Transportation Research Part B 44 (2010) 1166–1185

Contents lists available at ScienceDirect

Transportation Research Part B

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

Endogenous arrivals in batch queues with constant or variable capacity

Amnon Rapoport [a,b,*], William E. Stein [c], Vincent Mak [d], Rami Zwick [b], Darryl A. Seale [e]

[a] Eller College of Management, University of Arizona, Tucson, AZ 85721, United States
[b] Department of Management and Marketing, University of California, Riverside, Riverside, CA 92521, United States
[c] Mays Business School, Texas A&M University, College Station, TX 77843, United States
[d] Cambridge Judge Business School, University of Cambridge, Cambridge CB2 1AG, United Kingdom
[e] Department of Management, University of Nevada Las Vegas, Las Vegas, NV 89154, United States

ARTICLE INFO

Article history:
Received 19 March 2009
Received in revised form 19 December 2009
Accepted 7 January 2010

JEL classification:
C72
C92

Keywords:
Batch queueing
Game theory
Constant and variable capacity
Experiments

ABSTRACT

We study batch queueing systems with continuous time, finite commuter populations, single server, and endogenously determined arrival times. Symmetric equilibrium solutions in mixed strategies are constructed and subsequently tested in two experiments that examine two different batch queueing models, one with a fixed server capacity, and the other with a variable server capacity. With experience in playing the stage queueing game repeatedly, experimental results from groups of 20 subjects support equilibrium play on the aggregate level when the server capacity is fixed and commonly known. When it is known to be variable, randomly changing from round to round, subjects diverge from equilibrium play and increase their individual payoffs substantially by significantly shortening their waiting time.

1. Introduction

Although queues have considerable economic significance, little is known about the behavioral patterns that govern the decisions to join queues or stay out of them. The present study investigates decisions in *batch queues*, where commuters arrive at a facility that serves them in a batch of fixed size. Arrival times are endogenously determined. In addition to a fixed joining (entry) cost, there is also a variable waiting cost that depends on the individual waiting time. Ferry services, inter-urban buses, and air shuttles share these characteristics.

Batch service queueing systems have been studied in several forms in the operations research, economics, and transportation literature. The models that have been proposed differ from one another on several dimensions including scheduled vs. unscheduled service times, single vs. multiple service cycles, and observable vs. unobservable queues (e.g., Hassin and Haviv, 2003). Batch queueing systems with stochastic service time have been studied theoretically in operations research by Bailey (1954), Chaudhry and Templeton (1983), Mehdi (1975), Neuts (1967), and Weiss (1981) among others. Batch queueing systems with deterministic service time have been studied by Barnett (1973), Chaudhry and Templeton (1983), Kosten (1973), and Glazer and Hassin (1987). Hassin and Haviv (2003) survey equilibrium solutions for several queueing systems of interest, including batch queueing systems.

* Corresponding author. Address: Department of Management and Marketing, University of California Riverside, Riverside, CA 92521, United States. Tel.: +1 951 398 7108.
E-mail address: amnon@u.arizona.edu (A. Rapoport).

0191-2615/$ - see front matter
doi:10.1016/j.trb.2010.01.005

Previous analysis of batch queueing models is theoretical; the arrival time distribution is typically assumed to be exogenously determined, and the calling population of commuters is assumed to be infinite. We depart from these assumptions in three major ways. First, we study two queueing models where arrival times are *endogenously determined*. In one model the server capacity is fixed and in the other it is variable. Viewing the arrival times as an interactive decision process, we derive the equilibrium solutions to these two cases and illustrate them. Second, our two models assume a *finite* and commonly known number of commuters rather than infinitely large populations. A finite population better describes situations where a small group of travelers repeatedly use the same batch queueing facility. For example, a university provides its faculty and staff with a subsidized interurban bus that leaves the campus at a pre-designated time after working hours. An employee knows the size of the potential demand; however, she is uncertain as to how many other employees intend to use alternative modes of transportation (i.e., not joining the queue) at any given day and at what time they intend to join the queue. Due to the repeated nature of the interaction, opportunities for learning and coordination abound. For another example, if multiple stops occur on the way then both cases of fixed and variable server capacity occur in scheduled interurban bus and ferry services. The vehicle capacity is fixed for commuters waiting to be served in the first stop. Because the number of commuters embarking on the vehicle along the way is unknown, random capacity occurs thereafter.

The third departure reflects a major purpose of the present study to identify and explain *behavioral regularities* that emerge when financially motivated players have to decide independently whether to join the queue, and if so, at what time. To achieve this purpose, using standard methods of experimental economics (e.g., Camerer, 2003; Kagel and Roth, 1995), we construct two different batch queueing games, have subjects participate in these games for financial rewards contingent on their performance, and—using the symmetric mixed-strategy equilibrium solution as a benchmark that drives the analysis—compare observed with equilibrium behavior. Except for some very simple games (mostly ones with dominated strategies) in which players can arrive at equilibrium play by introspection, the consensus is that equilibrium behavior, if reached at all, is learned with experience. Consequently, to assess the descriptive power of the equilibrium solutions, the stage game in our experiment is iterated multiple times with the same population of subjects. Multiple iterations of the game allow us to determine if players converge to or diverge from equilibrium play with experience.

1.1. Previous research

Several studies are related to the present study. The first is by Holt and Sherman (1982), who proposed a simple model of waiting-line auctions formulated as a non-cooperative game with incomplete information. Similar to our two models, in their model time is continuous, cost of waiting is linear, and time of arrival is a choice variable. However, their model is not concerned with batch queueing. Glazer and Hassin (1987) considered batch queueing service systems with deterministic service times and fixed time intervals (cycles) between any two successive service starts. In their model, if the queue length exceeds the server capacity s at the scheduled service time, then s customers are served and the remaining customers have to wait for the next cycle. Their model, too, differs from ours in several respects. First, it places no limit on the population size. Second, although arrival time is a choice variable, Glazer and Hassin assume that the number of customer arrivals during a cycle follows a Poisson distribution with a known mean. Focusing on endogenous arrivals, our model makes no such assumption. Further, both studies by Holt and Sherman and by Glazer and Hassin are not concerned with empirical behavior. A third study is by Stein et al. (2007), who investigated the decisions players make in two batch service queueing games with endogenously determined arrivals, one allowing for balking and the other does not. Stein et al. constructed and experimentally tested the symmetric mixed-strategy equilibrium solutions of each of these two games under two information conditions, one with private information and the other with public information about previous outcomes. With repeated iterations of the stage game, all four experimental conditions (private/public information by balking/no balking) resulted in aggregate behavior approaching equilibrium play. Additional experimental studies by Rapoport et al. (2004), Seale et al. (2005), Pazgal and Radas (2008), Ziegelmeyer et al. (2008), and most recently Daniel et al. (2009) have also studied endogenous arrivals in networks and reported support for equilibrium play. In contrast to the present study, they investigated different queueing service systems in which customers are served one at a time.

There is a large body of research in public transportation that assumes endogenous departure times and is closely related to our study. In his classical and influential paper, Vickrey (1969) proposed a highway bottleneck model in which commuters face a tradeoff between the cost of arriving at work at other time than the most preferred time and the cost of time spent waiting in the queue that forms behind the bottleneck. Subsequent work by Arnott et al. (1989, 1993) have contributed to the development of Vickrey's model. These models assume endogenous departure times, as in our present model, but do not restrict arrivals to be in batch. More relevant to our study are transportation models on commuters' departure time choice for urban mass transit services with a single origin and a single destination. De Palma and Lindsey (2001) analyze the optimal timetable for public transport vehicles on a single transit line with no intermediate stops. Kraus and Yoshida (2002) and Kraus (2003) provide economic analysis of the commuters' time-of-use decision, the optimal pricing, and the service in an urban transit system with limited capacity. More recently, Tian et al. (2007) developed an equilibrium model of peak-period commuting for a mass transit line with multiple origins and a single destination. All of these models differ from our model in important details and none of these papers is concerned with experimentally assessing the descriptive power of their models.

The present study departs from the previous experimental work of Rapoport et al., Seale et al., and Stein et al. on arrival times in queues in three significant ways. First, to avoid the need for a tie-breaking rule for multiple arrivals at the same time, we propose two models that assume continuous rather than discrete time. Second, our experiments are conducted in "real time". Consequently, subjects could experience the time pressure before joining the queue and waiting time after arrival (if at all). This element greatly enhances the realistic nature of the task and as such contributes to the external validity of our study. Our third and most important departure is an investigation of batch queues with variable, rather than fixed server capacity. As we show below, we observe convergence to equilibrium play when the server capacity is constant but divergence from it when it is variable.

The paper is organized as follows: Section 2 presents two batch queueing models and their equilibrium solutions. Both models assume continuous arrival time, finite and commonly known populations, and no balking. In the first model, server capacity is constant and in the second it is a random variable. Section 3 describes two experiments designed to test the *constant* and *variable* capacity models. Section 4 concludes with a discussion of the main results. All the proofs are relegated to an appendix.

2. Two batch queueing systems

A population of n players wish to receive service scheduled at time T. Service time is assumed to be 0, arrival time is continuous, and reneging and balking are prohibited (queue is unobservable[1]). Each player must decide independently whether to join the queue and if so at what time to arrive. The queue discipline is First Come First Served (FCFS).

- *Service capacity*: We study two systems, one where the service capacity is fixed and commonly known to be s, and the other where s is a random variable with a commonly known distribution.
- *Payoff structure:* Each player who joins the queue is charged a fixed entry fee d, and a variable cost c per minute of waiting in the queue until time T. If she receives service, then she is awarded a fixed payoff r. A player who stays out of the queue receives a fixed payoff g ($g < r - d$). The payoff function, the same for each player i ($i = 1, \ldots, n$), takes the form:

$$H_i = \begin{cases} g, & \text{if player } i \text{ stays out of the queue}, \\ -d - ct_i, & \text{if player } i \text{ waits } t_i \text{ minutes without completing service}, \\ r - d - ct_i, & \text{if player } i \text{ waits } t_i \text{ minutes and completes service}, \end{cases}$$

where t_i is the time player i waits for service. The parameter values c, d, and r are assumed to be positive, whereas the payoff g for staying out may assume any real value.

In contrast to the bottleneck models in transportation research (e.g., Arnott et al., 1989, 1993) and mass transit transportation (e.g., Kraus and Yoshida, 2002), our model provides the players with the option of staying out of the queue and select an alternative mode of transportation.

2.1. Constant server capacity

The unique mixed-strategy equilibrium for the constant capacity model (see Appendix for proof) has a positive density on the time interval from $T_{\min}$ to T, where $T_{\min}$ is the earliest feasible arrival time. It can be shown that $T_{\min} = T - \frac{rw}{60c}$ with $w = 1 - (d + g)/r$. For any t in the range $T_{\min} \leqslant t \leqslant T$, $\tau(t) = \frac{t - T_{\min}}{T - T_{\min}}$ defines a time scaled to be between 0 and 1. Then, the equilibrium CDF, denoted by F, is given by

$$F(t) = G^{-1}_{s,n-s}(wt),$$

where $G_{s,n-s}$ denotes the *Beta(s, n−s)* CDF. The probability of staying out is given by:

$$P(\text{staying out}) = 1 - F(T) = 1 - G^{-1}_{s,n-s}(w).$$

Example 1. Let $n = 20$, $s = 14$, $T = 12{:}00$, $g = 60$, $d = 40$, $r = 340$, and $c = 4$. These are the parameter values used in the fixed capacity experiment (Section 3). Then, $w = 1 - (d + g)/r = 0.7059$, $T_{\min} = 11{:}00$, and $P(\text{staying out}) = 0.2397$. Fig. 1A exhibits the equilibrium probability density function. Although Fig. 1A may give the impression that there is a positive probability of simultaneously arriving at 11:00, this is not the case. In fact, there is a probability 0.32 of joining the queue within one second after 11:00.

