NBER WORKING PAPERS SERIES

THE APPOINTMENT-BOOK PROBLEM AND COMMITMENT, WITH APPLICATIONS TO REFEREEING AND MEDICINE

Daniel S. Hamermesh

Working Paper No. 3928

National Bureau of Economic Research
1050 Massachusetts Avenue
Cambridge, MA 02138
December 1991

Jeff Biddle, George Borjas, Peter Schmidt, Rachel Willis, and participants at seminars at several universities in Australia, Italy and the U.S., and at the National Bureau of Economic Research, offered helpful discussions of the ideas. I am deeply grateful to the editors and administrative assistants who kept and provided the records that made most of the empirical work possible. The data are available from the author on diskette in a LOTUS work file. This paper is part of NBER's research program in Labor Studies. Any opinions expressed are those of the author and not those of the National Bureau of Economic Research.

NBER Working Paper #3928
December 1991

THE APPOINTMENT-BOOK PROBLEM AND COMMITMENT, WITH APPLICATIONS TO REFEREEING AND MEDICINE

ABSTRACT

Markets that involve customers waiting for services or goods in queues whose length they cannot observe are studied. In these markets suppliers truncate queues that become so long that they jeopardize the supplier's future relations with the customer. The length of the queue and the probability of truncation increase with the quality of the supplier, and this implicitly defines the price that customers are willing to pay for quality. Queue-jumping or nontruncation can occur if monetary payments are made or if nonmonetary specific commitments exist between a customer and a supplier. The predictions apply to any activity where the queue is unobservable and transactions costs make contracts or spot pricing uneconomic.

The theory is examined on a random sample of refereeing requests by seven economics journals. Quality, measured by experience and citations to the referee's work, lengthens the queue and increases the probability of truncation. Monetary bribes affect queue discipline in the expected way; and specific commitments, measured by past publication in the journal and location at the editor's institution, greatly affect the truncation rate, but have no impact on the rate of servicing the queue. The implications for truncation are also examined on a set of data describing doctors' willingness to accept new patients, with much the same results as in the sample of referees.

Daniel S. Hamermesh
Department of Economics
Michigan State University
East Lansing, MI 48824
and NBER

Who more busy than he that hath least to do? (Thomas Draxe, 1633)

It's [Toots Shor's restaurant] so busy that nobody goes there any more. (Attributed to Yogi Berra)

I. Introduction and Motivation

A wide variety of economic activity consists of services offered by individuals or firms who are unable to ration completely using prices. In each case the supplier is approached by a customer demanding the service at a fixed price that may be below that which clears the spot market. The purpose of the analysis here is to examine the incentives facing suppliers, to infer how differences in their characteristics affect how they treat demand, and to study how their relations to their customers affect their behavior.

There has been some study of markets where non-price rationing of goods by the value of customers' waiting time takes place. Deacon and Sonstelie (1985) analyzed how customers choose whether to queue to purchase low-priced gasoline; Lindsay and Feigenbaum (1984) studied the case of the British National Health Service, and Nichols et al (1971) examined the demand for publicly-provided services. In those cases and numerous others, including waiting for bank tellers, obtaining a table at a crowded restaurant, etc., the market allocates the scarce good or service partly along the dimension of the value of the time of actual and potential consumers. Consumers choose whether or not to enter the queue by comparing their desire for the service to its observable full price that includes its time cost.

In a variety of other markets that have not been studied the nature of the available information prevents the customer from estimating the time cost of the service. This gives suppliers the ability to allocate places in the queue for the service. Moreover, it offers them a second margin of choice, whether to deny customers access to the queue, in addition to the choice about the rate of service that they have when the customer can observe the time cost.

This phenomenon, the appointment-book problem, characterizes behavior in a fairly broad range of markets. Physicians in private practice ration their services to potential new customers by imposing long waits for health examinations. They may also truncate the queue by having callers informed that they are not accepting new patients. Senior attorneys perform work for some new clients, but because their hourly billing rates are hard to change, the market cannot ration their scarce time fully. The time to completion of a fairly routine legal transaction by a senior attorney may be much greater than by a junior lawyer. The senior attorney may also assign the task to a junior colleague or refer the client elsewhere.

Markets outside professional services also provide examples of waiting and truncation. Very high-quality hair-dressers and barbers schedule appointments with longer lead times than their less well-known fellows. Airlines maintain no-charge (to the caller) telephone lines for making inquiries about reservations or flights: The inquiry may be served immediately; the caller may be put on hold and told to wait for the next available agent; or a busy signal may be heard.[1] This last response is equivalent to truncation from the queue.

The economic questions of interest in this phenomenon are what determine differences among suppliers in the length of the queue they maintain and the rate at which they truncate the queue. To what extent does suppliers' <u>specific commitment</u> to their customers affect which ones are allowed into the queue? To what extent does it determine the waiting time? How are these affected by the inherent quality of the supplier?

I focus here on what are in some sense small questions. The larger questions --- why suppliers fail to ration fully by price, and why, given that failure, they ever cut off the queue --- are not answered. Why not let anyone who wishes remain in the queue, thus reducing the fraction of periods when the queue is empty, and thus when the service is not utilized? This may reflect a long-run profit-maximizing desire to avoid antagonizing current or potential clients; or it may be explained by suppliers' commitment to their profession/skill.[2]

In the next Section I examine the behavior of a utility-maximizing supplier of services who has two margins of choice by which to affect the queue. Of particular interest is the derivation of the effects of supplier's quality on the expected length of the queue (or waiting time) and the fraction of customers

---

1. Frances Hamermesh and Paul Chen provided some of these examples.

2. This is the kind of altruistic commitment discussed in Frank (1988).

denied service. Section III describes a new set of data, on the behavior of referees for a group of economics journals, that was collected especially for this project. Section IV examines how well those data are described by the predictions of the appointment-book model, and Section V discusses queue-jumping and examines the implicit market for quality. Section VI studies the truncation issue on yet another set of data collected for this study, this one describing physicians in a pre-paid health plan. The conclusion indicates some extensions and considers the larger question of why truncation arises.

II. The Appointment-Book Model of Rationing

Assume the supplier's time is divided between providing the service and all other activities. Each time period has T units, and the supplier's expected utility is described by the function:

(1) $U = U(T - \mu, \mu)$,

where μ is the average number of units in the time period devoted to providing the service, and $U_i > 0$, $U_{ii} < 0$.[3] I assume throughout this Section that each customer requires the same amount of time from the supplier. Clearly, a more general model could expand this to make the quality of service, measured by the amount of time devoted to a customer, subject to choice. With this simple utility

3. The model is couched in terms of utility instead of profit maximization. This seems more consistent with the importance of time use and the nature of many of the examples of professional services that motivate the paper. Nonetheless, if one assumes suppliers are price-takers, the model can be revised mutatis mutandis to yield similar results based on profit maximization so long as there is a constraint on the supplier's capacity to serve clients.

function the supplier just sets the marginal rate of substitution between the two uses of time equal to one.

The rate of arrival of customers is exponential and is described by the paramter λ. The rate depends on the quality of the supplier, Z, so that:

(2) $\lambda = \lambda(Z)$, $\lambda' > 0$,

and by assumption $\lambda < \mu$.[4] In this simplest model λ has no effect on the expected service rate, since the latter is chosen to maximize (1). Without any denial of entry to the queue, L^*, the average waiting time in the queue, is:

$$L^* = \lambda(Z)/[\mu - \lambda(Z)].$$

Even a very simple model generates the prediction that the average waiting time is an increasing function of quality.

