Information and Usage of Free–Access Congestible Facilities

Richard Arnott
and
André de Palma
(University of Geneva)
and
Robin Lindsey
(University of Alberta)

Working Paper No. 238
October 1992

INFORMATION AND USAGE OF FREE-ACCESS CONGESTIBLE FACILITIES

Richard Arnott*
André de Palma**
Robin Lindsey***

October 1992

* Boston College, Chestnut Hill, Mass., U.S.A 02167-3806

** University of Geneva, Geneva 4, CH-1211, Switzerland.

*** University of Alberta, Edmonton, Alberta, Canada T6G 2H4

Financial support from the Social Sciences and Humanities Research Council of Canada, the Natural Sciences and Engineering Research Council of Canada, and NSF under grant SES-8912335 is gratefully acknowledged. Helpful comments were received from Qin Liu, Ken Small and colleagues at the University of Alberta.

Abstract

We investigate the effect of information on participation and time-of-use decisions in a free-access congestible facility subject to fluctuations in capacity and demand. Expected welfare is greater with perfect than with zero information, while optimal design capacity is greater if and only if demand elasticity is greater than one. Imperfect information can reduce welfare by inducing concentration in the arrival times of users at the facility. This suggests that information dissemination schemes such as advanced traveller information systems for automobile travel should be designed and implemented with care.

Notational Glossary

Greek characters

α	Unit cost of travel time
β	Unit cost of arriving early
γ	Unit cost of arriving late
$\Gamma(y)$	Cumulative distribution function (c.d.f.) of days
δ	$\equiv \beta\gamma/(\beta+\gamma)$
ε	Price elasticity of demand
η	Elasticity of average cost with respect to level of usage
θ	n/σ
ν	Certainty-equivalent intensity of demand
ξ	Scale factor for average cost function
$\rho(t)$	Proportional arrival rate of users at facility
σ	Ratio of realized capacity to design capacity
ϕ	N/s
$\phi^Q(t)$	Largest ϕ such that queueing never occurs at time t
$\phi^*(t)$	ϕ such that user arriving at facility at time t is served at t^*

English characters

$C(t)$	Cost for user arriving at facility at time t
CS	Consumers' surplus
e	(superscript) denoting equilibrium value
$F(n,\sigma)$	Joint c.d.f. of n and σ
$G(n)$	c.d.f of n
$H(\sigma)$	c.d.f. of σ
I	(superscript) denoting imperfect information régime
$J(\phi)$	c.d.f. of ϕ from perspective of users
$J^\theta(\theta)$	c.d.f. of θ from perspective of users
k	Marginal cost of capacity expansion
m	Index of messages
M	c.d.f. of m
MCS	Marginal consumers' surplus from expanding design capacity
n	Intensity of demand
N	Number of users
p	Price (= user cost) of using facility
P	(superscript) denoting perfect information régime
$q(t)$	Queueing time at time t
$q(t,\phi)$	Queueing time at time t when $N/s = \phi$
$Q(t)$	Number of users in queue at time t
$r(t)$	Arrival rate of users at facility at time t
$R(t)$	c.d.f. of arrivals at facility (integral of $\rho(t)$)
s	Realized capacity of facility
$\hat{s}$	Design capacity of facility
SB	Social benefit
t	Arrival time at facility

t^*	Preferred time of usage
t_L	Latest arrival time at facility
t_n	Perfect information régime: Arrival time at facility for which user served at t^* Imperfect information régime: Latest arrival time at facility for which user is never served late
t_o	Earliest arrival time at facility
t_1	Latest service time
$T(t)$	Total time required to use facility for user arriving at time t
$\bar{T}$	Time spent using facility in absence of a queue
TC	Total costs
u_i	Indicator variable (= 1 if individual i is a prospective user, = 0 otherwise).
U	(superscript) denoting user
y	Index of days
Z	(superscript) denoting zero information regime
$Z^Q(t)$	Rate of increase in expected queueing time wrt arrival time.
$Z^E(t)$	Rate of decrease in expected early time wrt arrival time.
$Z^L(t)$	Rate of increase in expected late time wrt arrival time.

A little learning is a dangerous thing[1]

1. Introduction

Congestion occurs when the (social) marginal cost of using a facility exceeds the user cost. Although congestion is inherently a dynamic phenomenon, the original peak-load model of congestion (Steiner (1957), Williamson (1966), Mohring (1970), *inter alios*) is essentially static. In its simplest form, there are two demand periods, peak and off-peak, that are linked only by a common capacity. Where interdependence in demand across periods is allowed for, it takes the form of parametric cross-price elasticities. Empirical studies (notably of electricity demand) generally adopt such a method. This approach has several ambiguities (Arnott *et al.* (1991b)) which derive from a failure to model explicitly user behavior and the congestion technology of the particular facility.

A second limitation of the original peak-load model is that the timing and intensity of peak and off-peak demands are assumed known, and capacity is assumed (usually implicitly) to be constant over time and perfectly reliable. Yet many real-world facilities are subject to substantial fluctuations in demand and capacity of a predictable or unpredictable nature.

Some progress has been made to rectify these two limitations of the original model. The first systematic treatment of time-of-use decisions was undertaken by Vickrey (1969) in the context of transportation. Vickrey used a deterministic queueing model to describe the evolution of congestion during the morning-rush-hour traffic. A number of authors have applied and extended his model, though they have continued to work in a nonstochastic setting.

Demand uncertainty was introduced into the peak-load model by Brown and Johnson (1969), whose work encouraged a series of further papers. De Vany and Saving (1977, 1980), Kraus (1982) and D'Ouville and McDonald (1990) have considered uncertainty in demand in the context of road transportation, but in a steady-state framework.[2] Some work has also been done in transportation in developing dynamic models which allow for stochasticity in individual driver behavior, but not stochasticity at the aggregate level. To our knowledge, the only non-steady-state model with macroscopic stochasticity is that of Gaver (1968), which solved for the optimal departure time of a commuter faced with random travel delays. The paper did not, however, derive an *equilibrium* departure time distribution or specify whether the source of randomness originates with demand or capacity.

In summary, while there is a considerable body of literature that addresses various facets of the dynamic peak-load problem under uncertainty, it has not yet received a conceptually integrated treatment.

This paper has two aims. First, we develop a model of individuals' participation and time-of-usage decisions at a free-access congestible facility. In allowing for both predictable and unpredictable fluctuations in both demand and capacity, the model is more general in its treatment of uncertainty than the previous literature.

The second objective of the paper is to investigate the welfare effects of information provided to users that reduces their uncertainty about demand and/or capacity. There are several margins of user behavior that can be affected by information, including participation, intensity of use, time of use, and choice of facility. Information can therefore alter both the temporal and spatial pattern of congestion.

The effect of information on usage of congestible facilities is a relatively unresearched topic. It is also of more than purely intellectual interest. Research has been underway for over a decade on the design and operation of Advanced Traveller Information Systems (ATIS) which provide drivers with advice on route selection, as well as information on the location and severity of delays on the road network.[3]

Given the substantial costs of implementing and operating ATIS, it is important to assess their potential benefits in advance. Traffic engineers distinguish between the user equilibrium and the system optimum. The system optimum for a fixed set of travelers consists of a departure time and route for each driver that minimizes aggregate travel costs for the system. Better information can only improve the system optimum.

The user equilibrium, in contrast, is a Nash equilibrium in which no driver can unilaterally do better by changing his departure time or route. Congestion in the user equilibrium is an *uninternalized* externality. Information may induce changes in driver behavior that increase the deadweight loss associated with unpriced congestion, thereby making information welfare-reducing. This can happen when drivers are induced to make more similar route and/or departure time choices. Ben-Akiva, de Palma and Kaysi (1991) have called this phenomenon *concentration*. A growing number of studies have demonstrated that concentration can lead to higher system costs under plausible conditions.[4] Schelling (1978, 229-230) discusses the possibility in the case of route choice. Arnott *et al.* (1991a) provide a formal analysis of concentration in departure time. Concentration in both departure time and route choice has recently been demonstrated through simulation by, *inter alios*, Mahmassani and Jayakrishnan (1991), Reiss, Gartner and Cohen (1991), Halati and Boyce (1991) and Koutsopoulos and Xu (1992).

There is, then, some reason to believe that supplying information to users of a congestible facility can be welfare-reducing even before accounting for the costs. It remains to identify the precise circumstances in which this can occur, in particular the nature of the congestion technology and the quality of the information provided.

Rather than attempt a general analysis of the dynamic peak-load problem with uncertainty and information, we provide a reasonably thorough analysis of a pair of related models with the following characteristics.

C-1: The congestible facilities we treat are *delay systems*, in which all users who wish to receive service do, but at a quality that degrades with aggregate usage.

Examples of delay systems include roads, indoor and outdoor recreational activities, and interactive computer systems. Depending on the facility and the circumstances, declining quality can manifest itself in queues, flow congestion or crowding. By contrast, *loss systems* provide all-or-nothing service at a (generally) constant quality. Electricity (except for brown outs) and natural gas are examples.[5]

C-2: We examine *free-access* facilities; those for which neither price nor non-price rationing is implemented as a policy.

Roads and public recreational facilities are the main examples of free-access facilities. Most roads are publicly provided and free of charge, at least in North America. Similarly, admission to public recreational areas in the U.S. has traditionally been provided either free or at a nominal fee.

Since most roads and public recreational facilities are free-access delay systems, our analysis is directly primarily at them.[6]

C-3: The treatment of uncertainty assumes that *nature is stochastic* and that users have *rational expectations*. Systematic or idiosyncratic errors in perception are not considered.[7]

C-4: There are both predictable (recurring) and unpredictable (nonrecurring) fluctuations in both demand and capacity *across* periods of use (*e.g.* morning rush hours). But demand and capacity remain constant *within* each period of use. This rules out, for example, traffic accidents which reduce capacity for only a portion of a rush hour, and precludes consideration of the value of in-vehicle information.

C-5: We treat a single congestible facility in isolation. This rules out examination of route and mode choice, and interdependence among congestible elements in a network.

C-6: Users are identical.

Our analysis consists of two parts. In Section 2 we adopt a static model in which a user decides *whether* to use a facility, but not *when*. This approach is reasonable if individuals lack strong time-of-use preferences, or if service entails a 'batch' operation, as is the case with entertainment or sporting events, or once-a-year parades. Also, as shown in Section 3, when users have full information, the static model is a reduced form of our dynamic model. Hence all the results from the static model with perfect information extend to the dynamic model with perfect information when time-of-use is made endogenous.

On the assumptions that user costs are isoelastic with respect to the ratio of usage to capacity and that demand is isoelastic, we derive the optimal design capacity for a facility, and show how it varies with the joint probability distribution of demand and capacity availability and with the information available to users. We also show that better information is welfare-improving in a sense that will be made precise. However, this result is not robust to the functional form of the cost and demand curves.

In Section 3 we extend consideration to time-of-use decisions by adopting Vickrey's (1969) model of queueing behind a bottleneck. Perfect information again turns out to be welfare-improving but, in contrast to the static model, imperfect information need not be.

2. Static Equilibrium

2.1 The Model

The static equilibrium model is based on conventional supply and demand analysis. As stated in the introduction, we consider an isolated delay system under free access. The idea that congestion is manifest as delay is formalized by the assumption

A-1: Average (*i.e.* user) cost is increasing in demand and decreasing in capacity:

$$AC = AC(N,s), \quad \frac{\partial AC}{\partial N} > 0, \quad \frac{\partial AC}{\partial s} < 0,$$

where AC is user cost, N is demand, and s is capacity.

And we characterize free access by:

A-2: The price, p, of facility use equals user cost:

$$p = AC.$$

Throughout most of the paper, we shall employ specific assumptions concerning the forms of demand and of user cost.

A-3: Average cost exhibits constant elasticity with respect to the ratio of demand to capacity:

$$AC = \xi\left(\frac{N}{s}\right)^{\eta}, \tag{2.1}$$

where ξ is a parameter, and $\eta > 0$ is the elasticity of average cost with respect to the demand-capacity ratio.

At this point, we do not specify the form of congestion (queueing, flow congestion, crowding, *etc.*), but in Section 3 we shall work with an explicit queueing congestion model.

A-4: Demand is isoelastic:

$$N = np^{-\varepsilon}, \tag{2.2}$$

where n is a parameter characterizing demand intensity and ε is the demand elasticity.

n and s are assumed to vary across periods of use. The fluctuations can range from being fully predictable to completely unpredictable. In the case of recreational capacity, seasonal openings and closures of hiking trails and campsites may follow a fixed timetable easily learned by prospective visitors. Traffic lane closures due to maintenance and repairs are predictable if publicized well in advance and completed according to schedule. Seasonal variations in lighting, that affect travel and outdoor activities, are deterministic. Predictability of weather depends on the accuracy of weather forecasts, and how speedily and widely forecasts are disseminated. Accidents, vehicular breakdowns and signal failures are unpredictable, as are mechanical breakdowns in public utilities and other facilities.

If individual usage decisions are statistically independent, then by the law of large numbers fluctuations in aggregate demand will be insignificant in proportional terms. If, alternatively, individual usage decisions are governed by common causal factors, there will be predictable daily, weekly, seasonal and/or holiday cycles. In either case, there should be no demand surprises. Yet surprises do occur. Previous attendance at annual events can be forgotten, or become obsolete through changing demographics, incomes, tastes, *etc.* And for relatively rare or distinctive events such as transit strikes, periods of gasoline rationing or epidemics, or for singular events such as world fairs, there may be little or no precedent on which to base predictions.

To formalize these ideas, we assume for concreteness that the period of use is a day, and that demand intensity and capacity are determined before daily participation decisions are made. The realized values of demand intensity and capacity are assumed constant until the following day. Fluctuations in capacity are modeled by writing $s = \sigma\hat{s}$, where $\hat{s}$ is design capacity (capacity under ideal conditions) and $\sigma \leq 1$ is the fraction of capacity available. Fluctuations in demand are captured through changes in demand intensity, n. Thus, n and σ are random variables.

To allow for systematic fluctuations in n and σ, we let y be an index of 'days'. For each y, there is a distinct joint frequency distribution of n and σ. Let $\Gamma(y)$ denote the cumulative distribution function of y, and $F_y(n,\sigma)$ be the joint c.d.f. of n and σ on day y. Both $\Gamma(y)$ and $F_y(n,\sigma)$ are assumed to be upper hemicontinuous (they can have mass points), and to be independent of design capacity. n/σ is assumed to have a finite upper bound and to be nonzero with positive probability. Finally, $\Gamma(y)$ and $F_y(n,\sigma)$ are assumed to be common knowledge among users.

We will be concerned with three information régimes: perfect information, zero information, and imperfect information. Under perfect information, users learn the precise realizations of n and σ before making their participation decisions. Under zero information they learn nothing, and have only $F_y(n,\sigma)$ to go on. Both perfect and zero information are limiting cases of imperfect information, which we discuss next.

In the imperfect information régime, there is a message system that reports daily on demand intensity and capacity after they are realized but before usage decisions are made. This message system is an idealized representation of weather forecasts, early morning traffic reports, reports on ski conditions, *etc.*[8] Let m be an index of messages, and $M(m|y,n,\sigma)$ be the c.d.f. of m conditional on y and the realization (n,σ). Messages convey information about n and/or σ in the form of a c.d.f. $\hat{F}_{ym}(n,\sigma)$. It is assumed that

A-5: $\hat{F}_{ym}(n,\sigma)$ coincides with the true probability distribution of n and σ conditional on y and m, $F_{ym}(n,\sigma)$:

$$\hat{F}_{ym}(n,\sigma) = F_{ym}(n,\sigma). \tag{2.3}$$

A-5 is a condition of unbiasedness or irrefutability, in the sense that forecast probabilities of states coincide with their realized frequencies. This seems reasonable in the long run, since forecasters will have had the opportunity to correct biases in their forecasts, and users to compare the frequencies of forecasts and realizations.[9] Consequently, the superscript ^ on the c.d.f. conditional on messages is henceforth omitted.

The probability function $F_{ym}(n,\sigma)$ has been defined from the perspective of the analyst or outside observer. A subtle but crucial point, which is well-known in queueing theory,[10] is that this distribution generally differs from the corresponding distribution of individuals who consider using the facility. The reason in the present context is that demand intensity varies from day to day.

Before showing this formally, we illustrate the point with a stylized example. Suppose that buses on a certain route are either full or half full, with equal probability. From the perspective of an outside observer the two states do indeed each occur with probability 1/2. But for bus riders, the respective probabilities of traveling on a full bus and on a half-full bus are 2/3 and 1/3, because twice as many people travel on full buses. The user's probability distribution is weighted toward high demand states. Of course, users may know which state is in effect because of regularity in their travel behavior *vis-à-vis* other riders. This can be captured in the model by making the probability distributions of demand day-specific.

We now derive the relationship between the probability distribution of n and σ for the outside observer, $F_{ym}(n,\sigma)$, and that of users, $F^{u}_{ym}(n,\sigma)$. We shall refer to individuals who ever wish to use the facility on day y as *potential users* on day y, and individuals who actually wish to use the facility on a given occasion as *prospective users*.[11] Let $\bar{N}_y$ be the number of potential users on day y. Further, let u_i be an indicator variable, with u_i = 1 if individual i is a prospective user, and $u_i = 0$ otherwise. The number of prospective users is $\sum_{i=1}^{\bar{N}_y} u_i$, and realized demand intensity is proportional to this:

$$n = c \sum_{i=1}^{\bar{N}_y} u_i, \tag{2.4}$$

for some $c > 0$. Now

$$\Pr\left\{ u_i=1 \middle| n,\sigma,y,m \right\} = E\left\{ u_i \middle| c \sum_{j=1}^{\bar{N}_y} u_j = n,y,m \right\}, \quad i = 1\ldots\bar{N}_y, \tag{2.5}$$

which is independent of i by assumption C-6. Taking the conditional expectation of (2.4)

$$c \sum_{i=1}^{\bar{N}_y} E\left\{ u_i \middle| c \sum_{j=1}^{\bar{N}_y} u_j = n,y,m \right\} = n. \tag{2.6}$$

Together, (2.5) and (2.6) yield

$$\Pr\left\{ u_i=1 \middle| n,\sigma,y,m \right\} = \frac{n}{c\bar{N}_y}. \tag{2.7}$$

We now drop subscript i. By Bayes' rule[12]

$$dF^u_{ym}(n,\sigma) = \frac{\Pr\left\{ u=1 | n,\sigma,y,m \right\}}{\int_n \int_\sigma \Pr\left\{ u=1 | n,\sigma,y,m \right\} dF_{ym}(n,\sigma)} dF_{ym}(n,\sigma). \tag{2.8}$$

Substituting (2.7) into (2.8),

$$dF^u_{ym}(n,\sigma) = \frac{n}{\bar{n}_{ym}} dF_{ym}(n,\sigma), \tag{2.9}$$

where

$$\bar{n}_{ym} \equiv \int_n \int_\sigma n \; dF_{ym}(n,\sigma).$$

As is true of the bus example, the user's probability distribution is weighted towards high demand-intensity states. Naturally, if demand intensity is not variable, the distributions of the outside observer and of the user coincide.[13]

2.2 Perfect Information

In this section we first solve for equilibrium price and level of usage under perfect information for given realizations of demand and capacity. Then we derive optimal design capacity. Finally, we examine how optimal design capacity varies with the joint distribution of n and σ. Several other authors have conducted parallel analyses, using different models and assumptions. At the end of this section we make a brief comparison.

Let n and s be the demand intensity and capacity realized on a given day. Since access is free, the full price equals user cost. Equilibrium price, p^e, and usage, N^e, can hence be solved by equating AC and p, given respectively by (2.1) and the inverse demand curve defined by (2.2):

$$p^e = \left(\frac{n}{N^e}\right)^{\frac{1}{\varepsilon}} = AC = \xi\left(\frac{N^e}{s}\right)^{\eta}. \tag{2.10}$$

This yields

$$p^e(n,s) = \xi^{\frac{1}{1+\varepsilon\eta}}\left(\frac{n}{s}\right)^{\frac{\eta}{1+\varepsilon\eta}}, \tag{2.11}$$

$$N^e(n,s) = \xi^{-\frac{\varepsilon}{1+\varepsilon\eta}}\, n^{\frac{1}{1+\varepsilon\eta}}\, s^{\frac{\varepsilon\eta}{1+\varepsilon\eta}}. \tag{2.12}$$

Consumers' surplus in state (n,s) is

$$CS(n,s) = \int_{p^e(n,s)} N(p)dp.$$

Expected consumers' surplus, taking into account both systematic and unpredictable variations in demand and capacity, is

$$\bar{CS}^P(\hat{s}) = \int_y \int_n \int_\sigma \int_{p^e(n,\sigma\hat{s})} np^{-\varepsilon}dp \; dF_y(n,\sigma)\; d\Gamma(y), \tag{2.13}$$

where superscript P denotes perfect information. The increase in consumers' surplus from a marginal capacity expansion, which we call marginal consumers' surplus, is

$$\overline{MCS}^P(\hat{s}) \equiv \frac{d\overline{CS}^P(\hat{s})}{d\hat{s}}.$$

From (2.11) and (2.13):

$$\overline{MCS}^P(\hat{s}) = \frac{\eta}{1+\varepsilon\eta}\,\xi^{\frac{1-\varepsilon}{1+\varepsilon\eta}}\,(\hat{s})^{-\frac{1+\eta}{1+\varepsilon\eta}} \int_y\int_n\int_\sigma n^{\frac{1+\eta}{1+\varepsilon\eta}}\,\sigma^{\frac{\eta(\varepsilon-1)}{1+\varepsilon\eta}}\,dF_y(n,\sigma)\,d\Gamma(y). \qquad (2.14)$$

If capacity costs are ks, then optimal capacity is given by[14]

$$\hat{s}^P_* = \text{argmax}\left[\overline{CS}^P(\hat{s}) - k\hat{s}\right].$$

Using the first-order condition $\overline{MCS}^P(\hat{s}) = k$, one obtains

$$\hat{s}^P_* = \left[\frac{\eta}{(1+\varepsilon\eta)k}\right]^{\frac{1+\varepsilon\eta}{1+\eta}}\xi^{\frac{1-\varepsilon}{1+\eta}}\,\nu^P, \qquad (2.15a)$$

where

$$\nu^P \equiv \left[\int_y\int_n\int_\sigma n^{\frac{1+\eta}{1+\varepsilon\eta}}\,\sigma^{\frac{\eta(\varepsilon-1)}{1+\varepsilon\eta}}\,dF_y(n,\sigma)\,d\Gamma(y)\right]^{\frac{1+\varepsilon\eta}{1+\eta}}. \qquad (2.15b)$$

ν^P can be interpreted as a certainty-equivalent intensity of demand with perfect information, *viz.* the constant demand intensity for which under ideal conditions ($\sigma = 1$) design capacity is $\hat{s}^P_*$. In the limiting case where neither systematic nor random fluctuations in capacity or demand occur, ν^P reduces to the (constant) demand intensity n.

