CONGESTION PRICING:
A PARKING QUEUE MODEL

Richard C. Larson¹ and Katsunobu Sasanuma²

¹Center for Engineering Systems Fundamentals
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139
rclarson@mit.edu

²Massachusetts Institute of Technology
Cambridge, Massachusetts 02139
sasanuma@mit.edu

August 2007
Congestion Pricing: A Parking Queue Model

Richard C. Larson*
Katsunobu Sasanuma
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139 USA
August, 2007

ABSTRACT

Congestion pricing imposes a usage fee on a public resource during times of high demand. **Road pricing** involves cordoning off a section of the city and imposing a fee on vehicles that enter it. **Parking pricing** increases the costs of on-street and perhaps off-street parking. Following an historical review, we develop a new queueing model of the parking pricing problem, recognizing that many urban drivers are simply looking for available on-street parking. Often, reducing the number of such “cruising drivers” would reduce urban road congestion dramatically, perhaps as effectively as cordoning off the center city.

*Corresponding Author. Contacting information: Massachusetts Institute of Technology, Room E40-231b, Cambridge, Massachusetts 02139. E-mail: rclarson@mit.edu. Phone: 617 253 3604; Fax: 617 258 7733.

**Acknowledgements**
We are grateful for helpful discussions and valuable information from Amedeo R. Odoni, Fred Salvucci, Akira Hibiki, and Robert Irwin. We also thank Cordell Hull for awarding a research grant to MIT to help support this research. In addition, we thank the MIT-Portugal Program, which supported this research, as well. Finally, we thank Judith Stitt for her professional excellence in helping to edit this paper.
1. Introduction

Consider a typical recent study of street traffic congestion in urban America. This particular report\(^1\) is from the Park Slope section of Brooklyn, New York, a thriving commercial and residential zone. The purpose of the study was “…to ascertain the extent of the neighborhood’s ever-worsening traffic and parking problems and to propose solutions to both.” Based on data collected early in 2007, “…the study reveals an overwhelming amount of traffic is simply circling the block “cruising” for parking, while the curbside itself is nearly 100% filled with parked vehicles.” The researchers found that 45% of total traffic and 64% of local traffic is cruising for a parking space. And the average curb occupancy rate is 94%, with “…nearly 100% occupancy at metered spaces during peak periods.”

Street congestion in a growing city is unavoidable. Incentives to use cars for ease, comfort, and other pleasures increase as cars become more affordable, even when an extensive public transportation system is available. As a result, people drive more frequently than necessary, leading to overuse of limited road resources. In rural areas, sufficient road networks can be provided to prevent road congestion because land is abundant and demand is less. For large cities, however, supplying new roads is more difficult because of the huge demand for land resources. As a result, major city streets become crowded and inconvenient despite the availability of public transportation, resulting in economic losses to the community.

Urban congestion is a challenge within a complex system that requires the simultaneous consideration of many options. In most cases, the solution chosen combines several initiatives such as the congestion pricing (CP) schemes described in this paper and other incentives to motivate people’s serious consideration of substituting public transportation, bicycles, or walking for personal motorized vehicles. The particular combination of options chosen depends on a number of factors, notably the city of concern and those areas most
affected within it.

Road Pricing (RP), one economic measure used in traffic management, has attracted the attention of many mayors of heavily congested cities, especially since the successful implementation of RP schemes in Singapore and London. However, current RP schemes present two serious challenges:

(1) Schemes billed as “RP” that are now planned or implemented in many cities are not actually road pricing, but rather cordon pricing (which approximates area pricing), and thus less effective than a full-scale electronic RP scheme: Theoretically, a congestion charge (CC) should apply to every road within a charged zone in a full-scale RP scheme rather than to only those roads crossing cordon lines, as currently is often the case. Otherwise, RP is ineffective, particularly when drivers make many trips wholly within the inner city.

(2) Any type of RP is costly because a new technology infrastructure (e.g., electronic gantries along cordon lines) need to be installed to detect traffic flow, and a management organization must be established to oversee toll collection. Application of RP has therefore been limited, and has been implemented effectively only in the downtown areas of large cities.

To address these issues, we consider parking pricing (PP), which can improve traffic control by (1) increasing the average parking price in an area, effectively imposing a congestion charge on parkers, and (2) raising on-street parking prices to improve traffic flow. Since on-street parking prices influence drivers’ decisions about how much time to spend searching for available street parking spaces, these prices are especially important in congested city centers and commercial districts (such as Brooklyn’s Park Slope), where many drivers look for spaces. Parking pricing offers the following practical benefits: (1) PP does not necessarily require an additional toll-collection organization, making it cost-
effective for medium- and small-scale cities for which RP is not affordable, and (2) PP can be extended as its affected area expands without costly additional infrastructure or the risk of increasing the number of exempted residents, which could significantly reduce the effectiveness of RP. Thus, PP is not only flexible and effective, but again can be applied even in medium- and small-scale cities that are grappling with disconnected or unevenly congested areas.

Good examples of the efficacy of parking pricing can be found in Japan,^2 where at one time only police enforced parking regulations. Although on-street parking officially has not been allowed on most roads in Japan, until recently people parked almost anywhere because the number of police was insufficient to check for violators, and officers permitted a grace period of 15 minutes or more before issuing tickets. Since June 1, 2006, however, Japan has enforced a strict parking regulation in Tokyo, Osaka, and other cities that has proved equivalent to eliminating much “free” street parking. In particular, a June 1 revision of the Road Traffic Law has enabled private vendors to enforce parking regulations by issuing tickets immediately after identifying violators, without any grace period. With no grace period and such improved enforcement, the 15 minutes or more of “free” parking that had been granted drivers on most roads as a grace period (or otherwise overlooked) has risen to an expensive on-street parking which costs 10,000 yen (US$86) (as a penalty fee) for parkers. This implementation has effectively increased the on-street parking prices for drivers and as a result, increased the average parking prices in major cities in Japan at the same time. Three months after the June 1, 2006, implementation, the National Police Agency reported a 27.2% decrease in the average length of traffic jams and a 9.5% decrease in average travel time on the main streets of Tokyo, comparable to results achieved with London’s RP scheme. In addition, a modal shift from cars to public transportation was observed. In fact, such improvements were observed in not only Tokyo but also medium-sized cities throughout Japan where strict parking regulations were enforced. Tokyo’s example thus demonstrates that PP can be as effective as RP, and a cost-effective alternative for cities of all sizes.
2. Traffic Congestion

2.1 Cost of Congestion

As countries develop and the number of cars traversing them increases, the cost of congestion greatly impedes their cities’ development. It has often been observed that congestion increases as cities grow: In the United States, for example, the Texas Transportation Institute (TTI) estimated the annual delay per peak-period traveler in very large urban areas with populations of more than 3 million to be 61 h for the year 2003, which is much larger than 13 h in small metropolitan areas with populations less than 0.5 million. The average annual delay for all cities has grown from 16 h to 47 h since 1982. According to TTI’s estimate, congestion costs Americans $63.1 billion a year, based on considerations of only time and fuel wasted. The total cost of congestion should be considered in at least the following four categories, however:

(1) Waste of time: Congestion deprives businesses and individuals of work hours by increasing commuting time. According to TTI, total delays reached 3.7 billion hours in 2003, a significant part of the total loss attributed to congestion. The simplest estimation multiplies hours lost on congested roads by wage/hour. TTI used US$12.85/h for the cost of time wasted in congestion, close to the average hourly wage in the US ($11.48) for all goods-producing workers. However, this approach may overestimate the cost of delay because delay cannot be completely eliminated from business activity. Another method of estimation calculates net gains from reducing congestion. The former estimation provides an upper bound for losses due to congestion; the latter method yields a lower bound.

(2) Waste of resources and associated costs: Congestion wastes gasoline and damages
pavement. According to TTI, engines idling in congested traffic wasted 2.3 billion gallons of fuel in 2003. Gasoline wastage also contributes to the urgency of U.S. strategic interests in the Middle East oil supply, which is much costlier in many ways than the simple loss of material resources.

(3) Loss of environmental quality and associated costs: Congestion produces more air pollution and noise than does smoothly flowing traffic, degrading the environmental quality of roadside areas and consequently negatively impacting people’s health. Congestion also produces excessive carbon dioxide (CO$_2$) emissions that contribute to global warming. The reduction in CO$_2$ emissions from improved traffic flow can range from several hundreds to thousands of tons, depending on city size. The cost associated with emission credits is small, however, compared to the losses that can be considered in terms of time and fuel.

(4) Loss of business: This loss to congestion is hard to estimate because congestion is a by-product of business activity. Although many agree that excessive congestion leads to inefficiencies and reduces a city’s attractiveness, many also believe that suppressing the inflow of people might harm a city more than doing nothing since it may reduce the number of people in the city and decrease business activity. This is often the case when insufficient public transportation is offered to those who stop driving into the city after RP is implemented. In fact, business leaders have raised the most opposition to RP in London and New York.