A model with exogenous <u>general commitment</u> requires that all suppliers have (the same) commitment to deny entry into the queue to those customers whom they cannot expect to serve in less than t^* time periods. With this cut-off we obtain the customer's average waiting time:

(3) $L^* = a\int_0^{t^*} t\exp(-at)dt/[1 - \exp(-at^*)]$,

and the supplier's mean rate of refusal:

(4) $\rho^* = \exp(-at^*)$,

4. Implicit in (2) is the assumption that all suppliers are constrained to charge the same (below-market) price. The model could be modified to allow price to be correlated with Z, so long as the negative effect of Z on λ through price is less than its direct positive effect.

where $a = [\mu - \lambda]/\mu$.[5] Since μ does not change with changes in λ, while higher Z increases λ (though not by enough so that it exceeds μ), it follows immediately that higher-quality suppliers have higher L^* and ρ^*.

What if the degree of general commitment is subject to suppliers' choice, and:

(1') $U = U(T - \mu, \mu, t^*)$,

with $U_3 < 0$ to capture the notion that suppliers' utility or profits decrease with the maximum time a customer waits in the queue?[6] Together with the arrival rate $\lambda(Z)$, this assumption generates an optimizing set of values of μ, t^*, L and ρ. Specifying general commitment in (1'), though, makes the optimizing μ and t^*, and the resulting L^* and ρ^*, functions of Z. Without any <u>a priori</u> restrictions one cannot guarantee that higher Z continues to increase L^* and ρ^*, as in the model with t^* exogenous and the resulting independence of μ from the arrival rate. The effects of Z on L^* and ρ^*, though, will always be in the same direction, given the assumptions. If, as seems reasonable, the effects of higher Z on μ and t^* are sufficiently smaller than the direct positive effect on ρ, we will continue to observe in the steady state that higher-quality suppliers exhibit longer waiting times and higher refusal rates.

---

5. The derivations are based on Karlin and Taylor (1975).

6. (1') could be written as a function of L instead of t^* with the same resulting ambiguous conclusions.

One might assume further that suppliers' commitment means that their utility is affected differently by changes in L and ρ that result from altering t^*. If so, the standard model of exponential arrival rates and fixed service times no longer applies. No formal steady-state queuing result is possible. The response depends on the extent of asymmetry in suppliers' attitudes about refusal and waiting time.

I have said nothing thus far about specific commitment --- who gets served. The commitment not to make customers wait too long has been general. Yet with the ability to ration customers, suppliers can discriminate in favor of those whose characteristics they find more desirable. For a given level of quality, if specific commitment exists we should find that prior supplier-customer ties will reduce the customer's waiting time and probability of denial of entry to the queue.[7] The problem is similar to that noted by Wilson (1989) in the demand for electric power, except that here the commitment is not forged by explicitly priced contracts, but instead by nonpriced arrangements.

This discussion does not allow for the possibility that consumers might circumvent the below-market price and/or specific commitment by bribing the supplier to enter a closed queue or to jump position in an existing queue. It is difficult to draw many inferences about queue-jumping. For a fixed bribe, though, we should expect that the amount of queue-jumping that a supplier

---

7. A simple example is airlines' practice of offering preferred (frequent) customers special toll-free numbers for booking tickets or obtaining flight information.

allows will be the minimum consistent with obtaining the bribe. Larger bribes will generate larger jumps ahead in the queue and larger drops in the refusal rate. For a smaller bribe one should expect smaller changes.

I have assumed that the arrival rate is independent of L^* and ρ^* and depends only on the supplier's quality. What if, though, customers recognize that higher-quality suppliers will generally have longer queues and higher truncation rates? The utility-maximizing supplier then chooses μ and t^* taking account of customers' reactions. In the steady state this generates an equilibrium in which differences in L^* and ρ^* reflect both suppliers' and customers' behavior and imply a market price of quality. The gradients of L^* and ρ^* with respect to quality will be flatter than they would be if customers did not account for suppliers' responses to differential arrival rates.

III. Refereeing as an Appointment-Book Problem

The main specific empirical example that I present is that of refereeing scholarly papers for academic journals (in economics). This activity clearly fits the appointment-book problem described above. In most cases suppliers are not paid for their services, so that any rationing of their time must be done by non-price methods. Customers (journal editors) cannot observe the length of referees' queues. The referees can be viewed as members of a community (of professional economists), so that there may be general commitment to the community (to the profession itself). The degree of specific commitment may vary depending on which journal requests the referee's services. Finally, there are objective measures of quality among professional economists that lend themselves to representing Z.

There has been a huge amount of research by sociologists, and more recently by economists, on the refereeing process. Most of that study has been of the fairness and quality of the reviews. Studies involving correlations between referees' opinions on the same article or proposal and the re-refereeing of articles and proposals (Cole et al, 1978; Peters and Ceci, 1982) have been designed to determine whether the refereeing process can distinguish quality among submissions. Blank (1991) and many earlier studies (e.g., Crane, 1967) have examined whether refereeing outcomes differ between single- and double-blind trials; and Laband (1990) studied the productivity of refereeing in terms of subsequent citations to the published article.

There has been very little study of the refereeing activity itself. Evidence from physics (Zuckerman and Merton, 1971) indicates that referees tend to be of higher quality on average than the authors whose papers they read; but no research has been conductedinto the allocation of time by referees. In addition to examining a novel and fairly broad area of economic behavior, the evidence here and in the next two Sections thus fills a gap in the study of the organization of scholarly publishing.

There are no secondary data on referees' time use. I therefore asked editors of eleven journals to participate in the following exercise. Based on a tabulation sheet (Appendix Table

A1), the editorial office was to keep a record of a very few characteristics of referee reports resulting from the next fifty requests on initial submissions of articles that were sent out. The information allows me to determine L_i^* for each refereeing request i, and whether $t_i \geq t^*$ for the particular request.[8] The length of the paper, a measure of heterogeneity among the referee's "customers," was also obtained, as was the referee's name.

Seven editors agreed to participate in the project, which began in November 1989. Their journals include four that publish articles in a variety of subspecialties (general journals, G1 through G4), and three that publish in only one subspecialty each (specialized journals, S1 through S3). I rank them in descending order according to the rate at which articles in them are cited.[9] By November 1990 all seven journals had returned the recording forms. Of the 350 possible data points, 343 were usable: On three the editor did not seek a report; one paper was withdrawn by its author; one form was blank; and on one no page-length was recorded. For one other request I was unable to construct one of the crucial variables in the vector Z. Given the nature of the sampling procedure and the nearly complete sample, the data set is a random sample of refereeing at these seven journals.

---

8. L_i^* was calculated as the actual number of days elapsed from the time the manuscript left the editorial office to the time it was received back, minus seven days if the Christmas-New Year's holiday intervened.

9. From Leibowitz and Palmer (1984). I have designated the one journal that was too new to be rated in that study as S1, based on its immediate success in attracting attention from the profession generally.

