. This concurs with Lemma 3. We only plot the arrival times for $v \geq 10$. For $v < 10$ the time to join the 0-th queue increases rapidly, however, it does remain upper bounded by the optimal joining time if there was a single queue.

To illustrate the thresholds at NE, the optimal joining costs, $c_l(v)$, for each queue is plotted as a function of v in Figure 2. A customer will join Queue l if it offers it the least optimal joining cost. Thus, the crossing points between the optimal joining costs determine the thresholds; this is observed clearly in Figure 2.

4.1 Discussion

We now explore how the NE strategy depends on the system parameters, specifically on the prices. We consider a two-queue boarding game with customer types uniformly distributed over $[5, 15]$. $n_v = v$ and m_v is given by (19).

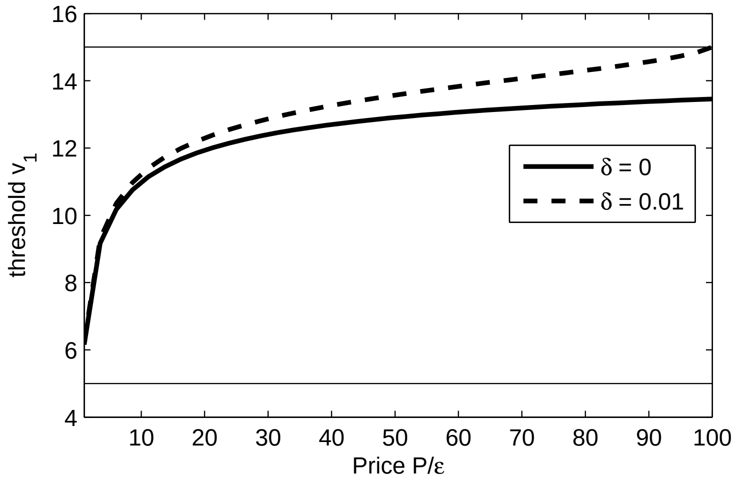

Fig. 3 Threshold v_1 as a function of price P for the two-queue boarding game ($v \sim \mathcal{U}[5, 15]$).

Figure 3 plots the threshold v_1 as a function of the price, P, of the priority queue. Observe that v_1 increases with P. This is intuitive because the incentive to join the high priority queue is diminished as the price increases. In fact, for $\delta = 0.01$, we find that $v_1 = B$ when $P = 100\epsilon$. However, this is not true when $\delta = 0$, in which case v_1 keeps increasing to B, but is never equal to it. Investigating this further, since $NF_1(0)$ is the number of customers joining Queue 1, revenue earned by the server would be $R(P) \triangleq NF_1(0)P$. This is not a linear function of P as $F_1(0) = \frac{B-v_1}{B-A}$ and depends on P.

In Figure 4, we plot the revenue (normalized by $N\epsilon$) as a function of the price P (normalized by ϵ). For $\delta = 0$, we observe that the revenue keeps increasing as P is increased, while for $\delta = 0.01$ there is an optimal choice of P at which the revenue attains its maximum. Note that $\delta = 0$ (resp. $\delta > 0$,) corresponds to the case when $m_B = 0$ (resp. $m_B > 0$). Hence, the intuition is that when $m_B = 0$, there are always customers (close to customers of type B) who do not value money much. Therefore, the service provider can always increase her revenue by setting a higher price for the priority queue. On the contrary, when $m_B > 0$, there exists a price P, large enough, that customers of type B cannot afford. The latter decide to join Queue 0.

In the previous discussions, we have considered m_v that decreases more rapidly for smaller v than it does for larger v. Now consider a system where m_v decreases linearly, i.e.,

$$m_v = \frac{N}{\epsilon}\left(\frac{B-v}{B-A}\right)\log(B/A) + \frac{N\delta}{\epsilon}. \tag{20}$$

We now investigate the two-queue boarding game with m_v as above, $n_v = v$, and $g(t) = t$, and customer types uniformly distributed over $[A, B]$. For this

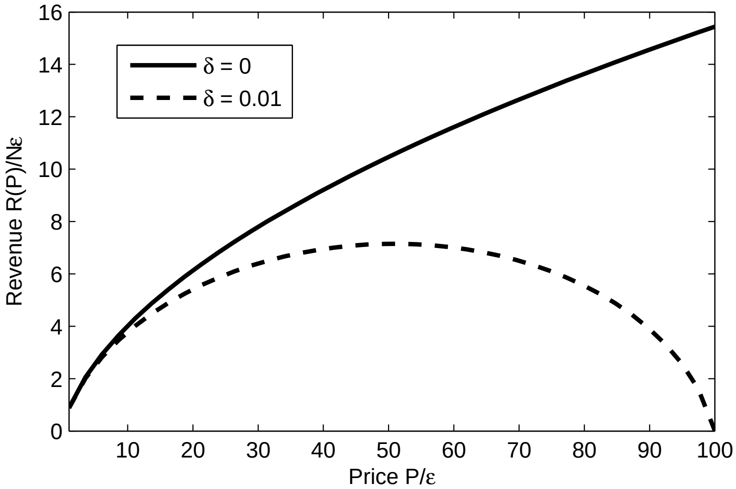

Fig. 4 Revenue earned by the service provider, $R(P)$, as a function of the price P for the two-queue boarding game ($v \sim \mathcal{U}[5, 15]$).

system $y(v)$ is

$$y(v) \; = \; \frac{1}{-m'_v} \int_v^B \frac{1}{B-A} \frac{dx}{x} \; = \; \frac{\epsilon}{N} \frac{\log\left(\frac{B}{v}\right)}{\log\left(\frac{B}{A}\right)}.$$

For $y(v)$ to be bounded we require $A > 0$, in which case $y(v) \leq \frac{\epsilon}{N}$. This upper bound is achieved when $v = A$. As observed previously, if m_A is sufficiently large then for (A2) it is sufficient if $P_{l+1} - P_l > \epsilon$. Further, for $\int_v^B \frac{d\mathcal{G}(x)}{n_x}$ to be bounded, we require $B < \infty$. The thresholds and revenue behave as in the previous example

