Pricing Urban Congestion*

Ian W.H. Parry

Resources for the Future, Washington, D.C. 20036; Web site: http://www.rff.org/parry.cfm; email: parry@rff.org

Annu. Rev. Resour. Econ. 2009. 1:461–84

First published online as a Review in Advance on June 5, 2009

The *Annual Review of Resource Economics* is online at resource.annualreviews.org

This article's doi:
10.1146/annurev.resource.050708.144226

1941-1340/09/1010-0461$20.00

Key Words

traffic congestion, externality, peak-period fee, congestion toll incidence

Abstract

This paper reviews literature on the optimal design of pricing policies to reduce urban automobile congestion. The implications of a range of complicating factors are considered; these include traffic bottlenecks, constraints on which roads and freeway lanes in the road network can be priced, driver heterogeneity, private toll operators, other externalities besides congestion, and interactions between congestion taxes and the broader fiscal system. I also briefly discuss the incidence of congestion taxes and experience with this policy in the United States and elsewhere. Although the economics literature on congestion pricing has advanced considerably over the past 20 years, research is still needed on the empirical measurement of second-best efficient tolls for urban centers and whether alternative design features have substantial implications for efficiency. More research is also needed on the design of schemes to promote feasibility by compensating adversely affected groups with minimal loss in economic efficiency.

1. INTRODUCTION

Given relentless growth in population and real income, expanding demand for automobile travel in the United States continues to outpace road construction, causing worsening urban congestion. Between 1980 and 2003, for example, urban vehicle miles traveled increased by 111%, against an increase in lane-mile capacity of only 51% (Bureau of Transportation Statistics 2006, tables 1.6 and 1.33). According to Schrank & Lomax (2007), the average traveler across the 437 largest urban areas in the United States lost 38 h to traffic delays in 2005, up from 14 h in 1982. Delays are most severe in Los Angeles, where the average traveler lost 72 h to congestion in 2005. The next most congested cities, with average delays of approximately 60 h per year, were San Francisco, California; Washington, D.C.; Atlanta, Georgia; Dallas, Texas; and Houston, Texas (Schrank & Lomax 2007, table 1). Nationwide, Schrank & Lomax (2007) put the annual costs of congestion, including wasted fuel, at $78.2 billion in 2005, up from $16.2 billion in 1982 (in 2007 dollars).[1]

Despite higher fuel prices, the trend of rising urban congestion is set to continue. The Department of Transportation (2008) projects an increase in automobile vehicle miles of 50% between 2010 and 2030. Meanwhile, because of environmental constraints, neighborhood opposition, and high land acquisition costs, new road construction is increasingly difficult. In any case, expanding capacity is partly self-defeating as it encourages new driving trips (see, for example, Downs 1992, Standing Advisory Committee on Trunk Road Assessment 1994, Goodwin 1996, Mackie 1996, Noland 2001, Litman 2006).

The limitations of other traditional approaches have also become apparent. Expanding transit and subsidizing fares has limited impacts on automobile congestion, given relatively modest own-price elasticities for transit.[2] In fact, the convenience of auto travel, particularly relative to traditional hub-and-spoke rail networks, may be increasing as places of work become more dispersed, rather than concentrated in the downtown area. Fuel taxes have limited effect as they do not differentiate between urban and rural driving or between peak and off-peak travel, and much of the long-run behavioral response comes from improved fuel economy rather than reduced vehicle mileage. Furthermore, political opposition to fuel taxes is intense in the United States, where auto and oil companies have substantial political influence and where per capita gasoline consumption is several times higher than in Western Europe.

It is therefore not surprising that U.S. policy makers are looking for more effective congestion policies (e.g., De Corla-Souza 2004, National Surface Transportation Policy 2007, Department of Transportation 2006). In theory, peak-period road pricing (sometimes called value pricing) is the ideal policy in this regard because it exploits all behavioral responses for reducing congestion, such as reduced overall travel; increased

[1]This figure omits some broader costs of congestion, such as the costs of people deviating from their preferred travel times to avoid the rush-hour peak. Conversely, there are limits to the costs of congestion as people cut back on peak period trips, change housing and job location, etc., as congestion becomes more severe.

[2]A typical estimate for the own-price elasticity for transit is approximately -0.4 (Pratt et al. 2000). Nonetheless, urban transit fares are heavily subsidized. Fare subsidies for the 20 largest transit authorizes in the United States, expressed as the difference between agency operating costs and revenues from passenger fares, vary from 30% to 90% of operating costs for rail systems and from 60% to 90% for bus systems (Parry & Small 2008, table 1). Despite these subsidies, transit accounted for only 4.4% of nationwide commuting trips in 2005, whereas automobiles made up 88.4% (Bureau of Transportation Statistics 2006, table 1-38). Improving service quality (e.g., increasing transit speed, reducing wait times at stops, and improving transit access) may be more effective in deterring automobile use (Litman 2007a).

carpooling; and shifting trips to off-peak periods, to transit, and to less congested routes. Moreover, the feasibility of peak-period pricing has greatly improved with recent developments in electronic metering technology. Fees can now be deducted electronically by in-vehicle transponders, thus reducing bottlenecks at manual tollbooths, or by direct billing with onboard global-positioning systems. Congestion fees may also help bridge the growing funding gap for financing upgrades of the aging transportation infrastructure, given that real fuel-tax revenue per automobile mile has declined with greater fuel economy and the failure of nominal tax rates to keep pace with inflation (Transportation Research Board 2006).