The editorial records were linked to a variety of indicators that might describe referees' quality. These included X, years of experience since the Ph.D., based in most cases on self-reported information in the American Economic Association Membership Directory, 1989.[10] This standard measure of productivity should proxy the same thing among these professionals. The names were also linked to the Social Science Citation Index for 1989, and each referee's citations by other scholars in 1989 were included in the data set. This measure of quality seems to be a more important determinant of one outcome of quality differences among economists --- dispersion in salaries --- than either counts of publications or the status of the outlets of one's research (Hamermesh et al, 1982). The density of citations is highly skewed; thus in subsequent discussion I report this measure in dummy-variable form, with CITS10-49 equalling one if the referee was cited 10 through 49 times in 1989 (was well-cited), and CITS50 equalling one if he or she received 50 or more citations (was heavily-cited). A final possible set of proxies for Z is a vector of dummy variables indicating the referee's employment. It includes: Top 20, if the referee's school was rated among the twenty best graduate economics departments in Boddy (1982); Cartel, if the school is included among the roughly one hundred that exchange information on salaries and employment; and Other academic, if not. (The excluded category is non-academic employment.)

10. For those individuals not listed in the Directory, indexes of other organizations were consulted or the editors were telephoned to supply the information.

The commitment measures were constructed to reflect the referee's ties to the particular journal or to the profession generally. The former, which can be viewed as measuring specific commitment, include whether the author is at the same school as the journal's editor, and whether he or she published in that journal during the quinquennium 1986-1990. Measures describing possible differences in general commitment to the profession (as represented by these American scholarly journals) are the referee's sex and whether or not the referee is located in North America.

Statistics describing the raw data and the constructed variables are presented in Table 1 for each of three categories of referees: Those who do the job (for whom $t_i < t^*$); those who refuse to referee (for whom $t_i \geq t^*$), and those who "lose" the paper. Seventy-eight percent of those asked to referee comply. Among "refusers" (seventeen percent of the sample) the mean time to refusal is only 23 days (with a median of 17 days, and a seventy-fifth percentile of 27 days), implying that truncation from the queue occurs shortly after entry. The remaining five percent of the sample who are "losers" pose some small difficulty for interpreting the results. This group includes only those from whom the editor requested the paper back or, in some of the seventeen cases, who had not refereed the paper in at least eight months at the time the recording sheets were returned from the editorial offices.[11]

---

11. Whether the paper should be viewed as having been truncated from a very long queue that would eventually have been served, or as in a Kafkaesque predicament without possibility of escape, is unclear. The relatively small proportion of the sample that is in this category means that it does not affect most of the

Many of the results on the determinants of ρ^* can be seen from the differences among the means in Table 1. Doers have less experience than refusers (though more than losers); they are less likely to be well- or heavily-cited than the other groups, and are less likely to be employed at the top 20 schools. Doers are more likely to be employed by the school where the journal's editor works and much more likely to have published in the journal than refusers or losers. Also, women are less likely to refuse to referee, or to lose the paper, than are men. Finally, papers that are refereed are shorter than those that are refused or lost.

Differences among journals are also evident from the means in Appendix Table A2. Most interesting, except that G4 is out of place, higher-quality journals use higher-quality referees. A similar quality gradient exists with respect to the institutions where the referees are employed. The Table also shows huge differences among journals in some of the variables measuring specific commitment: Journals G2, S2 and S3 rely very heavily on their own authors, while G4 obtains much of the refereeing from faculty employed around the editor's office. Other control measures also differ significantly among journals, e.g., the length of submissions to S1 and S2.

IV. The Effects of Quality and Commitment

Before testing the model's predictions about the determinants of L^*, time spent in the queue, and ρ^*, the probability of

---

conclusions; but the issue is not unimportant generally, as there are obvious analogies in other appointment-book problems.

Table 1. Means and Standard Deviations, by Category of Referee*

VARIABLE	CATEGORY		
	DOER	REFUSER	LOSER
Dummy Variables:			
Same School	0.130	0.053	0.118
Published in Journal	0.368	0.140	0.118
Foreign	0.037	0.035	0.059
Male	0.881	0.965	0.941
Top 20	0.301	0.368	0.471
Cartel	0.424	0.351	0.412
Other Academic	0.164	0.140	0.059
CITS10-49	0.368	0.474	0.588
CITS50	0.104	0.211	0.118
Continuous Variables:			
Citations	20.94 (46.69)	29.97 (35.56)	19.71 (19.21)
Ph.D. Experience	16.24 (9.08)	18.28 (8.87)	14.47 (6.76)
Pages	26.40 (10.32)	27.05 (9.03)	31.12 (10.51)
Days	54.98 (46.41)	22.91 (15.49)	----
Number of Observations	269	57	17

*Standard deviations in parentheses below the continuous variables.

truncation from the queue, we need to show that the basic assumption, $\lambda'(Z) > 0$, is correct for this sample. Direct evidence on a sample of 41 economists in one department based on variables like those in the data set supports this assumption. The following regression was estimated for the academic year 1987-88:

$$\text{Referee tasks} = \underset{(1.34)}{2.08} + \underset{(.20)}{.31X} - \underset{(.0052)}{.0095X^2} + \underset{(1.43)}{1.76\text{CITS10-49}} + \underset{(2.92)}{16.98\text{CITS50}}, \quad \bar{R}^2 = .52,$$

where standard errors are in parentheses. The mean number of articles refereed was five; the (two) heavily-cited faculty refereed four times the department's average, while the (eleven) well-cited faculty refereed two more articles than their typical colleague. Perhaps most interesting, there is a significant inverse quadratic in experience, with the peak refereeing at 16 years, not much different from the average experience in the sample (14 years).

A. Effects on Truncation

The results on the effects of the quality and commitment variables on ρ^* are presented in Table 2, columns (1) and (2), which show the parameters of a multinomial logit relating category of response to the variables listed. Well- or heavily-cited researchers are significantly less likely to be doers, other things equal. Additional estimates also included the vector of three variables denoting the referee's affiliation. For none of them did the absolute value of the t-statistic exceed .5. It is the individual referee's achievements that affect ρ^* and thus that

Table 2. Limited Dependent Variable Models, Referee Status*

VARIABLE	Multinomial Logit		Probit	Probit
	CATEGORY			
	DOER	LOSER	DOER	DOER
Same School	1.119 (.630)	0.868 (.985)	0.707 (.302)	0.882 (.364)
Published in Journal	1.435 (.416)	-0.406 (.858)	0.627 (.222)	0.585 (.241)
Foreign	0.490 (.818)	0.900 (1.30)	0.198 (.425)	0.313 (.479)
Male	-1.007 (.767)	-0.459 (1.288)	-0.514 (.348)	-0.559 (.411)
CITS10-49	-0.813 (.353)	0.505 (.695)	-0.472 (.200)	-0.320 (.217)
CITS50	-1.123 (.456)	-0.315 (.993)	-0.464 (.278)	-0.398 (0.297)
Ph.D. Experience	-0.0299 (.055)	0.012 (.147)	-0.0056 (.031)	-0.012 (.034)
Ph.D. Experience^2	0.0006 (.0012)	-0.0022 (.0043)	0.00021 (.0008)	0.00018 (.0007)
Pages	-0.010 (.015)	0.045 (.028)	-0.015 (.0093)	-0.013 (.010)
Journal Dummy Variables	No		Yes	Yes
Log-likelihood	-195.34		-146.11	-121.41
χ^2 df =	46.80 18		65.51 15	59.37 15
χ^2 (against dummies only) df =			27.98 9	22.14 9
N =	343		343	326

*Standard errors in parentheses. Each equation also includes a constant term.

signal quality to the customer. Once these are accounted for, proxies for quality that are based on external factors have no impact. While there is some evidence of the expected quadratic in experience, with additional experience first reducing, then increasing the probability that the job is done, neither term nor the pair jointly is significantly different from zero.