5 Revenue Maximization

In the previous section, the admission price to each of the L queues was assumed given and we provided a complete characterization of the NE strategy. In this section, we investigate how the service provider can maximize its revenue by appropriately setting the prices P_l. Our motivational setting is the airplane boarding system where customers have already purchased their tickets. Hence it is reasonable to assume that there is one queue, the lowest priority queue, where there is no premium, i.e., $P_0 = 0$. Thus although the customers cannot balk, the existence of the a 'free queue' ensures that the revenue is not an increasing function of the premiums in the higher priority queues and there will be an optimum value for these prices. Further, increased service differentiation allows for increased revenues. In this section we first determine the optimum prices for a given L and then numerically analyze the effect of increasing L.

Table 1 Revenue maximizing prices, and the corresponding thresholds, for two, three, and four queues ($v \sim \mathcal{U}[0, 150]$).

L	v_1	v_2	v_3	P_1	P_2	P_3
2	135.28	-	-	76.73	-	-
3	134.35	143.46	-	71.68	79.75	-
4	134	139.22	142.52	69.92	73.65	76.58

For the L queue system, $F_l(0)$ is the fraction of customers that joined Queue l and, hence, $NF_l(0)$ is the total number of customers that join Queue l. The revenue earned by the service provider is given by

$$R(P_1, P_2, \ldots, P_{L-1}) = \sum_{l=1}^{L-1} NF_l(0)P_l. \tag{21}$$

We know that, at NE, $F_l(0) = \mathcal{G}(v_{l+1}) - \mathcal{G}(v_l)$. Also, (18) gives a bijective relation between P_l and v_l. Using these, we can express revenue as follows:

$$\frac{1}{N}R(P_1, P_2, \ldots, P_{L-1}) = \sum_{j=1}^{L-1} u\left(v_{L-j}, v_{L-j+1}\right), \tag{22}$$

where

$$u\left(v_l, v_{l+1}\right) = \left[\frac{\mathcal{G}(v_{l+1}) - v_l \int_{v_l}^{v_{l+1}} \frac{d\mathcal{G}(x)}{x}}{\frac{1}{N}m_{v_l}}\right](1 - \mathcal{G}(v_l)) - \left[\frac{\mathcal{G}(v_l)}{\frac{1}{N}m_{v_{l+1}}}\right](1 - \mathcal{G}(v_{l+1}))\,.$$

This is derived in Appendix G.

Thus, maximizing $\frac{1}{N}R(P_1, P_2, \ldots, P_{L-1})$ can be solved as a finite horizon dynamic program (FHDP). Notice that the first term in the expansion depends only on v_{L-1} (action at time 1 in the FHDP), the second term depends only on v_{L-2} and v_{L-1} (action and state of the system at time $t = 2$), the third term depends only on v_{L-3} and v_{L-2} (action and state of the system at time $t = 3$), and so on. We assume that the FHDP has a solution and verify this numerically.

We now numerically evaluate the revenue maximizing prices by solving the FHDP. As before, $\mathcal{V} = [A, B]$ and type is uniformly distributed. There are L queues, $N = 10$, $n_v = v$, and m_v is as in (19) with $\delta = 0.05$.

Table 1 lists the revenue maximizing prices and thresholds when $A = 0$ and $B = 150$ for $L = 2$, 3, and 4. We observe that as the number of queues is increased, for maximum revenue, P_l has to be decreased at each Queue l. This has been consistently observed for other system parameters.

In Table 2, we compare the maximum revenue for three different type distributions when $L = 2$, 3, and 4. We observe that the revenue increases as more queues are added, regardless of the underlying type distribution. Most

Table 2 Maximum revenue as a function of number of queues and type distribution

Population distribution	Revenue		
	$L = 2$	$L = 3$	$L = 4$
$\mathcal{U}[0, 20]$	2.26	2.46	2.50
$\mathcal{U}[0, 150]$	7.53	7.83	7.87
$\mathcal{U}[20, 150]$	7.41	7.65	7.68

importantly, the increase in this revenue from three queues to four queues is as small as 1.6%, 0.5%, and 0.4% for the type distributions $\mathcal{U}[0, 20]$, $\mathcal{U}[0, 150]$, and $\mathcal{U}[20, 150]$, respectively. However, for the same type distributions the increase in maximum revenue is 8.9%, 4%, and 3.2%, respectively. This tells us that having three, may be even just two, queues shall get us very nearly the maximum revenue.

We also remark here that this observation of a small number of classes yielding nearly the maximum revenue is also seen in other similar settings that serve a heterogenous population. For example, in [20], the authors consider dividing link capacity into multiple classes of service, each with its own price, an instance of Paris Metro pricing. Utility maximising users choose the class of service and it is shown that for a large range of utility functions, the loss of revenue is small even with a small number of service classes.

Another observation from the results of Table 2 is that the revenue is larger when the type distribution is $\mathcal{U}[0, 150]$. Numerically, it is consistently observed that if $\tilde{\mathcal{V}} \subset \mathcal{V}$, then the boarding game with types in $\mathcal{V}$ yields greater maximum revenue than with $\tilde{\mathcal{V}}$; a larger diversity of types provides a larger revenue. This property is known in the pricing literature, e.g., [3]. Larger diversity provides for more customers with higher valuations for quality of service (QoS). Such a user population allows the service provider to extract more consumer surplus by tuning service quality and price to each type to maximise profit.

6 Concluding Discussion

The system that we have introduced in this paper can be seen to belong to a class of queueing systems where the arrival times are endogenously determined by the customers; unlike the exogenously determined arrival times in traditional queueing systems. Clearly, there are several potential applications for systems with endogenous arrivals in modeling waiting systems at airports, bus and train terminals, concert halls, etc. A better understanding of these systems can help in a more informed sizing of waiting facilities.