### 2.2 Reasons for Congestion and Measures to Reduce It

(1) Demand-side problem: One obvious reason for congestion is people’s persistent demand for private automotive transportation, which can be reduced by means of either fuel tax or congestion pricing (CP) schemes—demand-side efforts. Fuel tax is not as effective as congestion pricing in this regard because it reduces car usage uniformly rather than coping with local and time-specific forms of congestion. Congestion pricing,
however, can be applied effectively to specific areas and also made time-dependent if the extra fee, or “congestion charge” (CC), is adjusted locally and dynamically in accord with real-time traffic situations. According to some economic theories, the CC level is set equal to the marginal external cost, as discussed in the second half of this paper.

Congestion pricing can be implemented via RP by imposing fees on drivers crossing cordoned lines, and/or via PP by imposing fees on drivers who park within the cordoned area. Introducing truly local, dynamic CC is costly and technically difficult, though, especially for RP: Currently available cordon-line-type RP cannot control trips taken within cordoned areas effectively since road users are charged only once per day and residents in cordoned areas are often exempted from paying full CC. Full-scale electronic RP,\textsuperscript{8} which essentially prices every road in a city, requires advanced technologies or numerous gantries in cordoned areas, making RP less cost-effective for medium-scale cities than for larger cities. For the usual RP scheme, drivers are charged fixed or pre-determined CCs, but its effectiveness is compromised. In contrast, PP can effectively reduce trips taken within a cordoned area since the numbers of trips are related to parking behaviors. Since parking lots are distributed throughout cities, PP is locally and dynamically applied much more easily than is RP. PP can be applied to disconnected, congested areas where cordon lines are hard to draw, and can be adapted smoothly as congested areas expand.

(2) Supply-side problem: Excessive demand is not the only reason for congestion. Insufficient parking capacity and inappropriate parking pricing are other reasons, requiring supply-side efforts to counteract. Two observations\textsuperscript{9} in New York City have especially interested us: (1) a recent survey conducted by Bruce Schaller, principal of Schaller Consulting, showed that 28% of drivers in the SoHo district in Manhattan were searching for on-street parking, and (2) as cited in this paper’s Introduction, a second survey, by Transportation Alternatives, showed 45% of drivers were searching for on-
street parking in the Park Slope neighborhood in Brooklyn. This, of course, is not always the case, but often is during busy times in city centers, where most drivers try to find a place to park. Historical data on the percentage of traffic cruising in selected cities are displayed in Table 1 below. While these data are somewhat dated, we might well find that the same cruising behavior observed downtown today has been in evidence since the 1920s.

Table 1.

Percentage of traffic observed to be “cruising” for parking in selected cities.

<table>
<thead>
<tr>
<th>Year</th>
<th>City</th>
<th>Share of traffic cruising</th>
</tr>
</thead>
<tbody>
<tr>
<td>1927</td>
<td>Detroit (1)</td>
<td>19%</td>
</tr>
<tr>
<td>1927</td>
<td>Detroit (2)</td>
<td>34%</td>
</tr>
<tr>
<td>1960</td>
<td>New Haven</td>
<td>17%</td>
</tr>
<tr>
<td>1977</td>
<td>Freiburg</td>
<td>74%</td>
</tr>
<tr>
<td>1985</td>
<td>Cambridge</td>
<td>30%</td>
</tr>
<tr>
<td>1993</td>
<td>New York</td>
<td>8%</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>30%</td>
</tr>
</tbody>
</table>

*aThe numbers refer to different locations within the same city.

Source: Shoup (2005); selected data.

The underlying problem is inappropriate PP when on-street parking capacity cannot accommodate all who hope to park. In Manhattan, off-street parking (averaging US$24.42 per person per day) costs 14 times more than on-street parking (which averages US$1.73 per person per day). If prices for on-street parking are much lower than those charged by off-street parking lots, drivers have strong incentive to search for parking on the street, creating extra traffic and congestion. Well-planned PP dramatically improves traffic flow in cities and decreases congestion without imposing CC on all drivers. We show the effect of
3. Implementation

3.1 Examples of RP and PP Successes Around the World

3.1.1 London, England\textsuperscript{12} (RP)
Congestion charges went into effect in London on February 17, 2003, at an initial charge of 5 UK pounds (approximately US$10) per vehicle per day. Drivers paid the charge if their vehicles entered a congestion-charging zone anytime between 7 a.m. and 6:30 p.m. weekdays. Six hundred eighty-eight cameras in 203 locations within an 8-square-mile (21-square-km) area captured the license plate numbers of about 250,000 vehicles daily. The number of vehicles entering central London during charging hours declined about 25% the day congestion charges were introduced. Since the time the RP scheme first went into effect, vehicle delays due to traffic congestion have dropped about 30% and carbon dioxide emissions have decreased more than 15%. In 2005-06, London’s RP scheme cost 230 million UK pounds (about US$460 million) to implement, while its annual operating costs were around 88 million UK pounds (about US$176 million), its net revenue was 122 million UK pounds (about US$244 million),\textsuperscript{13} most of which was spent improving bus services (London put 300 additional buses into service before introducing the congestion charges). Bus passengers entering the charging zone during morning rush hour the first year increased 37%\textsuperscript{14}.

Because of such successes as well as a need to reduce traffic further, the City of London raised its congestion charge to 8 pounds (US$16) per vehicle per business day in July 2005. Since February 2007, the congestion-charging zone has extended west. Residents in the zone can apply for a 90% discount price of 4 pounds (US$8) per business week.
The success of London’s RP is usually explained as follows: First, the center city had been heavily congested, and the citizens and mayor recognized the RP scheme should be implemented. Second, the technology for gantries to automatically read license plates was provided by the city, so citizens would not need to bear any related monetary burden such as for the in-vehicle units (IVUs) that are required in Singapore’s case. Third, even before implementation of RP in London, 85% of commuters used public transportation: more than 1 million riders per day. Hence, the additional expenses required to reinforce the public transportation system in preparation for RP were minimal, and most of the commuting public favored RP, expecting that it would improve public transportation.

3.1.2 Singapore\(^{15}\) (RP)
Congestion pricing was first adopted by the city of Singapore in 1975, using a paper license scheme for want of a more reliable technology; the planners understood the scheme’s limited effectiveness. Every vehicle containing three or fewer people was charged Singapore$3 (about US$2) per business day on any given weekday upon initial entry to the 2.3-square-mile central area of the city between 7:30 a.m. and 10:15 a.m. This scheme reduced the total peak-period traffic each business day by 45%. An electronic toll collection system (Electronic RP, or ERP) using IVUs replaced paper licenses in April 1998 to better control traffic. IVUs simplified the task of varying tolls by time of day or location. Singapore initially did not change toll levels, hours, or boundaries to minimize controversy over the charges, but gradually did start to vary tolls according to time and place; signs on gantries now inform motorists of the toll in effect. Currently Singapore’s RP scheme is closer than any other city’s to the ideal dynamic RP since the CC level changes with time and location, using an ERP system.

3.1.3 Stockholm, Sweden\(^{16}\) (RP)
An RP scheme was applied in Stockholm on January 3, 2006. The system was scheduled to run for seven months, then a vote on whether to continue was held on September 17, 2006. In that referendum, the citizens of Stockholm voted for a congestion-charging scheme:
51.7% in favor, 45.6% against.\textsuperscript{17} All parties in the city council promised to abide by the results. A fee of 10 to 60 kronor (approximately US$1.4 to US$8.5) was charged vehicles entering the inner city on weekdays between 6:30 a.m. and 6:30 p.m., payable by direct debit. Traffic volumes were reduced 25%, and the number of vehicles during peak hours fell by 100,000. At the same time, public transit rides increased by 40,000 per day. Retail sales in central Stockholm shops also rose after the congestion-pricing scheme was introduced, as people bought more locally rather than drive to suburban stores. The system used cameras, but drivers were also encouraged to install radio-frequency identification (RFID) transponders in their cars. The permanent RP phase just started in Stockholm on August 1, 2007, employing the same system used during the 7-month trial RP in 2006.

According to Prud’homme and Kopp,\textsuperscript{18} however, the Stockholm urban toll (UT) scheme does not satisfy conditions for a successful RP scheme because (1) road congestion in Stockholm is not very severe; (2) the implementation cost of RP is too high; and (3) the marginal costs for public transportation improvements are high.