The commitment measures generate fascinating results. Measures of specific commitment sharply affect whether the task is completed. Being at the same school as the editor has the expected positive effect, and having published in the journal significantly increases the likelihood that the task will be done. The two measures of general commitment have insignificant effects on the categorization (although women are somewhat more likely than men to be doers). Finally, there is some evidence of the effect of heterogeneity among customers. Papers that are longer are more likely to be "lost," though length does not significantly affect whether the task is done.

As shown earlier, there are substantial differences among journals in many of the measures; but because none of the "only" 17 losers was at several of the journals, dummy variables for individual journals could not be included in the multinomial logit. The insignificance of the parameters distinguishing losers from refusers suggests that for purposes of examining ρ^* these categories can be combined. Dummy variables for the journals are included in simple probits describing whether the task is done (the probability $1 - \rho^*$).

The results are presented in column (3) for the entire sample, and in column (4) for doers and refusers only. They generally corroborate the estimates of the multinomial logit. The magnitudes of the specific commitment effects are substantial: Referees at the same school as the editor are 12 percentage points more likely, and those who have published in the journal are 11 percentage points more likely than others to referee the paper.[12] (With a mean of $1 - \rho^* = .78$, these effects are very large.) The commitment and quality measures are jointly significant against an equation containing only journal dummy variables.

B. Effects on Waiting Time

Ordinary least squares estimation of the effects of the commitment, quality and other variables on the refereeing lag is the most familiar way to examine the effects on L^*. These estimates are presented in column (1) of Table 3 for the sample of doers only. I do not discuss them here, mainly because they do not reflect the analysis in Section II. In a queuing model the interesting questions involve the determinants of waiting time and survival. These are described by models of hazard rates, where the hazard is $h(t) = -d\ln(S(t)/dt$, the escape rate from the queue (containing at time t the fraction $S(t)$ of the papers that arrived

12. An additional measure of specific commitment, whether the referee was on the journal's editorial board, was also included in some of the estimates. Its quantitative impact on the probability of doing the task was tiny and insignificant, as was its effect in the survival models estimated in the next Subsection. A similarly small and insignificant impact is produced by an interaction term that tests whether the effect of the editor being at the same school is greater for younger referees.

Table 3. The Determinants of L^* (Doers Only)*

VARIABLE	OLS	Proportional Hazards	Weibull Accel. Failure
Same School	8.469	-0.253	-0.169
	(8.53)	(.201)	(.135)
Published in Journal	1.292	-0.078	-0.082
	(6.21)	(.154)	(.114)
Foreign	23.36	-0.527	-0.324
	(14.63)	(.349)	(.328)
Male	1.168	-0.043	-0.063
	(8.72)	(.199)	(.147)
CITS10-49	5.634	-0.071	-0.051
	(6.33)	(.151)	(.107)
CITS50	15.77	-0.281	-0.269
	(9.69)	(.224)	(.132)
Ph.D. Experience	0.64	-0.023	-0.017
	(.92)	(.021)	(.016)
Ph.D. Experience^2	-0.026	0.00093	0.00067
	(.019)	(.00045)	(.00033)
Pages	0.789	-0.020	-0.015
	(.292)	(.0071)	(.005)
$\hat{\gamma}$			0.0171**
			(.0008)
$\hat{\alpha}$			1.528**
			(.091)
$\bar{R}^2$ or Log-Likelihood	0.111	-1220.8	-298.45
$\chi^2(15)$	47.05	43.18	42.24
$\chi^2(9)$ (against dummies)	16.46	19.40	27.50

*Standard errors in parentheses. Dummy variables for the journals are included in all the estimates.
Evaluated at the means of the **x.

at time zero). In this model the hazard rate can be used to infer the waiting time until refereeing. It thus provides an empirical approach that is based in queuing.[13]

The empirical survivor curve, the relation of S(t) to time measured in weeks, is shown for the sample of doers by the ■ marks in Figure 1 (leaving off weeks after 34, to which only five referees "survive"). It is noteworthy that the median waiting time is only 6-1/2 weeks, while the seventy-fifth percentile is 11 weeks.[14] Among the 78 percent who complete the task, the typical referee does so in fairly short order. Figure 1 also graphs the Kaplan-Meier estimates of the hazard rate (denoted by (+)). It is clear that the hazard is not monotonic. Instead, it is increasing up to the point where half the sample of doers has exited, and fluctuating around an essentially constant value thereafter.

The nonmonotonicity of the hazard rate and the absence of a single peak suggest that any simple parametric specification of the baseline hazard will be incorrect. I therefore estimate a proportional hazards model. Column (2) of Table 3 shows estimates of the parameter vector β in:

5) $h(t,\mathbf{x}) = h(0,t)e^{\beta'\mathbf{x}}$,

---

13. For a good discussion of hazard models, see Kiefer (1988). These models have been used by economists mostly to examine the duration of unemployment (most recently, Meyer, 1990). Additional applications have been to strike duration (e.g., Lancaster, 1972) and to job tenure (Dolton, 1991).

14. Given lags in the mails, the actual time that the paper spends in the referee's hands is probably one week less than the Figure indicates.

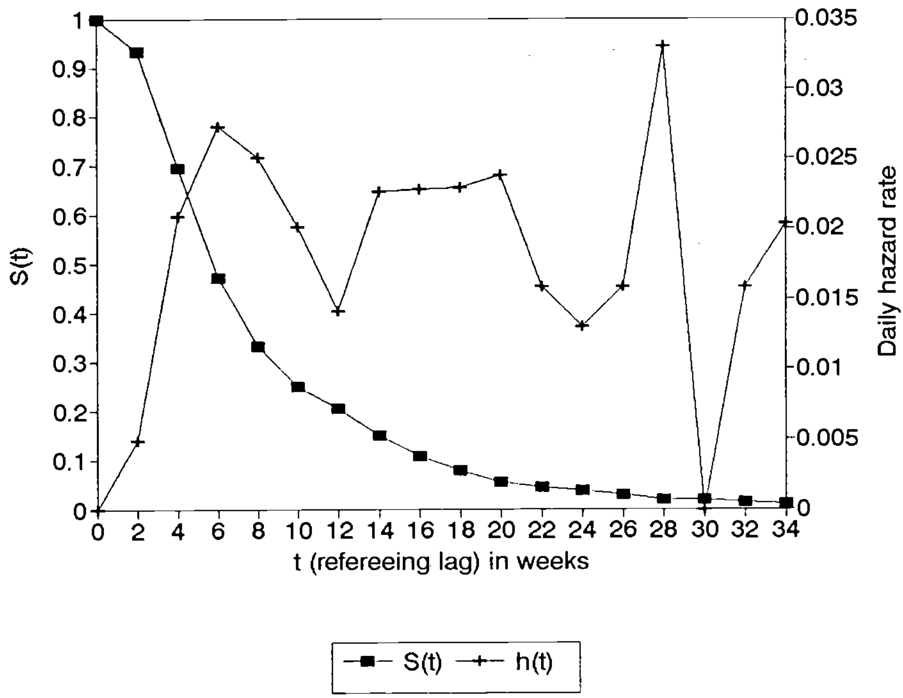

Figure 1. Survivor Function and Daily Hazard Rate, Doers

where **x** is the vector of variables included as determinants of the hazard rates. These estimates assume that each independent variable has the same proportional effect on h regardless of the duration in the queue. For any particular set of values of **x** the hazard rate can vary freely. A negative coefficient implies that increasing the particular variable reduces the hazard rate (increases the waiting time).

