

ERNEST KOENIGSBERG

**Queue systems with balking : a stochastic model of price discrimination**

*RAIRO. Recherche opérationnelle*, tome 19, n° 3 (1985), p. 209-219

[http://www.numdam.org/item?id=RO\\_1985\\_\\_19\\_3\\_209\\_0](http://www.numdam.org/item?id=RO_1985__19_3_209_0)

© AFCET, 1985, tous droits réservés.

L'accès aux archives de la revue « RAIRO. Recherche opérationnelle » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques  
<http://www.numdam.org/>

## QUEUE SYSTEMS WITH BALKING: A STOCHASTIC MODEL OF PRICE DISCRIMINATION (\*)

by Ernest KÆNIGSBERG <sup>(1)</sup>

---

*Abstract. — Queueing Systems with balking have been used to explain the role of uncertainty and capacity in behavior of a monopoly (De Vany, 1976) and a duopoly (Kænigsberg, 1980). In both these models service is on a FIFO basis and there is no distinction between customers. In this paper customers are distinguished by their impatience, by their expectation of alternative service and by priority classifications which depend on their own utility functions and the loss of customers in each class is determined under simple behavior rules. The utility functions consider both price and the quality of service. The models, assuming profit maximizing behaviors by the firm, result in the usual conditions for price discrimination by the monopolist. While the models hold for a wide range of distributions of service time, only the M/M/1 queue is discussed in this paper.*

**Keywords:** Queue models; Balking; Price Discrimination.

*Résumé. — Les systèmes de files d'attente avec obstruction ont été utilisés pour expliquer le rôle de l'incertitude et de la capacité dans le comportement d'un monopole (De Vany, 1976) et d'un duopole (Kænigsberg, 1980). Dans ces deux modèles le service est du type FIFO, et l'on ne fait aucune distinction entre les clients. Dans le présent article, les clients se différencient par leur impatience, par leur attente d'un autre service, et par des classifications de priorité qui dépendent de leur propres fonctions d'utilité; la perte de clients dans chaque classe est déterminée sous des règles de comportement simples. Les fonctions d'utilité considèrent à la fois le prix et la qualité de service. Ces modèles, supposant que la firme cherche à maximiser son profit, donnent les conditions usuelles pour la discrimination du prix par le monopoliste. Bien que les modèles soient valables pour un large éventail de lois de probabilité du temps de service, seule la file d'attente M/M/1 est examinée dans cet article.*

Queueing systems with balking have been used by De Vany (1976, 1977) as a possible model to explain the role of uncertainty and capacity in the behavior of a monopoly. In the De Vany model, service is on a first come-first served basis with no distinction between customers. Here we introduce classes of customers and several different queue disciplines and examine the resulting loss of customers by the monopolist.

Customers can be distinguished by classes because they value the service differently, because they have different alternatives to the offered service or because they have different cost for waiting time. Because of these factors

---

(\*) Received November 1984.

(<sup>1</sup>) Schools of Business Administration, University of California, Berkeley, CA.

each customer class chooses a "balking" number or, perhaps, a bribe to pay to advance their position in the queue.

The priority ordering of customers can arise from price differences—i. e., a form of price discrimination by the monopolist. Customers who pay a higher price receive better service, i. e., have a shorter waiting time, than those who pay a lower price. We assume that higher priority customers preempt lower priority customers in service and go to the head of the line when a high priority customer is in service. As De Vany pointed out in his paper, customers with a high value of waiting time are willing to pay a premium for better service.

## I. MULTIPLE CUSTOMER CLASSES

We consider queueing systems with  $M$  classes of customers in which a customer with lower class index  $m$  is more impatient than all customers with higher index  $m'$  ( $m = 1, 2, \dots, M$ ). The impatience of class  $i$  is expressed in a balking number  $B_i$ , with  $B_m > B_{m'}$  or a reneging rate  $k_i Y$  (with  $k_m > k_{m'}$ ) or a bribe  $b_i$  (with  $b_m > b_{m'}$ ). We will discuss the balking case in detail and only digress on the other forms of impatience.

