Available online at www.sciencedirect.com
ScienceDirect
Games and Economic Behavior 59 (2007) 345–363
GAMES and Economic Behavior
www.elsevier.com/locate/geb

Batch queues with choice of arrivals: Equilibrium analysis and experimental study

William E. Stein [a,*], Amnon Rapoport [b,c], Darryl A. Seale [d], Hongtao Zhang [c], Rami Zwick [c]

[a] *Mays Business School, Texas A&M University, College Station, TX 77843, USA*
[b] *Department of Management and Policy, University of Arizona, Tucson, AZ 85721, USA*
[c] *Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong*
[d] *Department of Management, University of Nevada Las Vegas, Las Vegas, NV 89154, USA*

Received 27 May 2006
Available online 15 November 2006

Abstract

We study the decisions agents make in two queueing games with endogenously determined arrivals and batch service. In both games, agents are asked to independently decide when to join a queue, or they may simply choose not to join it at all. The symmetric mixed-strategy equilibrium of two games in discrete time where balking is prohibited and where it is allowed are tested experimentally in a study that varies the game type (balking vs. no balking) and information structure (private vs. public information). With repeated iterations of the stage game, all four experimental conditions result in aggregate, but not individual, behavior approaching mixed-strategy equilibrium play. Individual behavior can be accounted for by relatively simple heuristics.

JEL classification: C90; Z13

Keywords: Batch queueing; Equilibrium solution; Experimental economics

* Corresponding author. Fax: +1 979 845 5653.
E-mail address: wstein@tamu.edu (W.E. Stein).

0899-8256/$ – see front matter
doi:10.1016/j.geb.2006.08.008

0. Introduction

This paper reports the results of an experimental study of the decision if and when to join a queue with batch service of constant size. Queueing systems with batch service are common in transportation markets and have considerable economic implications. Ferries, buses, cable cars, and intra-campus shuttles that run on schedule can carry a maximum number of passengers at a time; the duration of the trip and departure times are fixed and commonly known (Glazer and Hassin, 1987). In these queueing situations, agents typically have the option of staying out of the queue. If the queue is joined, waiting time depends on the agent's arrival time, arrival times of the other customers, capacity of the server, queue discipline, and whether the queue is observable.

Batch queueing is closely related to the economic study of traffic congestion in roads with bottlenecks (see, e.g., Hall, 1991, on bulk service in queueing networks). Vickrey (1963, 1969) might have been the first (Arnott et al., 1990) to consider an interactive decision situation where a fixed number of symmetric commuters have to travel from a single origin to a single destination along a single road in rush hour. The road has a single bottleneck with a fixed capacity or service rate. If the arrival rate at the bottleneck exceeds its capacity a queue develops. Therefore, a commuter who has to determine when to leave the origin faces the tradeoff between time of departure and costly waiting. Vickrey (1969) constructed the no-toll equilibrium, determined the social optimum, and also solved for the toll that decentralizes it.

Operations research studies of batch queueing systems usually assume an infinite stream of identical customers who arrive according to a Poisson process at a service facility with a single server that serves them in batches. The size of each batch has a (possibly infinite) limit. A fixed service fee is incurred as well as waiting cost that is an increasing function of the waiting time. Batch queueing systems with random service times have been studied by Bailey (1954), Chaudhry and Templeton (1983), Medhi (1975), and others. Batch queueing systems with deterministic service times have been discussed by Barnett (1973), Chaudhry and Templeton (1983), and others. An excellent reference is Hassin and Haviv (2003) that surveys the equilibrium behavior of customers in different queueing systems. These studies try to establish optimal service policies or compute efficient control limits (Glazer and Hassin, 1987). They leave no room for the customer to make decisions, as the arrival pattern is typically assumed to be exogenously determined. Nor do they seriously care about human behavior in queues. But as Saaty reminded us more than forty years ago, most queueing systems are designed for human beings. "Studying causes and remedies of queueing problems cannot be completely divorced from consideration of human factors and their influence on the problem" (Saaty, 1961, p. 4).

In line with this observation, our approach and goals differ from these studies in three major respects. First, our goal is descriptive not normative. We wish to identify and explain whatever behavioral regularities emerge when financially motivated members of large groups have to decide independently whether to join a queue, and if so at what time to arrive. For this purpose, we create two batch queueing games, derive and fully characterize the equilibrium solutions, and then compare observed to predicted behavior in a fully controlled environment. Since it is unlikely that subjects will adhere to equilibrium play in a single-shot game, we iterate the stage game a large number of times to determine the effects of experience. Second, we study batch queueing systems where players are required to decide whether to join the queue, and if so at what time to arrive, without making additional assumptions about inter-arrival times or the distribution of number of customer arrivals. Third, we assume finite and commonly known number of customers rather than infinitely large populations. In addition to the practical consideration of not being able to simulate an infinite population model in the laboratory, the finite and known population size better describes situations where a small group of customers repeatedly use the same queuing system.

Queues that form to embark on ferries are typically observable as it is possible upon arrival to count how many cars are already in the queue and compare this number to the ferry's known capacity. The queue's length in this case is not expected to exceed the service capacity because late arrivals will balk. Late arrivals incur the cost of attempting to join the queue, but are not charged with waiting cost, nor rewarded by the service. 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. We consider both cases of observable and unobservable queues in our study. Reneging (leaving the queue after it has been joined) is often observed in queues where customers are served one at a time and service time fluctuates greatly from one customer to another or when service is known to deteriorate with time. Customers in these queues can learn and revise their beliefs about the timing and probability of service and might find it beneficial to renege. However, in a batch queueing system no further information to what has already been available upon arrival is provided; once the decision has been made to join the queue, no reneging is expected. Finally, we study the effects of information each agent might have about the decisions and outcomes of others in previous rounds of play, as this information may prove important in changing her own behavior over time. We consider two extreme cases. In the public information condition, each player is fully informed at the end of each round about the decisions of others, whereas in the private information condition no information about the decisions of others is revealed.

Rapoport et al. (2004) and Seale et al. (2005) conducted two queueing experiments that focused on endogenous arrival and staying out decisions. However, unlike the present study where players are served in batches with capacity that is smaller than the group size, in both of these earlier studies service time was positive and fixed and customers were served one at a time. Comparison of the equilibrium solutions in these two previous studies with the ones constructed in the present paper shows that queueing behavior with or without batch service is quite different. In particular, when customers are served one at a time they do not face the same type of tradeoff between time of departure and costly waiting. Rapoport et al. and Seale et al. reported consistent and replicable patterns of behavior that were accounted for quite accurately on the aggregate, but not individual, level by a symmetric mixed-strategy equilibrium. These findings give rise to the hypothesis that even if aggregate behavior in queues may be accounted for by the symmetric mixed-strategy equilibrium solution, individual play may not.

The rest of the paper is organized as follows. Section 1 presents the two batch queueing games and characterizes their equilibrium solutions, and Section 2 presents the experimental design and the results. Section 3 summarizes the main conclusions.

1. The batch queueing games

1.1. The model

The batch queueing game is presented as a choice of arrival time to a ferry that departs once a day at a commonly known time. Players have the option of staying out and receiving a fixed reward. The assumptions of the game are as follows:

Service Time: The ferry departs at a fixed time T. Service time is 0.

Calling Population: The calling population is finite of commonly known size n.

Arrival Pattern: Each player must decide independently whether to join the queue at or before time T, and if so at what time to arrive. Decisions are made at 5-minute intervals.

Tie-Breaking Rule: If multiple players arrive at the same time interval, their order of arrival is determined randomly. Once ties are resolved, the players are informed immediately of the outcome.

