# Congestion Pricing and Queueing Theory

## Congestion Pricing

Congestion costs due to any specific user have 2 components:

(1) Cost of delay to that user (internal)

(2) Cost of delay to all other users caused by that user (external)

--At congested airports (and congested facilities, in general) this second component can be very large

--A congestion toll can be imposed to force users to experience this cost component (to internalize the external costs)

## Economic principle

Optimal use of a transportation facility cannot be achieved unless each additional (marginal) user pays for all the additional costs that this user imposes on all other users and onthe facility itself. Thus, a congestion toll not only contributes to a socially desirable result, but is necessary to reach such a result. (Vickrey, 1967; Carlin + Park, 1970)

## Congestion Pricing and Queueing Theory

In practice it is very hard to:

(1) Estimate external marginal delay costs (extensive data analysis or difficult simulation is typically needed);

(2) Determine equilibrium congestion tolls (trial-and error approach that may take long time to converge is used sometimes).

Queueing theory has much to offer in these two respects under certain conditions.

**The Principal Observation**

Consider a queueing facility with a single type of customer in steady-state.

Let

$c$ = delay cost per unit time per customer

$C$ = total cost of delay per unit time incurred in the system at equilibrium

Then: $C = cL_q = c\lambda W_q$

and the marginal delay cost, *MC*, imposed by an additional ("marginal") customer is given by:

$$MC = \frac{dC}{d\lambda} = c\ W_q + c\lambda \frac{dW_q}{d\lambda}$$

Note that the first term on the right is the "internal cost" experienced by the marginal customer and the second term is the "external cost" (s)he imposes!

These ideas can be extended to cases with multiple types of customers and to systems with priorities.

**Some Definitions**

| | |
|---|---|
| $\lambda_i$ | -demand rate of type *i* customers |
| $\lambda = \sum_{i=1}^{m} \lambda_i$ | -total demand rate |
| $S_i$ | -service time for type *i* customers |
| $\mu_i$ | -service rate for type *i* customers, $\mu_i^{-1} = E[S_i]$ |
| $S$ | -overall service time of customers |
| $\frac{1}{\mu} = E[S] = \sum_{i=1}^{m} \left( \frac{\lambda_i}{\lambda} \times \frac{1}{\mu_i} \right)$ | -expected overall service time |
| $\rho = \frac{\lambda}{\mu} = \sum_{i=1}^{m} \rho_i = \sum_{i=1}^{m} \frac{\lambda_i}{\mu_i}$ | -overall utilization ratio |
| $c_i$ | -delay cost per time unit for type *i* customers |
| $c = \sum_{i=1}^{m} \left( \frac{\lambda_i}{\lambda} c_i \right)$ | -average delay cost per unit time per customer |

$L_q$ -expected number of customers in queue at equilibrium ("steady-state")

$W_q$ -expected queueing time per customer at equilibrium

$C$ -total cost of delay per unit time incurred in the system in equilibrium

**A Simple Observation (No priorities among customer types)**

For a queueing system in equilibrium (need $\rho < 1$):

$L_q = \lambda W_q$ [Little's Law]

Therefore, $C = cL_q = c\lambda W_q$

and

$$MC(i) = \frac{dC}{d\lambda_i} = c_i W_q + c\lambda \frac{dW_q}{d\lambda_i} \qquad (1)$$

**Implication of (1)**

For many types of queueing systems explicit expressions for $W$ are available. (1) can then be used to compute MC and marginal external costs.

Example: M/G/1 system

$$MC(i) = \frac{dC}{d\lambda_i} = c_i \frac{\lambda \cdot E[S^2]}{2(1-\rho)} + c\lambda \frac{(1-\rho)E[S_i^2] + \frac{\lambda}{\mu_i} E[S^2]}{2(1-\rho)^2} \qquad (2)$$

**Extension 1**

• Similar analysis can be applied (and closed form results obtained) in cases in which customers are assigned priorities for service depending on their type.

• Note that in this case, each type of customer $i$ experiences a different expected time in the system, $W_{qi}$, depending on their priority.

• Many important practical applications.

**Extension 2**

• Let $\lambda_i(x_i)$ be the demand rate by type $i$ customers when the total cost of using the facility (internal costs plus external costs) is equal to $x_i$.

• If the functions $\lambda_i(x_i)$ are known for all *i,* then we can compute the *equilibrium congestion tolls* by solving a system of *m* equations of the form

$$x_i = c_i W_{qi}(\hat{x}) + \left( \sum_{j=1}^{m} c_j \lambda_j(x_j) \right) \frac{dW_q(\hat{x})}{d\lambda_i(x_i)} + K_i \qquad \forall i \qquad (3)$$

where $\hat{x} = \{x_1, x_2, ..., x_m\}$ and the $K_i$ are constants.

**Some Additional Issues**

Toll may vary in time and by location

Facility users may be driven by "network" considerations

"Social benefit" considerations

May have to achieve revenue targets

Politics

**References**

Carlin, Alan and R. E. Park, "Marginal Cost Pricing of Airport Runway Capacity", American Economic Review, 60, pp. 310-318 (1970).

Vickrey, William, "Congestion Theory and Transport Investment", American Economic Review Proceedings, 59, pp. 251-260 (1969).