Following De Vany (1976) we assume that a customer of priority class  $i$  chooses the balking number  $B_i$  so that his expected gain  $G_n^i$  by remaining in the monopolistic line when he is the  $n$ -th customer is greater than or equal to his expected gain  $G_a^i$  by seeking service from an alternative supplier:

$$G_n^i = R(p_i) - w_i n_i / \mu_i, \quad (1.1)$$

$$G_a^i = R^a(p_i^a) - c - w_i t - w_i (L^a + 1) / \mu^a, \quad (1.2)$$

where  $p^a$ ,  $L^a$  and  $\mu^a$  are the expected price, line length, and capacity at the alternate source,  $R(p_i)$  is the return to the customer,  $w_i$  is the waiting cost per unit time,  $t$  is the time to change sources and  $c$  is the fixed cost of a change.

A customer could select  $B_i$  so that the expected gain when the balking number is  $B_i$  is just greater than the expected gain using an alternate source. If this is done the expected gain using the monopolist is always greater than the expected gain using the alternate source. For the customer  $i$ :

$$B_i \leq (\mu_i / w_i) [R(p_i) - R(p^a) + c + w_i (t + (L^a + 1) / \mu^a)]. \quad (1.3)$$

Customers differ in their return functions  $R_i$ , their unit cost of waiting and the expected cost of changing to an alternate source and perhaps in the characteristics of their alternate source. We expect the following to hold:

$$\partial B_i / \partial p_i < 0, \quad \partial B_i / \partial w_i < 0, \quad \partial B_i / \partial p^a > 0 \quad \text{and} \quad \partial B_i / \partial (L^a / \mu^a) > 0.$$

Even when there is a single price (and a single priority class) customers will select a balking number (or renegeing rate or bribe) according to their gain functions. If there are two groups of customers with distinctly different gain functions each will have a unique balking number  $B_i$  [as given for example, by Equation (1.3), and demand function  $(\lambda_i(p_i))$ .

A related system in which customer departures without service are probabilistic rather than deterministic has been reported by Koenigsberg (1980). Within the problem framework, the probability that a customer will depart at time  $t$  given  $n$  in the queue at time zero is  $(1 - e^{-k(n-1)t})$ . In the queueing literature this departure process is known as "renegeing". (We note that the customer being served cannot leave the system by this process.)

We assume that customers of type  $i$  have a renegeing rate  $k_i$ . The renegeing rate can be examined in economic terms. We could write (for example):

$$k_i \propto \frac{w_i}{(p_a - p_i) + w_i L_a / \mu_a + c}, \tag{1.4}$$

where  $p_a$  is price,  $\mu_a$  the service rate and  $L_a$  the expected line length at the alternate server for class  $i$  customers.  $w_i$  is the unit waiting time cost for class  $i$  customers. The exact form of  $k_i$  is not important, but the sign of the derivatives should be:

$$\partial k_i / \partial w_i > 0, \quad \partial k_i / \partial p_i > 0, \quad \partial k_i / \partial p_a < 0, \quad \partial k_i / \partial (L_a / \mu_a) < 0.$$

We expect the renegeing rate of the more impatient customer to be greater than that of the less impatient ( $k_1 > k_2$ ). This would arise if the unit waiting cost  $w_1 > w_2$  or if  $(p_a - p_1) < (p_a - p_2)$ .

An equivalent class structure can be obtained if a class of entering customers can "buy" relative priority by means of a bribe (Kleinrock, 1957). If a portion  $\alpha$  of customers pay a fixed bribe to achieve priority status [e. g.  $\lambda_1 = \alpha\lambda$ ,  $\lambda_2 = (1 - \alpha)\lambda$ ] and each customer retains its balking numbers, the system will be identical to the ones mentioned previously [see Kleinrock (1976), p. 135ff].

## II. CONSERVATION LAWS

Kleinrock (1976) pointed out that "so long as the queueing discipline selects customers in a way that is independent of their service time (or any measure of their service time) then the distribution of the number in the system will be invariant to the order of service: the same will also be shown to be true for the average waiting time of customers." For single server

systems with Poisson arrivals and any general service time distribution ( $M/G/1$  systems) the general conservation law yields, (without balking):

$$\sum_i \lambda_i W_i^q = \lambda W_0^q / (1 - \rho). \quad (2.1)$$

Here  $W_i^q$  is the average waiting time for a customer of class  $i$ ,  $\lambda = \sum_i \lambda_i$  is the total arrival rate and

$$W_0^q = \lambda x^2 / 2,$$

where  $x^2$  is the variance of the service time.