Among the quality measures, the hazard rate is lower among heavily-cited referees, and somewhat lower, though quite insignificantly so, among well-cited referees. There is an inverse-quadratic effect of experience on the hazard rate, with the lowest hazard observed for referees with twelve years of experience (somewhat below the sample mean for this variable).[15] As the theory predicted, $\partial L^*/\partial Z > 0$.

The negative coefficients on the two measures of specific commitment contrast sharply to their effects on the refusal rate and cannot be predicted by a simple queuing model. One way of rationalizing them is to note that these commitment measures may also proxy heterogeneity in service times: Where the referee has published in the journal, or works with the editor, the paper may receive more careful attention and thus generate a lower escape rate. An alternative explanation is based on the possible observability of L^* (or at least of the response time to a specific customer) and referees' response to it: By showing a willingness

15. A test of the joint significance of the two experience variables yields $\chi^2(2) = 7.00$; $\chi^2_{.95}(2) = 5.99$. Also, the variables based on the referee's affiliation have no impact ($\chi^2(3) = 2.2$).

to referee where specific commitment exists, but doing so without undue alacrity, the referee slows the arrival rate from those journals.

The general commitment variables also have negative effects on waiting time. Conditional upon completing the task, there is essentially no difference in the hazard between male and female referees, while the hazard rate is lower among referees outside North America. The difference may reflect lack of commitment to the profession in its North American incarnation. Alternatively, it may be the mechanical result of postal delays: The twenty-three day difference implied by the OLS estimate is not far from the round-trip time of an air-mail letter.[16] The proportional hazards model also indicates the importance of heterogeneity in the refereeing queue: The hazard is lower for longer papers. Those that require more service time take longer to move through the queue.[17]

---

16. Another possible test of general commitment might be constructed by comparing the coefficients of the dummy variables for the S and G journals. Other things equal, one might expect greater commitment to the field journals. With only seven journals this is a rather stringent test; and the hazard rates are not significantly lower at the general journals. However, the same test on the coefficients of the dummy variables in the probits in column (3) of Table 2 does suggest a slightly higher ρ^* among the general journals.

17. Including losers as censored as of the date the paper was recalled from them (as of the end of the sample period in some cases) had only tiny effects on the $\hat{\beta}$. Similarly tiny changes were produced if I treated all 17 losers as having refereed as of that date, or if I viewed them as censored as of the longest duration observed. Censoring may be nonrandom; and it clearly affects parametric estimates of the hazard function. It does not, though, influence the estimates of the impact of the quality and commitment variables.

The proportional hazards model is a relatively unrestricted way of examining the effects of $\mathbf{x}$ on the hazard rate of refereeing. As a comparison to a more familiar form, though, column (3) of Table 3 presents estimates of the vector β from a model in which the baseline hazard is assumed to have a Weibull distribution. The estimated $\hat{\beta}$ in (3) are quite close to those of the proportional hazards model and merit little comment. The baseline hazard suggests that there is positive duration-dependence in refereeing: As time passes, the probability that a referee will do the job, <u>conditional on eventually doing it</u>, rises. Once losers are included in the Weibull estimates, though, the positive duration-dependence is not significant. This suggests, and formal tests verify, that the Weibull distribution does not provide a satisfactory fit to this hazard. As Figure 1 shows, there is no uniform duration dependence in this process; indeed, it is not described well by any standard distribution.[18]

A final econometric issue is whether the variables in $\mathbf{x}$ capture referee-specific heterogeneity. If we had two or more observations on each referee, or on many of the referees, we could estimate the hazard function using individual effects. We cannot do that satisfactorily here; nor could the kind of random sampling

18. The Kolmogorov-Smirnov statistic for this parameterization was 1.31, just above the 99-percent critical level for a Weibull distribution with unknown parameters (D'Agostino and Stephens, 1986, p. 147). A Gompertz distribution, a normal and a logistic were also estimated. The first two fit much worse than the Weibull distribution, while for the latter the Kolmogorov-Smirnov statistic had the same significance as that from the Weibull.

that is crucial to this project generate such data. Fifteen referees are, though, included twice in the sample of doers, and one is included three times. A vector of dummy variables for these people was included in reestimates of column (1), with a resulting $F(15, 238) = .87$, insignificantly different from zero at any conventional level. This admittedly very partial evidence suggests that this kind of heterogeneity among suppliers is not a problem in this sample.

One interpretation of the estimates of $\partial\rho^*/\partial Z$ and $\partial L^*/\partial Z$ is that they reflect referees trying to ingratiate themselves with editors by doing the task promptly as an investment in their careers. If so, L^* and ρ^* would increase steadily, though perhaps nonlinearly, with experience. That the quadratics in experience in both Tables 2 and 3 have extrema near the sample mean of X suggests this explanation is incorrect. Another interpretation is based upon sample selection. If editors learn which referees respond slowly (or refuse) and do not request their services, the fraction of rapid respondents among referees included in the sample would increase with experience. That we observe the probability and speed of response falling over the first half of the distribution of referee's experience implies that this interpretation is also wrong.

A more subtle and more serious potential problem arises from the possibility that referees' quality reflects the heterogeneity of customers because editors match papers that require more effort with higher-quality referees. There is no way of examining this effect in these data. However, I obtained another set of data from journal G1, a listing of all the referee-author matches over a roughly one-month period in 1991. A contingency table showing the citation counts for 1990 on each of the 129 matches (of 80 separate submissions to the journal) is presented in Table 4.

These data show that the referees are far more widely-cited than authors. Moreover, χ^2-tests over the entire Table reject the hypothesis that the author-referee matches are random. The rejection, though, is based solely on the very few extremely heavily-cited (more than 100 citations per year) authors. When these 6 matches are deleted, the matching process appears random. Whether the same conclusion would hold if data were available from other journals is unclear. But the higher quality of referees at G1 than at the other journals suggests the scope for matching elsewhere is less than at G1. If other editors behave as at G1, matching is even less important than the data in Table 4 indicate. Heterogeneity in matching does not seem to be important, at least along this dimension.

V. Queue Discipline, and the Price of Quality

A. Monetary Bribes to Change Suppliers' Behavior

Because one of the seven journals offers prompt referees a small monetary incentive, we can infer the <u>relative</u> effect of monetary incentives on waiting time. (Payments of this sort are offered by no more than ten percent of professional journals in economics.) The journal (G1) pays for completion within a nominal

Table 4. Author-Referee Matches by Citation Count

Author	Referee 0-4	5-9	10-49	50-99	100+	TOTAL
0-4	21	3	37	12	6	79
5-9	2	3	4	4	2	15
10-49	3	2	14	4	2	25
50-99	2	0	1	1	0	4
100+	0	1	0	2	3	6
TOTAL	28	9	56	23	13	129

χ^2-statistics

1. $\chi^2(9) = 17.58$; all observations, categories 0-4, 5-9, 10-49, 50+; $p = .04$.

2. $\chi^2(4) = 8.32$; all observations, categories 0-9, 10-49, 50+; $p = .08$.

3. $\chi^2(4) = 2.47$; excludes 6 matches on authors with CITES ≥ 100; $p > .10$.