2.2. Variable server capacity

Assume that the server capacity S is a random variable with a commonly known distribution: Let $b_s = P(S = s)$, $s = 1, \ldots, n$. In the Appendix, we prove that in equilibrium:

[1] Queues for large batch services such as free concerts in a stadium are considered non-observable because counting how many people are ahead in the queue is not practical. In such cases, those who join the queue, unless they are clearly at the head of the queue, learn if they can enter the stadium only after the gates are open. They incur joining and waiting costs whether or not service is awarded. Such queues are characterized as no-balking.

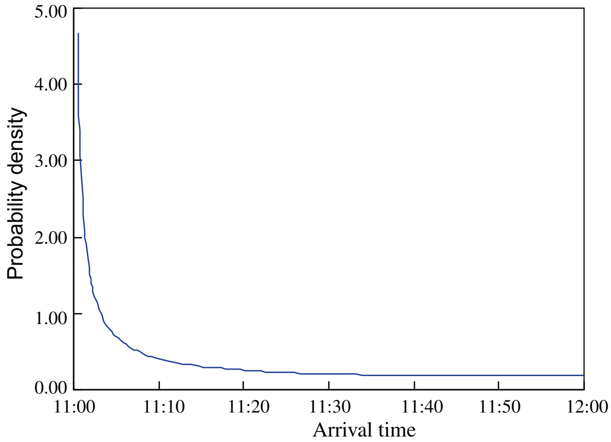


Fig. 1A. Equilibrium density function of arrival times: Experiment CON.

$$F^{-1}(a) = \frac{1}{w}\sum_{s=1}^{n-1} b_s G_{s,n-s}(a)$$

for $0 \leqslant a \leqslant \min(1, F(1))$. The right-hand side of the above equation is evaluated numerically, yielding F.

Example 2. Let $n = 20$, $T = 12{:}00$, $g = 60$, $d = 40$, $r = 340$, and $c = 4$ as in Example 1 with b_s taking one of two values with equal probability: $b_{10} = 0.5$ and $b_{16} = 0.5$. These are the parameter values used in one of the two conditions in the variable capacity experiment (see Section 3). Fig. 1B (top panel) displays the resulting probability density function, which is slightly bimodal. In this case, P(staying out) = 0.21.

Example 3. Same parameters as in Example 2 except that $b_8 = 0.5$ and $b_{18} = 0.5$. These are the parameter values used in the second condition in the variable capacity experiment. Fig. 1B (bottom panel) displays the resulting probability density function, which is evidently bimodal. In this case, P(staying out) = 0.1025.

3. Experiments

The experiments reported in this section were designed to assess the descriptive power of the equilibrium solutions for the constant and variable service capacity models, and uncover replicable behavioral regularities that govern arrival and staying out decisions. Except for the difference between fixed and variable capacity, they used the same experimental procedures.

3.1. Methods

3.1.1. Subjects

Two hundred and forty subjects volunteered to participate in an interactive decision making experiment with payoff contingent on performance. The subjects were divided into 12 groups of 20 members each. Each group of 20 subjects participated in a single computer-controlled session that lasted about 100 minutes and included 50 iterations of the stage game. Four groups participated in the constant capacity experiment (Experiment CON) and eight other groups (four in each of two conditions) in the variable capacity experiment (Experiment VAR).

3.1.2. Procedure

Upon arrival at the laboratory, group members were randomly assigned to separate cubicles and handed written instructions.[2] In both experiments, the batch queueing game was described to the subjects as a transportation problem with a ferry of limited capacity that departs once a day at a scheduled time (12:00) and a commuter population of fixed size ($n = 20$) that exceeds (any) ferry capacity. The capacity in Experiment CON was set at $s = 14$. There were two conditions in Experiment VAR. In Condition VAR(10, 16), the subjects were informed that one of two ferries, a small ferry with capacity $s_S = 10$ or a large ferry with capacity $s_L = 16$, would arrive on any particular round with equal probability. Condition VAR(8, 18) had the same expected capacity as Condition VAR(10, 16) except that the small and large capacities were set at $s_S = 8$ and $s_L = 18$.

Each round was structured in the same way. A computer timer displayed how many minutes remained before the ferry's departure ($T = 12{:}00$). The timer began with 90 minutes remaining and then ran backwards to 0. Time was compressed to

[2] The subject instructions are presented in an on-line appendix. They also may be obtained directly from the authors.

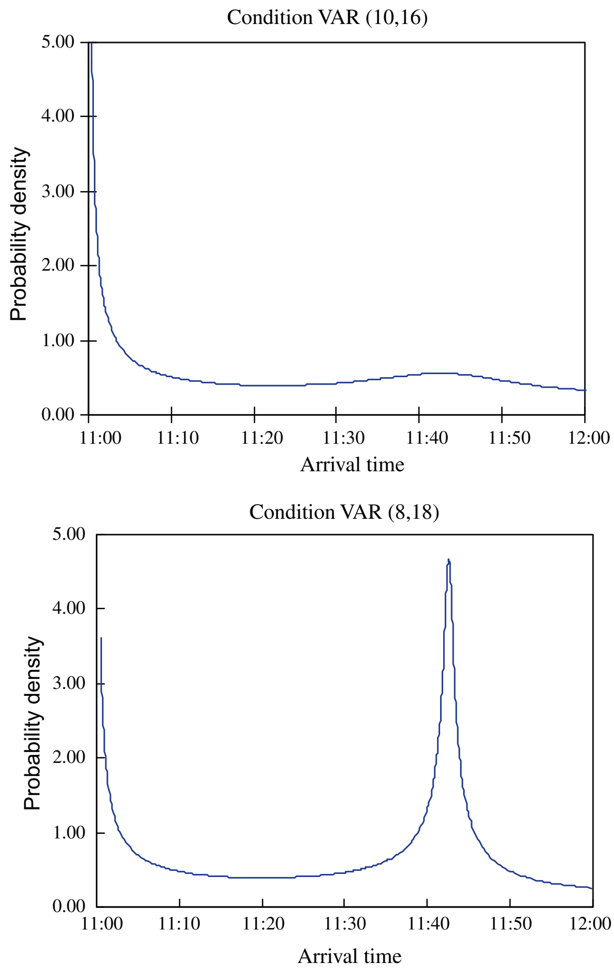


Fig. 1B. Equilibrium density function of arrival times: Experiment VAR.

1 minutes passing for every ½ s. Consequently, the timer ran from 90 to 0 minutes in about 45 s. A subject wishing to join the queue indicated the time of arrival by moving the mouse pointer out of a designated area on the computer screen. We implemented this procedure to prevent the noise of clicking on the mouse from informing other subjects of this subject's entry time decision. A subject wishing to stay out of the queue on any particular round was instructed to leave the mouse untouched.

At the end of each round, the computer screen displayed the following information on each of the 20 individual PC screens:

- The ferry capacity (only in Experiment VAR).
- The number of subjects who stayed out.
- Complete information about arrival times of subjects joining the queue and success in embarking on the ferry.[3]
- The subject's payoff for the round.
- The subject's cumulative payoff from round 1.

The parameters assumed the same values for all 50 rounds as in Examples 1–3 above: $T = 12{:}00$, $n = 20$, $g = 60$, $r = 340$, $d = 40$, and $c = 4$ (per minute). Subjects could arrive as early as 10:30. Payoffs and costs were stated in terms of a fictitious currency called "francs". At the end of the session, subjects were paid a \$5 show-up bonus plus their cumulative earnings across all 50 rounds converted into US dollars at the rate 10 francs = \$0.06. The mean payoff was \$16.50 in Experiment CON; in Experiment VAR, it was equal to \$17.89 and \$29.00 in Conditions VAR(10, 16) and VAR(8, 18), respectively.

[3] All arrival times were displayed in a descending order, with the times of the subjects who embarked on the ferry appearing on the screen in blue, and the times of the subjects who were denied service (if any) appearing in black. Individual players were not identified by name or subject number.