When balking occurs, the proportion of customers of type  $i$  "lost" depends on the queue discipline, but the unfinished work in the system during any busy period  $U(t)$  is independent of the order of service, where:

$$E(U(t)) = U = N_0^q + \sum p_i W_{ii}^q. \quad (2.2)$$

Thus, given a set arrival rates  $\lambda_i$  and of balking numbers,  $B_i$  the total awaiting time (over all classes  $i$ ) will be independent of the queue discipline.

The conservation law holds for a wide range of conservative systems (such as  $G/G/1$ ,  $M/M/c$ , etc.). Here we concentrate on systems of the type ( $M/M/1$ ) and priority on a pre-emptive non-resume basis.

### III. CUSTOMER DECISIONS

In these models the queue discipline system is established by the firm when it selects the prices and the priority rules. Given the prices for each priority class the customers compute their balking numbers by using, say, Equation (1.3) for each priority class: the selected class for a customer group would be the highest class (with a positive  $B_i$ ) for which the expected gain exceeds the gain from the alternate source.

By changing prices for the customer classes the monopolist changes the customers' balking numbers (or renegeing rates) and, in effect, changes the arrival rates for the several classes.

Even when there is a single price (and a single priority class) customers will select a balking number (or renegeing rate). If there are two groups of customers with distinct gain functions, each will have a distinct balking number  $B_i$  [as given, for example, by Equation (1.3)], and demand function. The customer decision is the value of  $B_i$  or  $k_i$  or the magnitude of the bribe paid to improve the position in the queue.

**(a) Priority Systems**

We consider a queueing system with  $M$  classes of customers where a customer with lower index  $m$  has priority for service over all customers with higher index  $m'$  ( $m = 1, 2, \dots, M$ ). If a customer with lower index  $m$  arrives when a customer with a higher index is being served, the higher index customer leaves the service center without completing service and rejoins the queue. The "preempted" customer begins service *anew* after all higher priority customers who arrive before he reenters service are served. The priority system is thus "pre-emptive non-resume". Customers of class  $i$  have a balking number  $B_i$ .

Thus the system is a single server pre-emptive non-resume priority system with balking. Given the fact that balking occurs, the firm providing the service will lose some customers of each class and hence lose revenue. We are interested in determining for a given set of values  $(\lambda_m, \mu_m, B_m)$  what numbers  $\lambda'_m (< \lambda_m)$  of each class of customers will be served. Latter we will consider ways in which the firm may set its "control" variables  $\lambda_m$  and  $\mu_m$  so that some figure of merit (say total profit) is optimized.

The customers of priority class  $m = 1$ , because they have a pre-emptive priority over all other customers, see a system with arrival rate  $\lambda_1(p_1)$ , service rate  $\mu_1$  and balking number  $B_1$  (and hence system utilization  $\rho_1 = \lambda_1(p_1)/\mu_1$ ). The results for this system are well known [Gross and Harris (1974)]:

$$P_0^1 = (1 - \rho_1)/(1 - \rho_1^{B_1+1}), \tag{3.1}$$

$$L_1 = \bar{n}_1 = \rho_1/(1 - \rho_1) - (B_1 + 1) \rho_1^{B_1+1}/(1 - \rho_1^{B_1+1}), \tag{3.2}$$

$$\lambda'_1 = \lambda_1 (1 - \rho_1^{B_1})/(1 - \rho_1^{B_1+1}), \tag{3.3}$$

$$W_1 = L_1/\lambda'_1 = \rho_1 [(1 - \rho_1^{B_1}) - B_1 \rho_1^{B_1} (1 - \rho_1)]/\lambda (1 - \rho_1^{B_1})(1 - \rho_1), \tag{3.4}$$

where  $P_0^1$  is the probability that a customer of class 1 will have to wait before being served,  $L_1$  is the mean number of class 1 customers in the queue,  $\lambda'_1$  is the number of class 1 customers served per period ( $\lambda'_1 < \lambda_1$ ) and  $W_1$  is the expected system time of class 1 customers. The expected waiting time for class 1 customers is  $W_1^q = W_1 - 1/\mu_1$ .