While the model underlying (2.15) is obviously a simple one, it has the merit of incorporating both demand and cost parameters, as well as systematic and nonsystematic fluctuations in demand and capacity. This contrasts with some formulas, or rather rules of thumb, used in capacity design. A good illustration of this is in the design of road capacity as described by the Institute of Transportation Engineers (1982), and the Highway Capacity Manual of the Transportation Research Board (1985).[15]

The effect of capacity and demand fluctuations on optimal design capacity as given in (2.15) is summarized in the following proposition (for a precise statement and proof, see Appendix 1):

PROPOSITION 1: Under A-1 to A-4 and with perfect information, if $\varepsilon < 1$ then, for any mean demand intensity, optimal design capacity is greater when shortfalls in capacity can occur than when capacity is always at its design level. Furthermore, optimal design capacity is the larger: (a) with a mean-preserving spread in demand, (b) with a mean-preserving spread in capacity availability, (c) the lower the ratio of generalized mean capacity to design capacity, and (d) the lower the generalized correlation between demand and capacity. The opposite is true of all the above when $\varepsilon > 1$.

Prop. 1 establishes that if demand for trips is inelastic (which is normally the case with commuting)[16] greater investment in capacity is warranted if loss of capacity can occur than if it cannot. If demand is elastic, less capacity is optimal in the stochastic case. Intuition can be gleaned from the following analogy. Suppose that tomatoes remain fresh if refrigerated, but if left unrefrigerated spoil according to some stochastic process. Consider an individual who ~~c~~shops once a week for tomatoes. If he has a fridge he maximizes utility by purchasing m pounds. If he does not have a fridge, should he purchase more or less than m pounds per week? If tomatoes are a necessity (inelastic demand) he should purchase more than m pounds, while if there is a reasonably close non-perishable substitute (elastic demand) he should purchase less than m pounds.

Prop. 1 also states that, if $\varepsilon < 1$, optimal design capacity increases with a mean-preserving spread (MPS) in the distribution of demand intensity or a MPS in the distribution of capacity availability. Optimal capacity is also greater the more states of *high* demand coincide with *low* capacity.

Our result concerning the effect of a MPS in demand intensity may be compared with related results of other authors. Brown and Johnson (1969) considered a public utility which sets price before demand intensity is known and then adjusts output to meet demand at the set price within its fixed capacity limits. If demand exceeds capacity, rationing necessarily occurs. (Thus, theirs is a loss system with state-independent pricing, whereas ours is a delay system with free access.) On the assumptions that demand is linear and that supply is rationed to users with the highest willingness to pay, Brown and Johnson showed that optimal capacity is unambiguously increased by demand uncertainty. Visscher (1973) later showed that with random rationing, or rationing to users with the lowest willingness to pay, optimal capacity may be lower with demand variability.

Kay (1979) considered an electric utility faced with uncertain demand that sets welfare-maximizing rates conditional on service. He showed (p.611) that, with isoelastic demand, optimal capacity under uncertainty is greater than riskless capacity if, but not only if, $\varepsilon < 1$. Capacity may be lower under uncertainty if $\varepsilon > 1$ if incremental demand is concentrated at peak periods, and if the costs of denied service are modest.

In the context of transportation, Kraus (1982) assumed that demand is constant, but known only imperfectly to a planner. Given unit demand elasticity he showed that optimal (road) capacity is increased by the planner's uncertainty about demand. By contrast, D'Ouville and McDonald (1990) assumed that travel demand is stochastic. On the assumption that individual travel costs vary quadratically with the number of drivers they showed that, again, optimal capacity is increased by demand variability.

The diversity of these results illustrates the fact, well-known in the economics of uncertainty, that the effects of uncertainty are sensitive to the nature of the uncertainty and to the characteristics of the system, here the congestion technology.

2.3 Imperfect Information

In the preceding section it was assumed that users learn demand and capacity before making participation decisions. In this section, we consider the more realistic situation in which users know the systematic fluctuations, but have only imperfect information about nonsystematic fluctuations. First we solve for the expected equilibrium price on a given day with a given message. Next we solve for optimal design capacity, and recover the limiting cases of perfect information and zero information. Finally, we rank optimal design capacities and consumer welfare under imperfect information and zero information.

Let $\bar{p}^u_{ym}$ be the price expected by a prospective user when message m is received on day y. We assume

A-6: Under uncertainty, users base their demand on the expected price. Thus

$$N_{ym}(n) = n\left(\bar{p}^u_{ym}\right)^{-\varepsilon}. \tag{2.16}$$

Let $p(n,\sigma|y,m)$ be the resulting price, given realizations n and σ.
By (2.10), this is

$$p(n,\sigma|y,m) = \xi\left(\frac{N_{ym}(n)}{\sigma\hat{s}}\right)^{\eta}, \tag{2.17}$$

Substituting (2.16) into (2.17):

$$p(n,\sigma|y,m) = \xi\left(\frac{n\left(\bar{p}^u_{ym}\right)^{-\varepsilon}}{\sigma\hat{s}}\right)^{\eta}. \tag{2.18}$$

We further assume

A-7: Prospective users know $F^{u}_{ym}(n,\sigma)$ and use it to compute $\bar{p}^{u}_{ym}$.

Then

$$\bar{p}^{u}_{ym} = \int_n \int_\sigma p(n,\sigma|y,m)\ dF^{u}_{ym}(n,\sigma). \tag{2.19}$$

Combining (2.9), (2.18), and (2.19) yields

$$\bar{p}^{u}_{ym} = \int_n \int_\sigma p(n,\sigma|y,m)\ \frac{n}{\bar{n}_{ym}}\ dF_{ym}(n,\sigma)$$

$$= \int_n \int_\sigma \xi\left(\frac{n\left(\bar{p}^{u}_{ym}\right)^{-\varepsilon}}{\sigma\hat{s}}\right)^{\eta} \frac{n}{\bar{n}_{ym}}\ dF_{ym}(n,\sigma)$$

$$= \left[\frac{\xi(\hat{s})^{-\eta}}{\bar{n}_{ym}}\int_n \int_\sigma n\ \left(\frac{n}{\sigma}\right)^{\eta} dF_{ym}(n,\sigma)\right]^{\frac{1}{1+\varepsilon\eta}}. \tag{2.20}$$

Consumers' surplus is more tricky to evaluate than under perfect information. The number of users, given by (2.16), is proportional to n, and independent of σ. Gross social benefit, SB, with demand intensity n equals the area under the corresponding inverse demand curve up to the corresponding number of users. This area is the sum of the areas above and below $\bar{p}^{u}_{ym}$. Integrating, one obtains

$$\overline{SB}^{I} = \int_y \int_m \left[\int_n \int_\sigma \left\{ \int_{\bar{p}^{u}_{ym}} np^{-\varepsilon}dp + n\left(\bar{p}^{u}_{ym}\right)^{1-\varepsilon}\right\} dF_{ym}(n,\sigma)\right] dM(m|y)\ d\Gamma(y), \tag{2.21}$$

where the superscript I denotes imperfect information. From gross social benefit must be deducted expected aggregate user costs, $\overline{TC}^{I}$, to arrive at expected consumers' surplus. Now

$$\overline{TC}^{I} = \int_y \int_m \overline{TC}^{I}_{ym}\ dM(m|y)\ d\Gamma(y), \tag{2.22}$$

where (using p = AC)

$$\overline{TC}^I_{ym} = \int_n\int_\sigma TC(n,\sigma|y,m)\ dF_{ym}(n,\sigma) = \int_n\int_\sigma p(n,\sigma|y,m)\ N(n,\sigma|y,m)\ dF_{ym}(n,\sigma)$$

$$= \int_n\int_\sigma p(n,\sigma|y,m)\ n\left(\bar{p}^u_{ym}\right)^{-\varepsilon} dF_{ym}(n,\sigma)\ \text{(by (2.16))}$$

$$= \int_n\int_\sigma n\left(\bar{p}^u_{ym}\right)^{1-\varepsilon} dF_{ym}(n,\sigma)\quad \text{(by (2.9) and (2.19))}. \tag{2.23}$$

Substituting (2.23) into (2.22) and subtracting the result from (2.21) yields finally

$$\overline{CS}^I(\hat{s}) = \int_y\int_m \left[\int_n\int_\sigma \left\{\int_{\bar{p}^u_{ym}} np^{-\varepsilon}dp\right\} dF_{ym}(n,\sigma)\right] dM(m|y)\ d\Gamma(y) \tag{2.24}$$

$$= \int_y\int_m \left[\int_{\bar{p}^u_{ym}} \bar{n}_{ym}\ p^{-\varepsilon}dp\right] dM(m|y)\ d\Gamma(y).$$

Thus, for any y and m, consumers' surplus is calculated in the usual way, as the area below 'the' demand curve and above the price line, where 'the' demand curve is the mean-demand-intensity demand curve, and the price is users' expected price. Note that consumers' surplus is unbounded when $\varepsilon \le 1$.

Define expected marginal consumers' surplus with imperfect information as

$\overline{MCS}^I(\hat{s}) \equiv \dfrac{d\overline{CS}^I(\hat{s})}{d\hat{s}}$. Differentiating[17], and using (2.20):

$$\overline{MCS}^I(\hat{s}) = \frac{\eta}{(1+\varepsilon\eta)\hat{s}}\int_y\int_m \bar{n}_{ym}\left(\bar{p}^u_{ym}\right)^{1-\varepsilon} dM(m|y)\ d\Gamma(y). \tag{2.25}$$

Following the same procedure as earlier, optimal capacity is found to be

$$\hat{s}^I_* = \left[\frac{\eta}{(1+\varepsilon\eta)k}\right]^{\frac{1+\varepsilon\eta}{1+\eta}} \xi^{\frac{1-\varepsilon}{1+\eta}}\ \nu^I, \tag{2.26a}$$

where

$$\nu^I = \left[\int_y\int_m \left(\bar{n}_{ym}\right)^{\frac{\varepsilon(1+\eta)}{1+\varepsilon\eta}} \left\{\int_n\int_\sigma \left(\frac{n}{\sigma}\right)^\eta ndF_{ym}(n,\sigma)\right\}^{\frac{1-\varepsilon}{1+\varepsilon\eta}} dM(m|y)d\Gamma(y)\right]^{\frac{1+\varepsilon\eta}{1+\eta}}. \tag{2.26b}$$

As before, ν^I can be interpreted as a certainty-equivalent demand intensity. In the case of perfect information, (2.26b) reduces to (2.15b). To see this, note that a perfect message system maps messages into unique states $M(m|y)$: $m \to (n,\sigma)$. The integral with respect to m in (2.26b) is thus replaced by a double integral with respect to n and σ, while the double integral inside the braces disappears.

The opposite extreme of zero information obtains when messages are completely uninformative. Optimal design capacity is given by (2.26) with the (redundant) integration and subscripts with respect to m omitted:

$$\hat{s}^Z_* = \left[\frac{\eta}{(1+\varepsilon\eta)k}\right]^{\frac{1+\varepsilon\eta}{1+\eta}} \xi^{\frac{1-\varepsilon}{1+\eta}} \left[\int_y \left(\bar{n}_y\right)^{\frac{\varepsilon(1+\eta)}{1+\varepsilon\eta}} \left\{\int_n \int_\sigma \left(\frac{n}{\sigma}\right)^\eta n dF_y(n,\sigma)\right\}^{\frac{1-\varepsilon}{1+\varepsilon\eta}} d\Gamma(y)\right]^{\frac{1+\varepsilon\eta}{1+\eta}}, \tag{2.27}$$

where the superscript Z denotes the zero information régime.

2.4 Comparison of the Information Régimes

In this subsection we rank information régimes with respect to optimal design capacity and efficiency. Régime A is said to be more efficient than régime B if

$$\overline{CS}^A(\hat{s}) > \overline{CS}^B(\hat{s}) \quad \forall\ \hat{s} > 0.$$

To rank efficiency, use is made of the following proposition (proved in Appendix 2):

PROPOSITION 2: Under A-1 to A-7, information régimes can be unambiguously ranked in terms of efficiency, *viz.* $\overline{CS}^A(\hat{s}) > \overline{CS}^B(\hat{s}) \Rightarrow \overline{CS}^A(\hat{s}') > \overline{CS}^B(\hat{s}')$ for all $\hat{s}'$. If $\varepsilon < 1$ (resp. $\varepsilon > 1$) the more efficient information régime has a lower (resp. higher) marginal consumers' surplus for *any* design capacity.

Since a lower marginal consumers' surplus implies a lower optimal design capacity, with $\varepsilon < 1$ (resp. $\varepsilon > 1$) the more efficient régime has the lower (resp. higher) design capacity. Both the efficiency and optimal design capacities of two information régimes can therefore be ranked by comparing the respective marginal consumers' surplus from capacity expansion. A relatively general ranking of information régimes is possible by using the idea of a *refinement*, defined as follows:

Definition Let I and I′ be two imperfect information régimes with respective message distributions $M(m|y)$ and $M'(m'|y)$ on days of type y. I′ is said to be a refinement of I if, for every message m from I, there is a corresponding set of messages $S(m)$ from I′, with more than one m′ in $S(m)$ for at least some m, such that m and $S(m)$ give the same probability distribution over (n,σ). Formally,

$$dF_{ym}(n,\sigma) = \int_{m'\in S(m)} dF_{ym'}(n,\sigma)\frac{dM'(m'|y)}{dM(m|y)} \qquad \forall\ n,\sigma,m,y. \tag{2.28}$$

Intuitively, I′ is more informative than I because it provides more messages. There are three limiting cases of (2.28) of interest.

(a) Imperfect information is a refinement of zero information:

$$dF_{y}(n,\sigma) = \int_{m'} dF_{ym'}(n,\sigma)\ dM'(m'|y)\ . \qquad \forall\ n,\sigma,y. \tag{2.28a}$$

(b) Perfect information is a refinement of zero information:

$$dF_{y}(n,\sigma) = 1\circ dF_{y}(n,\sigma) + 0(1 - dF_{y}(n,\sigma)) \qquad \forall\ n,\sigma,y. \tag{2.28b}$$

(c) Perfect information is a refinement of imperfect information:

$$dF_{ym}(n,\sigma) = 1\circ dF_{ym}(n,\sigma) + 0(1 - dF_{ym}(n,\sigma)) \qquad \forall\ n,\sigma,m,y. \tag{2.28c}$$

A ranking of information régimes is given by

PROPOSITION 3: Under A-1 to A-7, if information régime B is a refinement of régime A, then régime B is the more efficient. Furthermore, if $\varepsilon < 1$ (resp. $\varepsilon > 1$) then optimal design capacity is smaller (resp. larger) in régime B than régime A.

(Proof: See Appendix 3.) Prop. 3 establishes that if demand is inelastic, better information reduces the benefit from capacity expansion. There are two opposing forces at work. The fact that better information increases efficiency tends to lower the benefit of investment. But better information also reduces the efficiency loss from unpriced congestion. If $\varepsilon < 1$ the first factor dominates, and if $\varepsilon > 1$ the second does.

Prop. 3 is analogous to a result derived by Arnott *et al.* (1991b). There, we applied the Vickrey queueing model (see Section 3) in which users choose when as well as whether to use a facility. For a deterministic setting in which capacity and demand are nonstochastic we considered four pricing régimes; in order of increasing efficiency: no toll, a time-invariant toll, a step-function toll and a continuously time-varying toll. We showed that optimal design capacity decreases with the efficiency of the pricing régime if $\varepsilon < 1$, and *vice versa* if $\varepsilon > 1$.

Pricing in that paper serves a role similar to information here as a means of improving the efficiency with which a given facility is used. Pricing and better information both reduce the return from capacity expansion when demand is inelastic, and increase it when demand is elastic.

2.5 Robustness

We have assumed that user costs are homogeneous of degree zero in demand and capacity, and that demand is isoelastic. If either assumption is relaxed, Props. 1-3 no longer hold. To illustrate this lack of robustness, we shall show that consumers' surplus may be lower with perfect information than with zero information if either assumption is relaxed. Let N(p) be an arbitrary demand curve, and AC(N,s) the user cost curve. Using (2.24), the welfare change in going from zero to perfect information can be written as

$$\overline{CS}^P(\hat{s}) - \overline{CS}^Z(\hat{s}) = \int_y \int_n \int_\sigma \left\{ \int_{p(n,\sigma)}^{\bar{p}^u_y} N(p)dp \right\} dF_y(n,\sigma)\; d\Gamma(y), \tag{2.29}$$

where $\bar{p}^u_y$ is users' expected price on day y given no message. Consider a specific example in which there are no day-specific fluctuations, demand intensity is fixed, and capacity is either higher, s_H, or low, s_L, with equal probability. Let $p(s_H)$ and $p(s_L)$ denote the perfect-information prices with s_H and s_L. Then for this example, (2.29) reduces to

$$\overline{CS}^P - \overline{CS}^Z = \frac{1}{2}\left[\int_{p(s_H)}^{\bar{p}^u} N(p)dp - \int_{\bar{p}^u}^{p(s_L)} N(p)dp \right].$$

With zero information, the user will experience the two states with equal probability. Thus, $\bar{p}^u$ is calculated as that point on the demand curve which is vertically half-way between $AC(N(\bar{p}^u),s_L)$ and $AC(N(\bar{p}^u),s_H)$ - see Fig. 1. $\overline{CS}^P - \overline{CS}^Z$ equals one-half the difference between the diagonally shaded area and the vertically shaded area in Figure 1. In panel (a), with homogeneous costs and isoelastic demand this difference is positive *per* Prop. 3. But in panel (b), with nonhomogeneous costs, and in panel (c), with non-constant demand elasticity, the difference as drawn is negative. In both cases it is the shapes of the demand curve and the average cost curve in the neighborhood of the zero information equilibrium that determine how information affects welfare. This negative result is of sufficient interest that we state it as:

PROPOSITION 4: If either A-3 or A-4 is relaxed, zero information may be more efficient than perfect information.

The welfare effect of perfect information in the example is illustrated another way in Figure 2 (the average cost curves are made linear for simplicity). The marginal cost curves for the high and low capacity states are $MC(N,s_H)$ and $MC(N,s_L)$. With zero information, usage is $N(\bar{p}^u)$. Deadweight loss from overuse in the two capacity states is indicated by the heavily bordered areas (underusage is possible with high capacity). Given perfect information, usage expands to N_H when high capacity occurs and falls to N_L when low capacity occurs. The corresponding efficiency loss and efficiency gain relative to zero information are indicated by the vertically and diagonally shaded areas. Again, the shapes of the cost curves and the demand curve determine whether the net gain from information is positive or negative.

Two comments are worth making.

1. While Props. 1-3, as well as equations (2.15), (2.26), and (2.27), rely on assumptions A-3 and A-4, the method of analysis is general and can be applied to determine optimal capacity even when those assumptions do not hold.

2. Despite their lack of robustness, Props. 1-3 and equations (2.15), (2.26), and (2.27) are of interest. First, the results they provide are useful benchmarks. Second, in most practical contexts, there are insufficient data to determine whether the demand function is isoelastic or whether the user-cost functions has the form assumed in A-4. The propositions and formulae

given here could be applied in these contexts once estimates of the (local) demand and cost elasticities are obtained.

3. Dynamic Equilibrium

In the preceding section, we examined the effects of uncertainty and information in a static model of a congestible facility, in which individuals decided only whether to use the facility. In this section, we present a dynamic model in which individuals have both a participation and time-of-use decision. One might suppose, as indeed the traditional peak-load literature has, that this extension entails no more than adding time subscripts and generalizing the demand function to treat use of the facility at different points in time as different commodities. So supposing, one would conjecture that the Propositions of the previous section extend when time of use is considered. It turns out, however, that some parts of the Propositions do not extend. The fallacy in the traditional reasoning stems from neglect of the relationship between congestion at different points in time within the period of use.

To keep the analysis manageable, we shall employ an even more particular model than we did in the previous subsection. We shall extend a variant of Vickrey's (1969) bottleneck queueing model.

The facility is assumed to have a bottleneck with a maximum service rate of s users per unit time. (In the commuting context the facility is a road or freeway, and the bottleneck a bridge, tunnel, lane drop, *etc.*) If the arrival rate of users at the bottleneck exceeds s, a queue develops. The time taken to use the facility for a user arriving at time t is

$$T(t) = \bar{T} + \frac{Q(t)}{s},$$

where $\bar{T}$ is usage time in the absence of a queue. Without loss of generality we assume $\bar{T}$ is zero. Q(t) is the length of the queue, equal to

$$Q(t) = \int_{\hat{t}}^{t} r(\tau)d\tau - s(t-\hat{t}), \tag{3.1}$$

where $\hat{t}$ is the time at which the queue was last zero, and r is the arrival rate at the bottleneck.[18] Queueing time is

$$q(t) \equiv \frac{Q(t)}{s}. \tag{3.2}$$

Time of use matters. For example, commuters typically incur penalties from arriving at work after the official starting time. They also prefer not to arrive very early, because doing so entails an early morning rise, wasted time before work begins, and difficulties in scheduling household and leisure activities.

To model time-of-use preferences, we assume that individuals have a common preferred time, t^*, for being served - that is, for *ending* usage of the facility, and incur a 'schedule delay cost' (SDC) if served at a different time.[19] For an individual arriving at the facility at time t, user cost, the sum of queueing and schedule delay cost, is assumed to be

$$C(t) = \alpha q(t) + \beta \, \mathrm{Max}[0, t^* - t - q(t)] + \gamma \, \mathrm{Max}[0, t + q(t) - t^*]. \tag{3.3}$$

The parameter α measures the unit cost (time cost and money cost) of queueing time. β is the unit cost of being early, and γ the unit cost of being late. To admit differences in the cost of earliness and lateness we allow $\beta \neq \gamma$; for morning commuters with fixed work hours one expects $\beta << \gamma$.

3.2 Equilibrium with perfect information

In deciding when to use a facility, individuals face a trade-off between queueing time and schedule delay. With perfect information they know the distribution of arrival times. In equilibrium, no one can reduce his user cost by arriving at a different time. With identical individuals this means that costs are constant during the period when usage occurs.

Equilibrium for the Vickrey model in a deterministic setting, which is equivalent to that under perfect information, has been exposited elsewhere (*e.g.* Hendrickson and Kocur (1981), Arnott *et al.* (1990a, 1991b)). The equilibrium is shown in Figure 3. The number of individuals in the queue is measured by the vertical distance between the piecewise linear schedule depicting cumulative arrivals at the facility, and the straight line with slope s depicting the cumulative number served. Arrivals occur over the time interval $[t_0, t_1]$. A derivation of the equilibrium arrival rate and number of users follows.