3.1.4 Trondheim, Norway (RP)
In the late 1980s an RP scheme was applied in Oslo, Bergen, and Trondheim in Norway. Trondheim has used the RP scheme for nearly 20 years, since 1988. Its objectives are to not only fund new ring roads, but also improve the public transportation system and pedestrian ways, and invest in environmental measures. Currently the toll price is 7.5 to 25 Norwegian Kroner (about US$1 to US$4)—much less than the 8 UK pounds (about US$16) charged in London. Since pricing varies 24 hours each day from highs during peak hours to lows during off-peak hours, this scheme controls traffic flow to prevent congestion.\textsuperscript{19}

3.1.5 Tokyo, Japan (PP)
The city of Tokyo has been investigating RP schemes since 1999, when the current governor of the Tokyo District, Shintaro Ishihara, was first elected. One challenge has been that Tokyo is so large and congested everywhere. Many roads are congested not because of
a large inflow from the suburbs but because of numerous intracity trips: according to one estimate made by the Tokyo Metropolitan Government, about 40% of total trips in Tokyo are intracity trips.\textsuperscript{20} Therefore, RP would not be effective if cordon-tolls similar to London’s RP were implemented. PP would be more appropriate in Tokyo than a simple cordon-line RP because PP can vary charges as needed to regulate intracity trips.

The effectiveness of PP in major Japanese cities can be illustrated by the following example. Since June 1, 2006, parking violations have been regulated in such a way as to drastically reduce congestion in major cities in Japan. Earlier, parking regulations had been enforced only by police, who periodically would check to see how long cars had been parked in banned areas, marking cars’ positions with chalk and ticketing them if parked longer than 15–30 minutes. Since June 1, private firms have been consigned to issue tickets for parking violations regardless of the number of minutes vehicles have been parked in banned locations. Ticket fines range from 10,000 yen to 18,000 yen (US$86–155) for regular cars, depending on the violation.

According to the National Police Agency,\textsuperscript{21} results of the first three months of strict parking policy showed that illegal parking at main roads in Tokyo and Osaka had been reduced by 73.9\% and 73.3\%, respectively, congestion length on their main roads at 2p.m.-4p.m. had been shortened by 27.3\% and 23.1\%, respectively, and travel speed at 2p.m.-4p.m. had increased by 9.5\% and 11.8\%, respectively, compared to the same period the previous year. The agency estimated economic benefits of this policy to be 181 billion yen (US$1.6 billion) and the reduction in CO\textsubscript{2} emissions to be 15.2 thousand tons/yr. Retail shops with parking lots have also attracted more people since implementation, while popular restaurants without parking lots are said to have experienced a decrease in customers. Increases in the average prices of off-street parking lots are also observed, due to heightened demand for off-street parking. Also, more people have been using taxis and buses to reach restaurants since the strict parking policy was implemented.
3.1.6 Los Angeles, USA (PP)
Shoup examined eight case studies conducted from 1993 through 1995 on the effect of a “cash-out” scheme in the Los Angeles region\textsuperscript{22}. A cash-out scheme is one form of PP that gives employees a choice between free parking and its cash equivalent, introducing the market mechanism to companies’ free parking. This scheme does not remove a benefit from employees since they can either continue driving to work or receive a cash benefit by using public transportation. The results in Los Angeles were remarkable: after a cash-out scheme was introduced, the number of solo drivers fell 17% while the number of carpoolers rose 64%, and public transit ridership increased 50%. The number of miles traveled by private vehicles declined 12%. This program has reduced the number of cars used to commute without sacrificing the number of persons commuting, and, according to surveys taken, has increased both employers’ and employees’ satisfaction.

3.2 Examples of RP and PP Under Review or Rejected
Some cities have experienced difficulties introducing an RP scheme, or at least have required much time to consider doing so.

3.2.1 Edinburgh, Scotland (RP)
In February 2005, about 290,000 voters in and around the city of Edinburgh were asked whether the city should implement an RP scheme similar to London’s. The plan proposed was to charge 2 UK pounds (US$4) to enter the cordon area and 60 UK pounds (US$120) for violations—amounts much lower than those in London’s scheme. More than 74% of those queried rejected the scheme\textsuperscript{23}. Typical reasons Edinburgh citizens gave for rejecting it follow:\textsuperscript{24}

(1) Distrust of local government: Citizens regarded the RP scheme as a mere tool for raising revenue rather than reducing congestion. Many thought alternative ways to reduce congestion should be sought before additionally burdening citizens.
(2) Currently inadequate public transportation: Edinburgh has a poor public transport system. Citizens described it as expensive, dirty, and unreliable: an inadequate transport framework in which to implement the RP scheme.

(3) Two cordons: Because the Edinburgh proposal recommended two cordoned areas, the number of people impacted was greater than would have been the case with a simpler one-cordon plan. Even though computer simulation showed the increased congestion at two cordon lines would be minimal, people distrusted the results, worrying especially about traffic increases in residential areas and around schools between the two cordon lines.  

3.2.2 Hong Kong (RP)
Road pricing was first attempted in Hong Kong in the 1980s, using an electronic charge to control traffic. In 1983, an experiment involved 2500 vehicles for five days. Full-scale implementation was planned in 1985, but failed for two main reasons. First, people feared being identified because the Chinese government might utilize RP as a tracking tool. Second, the electronic charging system was underdeveloped.

3.2.3 New York City, USA (RP)
New York City has been considering RP schemes for many years, but none have been fully implemented. However, a recent news article indicated that Mayor Michael R. Bloomberg’s RP plan received strong support from Governor Eliot Spitzer and the Bush administration. The mayor’s plan would charge US$8 for cars and US$21 for commercial trucks entering Manhattan below 86th street between 6 a.m. and 6 p.m. on weekdays, or US$4 for all drivers within the congestion zone. However, the following concerns remain:

(1) New York might be hurt economically. A report by the Queens Chamber of Commerce released in February 2006 estimated that a US$14 congestion charge similar to London’s would reduce by 40,000 the number of people entering
Manhattan’s central business district each weekday, causing a loss of US$2.7 billion in economic output per year.

(2) There could be equity issues. Low-income drivers often lack flexibility to change a given traffic pattern because they generally have fixed work schedules and consequently difficult-to-change travel patterns. If they must drive to their required destinations, they may have no option other than to pay any CC that is imposed. However, a survey conducted in 2003 by Schaller Consulting for Transportation Alternatives and the NYPIRG Straphangers Campaign showed that most people who drive into Manhattan are wealthy. Specifically, Schaller Consulting conducted a survey regarding the East River bridges connecting Manhattan with other parts of the city. The East River bridges are inexpensive or even free, and therefore heavily congested. In order to estimate the effect of charging fees to drivers crossing the bridges, Schaller Consulting investigated the equity issue and found that lower-income people are far more likely to take transit than to drive themselves across the bridges. Since drivers crossing the bridges tend to be in the upper income ranges anyway, therefore, Schaller concluded that a toll would have little impact on lower-income drivers.

3.2.4 Boston, USA (PP)
In 1975, Boston capped the number of off-street parking spaces available downtown at 35,500 spaces, in part to reduce people’s incentive to drive downtown. As a result, Boston’s off-street parking price (averaging US$11 for the first hour) is now one of the nation’s highest, next to New York City’s and Chicago’s. However, since the price gap between on-street and off-street parking is wide, the incentive for drivers to find inexpensive street parking is high—creating extra traffic. The average savings realized by finding on-street parking in Boston is an estimated US$10/hour, again among the nation’s highest, next to New York City and Chicago. Although no data exist concerning what percentage of total
traffic is actually searching for street parking, the current pricing policy certainly does affect traffic conditions negatively on most congested Boston streets.

3.3 Issues of RP and PP
Having reviewed several examples of RP and PP, we now can analyze the main issues pertaining to these two sorts of pricing schemes as *functional* issues and *other* issues such as stakeholder issues and equity issues. Functional issues can be addressed relatively simply because they derive from RP and PP design.

Table 2.

Comparison of RP and PP

<table>
<thead>
<tr>
<th>Type</th>
<th>Road Pricing (cordoned-area RP)</th>
<th>Road Pricing (ful-scale RP)</th>
<th>Parking Pricing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scheme used</td>
<td>Congestion Charge</td>
<td>Congestion Charge</td>
<td>Congestion Charge, Price differential</td>
</tr>
<tr>
<td>Fee structure</td>
<td>Flat rate over charged area, Per day fee</td>
<td>Locally adjusted fee, Per trip fee</td>
<td>Locally adjusted fee, Per trip fee</td>
</tr>
<tr>
<td>Excluded (Exempted) groups</td>
<td>Residents in charged area, Public transportation</td>
<td>Public transportation</td>
<td>Through traffic, Public Transportation</td>
</tr>
</tbody>
</table>

Both RP and PP can be implemented efficiently if no groups are excluded (exempted) and congestion is uniform over the charged area. However, as the target area expands, the number of residents within it increases and congestion within the area becomes less uniform. In such a case, cordoned-area RP becomes less effective and full-scale RP
becomes more appropriate although associated administrative costs are huge. PP does not have these issues: residents are not exempted, and parking fees can be applied locally (i.e., varied by locale). The biggest problem with PP is when through traffic is responsible for most of the congestion, since PP cannot impose a congestion charge (CC) on each driver if he does not park. In reality, though, most large cities are serviced by an extensive network of highways. Therefore, in most cases, drivers with remote destinations never even enter local congested roads.\(^3\) (Our presentation of a PP model in the following section does not even consider through traffic.)