Several extensions are possible. An extension of immediate interest is to develop a learning algorithm to obtain the revenue maximizing prices. Specifically, since such a game would be played out repeatedly, the outcome of each instance may be used to adapt the prices to maximize the revenue. A second extension would be to allow customers to balk if the cost is higher than the value of obtaining service.

A Proof of Lemma 1

The proof uses the optimality of $T^{\mathrm{NE}}(v)$ and $T^{\mathrm{NE}}(v+h)$ for $c_v(\cdot)$ and $c_{v+h}(\cdot)$, respectively, and the structure of the cost function.

$$\begin{aligned}
c_v(T^{\mathrm{NE}}(v)) &\leq c_v\left(T^{\mathrm{NE}}(v+h)\right) \\
&= NF\left(T^{\mathrm{NE}}(v+h)\right) + n_v g\left(T^{\mathrm{NE}}(v+h)\right) \\
&= c_{v+h}\left(T^{\mathrm{NE}}(v+h)\right) - n_{v+h} g\left(T^{\mathrm{NE}}(v+h)\right) + n_v g\left(T^{\mathrm{NE}}(v+h)\right).
\end{aligned}$$

Similarly, we can get

$$c_{v+h}(T^{\mathrm{NE}}(v+h)) \leq c_v\left(T^{\mathrm{NE}}(v)\right) - n_v g\left(T^{\mathrm{NE}}(v)\right) + n_{v+h} g\left(T^{\mathrm{NE}}(v)\right).$$

Adding the two we obtain

$$(n_{v+h} - n_v)\left(g\left(T^{\mathrm{NE}}(v+h)\right) - g\left(T^{\mathrm{NE}}(v)\right)\right) \leq 0.$$

Note that $n_{v+h} > n_v$, as n_v is a strictly increasing function. Therefore,

$$g\left(T^{\mathrm{NE}}(v+h)\right) \leq g\left(T^{\mathrm{NE}}(v)\right).$$

This implies that $T^{\mathrm{NE}}(v+h) \leq T^{\mathrm{NE}}(v)$, as $g(\cdot)$ is a strictly increasing function.

B Proof of Theorem 1

First note that $F(T^{\mathrm{NE}}(v)) = \mathcal{G}(v)$. This is derived using (2) as follows.

$$F(T^{\mathrm{NE}}(v)) = \int \mathbb{1}_{T^{\mathrm{NE}}(x) > T^{\mathrm{NE}}(v)} d\mathcal{G}(x) = \int \mathbb{1}_{x<v} d\mathcal{G}(x) = \mathcal{G}(v), \tag{23}$$

because $\mathcal{G}$ is a non-atomic distribution, and the second equality follows from Lemma 1.

Now, since $T^{\mathrm{NE}}(v)$ is a NE, by first order condition on (7), we have

$$\begin{aligned}
0 &= \left.\frac{dc_v(t)}{dt}\right|_{t=T^{\mathrm{NE}}(v)} \\
&= N\left.\frac{dF(t)}{dt}\right|_{t=T^{\mathrm{NE}}(v)} + n_v g'\left(T^{\mathrm{NE}}(v)\right),
\end{aligned} \tag{24}$$

where $g'(x)$ denotes derivative g with respect to x. Differentiating (23) with respect to v we get

$$\frac{d\mathcal{G}(v)}{dv} = \frac{dF(T^{\mathrm{NE}}(v))}{dv} = \frac{dT^{\mathrm{NE}}(v)}{dv} \times \left.\frac{dF(t)}{dt}\right|_{t=T^{\mathrm{NE}}(v)}. \tag{25}$$

Substituting (25) in (24), we obtain

$$-\frac{N}{n_v}\frac{d\mathcal{G}(v)}{dv} = g'\left(T^{\mathrm{NE}}(v)\right)\frac{dT^{\mathrm{NE}}(v)}{dv} = \frac{dg\left(T^{NE}(v)\right)}{dv}.$$

Integrating both sides with respect to v, we get

$$\int_v^B dg\left(T^{NE}(x)\right) = -\int_v^B \frac{N}{n_x} d\mathcal{G}(x). \tag{26}$$

This implies

$$g\left(T^{NE}(B)\right) - g\left(T^{NE}(v)\right) = -\int_v^B \frac{N}{n_x} d\mathcal{G}(x). \tag{27}$$

Note that $T^{NE}(v)$ is a non-increasing function. Thus, users of type B arrive last in the queue. If $T^{NE}(B) > 0$ then the cost for user of type B can be reduced by decreasing $T^{NE}(B)$ to 0, which contradicts the definition of $T^{NE}(v)$. Therefore, $T^{NE}(B) = 0$. This implies that $g\left(T^{NE}(B)\right) = 0$ due to assumptions on g. (27) then implies the result.

C Proof of Lemma 2

For notational simplicity, we shall denote $D(l,t)$ to mean

$$NF_l(t) + \sum_{j=l+1}^{L-1} NF_j(0), \tag{28}$$

in this section. Thus, the cost of a customer v is given by

$$c_v(l,t) = D(l,t) + m_v P_l + n_v g(t). \tag{29}$$

Comparing the costs of customer v and $v+h$, we have

$$c_{v+h}(l,t) = c_v(l,t) - m_v P_l - n_v g(t) + m_{v+h} P_{l+1} + n_{v+h} g(t). \tag{30}$$

Substituting $t = T_l^{\mathrm{NE}}(v)$ in (30), we obtain

$$\begin{aligned} c_{v+h}\left(l, T_l^{\mathrm{NE}}(v)\right) &= c_v\left(l, T_l^{\mathrm{NE}}(v)\right) - m_v P_l - n_v g\left(T_l^{\mathrm{NE}}(v)\right) + m_{v+h} P_{l+1} \\ &\qquad + n_{v+h} g\left(T_l^{\mathrm{NE}}(v)\right) \\ &\geq c_{v+h}\left(l, T_l^{\mathrm{NE}}(v+h)\right). \end{aligned} \tag{31}$$