4. $\chi^2(2) = 1.49$; excludes 10 matches on authors with CITES ≥ 50; $p > .10$.

one month (in actuality, if the completed report is received at the editorial office within six weeks of the date it was sent out).[19]

If the discussion in Section II is correct, we should expect the prospective payment to have its biggest effect on queue discipline at the margin. That is, if the queue is very long, so that $t_i >> 42$ days (six weeks), the incentive should be unimportant. In that case either $t_i \geq t^*$, so that the request is truncated from the queue; or the paper remains in its place in the queue. Obversely, if the queue is so short that the paper will be refereed almost immediately anyway, the incentive will also be unimportant. Only if the queue is such that the paper is on the margin of being eligible to qualify the referee for payment will the bribe induce a change in queue discipline.

Figure 2 shows Kaplan-Meier estimates of the daily hazard rates for G1, and for the aggregate of the other six journals, for completion in 0-14 days, 15-28 days, etc., through 85-98 days.[20] (With only three manuscripts outstanding at journal G1, the hazard beyond 98 days is uninteresting.) Remarkably, the hazard rates are nearly identical in the first two-week interval, and essentially the same in all three biweekly intervals containing weeks 9 through

---

19. Based on conversation with the editor and editorial assistant, December 28, 1990. A small survey of economists suggests that the median time spent refereeing is such that the hourly payment offered by G1 is only slightly above the minimum wage rate in the United States.

20. The hazards for the other six journals in Figure 2 are based on a sample that includes losers as censored observations. If we include only doers, the hazard rate lies even closer to that of G1.

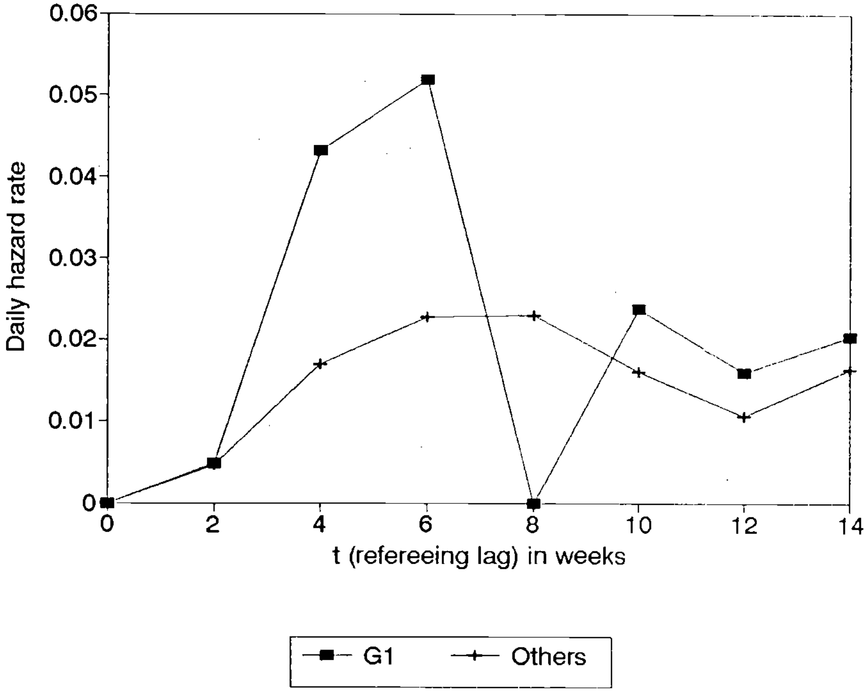

Figure 2. Hazard Rates at Journals G1 and Others

14. The hazard rate at G1 is higher in the second and third biweekly intervals (days 15-42) and lower in the fourth interval (days 43-56). Moreover, these three hazards, and only these, are more than one standard deviation apart, even with the very small sample size at G1.[21] Paying to induce queue-jumping shifts customers (refereeing tasks) just far enough ahead in the queue to qualify the supplier (referee) for the bribe (honorarium).

The comparisons in Figure 2 do not account for the higher quality of referees used by journal G1 than by other journals. Holding quality (and the other variables included in Table 3) constant, the hazard rate at G1 is higher than at every other journal. The effect is not small: Compared to the average of other journals, conditional upon submitting a report the same referee responds two weeks more quickly to G1.

Figure 2 reflects the gross effects of bribing referees. Presumably speed increases partly because referees increase L^* and ρ^* in their other (unpaid) refereeing tasks, partly because they increase μ overall. If all journals paid referees, only this latter, probably small scale effect would be generated. Each journal wishing to bribe referees confronts the problem of setting optimal prices in the face of unknown responses by suppliers of refereeing services to its own and others' prices. The minimal

21. That the effect works by shifting the hazard around a margin is analogous to Kennan's (1980) results on the effects of paying unemployment insurance benefits to strikers after the strike has lasted for some length of time, and to changes in hazard rates out of unemployment around the time benefits are exhausted (e.g., Meyer, 1990).

reliance on such bribes thus far means that we are far from being able to infer these optimal prices.

B. The Market for Quality Refereeing

Assume that the quality of each referee is known to the customer-editor who assigns papers for refereeing, and that the editor seeks to obtain a high-quality, timely evaluation of the paper. This may result from a desire to produce a high-quality journal of current interest as well as from the editor's role as an agent for gathering information for authors. Assume that the information contained in the proxies for quality, Z, is the same as that available to editors. Then we can interpret the effect of an increase in the value of one of these proxies as the price of quality in this implicit market.

To derive the implicit price, ask the question: What is the total effect of an increase in quality on the probability distribution of time until service (by all suppliers, not just the first referee to whom the paper is sent)? This distribution is affected both by changes in waiting time and by changes in refusal and loss rates, i.e., by $\partial L^*/\partial Z$ and by $\partial \rho^*/\partial Z$. I use the estimates of the multinomial logit in columns (1) and (2) of Table 2, and of the accelerated failure time model in column (3) of Table 3, to infer these derivatives. (For all but the citations and experience measures the variables in $\mathbf{x}$ are held constant at their means.) I assume that a paper that is returned by a refuser is sent out for a second time 5 weeks after the first time (2 weeks in the editorial office plus the 3-week mean time until refusal). Each paper handled by a loser is sent to a new referee after 37 weeks (2 weeks in the editorial office plus the 35-week mean time until the paper is viewed as lost).

Table 5 presents statistics describing the distribution of completed service times until an editor receives one report. Column (1) lists means and order statistics for the distribution when the mean values of all variables are used. It shows that one report can be obtained on the median paper in less than two months; and for only ten percent of articles does it take longer than four months to obtain a report.[22] The long upper tail of the distribution of waiting times is generated by the (small) probability that the paper is sent to a loser.

The remaining columns in the Table evaluate the distributions by comparing little- to heavily-cited referees, and fresh Ph.D.s (experience of zero) to those with fifteen years of experience (for whom the estimates in Tables 2 and 3 showed that L^* and ρ^* were around the highest in the sample). Consider columns (2) and (3), where only the referee's citations vary. Reliance on heavily-cited referees increases the mean time until completion by 3 weeks. If we add differences in experience (as in the comparison between columns (4) and (5)), the implicit price of quality is a difference in mean waiting times of 3-1/2 weeks.

Accounting for a one-week round-trip by mail, the difference between columns (4) and (5) implies that a policy of using highest-

---

22. If we assume that editors seek two reports ab initio, the median waiting time to receive both reports rises only to 80 days, and the ninetieth percentile rises to 6.5 months.