Stakeholder issues often create political hurdles for implementing RP/PP schemes, and conflicts of interest affect the choice of schemes. We summarize the primary stakeholders and their respective roles below:

(1) Government: Federal and local governments might be less interested in PP schemes since these do not generate much extra revenue compared to RP. Similarly, public transportation companies prefer RP because revenue generated by RP is used to improve an affected area’s public transit system.

(2) Residents: People living in cordoned areas oppose PP because they need to pay additional fees for parking. Even though they can use on-street parking spaces reserved for residents, especially after business hours, they must pay additional fees for parking every time they make a trip outside their residential area: even residents in the zone are not exempt from PP. Therefore, residents of cordoned areas prefer RP because they are exempt while still benefiting from reduced congestion.

(3) Business: A scheme’s impact on business varies with the type of business and sometimes even the industry segment. For example, the London Chamber of Commerce reported in its retail survey published in 2005 that the RP scheme in London was negatively affecting retail business.\(^3\) According to the report, 79% of Central London
retailers had experienced a fall in receipts and over half (56%) had seen a drop in number of customers. Forty-two percent of respondents indicated they felt the scheme was all or mostly to blame. London First, however, whose members account for 17% of all employees in London and contribute 22% to the city’s gross domestic product, has viewed the scheme positively. According to London First’s survey in London in 2003, 68% believed the scheme was working.

(4) Shops: Retail shops often benefit from congested roads and fully occupied parking lots, so their owners might not view favorably the elimination of street parking nearby.

Last, we consider equity issues. All economic measures are discriminatory policy because they try to exclude less productive people from using limited resources in order to maximize the “social surplus”. The major distributional equity issues follow:

(1) Poor and Rich: Unavoidably, RP (and, to a lesser extent, PP) deprive the poor of opportunities to drive cars, in order to increase the efficiency of utilization of limited road resources. However, this equity problem can be alleviated significantly by improving any public transportation systems currently provided. Revenue from RP (and PP) can be used not only to improve public transportation but also to install new affordable public transportation services.

(2) Suburbanites commuting to a cordon area and urban commuters living inside it: Most current RP (specifically, cordon area pricing) effectively distributes suburban commuters’ money to urban commuters when the latter are exempt from paying CC, and the inequities increase as a cordon area expands. Urban commuters therefore benefit from uncrowded roads after RP implementation without an appropriate CC burden. This inequity is difficult to resolve by means of RP alone because urban commuters have the power to reject a CP scheme if they are not exempted from paying a CC. However, this equity issue can be corrected by using PP to collect CC rather than
by using RP.

3.4 Key Factors for Successful Implementation

Both successful and unsuccessful implementations of RP/PP indicate the importance of quality public transportation systems as well as parking policies.

3.4.1 Enhancement of Public Transportation (PT)

Building up a quality PT system before RP/PP implementation is important because RP/PP shifts drivers to PT commutation. If the current PT is poor, people are likely to disapprove RP/PP. The quality of a public transport system includes its vehicles’ speed, punctuality, accessibility, network coverage, cleanliness and safety. For example, before implementing RP, London introduced about 300 additional buses,38 set new bus routes, increased the frequencies of bus operation. London also has enforced traffic rules strictly with police cooperation. London currently has 130 km of priority bus lanes, and bus service 24 h/day. Tokyo, too, is famous for its high-quality PT system. To compensate for its less than punctual bus system, a GPS bus-locator system has become common in Japan so users can check buses’ current location by Internet or cell phone.39 Trains in Japan are reliable and their network is extensive. Hence, Japanese commuters can often correctly estimate within minutes the time they will reach a destination—even if their itinerary includes ten transfers.40 In contrast, Edinburgh’s citizens were generally dissatisfied with their city’s PT, and as a result, roundly rejected the prospect of RP when that was raised. One important difference between London/Tokyo and U.S. cities should be noted: U.S. cities are less densely populated; therefore, providing extensive PT in the U.S. is more costly. A park-and-ride system can therefore be especially important in the U.S.

3.4.2 Parking Design Improvement

(1) Increase on-street parking prices

Inexpensive street parking creates congestion or adds to it not only by attracting more people to use cars but also by adding traffic to congested roads as cars queue up in
search of available street parking spaces—often so average wage earners can save money. Queueing theory suggests that just a few percentage points’ increase in traffic on almost fully congested roads significantly delays traffic. Real-world estimates are that between 8% and 74% of traffic may be cruising in search of available street parking in major US cities, with an average time required to find a vacant spot ranging from 3.5 to 14 minutes. These numbers can block nearly congested roads.

To make matters worse, one can also observe counteracting measures such as the following parking policy regarding New York City’s often packed Theater District in Manhattan; this particular ad can be spotted on a prominent banner near the top of the New York City Department of Transportation website:

**Driving to the Theater District?**

**Use On-Street Parking – Only $2.00 per hour**

**Evenings & Saturdays at Muni-Meters throughout the Theater District**

Thus, on-street parking spaces are only $2 per hour on weekdays 6 p.m.–12 a.m. and on Saturdays 8 a.m.–12 p.m., and free on Sundays. When one of this paper’s authors visited the area recently, on-street parking spaces were full even before traffic had become congested; some cars were double-parked in front of off-street parking lots as their drivers waited for a space to open up close by. People able to find on-street parking were either extremely lucky or patient enough (and possessing sufficient spare minutes) to spend a long time cruising or double-parking. The on-street parking capacity was obviously insufficient; therefore, the extremely inexpensive parking policy—“Only $2 per hour”—exacerbates congestion in the Theater District every evening and also on Saturdays—not to mention Sundays, when parking is free. When congestion is expected, street parking should be eliminated or its price level increased to that of nearby off-street parking.
(2) Eliminate parking subsidies

Subsidizing employees’ commuting expenses with free or discounted parking in lots is popular with employers but counteracts PP’s effectiveness. Census data for the year 2000 show that more than half (53%) of total commuters (about 230,000 people) driving into congested Manhattan each workday come from New York’s five boroughs. The data also show that 35% of government workers in Manhattan drive to work mainly because they have free parking. This problem could be solved by employers giving employees the cash equivalent of parking fees to spend on using an alternate mode of transportation. In California, for example, a law was passed in 1992 (although it has not been enforced) requiring all employers to make such cash-out options available to employees (Downs, 2004).
4. Queueing Model for Parking Pricing

In the following, we develop a model that depicts patrolling drivers seeking on-street metered or free parking. The model is motivated by recent data from Park Slope, Brooklyn and by extensive earlier analyses by Donald C. Shoup.

We assume that all parking spaces are occupied almost all of the time that would-be parkers are seeking parking spaces. Drivers seeking parking spaces are assumed to be driving around through the streets seeking the first available spot. As soon as one opens up, meaning a parked car is driven away, the next patrolling car virtually immediately occupies that spot. The platoon of patrolling cars is a moving queue serviced in random order. Not all would-be parkers are served in this queue, as the arrival rate of would-be parkers exceeds the departure rate of parked cars. So, we allow drivers in the patrolling queue to become discouraged, leave the queue and presumably settle for more expensive off-street parking (for instance, in a parking garage or in a parking lot).

For modeling purposes we assume an infinitely large homogeneous city with \( S \) parking spaces per square mile. We assume that the statistics of parking space availability and desirability are uniform over the city. We assume that the time any given parker occupies a parking space is a random variable \( W \) with probability density function \( f_W(x) \) and mean \( E[W]=1/\mu \). Prospective or would-be parkers appear in a Poisson manner at rate \( \lambda A \)/hour, where \( A \) is defined to be the size of the area being considered (in sq. mi.). Prospective parkers will patrol looking for the first available parking space. Any unsuccessful would-be parker can become discouraged. We model this process by assuming that any would-be parker will leave the queue of patrolling would-be parkers at an individual Poisson rate of \( \gamma \)/hr.

There are two “large numbers” features in this system that allow us to model the queue as a Markovian system. First, regardless of the details of the probability density function (pdf)
the aggregate process of parked cars leaving parking spaces is accurately modeled as a Poisson process with rate $A\bar{S}\mu/\text{hr}$. This is because the departure process from any given parking space is seen as a renewal process with inter-renewal pdf $f_W(x)$. As is well known, the merger or pooling of a large number of (sufficiently well-behaved) renewal processes converges to a Poisson process (Cox and Smith, 1954). We assume that the number of parking spaces we are considering is sufficiently large so that this approximation is very accurate. Second, the time until reneging of any would-be parker could be any well-behaved random variable having mean $1/\gamma$, not necessarily a negative exponential random variable. But, if the moving queue of patrolling would-be parkers is sufficiently large, we again have the pooling of many renewal processes --- each having the same probability density function of time until “renewal” and each starting at a random time. Such pooling will result in the aggregate process of $N$ would-be parkers leaving the queue becoming a Poisson process with rate $N\gamma$, where $N$ is typically large enough so that the Poisson assumption is valid.