The last inequality follows because $T_l^{\mathrm{NE}}(v+h)$ minimizes $c_{v+h}(l,\cdot)$. Similarly, substituting $t = T_l^{\mathrm{NE}}(v+h)$ in (30) we get

$$\begin{aligned} c_{v+h}\left(l, T_l^{\mathrm{NE}}(v+h)\right) = c_v\left(l, T_l^{\mathrm{NE}}(v+h)\right) - m_v P_l - n_v g\left(T_l^{\mathrm{NE}}(v+h)\right) + m_{v+h} P_{l+1} \\ + n_{v+h} g\left(T_l^{\mathrm{NE}}(v+h)\right). \end{aligned} \tag{32}$$

Since, $c_v\left(l, T_l^{\mathrm{NE}}(v+h)\right) \geq c_v\left(l, T_l^{\mathrm{NE}}(v)\right)$, we obtain

$$\begin{aligned} c_{v+h}\left(l, T_l^{\mathrm{NE}}(v+h)\right) \geq c_v\left(l, T_l^{\mathrm{NE}}(v)\right) - m_v P_l - n_v g\left(T_l^{\mathrm{NE}}(v+h)\right) + m_{v+h} P_{l+1} \\ + n_{v+h} g\left(T_l^{\mathrm{NE}}(v+h)\right), \end{aligned} \tag{33}$$

Adding the two inequalities, namely, (31) and (33), we get

$$(n_{v+h} - n_v)\, g\left(T_l^{\mathrm{NE}}(v)\right) \geq (n_{v+h} - n_v)\, g\left(T_l^{\mathrm{NE}}(v+h)\right). \tag{34}$$

Since, n_v is an increasing function of v, (34) reduces to $g\left(T_l^{\mathrm{NE}}(v)\right) \geq g\left(T_l^{\mathrm{NE}}(v+h)\right)$, which is nothing but

$$T_l^{\mathrm{NE}}(v) \geq T_l^{\mathrm{NE}}(v+h),$$

as $g(\cdot)$ is an increasing function. Thus, $T_l^{\mathrm{NE}}(v)$ is a decreasing function in v.

D Proof of Lemma 3

By Definition 2,

$$F_l\left(T_l^{\mathrm{NE}}(v)\right) = \int \mathbb{1}_{T_l^{\mathrm{NE}}(x) \geq T_l^{\mathrm{NE}}(v)} q_l(v) d\mathcal{G}(v) = \int \mathbb{1}_{x \leq v} q_l(v) d\mathcal{G}(v), \tag{35}$$

where the last equality follows due to Lemma 2. Now define

$$G_l(v) \triangleq \int_A^v q_l(v) d\mathcal{G}(v). \tag{36}$$

Then, we have $F_l\left(T_l^{\mathrm{NE}}(v)\right) = G_l(v)$; also note that $G_l(v) \leq \mathcal{G}(v)$ for every v. Differentiating this w.r.t. v, we obtain

$$\left.\frac{dF_l(t)}{dt}\right|_{t=T_l^{\mathrm{NE}}(v)} \frac{dT_l^{\mathrm{NE}}(v)}{dv} = \frac{dG_l(v)}{dv}. \tag{37}$$

Also, the first order derivative condition for the optimality of $T_l^{\mathrm{NE}}(v)$, namely, $\frac{dc_v(l,t)}{dt} = 0$, gives

$$0 = \left[N \frac{dF_l(t)}{dt} + n_v \frac{dg(t)}{dt}\right]_{t=T_l^{\mathrm{NE}}(v)}. \tag{38}$$

Substituting (37) in (38), we get

$$\frac{dg\left(T_l^{\mathrm{NE}}(v)\right)}{dv} = -\frac{N}{n_v} \frac{dG_l(v)}{dv}. \tag{39}$$

This can be simplified to

$$T_l^{\mathrm{NE}}(v) = g^{-1}\left(N \int_v^B \frac{dG_l(x)}{n_x}\right) \leq g^{-1}\left(N \int_v^B \frac{d\mathcal{G}(x)}{n_x}\right), \tag{40}$$

where, while the second inequality follows from $G_l(v) \leq \mathcal{G}(v)$, the first equality uses (13) and the same arguments that are used in Appendix B to arrive at the explicit expression of $w^{\mathrm{NE}}(v)$.

E Proof of Theorem 2

For the ease of presentation, we denote

$$D(l,t) \triangleq N F_l(t) + N \sum_{j=l+1}^{L-1} F_j(0). \tag{41}$$

In the following lemma, we prove that $c_l(A)$ increases, while $c_l(B)$ decreases in l.

Lemma 4 *$c_l(A)$ strictly increases and $c_l(B)$ strictly decreases in l.*

Proof The optimal cost to join Queue l is given by

$$c_l(v) = D\left(l, T_l^{\mathrm{NE}}(v)\right) + m_v P_l + n_v g\left(T_l^{\mathrm{NE}}(v)\right).$$

At $v = A$,

$$\begin{aligned} c_l(A) &= D\left(l, T_l^{\mathrm{NE}}(A)\right) + m_A P_l + n_A g\left(T_l^{\mathrm{NE}}(A)\right) \\ &= N \sum_{j=l+1}^{L-1} F_j(0) + m_A P_l + n_A g\left(T_l^{\mathrm{NE}}(A)\right), \end{aligned}$$

where the last equality follows from (14). We therefore have

$$
\begin{aligned}
c_{l+1}(A) - c_l(A) &= -NF_{l+1}(0) + m_A\left(P_{l+1} - P_l\right) - n_A\left(g\left(T_l^{\mathrm{NE}}(A)\right) - g\left(T_{l+1}^{\mathrm{NE}}(A)\right)\right) \\
&\geq -N + m_A\left(P_{l+1} - P_l\right) - n_A g\left(T_l^{\mathrm{NE}}(A)\right) \\
&\geq -N + m_A\left(P_{l+1} - P_l\right) - N n_A \int_A^B \frac{d\mathcal{G}(x)}{n_x},
\end{aligned} \tag{42}
$$

where the last inequality follows from Lemma 3 and the fact that $g(\cdot)$ is an increasing function. Using the fact that n_x is a decreasing function and $\int_A^B \frac{d\mathcal{G}(x)}{n_x} < \infty$ from (A1), we have

$$
\int_A^B \frac{n_A}{n_x} d\mathcal{G}(x) \leq \int_A^B d\mathcal{G}(x) = 1.
$$

Using this in (42) gives

$$
c_{l+1}(A) - c_l(A) \geq -2N + m_A\left(P_{l+1} - P_l\right) > 0,
$$

where the last inequality follows from (A2).