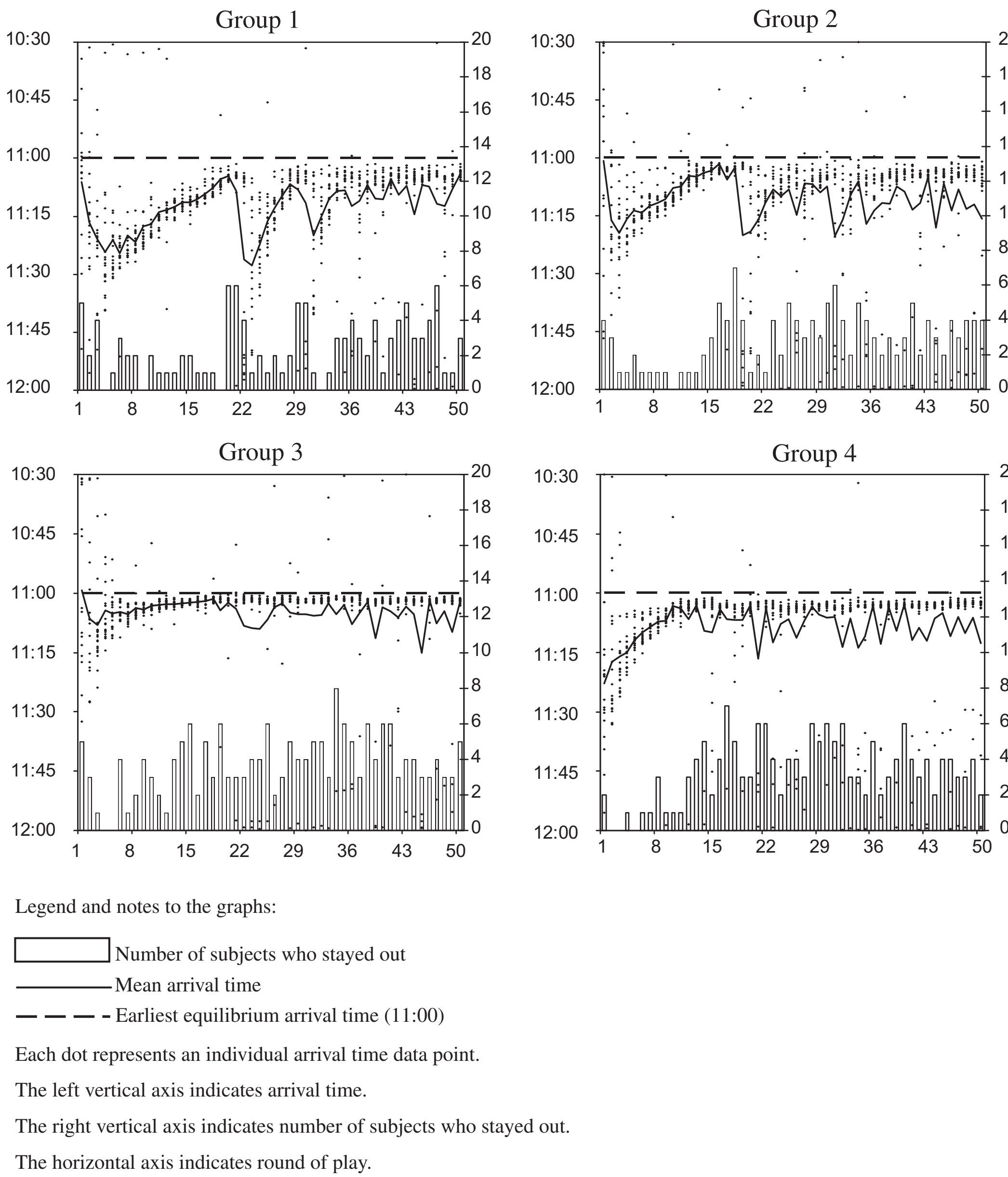
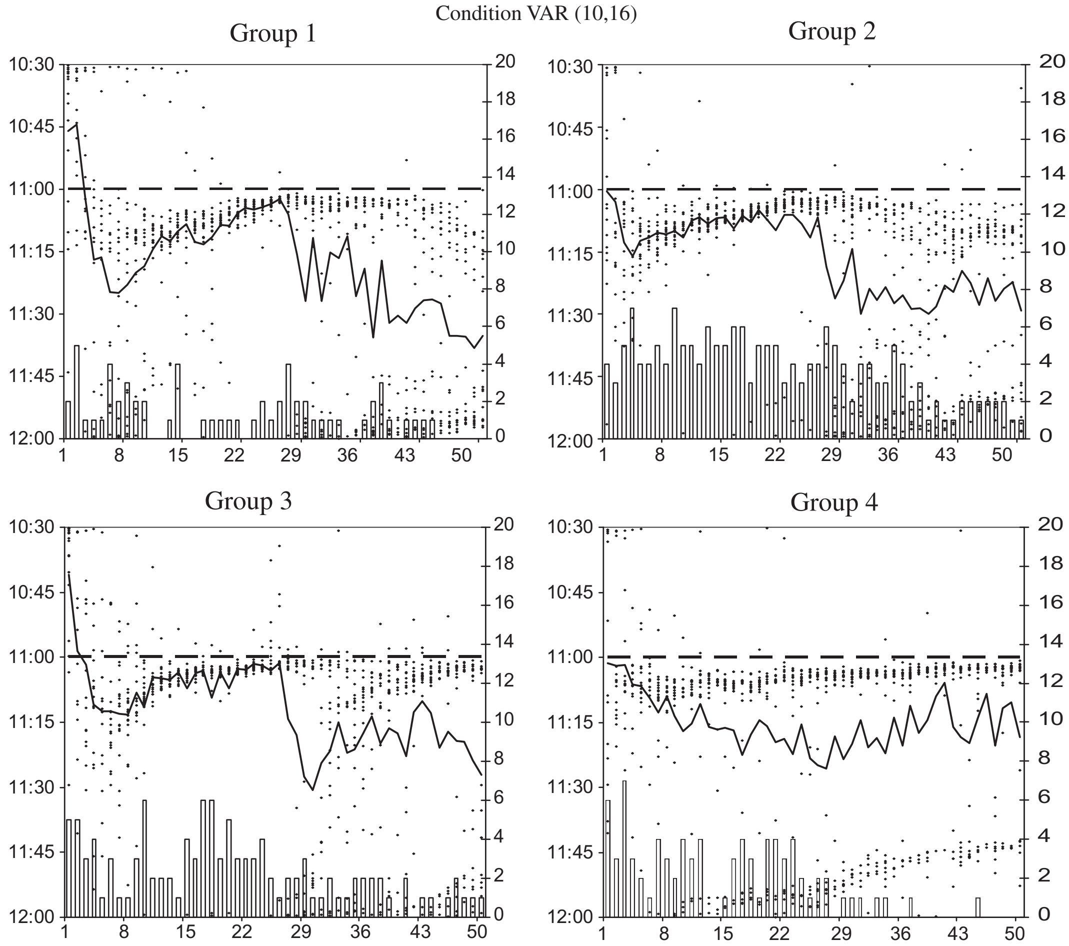


Fig. 2. Arrival times and decisions to stay out of individual subjects by group Experiment CON.

3.2. Results

We begin this section with a general description of the raw aggregate data. Figs. 2, 3A and 3B exhibit the arrival times and number of subjects who stayed out (bars) by experiment and condition (Fig. 2: Experiment CON, Fig. 3A: Condition VAR(10, 16), Fig. 3B: Condition VAR(8, 18)), round (1–50), and group (1–4). The horizontal axis depicts the round number (1–50). The left vertical axis shows the time arrival with 12:00 at the bottom and 10:30 at the top. The right vertical axis exhibits the frequency of subjects who stayed out in each round. Each dot represents the arrival time of a single subject. The mean arrival times (that can easily be translated to waiting times) are represented by solid lines. The dash lines at 11:00 represent the earliest equilibrium arrival time.[4] Figs. 2, 3A, and 3B show that only very few arrival times (less than

[4] If a player queues at exactly 11:00, then she is guaranteed receiving at most 60 "francs" (earn 340-40-60 $\times$ 4, if she is successful in embarking on the ferry, and lose 280, otherwise). Arrival times before 11:00 are clearly "errors," yielding a payoff that is smaller than the sure payoff g (=60) associated with a decision to stay out.



Legend and notes to the graphs:

Number of subjects who stayed out

Mean arrival time

Earliest equilibrium arrival time (11:00)

Each dot represents an individual arrival time data point.

The left vertical axis indicates arrival time.

The right vertical axis indicates number of subjects who stayed out.

The horizontal axis indicates round of play.

Fig. 3A. Arrival times and decisions to stay out of individual subjects by group Experiment VAR Condition VAR(10, 16).

5% in all cases) violated the earliest equilibrium arrival time predictions (dots above the horizontal dash line). Most of them occurred in the early rounds. Early arrivals before 11:00 occurred in all the 12 groups in both experiments.

We find very small differences between the four groups in each condition, but quite different patterns of behavior between the two treatments. Mean arrival time in Experiment CON (Fig. 2) initially falls in the first few rounds (except for Group 4) and then rises and continues to fluctuate in the 11:00–11:15 time interval. On the other hand, mean arrival time in Experiment VAR (Figs. 3A and 3B) exhibits an overall downward trend for the entire duration of the session.[5] The aggregate patterns of the decisions to stay out of the queue are also different between the two treatments. In Experiment CON (Fig. 2), fewer subjects decided to stay out in the beginning of the session compared to the same number in the later rounds. Exactly

[5] These trends refer to the two graphs. Because of the way the vertical axes in Figs. 2 and 3 are labeled, a down trend on the graph indicates later arrival times.

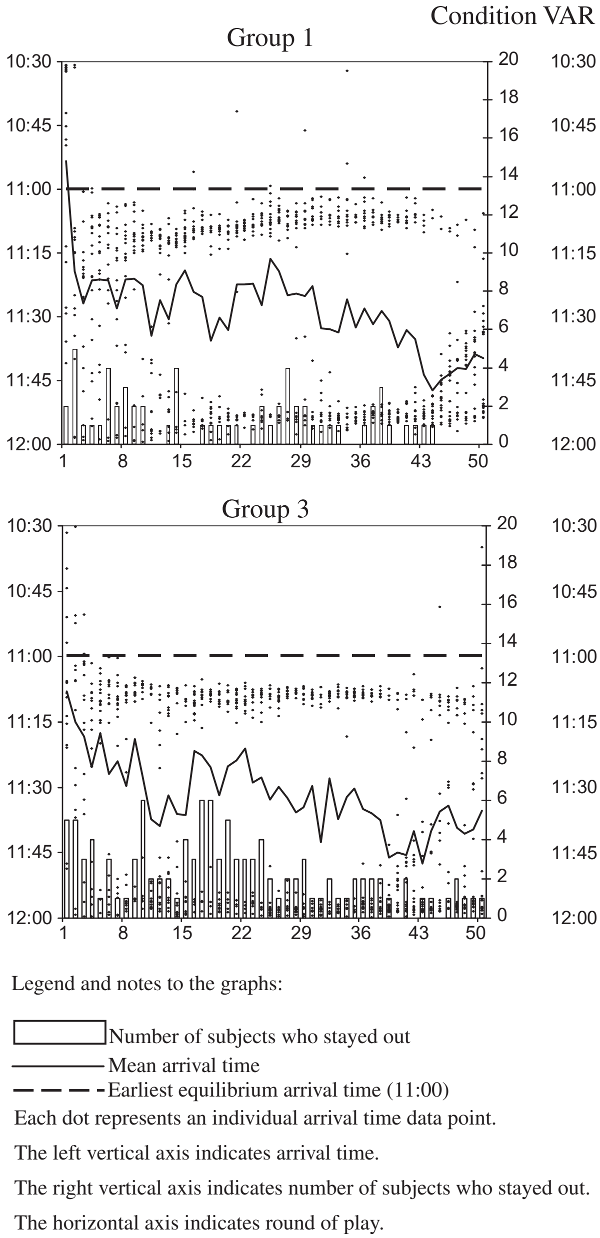


Fig. 3B. Arrival times and decisions to stay out of individual subjects by group Experiment VAR Condition VAR(8, 18).

the opposite trend is exhibited in Conditions VAR(10, 16) and VAR(8, 18) (Figs. 3A and 3B), where fewer subjects stayed out of the queue in the later rather than earlier rounds.

3.2.1. Observed vs. equilibrium behavior: aggregate results

To facilitate comparison between observed and equilibrium behavior, arrival times were divided into intervals of 5 minutes each ([10:55–11:00), [11:00–11:05), ..., [11:55–12:00]). The only exception was the 25-minutes [10:30–10:55) interval that, in addition to the [10:55–11:00) interval, was classified as "errors". Given the above categories, subjects could choose one of 15 strategies (14 arrival time intervals and a single decision to stay out). Denote any of these 15 strategies by j and its relative frequency within a group on round t ($t = 1, \ldots, 50$) by p_{jt}. Denote the probability of choosing strategy j under the mixed-strategy equilibrium solution by p_j^*. For each group and each round separately, we computed the square root of the sum of squared deviations $d_t = \sqrt{\sum_{j=1}^{15}(p_{jt} - p_j^*)^2}$ between the observed and equilibrium proportions. Figs. 4A and 4B exhibit the five-round moving average of d_t by group and condition. With experience in playing the batch queueing game, the deviation index (starting at $d_i \geqslant 0.60$) *decreases* for each of the four groups in Experiment CON (from about 0.62 to about 0.20 to 0.24). It also decreases for Groups 3 and 4 in Condition VAR(10, 16) (from about 0.45 to about 0.23), and with Groups 1 and 2 in that condition too, except for a slight increase towards the end of the game. The index *increases* for each of the four groups in Condition VAR(8, 18) from 0.40 to about 0.48. In general, aggregate behavior approaches equilibrium in Experiment CON (see Fig. 4A), and diverges from it in Condition VAR(8, 18) (see Fig. 4B, lower panel). The aggregate behavior in Condition VAR(10, 16) (see Fig. 4B, upper panel) oscillates over time and extrapolation of the trend beyond 50 rounds is difficult.

