Buying Into the Bypass: Allowing Trucks to Pay to Use the Ramp Meter Bypasses

This paper analyzes the opening of a high occupancy vehicle (HOV) ramp meter bypass to trucks that pay a toll. Trucks are similar to HOVs as both have higher values of time than single-occupant cars. Thus a time saving for these vehicles benefits the transportation system. Using queuing analysis, tolls maximizing user benefit, toll authority profit, and system benefit were estimated. It is found that to maximize system welfare, trucks should be allowed free use of the bypass, but that raises some equity and operational issues. However, a toll that allows trucks to use the bypass improves the welfare over prohibiting them.

by Satyanarayana Muthuswamy and David M. Levinson

Ramp meters have been used to manage traffic entering the freeway in order to maximize flow through bottlenecks. However, meters are crude devices that cannot discriminate based on trip purpose or the value of the vehicle in the queue. The year 2000 saw unprecedented public opposition to the ramp metering system in the Twin Cities of Minneapolis and St. Paul. This resulted in the Minnesota state legislature mandating a shutdown of the meters as part of an experiment to test the effectiveness of the ramp metering system. Cambridge Systematics (2001) and Levinson et al. (2002) concluded from the eight-week shutdown that the effectiveness of ramp metering varied by location. In 2001, similar controversy arose over the effectiveness of the high occupancy vehicle (HOV) lanes. The controversies have resulted in consideration of creative ideas to improve the effectiveness of the highway system. One such idea is allowing high value vehicles, especially trucks, to buy into HOV capacity.

Congestion is a transportation externality caused by the difference in the delay driver's face and what they impose on others. The primary objective of congestion pricing is to internalize this cost by imposing a toll to optimize road usage (Levinson 2002). Allowing high value vehicles to buy their way out of congestion is a form of value pricing, which sorts vehicles by their value of time (VoT).

Every vehicle has a value of time determined by a set of factors that are specific to the driver, passengers and cargo of that vehicle. In general, trucks have a higher VoT than passenger cars. While both have drivers, the driver of a truck is paid a wage and the truck is carrying valuable time-sensitive goods; some goods have a very high value of time, especially freight like perishables, medicines, and computer chips.

Tolling and congestion pricing are closely related. Although tolling has been the primary form for funding private and some public roads, tolling serves as a means to implement congestion pricing. A direct con-

Journal of the Transportation Research Forum. Published in *Transportation Quarterly*,
Vol. 57, No. 1, Winter 2003 (81–92).

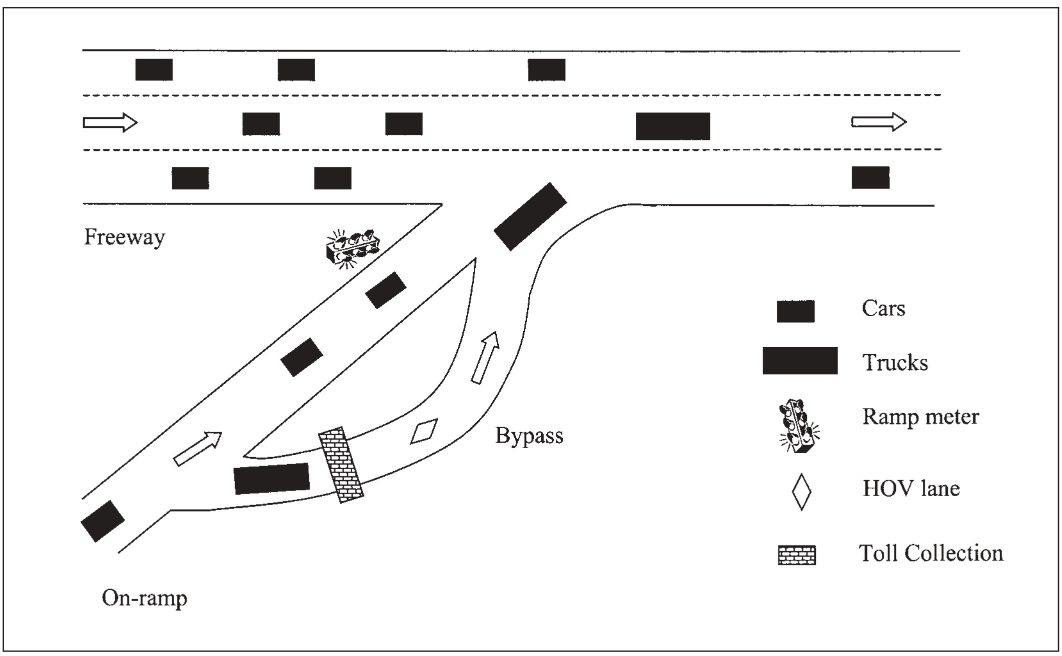
sequence of pricing of selected lanes is a transfer of the waiting time across users. The benefits to a system from pricing are the reduced delays for the high VoT users, but the reduced delays for the paying vehicles may be offset by the increased costs for other users.