Reneging and Balking: Reneging is prohibited. We study two conditions, one (Condition WB) that allows for balking and the other (Condition NB) that prohibits it.

Queue Discipline: FCFS.

Number of Servers: One.

Service capacity: $s(s < n)$.

Payoff Structure: Assume no balking (unobservable queue). Then, each player who joins the queue is charged (1) a fixed entry fee f, and (2) a variable cost c per minute of waiting in the queue until the ferry departs. If he completes the service successfully (i.e., embarks on the ferry), he receives a fixed reward r. Each player who stays out of the queue receives a fixed payoff $g(g < r - f)$. The resulting payoff function, the same for each player $i(i = 1, \ldots, n)$, takes the form:

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

In the payoff function above, w_i is the time (in minutes) player i waits for the departure of the ferry. The waiting cost c, entry fee f, and reward r are assumed to be positive, whereas the payoff g for staying out can be positive or negative.

Next, assume that balking is allowed. This is the case if the queue is observable. In our experiment balking occurs if s or more players precede player i in the queue. In this case, a player does not incur the variable waiting cost unless he is assured of completing the service successfully. The payoff function, which is the same for each player, takes the form:

$$H_i = \begin{cases} g, & \text{if player } i \text{ stays out of the queue,} \\ -f, & \text{if player } i \text{ is assured of not completing service,} \\ r - f - cw_i, & \text{if player } i \text{ waits } w_i \text{ minutes and completes service.} \end{cases}$$

1.2. Pure and mixed equilibrium solutions

Table 1 presents the symmetric mixed-strategy equilibria for the parameter values $n = 20$, $f = 40$, $c = 4$ (per minute), $r = 340$, $g = 60$, $T = 12:00$, and capacity s ranging from 5 to 19. In our experiment, $s = 14$ and that portion of Table 1 appears in bold. The equilibrium strategies for the case of no balking are displayed in the top panel of Table 1; those for the case where balking is allowed are shown in the bottom panel. For all values of s, the earliest arrival time in equilibrium is 11:05 (with an associated payoff of 80 if service is obtained). Entering the queue at 11:00 and receiving service yields a payoff of 60 that is the same as the one for staying out. Table 1 shows that when balking is prohibited (top panel), as s increases from 5 to 18 the probability of entering at 11:05 increases and the probability of staying out decreases. When $s = 14$, as in our experiment, each player should enter the queue at 11:05 with probability 0.702, stay out with probability 0.236, and enter at times 11:45, 11:50, 11:55 and 12:00 with respective probabilities of 0.024, 0.006, 0.025, and 0.007. When balking is allowed (bottom panel), the pattern is similar

Table 1

Symmetric mixed-strategy equilibria for balking/no balking where $n = 20$, $f = 40$, $c = 4$ (per minute), $g = 60$, $r = 340$, $T = 12{:}00$ and $s = 5, \ldots, 19$*

Balking prohibited

Server capacity

Time	$s = 5$	$s = 6$	$s = 7$	$s = 8$	$s = 9$	$s = 10$	$s = 11$	$s = 12$	$s = 13$	$s = 14$	$s = 15$	$s = 16$	$s = 17$	$s = 18$	$s = 19$
11:05	0.180	0.233	0.288	0.345	0.402	0.461	0.520	0.580	0.641	**0.702**	0.764	0.827	0.889	0.952	0.000
11:25	0.025	0.013	0.004	0.000	0.000	0.000	0.000	0.000	0.000	**0.000**	0.000	0.000	0.000	0.000	0.000
11:30	0.005	0.018	0.033	0.021	0.008	0.000	0.000	0.000	0.000	**0.000**	0.000	0.000	0.000	0.000	0.000
11:35	0.024	0.012	0.000	0.012	0.026	0.028	0.013	0.000	0.000	**0.000**	0.000	0.000	0.000	0.000	0.000
11:40	0.005	0.018	0.031	0.021	0.007	0.006	0.021	0.030	0.013	**0.000**	0.000	0.000	0.000	0.000	0.000
11:45	0.024	0.013	0.001	0.012	0.026	0.028	0.013	0.003	0.019	**0.024**	0.003	0.000	0.000	0.000	0.000
11:50	0.005	0.019	0.032	0.021	0.008	0.006	0.021	0.030	0.012	**0.006**	0.025	0.005	0.000	0.000	0.000
11:55	0.026	0.014	0.002	0.014	0.028	0.029	0.013	0.003	0.020	**0.025**	0.003	0.020	0.000	0.000	0.000
12:00	0.008	0.022	0.036	0.024	0.009	0.008	0.023	0.032	0.014	**0.007**	0.026	0.006	0.022	0.000	1.000
Stay Out	0.700	0.638	0.578	0.531	0.486	0.435	0.377	0.322	0.282	**0.236**	0.179	0.142	0.089	0.048	0.000

Balking allowed

Server capacity

Time	$s = 5$	$s = 6$	$s = 7$	$s = 8$	$s = 9$	$s = 10$	$s = 11$	$s = 12$	$s = 13$	$s = 14$	$s = 15$	$s = 16$	$s = 17$	$s = 18$	$s = 19$
11:05	0.260	0.324	0.389	0.454	0.518	0.583	0.647	0.711	0.774	**0.837**	0.899	0.960	1.000	0.000	0.000
11:25	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**0.000**	0.000	0.000	0.000	1.000	0.000
11:35	0.005	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**0.000**	0.000	0.000	0.000	0.000	0.000
11:40	0.011	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**0.000**	0.000	0.000	0.000	0.000	0.000
11:45	0.003	0.016	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**0.000**	0.000	0.000	0.000	0.000	0.000
11:50	0.010	0.000	0.007	0.000	0.000	0.000	0.000	0.000	0.000	**0.000**	0.000	0.000	0.000	0.000	0.000
11:55	0.002	0.011	0.006	0.000	0.000	0.000	0.000	0.000	0.000	**0.000**	0.000	0.000	0.000	0.000	0.000
12:00	0.009	0.000	0.005	0.007	0.000	0.000	0.000	0.000	0.000	**0.000**	0.000	0.000	0.000	0.000	1.000
Stay Out	0.700	0.649	0.593	0.539	0.482	0.417	0.353	0.289	0.226	**0.163**	0.101	0.040	0.000	0.000	0.000

* Only times (rows) with positive entry probability for some s values are included.

with smaller probabilities of entering the queue after 11:05. However, if $9 \leqslant s \leqslant 6$ (and, in particular, $s = 14$), then players should either enter at 11:05 or stay out but never enter after 11:05. A comparison of the equilibrium solutions for queues with or without balking shows that for all server capacity values, players opting to join the queue are expected to do so earlier when balking is allowed.

Characteristic of both games is that the equilibrium solutions are inefficient (Hassin and Haviv, 2003) in the sense of maximizing the overall social welfare. In both games, social optimum is achieved when s players enter the queue at exactly 12:00 and $n - s$ players stay out. This results in group payoff of $s(r - f) + (n - s)g$. When n is large and communication is prohibited, this coordination can only be achieved by luck. Another, more plausible, strategy that does not require this sort of coordination calls for all the n subjects to arrive at 12:00; it is very easy to implement even without communication. It, too, outperforms the equilibrium payoff by a large margin. With the parameter values used in our study, the social optimum group payoff is $s(r - f) + (n - s)g = 4560$ compared to the group equilibrium payoff of ng=1200, a 3.8 : 1 ratio. If all the n subjects arrive at 12:00, thereby saving on waiting costs and maintaining symmetry between players, the resulting group payoff is $s(r - f) - (n - s)f = 3960$, a 3.3 : 1 ratio. The latter ratio provides a strong incentive for tacit collusion.