Tables 1A and 1B present the observed proportions of arrival time and the proportion of the decisions to stay out of the queue by group and condition in the last 10 rounds. The corresponding equilibrium proportions are displayed in bold font.

Once the subjects in Experiment CON have gained considerable experience with the game, the modal observed and predicted distributions of arrival times mostly fall in the [11:00–11:05) time interval (mean proportion across groups is 0.576 compared to the equilibrium probability of 0.533), with relatively few arrivals falling between 11:10 and 12:00. The two-sided Kolmogorov–Smirnov (KS) test of goodness of fit was used to compare the observed and predicted CDFs (df = 20 for each group). In each case, the null hypothesis of equality of observed and theoretical CDFs could not be rejected at the 0.05 level. We conclude that with experience in playing the constant capacity batch queueing game aggregate behavior in each group converged to equilibrium play.

The equilibrium distribution of arrival times in Experiment VAR is bimodal (although only slightly so in Condition VAR(10, 16)), with two modes at the 11:00–11:05 and 11:40–11:45 time intervals (see Table 1B). In agreement with this prediction, bimodality is observed in each of the eight groups. However, in each case the arrival distribution is shifted in the direction of the termination time $T = 12{:}00$. Even after considerable experience in playing the game, the two modes of the observed distributions of arrival time fall in the (11:05–11:15) and (11:50–12:00) time intervals. Additionally, the relative frequency of staying out decisions was considerably lower than predicted, as the mean observed proportion of staying out under Condition VAR(10, 16) (VAR(8, 18)) is 0.143 (0.039), compared with the equilibrium probability of 0.210 (0.103). The two-sided KS test was invoked again to compare observed and equilibrium CDFs and rejects the null hypothesis for VAR(8, 18) ($D = 0.30$, 0.36, 0.38, and 0.39 for Groups 1, 2, 3, and 4, respectively, $p < 0.05$ in each case). We conclude that the observed arrival times support the theoretical bimodality, but in each group they are shifted to later arrival time intervals (more than 40% of arrivals occurred in the last 10 minutes before departure). The two-sided KS test also rejects the null hypothesis for Groups 1 and 2 in Condition VAR(10, 16) with $D = 0.26$ ($p \approx 0.1$) and $D = 0.31$ ($p < 0.05$) for Groups 1 and 2, respectively. But the test cannot reject the null hypothesis for Groups 3 and 4 in that same condition ($D = 0.16$ and $p > 0.1$ in both cases). In Condition VAR(10, 16), 33% of arrivals occurred in the last 10 minutes before departure.

3.2.2. Collusive behavior?

The equilibrium solution of the batch queueing game is Pareto inefficient (Naor, 1969; Hassin and Haviv, 2003). Given the parameter values in the present experiment, the expected payoff is $ng = 60 \times 20 = 1200$ in every experiment/condition. Group welfare (social optimum) in Experiment CON is maximized when s players arrive simultaneously at 12:00 and $n - s$ players stay out. This asymmetric group behavior results in a considerably higher group payoff of $s(r - d) + (n - s)g = 4560$ (3.8 time higher than in equilibrium). Similarly, social optimum is achieved in Experiment VAR if s_L players arrive at 12:00 and $(n - s_L)$ players stay out. This behavior results in an expected group payoff of

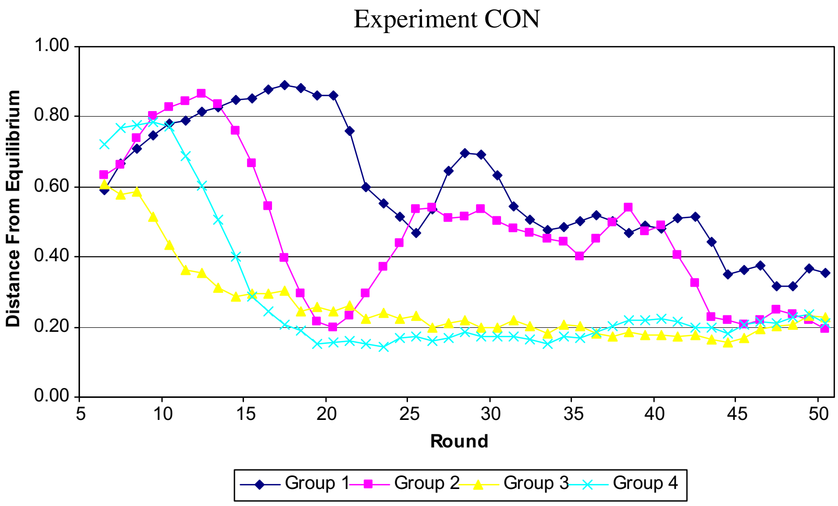
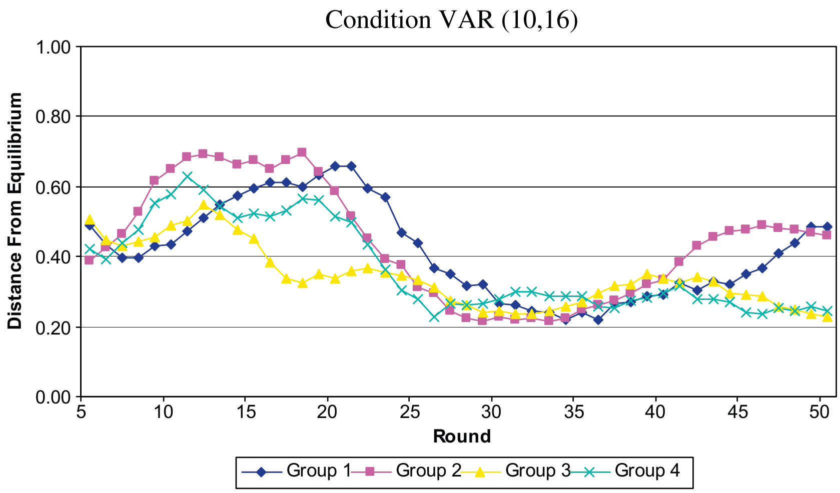


Fig. 4A. Five-round moving average of mean squared deviations between observed and equilibrium arrival and staying out decisions Experiment CON.



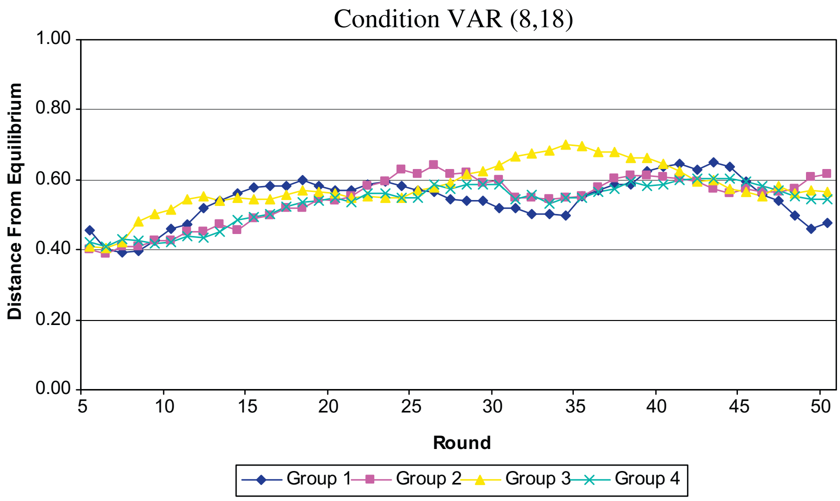


Fig. 4B. Five-round moving average of mean squared deviations between observed and equilibrium arrival and staying out decisions Experiment VAR.

Table 1A
Observed and equilibrium proportion of decisions in the last 10 rounds by group Experiment CON.

Arrival time	Group				Equil.
	1	2	3	4	
10:30–10:55	0.005	0.000	0.010	0.000	**0.000**
10:55–11:00	0.005	0.005	0.000	0.005	**0.000**
11:00–11:05	0.430	0.475	0.695	0.705	**0.533**
11:05–11:10	0.295	0.180	0.020	0.005	**0.043**
11:10–11:15	0.020	0.020	0.010	0.000	**0.029**
11:15–11:20	0.000	0.005	0.000	0.000	**0.023**
11:20–11:25	0.005	0.020	0.000	0.000	**0.020**
11:25–11:30	0.005	0.010	0.010	0.015	**0.017**
11:30–11:35	0.005	0.000	0.000	0.010	**0.016**
11:35–11:40	0.015	0.000	0.010	0.030	**0.016**
11:40–11:45	0.005	0.025	0.005	0.005	**0.015**
11:45–11:50	0.000	0.025	0.010	0.005	**0.015**
11:50–11:55	0.010	0.000	0.005	0.010	**0.015**
11:55–12:00	0.035	0.060	0.035	0.045	**0.016**
Staying out	0.165	0.175	0.190	0.165	**0.240**

Note: Cells with at least 10% of the observations are shaded.

Table 1B
Observed and equilibrium proportion of decisions in the last 10 rounds by group Experiment VAR.

Arrival time	Condition VAR(10, 16)					Condition VAR(8, 18)				
	Group				**Equil.**	Group				**Equil.**
	1	2	3	4		1	2	3	4	
10:30–10:55	0.005	0.015	0.010	0.005	**0.000**	0.000	0.000	0.010	0.005	**0.000**
10:55–11:00	0.000	0.005	0.025	0.000	**0.000**	0.000	0.000	0.000	0.000	**0.000**
11:00–11:05	0.100	0.035	0.425	0.525	**0.368**	0.025	0.000	0.010	0.000	**0.270**
11:05–11:10	0.130	0.225	0.115	0.035	**0.050**	0.115	0.190	0.110	0.010	**0.048**
11:10–11:15	0.105	0.210	0.005	0.005	**0.038**	0.015	0.115	0.115	0.110	**0.038**
11:15–11:20	0.055	0.055	0.000	0.010	**0.035**	0.005	0.060	0.030	0.145	**0.033**
11:20–11:25	0.030	0.030	0.005	0.005	**0.033**	0.005	0.025	0.010	0.065	**0.033**
11:25–11:30	0.005	0.000	0.005	0.005	**0.035**	0.015	0.000	0.030	0.040	**0.035**
11:30–11:35	0.005	0.005	0.010	0.010	**0.038**	0.060	0.025	0.020	0.040	**0.045**
11:35–11:40	0.005	0.030	0.020	0.010	**0.043**	0.080	0.010	0.025	0.015	**0.073**
11:40–11:45	0.025	0.015	0.015	0.125	**0.045**	0.085	0.015	0.030	0.010	**0.208**
11:45–11:50	0.075	0.070	0.010	0.070	**0.043**	0.120	0.000	0.075	0.015	**0.063**
11:50–11:55	0.095	0.125	0.080	0.010	**0.035**	0.395	0.030	0.040	0.145	**0.033**
11:55–12:00	0.245	0.065	0.125	0.000	**0.030**	0.060	0.450	0.445	0.395	**0.023**
Staying out	0.120	0.115	0.150	0.185	**0.210**	0.020	0.080	0.050	0.005	**0.103**

Note: Cells with at least 10% of the observations are shaded.

$$0.5[s_L(r-d)+(n-s_L)g]+0.5[s_S(r-d)-(s_L-s_S)d+(n-s_L)g]$$

which is equal to 4020 in Condition VAR(10, 16) and 3820 in Condition VAR(8, 18).[6]