One of the more recent policies to improve transportation system management is encouraging the use of HOVs. These vehicles, by virtue of their increased occupancy, have a higher value of time than the single occupant vehicles (SOV). There are incentives to increase their usage such as lanes open only to high occupancy vehicles. In the Twin Cities area, in the 430 metered on-ramps, HOVs are given bypasses at 73 on-ramps. However, Dahlgren (1998) argues that HOV lanes are in general underutilized and it is better to open them to the public as a tolled lane. In addition, Rogers (1985) also concludes that the presence of a ramp meter bypass did not result in a significant increase in the number of high occupancy vehicle users and that it resulted in a shift in the number of users from the ramps that did not have bypasses.

Currently, at ramp entrances, high occupancy vehicles are allowed to use a ramp meter bypass that leads directly onto the freeway without delay. The underlying concept is that, since HOVs carry more people, the system gains if the delay of such cars is reduced. Like high occupancy vehicles, trucks have a larger loss of utility (welfare) than SOVs for the same delay since they have a higher value of time Therefore, the trucks and the system should benefit if these vehicles were allowed to bypass the queue at the on-ramp.

As seen in Figure 1, the arriving trucks would be allowed to choose between the tolled bypass and the untolled on-ramp. If the truck decides to take the bypass, its delay is reduced but it imposes an additional delay to the vehicles in the queue or to the vehicles

Figure 1: Ramp Meter Bypass



Note: Current practice only permits HOVs to use the ramp meter bypasses on the freeway. The bypass actually serves as a shortcut in the queue. The truck that paid the toll and entered the freeway essentially jumps the queue from the back to the front.

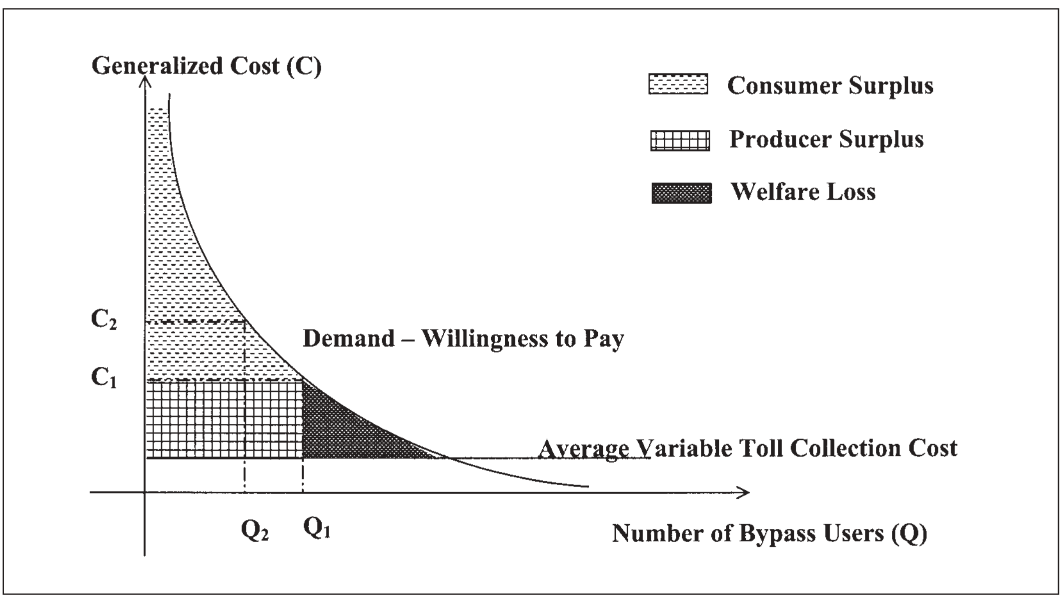
This paper argues that since HOV lanes are underutilized and trucks have a higher value of time than single occupant vehicles, that trucks be permitted to pay for access to these high occupancy vehicle lanes and ramp meter bypasses. This is a similar, but more specialized version of the high occupancy toll lanes that have been proposed and used on I-15 in San Diego (Fielding and Klein 1993; Poole and Orski 1999). This paper has been organized as follows. The first section of the paper develops the theory behind congestion pricing and a description of the optimal toll estimation based on the different scenarios. The following section describes the methodology to determine the toll. The queuing model used and the assumptions are discussed in the subsequent sections. The last two sections are the results and the conclusions.