In addition to the mixed-strategy equilibria, other equilibria exist. If $s = 14$ and balking is allowed, 16 arriving at 11:05 and 4 staying out is an asymmetric pure strategy equilibrium resulting in unequal expected payoffs of 65 and 60, respectively. If $s = 14$ and balking is not allowed, there is no pure strategy equilibrium. Since the players are symmetric, we will focus on symmetric mixed-strategy equilibria (Table 1) which are affected by the balking/no-balking distinction but not the information structure. Derivation of the mixed-strategy equilibrium can be found in the supplementary material associated with this article in the on-line version (see Supplementary Appendix).

2. Experiment

2.1. Method

Subjects. Two hundred and forty subjects participated in the experiment. All the subjects volunteered to take part in a decision making experiment with payoff contingent on performance. All of them were undergraduate or graduate students at the Hong Kong University of Science and Technology (HKUST). The subjects were divided into 12 equal-size groups, three groups in each of four different experimental conditions (see below). Each group participated in a single session that lasted, on average, 90 minutes.

Design. We used a 2×2 *information structure* (no information vs. full information) by *game type* (balking allowed vs. balking prohibited) between-subject experimental design with three independent groups in each condition for a total of 12 groups. Hereafter, we shall refer to these four experimental treatments as NINB (no information, no balking), NIWB (no information, with balking), FINB (full information, no balking), and FIWB (full information, with balking). The game type factor was introduced for studying the effects of balking allowed vs. balking prohibited on behavior; see Table 1 for the different equilibrium distributions. The information structure factor was introduced for studying the effects of outcome information on the dynamics of play across iterations. In the two full-information (public information). Conditions FINB and FIWB, at the end of each trial subjects were presented with a computer screen showing the number of players who joined the queue at different time intervals, the cumulative distribution of arrival times, the number of subjects who opted to stay out of the queue, the subject's payoff for the trial, and his cumulative payoff from trial 1. In the two no-information (private information) Conditions NINB and NIWB, at the end of the trial each subject was only informed of his payoff for the trial and his cumulative payoff. In the two "no balking." Conditions (unobservable queue) NINB and FINB, the variable waiting cost was deducted whether or not service was completed, whereas in the two "with balking" conditions (observable queue) the variable waiting cost was incurred only if service was completed. The values of n, s, g, r, f, and c were commonly known. With the ferry departing at $T = 12:00$, subjects were allowed to join the queue at or after 10:00. Arrival times were restricted to 5-minute intervals (i.e., 10:00, $10:05, \ldots, 12:00$). The subject instructions for Condition FINB can be found in the supplementary material associated with this article in the on-line version (see Supplementary Appendix). Instructions for the other conditions are similar.

Payoffs were stated in terms of a fictitious currency called "francs." At the end of the session, the cumulative payoffs were computed and converted into money at the rate of 40 francs = HK\$1.00 (US\$1.00 = HK\$7.78). Individual payoffs ranged between 0 and HK\$130 with mean payoff of HK\$54.00.

Procedure. The experiment was conducted in a computer laboratory at HKUST that includes 80 networked PC terminals. Upon arrival at the laboratory, the subjects were seated as far away from one another as possible and provided with sets of instructions that they read at their own pace. Any form of communication between them was forbidden. Questions asked about the game (there were only a few) were privately answered by the experimenter.

Each stage game was repeated 60 times. Once the trial number was displayed on the PC, each subject was asked to choose whether to join the queue. If opting to do so, he or she was further asked to choose his/her arrival time. The decisions were made anonymously and independently. Identification of individual subjects was not possible. No time pressure was imposed. Once all the group members typed in their decisions, a "Results" screen was displayed informing the subject of (1) his decision, (2) success or failure in embarking on the ferry, (3) the subject's payoff for the trial, and (4) the subject's cumulative payoff. Subjects in the two full-information conditions were also informed of how many players joined the queue at each time interval. This information was also displayed as cumulative frequency distribution. Hence, information on the latest arrival time (on the just completed trial) that still allowed players to embark on the ferry was easily accessible.

2.2. Results

2.2.1. Aggregate behavior: first vs. last block

Table 2 presents distributions of the proportions of entry time and staying out decisions by condition and group. The results are presented separately for the first block of five trials when subjects had very little experience with the queueing game, and for the last block of five trials after they have gained considerable experience. To facilitate comparison with the equilibrium predictions, the proportions are classified in terms of five categories: Entering at or before 11:00; Entering at 11:05; Entering between 11:10 and 11:40 (inclusive); Entering between 11:45 and 12:00 (inclusive); Staying out. Note that the payoff for entering at 11:00, conditional upon completion of service, is identical to the one guaranteed for staying out. However, entry at 11:00 does not necessarily result in successful service. Entry before 11:00 is irrational as it yields a smaller payoff than staying out.

When the same n players interact repeatedly over time, the appropriate statistical unit of analysis for testing the effect of the independent variables (i.e., Balking and Information) is the group. Although three data points per cell allows for a statistical test with low power, some effect sizes were large enough to reach significance.

Table 2
Proportions of entry and staying out decisions by group for the first (trials 1–5) and last (trials 56–60) blocks in the session

Condition	Entry time	Trials 1–5			Trials 56–60			Equilibrium
		Group			Group			
		1	2	3	1	2	3	
FINB	⩽11:00	.12	.02	.04	.32	.00	.29	**0**
	11:05	.09	.12	.03	.42	.52	.46	**.702**
	11:10–11:40	.67	.74	.80	.00	.19	.00	**0**
	11:45–12:00	.02	.00	.01	.10	.10	.12	**.062**
	Stay out	.10	.12	.12	.16	.19	.13	**.236**
		$\chi^2 = 17.968$		(0.021)	$\chi^2 = 71.182$		(<0.0001)	
NINB	⩽11:00	.03	.04	.02	.04	.08	.07	**0**
	11:05	.03	.08	.00	.30	.57	.40	**.702**
	11:10–11:40	.83	.77	.86	.18	.12	.28	**0**
	11:45–12:00	.02	.02	.06	.26	.02	.03	**.062**
	Stay out	.09	.09	.06	.22	.21	.22	**.236**
		$\chi^2 = 14.038$		(0.081)	$\chi^2 = 52.638$		(<0.0001)	
FIWB	⩽11:00	.08	.07	.06	.30	.07	.29	**0**
	11:05	.07	.02	.08	.49	.43	.55	**.837**
	11:10–11:40	.69	.74	.81	.05	.27	.01	**0**
	11:45–12:00	.04	.11	.00	.02	.06	.01	**0**
	Stay out	.12	.06	.05	.14	.17	.14	**.163**
		$\chi^2 = 21.045$		(0.007)	$\chi^2 = 57.536$		(<0.0001)	
NIWB	⩽11:00	.05	.03	.05	.02	.05	.28	**0**
	11:05	.03	.01	.10	.70	.72	.50	**.837**
	11:10–11:40	.83	.92	.79	.14	.07	.05	**0**
	11:45–12:00	.05	.02	.02	.04	.01	.03	**0**
	Stay out	.04	.02	.04	.10	.15	.14	**.163**
		$\chi^2 = 14.034$		(0.081)	$\chi^2 = 47.291$		(<0.0001)	

Because the assumption of subject independence is invalid in a stage game that is repeated multiple times, statistical tests based on this assumption are questionable.[1] But they are valuable when combined with inspection of the group data. Eight chi-square tests (5 × 3 tables) were conducted to test for group differences within condition. The χ^2 values and the associated p values (d.f. = 8) are presented in Table 2 for each combination of condition and block of trials. Table 2 displays several discernible patterns.