Let t_n be the arrival time for which an individual gets served at t^*. During $[t_0, t_n)$, individuals are served early. Setting the derivative of (3.2) with respect to t equal to zero, one obtains for the equilibrium arrival rate:

$$r(t) = \frac{\alpha s}{\alpha - \beta}, \qquad t \in (t_0, t_n). \tag{3.4}$$

This is positive and finite if $\alpha > \beta$, which we assume.[20]

Over the arrival interval (t_n, t_1), individuals are served late. Setting $\dot{C}(t) = 0$ as before, one gets

$$r(t) = \frac{\alpha s}{\alpha + \gamma}, \qquad t \in (t_n, t_1), \tag{3.5}$$

which is positive and finite.

The first user to arrive naturally encounters no queue, and suffers only the cost of being served early. The last user also escapes queueing. (If he did encounter a queue, he could arrive slightly later and reduce queueing time without any change in schedule delay cost.) Equating the schedule delay costs of the first and last users:

$$\beta(t^* - t_0) = \gamma(t_1 - t^*). \tag{3.6}$$

Since the bottleneck operates at capacity throughout (t_0, t_1),

$$t_1 = t_0 + \frac{N}{s}. \tag{3.7}$$

Solving (3.6) and (3.7) for t_0, t_1 and the equilibrium user cost, C:

$$t_0 = t^* - \frac{\delta}{\beta}\frac{N}{s}, \qquad t_1 = t^* + \frac{\delta}{\gamma}\frac{N}{s}, \tag{3.8}$$

$$C = \delta \frac{N}{s}, \tag{3.9}$$

where $\delta \equiv \frac{\beta\gamma}{\beta+\gamma}$. Comparing (3.9) with (2.1), it is evident that the equilibrium cost, for a given number of users, is as given by (2.11) for the static model, with $\delta = \xi$ and $\eta = 1$:

$$p^e(n,\sigma) = \left(\frac{\delta n}{\hat{s}\sigma}\right)^{\frac{1}{1+\varepsilon}}. \tag{3.10}$$

The equilibrium number of users can hence be solved as in Section 2.2, and the analysis of that section, including Prop. 1, goes through for the dynamic model when users have perfect information.[21] This result is of sufficient interest and importance that we record it as

> PROPOSITION 5: Under perfect information, equilibrium in the dynamic model reduces to that in the static model. Hence, Prop. 1 extends to the dynamic model.

3.3 Equilibrium with imperfect information

In this subsection, we solve for the dynamic equilibrium under uncertainty. In the next subsection, we establish that imperfect information can raise expected user costs, even though perfect information is necessarily welfare-improving. In this sense there is a nonconcavity in the value of information. This result stands in contrast to that of the static model, where both perfect and imperfect information improve efficiency (Prop. 3).

In the imperfect information régime, at the time they make their participation decision, individuals face uncertainty about the capacity and/or the number of users. The equilibrium arrival schedule equalizes *expected* costs, conditional on the information available. User costs will not be equal on a particular occasion *ex post*. Equilibrium is more difficult to compute than under perfect information for three reasons.

i) The evolution of user costs over the period of use must be computed for every n and σ, conditional on m and y, and for every m and y. The equal expected user cost condition with message m on day y, $\bar{C}_{ym}$, is

$$\bar{p}_{ym} = \bar{C}_{ym} = \int_n \int_\sigma \tilde{C}_{ym}(t,n,\sigma)\ dF^u(n,\sigma) \quad \text{for all } t \in T_{ym},$$

where $\tilde{C}_{ym}(t,n,\sigma)$ is user cost at time t, conditional on y and m, and with realizations n and σ, and T_{ym} is the set of arrival times in equilibrium with y and m.

ii) In computing a particular user cost function, $\tilde{C}_{ym}(t,n,\sigma)$, account needs to be taken that the bottleneck may not operate at capacity throughout the arrival period, and that there may be a queue at the end of the arrival period.

iii) Given y and m, an individual who arrives at the facility before t^* may be served early some days, and late others, depending on the realizations of n and σ.

We shall characterize equilibrium in terms of a set of *normalized* arrival rates, $\rho_{ym}(t)$, where $\rho_{ym}(t)dt$ is the fraction of users who arrive between t and t+dt on day y with message m.[22] There is a separate equilibrium condition for each y and m. To economize on notation we suppress y and m subscripts for the remainder of this subsection.

To derive equilibrium, some additional notation is required. Let t_0 denote the time of the first arrival, t_L the time of the last arrival, and t_1 the time at which the last arrival is served. If the last arrival does not face a queue, $t_1 = t_L$; if he does face a queue, $t_1 > t_L$. Let $R(t) \equiv \int_{t_0}^{t} \rho(t')dt'$ be the normalized cumulative arrival function; clearly, $R(t_0) = 0$ and $R(t_L) = 1$.

Now consider two separate occasions, with the same y and m, for which the realizations (n_1, σ_1) and (n_2, σ_2) are such that $\frac{n_1}{\sigma_1} = \frac{n_2}{\sigma_2}$. Since expected price is the same on both occasions, $\frac{N_1}{s_1} = \frac{N_2}{s_2}$. From (3.1) and (3.2), $q_1(t) = \int_{\hat{t}}^{t} \frac{r_1(\tau)}{s_1}d\tau - (t-\hat{t})$. Since $r_1(\tau) = N_1\rho(\tau)$, $q_1(t) = \frac{N_1}{s_1}\int_{\hat{t}}^{t} \rho(\tau)d\tau - (t-\hat{t})$. Furthermore, since $\frac{N_1}{s_1} = \frac{N_2}{s_2}$, queueing time and queueing costs evolve in the same way on both occasions. So too do schedule delay costs, and hence user costs. The fact that user cost depends only on the *ratio* of N to s simplifies the analysis.

Define $\phi \equiv \frac{N}{s}$. Given the assumption that $\frac{n}{\sigma}$ has a finite upper bound and is positive with nonzero probability, ϕ has a finite upper bound, ϕ^{Max}. Let ϕ^{Min} be its lower bound. Denote the c.d.f. of ϕ, from the perspective of users, by $J(\phi)$. Since $\phi = \frac{n(\bar{p}^u)^{-\varepsilon}}{\sigma\hat{s}}$

$$J(\phi) = \int_{\sigma} \int_{n=0}^{\phi\sigma\hat{s}(\bar{p}^u)^{\varepsilon}} dF^u(n,\sigma).$$

Define $q(t,\phi)$ to be queueing time at time t with ϕ, and $C(t,\phi)$ to be the corresponding user cost.

Two critical values of ϕ will feature prominently. First, at each t there is a largest ϕ such that no queueing occurs. Denote this by

$$\phi^Q(t) \equiv \mathrm{Max}\{\phi | q(t,\phi) = 0\}, \text{ all } t. \tag{3.11}$$

$\phi^Q(t)$ describes the *queueing boundary* in t-ϕ space. Second, at each t for which both early and late arrival is possible, there is a smallest ϕ above which individuals arrive late. Denote this by $\phi^*(t)$, defined by the condition

$$t + q(t,\phi^*(t)) = t^*, \ t < t^*. \tag{3.12}$$

$\phi^*(t)$ describes the *on-time arrival locus*. After $t_n \equiv \phi^{*-1}(\phi^{Max})$, users are sometimes served late. We shall assume that $\phi^{Min} < \phi^Q(t^*)$, so that t^* is the last arrival time for which users can possibly avoid being late.

The qualitative characteristics of equilibrium are shown in Figure 4.[23] There are 4 regions in t-ϕ space, defined by whether or not users queue, and whether they are served early or late. In regions NE and QE, users are served early; in regions NL and QL, they are served late. In regions NE and NL they do not queue, whereas in QE and QL they do. If there is no queue, a user is served early if $t < t^*$, and late if $t > t^*$. Regions NE and NL thus abut at $t = t^*$. The boundary between regions QE and QL is defined by the on-time arrival locus $\phi^*(t)$. Regions NE and NL are separated from QE and QL by the queueing boundary $\phi^Q(t)$. $\phi^Q(t) < \phi^{Max}$ for all $t \in (t_0, t_L)$, since otherwise the last user would never queue, and could reduce his cost by arriving earlier. Moreover, $\phi^*(t) > \phi^Q(t)$ for $t < t^*$, and $\phi^*(t) < \phi^{Max}$ for $t > t_n$.

Expected costs in each region are:

$$C^{NE}(t,\phi) = \beta(t^*-t), \tag{3.13a}$$

$$C^{QE}(t,\phi) = \beta(t^* - t - q(t,\phi)) + \alpha q(t,\phi), \tag{3.13b}$$

$$C^{NL}(t,\phi) = \gamma(t-t^*), \tag{3.13c}$$

$$C^{QL}(t,\phi) = \gamma(t + q(t,\phi) - t^*) + \alpha q(t,\phi). \tag{3.13d}$$

As is clear from Figure 4, there are three arrival intervals to consider:

A. $t \in (t_0, t_n)$ in which users are always served early,

B. $t \in (t_n, t^*)$ in which, depending on ϕ, users may be served early or late,

C. $t \in (t^*, t_L)$ in which users are always served late.

In equilibrium, C(t) must be independent of t within each of these intervals, and also across intervals.

$\boxed{\text{A. } t \in (t_0,\ t_n)}$:

To see that this interval is non-empty, note that the first user must arrive before t^*, since otherwise he could arrive at t^* and escape both queueing and schedule delay, which is not possible for all users. By continuity of R(t) there is a nonzero time interval over which users are always served early.

Given (3.13a), (3.13b) and Figure 4, expected costs are

$$C(t) = \beta\left\{ \int_0^{\phi^Q(t)} (t^*-t)dJ(\phi) + \int_{\phi^Q(t)}^{\phi^{Max}} (t^*-t-q(t,\phi))dJ(\phi) \right\}$$

$$+\alpha \int_{\phi^Q(t)}^{\phi^{Max}} q(t,\phi)dJ(\phi)$$

$$= \beta(t^*-t) + (\alpha-\beta) \int_{\phi^Q(t)}^{\phi^{Max}} q(t,\phi)dJ(\phi). \tag{3.14}$$

Given $Q(t,\phi^Q(t)) = 0$,

$$\dot{C}(t) = -\beta + (\alpha-\beta) \int_{\phi^Q(t)}^{\phi^{Max}} \dot{q}(t,\phi)dJ(\phi).$$

Using $\dot{q}(t,\phi) = \dfrac{r(t)-s}{s} = \dfrac{N\rho(t)}{s} - 1 = \phi\rho(t) - 1$ for $\phi > \phi^Q(t)$, one can derive as a condition for equilibrium

$$0 = \dot{C}(t) = \alpha Z^Q(t) - \beta(1+Z^Q(t)), \tag{3.15}$$

where

$$Z^Q(t) \equiv \int_{\phi^Q(t)}^{\phi^{Max}} (\phi\rho(t) - 1)dJ(\phi). \tag{3.16}$$

$Z^Q(t)$ is the rate of increase wrt t in expected queueing time. A user who postpones arrival by dt incurs an increase in expected queueing time cost of $\alpha Z^Q(t)dt$, which explains the first term on the RHS of (3.15). Since users are also served early, arriving dt later decreases expected schedule delay cost by $\beta(1+Z^Q(t))dt$, which explains the second term on the RHS of (3.15). In equilibrium the two terms must balance. Appendix 4 establishes that (3.15) defines a constant value of $\rho(t)$. The departure rate is thus constant over (t_0, t_n).

$\boxed{\text{B. } t \in (t_n, t^*)}$:

In this arrival interval, users are sometimes served early and sometimes late. Given (3.13a), (3.13b), (3.13d) and Figure 4, expected cost is

$$C(t) = \beta\left\{ \int_{0}^{\phi^Q(t)} (t^*-t)dJ(\phi) + \int_{\phi^Q(t)}^{\phi^*(t)} (t^*-t-q(t,\phi))dJ(\phi) \right\}$$
$$+ \gamma \int_{\phi^*(t)}^{\phi^{Max}} (t + q(t,\phi) - t^*)dJ(\phi) + \alpha \int_{\phi^Q(t)}^{\phi^{Max}} q(t,\phi)dJ(\phi). \tag{3.17}$$

The equilibrium condition is

$$0 = \dot{C}(t) = \alpha Z^Q(t) - \beta Z^E(t) + \gamma Z^L(t), \tag{3.18}$$

where

$$Z^E(t) \equiv \int_{\phi^Q(t)}^{\phi^*(t)} (\phi\rho(t) - 1)dJ(\phi) + J(\phi^*(t)) \tag{3.19}$$

is the rate of decrease in expected early time and

$$Z^L(t) \equiv \int_{\phi^*(t)}^{\phi^{Max}} (\phi\rho(t) - 1)dJ(\phi) + 1 - J(\phi^*(t)) \tag{3.20}$$

is the rate of increase in expected late time. Equation (3.18) has an interpretation analogous to (3.15). Appendix 4 establishes that $\rho(t)$ is weakly decreasing for $t \in (t_n, t^*)$.

$\boxed{\text{C. } t \in (t^*, t_L)}$ (if $t_L = t^*$ this interval is degenerate)

In this arrival interval, users are always served late. Given (3.13c), (3.13d) and Figure 4, expected costs are

$$C(t) = \gamma\left\{ \int_{0}^{\phi^Q(t)} (t-t^*)dJ(\phi) + \int_{\phi^Q(t)}^{\phi^{Max}} (t+q(t,\phi) - t^*)dJ(\phi) \right\}$$

$$+ \alpha \int_{\phi^Q(t)}^{\phi^{Max}} q(t,\phi)dJ(\phi),$$

$$= \gamma(t-t^*) + (\alpha+\gamma) \int_{\phi^Q(t)}^{\phi^{Max}} q(t,\phi)dJ(\phi). \tag{3.21}$$

Equilibrium requires

$$0 = \dot{C}(t) = \alpha Z^Q(t) + \gamma(1+Z^Q(t)). \tag{3.22}$$

(3.22) has an interpretation analogous to (3.15). Again, $\rho(t)$ is weakly decreasing on the interval (see Appendix 4).

Conditions (3.15), (3.18) and (3.22) ensure that expected costs are constant in each arrival interval. Since $\phi^*(t_n) = \phi^{Max}$, and $\phi^Q(t^*) = \phi^*(t^*)$, expected costs are also continuous at t_n and t^*, and hence constant over the whole interval $[t_0,t_L]$. Since $\rho(t)$ is constant on (t_0,t_n) and weakly decreasing thereafter:

PROPOSITION 6: The normalized arrival rate $\rho(t)$ is weakly decreasing over the arrival period, and the cumulative arrivals distribution R(t) is concave.

The intuition for Prop. 6 is as follows: After t_n, the later users arrive, the more likely they are to be served late, and to suffer increased lateness from marginally postponing arrival, rather than reduced earliness. To compensate, expected queueing costs must grow at a decreasing rate and eventually decline. This requires that the arrival rate decrease over time. Prior to t_n, since users are always served early, expected queueing cost must increase linearly with t (*viz.* 3.15), which implies a constant arrival rate.

Given concavity of R(t), it follows that if queueing occurs on a given occasion it is over a continuous time interval beginning at t_0. Queueing time is

$$q(t,\phi) = \text{Max}\ [0,\ t_0 + \phi R(t) - t]. \tag{3.23}$$

t_n is defined implicitly by the condition

$$t_0 + \phi^{Max}R(t_n) = t^*,$$

while

$$\phi^Q(t) = \frac{t-t_0}{R(t)}, \tag{3.24}$$

$$\phi^*(t) = \frac{t^*-t_0}{R(t)}. \qquad (3.25)$$

With regard to the timing of arrivals, it turns out that $t_L \geq t^*$. To see that equilibrium cannot entail $t_L < t^*$, suppose the contrary. If the last arrival postpones his arrival to $t \in (t_L, t^*)$, his user cost falls; if $\phi < \phi^Q(t)$, he faces no queue and is served less early, while if $\phi > \phi^Q(t)$ his schedule delay cost is the same, but his queueing cost falls.[24] This leads to:

PROPOSITION 7: Given A-1 to A-7, in user equilibrium with imperfect information the last user arrives no earlier than t^*.

To describe the equilibria with $t_L > t^*$ and $t_L = t^*$ some further notation is required. Define:

$$\bar{\phi} \equiv \int \phi dJ(\phi),$$

the $\left(\frac{\alpha}{\alpha+\gamma}\right)$ fractile of the J( ) distribution:

$$\tilde{\phi} \equiv J^{-1}\left(\frac{\alpha}{\alpha+\gamma}\right),$$

and the mean of ϕ for $\phi \geq \tilde{\phi}$:

$$\hat{\phi} \equiv \int_{\tilde{\phi}} \phi dJ(\phi) \,/(1-J(\tilde{\phi})) = \frac{\alpha+\gamma}{\gamma}\int_{\tilde{\phi}} \phi dJ(\phi) > \bar{\phi}. \qquad (3.26)$$

The case $t_L > t^*$ is described in Lemma 1 (for the proof, see Appendix 5)

LEMMA 1:

$$\text{If } \tilde{\phi} > \frac{\gamma}{\beta+\gamma}\hat{\phi} \qquad (3.27)$$

then the first and last arrivals occur at

$$t_0 = t^* - \frac{\delta}{\beta}\hat{\phi}, \qquad (3.28)$$

$$t_L = t_0 + \tilde{\phi} > t^*. \tag{3.29}$$

Expected costs are

$$C = C(t_0) = \beta(t^* - t_0) = \delta\, \phi', \tag{3.30}$$

where $\phi' = \hat{\phi}$.

A heuristic derivation of (3.29) goes as follows. Suppose the last user arrives at $t_L = t_0 + \phi_L$, for some ϕ_L. The user escapes queueing if $\phi \le \phi_L$, which occurs with probability $J(\phi_L)$. If he delays arrival by dt, his time late costs increase by γdt with probability $J(\phi_L)$, while his queueing time cost decreases by αdt with probability $1 - J(\phi_L)$. In equilibrium these expected costs must balance; hence $\phi_L = J^{-1}\left(\frac{\alpha}{\alpha+\gamma}\right) = \tilde{\phi}$, and $t_L = t_0 + \tilde{\phi}$, which is (3.29).

Note that expected user costs, $C = \delta\hat{\phi}$, are by (3.30) and (3.26) independent of $J(\phi)$ for $\phi < \tilde{\phi}$. To see why, consider the last user again. If $\phi > \tilde{\phi}$ he encounters a queue and incurs a cost that rises with ϕ. But if $\phi \le \tilde{\phi}$ he escapes queueing and is served immediately; his costs are independent of the precise value of ϕ realized. In either case his expected costs are independent of the probability distribution of ϕ below $\tilde{\phi}$. Since in equilibrium all users incur the same expected cost, C is also independent of the ϕ distribution below $\tilde{\phi}$.

If (3.27) is not satisfied, equilibrium is described by Lemma 2 (for the proof, see Appendix 6):

LEMMA 2:

$$\text{If } \tilde{\phi} \le \frac{\gamma}{\beta+\gamma}\, \hat{\phi} \tag{3.31}$$

then t_0 is defined implicitly by the equation

$$J(t^*-t_0) + \frac{1}{t^*-t_0} \int_{t^*-t_0} \phi dJ(\phi) - \frac{\alpha+\beta+\gamma}{\alpha+\gamma} = 0, \tag{3.32}$$

$$t_L = t^*,$$

and

$$C = C(t_0) = \beta(t^*-t_0) = \delta\phi', \tag{3.33}$$

where

$$\phi' \equiv \frac{\beta+\gamma}{\gamma}(t^*-t_0) \tag{3.34}$$

By (3.32) the equilibrium is again independent of the left-hand tail of the ϕ distribution.

Lemmas 1 and 2 lead to:

PROPOSITION 8: Given any y, expected user costs with imperfect information are bounded by the highest and lowest costs realizable under perfect information:

$$\delta\phi_y^{Min} \le \delta\phi'_{ym} \le \delta\phi_y^{Max}, \quad \forall\ y,\ m,$$

where ϕ_y^{Min} and ϕ_y^{Max} are respectively the minimum and maximum points of the support of $F_y(n,\sigma)$.

(Proof: See Appendix 7.) In this sense, equilibrium costs with perfect information bound equilibrium costs under imperfect information. Lemmas 1 and 2 also lead to:

PROPOSITION 9: For any y, expected user costs are greater with imperfect information than with perfect information:

$$\int_m \delta\phi'_{ym}\ dM^u(m|y) \ge \int_\phi \delta\phi\ dJ_y(\phi), \quad \forall\ y.$$

(Proof: See Appendix 8.) Since the frequency distribution of y is independent of the information régime it follows as a corollary that expected user costs are greater with imperfect information than with perfect information for the whole frequency distribution of days.

3.4 Comparison of the Information Regimes

Prop. 2 of Section 2 concerned the relationship between consumers' surplus and marginal consumers' surplus for the static model. The Proposition carries over without alteration to the dynamic model (Proof: see Appendix 9). For completeness we repeat the proposition here as

PROPOSITION 2′: In the dynamic model information régimes can be unambiguously ranked in terms of efficiency, *viz.* $\overline{CS}^A(\hat{s}) > \overline{CS}^B(\hat{s}) \Rightarrow \overline{CS}^A(\hat{s}') > \overline{CS}^B(\hat{s}')$ for all $\hat{s}'$. If $\varepsilon < 1$ (resp. $\varepsilon > 1$) the more efficient information régime has the lower (resp. higher) marginal consumers' surplus for *any* design capacity.

In place of Prop. 3 we have a weaker result (Proof: see Appendix 10).

PROPOSITION 3′: In the dynamic model, perfect information increases efficiency relative to imperfect information. Furthermore, if $\varepsilon < 1$ (resp. > 1) then optimal design capacity is smaller (resp. larger) with perfect information than with imperfect information.

Prop. 3′ is weaker than Prop. 3 in that better information is not always welfare-improving. In particular, imperfect information is not necessarily welfare-improving relative to zero information.

We now provide an example in which imperfect information is welfare-reducing. Suppose that there are no systematic fluctuations (the subscript y is redundant) and that demand intensity is fixed, but capacity fluctuates with an (unconditional) p.d.f.

$$s = \begin{cases} \hat{s} \text{ with probability } 1-\pi \\ \\ \sigma\hat{s} \text{ with probability } \pi, \quad \sigma < 1. \end{cases} \tag{3.35}$$

In the case of commuting, π can be interpreted as the probability of events (*e.g.* road repairs, bad weather, accidents) that reduce freeway capacity below its design level for the duration of the morning commute. In Appendix 11 we show that, if $\sigma < \dfrac{\gamma}{\beta+\gamma}$, then $(\bar{p}^u)^{1-\varepsilon}$ varies with π as shown in Figure 5.[25]

Now, with n fixed, no systematic fluctuations, and $\eta = 1$, we have by (2.25)

$$\overline{MCS}^I(\hat{s}) - \overline{MCS}^Z(\hat{s}) = \frac{n}{(1+\varepsilon)\hat{s}}\left[\int_m \left(\bar{p}^u_m\right)^{1-\varepsilon} dM(m) - \left(\bar{p}^u\right)^{1-\varepsilon}\right]. \tag{3.36}$$

By Prop. 2

$$\overline{CS}^I(\hat{s}) \gtrless \overline{CS}^Z(\hat{s}) \text{ if } \left[\overline{MCS}^I(\hat{s}) - \overline{MCS}^Z(\hat{s})\right](\varepsilon - 1) \gtrless 0. \tag{3.37}$$

With zero information, the probability of reduced capacity from users' perspective is π every day. With perfect information, it is 0 with probability $1-\pi$, and 1 with probability π. It is clear from Figure 5 that (3.36) is negative if $\varepsilon < 1$, and positive if $\varepsilon > 1$. By (3.37), perfect information is thus welfare-improving, as proved generally in Prop. 3′.