CONGESTION PRICING

The direct consequence of tolling can be visualized as a shift in the traffic. We know that by tolling a route, the cost to users of taking that particular route increases, forcing some vehicles off the tolled route. Consider the situation shown in Figure 2. The demand for the ramp meter bypass decreases as generalized cost (toll plus travel time) rises. The supply curve is assumed constant (average variable collection cost per vehicle). The toll is generally set at any value above this collection cost. If the toll is set such that the generalized cost is at C_1, then corresponding demand is Q_1. If the toll is increased such that the generalized cost rises to C_2, then the number of users declines from Q_1 to Q_2. The amount of the demand reduction is a direct function of the toll increase (C_2- C_1) and the price elasticity of demand.

The presence of a slower untolled alternative to a toll road ensures that drivers have a choice and the toll is not forced on them. The toll is the key parameter that needs to

Figure 2: Tolling, Consumer and Producer Surplus



be estimated under a tolling scenario. This value is determined based on the overall objective of the congestion pricing scenario.

Two entities govern the overall objective. The first entity is users (SOVs, HOVs, toll paying trucks, and nontoll paying trucks), and the other entity is the toll authority. Therefore, there are three distinct possibilities. We could maximize the gains of either entity or both the entities. The gains are measured as the consumers' surplus and the producer's surplus.

Before we proceed to analyze the cases, we define the units that measure the two quantities—the consumers' surplus and the producer's surplus. Essentially the producer (the toll authority) has a small toll collection cost per vehicle, say, c_i, per vehicle 'i'. Then the *producer's surplus* (the profits) is equal to the *revenues* that are generated from the tolling scheme less the toll collection costs. The toll authority collects the toll (τ-i) from vehicle 'i' if it takes the tolled route. Now the *producer's surplus* can be measured as $\Sigma (\tau_i - c_i)$.

The consumers' surplus is the difference between the costs users incur and the costs that users are willing to pay. Since the actual amount that the users are willing to pay cannot be determined, we compare the base case (Scenario 1) where the use of the bypass is prohibited and the case where the bypass is permitted with payment of a toll (Scenario 2). The users do not have any choice in the base case but have two alternatives in the second case. If the choice-making process in the second case is modeled using a logit model, then the consumers' surplus can be measured using the logsum (Ben-Akiva and Lerman 1985).

$$(1) \quad Logsum = \frac{1}{\mu} \ln \sum_i e^{\mu U_i^2} - \frac{1}{\mu} \ln \sum_i e^{\mu U_i^1}$$

Where,

μ = scale parameter,

U_i^1 = the utility of alternative 'i' in Scenario 1,

U_i^2 = the utility of alternative 'i' in Scenario 2.

The logsum as defined in (1) is the difference among expected maximum utilities in the two scenarios. If utility is defined to be inversely proportional to travel time, then the logsum measures the scaled difference in travel times between the two alternatives. In other words, the consumers' surplus is represented as the scaled gain from the travel time reduction based on the vehicle type and value of time, less the toll paid. The scale parameter in the logsum is taken to be 1, so that the choice probability function is the logit model.

Based on the above discussion, we can define the objective function (F) for the three cases of Toll Authority Profit Maximization, User Benefit Maximization, and the System Benefit Maximization.

Toll Authority Profit Maximization

The first objective is to maximize the toll authority's profits. This is described in Figure 2. We are trying to find the toll such that the lower hatched region is maximized. The aim is to estimate the optimal toll (τ*) as a tradeoff between the number of users and the value of τ*. The fixed costs of the toll authority are ignored because the derivative of the fixed cost with respect to the toll collected is zero, so it does not influence the value of the optimal toll (τ*).

Maximize $F = \Sigma (\tau_i - c_i)$

Consumer Benefit Maximization

The toll is set such that the user's gain (scaled travel times) is maximized. Graphically, we are trying to maximize the upper shaded area (in Figure 2). The objective function is as follows.

Maximize F = Logsum
Logsum as defined in (1).

System Benefit Maximization

This is the case that involves setting (t) such that the welfare of both the users and the toll authority are maximized. There is a tradeoff between the users and the toll. As the toll increases, the number of users of the tolled bypass declines, and the number in the congested ramp queue increases.

Maximize F = $\Sigma (\tau_i - c_i)$ + Logsum

CHOICE MODEL

The queuing and choice-making process is modeled using simulation. Individual arrivals and choices are simulated, and based on the individual choices the objective function is calculated and a single toll for the peak period is estimated.

The most important part of the simulation is the choice-making process at the toll booth. In this paper, *random utility theory* is used to model the choice-making process (Ben Akiva and Lerman 1985). The utility is defined to be inversely proportional to the travel costs (toll and travel times). The toll (τ) is the cost charged to use the bypass and the delay (d) is scaled in money units using the value of time. Therefore, the utility of the tolled lane (U_τ) can be defined as shown in (2) and the utility of the untolled lane (U_τ) can be defined as shown in (3).