• *Group differences*. The three groups in Conditions NINB and in NIWB did not differ from one another in the first block. However, with experience the groups in each of these two conditions diverged: compare the χ^2 values in the first and last blocks for each of these two conditions. The groups in Conditions FINB and in FIWB already exhibited group differences in the first block. Similar to NINB and NIWB, the three groups in both of Conditions FINB and FIWB diverged further as their members gained more experience with the game. One would expect to see

[1] In several analyses players are assumed to be independent. In justification of this assumption, we list two reasons. First, as group size becomes larger the effect of any particular player becomes negligible. With groups of 20 players each, treating the group as a population is not unreasonable. Second, our experimental design does not allow for establishing reputation, as the identity of individual subjects is not revealed. This mitigates any effect that a particular player may have on the choices of other group members.

greater group differences under the full-information condition. The results for the first block are in line with this hypothesis. One would also expect larger group differences with experience, as players react to histories of previous play that become longer. This hypothesis, too, is supported by the results in Table 2.

• *Shift to equilibrium play*. Table 2 provides evidence that with experience aggregate behavior approaches the equilibrium predictions in most categories.

(i) The proportions of irrational entries (entering before 11:00) declined with experience from 3 percent in the first block to 0.5 percent in the last block. The arrival frequencies at 11:00 increased and actually moved away from equilibrium over time. In particular, 5 of the 12 groups had arrival rates above 0.20 in the final block while the equilibrium was 0. The arrivals at 11:05 increased in the direction of equilibrium but were still significantly lower than the equilibrium proportions. However, collapsing data over the first two categories and thereby considering all arrivals at or before 11:05, the arrival rates for 8 of the 12 groups in the last block were not significantly different (by t-test for proportions) from the equilibrium. (The exceptions are Group 2 in Condition FINB, Groups 1 and 3 in Condition NINB, and Group 2 in Condition FINB).

(ii) In all 12 groups the proportion of entry between 11:10 and 11:40 sharply decreased in the direction of equilibrium play. Across groups, the proportion of entry in this time interval decreased from 0.788 to 0.113 while the equilibrium was 0.

(iii) The time interval 11:45–12:00 yielded minimal change with experience. While the equilibrium was 0 in Conditions FIWB and NIWB, the observed proportion in each group was about 0.03 in block 1 and, if anything, tended to increase slightly in the last block in the case of no balking but stayed at a low level when balking was allowed.

(iv) With experience, in all 12 groups the proportion of decisions to stay out moved toward the equilibrium. The equilibrium proportion is 0.236 for the No Balking and 0.163 for the With Balking conditions, whereas the corresponding observed mean proportions are 0.19 and 0.14 in the last block.

We conclude that with experience the groups' aggregate behavior in all four conditions shifted in the direction of equilibrium play. Additional evidence in support of this conclusion is reported below.

• *Effects of balking*. (i) In equilibrium, players should enter the queue at 11:05 more frequently when balking is allowed (83.7 percent) than when balking is prohibited (70.2 percent). This prediction was supported in the no-information condition (last block): compare the three proportions in Condition NIWB (0.70, 0.72, 0.50) to the three proportions in Condition NINB (0.30, 0.57, 0.40). The difference between the two conditions is significant ($t(4) = 2.05$, one-tailed test, $p < 0.06$). In contrast, there was no difference between Conditions FIWB (0.49, 0.43, 0.55) and FINB (0.42, 0.52, 0.46) ($t(4) = 0.52$, ns).

(ii) In equilibrium, players should stay out more frequently when balking is prohibited (23.6 percent) than when it is allowed (16.3 percent). Consider first the no-information conditions. The proportions of staying out decisions (in the last block) in the three groups in Condition NINB (0.22, 0.21, 0.22) exceeded the ones in the three groups in Condition NIWB (0.10, 0.15, 0.14). Using the group as the statistical unit of analysis, this difference is significant ($t(4) = 5.54$, one-tailed test, $p < 0.01$). In contrast, there was no difference between Conditions FINB (0.16, 0.19, 0.13) and FIWB (0.14, 0.17, 0.14) ($t(4) = 0.5$, ns).

• *Effects of information*. We have already remarked that groups diverged from one another more under the full-information conditions in comparison to the no-information conditions. In the last block, early entry (at or before 11:00) was more frequent under the full-information than no-information conditions, although not significantly so. In particular, the proportions of early entry in the six groups in Conditions FINB and FIWB (0.32, 0, 0.29, 0.30, 0.07, 0.29) averaged 0.21 while those in Conditions NINB and NIWB (0.04, 0.08, 0.07, 0.02, 0.05, 0.28) averaged 0.09.

2.2.2. Aggregate dynamics across the session

Our experimental design calls for 24 possible entry decisions plus a single staying out decision. Denote any of these 25 strategies by j and its relative frequency within a group on trial $t\,(t = 1, \ldots, 60)$ by p_{jt}. Denote the probability of choosing strategy j under the mixed-strategy equilibrium solution by p_j^*. For each group and each trial separately, we computed a deviation index $d_t = \sqrt{\sum_{j=1}^{25}(p_{jt} - p_j^*)^2}$ between the observed and predicted proportions. This gives the Euclidean distance between these two probability distributions. Figure 1A exhibits the 5-step

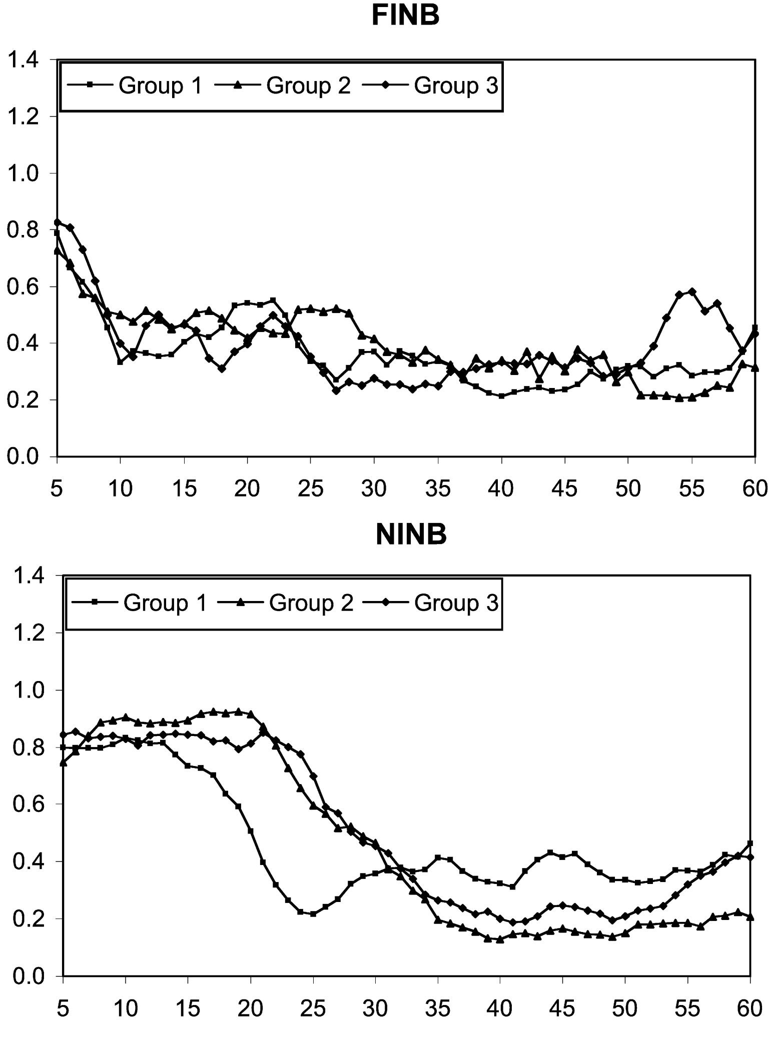

Fig. 1A. Five-trial moving average of Euclidean distance between observed and equilibrium proportions: No Balking condition.