Class 2 customers see a system which includes customers of classes 1 and 2. The arrival rate of class 1 customers is  $\lambda'_1$  and none of these class 1 customers leave the system. The arrival rate of class 2 customers is  $\lambda_2(p_2)$  and class 2 customers will depart if the line length exceeds  $B_2$ . Now if  $B_1 < B_2$  a customer of class 1 will never be in a priority line of length  $B_2$  and our assumption that only customers of class 2 leave the system is correct at all times. If

$B_1 \geq B_2$  we introduce a small error. Now, if  $\mu_2 = \mu_1 = \mu$  we have a system with arrival rate  ${}_2\lambda = \lambda'_1 + \lambda_2$ , a service rate  $\mu$ ,  ${}_2\rho = {}_2\lambda/\mu$ , and balking number  $B_2$ .

Using the same procedures as for the previous problem we have the same results as Equations (3.1) to (3.4) with  ${}_2\rho$  replacing  $\rho_1$ ,  ${}_2\lambda$  replacing  $\lambda_1$  and  $B_2$  replacing  $B_1$ . Note that this is for customers of both classes. We use a left subscript  $m$  (for class  $m$ ) for systems including all customer classes up to and including class  $m$ .

We have the conditions:

$$\lambda'_m = {}_m\lambda' - ({}_{m-1})\lambda', \quad (3.5)$$

$$L_m = {}_mL - ({}_{m-1})L, \quad (3.6)$$

$$W_m = ({}_mW_m \lambda' - ({}_{m-1})W_{(m-1)} \lambda')/\lambda'_m, \quad (3.7)$$

where the right subscript  $m$  indicates the values for class  $m$  alone. Using these conditions and computing priority class by priority class we have:

$${}_m\lambda = ({}_{m-1})\lambda' + \lambda_m, \quad (3.8)$$

$${}_mP_0 = (1 - {}_m\rho)/(1 - {}_m\rho^{B_m+1}), \quad (3.9)$$

$${}_mL = {}_m\rho/(1 - {}_m\rho) + (B_m + 1) {}_m\rho^{B_m+1}/(1 - {}_m\rho^{B_m+1}), \quad (3.10)$$

$${}_m\lambda' = {}_m\rho(1 - {}_m\rho^{B_m})/(1 - {}_m\rho^{B_m+1}), \quad (3.11)$$

$${}_mW = {}_mL/{}_m\lambda'. \quad (3.12)$$

Similar results are obtained for the related reneging system with reneging rates  $k_m$ . In this pre-emptive non-resume system, as in the balking system just described, a high priority customer sees a system which serves only class 1 arrivals. Using the results given by Koenigsberg (1980) we have:

$$P_0^1 = [{}_1F_1(1, b; y_1)]^{-1}, \quad (3.13)$$

$$L_1 = P_0^1 \rho_1 [{}_1F_1(2, b+1; y_1)] \quad (3.14)$$

$$\lambda'_1 = \lambda_1 (1 - P_0^1)/\rho_1, \quad (3.15)$$

where:

$$\rho_1 = \lambda_1/\mu_1, \quad b = \mu_1/k_1, \quad y = \lambda_1/k_1,$$

${}_1F_1(a, b; y)$  is the Confluent Hypergeometric Function (Abromowitz and Stegun, 1965) given by:

$${}_1F_1(a, b; y) = \sum_{n=0}^{\infty} \frac{\Gamma(a+n) \Gamma(b)}{\Gamma(a) \Gamma(b+n)} (y^n/n!), \quad (3.16)$$

and  $\Gamma(z)$  is the Gamma Function given by:

$$\Gamma(z+1) = \int_0^\infty u^z e^{-u} du = z \Gamma(z); \tag{3.17}$$

when  $z$  is an integer

$$\Gamma(z+1) = z! = z \Gamma(z). \tag{3.18}$$

Customers of class 2 see a system which includes customers of classes 1 and 2, just as in the balking case. The arrival rate of class 1 customers is  $\lambda'_1$  and no class 1 customers leave this "equivalent" system. The arrival rate of class 2 customers is  $\lambda_2$  and class 2 customers have a reneging rate  $k_2$ . Thus we have a system with arrival rate  ${}_2\lambda = \lambda'_1 + \lambda_2$ , a service rate  $\mu$ ,  ${}_2\rho = {}_2\lambda/\mu$  and reneging rate  $k_2$ .