At $v = B$, we have

$$
c_l(B) = D\left(l, T_l^{\mathrm{NE}}(B)\right) + m_B P_l + n_B g\left(T_l^{\mathrm{NE}}(B)\right) = N \sum_{j=l}^{L-1} F_j(0) + m_A P_l,
$$

where the last equality follows from (13) and the fact that $g(0) = 0$. Now,

$$
c_{l+1}(B) - c_l(B) = -NF_l(0) + m_B\left(P_{l+1} - P_l\right) < 0, \tag{43}
$$

from (A3). □

We now prove some properties of the optimal costs $c_l(v)$ which help us in arriving at the thresholds.

Lemma 5 *For the optimal costs, $c_l(v)$s, the following is true*

1. $\frac{dc_0(v)}{dv} > 0$.
2. *For all $l \in \{1, 2, \ldots L-2\}$, we have $0 > \frac{dc_l(v)}{dv} > \frac{dc_{l+1}(v)}{dv}$.*
3. *There exists a unique point v_l, for every l, at which the two costs, namely, $c_{l-1}(v)$ and $c_l(v)$ intersect. Also, $A < v_1 < \cdots < v_{L-1} < B$.*

Proof Taking the derivative of $c_l(v)$ w.r.t. v we get

$$
\frac{dc_l(v)}{dv} = m'_v P_l + n'_v g\left(T_l^{\mathrm{NE}}(v)\right) + \frac{dT_l^{\mathrm{NE}}(v)}{dv} \times \left[\frac{\partial D(l,t)}{\partial t} + n_v \frac{dg(t)}{dt}\right]_{t=T_l^{\mathrm{NE}}(v)}. \tag{44}
$$

Using the first order optimality condition for $T_l^{\mathrm{NE}}(v)$, which is

$$
\left[\frac{\partial D(l,t)}{\partial t} + n_v \frac{dg(t)}{dt}\right]_{t=T_l^{\mathrm{NE}}(v)} = 0, \tag{45}
$$

in (44) we get

$$
\frac{dc_l(v)}{dv} = m'_v P_l + n'_v g\left(T_l^{\mathrm{NE}}(v)\right). \tag{46}
$$

For $l = 0$, since $P_0 = 0$, we have

$$
\frac{dc_0(v)}{dv} = n'_v g\left(T_l^{\mathrm{NE}}(v)\right) > 0, \tag{47}
$$

which follows from the fact that n_v is an increasing function in v and $g(\cdot)$ always positive.

Now, for $l \geq 1$, using Lemma 3 in (46) and the fact that $g(\cdot)$ is an increasing function, we get

$$\frac{dc_l(v)}{dv} < m'_v P_l + n'_v N \int_v^B \frac{d\mathcal{G}(x)}{n_x} = \left(-m'_v\right) \left(Ny(v) - P_l\right) < 0, \tag{48}$$

where the last inequality follows from the fact that $-m'_v > 0$ and (A2): since $P_j - P_{j-1} > Ny(v)$ implies

$$P_l > Nly(v) \geq Ny(v).$$

Analyzing the difference between $\frac{dc_l(v)}{dv}$ and $\frac{dc_{l+1}(v)}{dv}$, we obtain

$$\begin{aligned}
\frac{dc_l(v)}{dv} - \frac{dc_{l+1}(v)}{dv} &= -m'_v \left(P_{l+1} - P_l\right) + n'_v g\left(T_l^{\mathrm{NE}}(v)\right) \\
&\quad - n'_v g\left(T_{l+1}^{\mathrm{NE}}(v)\right) \\
&> -m'_v \left(P_{l+1} - P_l\right) - n'_v g\left(T_{l+1}^{\mathrm{NE}}(v)\right) \\
&> -m'_v \left(P_{l+1} - P_l\right) - n'_v N \int_v^B \frac{d\mathcal{G}(x)}{n_x} \\
&= -m_v \left[\left(P_{l+1} - P_l\right) - Ny(v)\right] > 0,
\end{aligned} \tag{49}$$

where the last inequality follows from (A2), while the third in the third step we use Lemma 3 and the property that $g(\cdot)$ is increasing. This proves part 1 and 2 of the Lemma.

For part 3, take $l-1$, l, and $l+1$ in $\mathcal{L}$. First note that $c_{l-1}(A) < c_l(A)$ and $c_{l-1}(B) > c_l(B)$ from Lemma 4. Since, $c_l(v)$ is continuous and $\frac{dc_l(v)}{dv} > \frac{dc_{l+1}(v)}{dv}$, there exists exactly one v_l at which the two functions, namely, $c_{l-1}(v)$ and $c_l(v)$, meet. Further, note that $A < v_1$ because $c_0(A) < c_1(A)$ and, similarly, $v_{L-1} < B$ because $c_{L-2}(B) > c_{L-1}(B)$ by Lemma 4.

