

NOTES AND COMMENTS

ON THE OPTIMALITY OF FIRST-COME LAST-SERVED QUEUES<sup>1</sup>

BY REFAEL HASSIN

NAOR [4] was the first to show that individual behavior in queues is not in general socially optimal. A newly arriving customer examines the queue and chooses between joining it and balking. Since this person does not take into account the waiting costs that he imposes on future customers by joining the queue, he may choose to join even in instances when it is socially preferable that he balk. Naor [4] and Yechiali [5] also observed that in systems where service is exponentially distributed and where customers are risk neutral and identical, customers can be motivated to behave optimally by levying a fixed admission toll. A profit maximizing firm will tend to levy a toll that is higher than the optimal one.<sup>2</sup>

Consider a service system with exponential service distribution and identical risk neutral customers. Suppose that the queue organizer is free to assign positions in the queue to the new arrivals. Clearly, the externalities imposed by a newly arrived customer on others are highest if he is assigned the head of the queue. Under the first-come last-served discipline with preemption (FCLS) every incoming customer is placed at the head of the queue, moving back other customers who arrived earlier, including the one who is being served.

Surprisingly, this policy also leads to a socially optimal behavior by the customers; the relevant individual decision is now when *to leave* the queue rather than whether *to join* it. From the memoryless property of the exponential distribution it follows that the distributions of the customers' residual service are independent from the amount of service each of them has already received.<sup>3</sup> Since we assume identical customers, all of the waiting customers have identical distributions of residual service time, as well as waiting costs and benefits obtained from the service. It follows therefore that when a customer decides to balk there will be no other customer behind him. As everybody will be served before the person at the back of the queue, he imposes no externality and will take into account all the social costs, arriving at the socially desired decision.

There is a strategic difficulty associated with a FCLS queue. A person whose service has been preempted is motivated to balk and to re-enter the system, presumably disguised as a new person.<sup>4</sup> This behavior must be prevented administratively and may explain why we do not observe FCLS queues in reality.

The important properties of a FCLS discipline that lead to optimal individual behavior are that the balker is the person at the back of the queue and that he imposes no externality. This property is preserved by every policy where the newly arrived customer is placed anywhere but at the end of the queue. A particularly appealing policy is to assign him, whenever the server is busy, to the position before the last. This solution reduces the customers' incentives to balk and to re-enter as a new arrival since only the customer at the back of the queue will benefit from doing so, and the last customer's incentive to do so is weaker now. There is, however, another strategic difficulty associated with this

<sup>1</sup> This research was supported by the Office of Naval Research Grant ONR-N00014-79-C-0685 at the Center for Research on Organizational Efficiency, Stanford University. The author wishes to thank Kenneth Arrow, Amihai Glazer, and Jon Strand for their helpful comments.

<sup>2</sup> Edelson and Hildebrand [1] proved that the optimal tolls will be imposed also by a profit maximizing firm that can impose a two-part tariff, selling rights to check the queue, with an additional toll if service is rendered.

<sup>3</sup> This is also true for other distributions if service is lost upon preemption. However, in this case the disciplines which are discussed in this note may be highly inefficient.

<sup>4</sup> If waiting "at home" is less costly than waiting "in the system" it may be socially desirable that a customer returns to the system after balking. This question is discussed by Glazer and Hassin [2].

solution. If the customer at the end of the queue balks, the one positioned ahead of him will become the last one, and all future arrivals will be positioned ahead of him. Thus he may find it beneficial to offer the person at the end of the queue a payment so that he does not balk. Such side payments must be prevented to preserve optimal behavior; one way of doing so is to conceal from customers the identities of the others in the queue.

The latter solution has other advantages over FCLS: (i) Preemption may incur some loss of service and the latter solution is associated with fewer preemptions. (ii) Risk averse customers are worse off under a FCLS than under a first-come first-served system, as the first is associated with increasing uncertainty; more customers are continuously served without waiting in the queue, while others wait for a long period and finally balk without being served. Assigning newly arrived customers to the position before the last reduces this uncertainty.

My observations have an important application to the theory of priority queues. It is well known that when customers differ in their service distributions, waiting costs, or service value, social welfare can be increased by proper assignment of priorities. (These ideas are discussed in a survey about the economics of queues by Levhari and Sheshinski [3].) A direct consequence of my discussion is that assignment of priorities is beneficial even if customers are identical. Suppose that priorities are assigned randomly or according to some irrelevant basis. The customer at the end of the queue will usually have low priority, and may expect most future arrivals to be placed ahead of him. This decreases the externalities he imposes on others and increases his motivation to balk.

*Tel-Aviv University*

*Manuscript received May, 1983; revision received April, 1984.*

#### REFERENCES

- [1] EDELSON, N. M., AND D. K. HILDEBRAND: "Congestion Tolls for Poisson Queueing Processes," *Econometrica*, 43(1975), 81-92.
- [2] GLAZER, A., AND R. HASSIN: "Search Among Queues," Technical Report No. 406, IMSS, Stanford University, 1983.
- [3] LEVHARI, D., AND E. SHESHINSKI: "The Economics of Queues: A Brief Survey," in *Essays on Economic Behavior Under Uncertainty*, ed. by M. Balch, D. McFadden, and S. Wo. New York: Elsevier, 1974, pp. 195-212.
- [4] NAOR, P.: "The Regulation of Queue Size by Levying Tolls," *Econometrica*, 37(1969), 15-24.
- [5] YECHIALI, U.: "On Optimal Balking Rule and Toll Charges in the GI/M/1 Queueing Process," *Operations Research*, 19(1971), 349-370.