(2) $U \; = - \tau + \varepsilon$

(3) $U_\tau = - \alpha * d + \varepsilon$

Where,

α = VoT,

ε = error term/random variations in the utility.

Based on the utility associated with the two alternatives 'U_τ' and 'U_τ', the probability of choosing the tolled alternative ($P(\tau)$)[1] is defined as shown in (4). This is the logit model representation of the choice-making process. Using this probability function, the choice-making process can be modeled stochastically in the queue simulation.

(4) $P(\tau) = 1/(1 + e^{(U_t - U_t)})$
$P(t) = 1 - P(\tau)$

QUEUE SIMULATION

The basic pattern that is followed at the on-ramps is the First Come, First Serve (FCFS) queuing system. All vehicles using a ramp form queues and leave in the order of their arrival. Typically, the arrival rate is a random variable and it increases near the peak period and reduces as the peak period ends. In this paper, we assume that the arrivals are random about a mean value and the departure rate is the ramp meter rate.

Once a truck adopts the tolled route, it bypasses the queue and enters the freeway directly. The freeway conditions determine the ramp-metering rate. The consequence of this is an increased service time for the ramp meter, as there is an additional user on the freeway, and the ramp meter red time increases. Thus by tracking each truck and its choice, the departure rate and the delay for all other vehicles are determined.

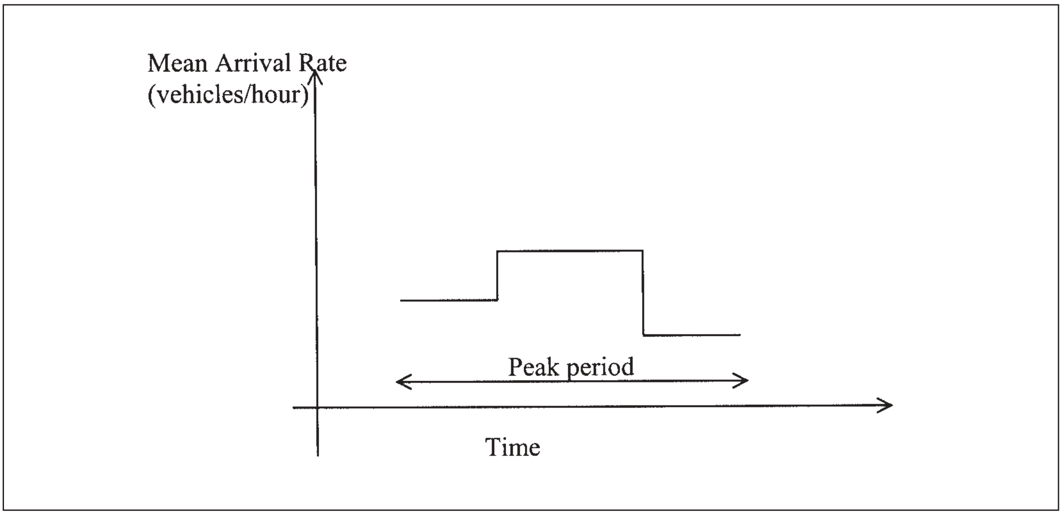
For the hypothetical on-ramp as described (Figure 1) a *stochastic queue* is used to represent the flow. The flow parameters on the freeway are assumed to be constant over the peak period, so the ramp-metering rate is held constant over the peak period.

Arrival Rate

(5) $h = \frac{1n(x)}{\mu}$

The arrival rates (headway) are randomly distributed about a mean value.

Figure 3: Step Shaped Mean Arrival Rate in Peak Period



Where,

h = headway,

x = random number between (0, 1),

μ = mean arrival rate.

In this paper, we have assumed a step shaped mean arrival rate, an approximation to reflect the peak period case where the arrival rate builds up and finally dissipates. This is depicted in Figure 3. The step function is divided into three equal intervals over the entire simulation period. The assumed values of the mean arrival rate (μ) are given in Table 1. The simulations were repeated for three different congestion levels—low, moderate, and high.

Service Rate

The service rate is the ramp-metering rate. The ramp meter red times used in the Twin Cities currently range from 2.2 seconds to 17 seconds. In this research, the red times of 5, 10, and 15 seconds have been used to represent low, moderate, and high congestion conditions. For every truck that takes the bypass, the ramp meter skips one cycle to maintain the overall departure rate of the queue.