Then:

$${}_2P_0 = [{}_1F_1(1, b_2; y_2)]^{-1}, \tag{3.19}$$

$${}_2L = {}_2P_0 {}_2\rho [{}_1F_1(2, b_2 + 1; y_2)], \tag{3.20}$$

$${}_2\lambda' = {}_2\lambda (1 - {}_2P_0) / {}_2\rho. \tag{3.21}$$

Equations (3.5) to (3.7) apply in computing  $P_0^2$ ,  $L_2$ ,  $\lambda'_2$  and  $W_2$ .

In the case of reneging there is no restriction on the relative magnitudes of  $k_1$  and  $k_2$ . Thus we could, for example, consider a system with only a small number of low priority customers who leave quickly when they discover that the waiting line is quite large (such customers represent opportunity traffic, such as "standby" passengers for an airline).

**(b) Non-priority systems**

Consider two classes of customers distinguished by demand rates  $\lambda_i$  ( $\rho_i$ ) and balking numbers  $B_i$  (with  $B_2 \geq B_1$ ) and a single priority class. Service is first-come first-served with an exponential distribution of service times. For this system we have:

$$P_0 = [(1 - \rho^{B_1+1}) / (1 - \rho) + \rho^{B_1} \rho_2 (1 - \rho^{B_2 - B_1}) / (1 - \rho_2)]^{-1}, \tag{3.22}$$

$$\lambda'_1 = \lambda_1 (1 - \rho^{B_1}) / (1 - \rho), \tag{3.23}$$

$$\lambda'_2 = \lambda_2 (1 - P_0 \rho^{B_1} \rho_2^{B_2 - B_1}), \tag{3.24}$$

$$L(< B_1) = P_0 [(\rho - B_1 \rho^{B_1} + (B_1 - 1) \rho^{B_1+1}) / (1 - \rho)^2], \tag{3.25}$$

$$L(\geq B_1) = P_0 \rho^{B_1} B_1 - (B_1 - 1) \rho_2 - (B_2 + 1) \rho_2^{B_2 - B_1 + 1} B_2 + \rho_2^{B_2 - B_1 + 2} / (\lambda - \rho_2), \tag{3.26}$$



$$L_2 = L(\leq B_1) + L(> B_1), \quad (3.27)$$

$$W_2 = (L_2/\lambda'_2) - 1/\mu, \quad (3.28)$$

where  $\rho = (\lambda_1 + \lambda_2)/\mu$ ,  $L(< B)$  is the mean number of customers in the system when there are fewer than  $B_1$  customers in the system (of both classes) and  $L(\geq B_1)$  the same when there are  $B_1$  or more customers in the system.

Table shows the results for the two systems (priority and non-priority) for the same values of  $\lambda_i$ ,  $B_i$ , and  $\mu$ . When customers of class 1 have priority they have shorter waiting times and a higher retention than under the non-priority system. Customers of class 2 show the reverse behaviour. We can also note that the total number of customers ( $\lambda' = \lambda'_1 + \lambda'_2$ ) is different. By imposing a priority system the firm retains more class 1 customers at the expense of losing a slightly larger portion of class 2 customers. The net balance is positive; i. e. the total number of customers retained is larger.

TABLE  
 $\lambda_1 = 1/3$ ,  $\lambda_2 = 2/3$ ;  $B_1 = 2$ ,  $B_2 = 6$ ;  $\mu = 3/2$

	Priority System	Non priority System
$\lambda'_1$	.320,4	.226,5
$\lambda'_2$	.637,8	.662,0
$\lambda' = \lambda'_1 + \lambda'_2$	.958,2	.888,5
$L_1$	.252,4	.447,9
$L_2$	1.276,7	1.140,9
$W_1^q$	.121,1	1.310,8
$W_2^q$	1.264,9	.828,7
$\lambda'_1 W_1^q + \lambda'_2 W_2^q$	.845,5	.845,5