We require one additional assumption in order to model this process efficiently. We assume that when there are zero cars patrolling in the modeled area, no parked cars leave their spaces. We know that this assumption is incorrect, but we are focusing on large queues of patrolling cars in which case the likelihood of zero patrolling cars is very small. If this assumption in an application setting is not valid, one can eliminate it by creating a larger Markovian model that includes the possibility of several or even many empty parking spaces.

In our work we will focus on a square area of the city having unit area (i.e., one square mile or one square kilometer). We will assume that this region is large enough for our saturation congestion theory to be valid. One might argue that in any actual city no would-be parker feels constrained to patrol within any arbitrary boundaries. This is true. But for every would-be parker who starts within our modeled square and then ventures out of it looking for an available parking space, there is statistically another equivalent would-be parker
who started in some near-by zone who ventures into our zone. Statistically, for everyone who leaves, there is someone who enters. We can take care of this by placing “reflecting barriers” around our zone, so that when anyone in the real system leaves, we simply reflect him or her back into the zone, creating a statistical equivalence to the real non-cordoned system.

We now can draw the state-rate-transition diagram for this queue, assuming one square mile of operation, as shown in Figure 1.

**Figure 1. State-Rate-Transition Diagram for Queueing System**

\[ \lambda \quad \lambda \quad \lambda \quad \lambda \quad \lambda \quad \lambda \quad \lambda \quad \lambda \quad \lambda \]

\[ S_\mu + \gamma \quad S_\mu + 2\gamma \quad S_\mu + 3\gamma \quad S_\mu + 4\gamma \quad S_\mu + n\gamma \quad S_\mu + (n+1)\gamma \quad S_\mu + (n+2)\gamma \]

By the usual process of “telescoping” balance of flow equations, we can express each steady state probability \( P_n \) in terms of \( P_0 \) and a product of upward transition rates (\( \lambda \)'s) divided by the product of downward transition rates between state \( n \) and state 0. The result is

\[
P_n = \frac{\lambda^n}{\prod_{i=1}^{n}(S_\mu + i\gamma)} P_0 \quad (1)
\]

Now, invoking the requirement that the steady state probabilities sum to one, we obtain

\[
(1 + \sum_{n=1}^{\infty} \frac{\lambda^n}{\prod_{i=1}^{n}(S_\mu + i\gamma)}) P_0 = 1,
\]

or,
\[ P_0 = \frac{1}{1 + \sum_{n=1}^{\infty} \frac{\lambda^n}{\prod_{i=1}^{n} (S\mu + i\gamma)}}. \]

Hence,

\[ P_n = \frac{\lambda^n}{\prod_{i=1}^{n} (S\mu + i\gamma)}, \quad n = 1, 2, 3, \ldots \quad (2) \]

For steady state to exist we require \( P_0 > 0 \), which always occurs. But we want \( P_0 \) to be very small for our approximations to be valid.

From the solutions obtained above, we can find all of the quantities of Little’s Law, \( L, L_q, W \) and \( W_q \). The basic Little’s Law relationship is, of course, \( L = \lambda W \). Here since “the system” is the queue only and service implies finding an empty parking space, we have the equivalences, \( L = L_q \) and \( W = W_q \). \( L \) is the time-average number of cars seeking parking spaces, or equivalently, the mean size of the patrolling queue of would-be parkers. \( W \) is the mean time that a patrolling car remains on patrol, until leaving either by finding a parking space or by frustration and reneging from the queue.

There are other performance measures of interest. The mean number of parking spaces becoming available per hour is \((1 - P_0)S\mu = S\mu \) since \( P_0 \ll 1 \). The mean number of renegers per hour is \( \lambda - (1 - P_0)S\mu = \lambda - S\mu \), assuming \( \lambda > S\mu \) (which is required for our approximations to be valid). For a random patrolling would-be parker, the probability of successfully getting a parking space is \( (1 - P_0)S\mu / \lambda = S\mu / \lambda \). This agrees with intuition. If say 100 parking spaces become available per hour and 250 would-be parkers arrive each hour, then 40% will succeed in finding a parking space and 60% will leave in frustration.
In the following we will assume that $0 < P_0 \approx 0$. This means that the queue of patrolling cars is, for all practical purposes, never empty. Under these conditions, we argue that the mean number of patrolling cars is

$$L = L_q = \frac{\lambda - S\mu}{\gamma}$$  \hspace{1cm} (3)

This is a fundamental result for our saturated on-street parking system. We argue its validity by changing the queue discipline from SIRO (Service In Random Order) to LCFS (Last Come, First Served). It is well known that $L$ and $L_q$ are invariant under the set of queue disciplines whose preferential orderings do not include customer-specific service times. The LCFS discipline is one such discipline. By LCFS here we mean the following: The next available parking space would be given instantaneously to that patrolling car that has been patrolling for the least amount of time. Usually this car would be the last to have arrived in queue. But it might be the case that the most recent car has already left the queue by reneging, in which case the next “youngest” patrolling car would be selected. The rate of successful parkings per hour is $S\mu$, and thus the fraction of would-be parkers who receive parking spaces virtually instantaneously upon arrival is $S\mu/\lambda$. The cars that do not get nearly instantaneous parking stay patrolling for an amount of time that is exponentially distributed with mean $1/\gamma$. For this revised queueing system $W_q$, the mean time patrolling can be written,

$$W_q \approx (0)(S\mu/\lambda) + (1/\gamma)(1 - S\mu/\lambda) = \frac{\lambda - S\mu}{\lambda \gamma}$$

Since $L_q = \lambda W_q$, we can write
\[ L_q = \frac{\lambda - S\mu}{\gamma}, \]

as was to be shown.

In the above argument we use “approximately equal to” signs instead of “equals signs.” This is due to the fact that there is a small but positive delay between a car’s arrival in the queue of patrolling cars and its selection as a recipient of a parking space. The mean delay between the arrival of a newly patrolling car and the emergence of a newly available parking space is \( 1/S\mu \), assumed to be very small in contrast to \( 1/\gamma \).

In the following two subsections we model explicitly two alternative ways of implementing the LCFS queue discipline, as discussed above. These analyses are to show the operational feasibility of the revised but highly fictional LCFS queue discipline. The “real system” at all times is still assumed to follow the SIRO queue discipline.

### 4.1 Random Walk

Assuming the postulated LCFS queue discipline, one can model the arrival of a newly patrolling car as an entry into “state 1” an infinite random walk on the non-negative integers, where state 0 implies that the car transitions to a trap state -- signifying successful assignment to a parking space. Transitioning to any higher state \( j+1, j \geq 1 \), indicates that the position in queue has been changed upward from \( j \) to \( j+1 \). Due to the LCFS discipline, higher states imply less likelihood of eventually receiving a parking space. If we define

\[ \beta_0 = P\{\text{car enters the trap state}\} = \\
P\{\text{car transitions down one state in the random walk}\} = \\
P\{\text{car obtains a parking space}\}, \]

then we can write
\[ \beta_0 = P\{\text{first transition is to trap state}\} + (1 - P\{\text{first transition is to trap state}\})\beta_0^2 \]

The reason for the term \( \beta_0^2 \) is the fact that if the car has transitioned into state 2, then to be awarded a parking space it must first transition down to state 1 and then eventually to state 0. Each transition down one state occurs with probability \( \beta_0 \), and the transition processes in each case are independent. The probability that the first transition is to the trap state is equal to the probability that a parking spot becomes available before the next arrival, and that is equal to \( S\mu/(S\mu + \lambda) \). Thus we can write,

\[ \beta_0 = \frac{S\mu}{S\mu + \lambda} + \frac{\lambda}{S\mu + \lambda} \beta_0^2 \]

The solution to this quadratic equation is \( \beta_0 = S\mu/\lambda \), and that agrees with our intuition and previous results.

There is a subtlety in the derivation, as the argument appears to ignore reneging. Since reneging can occur, the “cars” in the argument are in fact ordered slots: youngest slot in queue, 2nd youngest slot in queue, etc. The car occupant of any slot may change due to reneging. Once that is seen, the results are seen to be valid, even in the presence of reneging.