It only remains to show that $v_l < v_{l+1}$. Using part 1 and 2 of the Lemma it is clear that for $v < v_l$ we have $c_{l-1}(v) < c_l(v)$ and for $v > v_l$ we have $c_l(v) < c_{l-1}(v)$. Now, if $v_{l+1} \leq v_l$ then for all $v \in (v_{l+1}B)$, and hence for all $v \in [v_l, B)$, $c_{l+1}(v) < c_l(v)$. Thus, over the interval $[A, v_l)$ we shall have $c_{l-1}(v) < c_l(v)$ and over the interval $[v_l, B)$ we shall have $c_{l+1}(v) < c_l(v)$. This implies that no customer will join Queue l contradicting (A3). □

We now prove that joining Queue l is the best strategy for all $v \in (v_l, v_{l+1}]$. For this, it needs to be verified that $c_l(v) \leq c_j(v)$ for $v \in (v_l, v_{l+1}]$ for all $j \in \{0, 1, \ldots, L-1\}$. The following lemma provides some sufficient conditions for it. We later show that the optimal cost functions $c_l(v)$ satisfy these sufficient conditions.

Lemma 6 *If*

$$c_l(v) \leq c_{l-1}(v), \quad \textit{for all} \ \ v \geq v_l, \tag{50}$$

and

$$c_l(v) \leq c_{l+1}(v), \quad \textit{for all} \ \ v \leq v_{l+1}, \tag{51}$$

then $c_l(v) \leq c_j(v)$ *for* $v \in (v_l, v_{l+1})$, *for all* $j \in \{0, 1, \ldots, L-1\}$.

Proof Take $l \in \{0, 1, \ldots, L-1\}$ and a $v \in (v_l, v_{l+1}]$. Then

$$c_l(v) \leq c_{l+1}(v) \leq c_{l+2}(v) \leq \cdots c_{L-1}(v), \tag{52}$$

due to (51) and the fact that $v_j < v_{j+1}$. Also, by (50),

$$c_l(v) \leq c_{l-1}(v) \leq c_{l-2}(v) \leq \cdots c_0(v). \tag{53}$$

Thus, $c_l(v) \leq c_j(v)$ for all $j \in \{0, 1, \ldots, L-1\}$. □

We now show, by using Lemma 5, that the conditions of Lemma 6 are indeed satisfied and, thus, it is true that $c_l(v) \leq c_j(v)$ for $v \in (v_l, v_{l+1})$, for all $j \in \{0, 1, \ldots, L-1\}$.

1. Note that $c_0(v_1) = c_1(v_1)$ and $\frac{dc_0(v)}{dv} > 0$ and $\frac{dc_1(v)}{dv} < 0$, for $v \in (v_0, v_1)$. Take $v \in (v_0, v_1)$. By Taylor series expansion, for some $y \in (v, v_1)$, we have

$$c_0(v) = c_0(v_1) + (v - v_1) \left.\frac{dc_0(v)}{dv}\right|_{v=y} < c_0(v_1), \tag{54}$$

   where the last inequality follows because $(v - v_1) < 0$ and $\left.\frac{dc_0(v)}{dv}\right|_{v=y} > 0$. Similarly, for some $y \in (v, v_1)$, we have

$$c_1(v) = c_1(v_1) + (v - v_1) \left.\frac{dc_1(v)}{dv}\right|_{v=y} > c_0(v_1), \tag{55}$$

   where the last inequality holds as $(v - v_1) < 0$ and $\left.\frac{dc_1(v)}{dv}\right|_{v=y} < 0$. Thus, from (54) and (55),

$$c_0(v) < c_1(v),$$

   for all $v \in (v_0, v_1)$.
2. Take $l \in \{1, 2, \ldots, L-2\}$. Take a $v > v_l$. By mean value theorem, there exists a $z \in (v_l, v)$ such that

$$\frac{c_l(v_l) - c_l(v)}{c_{l-1}(v_l) - c_{l-1}(v)} = \frac{\frac{dc_l(z)}{dz}}{\frac{dc_{l-1}(z)}{dz}}. \tag{56}$$

   Since, $\frac{dc_l(z)}{dz} < \frac{dc_{l-1}(z)}{dz} < 0$ from Lemma 5, we have $\frac{\frac{dc_l(z)}{dz}}{\frac{dc_{l-1}(z)}{dz}} > 1$. Using this reduces (56) to

$$c_l(v_l) - c_l(v) > c_{l-1}(v_l) - c_{l-1}(v),$$

   which is nothing but

$$c_{l-1}(v) > c_l(v_l), \tag{57}$$

   since $c_l(v_l) = c_{l-1}(v_l)$.
   Similarly, if we take a $v < v_{l+1}$ there exists a $z \in (v_l, v)$, by the mean value theorem, such that

$$\frac{c_l(v) - c_l(v_{l+1})}{c_{l+1}(v) - c_{l+1}(v_{l+1})} = \frac{\frac{dc_l(z)}{dz}}{\frac{dc_{l+1}(z)}{dz}},$$

   which reduces to

$$c_l(v) - c_l(v_{l+1}) < c_{l+1}(v) - c_{l+1}(v_{l+1}), \tag{58}$$

   by using Lemma 5; but (58) is nothing but

$$c_l(v) < c_{l+1}(v), \tag{59}$$

   for all $v < v_{l+1}$.

This proves all the conditions of Lemma 6.