#### IV. ECONOMIC ANALYSIS

The firm's objective is to maximize its profit. The expected profit can be written as

$$E(\pi) = p_1 \lambda'_1 + p_2 \lambda'_2 - C(\lambda'_1, \lambda'_2, \mu). \quad (4.1)$$

We can note, that in the balking case with priorities:

$$\lambda'_1 = \lambda'_1(p_1, \mu, B_1) = \lambda_1(1 - \beta_1), \quad (4.2)$$

$$\lambda'_2 = \lambda'_2(p_2, \mu, B_2; p_1, B_1) = \lambda_2(1 - \beta_2), \quad (4.3)$$

where  $\beta_i$  is the fraction of customers of class  $i$  who leave the system without being served.

**(a) With priorities**

In the short run the firm can only control price and the optimizing conditions obtained from Equation (4.1) are:

$$\frac{\partial E(\pi)}{\partial p_1} = \lambda'_1 + \left( p_1 - \frac{\partial C}{\partial \lambda'_1} \right) \frac{\partial \lambda'_1}{\partial p_1} + \left( p_2 - \frac{\partial C}{\partial \lambda'_2} \right) \frac{\partial \lambda'_2}{\partial p_1} = 0, \quad (4.4)$$

$$\frac{\partial E(\pi)}{\partial p_2} = \lambda'_2 + \left( p_2 - \frac{\partial C}{\partial \lambda'_2} \right) \frac{\partial \lambda'_2}{\partial p_2} = 0. \quad (4.5)$$

From Equation (3.7) we obtain:

$$MC_2 = \frac{\partial C}{\partial \lambda'_2} = p_2 \left( 1 + \frac{1}{\epsilon'_2} \right), \quad (4.6)$$

where  $\epsilon'_2 = (\partial \lambda'_2 / \partial p_2) / (\lambda'_2 / p_2)$  is the price elasticity of the resulting demand (as indicated by actual sales) <sup>(2)</sup>. We substitute in Equation (4.4) and note that under our assumptions:

$$\partial C / \partial \lambda'_2 = \partial C / \partial \lambda'_1,$$

thus we have:

$$p_2 \left( 1 + \frac{1}{\epsilon'_2} \right) = p_1 \left( 1 + \frac{1}{\epsilon'_1} - \frac{p_2 \lambda'_2 \epsilon'_{21}}{p_1 \lambda'_1 \epsilon'_1 \epsilon'_2} \right) = \partial C / \partial \lambda'_1, \quad (4.7)$$

where  $\epsilon'_{21} = (\partial \lambda'_2 / \partial p_1) / (\lambda'_2 / p_1)$  is the cross-elasticity of demand. Because  $\epsilon'_{21} > 0$  and  $\epsilon_1 < 0$ ,  $\epsilon_2 < 0$  (in fact  $\epsilon_1 < -1$ ,  $\epsilon_2 < -1$ ) the term including cross-elasticity is positive.

The exact form of the price solution will differ somewhat between the balking and renegeing cases (in the form of  $\lambda'_i$  and  $\epsilon'_i$ ), but a short-run solution exists. The extension to more than two priority classes is straight-forward but may present computational difficulties, particularly for the renegeing case.

---

(2) We note that because  $\lambda'_i = \lambda_i (1 - \beta_i)$ :

$$\begin{aligned} \epsilon'_i &= (\partial \lambda'_i / \partial p_i) / (\lambda'_i / p_i) = p_i (\partial \beta_i / \partial p_i) / (1 - \beta_i), \\ \epsilon'_i &= \epsilon_i - p_i (\partial \beta_i / \partial p_i) / (1 - \beta_i), \end{aligned}$$

where  $\epsilon_i$  is the price elasticity of the initial demand (orders placed). Since  $\partial \beta_i / \partial p_i < 0$ , final demand is less elastic than initial demand.