The simulated queue has a combination of stochastic and deterministic queuing. In the first step, the mean arrival rate is less than the departure rate and hence we have stochastic queuing, but towards the center of the peak period, there is deterministic queuing, as the mean arrival rate exceeds the departure rate.

The arriving vehicle is randomly assigned as a truck or auto based on the fraction of trucks in the traffic stream (f). For every arriving truck, the *expected delay* is calculated as a function of the queue length ahead of it and the service rate. Based on the expected delay, the truck makes a choice of taking the tolled route, which is modeled stochastically. Using a random number sampled from the uniform distribution, and the choice probability calculated from the expected delay or toll, a choice is made for that truck. This represents the case wherein the driver looks at the conditions ahead and makes a choice. Thus each truck's choice process is modeled. If the vehicle is an auto, it does not have a choice and is added to the ramp.

Value of Time

The value of time for the vehicle is a function of its characteristics. There have been many studies to determine the VoT and there is no single value we can assign. The studies in the literature have estimated value of time based on trip type, vehicle type, and location. The following three studies include these estimates for autos. Beesley (1965) made one of the early estimates from a survey of a set of government employees in London and he concluded that it is typically 33% to 50% of the average hourly income. Using the Bureau of Labor Statistics (BLS) data, the US average hourly wage rate is around $15.36 (1999 dollars). So, the value of time would be in a range of $5/hour to $7.5/hour. The VoT as used by Highway Economic Requirements System (HERS) scaled to 1999 dollars using the Employment Cost Index (ECI) is $17.30/hour (Federal Highway Administration 2000). The study by Daniel McFadden (1974), where he developed a travel demand model to determine the mode choice between auto and transit after the introduction of the Bay Area Rapid Transit (BART), concluded that the in-vehicle value of time for autos is $1.23/hour (1973 dollars) and the waiting time is $2.34/hour, which is equivalent to $7.5/hour (1999 dollars). In our research, VoT for every arriving car was assigned from a uniform distribution between $7.75/hour to $17.75/hour.

Kawamura (2000) showed from a stated preference survey among various freight companies, that the value of time can be well replicated using a log normal distribution. The study was a survey conducted among different trucking agencies in California and the estimated values were not for a specific truck type, but for the general class of commercial trucks. For the truckers, it was estimated to have a mean of $23.40/hour and a standard deviation of $32/hour. This is implemented in the queuing model as a lognormal of equivalent mean 2.64 and standard deviation 1.139. The value of time for every arriving truck is sampled from this log normal distribution.

Toll Collection

There are a number of ways to implement toll collection. If this were to be implemented, electronic tolling is likely, eventually.

The system may begin with simpler technologies, particularly in a trial period. However, to use a conservative estimate of the toll collection costs, manual toll collection costs are used in the simulation. The manual toll collection costs were taken from Levinson (2002) as 8.5 cents per vehicle. These are costs from a tolling study in the San Francisco Bay area.

Simulation Details

For every additional vehicle arriving, the system parameters—the queue length on the ramp and the ramp-metering rate are updated. In addition, after each arrival has been modeled, the objective function (the toll authority profits, the benefit to the road user) is calculated. We start the simulation at t=0, and generate 1,000 arrivals. This process is repeated for the different values of the toll assumed.

In all the simulations, it is assumed that the high occupancy vehicles that would be using the bypass are ignored, since they are a very small proportion of the traffic and are unaffected by the additional trucks on the bypass. The simulations are repeated 50 times to average the stochastic variations and repeated for the three scenarios of low, moderate, and high congestion. Table 1 lists the assumed values of the simulation model.

RESULTS

For each of the three different scenarios (low, moderate, and high congestion) the variation of the toll authority profits, total user bene-

Table 1: Assumed Model Values

Item	Value
Fraction of trucks in the traffic stream (%)	10
VoT for car (Range) (dollars/hr)	7.75 – 17.75
VoT for truck (Mean) (dollars /hr)	26.80
VoT for truck (Standard Deviation) (dollars /hr)	32.00
Number of vehicles simulated	1000
Toll in steps of 0.5 (dollars)	0 - 19.5
Number of simulation runs per toll	50
Ramp Meter Red time (seconds)	
Scenario – I	5
Scenario – 2	10
Scenario – 3	15
Mean Arrival Rates of the 3 steps (vehicles/seconds)	
Scenario – I	0.06, 0.09, 0.03
Scenario – 2	0.1, 0.12, 0.05
Scenario – 3	0.2, 0.25, 0.1

fit and the system benefit have been plotted in Figures 4, 5, and 6.

First observations on the results—only the *toll authority profit maximization* scenario has produced a *nonzero optimal toll.* Both the *user benefit maximization* and the *system benefit maximization* have produced *zero optimal* tolls. Also, as we move across scenarios, we can see that the type of the variation in the toll authority's profit is similar for the three scenarios. As the red times of the ramp meter increase, the curves shift along the increasing direction, an indication that the optimal toll under toll authority profit maximization increases with congestion.