4.2 Queueing Newly Available Parking Spaces

If one does not wish to consider the LCFS policy analyzed above, perhaps due to unrealistic demands on tracking newly arriving cars, one can accomplish the same objective by using a queue discipline that we will call NCNS, Next Come, Next Served. In this scheme each newly available parking slot enters a queue of other newly available parking slots, and this queue is depleted by newly arriving cars seeking parking slots. Any driver in a car lucky enough to arrive when this queue of available parking slots is nonempty is immediately
given a slot. All others are denied slots forever, and they join the other patrollers who eventually renege after patrolling a random time having mean $1/\gamma$. This process can be modeled as an M/M/1 queue, with state $i$ indicating $i$ available parking slots ($i = 0, 1, 2, \ldots$), and with upward transition rates $S \mu$ and downward transition rates $\lambda$. Since $\lambda > S \mu$, we know that the queue is stable and possesses a steady state solution. Using well-known results from the M/M/1 queue, we immediately have,

$$P\{\text{an empty parking space is available at a random time}\} = 1 - P_0 = S \mu / \lambda < 1.$$ 

Since Poisson Arrival See Time Averages (PASTA), we have

$$P\{\text{a random arrival obtains a parking space}\} = 1 - P_0 = S \mu / \lambda < 1,$$

as expected.

In steady state, the mean number of free parking spaces is,

$$N_p = \sum_{n=1}^{\infty} nP_n = P_0 \sum_{n=1}^{\infty} n(S \mu / \lambda)^n = \frac{\lambda - S \mu}{\lambda} \sum_{n=1}^{\infty} n(S \mu / \lambda)^n = \frac{S \mu}{\lambda - S \mu}.$$

For example, if $\lambda = 2S \mu$, then $N_p = 1$ free parking space. One free parking space would remain free for an amount of time equal to the time of the next driver seeking a parking space, having mean $1/\lambda$. Usually this time is quite small in contrast other times in the system. More generally, in this instance Little’s Law states that $N_p = S \mu W_p$, so we have the mean time that a newly available parking space remains available is

$$W_p = \frac{1}{\lambda - S \mu}.$$

29
As an example, if \( \lambda = 100 \) cars per hour and \( S \mu = 40 \) cars per hour, then \( W_p = (1/60) \) hour = 1 minute. Again, this time is small in contrast to other times in the system, and all of our results are correct within acceptable “engineering approximations.”

In conclusion, we can feasibly implement a car-to-parking-space queue discipline that supports Eq.(3), using either LCFS or NCNS. But we remember that the actual or “real” discipline is still assumed to be SIRO.

4.3 The Distribution of Patrolling Cars

Using the above logic, we see that the entire system, conceptually augmented with either LCFS or NCNS queue discipline; can be viewed as a Poisson arrival queue with infinite number of servers, i.e., an \( M/G/\infty \) queue. “Service” occurs for any car the instant the car obtains a parking space or reneges from patrolling. The distribution of numbers of patrolling cars in the system is not affected by our augmented queueing discipline. Mean service time \( M \) can be written,

\[
M = (0) \frac{S \mu}{\lambda} + (\frac{\lambda - S \mu}{\lambda}) \frac{1}{\gamma} = \frac{\lambda - S \mu}{\lambda \gamma}.
\]

The Poisson process arrival rate is \( \lambda \). For the \( M/G/\infty \) queue having arrival rate \( \lambda \) and mean service time \( M \), the steady state probability distribution of the number \( N \) of customers in the system is well-known to be Poisson with mean \( \lambda M \), i.e.,

\[
P(N = n) = \frac{(\lambda M)^n}{n!} e^{-\lambda M}, \quad n = 0, 1, 2, ...
\]

In this case, we can write the probability that there are \( N \) cars patrolling for parking spaces is equal to
Here again we see that the mean number of patrolling cars is equal to \( \frac{\lambda - Su}{\gamma} \), the result of Eq. (3). But now we know that the entire distribution – assuming our saturation conditions – is Poisson. Finally, as saturation grows worse, that is as \( \lambda \) increases towards ever-greater congestion, the Poisson distribution becomes a Gaussian or Normal distribution.

The next step to take with this model is to place hourly prices on on-street parking and off-street parking. Then one makes certain model parameters dependent on these prices, especially the price difference between on-street and off-street parking. These ideas build on the suggestions of Shoup (2005). As the price difference between on-street and off-street parking becomes less, one should have the rate \( \gamma \) at which one leaves the queue of patrolling cars increase. That is, the desire to find an on-street parking space and the patience it requires in the patrolling queue will decrease as the price advantage of on-street parking decreases. Eventually as one gets closer to price parity, our approximate assumption of an endless queue of patrolling cars becomes invalid and we must modify the model accordingly. Shoup’s stated objective is to raise on-street prices so that one has roughly 15% of the on-street parking spaces available in steady state. For the model, this would require extending the state-rate-transition diagram down significantly into unsaturated states but still allowing the artifice of stopping at some left-most nonzero state that has very small steady state probability. We do not see the need to model the system all the way down to zero parking spaces being occupied.

4.4 Congestion Pricing and Queueing Theory

Congestion pricing theory is based on the following observation: The congestion cost caused by the entrance of a driver to a queueing system consists of the cost of delay to this driver (internal cost) plus the cost of additional delay to all other users caused by this driver.
(external cost). For example, if the driver enters into a congested road and experiences 5 minutes delay, the internal cost to him is the cost of 5 minutes. However, when the road is very congested, the entrance of this driver may delay 1 minute to 7 other drivers. Then the external cost generated by him is the cost of 7 minutes to the other drivers. In order to achieve the most efficient use of the road facility, this external cost should be burdened by each driver. In economic terms, the external cost should be internalized. This was first pointed out by Vickrey⁴⁵ and by Carlin and Park.⁴⁶ They claimed that “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 on the facility itself. A congestion toll not only contributes to maximizing social economic welfare, but is also necessary to reach such a result.” In 1959, William Vickrey, Columbia University economist and 1996 Nobel Laureate, proposed an electronic RP system in detail to the Joint Committee on Washington Metropolitan Problems.⁴⁷ At the time, he also pointed out the importance of a variable pricing system for on-street parking spaces in order to ensure some vacancy to accommodate the demand and avoid unnecessary traffic congestion caused by on-street parking shortages.

We follow economic principles to obtain the “optimal” congestion pricing. Consider a queueing facility with a single type of user in steady state and let

\[ \lambda = \text{demand rate per unit of time by road users.} \]
\[ c = \text{cost of delay per unit time per user.} \]
\[ C = \text{total cost of delay per unit time incurred by all users in the system.} \]
\[ L_q = \text{expected number of users in queue.} \]
\[ W_q = \text{expected delay time in queue for a random user.} \]

We can also assume that \( L = L_q \) and \( W = W_q \), as in our parking model.

Then the time-average total delay cost per unit time can be written,
\[ C = cL_q = c\lambda W_q, \]

where Little’s Law is used. The marginal delay cost \((MC)\) imposed by an additional road user can be obtained as,

\[ MC = \frac{dC}{d\lambda} = cW_q + c\lambda \frac{dW_q}{d\lambda}. \]

The first term on the right is the internal cost experienced by the additional road user, and the second is the external cost due to the increase in the expected delay, \(\frac{dW_q}{d\lambda}\), resulting from the increased traffic created by this user. Hence, we can write two components of the marginal delay cost \(MC\) as follows:

1. **Marginal internal cost**: \(MC_i = cW_q\)
2. **Marginal external cost**: \(MC_e = c\lambda \frac{dW_q}{d\lambda}\).

Vickrey suggested that the marginal external cost \(MC_e\) should be imposed on each road user in order to realize socially “optimal” utilization of road resources. In the most common cordoned-area RP scheme today, however, the fee for residents in cordoned areas is significantly discounted; also, the CC level is set to a constant fee per day regardless of the frequency of trips a driver makes. Therefore, imposing appropriate charges on each road user is difficult for RP, and consequently road resources become overused, which is partly the reason why implementing RP over a large area is difficult. PP, on the other hand, does not present such issues because a parking fee is charged all road users impartially (except for privileges granted to physically challenged people), per trip, regardless of whether or not they are residents of the charging area.
4.5 The Parking Pricing Model

From our previous work, for the parking process in saturation, the total delay cost per unit time and associated marginal delay cost are

\[ C = cL_q = (\lambda - S\mu) \frac{c}{\gamma} \]  

(4)

and

\[ MC = \frac{\partial C}{\partial \lambda} = \frac{c}{\gamma}. \]  

(5)

We can also obtain the marginal internal cost and marginal external cost,

\[ MC_i = cW_q = \left[1 - \left(\frac{S\mu}{\lambda}\right)\right] \frac{c}{\gamma} \]  

(6)

\[ MC_e = c\frac{\partial W_q}{\partial \lambda} = \frac{S\mu}{\lambda}, \frac{c}{\gamma}. \]  

(7)

The ratio of \(MC_e/MC_i\) is

\[ r = \frac{MC_e}{MC_i} = \frac{\frac{S\mu}{\lambda}, \frac{c}{\gamma}}{\left(1 - \frac{S\mu}{\lambda}\right), \frac{c}{\gamma}} = \frac{S\mu}{\lambda}, 1 - \frac{S\mu}{\lambda}. \]  

(8)

Here, we observe an interesting result. For a given \(c\), the marginal delay cost to society is dependent only on \(\gamma\) and does not depend on \(S\mu\) or \(\lambda\). In a sense, in saturation each additional would-be parker “brings with him” an average of \(1/\gamma\) of delay, to be incurred by somebody or some combination of people. However, the marginal internal cost, the marginal external cost and their ratio \(r\) are dependent only on \(\frac{S\mu}{\lambda}\), which is the success probability for would-be parkers to find on-street parking spaces. Eq. (7) shows that the
marginal external cost $MC_e$ is proportional to the parking success probability. $MC_e$ becomes larger when more would-be parkers expect they will find parking spaces. $MC_e$ decreases if we reduce the number of on-street parking spaces $S$, or increase the arrival rate $\lambda$, or increase the reneging rate $\gamma$. If half of would-be parkers will find a parking space, then marginal internal and external costs are equal. If 90% of all would-be parkers are denied parking, then the external cost $MC_e$ associated with one new would-be parker is only $0.1c/\gamma$, whereas the internal cost $MC_i$ is $0.9c/\gamma$. This is due to the fact that 90% of the time our new would-be parker arrives, he will be denied parking and will have to incur the mean patrolling time (cost) $1/\gamma$ almost all by himself; he denies others only 10% of the time.