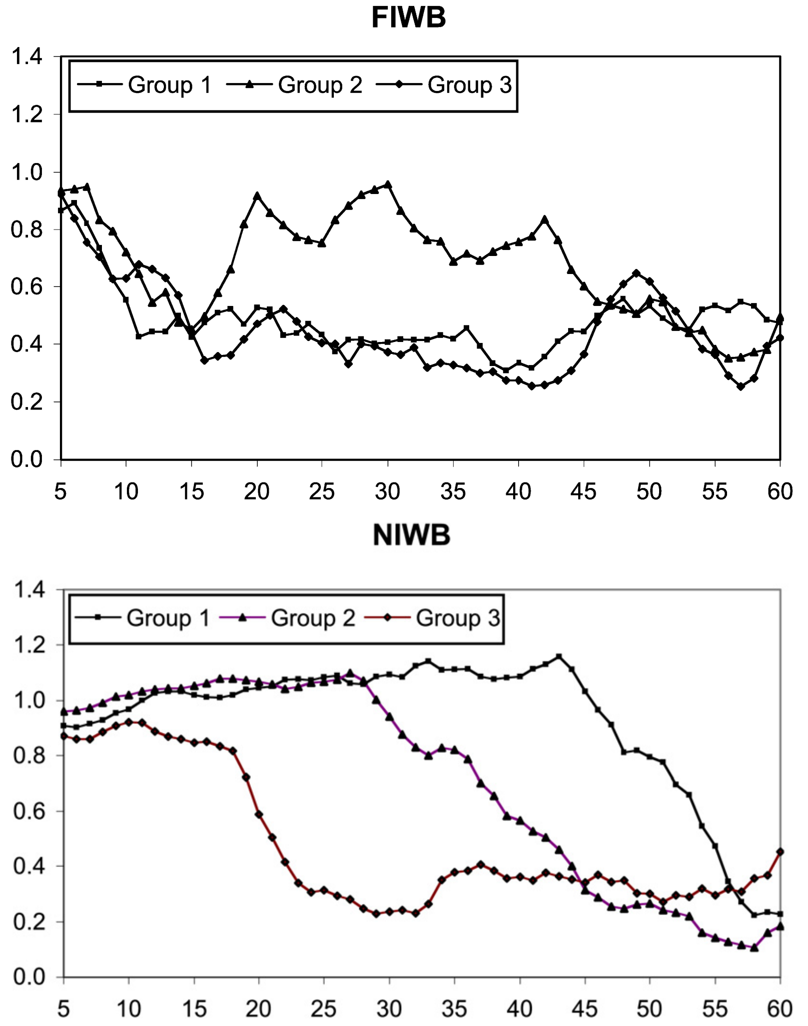

Fig. 1B. Five-trial moving average of Euclidean distance between observed and equilibrium proportions: With Balking condition.

moving average of d_t for the two no-balking conditions. The results are shown separately for each group. Figure 1B exhibits similar moving averages for the two with-balking conditions. Consistent with the conclusion drawn in the previous section, the deviation index decreases with experience for each of the 12 groups. Convergence is much faster in the full-information than in the no-information conditions. In the full-information conditions the d_t measure declines sharply almost from the beginning of the session and reaches its limiting value after about 15 trials (Group 2 in Condition FIWB is the only exception), whereas in the two no-information conditions in both Figs. 1A and 1B, the d_t index starts declining only after 15 to 40 trials. The much faster convergence in the full information condition is evidence that information feedback facilitated learning in the direction of the equilibrium. The effect of balking is also clearly evident from a comparison of Figs. 1A and 1B. With balking (Fig. 1B), the d_t scores tend to be larger than in the no balking case (Fig. 1A). However, at the end of the 60 trials the d_t scores in all four conditions are quite similar. As a point a comparison, if all players used the equilibrium strategy by randomizing over trials, d_t would average 0.10 for with balking and 0.13 for no balking. (It is not 0 since each trial consists of 20 decision points and we cannot come arbitrarily close to the equilibrium in this situation.)

2.2.3. *Individual differences*

As noted by Erev and Roth (1998) and others, analyses of individual data can reveal information that is totally lost in the averaging process. Figure 2 displays the frequency distributions of entry time and staying out decisions of all the 20 members of Group 1 in Condition FINB. We refer to them as *individual profiles*. Subject numbers from 1 through 20 are listed above each individual profile. The horizontal axis presents the trial number from 1 through 60 and the vertical axis shows the arrival time. The arrival times 11:00, 11:15, 11:30, 11:45 and 12:00 are indicated by horizontal grid lines. In addition, there is a horizontal line at 10:55. Absence of a bar indicates arrivals from 10:00 to 10:55 (there are very few of these). A short vertical line that extends below 10:55 indicates staying out. For example, subject 20 arrived at 11:00 on 22 of the 60 trials, at 11:05 on 14 trials, at 11:30 on a single trial, and stayed out on 23 trials. The individual data displayed in Fig. 2 is representative of the results of the other groups.

Figure 2 displays substantial differences between individual profiles and no support for mixing. It shows that 11 of the 20 subjects stayed out of the queue no more than five times. Most of the decisions to stay out are due to only 4 subjects in the group, namely, subjects 9, 13, 18, and 20. In contrast, 3 subjects (3, 10, 12) never stayed out. About half of the subjects (2, 3, 4, 5, 6, 12, 14, 17, and 19) joined the queue no later than 11:10 on the majority of the trials, whereas others (subjects 16 and 18) tended to join the queue late on a substantial proportion of the trials. In addition, we observe large individual differences in the incidence of switching decisions between trials. For example, subjects 2, 3, and 17 switched their decisions infrequently. Other subjects (e.g., subjects 8, 10, 18, and 20) tended to switch their decisions, on average, every 2–3 trials.

Figure 3 organizes the same individual decisions into frequency distributions, collapsing the decisions across trials and therefore providing no information about the dynamics. Here, the horizontal axis portrays 15 pure strategies (left to right: arriving at or before 10:55 and then from 11:00 to 12:00 at 5 minute intervals and, finally, the staying out strategy) and the vertical axis portrays the relative frequencies in units of 0.15 per interval. A major feature of Fig. 3 is that, with the possible exception of the individual profiles of subjects 8 and 14, none of the twenty individual frequency distributions resemble the aggregate distribution for condition FINB in Table 1. Therefore, whatever explanation is invoked to account for the aggregate results it cannot also be invoked to explain the variety of individual decision profiles.