Table 5. Distributions of Times to Obtain One Referee's Report (in Days)

		CHARACTERISTICS			
	Means	CITS < 10	CITS > 49	CITS < 10 X = 0	CITS > 49 X = 15
PERCENTILE					
10	15	14	22	13	24
25	31	27	43	25	45
50	55	50	72	46	76
75	86	79	107	76	116
90	128	115	153	124	166
95	234	163	195	288	226
MEAN	75	68	89	71	95
Probability of:					
Doer	0.784	0.854	0.652	0.883	0.644
Refuser	0.166	0.104	0.317	0.044	0.316
Loser	0.050	0.042	0.031	0.073	0.040

instead of lowest-quality referees requires the payment of 38 percent in additional time costs. This is an unbiased measure of the price of quality in this implicit market. If we assume that the market is in equilibrium, and assume too that editors are rational, we must conclude that this additional time cost is offset by the better refereeing job done by higher-quality referees. The additional lag is the price editors pay for obtaining better reports.[23]

VI. Truncating the Queue in Medical Practice

To examine whether the predictions of the theory of appointment books carry over into an entirely different endeavor, I obtained data describing physicians in private practice who participate in a prepaid health plan in a midwestern metropolitan area. The data describe only truncation, in particular, whether the physician was not taking new patients at the time the plan published its annual directory of participating members.

There are 321 primary-care physicians in the plan (family practitioners, internists, obstetrician-gynecologists, and pediatricians). Complete information was obtained on 264 (82 percent), who form the sample used in this Section. The measure of quality is experience, years since licensed in the state,

---

23. Of course, much of the burden of the lag is borne by authors, not by editors. Editors may know they face this trade-off and do so willingly. It is hard to believe that authors have such knowledge, though they may have information about the average quality of referees at different journals and the average refereeing lags. One might, though, view editors as the authors' agents in obtaining comments from the profession in a way that maximizes authors' utility.

essentially a measure of when the physician began practice.[24] Since I do not observe who the potential patients are, I cannot construct variables to distinguish between the effects of specific and general commitment. Instead, all of the other variables --- M.D. compared to osteopath, family practitioner or internist compared to others, female or male, outside or inside the central part of the metropolitan area, and practicing at least part-time at the local university medical school --- can be interpreted either way. Also, some may be viewed as reflecting differences among groups of physicians in arrival rates.

Columns 1 and 3 of Table 6 show the means of the variables. The mean of ρ^* is about 20 percent, but is nearly 30 percent for MDs alone. The average physician has been practicing around 16 years, but the range of experience is between 0 and 50 years. The large majority of physicians are men, are not affiliated with the university medical school, and are located in the central cities or their adjacent suburbs.

The main results of this Section are contained in the probits describing ρ^*, the probability that the physician is not taking new patients. Consider the estimates of the impact of experience on this probability, shown in columns 2 and 4 of Table 6. As a pair the coefficients on X and X^2 are jointly significant in the entire sample and especially among MDs alone. (Adding higher-order terms in X does not change the conclusion that ρ^* eventually decreases

24. The data are from American Medical Association Directory, 32nd edition, Chicago, 1990, and American Osteopathic Association, Yearbook and Directory of Osteopathic Physicians, Chicago, 1990.

Table 6. Means and Probit Estimates, Doctors' Refusal of New Patients*

	All Physicians		MDs	
	Mean	**Parameter**	**Mean**	**Parameter**
Variable	(1)	(2)	(3)	(4)
Family Practice or Internist	.77	1.16 (.30)	.73	1.23 (.33)
In Town	.73	.36 (.26)	.76	.69 (.31)
At University	.22	.55 (.25)	.19	.94 (.30)
Female	.19	.59 (.27)	.17	.64 (.33)
Experience	16.21	.0945 (.045)	17.14	.1274 (.055)
Experience^2	--	-.00158 (.00098)	--	-.00202 (.00115)
MD	.61	1.25 (.25)	1.00	--
Log-likelihood		-100.76		-76.62
N =		264		162
ρ^*	.201		.296	

*Standard errors in parentheses. Each equation also contains a constant.

with additional experience.) In the entire sample and among MDs alone 16 percent of the physicians are on the decreasing part of the truncation-experience relationship. The inverse U-shaped relationship is consistent with the view that experience is a proxy for quality that determines the arrival rate of customers to suppliers who are committed to offering a service without making customers remain in line too long. It is inconsistent with an interpretation that physicians build up a practice, begin turning away patients once they reach a full load that they somehow define, and continue to do so to allow their clientele to diminish so that they can reduce their labor supply as they near retirement.

Several of the control variables have significant impacts in ways that are consistent with arguments based on specific commitment (but also with heterogeneity in customers' demand). That in-town physicians and those associated with the medical school are more likely to truncate their queues may reflect their lesser likelihood of encountering their potential patients outside their practices. The greater probabilities of truncation by family practitioners and internists and by MDs are consistent with tighter markets for their services (and also with lower turnover of their patients). Greater truncation by female physicians may reflect their desire to allocate time to household production.[25]

Clearly, the results in this Section are not by themselves convincing evidence of the importance of appointment-book markets.

25. The higher truncation rate by female doctors is independent of specialty. Also, using separate dummy variables for all four specialties does not alter any of the results.

They do, though, corroborate the evidence from refereeing that the effect of quality on the arrival rate of customers combines with suppliers' commitment to generate differences in truncation rates. Moreover, because there is no intermediary who might match more difficult cases with higher-quality suppliers (these are all primary-care physicians), there is even less possibility that the matching of customers and suppliers is generating the results in this example than in the sample of referees.

VII. Conclusions, and Other Applications

I have pointed out the existence of what I have called appointment-book markets. In them suppliers choose their customers' waiting times and deny some customers entrance to the queue for services. The latter choice arises because the customers cannot obtain information about waiting times, and thus cannot make the utility-maximizing choices that the suppliers, who are generally committed to them, make on their behalf. The model predicts that higher-quality suppliers will exhibit longer waiting times and will be more likely to deny customers entry to the queue. They are less likely to deny entry to those customers with whom past contact has created specific commitment, and they will also serve those customers faster.

I examined this model using a unique set of data describing the behavior of referees for economics journals. Several measures of quality are positively correlated with the probability of denying entry to the queue (refusing to referee) and with waiting time (the lag in refereeing an article). Measures of specific commitment --- recent publication in the journal and location at the institution that houses the editorial office --- have the predicted negative effects on the refusal to referee, but have no effect on waiting time. This latter result cannot be explained within a simple queuing model. The results on the probability of denying entry to the queue are verified on a sample of physicians.

Consider first a specific normative implication of the empirical results on refereeing. They suggest that most referees are remarkably prompt, and that lags in the process result mainly from delays generated by a small percentage of referees whom I have called losers. Half (two-thirds) of the papers held longer than four (six) months have been sent to losers. To speed up the process with only a small loss in quality, journal editors might consider automatically truncating the long thin tail of waiting times by assuming that any referee who holds a paper for four (or six) months is a loser and sending the paper to another referee. Authors' welfare would be increased by the elimination of the low-probability, high-loss event of a very long wait; and editorial offices would obviate the serious headaches of dealing with losers.

More important than guidelines for scholarly publishing are the general implications of the appointment-book problem and the subtler issues of suppliers' behavior. Perhaps chief among these is the question of what generates general commitment --- why do suppliers truncate a queue that customers cannot see? Some impressionistic evidence that this is explained by businesses' desires to invest in long-term relations with customers is the failure to truncate queues for admission to certain tourist attractions. For examples, at Disneyland and at the Empire State Building, waiting occurs at a separate location from where tickets are purchased, there is no truncation, and the probability of a long-term relationship between customer and supplier is low. The same behavior may characterize how employers handle job applications: To what extent do employers bother notifying unsuccessful applicants that jobs have been filled? Further examination of this issue is crucial.