The toll authority profit maximization scenario has produced our expected result. The shape of the curve reflects the tradeoff between the toll and profits. The user benefit maximization scenario indicates that the user gains are a maximum at zero-toll. This is explained as follows. Since we are trying to maximize the consumers' surplus, the upper shaded area (Figure 2) is maximized at a zero cost, and hence we have a zero toll.

The more interesting result is the zero-toll for the system benefit maximization. The zero-toll can be explained as follows. In this scenario, we are trying to maximize the consumers' surplus plus producer's surplus. When there is no toll, the toll authority has no cost and no revenue. Therefore, the consumers' surplus governs the maximization process and it is maximized at a zero-toll. Introducing tolls and toll collection immediately results in a welfare loss (Figure 2).

One important point that needs to be noted is that even though the total time gains in the system are nearly zero (the total delay is nearly constant) the delays experienced by separate vehicles are scaled with different values of time. Hence the system gains when the higher VoT vehicles have more travel time savings than the lower VoT vehicles. In addition the toll paid is a transfer within the system across the users and the toll authority.

Figure 4: Toll Authority Profit for Various Tolls

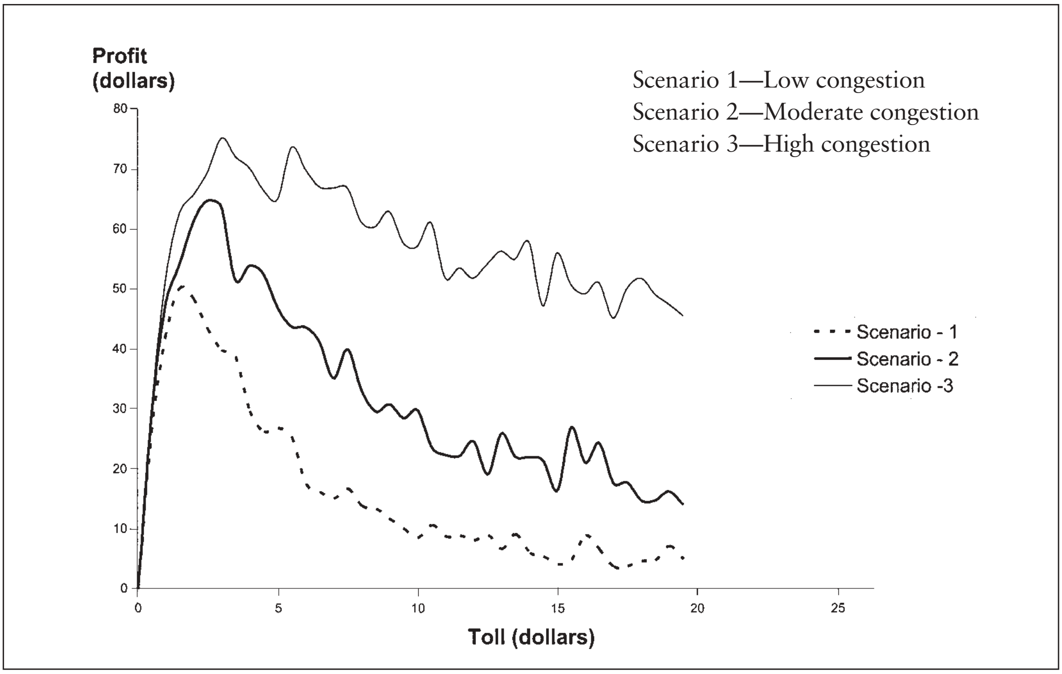


Figure 5: Total User Benefit for Various Tolls

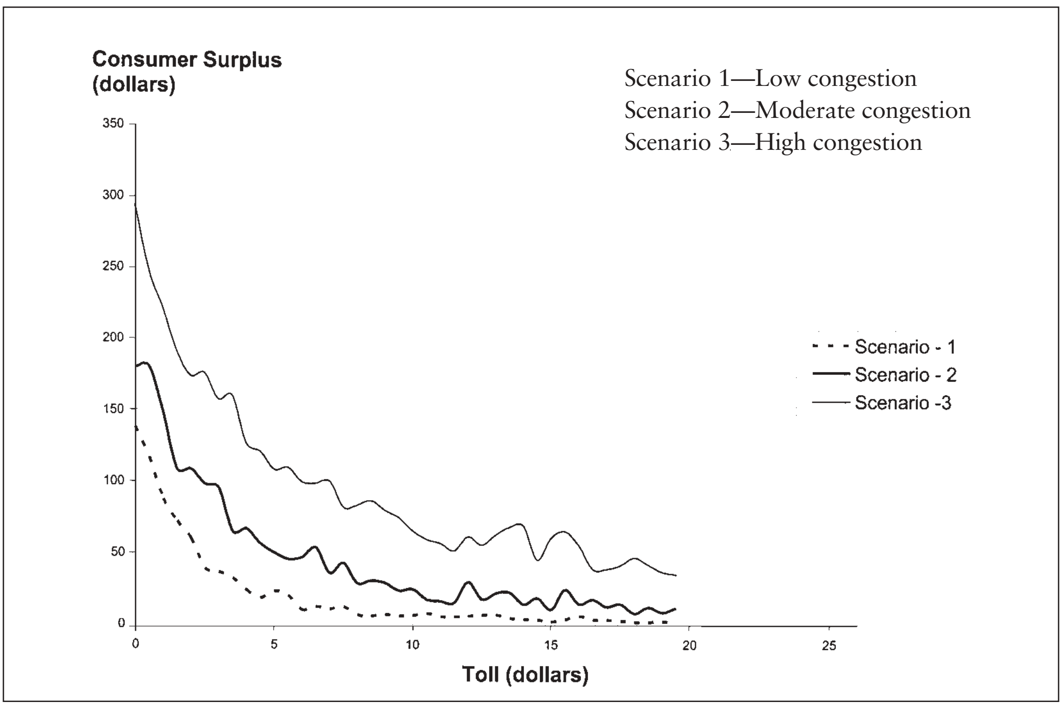
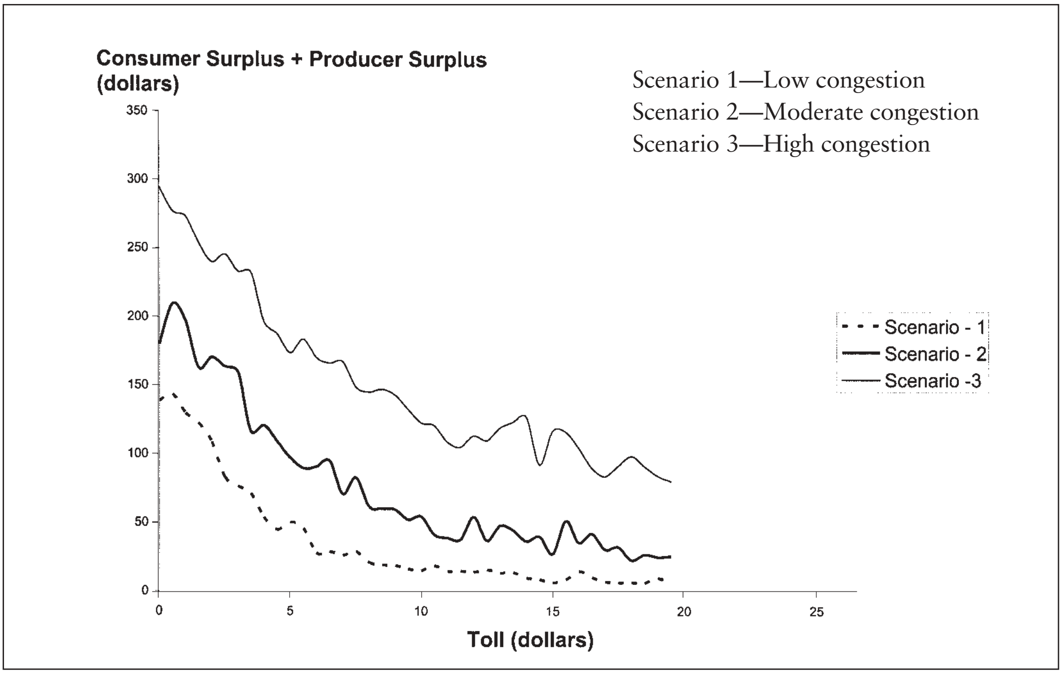


Figure 6. Total System Benefit for Various Tolls



Therefore the system gains are only the travel time savings scaled by the value of time.

The last feature of the results is comparison between the base case that prohibits the trucks from using the bypass and the case where they are allowed to pay a toll and use the bypass. Since the benefits of the bypass to the user are positive, it implies that the bypass at any toll is a better alternative to the base case, which prohibits trucks from using the bypass.