When n is large and communication is prohibited, as is the case in the present experiment, tacit coordination that calls for asymmetric play is unlikely. Another, considerably more plausible strategy that does not require asymmetric coordination calls for *all* the n players to arrive at 12:00, thereby completely eliminating waiting costs and at the same time maintaining symmetry between players. The resulting group payoff is $s(r-d)-(n-s)d = 3960$ in Experiment CON and $0.5[s_L(r-d)-(n-s_L)d] + 0.5[s_S(r-d)-(n-s_S)d] = 3620$ in both conditions in Experiment VAR. Although only slightly smaller than the social optimum, both values outperform the equilibrium payoff by a factor of 3:1. Such a large margin would be expected to provide a strong incentive for tacit collusion that maintains symmetry across players. Clearly, this did not happen in Experiment CON as can be attested by the relatively low frequencies of "last minute" arrivals and the relatively high frequency of staying out decisions (see Table 1A). A very different picture emerges in Experiment VAR, especially in Condition VAR(8, 18) where departure from constant capacity is most conspicuous. Figs. 3A and 3B show that the subjects arrived later than predicted as they gained more experience with the game. During the final ten rounds in Condition VAR(8, 18), a high proportion of "last minute" arrivals were registered (in the 11:55–12:00 time interval, except in Group 1) and very few cases of staying out decisions (see right panel of Table 1B). Consequently, on average, subjects in that condition earned twice as much as the equilibrium expected payoff. However, their mean payoff was still considerably lower than the expected payoff associated with last minute arrivals by *all* the subjects because even in Condition VAR(8, 18) fewer than 50% of the subjects arrived at the "last minute". Although the tacit coordination calling for last minute arrival is very simple to play and payoff attractive to the collective, the results in Table 1B and Figs. 3A and 3B suggest that for many subjects the temptation to precede the other group members and thereby increase the probability of successful service might have been too strong to resist.

3.2.3. Dynamics

Although the pattern of aggregate behavior in the last 10 rounds of Experiment CON is consistent with mixed-strategy equilibrium play, we find no evidence supporting randomization on the *individual* level. The same holds for the individual behavior in the two conditions of Experiment VAR. Randomization of strategies implies independent decisions on each round; in particular, the decisions and outcome of round t should not affect the subject's decision on round $t+1$. As we show below, sequential dependencies are present in all the conditions. If subjects do not randomize their decisions, then what drives their strategy selection?

Individual decisions could be accounted for by relatively simple heuristics that are contingent on the individual's personal history of play and group results on the previous round. Using different statistical procedures, in what follows we investigate if behavioral trends analogous to those identified by Stein et al. (2007) are present in our data. Separately for each condition, we list the effect of the variables presented in Table 2 on the decision to either join the queue or stay out and, in the former case, on the arrival time.

[6] Given the parameter values of the game, as long as the probability of arrival of the large ferry exceeds 5/17, this is the social optimum in both VAR conditions. If this probability is smaller than 5/17, then the social optimum is for s_S players to join the queue at 12:00 and the others to stay out.

Table 2
Definitions of variables in the dynamics model.

Experiment	Variable	Definition
Individual level variables		
Both	$D_{s,it-1}$	=1, if subject *i* joined the queue on round $t-1$ and boarded the ferry =0, otherwise
Both	$D_{f,it-1}$	=1, if subject *i* joined the queue on round $t-1$ and failed boarding the ferry =0, otherwise
Both	w_{it-1}	Subject *i*'s waiting time on round $t-1$ (conditional on her joining the queue)
Aggregate level variables		
Both	$N_{NQ,t-1}$	Number of subjects who stayed out on round $t-1$
CON	MW_{t-1}	Median waiting time[a] for subjects boarding the ferry on round $t-1$
VAR	MW^s_{t-1}	Median waiting time of the subjects who would have boarded the ferry, had the small ferry arrived on round $t-1$ (regardless of the realized capacity of the ferry that arrived on $t-1$)
VAR	MW^l_{t-1}	As above, had the large ferry arrived on round $t-1$
CON	mW_{t-1}	Minimum waiting time of the subjects who boarded the ferry on round $t-1$
VAR	mW^s_{t-1}	Minimum waiting time of the subjects who would have boarded the ferry, had the small ferry arrived on round $t-1$
VAR	mW^l_{t-1}	As above, had the large ferry arrived on round $t-1$
VAR	$Size_{t-1}$	=1, if the realized capacity of the ferry that arrived on round $t-1$ is 18 =0, otherwise

[a] We take this measure as a proxy for how long one had to wait to board the ferry on round $t-1$. Because the experiment has a full information design, this information is always available to every subject at the end of every round. Using mean successful waiting time yields similar estimation results.

Table 3A
Estimated coefficient (standard error) for the dynamics models: Experiment CON.

Variable	TQONTQ	Arrival Timing
Intercept	4.90** (1.19)	9.31** (4.10)
$D_{s,it-1}$	3.08** (0.13)	–
$D_{f,it-1}$	1.86** (0.17)	−0.76 (0.87)
$N_{NQ,t-1}$	0.056 (0.046)	−1.07** (0.20)
MW_{t-1}	−0.089** (0.023)	0.55** (0.084)
mW_{t-1}	−0.013** (0.0036)	−0.016 (0.016)
w_{it-1}	–	0.32** (0.024)

3.2.4. The TQONTQ (to queue or not to queue) model

We use the logit regression to test for the presence of sequential dependencies. We propose that subject *i* in Experiment CON joined the queue on round *t* if

$$y_{it} = \alpha_0 + \alpha_1 D_{s,it-1} + \alpha_2 D_{f,it-1} + \alpha_3 N_{NQ,t-1} + \alpha_4 MW_{t-1} + \alpha_5 mW_{t-1} + \varepsilon_{it} > 0,$$

and stayed out, otherwise. Similarly, subject *i* in Experiment VAR joined the queue on round *t* if

$$y_{it} = \alpha_0 + \alpha_1 D_{s,it-1} + \alpha_2 D_{f,it-1} + \alpha_3 N_{NQ,t-1} + \alpha_{4s} MW^s_{t-1} + \alpha_{4l} MW^l_{t-1} + \alpha_{5s} mW^s_{t-1} + \alpha_{5l} mW^l_{t-1} + \alpha_6 Size_{t-1} + \varepsilon_{it} > 0$$

and stayed out, otherwise. In both models, the α coefficients are to be estimated, and the iid random shock ε_{it} is assumed to have a Type I extreme value distribution. We thus have:

$$P(\text{Subject } i \text{ stayed out on round } t) = \frac{1}{1+\exp(u_{it})}, \quad \text{and}$$

$$P(\text{Subject } i \text{ joined the queue on round } t) = \frac{\exp(u_{it})}{1+\exp(u_{it})},$$

where $u_{it} = y_{it} - \varepsilon_{it}$. Tables 3A and 3B (the columns under TQONTQ) present the estimated coefficients for the pooled data using maximum likelihood estimation for the single condition of Experiment CON and the two conditions of Experiment VAR. The analyses reported on the TQONTQ and Arrival Timing decisions (to be reported in the next subsection) are based on data from rounds 11–50.[7]

[7] Examination of the raw data shows that individual decisions on the first 5–10 rounds were erratic. We attribute this erratic behavior to the considerable experience a player might have required to become familiar with the new real-time mechanism. With the timer moving rather quickly, subjects seemed to require several trials to learn adjusting their entry times by moving the mouse out of the designated area on the PC screen.

Table 3B
Estimated coefficient (standard error) for the dynamics models: Experiment VAR.

Variable	VAR(10, 16)		VAR(8, 18)	
	TQONTQ	Arrival Timing	TQONTQ	Arrival Timing
Intercept	6.80**	−2.26	1.99**	4.12*
	(1.41)	(6.53)	(0.75)	(2.41)
$D_{s,it-1}$	3.09**	–	3.20**	–
	(0.13)		(0.17)	
$D_{f,it-1}$	2.68**	−2.57**	2.46**	0.65
	(0.18)	(0.94)	(0.22)	(1.15)
$N_{NQ,t-1}$	0.20**	−0.42*	−0.046	0.90**
	(0.04)	(0.22)	(0.055)	(0.27)
MW^s_{t-1}	−0.16**	0.38**	−0.035**	0.046
	(0.03)	(0.14)	(0.016)	(0.052)
MW^l_{t-1}	0.022*	−0.073	−0.0069	−0.064**
	(0.013)	(0.070)	(0.0069)	(0.031)
mW^s_{t-1}	−0.0061	0.022	0.0024	0.050
	(0.0069)	(0.039)	(0.0074)	(0.034)
mW^l_{T-1}	−0.0020	0.12**	0.17**	0.44*
	(0.0029)	(0.02)	(0.054)	(0.23)
$Size_{t-1}$	−0.20*	1.98**	−0.60**	2.15**
	(0.12)	(0.68)	(0.17)	(0.92)
w_{it-1}	–	0.61**	–	0.60**
		(0.02)		(0.02)

Note: All estimates with ** are significantly different from zero at p < o.05.
All estimates with * are significantly different from zero at $p \approx 0.1$.
All other estimates have $p > 0.1$.

Now, define the following events for subject i:

- NQ_t – stayed out on round t.
- Q_t – joined the queue on round t.
- SQ_{t-1} – successfully boarded the ferry on round $t-1$.
- FQ_{t-1} – failed to board the ferry on round $t-1$.

Based on the estimated TQONTQ parameters in Tables 3A and 3B that are at least significantly different from zero at $p \approx .1$, we deduce the following statements for our data, which are true for all three cases (Experiment CON and the two conditions in Experiment VAR) unless otherwise stated:

(1) $P(Q_t|Q_{t-1}) > P(Q_t|NQ_{t-1})$, or equivalently, $P(NQ_t|NQ_{t-1}) > \text{Prob}(NQ_t|Q_{t-1})$.
(2) $P(Q_t|SQ_{t-1}) > P(Q_t|FQ_{t-1}) > 0.5$.
(3) $P(NQ_t)$ increases with MW_{t-1} in Experiment CON and with MW^s_{t-1} in Experiment VAR, Condition VAR(10, 16) (VAR(8, 18)); it also has a marginal dependence on MW^l_{t-1} in the opposite direction in Condition VAR(10, 16).
(4) $P(NQ_t)$ increases with mW_{t-1} in Experiment CON and with mW^l_{t-1} in Experiment VAR, Condition VAR(8, 18).
(5) $P(Q_t|\text{small ferry arrived at } t-1) > P(Q_t\ |\text{large ferry arrived at } t-1)$ in Experiment VAR.
(6) $N_{NQ,t-1}$ has no effect on the decision to join the queue or stay out in Experiment CON and Experiment VAR, Condition VAR(8, 18). $P(NQ_t)$ decreases with $N_{NQ,t-1}$ in Condition VAR(10, 16).

To illustrate the derivation of these statements, consider statement (2) as applied to Experiment CON. First, note that $P(\text{Subject } i \text{ stayed out on round } t) = 1/[1 + \exp(u_{it})] < 0.5$ if and only if $u_{it} > 0$. Now, if the subject joined the queue on round $t-1$, then either $\{D_{s,it-1} = 1, D_{f,it-1} = 0\}$ or $\{D_{s,it-1} = 0, D_{f,it-1} = 1\}$. We can use the estimated values of the coefficients plus the observation that $60 > MW_{t-1}$ ($\geqslant mW_{t-1}$ by definition) to deduce that:

$$u_{it} = 4.90 + 3.08D_{s,it-1} + 1.86D_{f,it-1} - 0.089MW_{t-1} - 0.013mW_{t-1} > 0.$$

Hence, $P(Q_t|Q_{t-1}) > 0.5$.