With imperfect information, the probability of low capacity conditional on the message received fluctuates around π. At $\pi_c \equiv \dfrac{\beta}{\alpha+\gamma}\dfrac{\sigma}{1-\sigma}$, there is a kink in the function $\left(\bar{p}^u(\pi)\right)^{1-\varepsilon}$. Suppose $\pi \cong \pi_c$, and messages are not very informative, in the sense that the probability of capacity reduction conditional on a message rarely differs much from π. By the unbiasedness of the message system and Jensen's inequality, (3.36) is then *positive* if $\varepsilon < 1$ and *negative* if $\varepsilon > 1$. Imperfect information is then welfare-reducing; there is a nonconcavity in the value of information.[26] This is the inspiration for the epigraph at the head of the paper: "A little learning is a dangerous thing".[27]

The local nonconcavity is generated by a régime change. For (perceived) $\pi < \frac{\beta}{\alpha+\gamma}\frac{\sigma}{1-\sigma}$, the last arrival always arrives late. For π somewhat larger than $\frac{\beta}{\alpha+\gamma}\frac{\sigma}{1-\sigma}$, the constraint that the last arrival cannot arrive early becomes operational, which increases the slope of $\frac{\partial \bar{p}^u}{\partial \pi}$. It is evident from this explanation that the result is not due to the linearity of the schedule delay cost function.[28]

The fact that imperfect information necessarily increases welfare in the static model, but can decrease it in the dynamic model, illustrates the importance of modeling all margins of user behavior. Neglect of users' time-of-use decisions can lead to qualitative errors in analysis, as well as overlooking of interesting and counterintuitive behavior.

4. Summary and Concluding Remarks

Most congestible facilities are subject to nonsystematic fluctuations in demand and/or capacity. In the case of transportation, recreational areas, telecommunications and other facilities, it may be possible to obtain real-time or near real-time information on the state of the system. Where the provision of information is a policy decision, as will presumably be the case with advanced traveler information systems for automobiles, it is important that the welfare effects of information be understood.

In this paper we investigated the effect of information on usage of a congestible free-access facility subject to stochastic demand and capacity. We considered three information régimes -- zero information, perfect information, and imperfect information which contains as limiting cases the other two régimes. We began with a static model in which only the aggregate level of usage is endogenous. We assumed that user cost is homogeneous of degree zero in usage and capacity, and that demand is isoelastic. We then considered a dynamic model in which individuals' times of usage as well as their participation decisions are endogenous. In the case of perfect information the reduced-form equilibrium cost function of the dynamic model is a special case of the static model.

For the static model with homogeneous costs and isoelastic demand, both imperfect and perfect information are welfare-improving relative to zero information. Optimal capacity is lower with information than without it if and only if the elasticity of demand is less than unity. Perfect information is welfare-improving relative to zero information for the dynamic model, but imperfect information need not be. And neither perfect nor imperfect information need be welfare-improving in either model for arbitrary user cost and demand functions. These results suggest that a cautious approach should be adopted in designing and implementing information dissemination schemes, at least absent tolls or other means of regulating usage.

There are several directions in which the analysis could be extended.

1. Perhaps the most important is to consider the use of pricing in conjunction with information as a policy tool. In the case of the bottleneck queueing model with perfect information it has been shown (*viz.* Vickrey (1969), Arnott *et al.* (1990a, 1991b)) that time-varying tolls can provide substantial

efficiency gains through induced changes in time-of-usage decisions. It remains to be seen how large these gains are in a stochastic environment, and how they vary with the quality of information.[29]

2. A single facility in isolation has been considered. This is unsatisfactory if the facility is part of a system, such as one road in a corridor with several roads, or a ski area close to competing facilities. Information may affect users' choice of facility as well as their decision whether to participate and when. Furthermore, with unpriced congestion, capacity expansion or other policies adopted at one facility will affect the efficiency loss due to congestion elsewhere in the system.[30]

3. In the bottleneck model, queue discipline is FIFO. This is reasonable for transportation and some recreational facilities. Other facilities have different congestion technologies and service disciplines. With time sharing, computer users are served quasi-simultaneously. On a telephone network, capacity is determined by the number of telephone lines, while the completion rate of calls is governed by the average length of a call. Too, in some exchanges users gain access to a line at random rather than in order of placing a call. In other facilities the quality of service may degrade with cumulative usage; *e.g.* due to buildup of heat in a museum, or wear and tear on nature trails over the course of a season. It remains to be seen whether the welfare effects of information are sensitive to these considerations.

4. The joint probability distribution of demand intensity (n) and capacity availability (σ) has been assumed independent of design capacity ($\hat{s}$). In practice, there may be a tradeoff between the utilization rate of a facility and the probability or magnitude of loss of capacity. For example, if highway

shoulders are used as travel lanes during peak hours, which effectively increases design capacity, they will be unavailable for stalled vehicles and rescue equipment, so that the average capacity utilization rate may fall. Furthermore, σ may depend on demand, as is the case for road traffic to the extent that the probability of incidents increases with traffic volume (Newbery (1990)). And demand may be a function of service reliability (Kay (1979), Saving and DeVany (1981), Coate and Panzar (1988)). A definition of capacity ought thus to include a reliability coefficient, which would be a factor in determining optimal design capacity, as well as maintenance and repair policy.

5. The facility has been assumed to be a delay system, in which users are neither turned away nor balk when congestion is heavy. This is reasonable for travel because of the substantial sunk costs and commitments generally involved. With other facilities, such as the telephone, the cost of attempting usage may be quite low, so that users can balk and try again later.

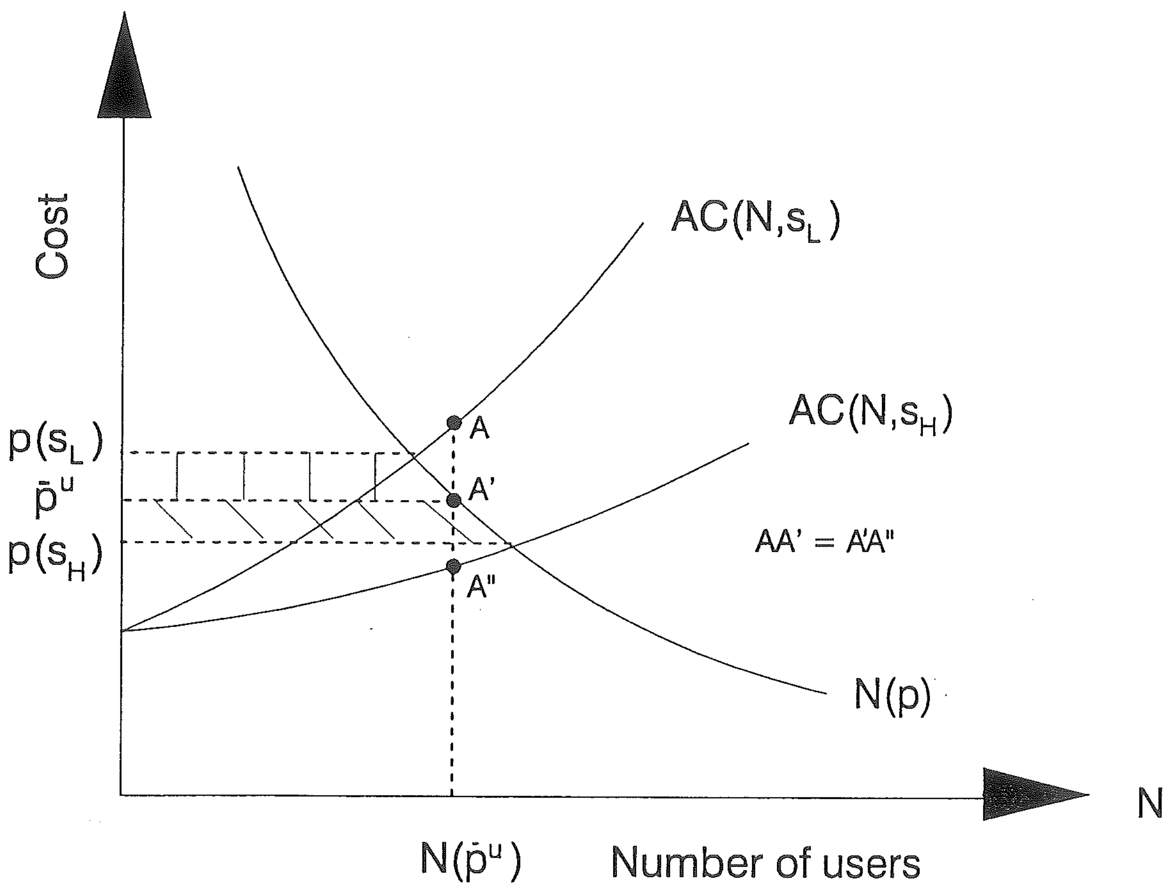
6. It has been assumed that all users have the same information. For various reasons some individuals tend to be better informed than others: greater experience, preferential access to information, superior cognitive capabilities, *etc.* In the case of automobile travel it has been theorized, and observed in simulation models, that the marginal effect on equilibrium of information diminishes with the fraction of drivers who are informed. Similarly, Haltiwanger and Waldman (1985) have shown that informed decision makers have an impact disproportionate to their fraction of the population in determining equilibrium in congestible systems. Given this, the fact that the *private* benefit to a driver from becoming informed generally differs from the

social benefit, and the high costs of installing and operating information systems in individual vehicles, the fraction of drivers to equip becomes a policy issue.

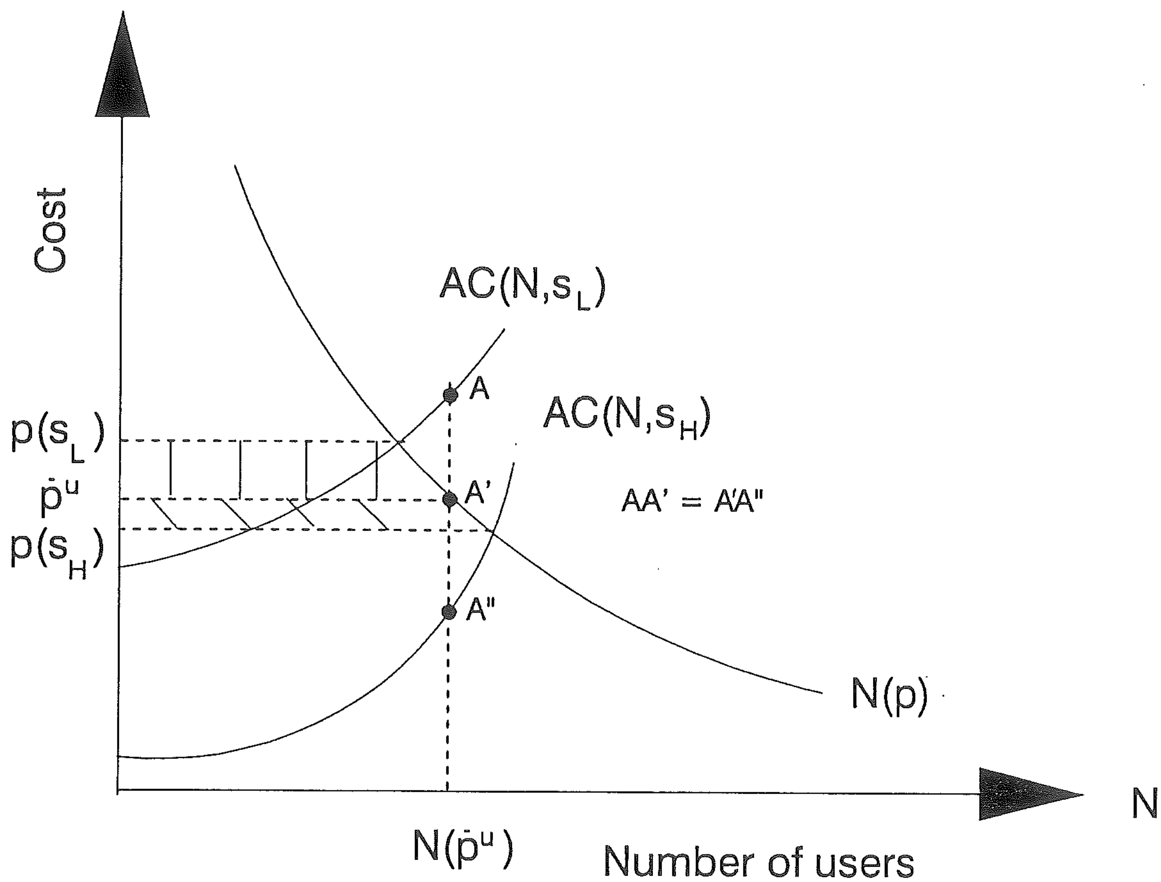
FIGURE 1

Welfare Comparison of Perfect Information and Zero Information

(a) Homogeneous user cost, constant elasticity demand



(b) Nonhomogeneous user cost



(c) Non-constant demand elasticity

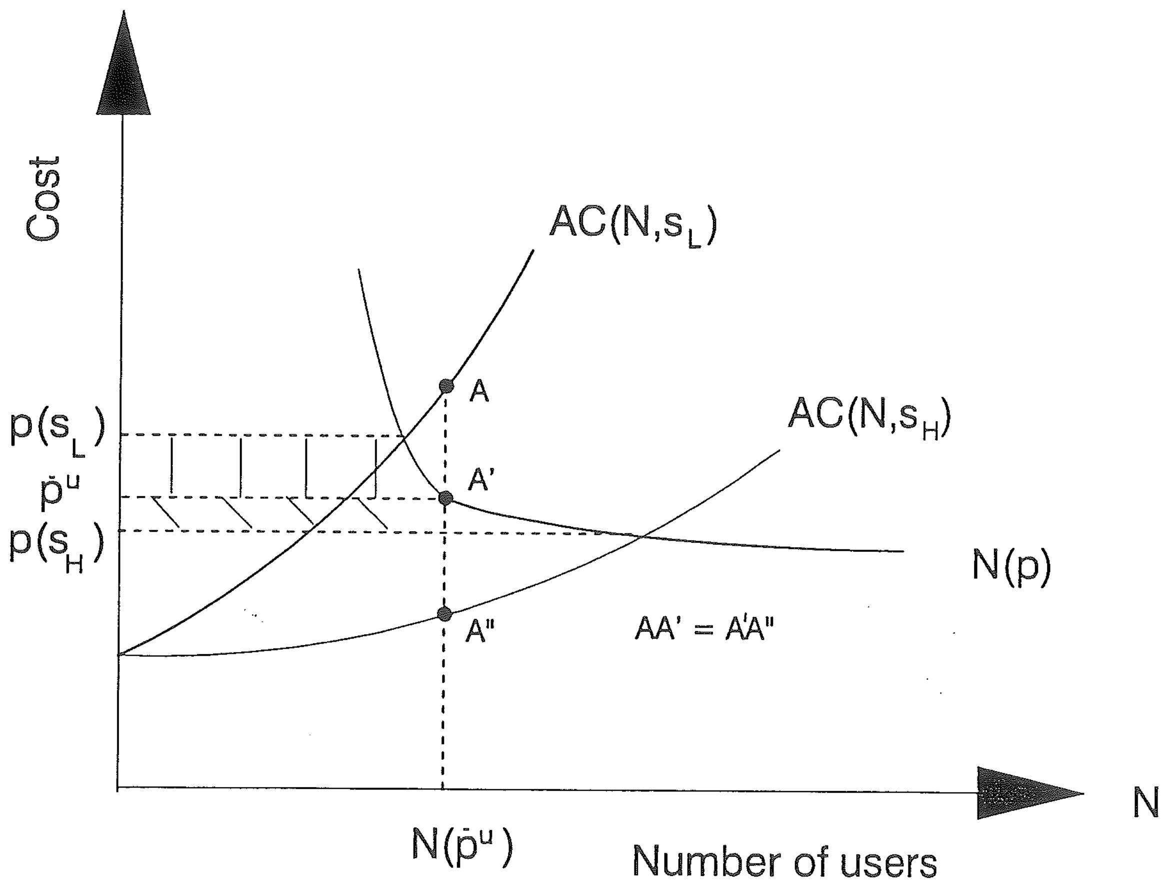
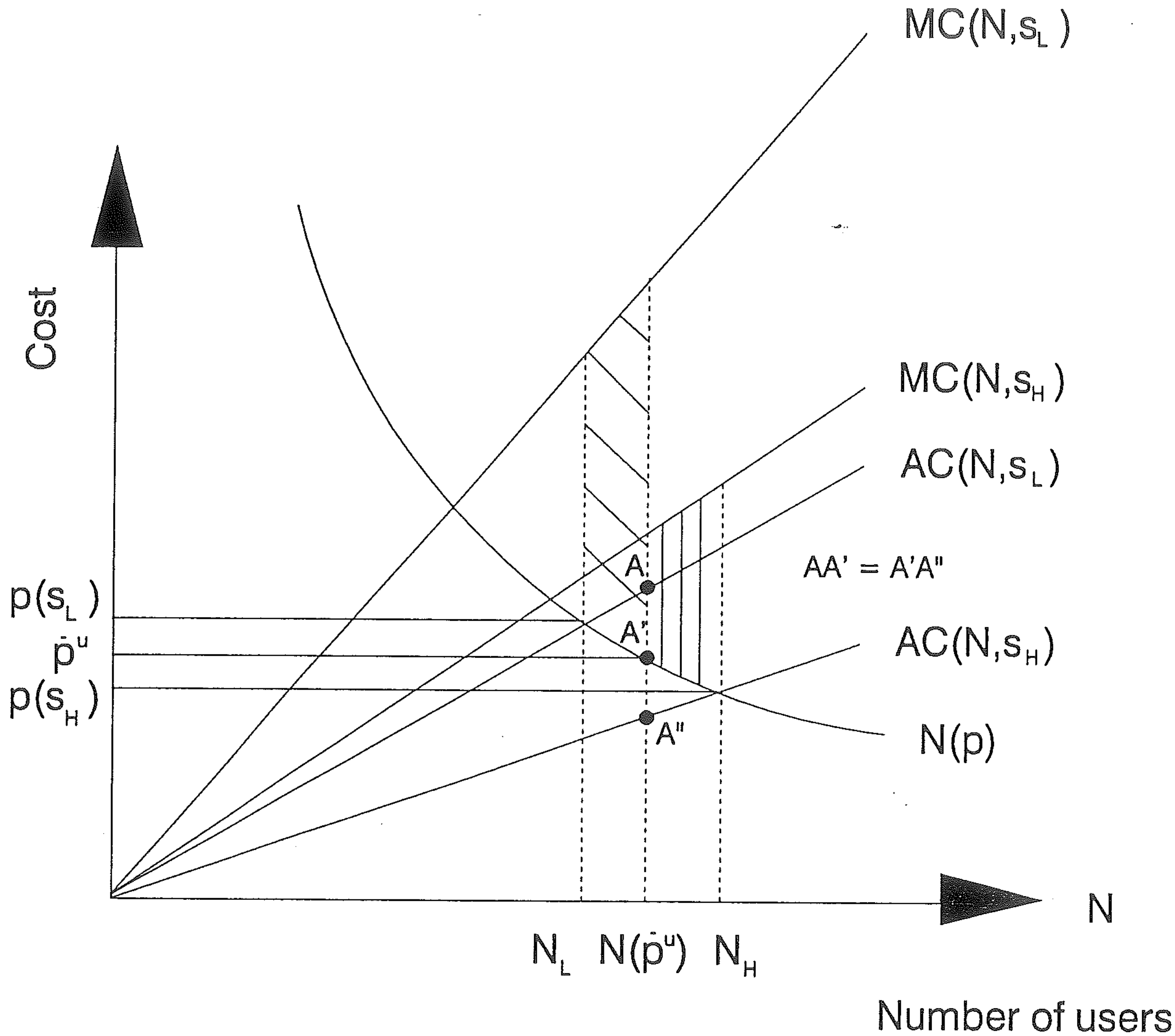