In the long run, the firm can adjust its capacity  $\mu$ . Thus we have the additional condition:

$$\frac{\partial E(\pi)}{\partial \mu} = \left( p_1 - \frac{\partial C}{\partial \lambda'_1} \right) \frac{\partial \lambda'_1}{\partial \mu} + \left( p_2 - \frac{\partial C}{\partial \lambda'_2} \right) \frac{\partial \lambda'_2}{\partial \mu} - \frac{\partial C}{\partial \mu} = 0. \quad (4.8)$$

Again,  $\partial C / \partial \lambda'_1 = \partial C / \partial \lambda'_2$ .

The short-run equations and solutions show that the priority problem is equivalent to the problem of price discrimination by a monopolist. The high priority customer pays a higher price to obtain a higher quality of service; i. e., a shorter expected waiting time. We would normally expect the demand function of the higher priority customer to be less elastic than that of the lower priority customer;  $\varepsilon_1 > \varepsilon_2 > \varepsilon_3 \dots$ . In the price discrimination case, the prices  $p_i$  and quantities  $\lambda'_i$  are set so that the marginal revenue is the same in all markets and the marginal revenue is equal to the marginal cost. That is indeed the case here.

Because the term containing cross-elasticity in Equation (4.7) is positive, we will find in all cases that  $p_1 > p_2$  if  $\varepsilon'_1 > \varepsilon'_2$ . The normal price discrimination results hold—i. e., the price is lower in the more elastic market. In this case the monopolist justifies the price differential by providing a higher standard of service to those willing to pay a higher price.

### (b) No priorities

In this case Equations (4.4) holds but Equation (4.5) becomes:

$$\frac{\partial E(\pi)}{\partial p_2} = \lambda'_2 + \left( p_1 - \frac{\partial C}{\partial \lambda'_1} \right) \frac{\partial \lambda'_1}{\partial p_2} + \left( p_2 - \frac{\partial C}{\partial \lambda'_2} \right) \frac{\partial \lambda'_2}{\partial p_2} = 0. \quad (4.5A)$$

Writing  $a_i = (\partial C / \partial \lambda'_i) = MC_i$  and  $\partial \lambda'_i / \partial p_j = b_{ij}$  we have, when  $MC_1 = MC_2$ :

$$MC_1 = p_1 + \frac{\lambda'_1 b_{22} - \lambda'_2 b_{12}}{b_{11} b_{22} - b_{21} b_{12}} = p_2 + \frac{\lambda'_2 b_{11} - \lambda'_1 b_{21}}{b_{11} b_{22} - b_{21} b_{12}} = MC_2, \quad (4.9)$$

we note that  $b_{ii} < 0$  and  $b_{ij} > 0$  so the numerator in Equation (4.9) is always negative. The denominator is positive when  $b_{11} b_{22} > b_{21} b_{12}$ . (e. g. only when the product of own elasticities exceeds the product of cross-elasticities).

The equilibrium equations for profit maximization hold for the priority, non-priority and bribing systems. The results (i. e. the selection of the prices and production rates) will depend on the form of the demand function.

All of the models present an alternate way of depicting the interaction between buyers and sellers. They introduce factors other than price. In

particular, the utility function of the buyer is composed of preferences with respect to both price and quality of service.

## REFERENCES

1. M. ABROMOWITZ and I. A. STEGUN, *Handbook of Mathematical Functions*, New York, Dover, 1965.
2. A. DE VANY, *Uncertainty, Waiting Time, and Capacity Utilization: A Stochastic Theory of Product Quality*, J. Polit. Econ., Vol. 84, June, 1976, pp. 523-541.
3. A. DE VANY and T. R. SAVING, *Product Quality, Uncertainty and Regulation: The Trucking Industry*, Amer. Econ. Rev., Vol. 67, September, 1977, pp. 583-594.
4. See, for example, D. GROSS and C. M. HARRIS, *Fundamentals of Queueing Theory*, New York, John Wiley and Sons, 1974.
5. L. KLEINROCK, *Optimum Bribing for Queue Position*, Operat. Res., Vol. 15, January/February 1967, pp. 304-318.
6. L. KLEINROCK, *Queueing Systems Vol. II: Computer Applications*, New York, Wiley-Interscience, 1976.
7. E. KÆNIGSBERG, *Uncertainty, Capacity and Market Share in Oligopoly: A Stochastic Theory of Product Quality*, J. Business, The University of Chicago, 1980, Vol. 53, No. 2, pp. 151-164).