We tested the hypothesis that the subject's decisions constitute a random sequence generated by probabilities $p_1, p_2, \dots, p_{25} (p_1 + p_2 + \dots, p_{25} = 1)$ associated with this subject's set of observed strategies. We imposed no further constraints on these probabilities and allowed them to vary from one subject to another. The interpretation of this hypothesis is that each subject has stable propensities for using each strategy, and that these propensities may vary across subjects, but for a given subject they are not adjusted over time. The hypothesis of randomly generated sequences was tested separately on the individual level using a bootstrap run test for all 60 trials and for the last 30 trials. For each subject, we simulated 10,000 sequences of 60 (30) decisions, given the subject's overall proportions of selecting each strategy. For each simulated sequence, we calculated the number of runs. The null hypothesis that the actual observed number of runs was generated by a random sequence was tested by comparing the observed number of runs

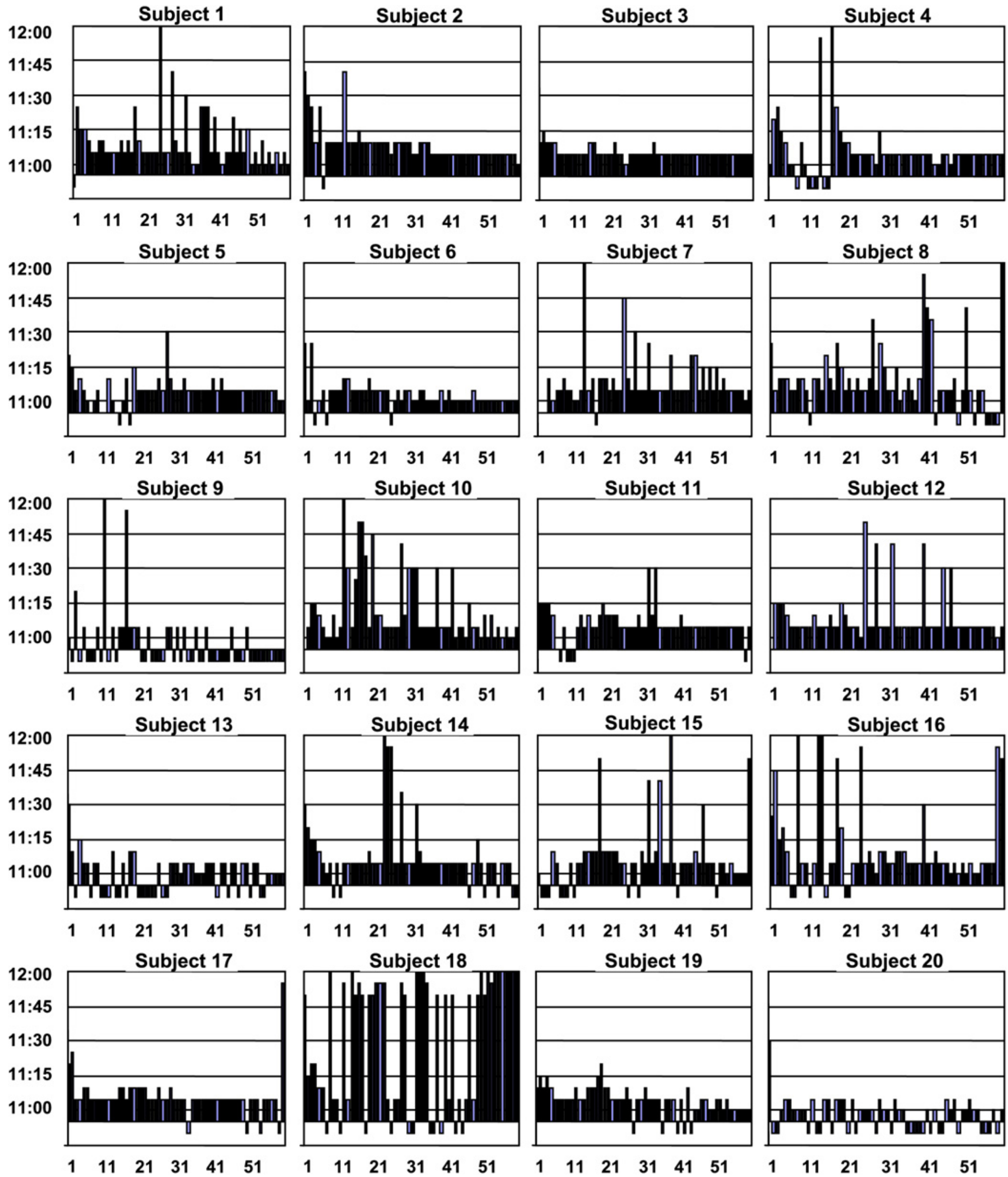

Fig. 2. Individual decisions by trial of all twenty members in Group 1 of Condition FINB.

to the distribution generated by the simulation. The null hypothesis could not be rejected for only 24 (out of 240) subjects using all 60 trials, and only 16 (out of 240) subjects using the last 30 trials (subjects for whom the hypothesis can not be rejected are distributed evenly among conditions and groups.) In most cases, the hypothesis was rejected because of too few runs. This is understandable given the observation that in many cases strategies that are selected quite frequently at the beginning of the session disappear as the session progress. Clearly, the vast