FIGURE 2

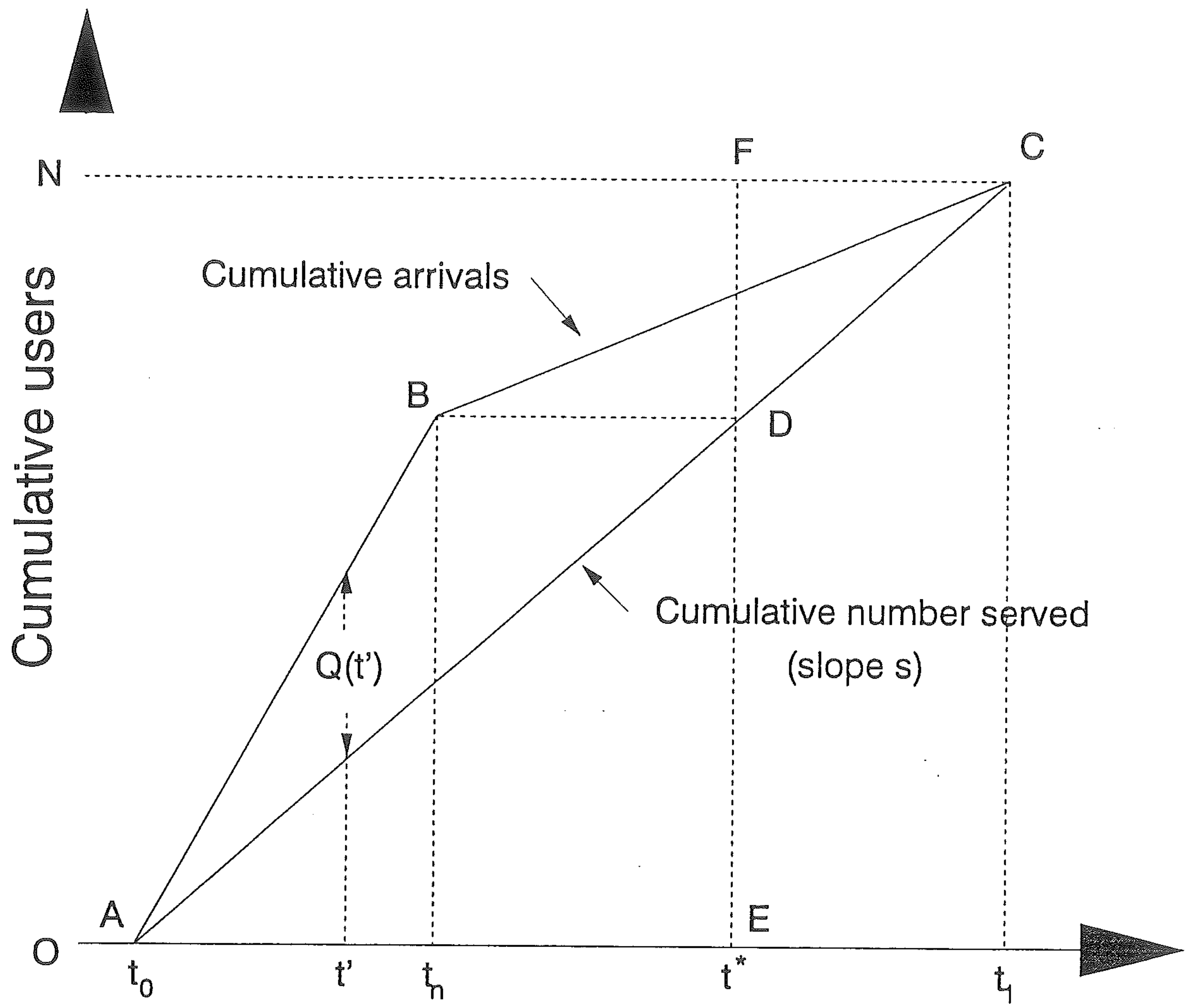


Efficiency gain from perfect information in low capacity state

Efficiency loss from perfect information in high capacity state

FIGURE 3

Dynamic Equilibrium with Full Information



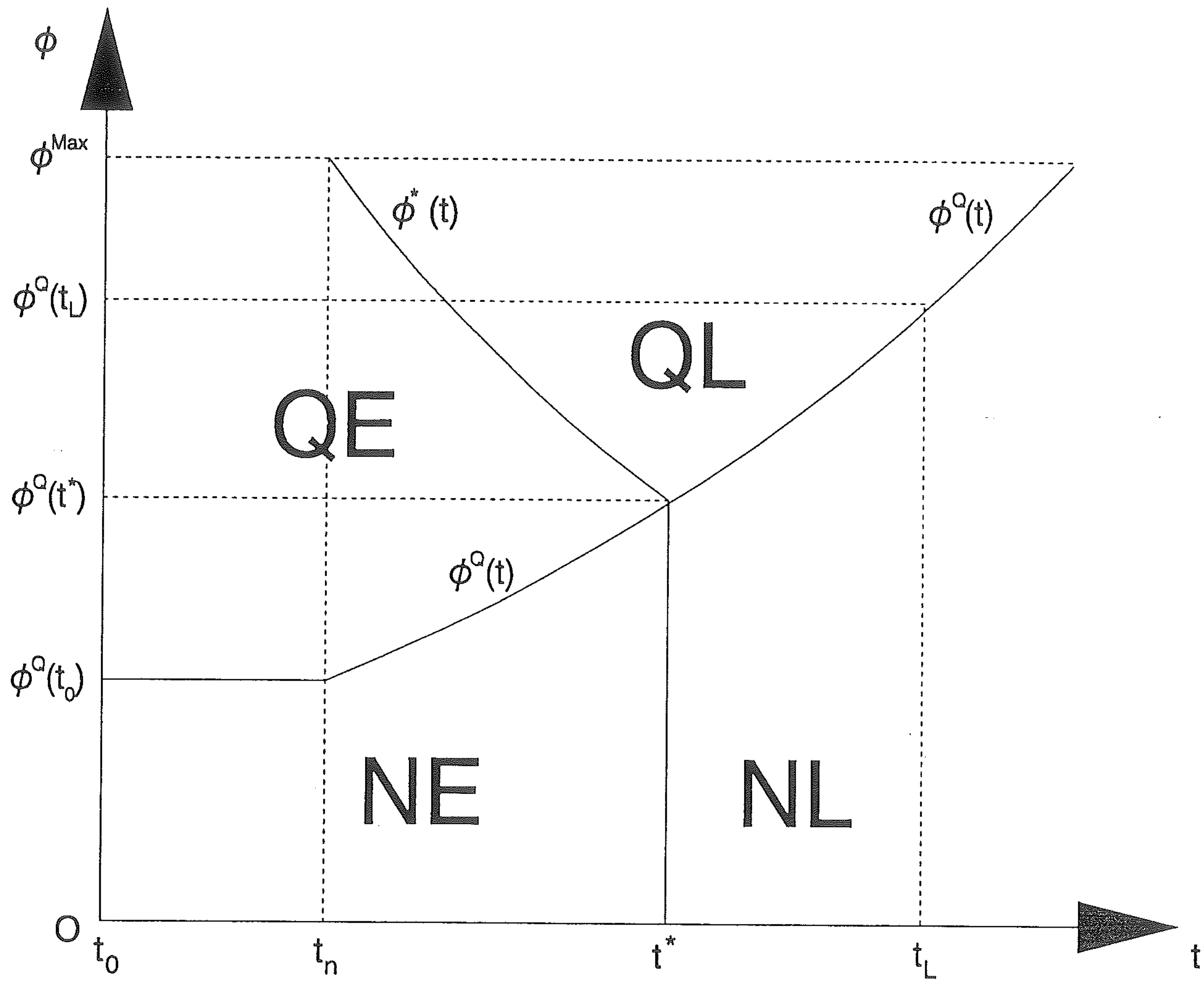
Total queueing time = ABCA

Total time early = ADEA

Total time late = CDFC

FIGURE 4

Queue Encountered and Time of Service as a Function of Arrival Time



Key: NE: No queueing, served early

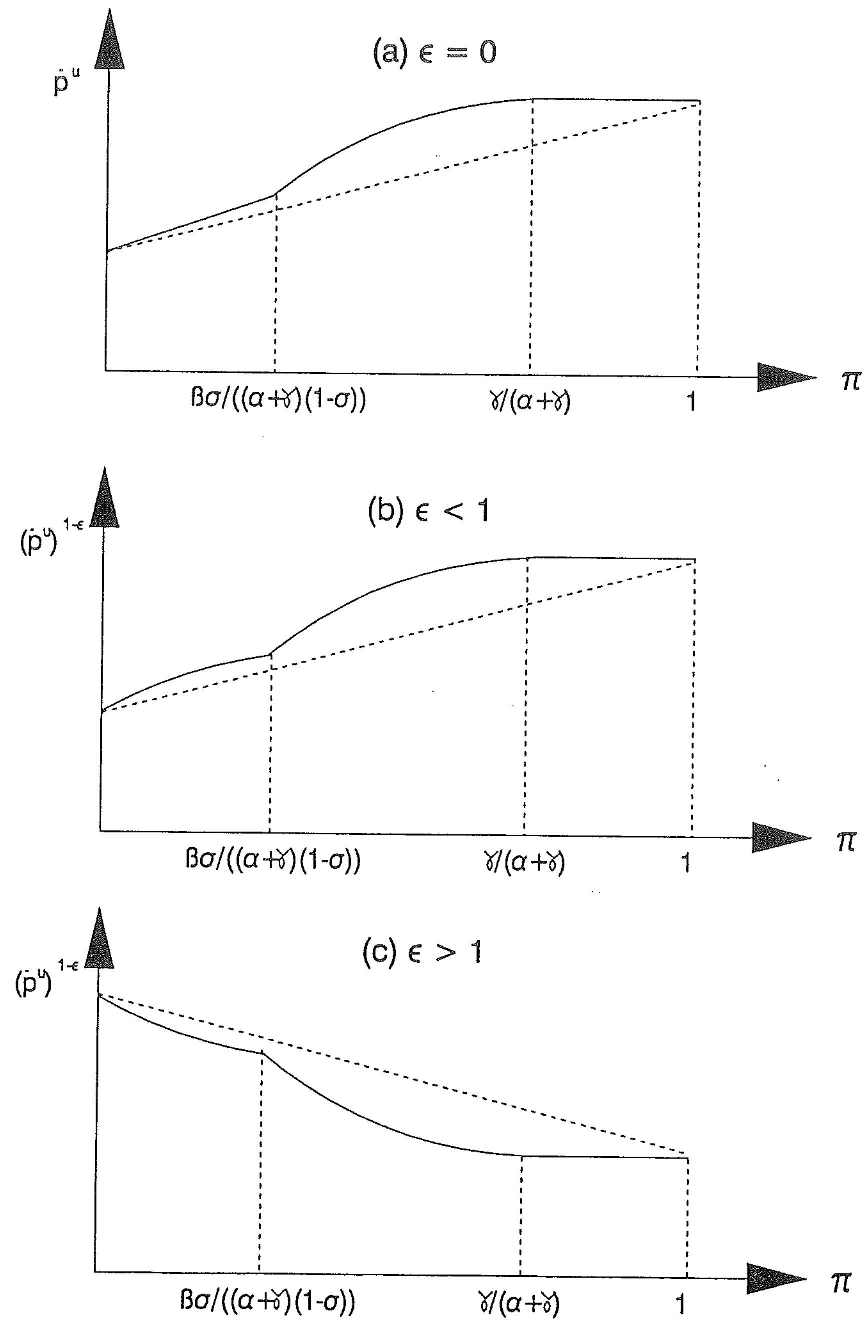
QE: Queueing, served early

NL: No queueing, served late

QL: Queueing, served late

FIGURE 5
Welfare Effect of Imperfect Information
(Demand fixed, capacity stochastic)

$\sigma < \gamma/(\beta+\gamma)$



APPENDIX 1

PROOF OF PROPOSITION 1

Proposition 1 concerns the ratio of optimal design capacity with anticipated (*viz.* predictable) fluctuations in capacity and demand, $\hat{s}_*^P$ given in (2.15), to optimal design capacity, $\hat{s}_*$, when capacity remains at its design value and demand intensity is fixed at its mean value.

Define $F(n,\sigma) \equiv \int_y F_y(n,\sigma)\, d\Gamma(y)$. Then

$$\frac{\hat{s}_*^P}{\hat{s}_*} = \frac{\nu^P}{\bar{n}} = \left[\int_n \int_\sigma n^{\frac{1+\eta}{1+\varepsilon\eta}} \sigma^{\frac{\eta(\varepsilon-1)}{1+\varepsilon\eta}} dF(n,\sigma) \right]^{\frac{1+\varepsilon\eta}{1+\eta}} \Big/ \bar{n} \,, \qquad \text{(A1.1)}$$

where $\bar{n}$ is mean demand intensity. The various statements in Proposition 1 are proved by a series of lemmas.

LEMMA 1: (First statement of Prop. 1)

$$\nu^P \gtreqless \bar{n} \quad \text{as } \varepsilon \lesseqgtr 1.$$

PROOF: Let x and z be variates. By Hölder's inequality

$$\int\int x^\alpha z^{1-\alpha} dF(x,z) \lesseqgtr (\bar{x})^\alpha (\bar{z})^{1-\alpha} \quad \text{as} \quad \alpha \begin{cases} \in (0,1) \\ = 0,\ 1 \\ < 0 \text{ or } > 1 \end{cases}$$

where $F(x,z)$ is an arbitrary nondegenerate joint c.d.f. Replacing x with σ, z with n, α with $\eta(\varepsilon-1)/(1+\varepsilon\eta)$ and $F(x,z)$ with $F(n,\sigma)$ yields

$$\nu^P \gtreqless (\bar{\sigma})^{\frac{\eta(\varepsilon-1)}{1+\eta}} \bar{n} \gtreqless \bar{n} \quad \text{as } \varepsilon \lesseqgtr 1. \quad \text{QED.}$$

Thus, optimal design capacity is greater in the stochastic case if $\varepsilon < 1$. The reason is that, with $\varepsilon < 1$, marginal consumers' surplus is a convex function of n and σ, so that by Jensen's inequality expected marginal consumers' surplus from capacity expansion is greater under stochasticity. If $\varepsilon > 1$, marginal consumers' surplus is concave in n and σ, and the inequality is reversed.

LEMMA 2: ((a) of Prop. 1)

Let $H(\sigma)$ be the c.d.f. of σ and $G(n|\sigma)$ be the c.d.f. of n conditional on σ. Then if $\varepsilon < 1$, $\nu^P/\bar{n}$ increases with a mean-preserving spread to $G(n|\sigma)$ for any set of σ of nonzero probability measure. The opposite is true if $\varepsilon > 1$.

PROOF: From (A1.1)

$$\nu^P = \left[\int_\sigma \sigma^{\frac{\eta(\varepsilon-1)}{1+\varepsilon\eta}} \nu(\sigma) dH(\sigma) \right]^{\frac{1+\varepsilon\eta}{1+\eta}}, \quad \text{where } \nu(\sigma) \equiv \int_n n^{\frac{1+\eta}{1+\varepsilon\eta}} dG(n|\sigma).$$

Let i_1 and i_2 be indexes. Assume that, for all σ, $G(n|\sigma, i_1)$ either coincides with $G(n|\sigma, i_2)$ or can be obtained by a mean-preserving spread, with the latter true for a subset of σ of positive probability measure. If $\varepsilon < 1$, $n^{\frac{1+\eta}{1+\varepsilon\eta}}$ is strictly convex in n, so that by Rothschild and Stiglitz (1970), $\nu(\sigma, i_1) \geq \nu(\sigma, i_2)$ for all σ, with a strict inequality for σ of positive probability measure. Then $\nu(i_1) > \nu(i_2)$. If $\varepsilon > 1$, $n^{\frac{1+\eta}{1+\varepsilon\eta}}$ is strictly concave in n, and by analogous reasoning $\nu(i_1) < \nu(i_2)$. The proof is completed by noting that the construction holds $\bar{n}$ fixed. QED.

LEMMA 3: ((b) of Prop. 1)

Let $G(n)$ be the c.d.f. of n, and $H(\sigma|n)$ be the c.d.f. of σ conditional on n. Then if $\varepsilon < 1$, $\nu^P/\bar{n}$ increases with a mean-preserving spread to $H(\sigma|n)$ for any set of n of positive probability measure. The opposite is true if $\varepsilon > 1$.

PROOF: By (A1.1)

$$\nu^P = \left[\int_n n^{\frac{1+\eta}{1+\varepsilon\eta}} w(n)dG(n) \right]^{\frac{1+\varepsilon\eta}{1+\eta}}, \quad \text{where } w(n) \equiv \int_\sigma \sigma^{\frac{\eta(\varepsilon-1)}{1+\varepsilon\eta}} dH(\sigma|n).$$

If $\varepsilon < 1$ (resp. > 1), $\sigma^{\frac{\eta(\varepsilon-1)}{1+\varepsilon\eta}}$ is strictly convex (resp. concave) in σ. The proof is analogous to that for Lemma 2. QED.

To establish the two remaining lemmas some preliminaries are required. Let x and z be variates with joint c.d.f. F(x,z) and marginal distributions H(x) and G(z). Define

$$\mathbb{M}_\alpha(x) \equiv \left\{ \int x^\alpha dH(x) \right\}^{1/\alpha}, \quad \mathbb{M}_\alpha(z) \equiv \left\{ \int z^\alpha dG(z) \right\}^{1/\alpha},$$

$$\mathbb{M}_\alpha(x,z) \equiv \iint x^\alpha z^{1-\alpha} dF(x,z),$$

and

$$\rho_\alpha(x,z) \equiv \left\{ \frac{\mathbb{M}_\alpha(x,z)}{(\mathbb{M}_\alpha(x))^\alpha (\mathbb{M}_{1-\alpha}(z))^{1-\alpha}} \right\} - 1. \tag{A1.2}$$

$\mathbb{M}_\alpha(x)$, $\mathbb{M}_\alpha(z)$ and $\mathbb{M}_\alpha(x,z)$ are generalized means, and $\rho_\alpha(x,z)$ is a generalized correlation coefficient. If x and z are independent, then

$$\mathbb{M}_\alpha(x,y) = \left[\int z^{1-\alpha} dG(z)\right]\left[\int x^\alpha dH(x)\right] = (\mathbb{M}_\alpha(x))^\alpha (\mathbb{M}_{1-\alpha}(z))^{1-\alpha},$$

so that $\rho_\alpha(x,z) = 0$. From (A1.1) and (A1.2)

$$\nu^P = \left[(\mathbb{M}_\alpha(\sigma))^\alpha (\mathbb{M}_{1-\alpha}(n))^{1-\alpha} (1 + \rho_\alpha(\sigma,n)) \right]^{1/(1-\alpha)}, \tag{A1.3}$$

where

$$\alpha = \frac{\eta(\varepsilon-1)}{1+\varepsilon\eta} < 1.$$

We can now proceed to lemmas 4 and 5.

LEMMA 4: ((c) of Prop. 1)

If $\varepsilon < 1$ (resp. > 1), $\nu^P/\bar{n}$ is larger the smaller (resp. larger) the α-mean of σ.

PROOF: Fix $\mathbb{M}_{1-\alpha}(n)$, $\bar{n}$ and $\rho_{\alpha}(\sigma,n)$. By (A1.3), $\nu^{P}/\bar{n}$ is larger the larger is $(\mathbb{M}_{\alpha}(\sigma))^{\alpha}$, and hence with $\varepsilon < 1$ $(\alpha < 0)$ the *smaller* is $\mathbb{M}_{\alpha}(\sigma)$. If $\varepsilon > 1$, $\nu^{P}/\bar{n}$ is larger the *larger* is $\mathbb{M}_{\alpha}(\sigma)$. QED.

LEMMA 5: ((d) of Prop. 1)

If $\varepsilon < 1$ (resp. > 1) $\nu^{P}/\bar{n}$ is larger the smaller (resp. larger) the correlation between n and σ.

PROOF: Fix $\mathbb{M}_{\alpha}(\sigma)$, $\mathbb{M}_{1-\alpha}(n)$ and $\bar{n}$. By (A1.3), $\nu^{P}/\bar{n}$ is larger the larger is $\rho_{\alpha}(\sigma,n)$. If $\varepsilon < 1$, $\alpha < 0$, and $\nu^{P}/\bar{n}$ is larger the *smaller* the correlation between σ and n. If $\varepsilon > 1$, $\alpha > 0$, and $\nu^{P}/\bar{n}$ is larger the *larger* the correlation. QED.

APPENDIX 2

PROOF OF PROPOSITION 2

Consumers' surplus with imperfect information is given by (2.24), which we write here with the order of integration altered:

$$\overline{CS}^{I}(\hat{s}) = \int_y \int_n \int_\sigma \left[\int_m \left\{ \int_{\bar{p}^u_{ym}} np^{-\varepsilon} dp \right\} dM(m|y,n,\sigma) \right] dF_y(n,\sigma) d\Gamma(y).$$

Let I and I′ be two imperfect information régimes with message distributions $M(m|y,n,\sigma)$ and $M'(m'|y,n,\sigma)$ respectively. The difference in expected consumers' surplus between the two régimes is

$$\overline{CS}^{I'}(\hat{s}) - \overline{CS}^{I}(\hat{s}) = \int_y \int_n \int_\sigma \left[\int_{m'} \left\{ \int_{\bar{p}^u_{ym'}} np^{-\varepsilon} dp \right\} dM'(m'|y,n,\sigma) - \right.$$

$$\left. \int_m \left\{ \int_{\bar{p}^u_{ym}} np^{-\varepsilon} dp \right\} dM(m|y,n,\sigma) \right] dF_y(n,\sigma) d\Gamma(y).$$

$$= \int_y \int_n \int_\sigma \left[\int_m \int_{m'} \left\{ \frac{n}{1-\varepsilon} \left((\bar{p}^u_{ym})^{1-\varepsilon} - (\bar{p}^u_{ym'})^{1-\varepsilon} \right) \right\} \right.$$

$$\left. dM'(m'|y,n,\sigma) dM(m|y,n,\sigma) \right] dF_y(n,\sigma) d\Gamma(y). \tag{A2.1}$$

Now $M\overline{CS}^{I'}(\hat{s}) - M\overline{CS}^{I}(\hat{s}) = \dfrac{d}{d\hat{s}} \left[\overline{CS}^{I'}(\hat{s}) - \overline{CS}^{I}(\hat{s}) \right]$

$$= \int_y \int_n \int_\sigma \left[\int_m \int_{m'} \frac{n}{1-\varepsilon} \frac{d}{d\hat{s}} \left(\left(\bar{p}^u_{ym}\right)^{1-\varepsilon} - \left(\bar{p}^u_{ym'}\right)^{1-\varepsilon} \right) dM'(m'|y,n,\sigma) \right.$$

$$\left. dM(m|y,n,\sigma) \right] dF_y(n,\sigma) d\Gamma(y). \tag{A2.2}$$

By (2.20)

$$\frac{d\bar{p}^u_{ym}}{d\hat{s}} \frac{\hat{s}}{\bar{p}^u_{ym}} = - \frac{\eta}{1+\varepsilon\eta}, \tag{A2.3}$$

whence

$$\frac{d}{d\hat{s}} \left(\bar{p}^u_{ym}\right)^{1-\varepsilon} = - \frac{\eta(1-\varepsilon)}{1+\varepsilon\eta} \frac{1}{\hat{s}} \left(\bar{p}^u_{ym}\right)^{1-\varepsilon}. \tag{A2.4}$$

Substituting (A2.4) and its counterpart for m' into (A2.2) and using (A2.1) one finds

$$\overline{MCS}^{I'}(\hat{s}) - \overline{MCS}^{I}(\hat{s}) = \frac{\eta(\varepsilon-1)}{1+\varepsilon\eta}\,\frac{1}{\hat{s}}\left[\overline{CS}^{I'}(\hat{s}) - \overline{CS}^{I}(\hat{s})\right].$$

If $\varepsilon > 1$, $\overline{MCS}^{I'}(\hat{s}) - \overline{MCS}^{I}(\hat{s})$ and $\overline{CS}^{I'}(\hat{s}) - \overline{CS}^{I}(\hat{s})$ have the same sign. If $\varepsilon <$ 1, they have the opposite sign. Furthermore, by (2.24)

$$\overline{MCS}^{I}(\hat{s}) = \frac{\eta}{(1+\varepsilon\eta)\hat{s}} \int_y \int_m \bar{n}_{ym}\left(\bar{p}^{u}_{ym}\right)^{1-\varepsilon} dM(m|y)\, d\Gamma(y).$$

Differentiating and using (A2.3)

$$\frac{d\overline{MCS}^{I}(\hat{s})}{d\hat{s}}\,\frac{\hat{s}}{\overline{MCS}^{I}(\hat{s})} = -\,\frac{1+\eta}{1+\varepsilon\eta}.$$

Since the elasticity of marginal consumers' surplus is constant, the ranking of $\overline{MCS}^{I}(\hat{s})$ and $\overline{MCS}^{I'}(\hat{s})$ is independent of $\hat{s}$, so that the relative efficiency of the two information régimes can be established by examining *any* level of design capacity. Since perfect information and zero information are limiting cases of imperfect information, Prop. 2 applies to all information régimes. QED.