F Proof of Theorem 3

Take $v \in (v_l, v_{l+1}]$, then

$$\begin{aligned} F_l\left(T_l^{\mathrm{NE}}(v)\right) &= \int \mathbb{1}_{T_l^{\mathrm{NE}}(x) \geq T_l^{\mathrm{NE}}(v)} q_l^{\mathrm{NE}}(v) d\mathcal{G}(v) \\ &= \int \mathbb{1}_{x \leq v} \mathbb{1}_{v_l < v \leq v_{l+1}} d\mathcal{G}(v) = \mathcal{G}(v) - \mathcal{G}(v_l), \end{aligned} \tag{60}$$

where the second equality follows from Lemma 2 and Theorem 2. Differentiating (60) w.r.t. v, we get

$$\frac{d\mathcal{G}(v)}{dv} = \frac{dF_l\left(T_l^{\mathrm{NE}}(v)\right)}{dv} = \frac{dT_l^{\mathrm{NE}}(v)}{dv} \times \left.\frac{dF_l(t)}{dt}\right|_{t=T_l^{\mathrm{NE}}(v)}. \tag{61}$$

For optimality, $T_l^{\mathrm{NE}}(v)$ must also satisfy the first order condition, namely, $\frac{dc_v(l,t)}{dt} = 0$. This gives

$$0 = \left[N\frac{dF_l(t)}{dt} + n_v\frac{dg(t)}{dt}\right]_{t=T_l^{\mathrm{NE}}(v)}. \tag{62}$$

Substituting (61) in (62), we get

$$\frac{dg\left(T_l^{\mathrm{NE}}(v)\right)}{dv} = -\frac{N}{n_v}\frac{d\mathcal{G}(v)}{dv}.$$

This gives

$$T_l^{\mathrm{NE}}(v) = g^{-1}\left(N\int_v^{v_{l+1}}\frac{d\mathcal{G}(x)}{n_x} + \alpha\right),$$

where α is the integration constant that can be shown to equal 0 using the same line of arguments as in Appendix B while deriving $w^{\mathrm{NE}}(v)$ in explicit form.

Further, since $T_l^{\mathrm{NE}}(v)$ is a decreasing function by Lemma 2, $T_l^{\mathrm{NE}}(v) = 0$ for all $v \geq v_{l+1}$. Hence, for $v \geq v_{l+1}$

$$\begin{aligned} F_l\left(T_l^{\mathrm{NE}}(v)\right) = F_l(0) &= \int \mathbb{1}_{T_l^{\mathrm{NE}}(x)\geq 0}\mathbb{1}_{v_l<v\leq v_{l+1}}d\mathcal{G}(v) \\ &= \int \mathbb{1}_{v_l<v\leq v_{l+1}}d\mathcal{G}(v) \\ &= \mathcal{G}(v_{l+1}) - \mathcal{G}(v_l). \end{aligned} \tag{63}$$

Also, note that at Queue l arrival of customer $v = v_l$ is the earliest. Thus, $F_l\left(T_l^{\mathrm{NE}}(v_l)\right) = 0$. Now, since $F_l\left(T_l^{\mathrm{NE}}(\cdot)\right)$ is an increasing function $F_l\left(T_l^{\mathrm{NE}}(v_l)\right) = 0$ for all $v < v_l$. Thus, any customer $v < v_l$ in order to minimize her waiting cost has to choose $T_l^{\mathrm{NE}}(v) = T_l^{\mathrm{NE}}(v_l)$. This proves the first part of Theorem 3 which characterizes $T_l^{\mathrm{NE}}(v)$ completely. Notice that, we have also shown that $T_l^{\mathrm{NE}}(v)$ satisfies

$$F_l\left(T_l^{\mathrm{NE}}(v)\right) = \begin{cases} F_l(0) = \mathcal{G}(v_{l+1}) - \mathcal{G}(v_l), & \text{for } \ v > v_{l+1} \\ \mathcal{G}(v) - \mathcal{G}(v_l), & \text{for } \ v_l \leq v \leq v_{l+1} \\ 0, & \text{for } \ v < v_l \end{cases}. \tag{64}$$

For obtaining the thresholds $\{v_l\}_{l=1}^{L-1}$ we look at the optimal queue joining cost. The optimal joining cost for Queue l is

$$c_l(v) = N\left[F_l\left(T_l^{\mathrm{NE}}(v)\right) + \sum_{j=l+1}^{L-1}F_j(0)\right] + m_vP_l + n_vg\left(T_l^{\mathrm{NE}}(v)\right),$$

which, by using (64), reduces to

$$c_l(v) = N(\mathcal{G}(v) - \mathcal{G}(v_{l-1})) + N(1 - \mathcal{G}(v_l)) + m_vP_l + n_vN\int_v^{v_l}.\frac{d\mathcal{G}(x)}{n_x}. \tag{65}$$

Now, since v_l is the point of intersection of $c_l(v)$ and $c_{l-1}(v)$, at $v = v_l$ we should have

$$\begin{aligned} 0 &= c_l(v_l) - c_{l-1}(v_l) \\ &= N(1 - \mathcal{G}(v_{l+1})) + m_{v_l}P_l + n_{v_l}N\int_{v_l}^{v_l}\frac{d\mathcal{G}(x)}{n_x} - N(1 - \mathcal{G}(v_{l-1})) - m_{v_l}P_{l-1}. \end{aligned}$$

This can be re-written as (18).

G Derivation of the Revenue Function

We can re-write (18) as

$$P_l - P_{l-1} = \frac{N}{m_{v_l}} \left[\mathcal{G}(v_{l+1}) - \mathcal{G}(v_{l-1}) - v_l \int_{v_l}^{v_{l+1}} \frac{d\mathcal{G}(x)}{x} \right], \tag{66}$$

for all $l \in \{1, 2, \dots L-1\}$. Adding (66) for all $l = 1$ to $L-1$, we get

$$P_{L-1} = \sum_{l=1}^{L-1} \frac{N}{m_{v_l}} \left[\mathcal{G}(v_{l+1}) - \mathcal{G}(v_{l-1}) - v_l \int_{v_l}^{v_{l+1}} \frac{d\mathcal{G}(x)}{x} \right]. \tag{67}$$

We use $\mathbf{P}$ to denote the price vector $(P_1, \dots, P_{L-1})$. Revenue function can be expanded as

$$\frac{1}{N} R(\mathbf{P}) = \sum_{l=1}^{L-1} \left[\mathcal{G}(v_{l+1}) - \mathcal{G}(v_l) \right] P_l = P_{L-1} - \sum_{l=1}^{L-1} (P_l - P_{l-1}) \, \mathcal{G}(v_l). \tag{68}$$