We turn next to explain these six statements. Statement (1) describes a tendency to repeat the same decision on consecutive rounds. This is further elaborated by (2), which states that a subject is inclined to join the queue again if she did so in the previous round regardless of whether the previous arrival was successful (both probabilities are higher than 0.5), but that she is more likely to join the queue after success than after failure. The observed propensity to repeat joining the queue following successful service (82.7%, 76.8%, and 70.8% of arrival were successful in Experiment CON, Condition VAR(10, 16), and Condition VAR(8, 18), respectively) sheds light on the lower than expected level of staying out decisions in both experiments.

Statement (3) indicates that, overall, the longer the expected waiting time for successful boarding, the more inclined the subjects are to stay out. Statement (4) indicates that the longer the last person who successfully boarded the ferry had to wait (i.e., the shortest successful wait), the more inclined subjects are to stay out, and that this minimum is considered under the "best case scenario" in Condition VAR(8, 18).

Statement (5), which only applies to Experiment VAR, can be attributed to the gambler's fallacy on the decision to join the queue. Subjects might have expected that, given the size of the ferry that arrived on round $t-1$, the other type of ferry is more likely to arrive on the next round. Consequently, after a small ferry arrived on round $t-1$, subjects were more likely to join the queue on round t, compared to when the large ferry arrived on round $t-1$, and vice-versa.

Statement (6) indicates that the number of subjects who stayed out on the previous round, $N_{NQ,t-1}$, had no effect on the decision to join the queue except for Condition VAR(10, 16). However, as we show below, it had a significant effect on the arrival time decisions.

3.2.5. Arrival timing model

Once the subject decides to join the queue on round t, she has to decide *when* to arrive, or equivalently, how long to wait. This decision is presumably based on information about the arrival times of the other group members that she has gathered up to round t, including her own history of play. One way to classify the history of play is according to when the subject last joined the queue: this might occur on rounds $t-1$, $t-2$, etc. We report below the analysis of the case in which the subject last joined the queue on round $t-1$. This accounts for more than 90% of all the arrival decisions. Let w_{it} denote subject i's waiting time on round t, and consider the following linear regression model applied to Experiment CON:

$$w_{it} = \beta_0 + \beta_1 D_{f,it-1} + \beta_2 N_{NQ,t-1} + \beta_3 MW_{t-1} + \beta_4 mW_{t-1} + \beta_5 w_{it-1} + \varepsilon_{it}.$$

Similarly, consider the following linear regression model applied to each of the two conditions of Experiment VAR:

$$w_{it} = \beta_0 + \beta_1 D_{f,it-1} + \beta_2 N_{NQ,t-1} + \beta_{3s} MW^s_{t-1} + \beta_{3l} MW^l_{t-1} + \beta_{4s} mW^s_{t-1} + \beta_{4l} mW^l_{t-1} + \beta_5 w_{it-1} + \beta_6 Size_{t-1} + \varepsilon_{it}.$$

In both models, the β coefficients are to be estimated from the data, and the iid random shock ε_{it} is assumed to have a zero-mean normal distribution. Because we are only looking at cases when subjects joined the queue on round $t-1$, $D_{s,it-1} = 1 - D_{f,it-1}$. Therefore, we do not need a dummy variable for successful arrival.

Tables 3A and 3B (the columns under Arrival Timing) present the estimated coefficients for the pooled data using maximum likelihood estimation for Experiments CON and VAR. Note that failure to board in the previous round bears no explicit impact on the decision when to arrive. This is due to the fact that $D_{f,it}$ is significantly correlated with w_{it}, the median successful waiting times (MW_t for Experiment CON and MW^l_t for Experiment VAR), and the minimum successful waiting times (mW_t for Experiment CON and mW^l_t for Experiment VAR) at $p < .0001$; hence, we conclude that these terms mediate the effect of failure.[8,9]

Denote as E the event that subject i joined the queue earlier on round t than on round $t-1$, i.e. $w_{it} - w_{it-1} > 0$, conditional on her joining the queue in both rounds. The following statements are consistent with those results reported in Tables 3A and 3B that are at least significantly different from zero at $p \approx .1$[10]:

(1) $P(E)$ decreases with w_{it-1}.
(2) In Experiment CON, $P(E)$ decreases with $w_{it-1} - MW_{t-1}$ over and above its dependence on w_{it-1};
(3) In Experiment VAR, $P(E)$ decreases with $w_{it-1} mW^l_{t-1}$ over and above its dependence on w_{it-1}.
(4) $P(E)$ decreases with $N_{NQ,t-1}$ except in Condition VAR(8, 18) where the dependence is opposite.
(5) In Experiment VAR, $P(E|\text{large ferry arrived on round } t-1) > P(E\ |\text{small ferry arrived on round } t-1)$.

Statement (1) means that if a subject joined the queue relatively early, she would tend to arrive later on the next round. Statements (2) and (3) are quite intuitive. The longer a subject waits compared to the median waiting time, the less likely she is to arrive earlier on the next round, or alternatively, the more likely she is to arrive later on the next round. The other dependences on previous successful arrival time in the two VAR conditions are largely consistent with this intuition (the opposite dependence on MW^l_{t-1} in Condition VAR(8, 18) being relative small in magnitude). Statement (3) as applied to Condition VAR(8, 18) also indicates that, similar to the decision to join the queue, subjects in that condition are mostly sensitive to the (actual or hypothetical) outcome information about the arrival of the large ferry.

When examined together with statement (6) from the TQONTQ analysis, statement (4) indicates that, in general, subjects might have estimated the expected "demand" for space on the ferry at the current round based on the previous round demand. In Experiment CON and Condition VAR(10, 16), the higher the previous round demand, the higher the expected demand is on the coming round for the subject, and thus the more likely the subject is to arrive earlier rather than later compared to the previous round queuing time. In Condition VAR(8, 18), the dependence is opposite, perhaps because subjects employ higher-order strategic thinking in the relatively more uncertain environment, so that, if demand on the previous

[8] These correlations are negative with $|r| > .1$ for all variables except the minimum successful waiting times, when the correlations are positive with $r \approx 0.1$.

[9] A regression of w_{it} on $D_{f,it-1}$ alone yields significantly negative coefficients for the latter in both conditions at $p < 0.0001$, further supporting the claim of mediation. See also the results reported at the end of this subsection.

[10] These trends are reversed when we consider the probability of the subject joining the queue later on round t than on round $t-1$, conditional on her joining the queue in both rounds.

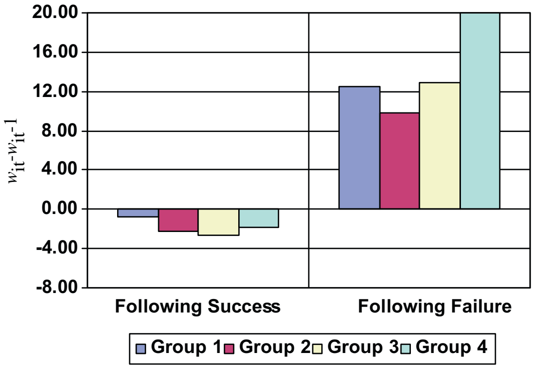


Fig. 5A. Mean change in arrival time following success or failure in boarding the ferry on the previous round Experiment CON.

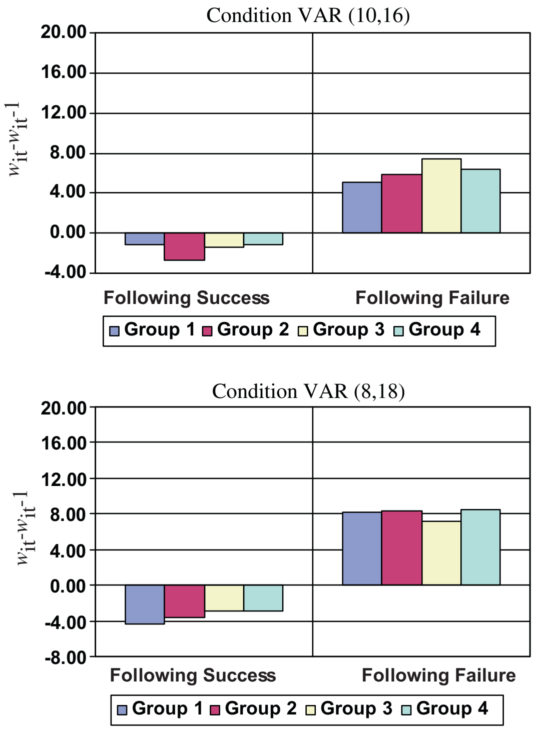


Fig. 5B. Mean change in arrival time following success or failure in boarding the ferry on the previous round Experiment VAR Condition VAR(10, 16).

round was high, they reasoned that not many subjects (who they assume use previous round demand as a direct proxy for expected demand) would join the queue on the coming round, and thus the expected demand should be low.

Finally, statement (5), which applies to Experiment VAR only, resembles statement (5) of the decision to join the queue as both reflect gambler's fallacy reasoning. Expecting that it is more likely that the other type of ferry will arrive on the present round, and believing that an early arrival is needed to board the smaller ferry, the subject is more likely to arrive earlier after the large ferry arrived compared to after the small ferry.

To complete the analysis of the dynamics of play, it is instructive to look at how success or failure in boarding the ferry on round $t-1$ affected arrival time on round t not only directionally (as was reported previously) but also quantitatively. In particular, joining the queue and failing to board the ferry resulted in an actual monetary loss. This loss might have increased the likelihood of not joining the queue in the next period or joining it earlier. On the other hand, boarding the ferry resulted in a monetary gain (if joining it after 11:00). Clearly, it was possible that other subjects joining the queue later would board the ferry and gain even more. Hence, arriving later would only have resulted in a potential gain. Because subjects are in general loss averse (e.g., Kahneman et al., 1991), we expect actual loss to affect future behavior more than foregone gain. Therefore, we calculate for each subject, his/her mean *change in arrival time* following a success or failure in boarding the ferry in the previous round of play; these means are then averaged over subjects in the same group and the results are displayed in

Figs. 5A and 5B. The vertical axis measures the *change* in arrival time in minutes between two successive rounds. In both conditions, the difference between the effect of success and failure in boarding the ferry on the previous round is quite dramatic. Subjects in all three experimental conditions joined the queue *considerably earlier* after failing to board the ferry on the previous round, but only *slightly later* after boarding it successfully. The effect is stronger in Experiment CON than Experiment VAR.

4. Summary and discussion

Our experimental evidence shows that, with experience gained in playing the stage queueing game repeatedly, subjects' aggregate behavior approaches equilibrium in Experiment CON but diverges from it in Experiment VAR. Non-equilibrium collusive behavior that is Pareto superior to the equilibrium outcomes is not observed. A closer examination of the dynamics of play reveals strong dependence of subject behavior on her history of play as well as on the aggregate outcome in the previous round. In general, subjects tend to repeat what they did in the previous round; this can be most clearly seen in our analysis of the decision to join the queue. Reinforcement learning is also evident, as in the finding that a subject is more likely to queue again if she successfully boarded the ferry in the previous round, compared with the case where she joined the queue but failed to board the ferry in the previous round. Moreover, the decision whether to join the queue is affected by the waiting cost required to board successfully in the previous round (as reflected in the negative relationship between the probability to queue and the median and minimum successful waiting times in the previous round). Whether the major patterns of the results would have been observed in longer sessions with more than 50 rounds of play is an open experimental question.