APPENDIX 3

PROOF OF PROPOSITION 3

By Prop. 2 the relative efficiency of two information régimes can be ranked by comparing their marginal consumers' surplus for any design capacity. Consider two imperfect information régimes with respective message distributions $M'(m'|y)$ and $M(m|y)$ such that M' is a refinement of M; by (2.28)

$$dF_{ym}(n,\sigma) = \int_{m'\in S(m)} dF_{ym'}(n,\sigma)d\tilde{M}'(m'|m,y) \qquad \forall\ n,\sigma,m,y, \tag{A3.1}$$

where

$$d\tilde{M}'(m'|m,y) \equiv \frac{dM'(m'|y)}{dM(m|y)}.$$

Denote the two information régimes I' and I. Then

$$\overline{MCS}^{I'}(\hat{s}) - \overline{MCS}^{I}(\hat{s}) = \int_y \left(\overline{MCS}_y^{I'}(\hat{s}) - \overline{MCS}_y^{I}(\hat{s})\right)d\Gamma(y), \tag{A3.2}$$

where by (2.25)

$$\overline{MCS}_y^{I'}(\hat{s}) - \overline{MCS}_y^{I}(\hat{s}) = \frac{\eta}{(1+\varepsilon\eta)\hat{s}}\left[\int_{m'} \bar{n}_{ym'}\left(\bar{p}^u_{ym'}\right)^{1-\varepsilon} dM'(m'|y)\right.$$
$$\left. - \int_m \bar{n}_{ym}\left(\bar{p}^u_{ym}\right)^{1-\varepsilon} dM(m|y)\right]$$

$$\equiv \int_m \left[\int_{m'\in S(m)} \bar{n}_{ym'}\left(\bar{p}^u_{ym'}\right)^{1-\varepsilon} d\tilde{M}'(m'|m,y) - \bar{n}_{ym}\left(\bar{p}^u_{ym}\right)^{1-\varepsilon}\right]dM(m|y). \tag{A3.3}$$

Substituting (2.20) and the relation

$$\bar{n}_{ym} = \int_n\int_\sigma ndF_{ym}(n,\sigma)$$

for the two information régimes into (A3.3) one has

$$\overline{MCS}_y^{I'}(\hat{s}) - \overline{MCS}_y^{I}(\hat{s}) \equiv \int_m \left\{ \int_{m'\in S(m)} \left[\int_n\int_\sigma ndF_{ym'}(n,\sigma)\right]^{\frac{\varepsilon(1+\eta)}{1+\varepsilon\eta}}\right.$$

$$\left[\int_n \int_\sigma \left(\frac{m}{\sigma}\right)^{\eta} n dF_{ym'}(n,\sigma)\right]^{\frac{1-\varepsilon}{1+\varepsilon\eta}} d\tilde{M}'(m'|m,y)$$

$$- \left[\int_n \int_\sigma n dF_{ym}(n,\sigma)\right]^{\frac{\varepsilon(1+\eta)}{1+\varepsilon\eta}} \left[\int_n \int_\sigma \left(\frac{m}{\sigma}\right)^{\eta} n dF_{ym}(n,\sigma)\right]^{\frac{1-\varepsilon}{1+\varepsilon\eta}} \Bigg\} dM(m|y). \qquad \text{(A3.4)}$$

The RHS of (A3.4) has the form

$$\int_x \int_z x^{\alpha} z^{1-\alpha} dB(x,z) - (\bar{x})^{\alpha}(\bar{z})^{1-\alpha},$$

where

$$x \equiv \int_n \int_\sigma n\, dF_{ym'}(n,\sigma) \quad , \quad z \equiv \int_n \int_\sigma \left(\frac{m}{\sigma}\right)^{\eta} n\, dF_{ym'}(n,\sigma),$$

and by (A3.1)

$$\bar{x} = \int_{m'} x\, d\tilde{M}'(m'|m,y) \ , \ \bar{z} = \int_{m'} z\, d\tilde{M}'(m'|m,y), \ \alpha \equiv \frac{\varepsilon(1+\eta)}{1+\varepsilon\eta},$$

and the integration with respect to x and z is taken along the curve $(x(m'|m,y),z(m'|m,y))$, so that $dB(x,z) = dB(x(m'|m,y),z(m'|m,y)) \equiv d\tilde{M}'(m'|m,y)$.

Applying Hölder's inequality as in Appendix 1 we have

$$\overline{MCS}_y^{I'}(\hat{s}) \gtreqless \overline{MCS}_y^{I}(\hat{s}) \qquad \text{as} \quad \varepsilon \gtreqless 1.$$

Combining this with (A3.2) and Proposition 2 yields

$$\overline{CS}^{I'}(\hat{s}) \geq \overline{CS}^{I}(\hat{s}),$$

and

$$\hat{s}_*^{I'} \gtreqless \hat{s}_*^{I} \quad \text{as} \quad \varepsilon \gtreqless 1. \qquad \text{QED.}$$

APPENDIX 4

PROOF OF PROPOSITION 6

We prove that $\rho(t)$ is constant for $t \in (t_0, t_n)$ and weakly decreasing for $t \in (t_n, t_L)$. The proof begins with two lemmas.

Lemma 1 $\phi^Q(t)\rho(t) \le 1$

Proof A necessary condition for there to be no queue at t is $r(t) \le s$, or $\rho(t) \le 1/\phi$. By definition, this is true of $\phi^Q(t)$, hence $\phi^Q(t)\rho(t) \le 1$. QED.

Lemma 2 $[\phi^Q(t)\rho(t) - 1]\dot{\phi}^Q(t) \le 0$.

Proof Immediate if $\phi^Q(t)\rho(t) = 1$. By Lemma 1 we need only consider $\phi^Q(t)\rho(t) < 1$. By definition, $\phi^Q(t)$ is the largest ϕ such that there is no queue at t, so with $\rho(t) < 1/\phi^Q(t)$ the queue must be positive just before t. Let $\hat{t}$ be the last time when the queue was zero. By (3.1) in the text

$$\int_{\hat{t}}^{t} \phi^Q(t)\rho(\tau)d\tau = t - \hat{t}.$$

Differentiating,

$$\dot{\phi}^Q(t) = \frac{1-\phi^Q(t)\rho(t)}{\int_{\hat{t}}^{t} \rho(\tau)d\tau} > 0,$$

and hence $[\phi^Q(t)\rho(t) - 1]\dot{\phi}^Q(t) < 0$. QED.

We now consider each of the 3 arrival intervals in turn.

$\boxed{(a)\ t \in (t_0, t_n)}$

By (3.15), $Z^Q(t)$ defined in (3.16) is constant. Differentiating, we find

$$\dot{\rho}(t) \int_{\phi^Q(t)}^{\phi^{Max}} \phi dJ(\phi) = (\phi^Q(t)\rho(t) - 1)\left.\frac{dJ(\phi)}{d\phi}\right|_{\phi^Q(t)} \dot{\phi}^Q(t). \qquad (A4.1)$$

By Lemma 2, the RHS of (A4.1) is nonpositive. The integral term on the LHS is strictly positive if $J(\phi^Q(t)) < 1$, which is the case since otherwise there would never be a queue at $t > t_0$, and arrival at t would be preferable to arrival at t_0. Hence $\dot{\rho}(t) \le 0$, $R(t)$ is concave and

$$\phi^Q(t) = \frac{t-t_0}{R(t)}. \tag{A4.2}$$

Now it is clear that (A4.1) has a solution with $\rho(t) = \rho(t_0)$, $R(t) = \rho(t_0)(t-t_0)$ and $\phi^Q(t) = 1/\rho(t_0)$ for $t \in (t_0,t_n)$. To see that this is the unique solution substitute (3.1), (3.2) and (A4.2) into (3.14)

$$C(t_0) = C(t) = \beta(t^*-t) + (\alpha-\beta) \int_{\frac{t-t_0}{R(t)}}^{\phi^{Max}} \left[\phi R(t)-(t-t_0)\right] dJ(\phi).$$

Since the RHS term is a strictly increasing function of $R(t)$, $R(t)$ is unique, and hence so is $\rho(t)$. This establishes that the arrival rate is constant for $t \in (t_0,\ t_n)$. QED.

$\boxed{\text{(b) } t \in (t_n,\ t^*)}$

Differentiating (3.18) and cancelling terms:

$$\left\{(\alpha-\beta) \int_{\phi^Q(t)}^{\phi^*(t)} \phi dJ(\phi) + (\alpha+\gamma) \int_{\phi^*(t)}^{\phi^{Max}} \phi dJ(\phi)\right\} \dot{\rho}(t) = \tag{A4.3}$$

$$(\alpha-\beta)[\phi^Q(t)\rho(t)-1]\dot{\phi}^Q(t) \frac{dJ(\phi^Q(t))}{d\phi} + (\beta+\gamma)\phi^*(t)\rho(t)\dot{\phi}^*(t)\frac{dJ(\phi^*(t))}{d\phi}.$$

The term in braces on the LHS of (A4.3) is strictly positive if $J(\phi^Q(t)) < 1$, which is the case by the same reasoning given for $t \in (t_0,\ t_n)$. $[\phi^Q(t)\rho(t)-1]\dot{\phi}^Q(t) \le 0$ by Lemma 2, so the first term on the RHS is nonpositive. The second term on the RHS is nonpositive if $\dot{\phi}^*(t) \le 0$. By definition of $\phi^*(t)$, $t + q(t,\phi^*(t)) = t^*$. Differentiating, and setting $\phi =$

$\phi^*(t)$ one has

$$\dot{\phi}^*(t) = -\frac{\phi^*(t)\rho(t)}{R(t)} < 0.$$

The RHS of (A4.3) is therefore nonpositive, which establishes that $\dot{\rho}(t) \leq 0$ for $t \in (t_n, t^*)$.

$\boxed{\text{(c) } t \in (t^*, t_L)}$

Differentiating (3.22):

$$\left\{ \int_{\phi^Q(t)}^{\phi^{Max}} \phi dJ(\phi) \right\} \dot{\rho}(t) = [\phi^Q(t)\rho(t)-1]\dot{\phi}^Q(t) \frac{dJ(\phi^Q(t))}{d\phi}. \tag{A4.4}$$

Since the LHS term in braces of (A4.4) is positive and the RHS is nonpositive

$$\dot{\rho}(t) \leq 0. \quad \text{QED.}$$

By arriving after t^*, a user is always served late, so the late service effect operating through changes in $\phi^*(t)$ is absent, leaving only the queueing effect operating through changes in $\phi^Q(t)$.

APPENDIX 5

PROOF OF LEMMA 1

Using (3.21) and (3.23), and the fact that with $t > t_L$, $R(t) = 1$ and hence $\phi^Q(t) = t - t_0$, expected costs for arrival after t_L are

$$C(t) = \gamma\left\{\int_0^{t-t_0} (t-t^*)dJ(\phi) + \int_{t-t_0}^{\phi^{Max}} (t_0+\phi-t^*)dJ(\phi)\right\} + \alpha\int_{t-t_0}^{\phi^{Max}} (t_0+\phi-t)dJ(\phi). \quad (A5.1)$$

If t_L is indeed the last arrival time, then $C(t)$ must be nondecreasing after t_L. Since $C(t)$ is a continuous function of t, it suffices to consider $\dot{C}(t)$ after t_L. From (3.22) and (3.16), the left-hand derivative (which is zero by construction) is:

$$\lim_{t\uparrow t_L} \dot{C}(t) = \gamma(1+\lim_{t\uparrow t_L} Z^Q(t)) + \alpha \lim_{t\uparrow t_L} Z^Q(t)$$

$$= \gamma - (\alpha+\gamma)(1-\lim_{t\uparrow t_L} J(t-t_0)) + (\alpha+\gamma)\lim_{t\uparrow t_L} \int_{t^*-t_0}^{\phi^{Max}} \phi\rho(t)dJ(\phi) = 0. \quad (A5.2)$$

From (A5.1), the right-hand derivative is

$$\dot{C}(t) = \gamma - (\alpha+\gamma)(1-J(t_L-t_0)). \quad (A5.3)$$

Since $\dot{C}(t)$ is nonincreasing in t, a necessary and sufficient condition for equilibrium is

$$\lim_{t\downarrow t_L} \dot{C}(t) = \gamma - (\alpha+\gamma)(1-J(t_L-t_0)) \geq 0$$

or

$$J(t_L-t_0) \geq \frac{\alpha}{\alpha+\gamma}. \quad (A5.4)$$

But since the last term in (A5.2) is nonnegative

$$\lim_{t\uparrow t_L} J(t-t_0) \leq \frac{\alpha}{\alpha+\gamma}. \quad (A5.5)$$

Thus

$$t_L = t_0 + \tilde{\phi}, \tag{A5.6}$$

(recall that $\tilde{\phi} = J^{-1}\left(\dfrac{\alpha}{\alpha+\gamma}\right)$) which is (3.29) in the text. A solution for t_0 and t_L now follows directly. From (3.14)

$$C(t_0) = \beta(t^* - t_0). \tag{A5.7}$$

From (3.21) and (3.23)

$$C(t_L) = \gamma(t_L - t^*) + (\alpha+\gamma)\int_{t_L - t_0}^{\phi^{Max}} (t_0 + \phi - t_L)dJ(\phi)$$

$$= \gamma(t_L - t^*) - (\alpha+\gamma)(t_L - t_0)(1 - J(t_L - t_0)) + (\alpha+\gamma)\int_{t_L - t_0}^{\phi^{Max}} \phi dJ(\phi)$$

$$= -\gamma(t^* - t_0) + (\alpha+\gamma)\int_{\tilde{\phi}}^{\phi^{Max}} \phi dJ(\phi) \quad \text{(using (A5.6))}.$$

Setting $C(t_0) = C(t_L) = C$:

$$C = \beta(t^* - t_0) = \beta\frac{\alpha+\gamma}{\beta+\gamma}\int_{\tilde{\phi}}^{\phi^{Max}} \phi dJ(\phi), \tag{A5.8}$$

which is (3.30), and immediately yields (3.28). The initial premise that $t_L > t^*$ is satisfied provided:

$$t_L - t_0 > t^* - t_0 \Leftrightarrow \tilde{\phi} > \frac{\alpha+\gamma}{\beta+\gamma}\int_{\tilde{\phi}}^{\phi^{Max}} \phi dJ(\phi),$$

which is condition (3.27). This completes the proof of Lemma 1.

APPENDIX 6

PROOF OF LEMMA 2

If $t_L = t^*$

$$\phi^Q(t_L) = \phi^Q(t^*) = \phi^*(t_L) = \phi^*(t^*) = t^*-t_0 \text{ and } R(t^*) = 1. \tag{A6.1}$$

As in Appendix 5, a necessary and sufficient condition for equilibrium is (A5.4):

$$J(t_L-t_0) \geq \frac{\alpha}{\alpha+\gamma}. \tag{A6.2}$$

Using (A6.1) and (3.17) one also has

$$C(t_L) = C(t^*) = (\alpha+\gamma) \int_{t^*-t_0}^{\phi^{Max}} (t_0+\phi-t^*)dJ(\phi). \tag{A6.3}$$

As before, $C(t_0) = \beta(t^*-t_0)$, which is (3.33). Equating $C(t_0)$ and $C(t^*)$ yields (3.32):

$$J(t^*-t_0) + \frac{1}{t^*-t_0} \int_{t^*-t_0}^{\phi^{Max}} \phi dJ(\phi) - \frac{\alpha+\beta+\gamma}{\alpha+\gamma} = 0. \tag{A6.4}$$

Now, the LHS of (A6.4) is strictly decreasing in t^*-t_0. Since $J(\phi)$ is monotonically increasing, (A6.2) is satisfied if the LHS of (A6.4) is nonnegative with $J(t^*-t_0) = \frac{\alpha}{\alpha+\gamma}$:

$$\frac{1}{\tilde{\phi}} \int_{\tilde{\phi}}^{\phi^{Max}} \phi dJ(\phi) - \frac{\beta+\gamma}{\alpha+\gamma} \geq 0, \quad \text{or}$$

$$\tilde{\phi} \leq \frac{\alpha+\gamma}{\beta+\gamma} \int_{\tilde{\phi}}^{\phi^{Max}} \phi dJ(\phi), \tag{A6.5}$$

which is condition (3.31).

APPENDIX 7

PROOF OF PROPOSITION 8

Conditional on y and m, expected user costs are $\bar{p}^u_{ym} = \delta\phi'_{ym}$, with ϕ'_{ym} given by (3.30) if $t_L > t^*$ and by (3.34) if $t_L = t^*$. With perfect information and a realization (n,σ), user costs are by (3.10)

$$p^e(n,s) = \left(\frac{\delta n}{\hat{s}\sigma}\right)^{\frac{1}{1+\varepsilon}},$$

Define $\theta = \frac{n}{\sigma}$, and let $J^\theta_y(\theta)$ be the c.d.f. for users of θ given y, and $J^\theta_{ym}(\theta)$ the c.d.f. given y and m. Then

$$\phi = \theta\frac{p^e(n,s)^{-\varepsilon}}{\hat{s}} = \delta^{\frac{-\varepsilon}{1+\varepsilon}}(\hat{s})^{\frac{-1}{1+\varepsilon}}\theta^{\frac{1}{1+\varepsilon}}. \tag{A7.1}$$

Define

$$\theta^{Min}_y \equiv \mathrm{Inf}\{\theta|\theta \in \mathrm{Supp}\ J^\theta_y(\theta)\},$$

$$\theta^{Max}_y \equiv \mathrm{Sup}\{\theta|\theta \in \mathrm{Supp}\ J^\theta_y(\theta)\},$$

$$\phi^{Min}_y \equiv \delta^{\frac{-\varepsilon}{1+\varepsilon}}(\hat{s})^{\frac{-1}{1+\varepsilon}}\left(\theta^{Min}_y\right)^{\frac{1}{1+\varepsilon}},$$

$$\phi^{Max}_y \equiv \delta^{\frac{-\varepsilon}{1+\varepsilon}}(\hat{s})^{\frac{-1}{1+\varepsilon}}\left(\theta^{Max}_y\right)^{\frac{1}{1+\varepsilon}}.$$

Prop. 8 states that

$$\delta\phi^{Min}_y \le \delta\phi'_{ym} \le \delta\phi^{Max}_y.$$

By assumption A-5 the support of $F^u_{ym}(n,\sigma)$ is contained in the support of $F^u_y(n,\sigma)$. To prove Prop. 8, it thus suffices to prove

$$\phi^{Min}_{ym} \le \phi'_{ym} \le \phi^{Max}_{ym}, \tag{A7.2}$$

where by (A7-1)

$$\phi^{Min}_{ym} \equiv \delta^{\frac{-\varepsilon}{1+\varepsilon}}(\hat{s})^{\frac{-1}{1+\varepsilon}}\left(\theta^{Min}_{ym}\right)^{\frac{1}{1+\varepsilon}}, \tag{A7.3}$$

$$\phi_{ym}^{Max} \equiv \delta^{\frac{-\varepsilon}{1+\varepsilon}}(\hat{s})^{\frac{-1}{1+\varepsilon}}\left(\theta_{ym}^{Max}\right)^{\frac{1}{1+\varepsilon}}, \tag{A7.4}$$

and θ_{ym}^{Min} and θ_{ym}^{Max} are defined relative to $J_{ym}^{\theta}(\theta)$. (A7.2) is proved by considering the cases $t_L > t^*$ and $t_L = t^*$ separately.

$\boxed{t_L > t^*}$: By (3.25)

$$\phi'_{ym} = \hat{\phi}_{ym} = \frac{\alpha+\gamma}{\gamma}\int_{\tilde{\phi}_{ym}} \phi dJ_{ym}(\phi) = \frac{\alpha+\gamma}{\gamma}\int_{\tilde{\theta}_{ym}} \theta \frac{\left(\delta\hat{\phi}_{ym}\right)^{-\varepsilon}}{\hat{s}} dJ_{ym}^{\theta}(\theta)$$

$$= \left[\frac{\alpha+\gamma}{\gamma}\int_{\tilde{\theta}_{ym}} \theta \frac{\delta^{-\varepsilon}}{\hat{s}} dJ_{ym}^{\theta}(\theta)\right]^{\frac{1}{1+\varepsilon}} = \delta^{\frac{-\varepsilon}{1+\varepsilon}}(\hat{s})^{\frac{-1}{1+\varepsilon}}\left(\hat{\theta}_{ym}\right)^{\frac{1}{1+\varepsilon}}, \tag{A7.5}$$

where

$$\hat{\theta}_{ym} \equiv \frac{\alpha+\gamma}{\gamma}\int_{\tilde{\theta}_{ym}} \theta \, dJ_{ym}^{\theta}(\theta).$$

Given (A7.3), (A7.4) and (A7.5), (A7.2) is satisfied if

$$\theta_{ym}^{Min} \le \hat{\theta}_{ym} \le \theta_{ym}^{Max}, \tag{A7.6}$$

which is the case because $\hat{\theta}_{ym}$ is a truncated mean of θ for the distribution $J_{ym}^{\theta}(\theta)$.

$\boxed{t_L = t^*}$: By (3.32) and (3.34), ϕ'_{ym} is defined implicitly by the equation

$$J_{ym}\left(\frac{\gamma}{\beta+\gamma}\phi'_{ym}\right) + \frac{\beta+\gamma}{\gamma\phi'_{ym}}\int_{\frac{\gamma}{\beta+\gamma}\phi'_{ym}} \phi dJ_{ym}(\phi) - \frac{\alpha+\beta+\gamma}{\alpha+\gamma} = 0.$$

Now $\phi = \theta\dfrac{\left(\bar{p}_{ym}^{u}\right)^{-\varepsilon}}{\hat{s}} = \theta\dfrac{\left(\delta\phi'_{ym}\right)^{-\varepsilon}}{\hat{s}}$.