### 4.6 Trading Off Cost Savings and Convenience

Economists like to speak of “optimal” charges for those imposing external costs, this problem being no exception. But it is difficult to operationalize this concept. What precisely is meant by optimal? Optimal is an absolute word requiring a precise and unambiguous objective function and set of constraints. We do not have those conditions in the context of on-street vs. off-street parking. And how do the fees collected get distributed to aggrieved parties? As operations researchers and not as economists, we tend to think of drivers as decision makers who weigh their options and act accordingly.

Without significant empirical research, it is not possible to know precisely how would-be parkers would behave in our “patrolling queue” situation. But we can make some plausible first-order assumptions, presented in a transparent manner for review and critique. First, it seems clear that some drivers would value their time more than others, and those would tend to leave the queue of patrolling drivers more quickly than others. Second, a driver’s willingness to spend time in the patrolling queue would rise or fall with the price differential between on-street and off-street parking, with higher price differentials meaning more willingness to spend time looking for less expensive on-street parking. Third, any unsuccessful patrolling driver will eventually become discouraged, “cut his losses,” and leave the queue for more expensive off-street parking.
We can develop a simple model reflecting these assumptions. Suppose there are \( D \) categories of drivers, where category \( d, 1 \leq d \leq D \), has a self-assessed value for time of \( W_d \) dollars per hour. We assume the categories are rank-ordered such that \( W_1 \geq W_2 \geq W_3 \geq \ldots \geq W_D \). Let \( p_d \) be the fraction of all would-be parkers belonging to category \( d, 1 \leq d \leq D \). Clearly, \( \sum_{d=1}^{D} p_d = 1 \). Let \( \Delta \) be the hourly parking price differential (in dollars) between off-street and on-street parking, with the on-street parking being less expensive. We now need a decision criterion for a patrolling driver to leave the queue and accept the more expensive off-street parking. One plausible criterion is this: When the value of the time already invested in patrolling for a less expensive on-street parking space equals the price differential between off-street and on-street parking, then the expected values of the respective options – when including sunk costs – become equal. But the variance of costs for continued patrolling is large, whereas the variance of cost associated with the off-street option is zero (a known, published parking fee). Thus, the decision rule is to leave the queue and switch to off-street parking when the sunk cost of time invested becomes equal to the parking price differential. This set of assumptions provides a basis for evaluating the resultant reneging parameter \( \gamma \) as a function of the price differential \( \Delta \). The mean time that a category \( d \) patrolling driver would remain patrolling is \( \frac{1}{\gamma_d} = \frac{\Delta}{W_d} \).

Including all \( D \) categories, weighed by their respective relative frequencies, the resulting relationship can be written,

\[
\frac{1}{\gamma} = \Delta \sum_{d=1}^{D} \frac{p_d}{W_d}. \tag{9}
\]

As a numerical example, consider one-hour parking with \( D = 3; p_d = 1/3 \) for \( d = 1, 2, 3 \); \( W_1 = 100, W_2 = 25, W_3 = 10 \). Then
If $\Delta = \text{US}$10/hr. then $(1/\gamma) = 0.5$ hr. = 30 minutes. If $\Delta = \text{US}$20/hr., then $(1/\gamma)$ is doubled to 60 minutes. One socially positive aspect of the driver behavior assumed in this model is that the successful on-street parkers are differentially more likely to be poorer people who value their time less than others. Those who value their time highly will tend to leave the queue more quickly and pay the higher off-street parking rates.

### 4.7 Extending the Model to Include Heterogeneous Drivers

In this section, we confirm the intuition that “poorer people are more likely to be successful on-street parkers than richer people”. Assume there are two types of drivers, or “would-be-parkers,” Type 1 and Type 2, whose corresponding arrival rates and reneging rates are $\lambda_i$ and $\gamma_i$ ($i=1, 2$), respectively. We construct a 2-dimensional state-rate-transition diagram for the Markovian queue created by two types of drivers. Assume that each state is represented by the ordered pair $n_1$ and $n_2$, which correspond to the respective numbers of Type 1 and Type 2 drivers in the system. The state-rate-transition diagram is shown in Figure 2.

*Figure 2. State-Rate-Transition Diagram for Queueing System with Two Types of Drivers*
As before, we continue to assume that \( 0 < P_{00} = 0 \), but now for this 2-dimensional system.

Again as before, we assume that the road is congested, with either type of driver able to fill all available parking spaces: \( \lambda_1 \geq S\mu \) and \( \lambda_2 \geq S\mu \).

We can write a set of balance-of-flow equations, where the balanced flows occur across complete horizontal cuts of the network of Figure 2,

\[
(\lambda_1 + \lambda_2)P_{00} = (S\mu + \gamma_1)P_{10} + (S\mu + \gamma_2)P_{01} = S\mu(P_{10} + P_{01}) + \gamma_1(P_{10} + 0P_{01}) + \gamma_2(0P_{01} + 0P_{10})
\]

\[
(\lambda_1 + \lambda_2)(P_{10} + P_{01}) = (S\mu + 2\gamma_1)P_{20} + (\frac{1}{2}S\mu + \gamma_1)P_{11} + (\frac{1}{2}S\mu + \gamma_2)P_{11} + (S\mu + 2\gamma_2)P_{02}
\]

\[
= S\mu(P_{20} + P_{11} + P_{02}) + \gamma_1(2P_{20} + P_{11} + 0P_{02}) + \gamma_2(2P_{02} + P_{11} + 0P_{20})
\]
\[(\lambda_1 + \lambda_2)(P_{20} + P_{11} + P_{02}) = (S\mu + 3\gamma_1)P_{30} + \left(\frac{2}{3}S\mu + 2\gamma_1\right)P_{21} + \left(\frac{1}{3}S\mu + \gamma_1\right)P_{12} + \left(\frac{1}{3}S\mu + \gamma_2\right)P_{21} + \left(\frac{2}{3}S\mu + 2\gamma_2\right)P_{12} + (S\mu + 3\gamma_2)P_{03} = S\mu(P_{30} + P_{21} + P_{12} + P_{03}) + \gamma_1(3P_{30} + 2P_{21} + P_{12} + 0P_{03}) + \gamma_2(3P_{03} + 2P_{12} + P_{21} + 0P_{30}) \]

\[\cdots\]

\[(\lambda_1 + \lambda_2)(P_{(n-1)0} + P_{(n-2)1} + \ldots + P_{0,(n-1)}) = S\mu(P_{n0} + P_{(n-1)1} + \ldots + P_{0n}) + \gamma_1(nP_{n0} + (n-1)P_{(n-1)1} + \ldots + 1P_{1(n-1)} + 0P_{0n}) + \gamma_2(nP_{0n} + (n-1)P_{1(n-1)} + \ldots + 1P_{(n-1)1} + 0P_{n0}) \]

\[\cdots\]

Adding up the countably infinite set of balance equations, we obtain

\[(\lambda_1 + \lambda_2)\left(\sum_{n,m=0}^{\infty} P_{nm}\right) = S\mu\left(\sum_{n,m=0}^{\infty} P_{nm} - P_{00}\right) + \gamma_1\left(\sum_{n,m=0}^{\infty} nP_{nm}\right) + \gamma_2\left(\sum_{n,m=0}^{\infty} mP_{nm}\right).\]

Using the assumption \(P_{00} = 0\), invoking the normalizing condition \(\sum_{n,m=0}^{\infty} P_{nm} = 1\), and using the definitions \(L_1 = \sum_{n,m=0}^{\infty} nP_{nm}\) and \(L_2 = \sum_{n,m=0}^{\infty} mP_{nm}\), we obtain

\[\lambda_1 + \lambda_2 = S\mu + \gamma_1 L_1 + \gamma_2 L_2.\]  \hspace{1cm} (10)