Substituting (66) and (67), we obtain

$$\begin{aligned} \frac{1}{N} R(\mathbf{P}) = \sum_{l=1}^{L-1} \frac{N}{m_{v_l}} & \left[\mathcal{G}(v_{l+1}) - \mathcal{G}(v_{l-1}) - v_l \int_{v_l}^{v_{l+1}} \frac{d\mathcal{G}(x)}{x} \right] \\ & - \sum_{l=1}^{L-1} \mathcal{G}(v_l) \frac{N}{m_{v_l}} \left[\mathcal{G}(v_{l+1}) - \mathcal{G}(v_{l-1}) - v_l \int_{v_l}^{v_{l+1}} \frac{d\mathcal{G}(x)}{x} \right], \end{aligned} \tag{69}$$

which can be written as

$$\frac{1}{N} R(\mathbf{P}) = \sum_{l=1}^{L-1} (1 - \mathcal{G}(v_l)) \, \frac{N}{m_{v_l}} \times \left[\mathcal{G}(v_{l+1}) - \mathcal{G}(v_{l-1}) - v_l \int_{v_l}^{v_{l+1}} \frac{d\mathcal{G}(x)}{x} \right]. \tag{70}$$

This can be simplified as

$$\begin{aligned} \frac{1}{N} R(\mathbf{P}) = \sum_{l=1}^{L-1} \frac{N}{m_{v_l}} & \left[\mathcal{G}(v_{l+1}) - v_l \int_{v_l}^{v_{l+1}} \frac{d\mathcal{G}(x)}{x} \right] (1 - \mathcal{G}(v_l)) \\ & - \sum_{l=1}^{L-1} (1 - \mathcal{G}(v_l)) \, \frac{N}{m_{v_l}} \mathcal{G}(v_{l-1}). \end{aligned} \tag{71}$$

Note that the first term in the second sum is 0. Hence, by increasing index of the second summation this can be re-written as

$$\begin{aligned} \frac{1}{N} R(\mathbf{P}) = \sum_{l=1}^{L-1} \frac{N}{m_{v_l}} & \left[\mathcal{G}(v_{l+1}) - v_l \int_{v_l}^{v_{l+1}} \frac{d\mathcal{G}(x)}{x} \right] (1 - \mathcal{G}(v_l)) \\ & - \sum_{l=1}^{L-2} (1 - \mathcal{G}(v_{l+1})) \, \frac{N}{m_{v_{l+1}}} \mathcal{G}(v_l). \end{aligned} \tag{72}$$

This is noting but $\frac{1}{N} R(\mathbf{P}) = \sum_{l=1}^{L-1} u(v_l, v_{l+1})$.

References

1. P. Afeche and H. Mendelson. Pricing and priority auctions in queueing systems with a generalized delay cost structure. *Management Science*, 50(7), July 2004.
2. K. R. Balachandran. Purchasing priorities in queues. *Management Science*, 18(5):319–326, 1972.
3. C. Courcoubetis and R. Weber *Pricing Communication Networks: Economics, Technology and Modeling* John Wiley and Sons, 2003.
4. S. Doroudi, M. Akan, M. Harchol-Balter, J. Karp, C. Borgs, and J. T. Chayes. Priority pricing in queues with a continuous distribution of customer valuations. Technical Report CM-CS-13-109, Computer Science Department, Carnegie Mellon University, May 2013.
5. A. Glazer and R. Hassin. Stable priority purchasing in queues. *Operations Research Letters*, 4(6):285–288, April 1986.
6. O. Gueant, J. M. Lasry, and P. L. Lions. *Mean field games and applications*. Springer: Paris-Princeton Lectures on Mathematical Finance, 2010.
7. H. Honnappa and R. Jain. Strategic arrivals into queueing networks: The network concert queueing game. *arXiv:1112.3052*, Dec. 2011.
8. R. Jain, S. Juneja, and N. Shimkin. The concert queuing problem: To wait or to be late. *Discrete Events Dynamic Systems*, 21(1):103–138, Mar. 2011.
9. S. Juneja, T. Raheja, and N. Shimkin. The concert queueing game with a random volume of arrivals. In *Proc. VALUETOOLS*, pages 317–325, Oct. 2012.
10. S. Juneja and N. Shimkin. The concert queueing game: strategic arrivals with waiting and tardiness costs. *Queueing Syst.*, 74(4):369–402, Aug. 2013.
11. M. A. Khan, K. P. Rath, and Y. Sun. On the existence of pure strategy equilibria in games with a continuum of players. *Journal of Economic Theory*, 76(1):13–46, Sep. 1997.
12. T. Kittsteiner and B. Moldovanu. Priority auctions and queue disciplines that depend on processing time. *Management Science*, 51(2):236–248, 2005.
13. L. Kleinrock. Optimal bribing for queue position. *Operation Research*, 15, 1967.
14. M. Lariviere and J. A. Van Mieghem, ‘ Strategically seeking service: How competition can generate Poisson arrivals
Manufacturing & Service Operations Management, 6(1), Winter 2004, pp.23–40.
15. F. Lui. An equilibrium model of bribery. *Political Economy*, 93, 1985.
16. M. Marchand. Priority pricing. *Management Science*, 20(7):1131–1140, 1974.
17. A. Mas-Colell. On a theorem of Schmeidler. *Journal of Mathematical Economics*, 13(3):201–206, Dec. 1984.
18. A. W. Matyszkiel. Existence of pure equilibria in games with nonatomic space of players. *Topological Methods in Nonlinear Analysis*, 16(2):339–349, Dec. 2000.
19. S. Rao and E. Petersen. Optimal pricing of priority services. *Operations Research*, 46(1):46–56, 1998.
20. S. Shakkottai, R. Srikant, A. Ozdaglar, and D. Acemoglu, The Price of Simplicity. *IEEE Journal on Selected Areas in Communication,* vol. 26, no. 7, Sept 2008.