Define θ'_{ym} by the equation

$$\phi'_{ym} = \frac{\theta'_{ym}\left(\delta\phi'_{ym}\right)^{-\varepsilon}}{\hat{s}}.$$

Then

$$\phi'_{ym} = \delta^{\frac{-\varepsilon}{1+\varepsilon}}(\hat{s})^{\frac{-1}{1+\varepsilon}}\left(\theta'_{ym}\right)^{\frac{1}{1+\varepsilon}}, \tag{A7.7}$$

and

$$Y\left(\theta'_{ym}\right) \equiv J^{\theta}_{ym}\left(\frac{\gamma}{\beta+\gamma}\theta'_{ym}\right) + \frac{\beta+\gamma}{\gamma\theta'_{ym}} \int_{\frac{\gamma}{\beta+\gamma}\theta'_{ym}} \theta dJ^{\theta}_{ym}(\theta) - \frac{\alpha+\beta+\gamma}{\alpha+\gamma} = 0. \tag{A7.8}$$

Given (A7.3), (A7.4) and (A7.7), (A7.2) is satisfied if

$$\theta^{Min}_{ym} \leq \theta'_{ym} \leq \theta^{Max}_{ym}, \tag{A7.9}$$

Now $Y\left(\theta'_{ym}\right)$ is a strictly decreasing function, so that (A7.9) holds if

$$Y\left(\theta^{Min}_{ym}\right) > 0 \text{ and } Y\left(\theta^{Max}_{ym}\right) < 0. \tag{A7.10}$$

Making the substitutions:

$$Y\left(\theta^{Min}_{ym}\right) = \frac{\beta+\gamma}{\gamma}\,\frac{\bar{\theta}_{ym}}{\theta^{Min}_{ym}} - \frac{\alpha+\beta+\gamma}{\alpha+\gamma} > 0.$$

$$\begin{aligned}
Y\left(\theta^{Max}_{ym}\right) &= J^{\theta}_{ym}\left(\frac{\gamma}{\beta+\gamma}\theta^{Max}_{ym}\right) + \frac{\beta+\gamma}{\gamma\theta^{Max}_{ym}} \int_{\frac{\gamma}{\beta+\gamma}\theta^{Max}_{ym}} \theta dJ^{\theta}_{ym}(\theta) - \frac{\alpha+\beta+\gamma}{\alpha+\gamma} \\
&\leq J^{\theta}_{ym}\left(\frac{\gamma}{\beta+\gamma}\theta^{Max}_{ym}\right) + \frac{\beta+\gamma}{\gamma}\left[1 - J^{\theta}_{ym}\left(\frac{\gamma}{\beta+\gamma}\theta^{Max}_{ym}\right)\right] - \frac{\alpha+\beta+\gamma}{\alpha+\gamma} \\
&\leqq 1 - J^{\theta}_{ym}\left(\frac{\gamma}{\beta+\gamma}\theta^{Max}_{ym}\right) - \frac{\gamma}{\alpha+\gamma} \\
&\leqq \tilde{\theta}_{ym} - \frac{\gamma}{\beta+\gamma}\theta^{Max}_{ym} < \frac{\gamma}{\beta+\gamma}(\hat{\theta}_{ym} - \theta^{Max}_{ym}) < 0.
\end{aligned}$$

This establishes (A7.10), and hence (A7.9) and (A7.2) for $t_L = t^*$.

APPENDIX 8

PROOF OF PROPOSITION 9

We wish to prove

$$\int_m \delta\phi'_{ym}\ dM^u(m|y) \geq \int_\phi \delta\phi\ dJ_y(\phi), \quad \forall\ y. \tag{A8.1}$$

Given (A7.1)

$$\int_\phi \delta\phi\ dJ_y(\phi) = \delta^{\frac{1}{1+\varepsilon}}(\hat{s})^{\frac{-1}{1+\varepsilon}}\int_\theta \theta^{\frac{1}{1+\varepsilon}}dJ^\theta_y(\theta). \tag{A8.2}$$

Given (A7.5) and (A7.7)

$$\int_m \delta\phi'_{ym}\ dM^u(m|y) = \delta^{\frac{1}{1+\varepsilon}}(\hat{s})^{\frac{-1}{1+\varepsilon}}\int_m \left(\theta'_{ym}\right)^{\frac{1}{1+\varepsilon}}dM^u(m|y) \tag{A8.3}$$

where

$$\theta'_{ym} = \hat{\theta}_{ym} = \frac{\alpha+\gamma}{\gamma}\int_{\tilde{\theta}_{ym}} \theta dJ^\theta_{ym}(\theta) \quad \text{if } t_L > t^*,$$

$$\theta'_{ym} \text{ is defined implicitly by (A7.8)} \quad \text{if } t_L = t^*.$$

Since $\hat{\theta}_{ym}$ is a mean of θ for $J^\theta_{ym}(\theta)$, with the left-hand tail truncated,

$$\hat{\theta}_{ym} > \bar{\theta}^u_{ym}. \tag{A8.4}$$

Condition (A7.8) can be written

$$\int_{\frac{\gamma}{\beta+\gamma}\theta'_{ym}} \left(\theta - \frac{\gamma}{\beta+\gamma}\theta'_{ym}\right)dJ^\theta_{ym}(\theta) - \frac{\beta}{\alpha+\gamma}\,\frac{\gamma}{\beta+\gamma}\,\theta'_{ym} = 0. \tag{A8.5}$$

Now,

$$\frac{\gamma}{\beta+\gamma}\,\theta'_{ym}\ J^\theta_{ym}\left(\frac{\gamma}{\beta+\gamma}\theta'_{ym}\right) + \int_{\frac{\gamma}{\beta+\gamma}\theta'_{ym}} \theta dJ^\theta_{ym}(\theta) \geq \bar{\theta}^u_{ym}. \tag{A8.6}$$

Subtracting (A8.5) from (A8.6) one obtains

$$\theta'_{ym} \geq \frac{\beta+\gamma}{\gamma} \frac{\alpha+\gamma}{\alpha+\beta+\gamma} \bar{\theta}^{u}_{ym} > \bar{\theta}^{u}_{ym}. \tag{A8.7}$$

It follows from (A8.2), (A8.3), (A8.4) and (A8.7) that (A8.1) is satisfied if

$$\int_{m} \left(\bar{\theta}^{u}_{ym}\right)^{\frac{1}{1+\varepsilon}} dM^{u}(m|y) - \int_{\theta} \theta^{\frac{1}{1+\varepsilon}} dJ^{\theta}_{y}(\theta) \geq 0. \tag{A8.8}$$

By Bayes' rule

$$\int_{m} dJ^{\theta}_{ym}(\theta) dM^{u}(m|y) = dJ^{\theta}_{y}(\theta),$$

and (A8.8) can be written

$$\int_{m} \left[\left(\int_{\theta} \theta dJ^{\theta}_{ym}(\theta) \right)^{\frac{1}{1+\varepsilon}} - \int_{\theta} \theta^{\frac{1}{1+\varepsilon}} dJ^{\theta}_{ym}(\theta) \right] dM^{u}(m|y) \geq 0. \tag{A8.9}$$

But

$$\left(\int_{\theta} \theta dJ^{\theta}_{ym}(\theta) \right)^{\frac{1}{1+\varepsilon}} - \int_{\theta} \theta^{\frac{1}{1+\varepsilon}} dJ^{\theta}_{ym}(\theta)$$

$$\cong \int_{\theta} \theta dJ^{\theta}_{ym}(\theta) - \left(\int_{\theta} \theta^{\frac{1}{1+\varepsilon}} dJ^{\theta}_{ym}(\theta) \right)^{1+\varepsilon}$$

which is nonnegative by Hardy, Littlewood and Polya (HLP) (1934, Prop. 2.9.1). (A8.9) is thus satisfied. QED.

APPENDIX 9

PROOF OF PROPOSITION 2′

The proof of Prop. 2 for the static model given in Appendix 2 carries overly largely unchanged to the dynamic model. Expected consumers' surplus is still given by (2.24). The only difference is that expected user cost, given by (2.20), must be replaced by its dynamic counterpart. To reduce the notational burden, subscripts y and m denoting the day and message are suppressed. The cases $t_L > t^*$ and $t_L = t^*$ must be considered separately.

$\boxed{\text{(a) } t_L > t^*}$:

From (3.26) and (3.30)

$$\bar{p}^u = C = \beta(t^* - t_0) = \frac{\beta(\alpha+\gamma)}{\beta+\gamma} \int_{\tilde{\phi}} \phi dJ(\phi)$$

$$= \frac{\beta(\alpha+\gamma)}{\beta+\gamma} \int_{\tilde{\theta}} \theta \frac{(\bar{p}^u)^{-\varepsilon}}{\hat{s}} dJ^\theta(\theta) = \left\{ \frac{\beta(\alpha+\gamma)}{\beta+\gamma} \frac{1}{\hat{s}} \int_{\tilde{\theta}} \theta dJ^\theta(\theta) \right\}^{\frac{1}{1+\varepsilon}} . \tag{A9.1}$$

Furthermore, $\dfrac{d\bar{p}^u}{d\hat{s}} \dfrac{\hat{s}}{\bar{p}^u} = -\dfrac{1}{1+\varepsilon}$, which establishes that (A2.3) and (A2.4) in Appendix 2 hold also for the dynamic model when $\eta = 1$.

$\boxed{\text{(b) } t_L = t^*}$:

In this case

$$\bar{p}^u = \beta(t^* - t_0) = \beta\phi^0 \tag{A9.2}$$

with ϕ^0 defined implicitly by (3.32)

$$J(\phi^0) + \frac{1}{\phi^0} \int_{\phi^0} \phi dJ(\phi) - \frac{\alpha+\beta+\gamma}{\alpha+\gamma} = 0. \tag{A9.3}$$

Making the change of variable from ϕ to θ and defining $\theta^0 \equiv \left(\bar{p}^u\right)^{\varepsilon}\hat{s}\phi^0$,

(A9.2) and (A9.3) can be rewritten

$$\bar{p}^u = \left\{\frac{\beta\theta^0}{\hat{s}}\right\}^{\frac{1}{1+\varepsilon}}, \tag{A9.4}$$

$$J^{\theta}(\theta^0) + \frac{1}{\theta^0}\int_{\theta^0}\theta dJ^{\theta}(\theta) - \frac{\alpha+\beta+\gamma}{\alpha+\gamma} = 0. \tag{A9.5}$$

Since the joint distribution of n and σ is assumed to be independent of $\hat{s}$, so is the distribution of θ, and θ^0. Thus, from (A9.4) $\frac{d\bar{p}^u}{d\hat{s}}\frac{\hat{s}}{\bar{p}^u} = -\frac{1}{1+\varepsilon}$, and (A2.3) and (A2.4) again hold for the dynamic model when $\eta = 1$.

APPENDIX 10

PROOF OF PROPOSITION 3′

$$\overline{MCS}^P(\hat{s}) - \overline{MCS}^I(\hat{s}) = \int_y \left(\overline{MCS}_y^P(\hat{s}) - \overline{MCS}_y^I(\hat{s})\right) d\Gamma(y), \tag{A10.1}$$

where by (2.25) with $\eta = 1$, and Bayes' rule

$$\overline{MCS}_y^P(\hat{s}) - \overline{MCS}_y^I(\hat{s}) \equiv \int_m \left\{ \int_n \int_\sigma np^e(n,\sigma)^{1-\varepsilon} dF_{ym}(n,\sigma) - \bar{n}_{ym}\left(\bar{p}_{ym}^u\right)^{1-\varepsilon} \right\} dM(m|y).$$

Substituting for p with (2.11),

$$\overline{MCS}_y^P(\hat{s}) - \overline{MCS}_y^I(\hat{s}) \equiv \int_m \left\{ \left(\frac{\delta}{\hat{s}}\right)^{\frac{1-\varepsilon}{1+\varepsilon}} \int_n \int_\sigma n\left(\frac{n}{\sigma}\right)^{\frac{1-\varepsilon}{1+\varepsilon}} dF_{ym}(n,\sigma) - \bar{n}_{ym}\left(\bar{p}_{ym}^u\right)^{1-\varepsilon} \right\} dM(m|y). \tag{A10.2}$$

For messages such that $t_L > t^*$ we have by (A9.1)

$$\begin{aligned} \bar{p}_{ym}^u &= \left\{ \frac{\beta(\alpha+\gamma)}{\beta+\gamma} \frac{1}{\hat{s}} \int_n \int_{\sigma \le n/\tilde{\theta}_y} \frac{n}{\sigma} dF_{ym}^u(n,\sigma) \right\}^{\frac{1}{1+\varepsilon}} \\ &> \left(\frac{\delta}{\hat{s}}\right)^{\frac{1}{1+\varepsilon}} \left\{ \int_n \int_\sigma \frac{n}{\sigma} dF_{ym}^u(n,\sigma) \right\}^{\frac{1}{1+\varepsilon}} \\ &= \left(\frac{\delta}{\hat{s}}\right)^{\frac{1}{1+\varepsilon}} \left(\bar{n}_{ym}\right)^{\frac{-1}{1+\varepsilon}} \left\{ \int_n \int_\sigma \frac{n^2}{\sigma} dF_{ym}(n,\sigma) \right\}^{\frac{1}{1+\varepsilon}}. \end{aligned} \tag{A10.3}$$

Call the integrand in (A10.2) $\Delta MCS(m)$. If $\varepsilon < 1$

$$\Delta MCS(m) < \left(\frac{\delta}{\hat{s}}\right)^{\frac{1-\varepsilon}{1+\varepsilon}} \left\{ \int_n \int_\sigma n\left(\frac{n}{\sigma}\right)^{\frac{1-\varepsilon}{1+\varepsilon}} dF_{ym}(n,\sigma) - \left\{ \int_n \int_\sigma n \, dF_{ym}(n,\sigma) \right\}^{\frac{2\varepsilon}{1+\varepsilon}} \left\{ \int_n \int_\sigma \frac{n^2}{\sigma} dF_{ym}(n,\sigma) \right\}^{\frac{1-\varepsilon}{1+\varepsilon}} \right\}. \tag{A10.4}$$

The RHS of (A10.4) has the form

$$\int_x \int_z x^{\alpha} z^{1-\alpha} dF_{ym}(n,\sigma) - (\bar{x}_m)^{\alpha}(\bar{z}_m)^{1-\alpha},$$

where

$$x \equiv m \ , \quad z \equiv \frac{n^2}{\sigma}, \ \alpha \equiv \frac{2\varepsilon}{1+\varepsilon} \ ,$$

$$\bar{x}_m = \int_n \int_\sigma x \ dF_{ym}(n,\sigma) \ , \ \bar{z}_m = \int_n \int_\sigma z \ dF_{ym}(n,\sigma).$$

Since $\alpha \in (0,1)$ when $\varepsilon < 1$, the RHS of (A10.4) < 0 by Hölder's inequality and hence $\Delta MCS(m) < 0$. If $\varepsilon > 1$, the inequality in (A10.4) is reversed and $\Delta MCS(m) > 0$.

For messages such that $t_L = t^*$ we have by (A9.4) and (A9.5)

$$\bar{p}^u_{ym} = \left\{ \frac{\beta\theta^0_{ym}}{\hat{s}} \right\}^{\frac{1}{1+\varepsilon}} ,$$

where

$$J^\theta_{ym}(\theta^0_{ym}) + \frac{1}{\theta^0_{ym}} \int_{\theta^0_{ym}} \theta dJ^\theta_{ym}(\theta) - \frac{\alpha+\beta+\gamma}{\alpha+\gamma} = 0.$$

By (A7.8) and (A8.7)

$$\theta^0_{ym} = \frac{\gamma}{\beta+\gamma} \theta'_{ym} \geq \frac{\gamma}{\beta+\gamma} \bar{\theta}^u_{ym}.$$

So

$$\bar{p}^u_{ym} \geq \left\{ \frac{\beta\gamma}{\beta+\gamma} \frac{\bar{\theta}^u_{ym}}{\hat{s}} \right\}^{\frac{1}{1+\varepsilon}}$$

$$= \left(\frac{\delta}{\hat{s}}\right)^{\frac{1}{1+\varepsilon}} \left(\bar{n}_{ym}\right)^{\frac{-1}{1+\varepsilon}} \left\{ \int_n \int_\sigma \frac{n^2}{\sigma} \ dF_{ym}(n,\sigma) \right\}^{\frac{1}{1+\varepsilon}}.$$

Since this is the same inequality as (A10.3), the results for $t_L > t^*$ apply also for $t_L = t^*$.

We thus have

$$\Delta MCS(m) \gtreqless 0 \quad \text{as} \quad \varepsilon \gtreqless 1 \qquad \forall\ m.$$

By (A10.2)

$$\overline{MCS}_y^P(\hat{s}) \lesseqgtr \overline{MCS}_y^I(\hat{s}) \quad \text{as} \quad \varepsilon \lesseqgtr 1 \qquad \forall\ y.$$

By (A10.1)

$$\overline{MCS}^P(\hat{s}) \lesseqgtr \overline{MCS}^I(\hat{s}) \quad \text{as} \quad \varepsilon \lesseqgtr 1.$$

Finally, by Prop. 2′

$$\overline{CS}^P(\hat{s}) \geq \overline{CS}^I(\hat{s}),$$

and

$$\hat{s}_*^P \gtreqless \hat{s}_*^I \text{ as } \varepsilon \gtreqless 1. \quad \text{QED.}$$

APPENDIX 11

In this appendix we show that for the bivariate distribution of capacity given in (3.35), $(\bar{p}^u)^{1-\varepsilon}$ varies with π as shown in Figure 5.

The first step is to determine when $t_L > t^*$. The condition given in (3.27) is $\tilde{\phi} > \frac{\gamma}{\beta+\gamma}\hat{\phi}$, which can be written

$$\tilde{\theta} > \frac{\alpha+\gamma}{\beta+\gamma}\int_{\tilde{\theta}} \theta dJ_y^{\theta}(\theta). \tag{A11.1}$$

If $\pi > \frac{\gamma}{\alpha+\gamma}$ then $\tilde{\theta} = \frac{n}{\sigma}$ and (A11.1) is always satisfied. If $\pi \le \frac{\gamma}{\alpha+\gamma}$ then $\tilde{\theta} = n$, and (A11.1) reduces to

$$\pi < \pi_c \equiv \frac{\beta}{\alpha+\gamma}\,\frac{\sigma}{1-\sigma}. \tag{A11.2}$$

Thus, $t_L > t^*$ if $\pi \in [0, \pi_c) \cup (\frac{\gamma}{\alpha+\gamma}, 1]$. The complementary interval $[\pi_c, \frac{\gamma}{\alpha+\gamma}]$, within which $t_L = t^*$, is nonempty if $\sigma < \frac{\gamma}{\beta+\gamma}$, a condition we assume holds.

Now, for $t_L > t^*$ we have by (A10.2)

$$\bar{p}^u = \left\{\frac{\beta(\alpha+\gamma)}{\beta+\gamma}\,\frac{1}{\hat{s}}\int_{\tilde{\theta}} \theta dJ^{\theta}(\theta)\right\}^{\frac{1}{1+\varepsilon}}. \tag{A11.3}$$

If $\pi < \pi_c$, (A11.3) reduces to

$$\bar{p}^u = \left\{\frac{\delta n}{\hat{s}\sigma}\right\}^{\frac{1}{1+\varepsilon}}\left\{1 - \frac{(1-\pi)(\alpha+\gamma)-\alpha}{\gamma}(1-\sigma)\right\}^{\frac{1}{1+\varepsilon}}, \tag{A11.4}$$

and if $\pi > \frac{\gamma}{\alpha+\gamma}$

$$\bar{p}^u = \left\{\frac{\delta n}{\hat{s}\sigma}\right\}^{\frac{1}{1+\varepsilon}}. \tag{A11.5}$$

If $\pi \in [\pi_c, \frac{\gamma}{\alpha+\gamma}]$, $t_L = t^*$ and by (A10.4) and (A10.5)

$$\bar{p}^u = \left\{\frac{\beta\theta^0}{\hat{s}}\right\}^{\frac{1}{1+\varepsilon}},$$

where θ^0 is defined by the equation

$$J_\theta(\theta^0) + \frac{1}{\theta^0}\int_{\theta^0} \theta dJ^\theta(\theta) - \frac{\alpha+\beta+\gamma}{\alpha+\gamma} = 0.$$

(A11.3) reduces to

$$\bar{p}^u = \left\{\frac{\delta n}{\hat{s}\sigma}\right\}^{\frac{1}{1+\varepsilon}}\left\{\frac{(\beta+\gamma)(\alpha+\gamma)}{\gamma[\beta+(\alpha+\gamma)\pi]}\pi\right\}^{\frac{1}{1+\varepsilon}}. \tag{A11.6}$$

Figure 5 follows by plotting (A11.4), (A11.6) and (A11.5) over the three respective intervals $[0,\pi_c)$, $[\pi_c, \frac{\gamma}{\alpha+\gamma}]$ and $(\frac{\gamma}{\alpha+\gamma}, 1]$. It is straightforward to establish

$$\lim_{\pi\downarrow\pi_c} \left.\frac{d(\bar{p}^u)^{1-\varepsilon}}{d\pi}\right|(A11.6) \gtrless \lim_{\pi\uparrow\pi_c} \left.\frac{d(\bar{p}^u)^{1-\varepsilon}}{d\pi}\right|(A11.4) \quad \text{as} \quad \varepsilon \lessgtr 1,$$

so that the curves in Figure 5 are kinked at $\pi = \pi_c = \frac{\beta}{\alpha+\gamma}\frac{\sigma}{1-\sigma}$ as shown.

ENDNOTES

[1]Alexander Pope, from "An Essay in Criticism", in Nims (1981, p.202).

[2]De Vany and Saving (1977) also allow for unpredictable fluctuations in road capacity.

[3]ATIS can provide both historical information on *recurring* patterns of congestion, and real-time information about *current* driving conditions. Information can be conveyed in the form of a list of *options* amongst which the driver is free to choose, or as *directions* on when to depart, which route to take, *etc.*

A number of field experiments have been conducted at various sites on the impact of ATIS on congestion and accidents. Among these can be mentioned the Autoguide project in London, the Comprehensive Automobile Traffic Control (CACS) study carried out by MITI in Japan, the ALI-SCOUT Destination Guidance System in Germany, the European PROMETHEUS and DRIVE projects, and the U.S. ETAK system which is being tested in the San Francisco and Los Angeles areas. Descriptions of these projects, and the technologies involved, are found in Boyce (1988), OECD (1988) and Hoffmann (1991).

Simulation studies of ATIS have also been conducted (Koutsopoulos and Lotan (1989), Sullivan and Wong (1989), Gardes and May (1990), Hamerslag and Van Berkum (1991), Mahmassani and Jayakrishnan (1991), *inter alios*). While these studies have been insightful, most have suffered from one or more limitations or conceptual problems. Some have focused on the impact of information on test vehicles while ignoring the general equilibrium effects of information on traffic overall. Others have overlooked the incentive compatibility constraint that drivers will follow advice only if they believe it will benefit them. A third problem is that the traffic engineering models employed are usually complex and require numerical solution, so that basic economic insights are obscured.

[4]Perverse effects may also occur if a large number of individuals react to information, by switching routes for example. Ben-Akiva, de Palma and Kaysi (1991) refer to such out-of-equilibrium behavior as *overreaction*. Overreaction has been produced in a simulation model by Mahmassani and Jayakrishnan (1991) in the case of one-shot adjustments, and by Ben-Akiva *et al.* (1986) in the case of oscillatory behavior.