We need to derive one more equation to solve for \(L_1\) and \(L_2\). In order to do this, consider the mean number of Type 1 and Type 2 renegers per hour, which are

\[\sum_{n,m=0}^{\infty} n\gamma_1 P_{nm} = \gamma_1 \left[\sum_{n,m=0}^{\infty} nP_{nm}\right] = \gamma_1 L_1\]

and

\[\sum_{n,m=0}^{\infty} m\gamma_2 P_{nm} = \gamma_2 \left[\sum_{n,m=0}^{\infty} mP_{nm}\right] = \gamma_2 L_2,\]

respectively. Using these, the steady state mean number of parking spaces available and taken by Type 1 and Type 2 parkers per hour are \(\lambda_1 - \gamma_1 L_1\) and \(\lambda_2 - \gamma_2 L_2\), respectively. Note that the sum of the mean number of parking spaces available and taken by Type 1 and Type 2 parkers per hour is

\[\lambda_1 - \gamma_1 L_1 + \lambda_2 - \gamma_2 L_2 = S\mu,\]

using Eq. (10). Note also that both \(\lambda_1 - \gamma_1 L_1\) and \(\lambda_2 - \gamma_2 L_2\) are
positive because the mean number of renegers $\gamma_1 L_1$ and $\gamma_2 L_2$ must be less than the arrival rate $\lambda_1$ and $\lambda_2$, respectively, in steady state.

We now argue that the proportion of parking spaces taken hourly by Type 1 (Type 2) drivers is equal to the proportion of cruising drivers who are Type 1 (Type 2). For if not, then Type 1 (Type 2) drivers would be more or less skilled than Type 2 (Type 1) drivers at finding parking spaces. Due to the SIRO queue discipline that rewards that driver, Type 1 or Type 2, who just happens to be closest to the newly available parking space, each type of driver is by definition equally skilled. And clearly the proportion of parking spaces taken per hour by Type 1 (Type 2) drivers is equal to the fraction of parking spaces occupied by Type 1 (Type 2) drivers. For if not, then the parking time statistics of the two types of drivers would differ, and this is not allowed in our model.

Invoking these results, we can write

$$\frac{L_2}{L_1} = \frac{\lambda_2 - \gamma_2 L_2}{\lambda_1 - \gamma_1 L_1}, \quad \text{or, simplifying,} \quad \gamma_2 - \gamma_1 = \frac{\lambda_2}{L_2} - \frac{\lambda_1}{L_1}$$

(11)

Combining Eqs. (10) and (11), we have

$$\gamma_2 - \gamma_1 = \frac{\lambda_2}{L_2} - \frac{\lambda_2}{\lambda_1 + \lambda_2 - S\mu} \frac{\gamma_2 - \gamma_1}{L_2}$$

(12)

and

$$\gamma_2 - \gamma_1 = \frac{\lambda_2}{\lambda_1 + \lambda_2 - S\mu} \frac{\gamma_1}{L_1} - \frac{\lambda_1}{L_1}$$

(13)

Since both $\lambda_1 - \gamma_1 L_1$ and $\lambda_2 - \gamma_2 L_2$ are positive, the denominators in Eqs. (12) and (13) are all positive. Therefore, unique positive solutions for both $L_1$ and $L_2$ are guaranteed in the above equations. Analytical solutions can be obtained for both $L_1$ and $L_2$ using the quadratic
formula. The method extends to three or any number of different types of drivers.

For simple illustrative purposes, consider a numerical example. Assume there are two types of drivers: 100 poor people per hour arrive to the system and their per-person reneging rate is 1/hr., and 300 rich people per hour arrive to the same system and their per-person reneging rate is 3/hr. Both types of drivers are trying to find on-street parking spaces which capacity is \( S\mu = 50 \) hr. In this case, one could argue that poor people value their time at a rate of 1/3 that of rich people. By placing numbers in Eqs. (12) and (13), we obtain

\[
3 - 1 = \frac{300}{L_2} - \frac{100}{100 + 300 - 50} - \frac{3}{1} \quad \text{and} \quad 3 - 1 = \frac{300}{3L_1} - \frac{100}{L_1} - \frac{1}{3}
\]

Solving, we have \( L_1 = 77 \) and \( L_2 = 91 \). Hence, the ratio of poor and rich in parking spaces are

\[
\text{Poor} : \text{Rich} = L_1 : L_2 = 77 : 91 = 46\% : 54\%.
\]

The interpretation is as follows: Even though poor people’s arrival is 25% of the total arrivals, poor people occupy nearly half of the on-street parking spaces because of their lower reneging rate, their greater “patience” while cruising for an available parking space. Furthermore, the success rate of finding available parking spaces for poor and rich are

\[
\frac{\lambda_1 - \gamma_1 L_1}{\lambda_1} = \frac{100 - 1 \cdot 77}{100} = 23\% \quad \text{and} \quad \frac{\lambda_2 - \gamma_2 L_2}{\lambda_2} = \frac{300 - 3 \cdot 91}{300} = 9\%,
\]

respectively. Therefore, in terms of distributional equity, the provision of on-street parking spaces can be seen as “good” because poor people tend to utilize inexpensive parking more often than rich people. However, the result also suggests that poor people are more apt to patrol than rich people, thereby maintaining levels of street congestion that may be found unacceptable. The way to fix that problem is to raise the price of on-street parking, and that would increase the reneging rate of poor people since the price advantage of patrolling for on-
street parking diminishes.

5. Conclusion

We reviewed current various road pricing (RP) and parking pricing (PP) schemes for implementing congestion pricing (CP). We found PP is not only a cost effective alternative to RP which can be implemented in a small city but also has a capability of controlling local and time-varying traffic congestion using the price differentials between on-street parking and off-street parking. Following the review, we developed a new queueing model of the parking problem. We found (1) the queueing delay is inversely proportional to reneging rate and the distribution of number of patrolling drivers follows Poisson distribution, (2) the marginal delay cost imposed by an additional road user becomes constant as a result of reneging when on-street parking spaces are full, and (3) the congestion charge (CC) is calculated as the marginal external cost. We then extend the homogeneous model to heterogeneous model with two types of drivers. We found that the successful on-street parkers are differentially more likely to be poorer people who value their time less than others.

Notes
2. See Section 3.1.5, ‘Tokyo, Japan (PP)’ of this article for more information and sources regarding these examples.


6. This purely economic interpretation of “wasted time,” equating its cost to the product of time and hourly wage, is questionable because some commuters enjoy aspects of driving that can be appreciated even during periods of congestion, such as listening to radio. It is also questionable to conclude that eliminating such “wasted time” from congestion might lead to additional production by workers.

7. The CO₂ market price in the European Union is 21.0 euro/ton as of 7/4/2007 (http://www.pointcarbon.com/). This price may increase, considering the accelerating trend of global warming.

8. Singapore’s RP, with many gantries installed over the city center, is currently the system closest to full-scale RP. Hong Kong tested a large-scale, peak-hour pricing system for six months but rejected it because of privacy concerns. A GPS-based system might be a viable solution to address Hong Kong’s concerns. One such system has been tried in Gothenburg, Germany, but many technical problems remain before a GPS-based system can be fully implemented.


10. Transportation Alternatives (February 27, 2007) ‘No Vacancy: Park Slope’s Parking Problem And How to Fix It’.


http://www.eltis.org/study_sheet.phtml?study_id=140
14. Ibid.
17. Rémy Prud'homme and Pierre Kopp (November 17, 2006) ‘Urban Tolls: The lessons of the Stockholm Experiment may not be what you think’, Distinguished Speaker Series lecture for the Center for Transportation and Logistics, Massachusetts Institute of Technology, Cambridge, Massachusetts. According to Professors Prud’homme and Kopp, the referendum actually showed 53% in favor in the central municipality, but surveys conducted the same day in 14 other municipalities showed only 40% in favor and 60% against. If these latter polls had been included, the overall rate in favor would have been only 47%. Questions asked by the referendum have also been criticized as being misleading.
27. Note that 1985 was before Hong Kong’s reunification with China.
33. The researchers also found that many drivers use free bridges to avoid tolls; free bridges are extremely congested as a result. The congestion mechanism triggered by the price differential resembles the PP problem we are considering here.
35. If a highway is congested, its users should also be charged a CC, but this paper does not consider CC for highway users.
http://www.londonchamber.co.uk/docimages/260.pdf
39. Several examples of Japanese bus locator Internet sites are:
http://info.entetsu.co.jp/navi/pc/location.aspx,
40. The following websites are commonly consulted to check transfer information:

References


http://web.mit.edu/urban_or_book/www/


http://shoup.bol.ucla.edu/Cruising.pdf