The typical subject determines her arrival time in the current round (conditional on joining the queue) according to four major reference criteria: (1) her arrival time in the previous round (conditional on joining the queue in the previous round); (2) some representative measure of the arrival time for successful boarding in the previous round; (3) the number of subjects who joined the queue in the previous round; and (4) in Experiment VAR, whether a large ferry or a small ferry arrived in the previous round. The last dependence indicates gambler's fallacy in a similar way as with the decision to queue; that is, subjects tend to think that a small ferry in round $t-1$ is likely to be followed by a large ferry in round t, and vice versa. Next, we find (Figs. 5A and 5B) that whether a subject succeeded or failed in boarding the ferry in the previous round had a strong asymmetric impact on her arrival time in the current round.

When the server capacity is fixed and commonly known (Experiment CON), the aggregate behavioral patterns converge to the mixed-strategy equilibrium. We have confidence in this finding as the same patterns have been reported by Stein et al. An intriguing observation is that subjects in Experiment VAR arrived significantly later than predicted. This is reflected in the dependence of subject behavior on the median/latest successful arrival time in the previous round had the ferry that arrived been a large one, as our estimation shows (statements (3) and (4) from the TQONTQ model estimation; statements (2) and (3) from the arrival timing model estimation); similar dependence related to the hypothetical arrival of a small ferry is non-significant in Condition VAR(8, 18). Apparently, when there is a high variance in ferry capacity, subjects assume by default that the next ferry is more likely to be a large one, and may only adjust this baseline assumption afterwards due to gambler's fallacy.

These latter findings in Conditions VAR(10, 16) and VAR(8, 18) are consistent with multiple studies in psychology that report an unrealistic optimism bias. In a review of the literature, Armor and Taylor (2002) reported that "By a number of metrics and across a variety of domains, people have been found to assign higher probabilities to their attainment of desirable outcomes than either objective criteria or logical analysis warranted" (p. 334). Table 1B shows a higher tendency to enter the queue and more late arrivals than predicted by equilibrium in both VAR conditions. But why is the discrepancy from equilibrium play small and not significant in Experiment CON and considerably higher and significant in Experiment VAR? The findings reviewed by Armor and Taylor are consistent with the answer that people are more optimistically biased under conditions of greater uncertainty. Citing earlier studies by Irwin (1953), and more recent studies by Buckley and Sniezek (1992), and by Dunning et al. (1989), Armor and Taylor noted that the tendency to expect success (the large ferry in Experiment VAR) unrealistically on a binary task entirely determined by chance is maximized when the odds for the desirable and undesirable events are even. This is exactly the case in both conditions of Experiment VAR.

Our interpretation can help explain the aggregate deviations from equilibrium in Experiment VAR, as evidenced in Table 1B. If subjects mentally exaggerate the probability that a large ferry arrives – or make their decisions in ways that are effectively equivalent to this bias – then the two modes of arrival timing in the equilibrium mixed strategies will increase, while the probability of staying out will decrease. In other words, the "best case scenario" approach keeps subjects' queueing rate generally high and waiting costs generally low with respect to the equilibrium predictions. As our analysis reveals, this is most conspicuously reflected when the difference between large and small ferries is considerable, as in Condition VAR(8, 18), while there is also some evidence of such in Condition VAR(10, 16). More investigation is required to understand the mechanisms behind the suggested bias.

Acknowledgements

We gratefully acknowledge financial support by NSF Grant No. SES-0135811 to Darryl A. Seale and Amnon Rapoport, and AFOSR/MURI contract F49620-03-1-0377 to the University of Arizona. We thank Maya Rosenblatt and Ryan O. Murphy for their help in data collection, and the Associate Editor and two reviewers for very constructive comments.

Appendix A. Derivation of the mixed-strategy equilibria

We construct equilibrium solutions to non-cooperative, complete information, n-person games of timing with endogenously determined arrivals. Although we frame the game in a specific context, it applies to endogenous arrivals in more general batch queues. A ferry is known to depart at time T, and n symmetric commuters who are interested in boarding it may independently choose arrival times at or before T. Upon arrival they join a FCFS queue in an attempt to occupy one of the s seats on the ferry. At time T, the first s commuters in line are allowed to board the ferry. Because waiting in line is costly, commuters may also refuse to join the queue. As with the experiment, we assume no balking, namely, that commuters joining the queue must stay there until time T. This corresponds to an unobservable queue (Hassin and Haviv, 2003).

The following parameters characterize the model: n players, s = capacity of the ferry ($0 < s < n$), c = cost per minute of waiting in the queue, d = entry fee, g = a payoff for staying out of the game, r = reward for commuters who board the ferry. All parameter values are assumed to be commonly known.

A.1. The modified game

We first define a *modified game* with n players, capacity s, and a cost parameter $0 < w < 1$. Each player chooses an arrival time $t \in [0, 1]$ or may opt to stay out. A player arriving at time t waits for time $1 - t$ at a cost of $w(1 - t)$. Now distinguish one of the n players, and call her the *critical player*; we will compute the expected payoff for this player for each t and then use the results to solve for the symmetric mixed-strategy equilibrium. Let F be the CDF of this mixed-strategy of arrival times and f its density function, if it exists. $F(t)$ is the probability that a player will arrive at or before time t. Then, $p_{out} = 1 - F(1)$ is the probability that a player chooses to stay out of the queue. Let $P(t)$ denote the probability of receiving the reward of 1 unit, if the critical player arrives at time t.

Assumptions of the modified game:

(a) The critical player earns a reward of 1 unit, if at most $s - 1$ of the other $n - 1$ players chose to arrive before t. (We will show that there is probability zero of a tie at t.)
(b) Waiting time cost is $w(1 - t)$ since the critical player waits a time $1 - t$ at w per unit time.
(c) Staying out of the queue results in a certain payoff $1 - w$.
(d) There is no cost for playing the game.

If the critical player chooses to arrive at t, then in equilibrium her expected payoff from joining the queue equals her payoff from staying out:

$$P(t) - w(1 - t) = 1 - w.$$

This reduces to $P(t) = 1 - wt$. This equality must hold in equilibrium for any time for which arrivals are possible. On the other hand, at any time the expected payoff from playing the game is at most the payoff from staying out. This leads to the inequality $P(t) \leqslant 1 - wt$.

Theorem 1. *The CDF F has no point mass and the density f exists for* $t \in [0, 1]$.

Proof. Note that the earliest possible arrival time is 0 since any arrival before this time would obtain less than $1 - w$ in expectation. There is some probability $F(0) \geqslant 0$ that each one of the other $n - 1$ players enters at 0. We will show that $F(0) = 0$ by assuming $F(0) > 0$ and then show that leads to a contradiction. Since there is a positive probability $F(0) > 0$ that each player enters at 0, there is a positive probability that all $n - 1$ do so. In this case, a lottery would be held among those arriving. In any case, if the critical enters at 0, it is not certain that she will find a space on the ferry. That is, $P(0) < 1$. Then, her expected payoff for an arrival at $t = 0$ is $P(0) - w$. So we have shown that this quantity is smaller than $1 - w$. But that violates the condition $P(t) = 1 - w$ for all t for which arrivals are possible. Therefore, we conclude that $F(0) = 0$ so that in equilibrium no arrivals at (or before 0) are possible.Previously, we showed that $P(t) = 1 - wt$. This shows that $P(t)$ is continuous at all t, $0 < t < 1$. If the arrival distribution F had a point mass at some point t, $0 < t < 1$, then P would have a discontinuity at t since this would decrease $P(t)$ at t.

Could there be a point mass at $t = 1$? From $P(t) = 1 - wt$ we see at "just before 1" $P(1-)$ must be $1 - w$ so if I arrived at $t = 1$ while no one else was permitted to do so, then there would be no chance of a lottery and $P(1) = P(1-) = 1 - w$. However, if there was a point mass at $t = 1$ then there would be a positive probability of holding a lottery and so $P(1)$ would strictly decrease. The expected payoff of a critical player if she arrives at $t = 1$ is then less than $1 - w$. This violates the equilibrium condition and shows no point mass at $t = 1$ is consistent with equilibrium. □

In equilibrium, the other $n - 1$ players will choose their arrival times independently from the same distribution F. Theorem 1 implies that we do not have to consider the possibility of ties. Therefore, we have

$$P(t) = \sum_{k=0}^{s-1} \binom{n-1}{k} F(t)^k [1 - F(t)]^{n-1-k}, \quad 0 \leqslant t \leqslant 1. \tag{1}$$

A.2. Converting to the modified game

In this section, we return to the original game and show how it can be converted to the modified game; once we do so, we only need to solve the modified game. Let $\widehat{F},\ \hat{f}$ and, $\widehat{P}$ denote the quantities in the original game that correspond to F, f, and P in the modified game. If the critical player arrives at time τ in the original game and at the equivalent time t in the modified game, then

$$\widehat{F}(\tau) = F(t) \quad \text{and} \quad \widehat{P}(\tau) = P(t).$$

In the original game, $\widehat{P}(\tau)r - 60c(T-\tau) - d$ is the expected payoff of the critical player and this must be a constant over all τ at the equilibrium. Assume for the moment that in equilibrium there is a positive probability of staying out. Players staying out receive payoff g, so this means that all arrivals must receive g in expected value. Therefore, the condition:

$$\widehat{P}(\tau)r - 60c(T-\tau) - d = g, \tag{2}$$

must hold whenever $\hat{f}(\tau) > 0$. Equivalently, the arrival time may be expressed as $x = T - \tau$ hours before T. Using the equilibrium, there will be an earliest feasible arrival time τ_m. As pointed out in the proof of Theorem 1, no player will arrive exactly at τ_m either. Therefore, $\widehat{F}(\tau_m) = 0$ and so $\widehat{P}(\tau_m) = 1$, which implies that $r - 60c(T-\tau_m) - d = g$. Solving this for $x_m \equiv T - \tau_m$, we obtain:

$$x_m = \frac{r-(d+g)}{60c}. \tag{3}$$

For a non-trivial problem, we must have $x_m > 0$, which requires $r > d + g$. We then write (2) as:

$$\widehat{P}(\tau)r - \frac{60cx}{x_m}\left[\frac{r-(d+g)}{60c}\right] = d+g, \quad \text{i.e. } \widehat{P}(\tau) - \frac{x}{x_m}\left[1-\frac{d+g}{r}\right] = \frac{d+g}{r}.$$

Now define $w = 1 - (d+g)/r$ and $t = 1 - x/x_m$ so that $0 < w \leqslant 1$ and $0 < t < 1$. Then,

$$P(t) - (1-t)w = 1 - w,$$

and from this we get the equilibrium condition

$$P(t) = 1 - wt, \quad \text{for all } 0 < t < 1.$$

This is the same condition as the one obtained for the modified game on [0, 1]. Thus, we have reduced the original game to the [0, 1] game.