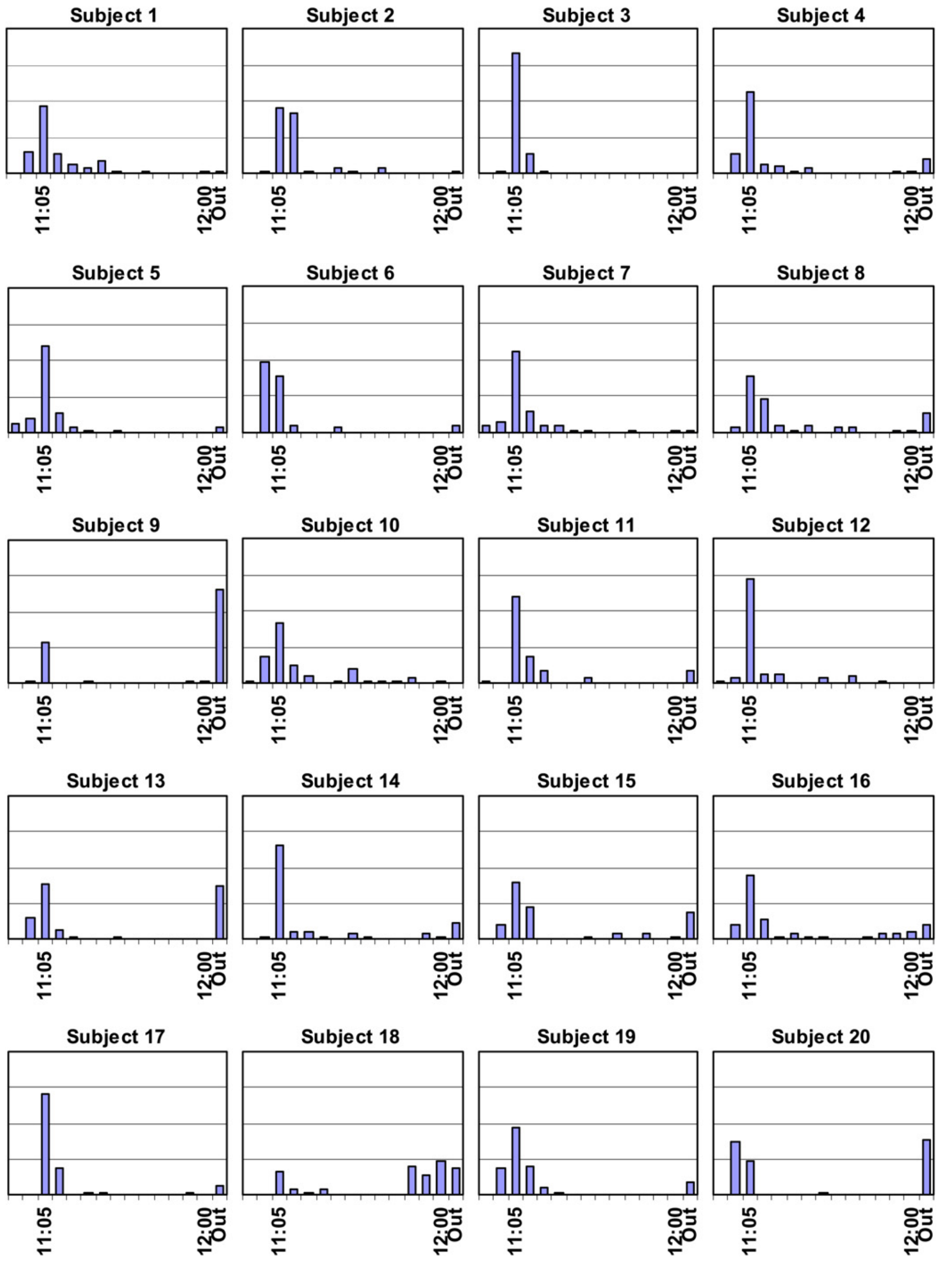

Fig. 3. Frequency distribution of individual decisions of all twenty members in Group 1 of Condition FINB.

majority of subjects do not randomize their decisions. If not randomization, then what drives strategy selection?

2.2.4. Qualitative hypotheses

To answer this question, we conducted several analyses that focus on the effects of the immediately preceding outcome on individual decisions. These analyses are in the spirit of learning direction theory (Selten and Stoecker, 1986; Ockenfels and Selten, 2005), which predicts the direction, but not necessarily the amount, by which decisions change from trial to trial. Although learning direction theory has not been fully specified, Camerer (2003) interprets it as a kind of belief learning that combines elements of Cournot dynamics with habit or inertia. Our first and simplest analysis considers the effect of the outcome on trial $t-1$ on the subject's decision on trial t. Define "success" as joining the queue and receiving service, and "failure" as joining the queue and not receiving service. Categorize the subject's decision on trial t in relation to her decision on the previous trials as

S: entering the queue at the same time as the previous trial,

E: entering the queue earlier than the previous trial,

L: entering the queue later than the previous trial,

O: staying out.

Four directional hypotheses about the subject's decision on trial t were tested:

H1: P(O | failure) > P(O | success),

H2: P(S | success) > P(S | failure),

H3: P(L | success) > P(E | success),

H4: P(E | failure) > P(L | failure).

Hypotheses H1 asserts that it is more likely to stay out of the queue after failure than after success, and hypothesis H2 that it is more likely to enter the queue at the same time after success than failure. These are qualitative hypotheses meant to capture the reaction of naïve subjects to the immediately previous outcome. Hypothesis H3 states that following success, a subject is more likely to delay her entry time than entering earlier. Hypothesis H4 asserts the opposite trend following failure.

Table 3 presents information relevant to these hypotheses. The results are organized by condition and hypothesis in 16 panels. Within each panel, SIG indicates results of a z-test for proportions that are significant at the 0.05 level, NS indicates non-significant results, and NT cases that could not be tested for lack of data. Thus, consider Group 1 in Condition FINB in the H1 panel. The null hypothesis P(O | failure) $\leqslant$ P(O | success) was rejected for one subject, not rejected for 15 subjects, and not tested for 4 subjects. Category NT mostly applies to hypothesis H1, as many subjects never stayed out of the queue.

Table 3 shows that hypothesis H1 accounted for the behavior of 44 of the 71 subjects (62 percent) in the two no-information conditions for whom this hypothesis could be tested, but only for 17 of the 95 subjects (17.9 percent) in the two full-information conditions for whom this hypothesis could be tested. Whereas the majority of the subjects in Conditions NINB and NIWB stayed out on trial t less frequently after success than failure, most of the subjects in Conditions FINB and FIWB were not affected by the information about the previous outcome when deciding to stay out. Hypothesis H2 was supported in all four conditions. Most subjects (107 out of 118

Table 3
Number of subjects for whom z-tests for proportions are significant (SIG), non-significant (NS), or cannot be tested (NT): Hypotheses H1–H4

Condition	Group	H1			H2			H3			H4		
		SIG	NS	NT	SIG	NS	NT	SIG	NS	NT	SIG	NS	NT
FINB	1	1	15	4	13	7	0	4	16	0	17	3	0
	2	2	15	3	12	8	0	3	17	0	14	5	1
	3	3	12	5	8	11	1	2	17	1	17	1	2
FIWB	1	4	12	4	13	7	0	8	12	0	19	0	1
	2	3	11	6	12	8	0	3	17	0	15	5	0
	3	4	13	3	10	10	0	6	14	0	14	5	1
NINB	1	8	6	6	17	3	0	11	9	0	18	1	1
	2	7	6	7	18	0	2	14	5	1	18	0	2
	3	9	4	7	17	3	0	13	7	0	18	1	1
NIWB	1	3	4	13	16	4	0	20	0	0	20	0	0
	2	6	5	9	19	1	0	17	3	0	18	1	1
	3	11	2	7	20	0	0	17	3	0	19	1	0

in the two no-information conditions, and 68 out of 119 in the two full-information conditions) entered the queue at the same time interval as before more frequently following a success than failure entry. Consistent with the previous test of hypothesis H1, the support for H2 was more pronounced in the two no-information conditions (90.7 percent) than in the two full-information conditions (57.1 percent). Stronger effects of information are observed in the results of the test of hypothesis H3, which was supported by 92 of the 119 subjects in the two no-information conditions (77.3 percent) but only by 26 of the 119 subjects in the two full-information conditions (21.8 percent). Finally, Table 3 shows that failure on trial $t-1$ had a stronger effect than success. Following failure in embarking on the ferry on trial $t-1$, 111 of the 115 subjects in the two no-information conditions (96.5 percent) and 96 of the 115 subjects in the two full-information conditions (83.5 percent) were more likely to enter the queue earlier.

Taken together, the results show that hypotheses H2 and H4 were supported by the entry decisions of most of the subjects in all four experimental conditions, with stronger support observed in the two no-information conditions than in the two full-information conditions. Hypotheses H1 and H3 were only supported by the majority of the subjects in the two no-information conditions who were provided no information about the decisions and outcomes of the other group members. Full information seems to have fostered a more sophisticated approach to the batch queueing game.

2.2.5. Local switching in entry times

By a "more sophisticated approach" we mean that subjects in Conditions FINB and FIWB might have fully used the additional information about entry times of all group members that was not available in the two other conditions. Consider a subject in Conditions FINB and FIWB, who entered the queue on trial $t-1$ at time interval k and successfully boarded the ferry. The outcome information she received at the end of trial $t-1$ may be classified into three categories: (1) At least one of the subjects entering at time interval k did not receive service. This occurred if one or more of the subjects entering at time k were not chosen to board the ferry. (2) All the subjects who entered the queue at time k received service, and there were no successful entries at a later time. (3) All the subjects who entered the queue at time k received service, and there was at least one more successful entry at a later time. We refer to these three cases as LAT (Latest Arrival Time), LGAT (Latest Guaranteed Arrival Time), and EA (Early Arrival), respectively.

Recall that hypothesis H3 was rejected for most of the subjects in Conditions FINB and FIWB. We hypothesize that this occurred because subjects reacted not just to the outcome on trials $t-1$ (success or failure) but also to the full outcome information. Refining hypothesis H3, we formulated and tested two additional hypotheses. These are: H3-A: P(L | EA) > P(E | EA) and H3-B: P(E | LAT) > P(L | LAT).