[5]The distinction between delay and loss systems is not altogether clear-cut. For example, in a delay system some potential users may choose not to participate if congestion is heavy. However, individuals often make a substantial commitment in time and/or money to use a facility before they observe the level of congestion, and may follow through with their plans even if they regret having made them. We shall assume so here; that is users do not balk.

[6]There is a strong parallel between highways and recreational areas in the nature of the physical interaction between users (Sandler and Tschirhart (1980, Table 2)). For example, flow congestion occurs on highways when vehicle densities exceed a critical level. Similarly, flow congestion occurs on crowded ski slopes, and on nature trails when hiking and backpacking parties meet or overtake one another. When flows are substantial, queueing can occur on highways at lane drops, bridges and other bottlenecks. Similarly, queues can develop at recreational facilities at entry points (*e.g.* ski lifts), intersections, and parking lots. These similarities are exemplified by the computer model Smith and Krutilla (1976) used to simulate movements of users through recreational areas. Indeed, the model has been described as a "traffic simulation model" (Cicchetti and Smith (1976, p.199)).

There are also parallels in recreation to Advanced Traveller Information Systems for drivers. State wildlife management agencies maintain data bases on hunting and fishing success rates by region. The U.S. National Weather Service keeps extensive historical data on wind patterns, temperatures, cloud formations and rainfall that can be used by recreationists to make long-range travel plans. And attempts to promote efficient usage of recreational areas have been made by disseminating current information to users (Shechter and Lucas (1978)).

[7]Traffic studies have found that nonrecurring congestion (due to incidents, traffic signal failures, *etc.*) contributes as much or more to traffic delay than does recurrent congestion (Ju *et al.* (1987), Lindley (1987), OECD (1988)). Thus, stochasticity is quantitatively important in the context of urban auto congestion. In studies of ATIS, this stochasticity has been captured by treating travel times on links of the road network as stochastic (*e.g.*, Tsuji *et al.* (1985)) or by modeling incidents or other shocks explicitly (Gardes and May (1990)).

Rush-hour congestion provides the principal motivation for our modeling approach. Since most drivers during rush hour are commuters, who have considerable experience with the system, the rational expectations assumption appears reasonable.

An alternative treatment of uncertainty is to assume that nature is deterministic but that individuals are imperfectly informed. Information is assumed to reduce the variance of users' errors in perception. Koutsopoulos and Lotan (1989) and Mahmassani and Jayakrishnan (1991) took this approach in modeling ATIS. Haltiwanger and Waldman (1985) adopted an intermediate treatment in which there are two groups -- the 'sophisticated' who have rational expectations, and the 'naive' who hold the same, incorrect beliefs.

[8]The form of the message system is similar to that considered by Nelson and Winter (1964) and Marschak and Miyasawa (1968). If demand or capacity is serially correlated, users can update the respective probability distributions using the recent history of realized states. Autocorrelation will not be treated explicitly here.

[9]While unbiasedness would appear to be a *sine qua non* of a message system, it does not always hold in practice. For example, it has been observed that weather forecasters often over-forecast precipitation. Broadus and Solow (1988) suggest that this could be because forecasters apply an asymmetric loss function to their forecast errors.

[10]See, for example, Cooper (1981, p.57).

[11]We say 'prospective' because if demand is elastic, individuals may be dissuaded from usage by a high (expected) cost.

[12]Throughout the paper, $\int x(n)dn$ and $\int_n x(n)dn$ indicate integration over the full range of n. $\int_z x(n)dn$ indicates integration from z to the upper limit of n.

[13]Note from (2.3) and (2.9) that if the message system is unbiased from the perspective of an outside observer, it is *not* unbiased from the perspective of users unless $\bar{n}_{ym} = n$, $\forall$ y, m, n. This condition is met only if the message system conveys perfect information about n.

[14]In the case of electricity demand, which has strong and regular cycles, it is optimal to have a 'diverse' technology (Chao (1983)). The base load is served by capacity with high fixed costs but low marginal cost, whereas demand peaks are met by capacity with low fixed but high marginal costs. The same may be true of other facilities. The assumption that there is only one dimension to capacity, with a constant marginal cost, is made for simplicity.

[15]We are grateful to Professor Stan Teply of the University of Alberta Civil Engineering Department for clarifying the rules used by traffic engineers for road capacity design. According to the Highway Capacity Manual (hereafter HCM), Section 2, it is standard practice in the U.S. to design roads to provide a given quality, or 'level', of service (there are 6 levels, ranked from A to F) at a particular rate of usage. For urban roads the usage rate typically chosen is that which occurs at the 10th or 20th busiest hour of the year. For rural roads the critical hour corresponds to a lower usage (*e.g.* the 50th busiest hour) because of the relatively low traffic experienced on rural roads over much of the year. Short-term variations in demand, such as those experienced during a rush hour, are accounted for by use of a peak-hour factor (PHF), defined as the ratio of the total volume occurring during an hour to the peak flow (expressed as an hourly flow rate) during a shorter time period, typically between 5 and 15 minutes (Institute of Transportation Engineers (hereafter ITE), p.475). The PHF, which by definition is ≤ 1, is used to scale down the peak measured service volume to a more realistic measure of sustainable design capacity.

This procedure for capacity design is flawed in several respects, as has been argued forcefully by Ben-Akiva *et al.* (1985). First, it allows for demand variability only by considering a particular fractile of the demand distribution. In contrast, the formula given in (2.15) indicates that the whole demand distribution is relevant. Use of such a rule of thumb might be justified by data collection and processing costs. However, data on the complete distribution of traffic flow is often assembled (HCM, Section 2).

Second, there is no clear indication in ITE or HCM why a particular target level of service is chosen. Third, the design procedures ignore capacity fluctuations. Incidents are a major cause of road capacity reductions, and as noted earlier may contribute more to delay than recurrent peak-period congestion. Yet the design of roads to meet a given level of service is based on the assumption of no incidents (HCM, 6-10). Road capacity is also affected by weather. Several studies have found that precipitation reduces effective capacity appreciably. To a lesser extent, so does poor visibility. Yet the HCM procedures do not take weather into account, though the manual does recommend this be done in areas where bad weather is common (HCM, 2-10, 6-15). Finally, the design procedure ignores the monetary costs of building capacity.

Rules analogous to the nth busiest hour of the HCM are used in the design of parking facilities: see Frantzeskakis (1982, p.22) and Smith (1983, p.441). Parallel rules are also used to choose reserve levels for storable outputs. For example, British Gas holds sufficient gas reserves to meet demand in a cold winter occurring once in 50 years (Cannon (1987)). Similarly, water utilities may construct sufficient reservoir capacity to meet a once-in-50-years drought (Crew and Kleindorfer (1986, p.260)). In cases such as these, where only extreme values of the distribution matter, decision rules based on a particular fractal of the distribution may be justifiable.

[16]Work by McFadden (1974), Pucher and Rothenberg (1976) and Small (1983) suggests that the elasticity is about 0.2.

[17]In Section 2.2 it was assumed that $F_y(n,\sigma)$ is independent of design capacity. We now assume this is true of $F_{ym}(n,\sigma)$ and $M(m|y)$.

[18]It is more common in the traffic engineering and transport economics literatures to assume flow rather than queueing congestion. However, several recent empirical studies have found that travel speed on freeways declines only slightly with flow until capacity is approached, and that the discharge rate of vehicles from a queue is equal to or only slightly below free-flow capacity (see, for example, Hurdle and Datta (1983), Hurdle and Solomon (1986), Banks (1990) and Hall and Hall (1990)). This is consistent with the properties of the bottleneck model that travel time not spent in queues is constant, and that the maximum service rate of a bottleneck is the same with and without a queue.

[19]The term 'schedule *delay* cost', which is standard in the literature, is perhaps misleading in that it refers to the cost of being early as well as late. An insightful discussion of schedule delay in air travel is found in Douglas and Miller (1974); for a welfare analysis see Panzar (1979).

In the case of commuting to work it is natural to express time-of-use preferences in terms of arrival time, when use of the road *ends*. In other contexts, *e.g.* the afternoon commute and perhaps long-distance telephone calling, preferences may be more strongly associated with the time use is initiated.

The assumption that everyone has the same t^* is made for simplicity. While the shape of the distribution of t^* affects equilibrium schedule delay costs, it does not affect the evolution of the queue provided the distribution is not too spread out; see Vickrey (1969), Newell (1987) and Arnott *et al.* (1988).

[20]The case $\alpha \le \beta$, for which an equilibrium can exist only when users arrive *en masse*, is discussed in Arnott *et al.* (1985).

[21]The perfect information equilibrium of the dynamic model can be treated using a static model whatever the form of the SDC function. Let $D(t - t^*)$ be an arbitrary SDC function. In equilibrium, $C = D(t_0 - t^*) = D(t_0 + \frac{N}{s} - t^*)$, which can be solved for t_0 and an equilibrium cost function $\bar{C}(\frac{N}{s})$.

[22]The assumption that $\rho_{ym}(t)$ is the same on all occasions given y and m can be justified by the law of large numbers if there are many users choosing arrival times independently.

[23]In Figure 4 it is assumed that $t_L > t^*$, although as will be shown $t_L = t^*$ is possible.

[24]Since this argument does not depend on the form of the schedule delay cost function, the result $t_L \ge t^*$ is true generally.

[25]Small (1982, Table 2, model 1) reports estimates of $\beta/\alpha = 0.61$ and $\gamma/\alpha = 2.38$, which yield $\frac{\gamma}{\beta+\gamma} \cong 0.80$. The behavior shown in Figure 5 then results if more than about 20% of design capacity is lost in a capacity reduction.

[26]The nonconcavity here is sharper than that identified by Radner and Stiglitz (1984), where the value of information *gross* of the cost of acquiring it is convex, but nonnegative.

[27]If $\varepsilon > 0$, and $|\pi-\pi_c|$ is small but positive (or if $\varepsilon = 0$, and $\pi - \pi_c$ is small but positive), it is possible that both weakly informative and fully informative message systems improve welfare, but a moderately informative system reduces it. The value of information can thus be multiple-peaked.

[28]Note that $\bar{p}^u(\pi)$ in Figure 5 is constant for $\pi \in [\frac{\gamma}{\alpha+\gamma}, 1]$. This follows from the fact that equilibrium costs are invariant to the distribution of ϕ below $\tilde{\phi}$. Changes in π within this range induce changes in the arrival *rate*, but not the arrival *interval*.

[29]Wilson (1989) has shown that efficient rationing can sometimes be effected by priority service pricing without need for either a spot market or for users to be informed in advance of consumption. Priority pricing is more suited to loss systems such as electricity service, where supply can be interrupted instantaneously and temporarily, than delay systems, where individuals often incur costs of commitment to usage before learning what state has been realized, and where the set of individuals using the facility and their relative preferences for service reliability are less stable over time.

[30]A conceptual analysis of multiple congested recreational areas has been undertaken by McConnell and Sutinen (1984). We have extended the bottleneck queueing model to a travel corridor with two routes in parallel in Arnott *et al.* (1990b, 1991a), and to two routes and heterogeneous users in Arnott *et al.* (1992).

REFERENCES

Arnott, R., A. de Palma and R. Lindsey (1985), "Economics of a Bottleneck", Queen's University, Institute for Economic Research, discussion paper #636.

Arnott, R., A. de Palma and R. Lindsey (1988), "Schedule Delay and Departure Time Decisions with Heterogeneous Commuters", Transportation Research Record 1197, 56-67.

Arnott, R., A. de Palma and R. Lindsey (1990a), "Economics of a Bottleneck", Journal of Urban Economics 27, 111-130.

Arnott, R., A. de Palma and R. Lindsey (1990b), "Departure Time and Route Choice for Routes in Parallel", Transportation Research B 24B(3), 209-228.

Arnott, R., A. de Palma and R. Lindsey (1991a), "Does Providing Information to Drivers Reduce Traffic Congestion?", Transportation Research A 25A(5), 309-318.

Arnott, R., A. de Palma and R. Lindsey (1991b), "A Structural Model of Peak-Load Congestion: A Traffic Bottleneck with Elastic Demand", American Economic Review, forthcoming.

Arnott, R., A. de Palma and R. Lindsey (1992), "Route Choice with Heterogeneous Drivers and Group-Specific Congestion Costs", Regional Science and Urban Economics 22, 71-102.

Banks, J.H. (1990), "Flow Processes at a Freeway Bottleneck", paper presented at the 69th Annual Meeting of the Transportation Research Board, Washington, D.C.

Ben-Akiva, M., M. Hirsh and J. Prashker (1985), "Probabilistic and Economic Factors in Highway Geometric Design", Transportation Science 19(1), 38-57.

Ben-Akiva, M., A. de Palma and P. Kanaroglou (1986), "Dynamic Model of Peak Period Traffic Congestion with Elastic Arrival Rates", Transportation Science 20, 164-181.

Ben-Akiva, M., A. de Palma and I. Kaysi (1991), "Dynamic Network Models and Driver Information Systems", Transportation Research A 25A(5), 251-266.

Boyce, D. (1988), "Route Guidance Systems for Improving Urban Travel and Location Choices", Transportation Research A 22(A), 275-281.

Broadus, J.M. and A.R. Solow (1988), "Towards a Policy for Global Climate Change", Association of Environmental and Resource Economics Newsletter 8(2), November, 3-8.

Brown, G. Jr. and M. B. Johnson (1969), "Public Utility Pricing and Output under Risk", American Economic Review 59 (March), 119- 128.

Cannon, C.M. (1987) "Peak-pricing and Self-Rationing of Gas", Energy Economics 9(2), 99-103.

Chao, Hung-Po (1983), "Peak Load Pricing and Capacity Planning with Demand and Supply Uncertainty", Bell Journal of Economics 14(1), 179-190.

Cicchetti, C.J. and V. Kerry Smith (1976) "The Measurement of Individual Congestion Costs: An Economic Application to Wilderness Recreation", in Steven A.Y. Lin (ed.) Theory and Measurement of Economic Externalities, New York: Academic Press, 183-200.

Coate, S. and J.C. Panzar (1988), "Public Utility Pricing and Capacity Choice under Risk: A Rational Expectations Approach", mimeo.

Cooper, R.B. (1981), Introduction to Queueing Theory 2nd ed. New York: North-Holland.

Crew, M.A. and P.R. Kleindorfer (1986) The Economics of Public Utility Regulation, Cambridge Mass: MIT Press.

De Vany, A. and T. Saving (1977), "Product Quality, Uncertainty, and Regulation: The Trucking Industry", American Economic Review 67, 583-594.

De Vany, A. and T. Saving (1980), "Competition and Highway Pricing for Stochastic Traffic", Journal of Business 53 (January), 45-60.

Douglas, G. and J. Miller (1974), Economic Regulation of Domestic Air Transport: Theory and Policy, Washington: Brookings.

d'Ouville, E.L. and J.F. McDonald (1990), "Effects of Demand Uncertainty on Optimal Capacity and Congestion Tolls for Urban Highways", Journal of Urban Economics 28, 63-70.

Frantzeskakas, J. (1982), "Capacity Analysis and Divergency of Off-street Parking facilities", Transportation Quarterly 36, 21-43.

Gardes, Y. and A.D. May (1990), "Traffic Modeling to Evaluate Potential Benefits of Advanced Traffic Management and In-Vehicle Information Systems in a Freeway/Arterial Corridor", ITS, U.C. Berkeley, PATH Research report UCB-ITS-PRR-90-5.

Gaver, D.P. (Jr.) (1968), "Headstart Strategies for Combating Congestion", Transportation Science 2, 172-181.

Halati, A. and D.E. Boyce (1992), "Simulation Evaluation of In-Vehicle Navigation Systems - Methodology and Application", paper presented at the 71st annual meeting of the Transportation Research Board, Washington D.C.

Hall, F.L. and L.M. Hall (1990), "Capacity and Speed Flow Analysis of the QEW in Ontario", paper presented at the 69th Annual Meeting of the Transportation Research Board, Washington, D.C.

Haltiwanger, J. and M. Waldman (1985), "Rational Expectations and the Limits of Rationality: An Analysis of Heterogeneity", American Economic Review 75, 326-340.

Hamerslag, R. and E.C. van Berkum (1991), "Effectiveness of Information Systems in Networks with and Without Congestion", paper presented at the 70th Annual Meeting of the TRB, Washington, D.C.

Hardy, G.H., J.E. Littlewood and G. Polya (1934), Inequalities, Cambridge.

Hendrickson, C. and G. Kocur (1981) "Schedule Delay and Departure Time Decisions in a Deterministic Model", Transportation Science 15(1), 62-77.

Hoffmann, G. (1991), "Up-to-the-minute Information as We Drive - How it can Help Road Users and Traffic Management", Transport Reviews 11(1), 41-61.

Hurdle, V.F. and P.K. Datta (1983), "Speeds and Flows on an Urban Freeway: Some Measurements and a Hypothesis", Transportation Research Record, 905, 127-137.

Hurdle, V.F. and D.Y. Solomon (1986), "Service Functions for Urban Freeways: An Empirical Study", Transportation Science 20, 153-163.

Institute of Transportation Engineers (1982), Transportation and Traffic Engineering Handbook 2nd ed., Englewood Cliffs: Prentice-Hall.

Ju, R-S, A.R. Cook and T.H. Maze (1987), "Techniques for Managing Freeway Traffic Congestion", Transportation Quarterly 41(4), 519-537.

Kay, J. (1979), "Uncertainty, Congestion and Public Utility Pricing", Review of Economic Studies 46, 601-611.

Koutsopoulos, H.N. and T. Lotan (1989), "Effectiveness of Motorist Information Systems in Roadway Traffic Congestion", Department of Civil Engineering, MIT. Paper presented at the IEEE Vehicle Navigation and Information Systems Conference, Toronto.

Koutsopoulos, H.N. and H. Xu (1992), "An Information Discounting Routing Strategy for Advanced Driver Information Systems (ATIS)", paper presented at the 71st annual meeting of the Transportation Research Board, Washington D.C.

Kraus, M. (1982), "Highway Pricing and Capacity Choice under Uncertain Demand", Journal of Urban Economics 12, 122-128.

Lindley, J.A. (1987), "Urban Freeway Congestion: Quantification of the Problem and Effectiveness of Potential Solutions", ITE Journal 57(1), 27-32.

Mahmassani, H. and R. Jayakrishnan (1991), "System Performance and User Response under Real-time Information in a Congested Traffic Corridor", Transportation Research 25A(5), 293-308.

Marschak, J. and K. Miyasawa (1968), "Economic Comparability of Information Systems", International Economic Review 9(2), 137-174.

McConnell, K.E. and J.G. Sutinen (1984) "A Conceptual Analysis of Congested Recreation Sites", in V.K. Smith and A.D. White (eds.) Advances in Applied Micro-Economics Vol.3, Greenwich: JAI Press, 9-36.

McFadden, D. (1974), "The Measurement of Urban Travel Demand", Journal of Public Economics 3, 303-328.

Mohring, H. (1970), "The Peak Load Problem with Increasing Returns and Pricing Constraints", American Economic Review 60, 693-705.

Nelson, R.R. and S.G. Winter (Jr.) (1964), "A Case Study in the Economics of Information and Coordination: The Weather Forecasting System", Quarterly Journal of Economics 78, 420-441.

Newbery, D.M.G. (1990), "Pricing and Congestion: Economic Principles Relevant to Pricing Roads", Oxford Review of Economic Policy 6(2), 22-38.

Newell, G. (1987) "The Morning Commute for Nonidentical Travelers", Transportation Science 21(2), 74-88.

Nims, J.F. (1981), The Harper Anthology of Poetry, New York: Harper and Row.

OECD (1988), "Road Transport Research: Route Guidance and In-car Communication Systems", Paris.

Panzar, J. (1979), "Equilibrium and Welfare in Unregulated Airline Market", American Economic Review (Papers and Proceedings), 69, 92-95.

Pucher, J. and J. Rothenberg (1976), "Pricing in Urban Transportation: A Survey of Empirical Evidence on the Elasticity of Travel Demand", mimeo.

Radner, R. and J.E. Stiglitz (1984), "A Nonconcavity in the Value of Information", in M. Boyer and R.E. Kihlstrom (eds.), Bayesian Models in Economic Theory, Amsterdam: North-Holland, 33-52.

Reiss, R.A., N.H. Gartner and S.L. Cohen (1991), "Dynamic Control and Traffic Performance in a Freeway Corridor: A Simulation Study", Transportation Research 25A(5), 267-276.

Rothschild, M. and J.E. Stiglitz (1970), "Increasing Risk: I: A Definition", Journal of Economic Theory 2, 225-243.

Sandler, T. and J. Tschirhart (1980), "The Theory of Clubs: A Survey", Journal of Economic Literature 43(4), 1481-1521.

Saving, T.R. and A.S. De Vany (1981), "Uncertain Markets, Reliability and Peak-Load Pricing", Southern Economic Journal 47(4), 908-923.

Schelling, T. (1978), Micromotives and Macrobehavior, Norton.

Shechter, M. and R.C. Lucas (1978), Simulation of Recreational Use for Park and Wilderness Management, Baltimore: Johns Hopkins University Press for Resources for the Future.

Small, K.A. (1982), "The Scheduling of Consumer Activities: Work Trips", American Economic Review 72(3), 467-479.

Small, K.A. (1983), "The Incidence of Congestion Tolls on Urban Highways", Journal of Urban Economics 13, 90-111.

Smith, W.S. (1983), "Automobile Parking Trends", Transportation Quarterly 37, 431-452.

Smith, V.K. and J.V. Krutilla (1976), Structure and Properties of a Wilderness Travel Simulator: An Application to the Spanish Peaks Area, Baltimore: Johns Hopkins University Press for Resources for the Future.

Steiner, P. (1957), "Peak Loads and Efficient Pricing", Quarterly Journal of Economics 71, 585-610.

Sullivan, E.C. and S. Wong (1989), "Development of a Dynamic Equilibrium Assignment Procedure for Network-Level Analysis of New Technology", ITS, U.C. Berkeley, PATH Research Report UCB-ITS-PRR-89-4.

Transportation Research Board (1985), Special Report 209: Highway Capacity Manual TRB, National Research Council, Washington, D.C.

Tsuji, H., R. Takahashi, H. Kawashima and Y. Yamamoto (1985), "A Stochastic Approach for Estimating the Effectiveness of a Route Guidance Systems and its Related Parameters", Transportation Science 19(4), 333-351.

Vickrey, W.S. (1969), "Congestion Theory and Transport Investment", American Economic Review (Papers and Proceedings), 59, 251-261.

Visscher, M. (1973), "Welfare-Maximizing Price and Output with Stochastic Demand", American Economic Review 63, 224-229.

Williamson, O. (1966), "Peak Load Pricing and Optimal Capacity under Indivisibility Constraints", American Economic Review 56, 810-827.

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