Note that $x_m = \frac{r-(d+g)}{60c} = \frac{r}{60c}\left[1-\frac{d+g}{r}\right] = \frac{rw}{60c}$, so that the solution depends on the parameters d, g, r, c only through the ratios $(d+g)/r$ and c/r. We could choose to express everything in terms of the cost by dividing d, g, r by c and then re-defining $c = 1$. Or we may always set $r = 1$, if we wish, and then scale d, g, c by r. Instead, we could choose x_m to be 1 unit. This is what we did in the modified game so that it is always a game on [0, 1].

In summary, starting with the original game, the equilibrium solution can be found by first solving a modified game with $w = 1 - (d+g)/r$ to obtain $F(t)$. Then, to find the probability of arriving at least x hours early, evaluate the solution to the modified game at time $t = 1 - x/x_m$. Or arriving at time $\tau \leqslant T$, we must evaluate the solution to the modified game at $t = 1 - (T-\tau)/x_m$.

A.3. Solution of the modified game

We now determine the solution to the modified game. We use the following well-known fact relating the binomial and cumulative beta distributions:

$$\sum_{k=s}^{n-1}\binom{n-1}{k}p^k(1-p)^{n-1-k} = \int_0^p \frac{(n-1)!}{(s-1)!(n-1-s)!}u^{s-1}(1-u)^{n-1-s}\,du. \tag{4}$$

Let $G_{s,n-s}(p)$ denote the *Beta(s, n − s)* CDF given on the right side of (4).

Theorem 2. *The equilibrium solution to the modified game on* [0, 1] *is*:

$$P(t) = 1 - wt, \quad \text{for all } 0 < t < 1 \tag{5}$$

or equivalently,

$$t = [1 - P(t)]/w. \tag{6}$$

and the explicit solution for F is:

$$F(t) = G^{-1}_{s,n-s}(wt). \tag{7}$$

Proof. Eq. (5) was derived previously. Numerically, the solution can be obtained by starting with any $F(t)$ value and then evaluating $P(t)$ from (1). Then solve (5) for t. More easily, we can use (6).

From (1)

$$P(t) = 1 - \sum_{k=s}^{n-1} \binom{n-1}{k} F(t)^k [1 - F(t)]^{n-1-k}.$$

Then, Eq. (4) becomes:

$$1 - P(t) = \int_{F(0)}^{F(t)} \frac{(n-1)!}{(s-1)!(n-1-s)!} u^{s-1}(1-u)^{n-1-s}\,du, \tag{8}$$

but Theorem 1 implies $F(0) = 0$. Thus, $1 - P(t) = G_{s,n-s}(F(t))$ and from (5) and (8):

$$F(t) = G^{-1}_{s,n-s}(wt). \tag{9}$$

completing the proof. □

As the inverse beta CDF is available in Excel and MATLAB, F can be computed easily. Note that the probability of staying out has a simple expression:

$$P(\text{Stay Out}) = 1 - F(1) = 1 - G^{-1}_{s,n-s}(w). \tag{10}$$

Moreover, this is always positive, verifying the assumption used to derive Eq. (2).

A.4. Search for other equilibria

We show in this section that there exist no other equilibria. In the previous derivation of the equilibrium we assumed that some positive probability was assigned to staying out. This allowed us to obtain Eq. (2). Is it possible that the parameters of the original problem are such that it is favorable to arrive with probability 1? The proof of Theorem 1 shows that even in this case there is zero probability of arriving exactly at the departure time. Returning to the original parameters of the game, the expected payoff at τ is

$$\widehat{P}(\tau)r - 60c(T - \tau) - d.$$

This must be a constant for all arrival times: $\widehat{P}(\tau)r - 60c(T - \tau) - d = \gamma$, for some γ with $\gamma > g$. This becomes $\widehat{P}(\tau) - \frac{x}{x_m}\left[1 - \frac{d+\gamma}{r}\right] = \frac{d+\gamma}{r}$ with $x_m = \frac{r-(d+\gamma)}{60c}$. With w defined as $w = 1 - \frac{d+\gamma}{r}$ and $t = 1 - x/x_m$, we get the same form of the modified game as we had previously, namely, $P(t) = 1 - wt$, $0 < t < 1$. However, we require that players arrive before $t = 1$ for certain. From (10) we see that this requires that $G^{-1}_{s,n-s}(w) = 1$. But $w \leqslant 1$ and so $w = G_{s,n-s}(1) = 1$ and this implies $d + \gamma = 0$. So we get only a trivial case.

We have shown that there are no interesting equilibria which exclude the possibility of staying out of the process. No matter how favorable we make the game by increasing r or decreasing c and d, the gains are dissipated by players selecting earlier arrival times. The increased direct payoff is offset by an increase in the total expense of waiting.

A.5. Unknown capacity

Finally, we extend the model by assuming that the capacity S is a random variable with a commonly known probability distribution. Let $b_s = P(S = s)$, $s = 1, \ldots, n$. Now, Eq. (1) is the *conditional* probability $\widehat{P}_s(\tau)$ of boarding the ferry. The unconditional probability is:

$$\widehat{P}(\tau) = \sum_{s=1}^{n} b_s \widehat{P}_s(\tau).$$

Using (8), this can be written as:

$$\widehat{P}(\tau) = \sum_{s=1}^{n-1} b_s[1 - G_{s,n-s}(\widehat{F}(\tau))] + b_n = 1 - \sum_{s=1}^{n-1} b_s G_{s,n-s}(\widehat{F}(\tau)).$$

From (5), we can simplify the above equation to:

$$wt = \sum_{s=1}^{n-1} b_s G_{s,n-s}(\widehat{F}(\tau))$$

Let $a = \widehat{F}(\tau) = F(t)$. Then, we have:

$$F^{-1}(a) = \frac{1}{w}\sum_{s=1}^{n-1} b_s G_{s,n-s}(a) \quad (11)$$

for $0 \leqslant a \leqslant \min(1, F(1))$. The right side of (11) can be evaluated for a range of a values on [0, 1]. This is an increasing function of a. The a values are values of F. Then, the right side of (11) yields the time to arrive to achieve the probability a in the equilibrium arrival CDF. In this way we can evaluate the function F numerically.

A special case: If $b_n = 1$, then the right side of (11) is 0 and we do not get a solution. In this case, we know the solution is for all to arrive at time 1 when the ferry departs. The reason that this happens is that the equilibrium condition is $P(t) = 1 - wt$. If there is ample room for certain, then $P(t) = 1$ and the equilibrium equation does not have a solution for all t. If $b_n < 1$, then there is some incentive to arrive before the departure time.

Appendix B. Supplementary material

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.trb.2010.01.005.

References

Armor, D.A., Taylor, S.E., 2002. When predictions fail: the dilemma of unrealistic optimism. In: Gilovich, T., Griffin, D., Kahneman, D. (Eds.), Heuristics and Biases: The Psychology of Intuitive Judgment. Cambridge University Press, Cambridge, pp. 334–347.

Arnott, R., de Palma, A., Lindsey, R., 1989. Schedule delay and departure time decisions with heterogeneous commuters. Transportation Research Record (1197), 137–153.

Arnott, R., de Palma, A., Lindsey, R., 1993. A structural model of peak-period congestion: a traffic bottleneck with elastic demand. American Economic Review 83 (1), 161–179.

Bailey, N.T.J., 1954. On queueing processes with bulk service. Journal of the Royal Statistical Society (Series B) 16 (1), 80–87.

Barnett, A., 1973. On operating a shuttle service. Networks 3 (4), 305–313.

Buckley, T., Sniezek, J., 1992. Passion, preference, and predictability in judgmental forecasting. Psychological Reports 70 (2), 1022.

Camerer, C.F., 2003. Behavioral Game Theory. Princeton University Press, Princeton.

Chaudhry, M.L., Templeton, J.G.C., 1983. A First Course in Bulk Queues. Wiley, New York.

Daniel, T.E., Gisches, E., Rapoport, A., 2009. Departure time in Y-shaped traffic networks with multiple bottlenecks. American Economic Review. 99 (5), 2149–2176.

De Palma, A., Lindsey, R., 2001. Optimal timetables for public transportation. Transportation Research Part B 35 (8), 789–813.

Dunning, D., Meyerowitz, J.A., Holzberg, A.D., 1989. Ambiguity and self-evaluation: the role of idiosyncratic trait definitions in self-serving assessments of ability. Journal of Personality and Social Psychology 57 (6), 1082–1090.

Glazer, A., Hassin, R., 1987. Equilibrium arrivals in queues with bulk service at scheduled times. Transportation Science 21 (4), 273–278.

Hassin, R., Haviv, M., 2003. To Queue or Not to Queue: Equilibrium Behavior in Queueing Systems. Kluwer, Boston.

Holt Jr., C.A., Sherman, R., 1982. Waiting-line auctions. Journal of Political Economy 90 (2), 280–294.

Irwin, F.W., 1953. Stated expectations as functions of probability and desirability of outcomes. Journal of Personality 21, 329–335.

Kagel, J., Roth, A.E., 1995. Handbook of Experimental Economics. Princeton University Press, Princeton.

Kahneman, D., Knetsch, J.L., Thaler, R.H., 1991. Anomalies: the endowment effect, loss aversion, and status quo bias. Journal of Economic Perspectives 5 (1), 193–206.

Kosten, L., 1973. Stochastic Theory of Service Systems. Pergamon Press, Oxford.

Kraus, M., 2003. A new look at the two-mode problem. Journal of Urban Economics 54 (3), 511–530.

Kraus, M., Yoshida, Y., 2002. The commuter's time-of-use decision and optimal pricing and service in urban mass transit. Journal of Urban Economics 51 (1), 170–195.

Mehdi, J., 1975. Waiting time distribution is a Poisson queue with bulk service. Management Science 21 (7), 777–782.

Naor, P., 1969. The regulation of queue size by levying tolls. Econometrica 37 (1), 15–24.

Neuts, M.F., 1967. A general class of bulk queues with Poisson input. Annals of Mathematical Statistics 38 (3), 759–770.

Pazgal, A.I., Radas, S., 2008. Comparison of customer balking and reneging behavior to queueing theory predictions: an experimental study. Computers & Operations Research 35 (8), 2537–2548.

Rapoport, A., Stein, W.E., Parco, J.E., Seale, D.A., 2004. Equilibrium play in single-server queues with endogenously determined arrival times. Journal of Economic Behavior and Organization 55 (1), 67–91.

Seale, D.A., Parco, J.E., Stein, W.E., Rapoport, A., 2005. Joining a queue or staying out: effects of information structure and service time on arrival and staying out decisions. Experimental Economics 8 (2), 117–144.

Stein, W.E., Rapoport, A., Seale, D.A., Zhang, H., Zwick, R., 2007. Batch queues with choice of arrivals: equilibrium analysis and experimental study. Games and Economic Behavior 59 (2), 345–363.

Tian, Q., Huang, H.-J., Yang, H., 2007. Equilibrium properties of the morning peak-period commuting in a many-to-one mass transit system. Transportation Research Part B 41 (6), 616–631.

Vickrey, W.S., 1969. Congestion theory and transport investment. American Economic Review 59 (2), 251–260.

Weiss, H.J., 1981. Further results on an infinite capacity shuttle with control at a single terminal. Operations Research 29 (6), 1212–1217.

Ziegelmeyer, A., Koessler, F., Boun My, K., Denant-Boèmont, L., 2008. Road traffic congestion and public information: an experimental investigation. Journal of Transport Economics and Policy 42 (1), 43–82.