To test these two hypotheses, we identified for each subject in the two full-information conditions the occasions on which she entered the queue and received service (trials $t-1$). We then categorized the information outcome she received on that trial as LAT, LAGT, and EA. We used the z-test for proportions and the $\alpha = 0.05$ level of significance. The results show that hypothesis H3-A was supported by the majority of the subjects in all six groups (78.7 percent across the six groups). Hypothesis H3-B was supported by the majority of the subjects in all groups except Group 2 in Condition FINB.

3. Conclusions

The present study has examined *decentralized* decision making in two different batch queueing games in which the number of players is finite and commonly known, the strategy space is discrete, the stage game is iterated 60 times, and reneging is prohibited. In both games, where balking is allowed and where it is prohibited, players were given the option of staying out of the queue. Our main result (Table 2, Figs. 1A and 1B) is that with experience gained in playing the game aggregate behavior moves in the direction of equilibrium play. Players moved quicker toward equilibrium play when balking was prohibited than when it was not (Figs. 1A and 1B) and when information was public rather than private. The information provided at the end of each round affected the dynamics of play. When the information was private, trial-to-trial changes in entry and staying out decisions largely followed very simple heuristics (hypotheses H1–H4). When it was public, trial-to-trial changes in entry time also took into account the decisions and outcomes of all the group members.

The symmetric mixed-strategy equilibria are inefficient; players could more than triple their earnings by tacitly agreeing not to wait in line at all. In a different context of routing traffic to optimize the performance of a congested network, Papadimitriu (2001) refers to the inefficiency inherent in a selfishly defined solution as the "price of anarchy." Our results show that when decisions are decentralized with $n = 20$ players in a group, players cannot avoid paying this price. Despite a major potential increase in payoff, Table 1 (entry time (11:45, 12:00)) shows only weak evidence in Conditions FINB and NINB and no evidence in Conditions FIWB and NIWB of successful attempts to maximize social welfare.

Our results are consistent with those reported by Rapoport et al. (2004) and Seale et al. (2005), who studied different queueing games in which arrivals are endogenous, service time is positive and fixed, customers are served by a single server one at a time, and the option of staying out is available. The mixed-strategy equilibrium solutions for the games studied in these two experiments are quite different from the ones constructed in the present study. In the study by Seale et al., where early arrivals before the service station opens are allowed, equilibrium arrival times are distributed more or less evenly over most of the strategy space. When early arrivals are prohibited in the study by Rapoport et al., there is a relatively high probability of arrival just when the service station opens, followed by zero probability of arrival over a short segment of the strategy space, and then arrival probabilities that slowly decrease over time as players approach the closing time of the station. In contrast, with the exception of small probabilities of arrival between 11:45 and 12:00 (Table 1) in the present study, under the mixed-strategy equilibrium players either arrive at 11:05 or stay out of the queue. Despite these sharp differences in the equilibrium solutions, we observe similar patterns that have been observed elsewhere (e.g., Rapoport et al. (2000, 2002); Selten et al., 2004; Seale and Rapoport, 2000) of large-scale behavior slowly approaching equilibrium play and patterns of individual behavior that do not support equilibrium play but can be accounted for by relatively simple heuristics.

There is one experimental question and one theoretical issue that remain unresolved. Individual results not reported here indicate considerable differences in both the frequency and magnitude of switching decisions between adjacent trials. Will players eventually converge to pure-strategy play so that with experience different players either join the queue at possibly different times or stay out? Answering this question requires iterating the stage game for a considerably larger number of trials than in the present study, possibly in the hundreds. The theoretical challenge is to proceed beyond direction learning theory by constructing a learning model that accounts *simultaneously* for the individual and aggregate patterns of results that we have documented above. A preliminary attempt in this direction has been made by Bearden et al. (2005), who proposed and tested a reinforcement-type adaptive learning model for the experimental results reported by Rapoport et al. and Seale et al.

Acknowledgments

We gratefully acknowledge financial support by NSF Grant No. SES-0135811 to D. A. Seale and A. Rapoport and by an AFOSR/MURI contract F49620-03-1-0377 to the University of Arizona. We also wish to thank Maya Rosenblatt and Vivian Lam for help in data collection and Tamar Kugler for constructive comments.

Supplementary Appendix

Supplementary material associated with this article can be found, in the on-line version, at doi: 10.1016/j.geb.2006.08.008.

References

Arnott, R., De Palma, A., Lindsey, R., 1990. Economics of a bottleneck. J. Urban Econ. 27, 111–130.

Bailey, N.T.J., 1954. On queueing processes with bulk service. J. Roy. Statistical Society (Ser. B) 16, 80–87.

Bearden, J.N., Rapoport, A., Seale, D.A., 2005. Entry times in queues with endogenous arrivals: Dynamics of play on the individual and aggregate levels. In: Rapoport, A., Zwick, R. (Eds.), Experimental Business Research, vol. II. Springer, Dordrecht, Holland, pp. 201–221.

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

Camerer, C.F., 2003. Behavioral Game Theory. Princeton Univ. Press, New York.

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

Erev, I., Roth, A.E., 1998. Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. Amer. Econ. Rev. 88, 848–881.

Glazer, A., Hassin, R., 1987. Equilibrium arrivals in queues with bulk service at scheduled times. Transp. Sci. 21, 273–278.

Hall, R.W., 1991. Queueing Methods for Services and Manufacturing. Prentice Hall, Englewood Cliffs.

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

Medhi, J., 1975. Waiting time distribution in a Poisson queue with a general bulk service rule. Management Science 21, 775–782.

Ockenfels, A., Selten, R., 2005. Impulse balance equilibrium and feedback in first price auctions. Games Econ. Behav. 51, 155–170.

Papadimitriu, C., 2001. Algorithms, games, and the Internet. In: Proceedings of the 33rd Annual ACM Symposium on the Theory of Computing. ACM, New York, pp. 749–753.

Rapoport, A., Seale, D.A., Winter, E., 2000. An experimental study of coordination and learning in iterated two-market entry games. Econ. Theory 16, 661–687.

Rapoport, A., Seale, D.A., Winter, E., 2002. Coordination and learning behavior in large groups with asymmetric players. Games Econ. Behav. 39, 111–136.

Rapoport, A., Stein, W.E., Parco, J.E., Seale, D.A., 2004. Strategic play in single server queues with endogenously determined arrival times. J. Econ. Behav. Organ. 55, 67–91.

Saaty, T.L., 1961. Elements of Queueing Theory with Applications. McGraw-Hill, New York.

Seale, D.A., Rapoport, A., 2000. Elicitation of strategy profiles in large group coordination games. Exper. Econ. 3, 153–179.

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. Exper. Econ. 8, 117–144.

Selten, R., Stoecker, R., 1986. End behavior in sequence of finite Prisoner's Dilemma supergames. J. Econ. Behav. Organ. 7, 47–70.

Selten, R., et al., 2004. Experimental investigation of day-to-day route-choice behavior and network simulations of Autobahn traffic in North Rhine-Westphalia. In: Schreckenberg, M., Selten, R. (Eds.), Human Behavior and Traffic Networks. Springer, Berlin, pp. 1–21.

Vickrey, W.S., 1963. Pricing in urban and suburban transport. Amer. Econ. Rev. 53, 452–465.

Vickrey, W.S., 1969. Congestion theory and transport investment. Amer. Econ. Rev. 59, 251–261.