CONCLUSIONS

Opening the HOV ramp meter bypass as a tolled lane for trucks—an identifiable class of vehicles in the traffic stream was proposed in this paper. A single peak period toll was proposed. The optimal toll was estimated under three different scenarios—toll authority profit maximization, user benefit maximization, and system benefit maximization. The toll authority profit maximization case resulted in a nonzero toll that increased with delays/congestion. System benefits were maximized at a zero toll. This implies that the bypass should be free to the trucks. This raises equity issues that need to be addressed. Society is better off by opening ramp meter bypasses to trucks at any toll. However, the road owner must find a toll that is politically acceptable, which may be between the profit-maximizing toll and zero. Also the use of toll revenues would need to be considered.

This paper has important implications for policies regarding the ramp meter HOV bypass lanes. As our research has indicated, it is more beneficial to open the underutilized lanes as Truck Toll lanes, so the ramp meter bypass could be used in a manner similar to a High Occupancy Truck Toll (HOTT) lane. Excessive traffic on the bypass will not only delay SOV commuters significantly and incur political opposition but also undermine the effectiveness of ramp meters. Opening underutilized high occupancy vehicle lanes, as tolled lanes would be appropriate for a small and easily discriminated portion of traffic such as trucks.

A proposal such as this should be welcome in the trucking industry, particularly for the express carriers, because they can save time spent on the metered ramps. The tolling scheme need not be dynamic in real time and an inexpensive administrative system such as prepaid passes may be workable.

However, the next question is the issue of equity. A truck toll scheme allows trucks to buy their way to the front of the line at the expense of those behind. We can alleviate this by delaying only those behind the truck and not the vehicles already in the queue, at a small cost to the traffic already on the freeway. The other important question is what to do with the collected money. Since this is being done on a small scale there is little money to be distributed. Returning the money to the people delayed on the ramp or investing in transportation facilities might be appropriate.

Endnotes

1. For the derivation of the logit choice probability, see Ben Akiva and Steve Lerman (1985).

References

Beesley. "Value of Time Spent in Traveling: Some New Evidence." *Economica* 32 (126) (1965): 174-185.

Ben Akiva, Moshe and Steve Lerman. *Discrete Choice Analysis.* M.I.T. Press, 1985.

Cambridge Systematics Inc. *Twin Cities Ramp Meter Evaluation.* Prepared for the Minnesota Department of Transportation, 2001.

Dahlgren, Joy. "High Occupancy Vehicle Lanes: Not Always More Effective Than General Purpose Lanes." *Transportation Research Part A* 32a (2) (1998): 99-114.

Federal Highway Administration, US DOT. "Highway Economic Requirement System, Technical Report." FHWA PL-00-028 (2000): 7-1.

Fielding, J. Gordon and D.B. Klein. *High Occupancy/Toll Lanes: Phasing in Congestion Pricing a Lane at a Time.* Reason Foundation, Policy Study 170(1993).

Kawamura, Kazuya. *Perceived Value of Time for Truck Operators.* TRB 79th Annual Meeting, (Pre Print CD-ROM). Paper No.00-0711, 2000.

Levinson, M. David. *Financing Transportation Networks.* Edward Elgar Publishers,

2002.

Levinson, M. David, Lei Zhang, Shantanu Das, and Atif Sheikh. *Evaluating Ramp Meters: Evidence from the Twin Cities Ramp Meter Shutoff.* TRB 81st Annual Meeting, Paper No.02-2167, 2002.

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

Poole, W. Robert and C. K. Orski. *Building a Case of HOT Lanes: A New Approach to Reduce Congestion.* Reason Foundation, Policy Study 257, 1999.

Rogers, A. Christy. "Effects of Ramp Metering with HOV Bypass Lanes on Vehicle Occupancy." In *Transportation Research Record* 1021, TRB, National Research Council, Washington, DC (1985): 10-15.

***Satyanarayana (Satya) Muthuswamy** is currently a traffic engineer with KLD Associates Inc., Huntington, New York, working on simulation models for toll plaza operations. Muthuswamy hails from India and earned a Bachelor's from the Indian Institute of Technology, Madras. He graduated from the University of Minnesota with a Master of Science in Civil Engineering. During graduate study, Muthuswamy worked on travel demand estimation for freeway sections using simulation and optimization. His areas of interest are eimulation and modeling, transportation economics, and intelligent transportation systems.*

***David M. Levinson** is an assistant professor in the Department of Civil Engineering at the University of Minnesota. Levinson recently completed the book* Financing Transportation Networks, *published by Edward Elgar. Levinson's Ph.D. in Transportation Engineering was earned at the University of California, Berkeley. He has conducted research into travel behavior and received the 1995 Tiebout Prize in Regional Science for "Location, Relocation, and the Journey to Work." Current research projects include: Measuring the Equity and Efficiency of Ramp Meters; Modeling Network Growth; Ramp Metering, Freeway Congestion, and Driver Acceptance; and Evaluating the Spring Load Restrictions Policy.*