Another issue worth studying is whether welfare maxima are reached in these markets. For example, instead of truncating, suppliers could simply announce the queue length and let customers ration the service by the value of their waiting times. Do transactions costs prevent this; or is the failure due to the monopolists' desires to control access to the service and offer favors to those customers to whom they are specifically committed? Comparisons of differences in rationing schemes that suppliers impose in response to different technical conditions could answer these questions. These in turn lead to studying how market structure affects waiting times and trunction, and whether suppliers invest in technical improvements that might enable customers to ration the service themselves.[26]

26. Some evidence against the monopoly argument is provided by the example of the Australian National Roads and Motorists' Association, whose automated phone-answering system announces the current response time to callers who are put on hold. The institution of automated branched phone-response systems is an example of a technical improvement designed to overcome the appointment-book problem.

Consumers' inability to see the full price of a service because they cannot discover its time price generates unusual behavior by suppliers. It allows them more discretion than in markets with only money prices, or with visible time prices, in that they can discriminate among customers. It also, though, may encourage them to turn away customers even though that may not be short-run profit-maximizing. It may not even be long-run profit-maximizing if their actions are affected by a general feeling of commitment to potential customers.

REFERENCES

Rebecca Blank, "The Effects of Double-Blind versus Single-Blind Reviewing: Experimental Evidence from the American Economic Review," American Economic Review, 81 (1991), forthcoming.

Francis Boddy, "Rating of Economics Graduate Programs - 1982," unpublished report, University of Minnesota, 1982.

Stephen Cole, Leonard Rubin and Jonathan Cole, Peer Review in the National Science Foundation. Washington: National Academy of Sciences, 1978.

Diana Crane, "Gatekeepers of Science: Some Factors Affecting the Selection of Articles for Scientific Journals," American Sociologist, 2 (1967): 195-201.

Ralph D'Agostino and Michael Stephens, Goodness-of-Fit Techniques. New York: Marcel Dekker, 1986.

Robert Deacon and Jon Sonstelie, "Rationing by Waiting and the Value of Time," Journal of Political Economy, 93 (1985): 627-647.

Peter Dolton, "Job Separations, Job Tenure and Earnings: An Empirical Analysis," unpublished paper, University of Newcastle, 1991.

Robert Frank, Passions Within Reason. New York: Norton, 1988.

Daniel Hamermesh, George Johnson and Burton Weisbrod, "Scholarship, Citations and Salaries," Southern Economic Journal, 49 (1982): 472-481.

Samuel Karlin and Howard Taylor, A First Course in Stochastic Processes. New york: Academic Press, 1975.

John Kennan, "The Effect of Unemployment Insurance Payments on Strike Duration," in National Commission on Unemployment Compensation, Unemployment Compensation: Studies and Research, Volume 2. Washington: NCUC, 1980.

Nicholas Kiefer, "Economic Duration Data and Hazard Functions," Journal of Economic Literature, 26 (1988): 646-679.

David Laband, "Is There Value-Added from the Review Process in Economics?: Preliminary Evidence from Authors," Quarterly Journal of Economics, 105 (1990): 341-353.

Tony Lancaster, "A Stochastic Model for the Duration of a Strike," Journal of the Royal Statistical Society, 135 (1972): 257-271.

Stanley Liebowitz and John Palmer, "Assessing the Relative Impacts of Economics Journals," Journal of Economic Literature, 22 (1984): 77-88.

Cotton M. Lindsay and Bernard Feigenbaum, "Rationing by Waiting Lists," American Economic Review, 74 (1984): 404-417.

Bruce Meyer, "Unemployment Insurance and Unemployment Spells," Econometrica, 58 (1990): 757-782.

Donald Nichols, Eugene Smolensky and T. Nicolaus Tideman, "Discrimination by Waiting Time in Merit Goods," American Economic Review, 61 (1971): 312-323.

Douglas Peters and Stephen Ceci, "Peer-Review Practices of Psychological Journals: The Fate of Published Articles, Submitted Again," in Stevan Harnad, ed. Peer Commentary on Peer Review. Cambridge: Cambridge University Press, 1982.

Robert Wilson, "Efficient and Competitive Rationing," Econometrica, 57 (1989): 1-40.

Harriet Zuckerman and Robert Merton, "Patterns of Evaluation in Science: Institutionalisation, Structure and Functions of the Referee System," Minerva, 9 (1971): 66-100.

Appendix Table A1. Questionnaire for the Study of Refereeing Lags

DATE PAPER SENT TO REFEREE: MONTH __________ DAY __________

LENGTH OF PAPER (total pages): __________

REFEREE'S NAME: __________

REFEREE'S AFFILIATION: __________

DATE PAPER RECEIVED BACK
FROM REFEREE: MONTH __________ DAY __________

THE REFEREE: DID: __________ DID NOT: __________ SUBMIT A REPORT

Appendix Table A2. Means of Variables, by Journal*

VARIABLE	JOURNAL DESIGNATION						
	G1	G2	G3	G4	S1	S2	S3
Same School	0.042 (.067)	0.143 (.163)	0.042 (.053)	0.306 (.379)	0.122 (.135)	0.140 (.163)	0.020 (.020)
Published in Journal	0.250 (.333)	0.490 (.488)	0.271 (.316)	0.061 (.069)	0.184 (.189)	0.420 (.465)	0.540 (.551)
Foreign	0.083 (.100)	0.082 (.070)	0 (0)	0.020 (.035)	0.061 (.054)	0.020 (.023)	0 (0)
Male	0.938 (.90)	0.939 (.930)	0.896 (.895)	0.939 (.966)	0.816 (.757)	0.820 (.791)	0.940 (.939)
Top 20	0.562 (.533)	0.286 (.256)	0.062 (.053)	0.184 (.138)	0.510 (.514)	0.400 (.419)	0.240 (.224)
Cartel	0.271 (.333)	0.449 (.465)	0.479 (.500)	0.633 (.759)	0.306 (.297)	0.340 (.279)	0.400 (.408)
Other Academic	0.062 (.033)	0.122 (.140)	0.292 (.289)	0.102 (.103)	0.143 (.135)	0.200 (.233)	0.160 (.163)
CITS10-49	0.500 (.500)	0.429 (.419)	0.208 (.184)	0.449 (.448)	0.367 (.297)	0.340 (.279)	0.480 (.469)
CITS50	0.292 (.200)	0.122 (.116)	0.021 (.026)	0.143 (.138)	0.122 (.108)	0.100 (.116)	0.060 (.061)
Citations	43.79 (39.67)	24.65 (24.79)	7.08 (6.89)	22.47 (22.24)	28.18 (28.46)	15.32 (15.40)	15.56 (15.41)
Ph.D. Experience	18.54 (16.33)	14.74 (14.56)	15.77 (15.45)	18.16 (19.83)	10.37 (9.27)	19.80 (19.91)	18.00 (18.20)
Pages	27.58 (25.43)	25.45 (25.26)	22.29 (21.97)	22.26 (22.00)	33.43 (32.73)	32.78 (33.28)	23.26 (23.20)
Days	(45.37)	(56.02)	(50.10)	(56.76)	(90.60)	(49.30)	(40.76)
N =	48 (30)	49 (43)	48 (38)	49 (29)	49 (37)	50 (43)	50 (49)

*Means in parentheses are for doers only; others describe the entire sample.