Although there are few successful congestion pricing schemes in the United States so far, it is most likely that congestion pricing will become more appealing as urban travel speeds continue to deteriorate. The relative success of area license fees in London further suggests that public opposition is not insurmountable. Thus, guidance from transportation economists on how congestion pricing policies should be designed, and in ways to reduce public opposition, could not be more timely.

Unfortunately, even at a conceptual level, designing taxes to reduce congestion can be far more complicated than addressing other externalities like household garbage, drunk driving, and smokestack pollution. Real-time variation in the toll rate within the peak period is needed to optimize road capacity by partially flattening the distribution of trip departure times. Moreover, because of political or other constraints, congestion pricing is emerging piecemeal, typically on one lane of one highway at a time. In assessing the appropriate toll, account must be taken of changes in congestion on parallel (unpriced) freeway lanes and on other links in the urban network, as the toll induces drivers to alter their travel routes.

The congestion pricing literature has advanced considerably over the past 20 years, providing valuable insights on various design issues. More research is needed, however, on the empirical measurement of efficient tolls for major urban centers and on the efficiency implications of design features—such as toll variation within the peak period, supplementary pricing on other links in the network, and toll exemption provisions (e.g., for taxis or clean-fuel vehicles). Moreover, not much literature addresses the design of compensation schemes to advance political feasibility, without substantial loss of economic efficiency.

Drawing on several other reviews, this paper distills some key findings from the congestion pricing literature and issues in need of further study, and discusses experience with congestion tolls to date and prospects for more widespread policy implementation in the United States.[3] Section 2 briefly discusses alternative models of congestion. Section 3 discusses complicating factors in the design of congestion tolls. Section 4 considers practical obstacles to, and prospects for improving the feasibility of, congestion pricing. Section 5 summarizes experience with congestion pricing to date. Section 6 offers concluding thoughts.

[3]Other recent reviews include Arnott et al. (2005, ch. 1), Lindsey (2006, 2007), Litman (2008), Santos (2004a), and Small & Verhoef (2007). As Arnott et al. (2005) emphasized, economists need to examine policies that may complement congestion pricing, such as appropriately pricing freight and mass transit, staggering work hours for government employees, encouraging biking and walking, and improving the design of roads and intersections to improve traffic flow. Reforming the pricing of parking space is especially important, given that owner- and employer-provided public parking is currently heavily subsidized (Shoup 2005). These broader policies, however, are beyond the scope of this article.

2. ALTERNATIVE APPROACHES TO MODELING CONGESTION

In this section, I outline the two main theoretical approaches to modeling congestion: (*a*) the static, speed/flow model of highway congestion and (*b*) the dynamic model of traffic bottlenecks. I also comment on empirical implementation of these models.

2.1. Basic Model of Highway Congestion

Beckmann et al. (1956), Walters (1961), and Vickrey (1963) developed the basic model of highway congestion, and Hau (2005a,b) provides a recent exposition and various extensions to the basic model. Here, I begin with the basic model of congestion along a uniform segment of an urban freeway. This model makes a number of simplifying assumptions (the implications of relaxing them are discussed below):

- No traffic bottlenecks at highway entry and exit points or intersections
- A uniform flow of incoming traffic across the period
- Uniform traffic flows, and congestion tolls, across all lanes of the freeway
- All motorists have the same value of travel time
- No linkages with congestion on other network links
- No interactions between congestion tolls and distortions in the broader economy
- No consideration of recycling congestion toll revenues

2.1.1. Basic traffic-engineering relationships. Underlying the model is the fundamental diagram of traffic congestion (shown in **Figure 1**). The upper-left quadrant indicates the relation between vehicle density (D)—that is, the average number of vehicles along the highway at a given time—and average speed (S). Speed declines with higher density because drivers slow down to maintain a comfortable separation from the vehicle ahead given shorter distances between vehicles. Traffic volume, or flow (V), is the number of vehicles passing through the highway segment per unit of time. Volume (vehicles/hour) is the product of density (vehicles/mile) and speed (miles/hour). Higher density is initially associated with higher vehicle flows, or completed highway trips per hour (lower-right quadrant of **Figure 1**). However, at some point, the contribution of an extra vehicle to traffic flow is offset by the reduction in flow attributable to existing vehicles traveling at slightly slower speeds to accommodate more traffic. This point represents the maximum carrying capacity. Beyond this point, the highway is said to experience hypercongestion, because additional vehicle density reduces the overall vehicle flow. This implies an inverted-U relation between speed and volume (the upper-right quadrant of **Figure 1**).

Table 1 shows the approximate relationship between density, speed, and volume under ideal highway conditions (e.g., with no intersections and bends). The flow peaks at roughly 1850–2000 vehicles/hour/lane of highway at speeds of approximately 30–46 mph (i.e., at a particular point on the highway, one vehicle passes approximately every 2 s).

2.1.2. Basic economic analysis. The traditional economic analysis of congestion tolls uses these relationships to plot average and marginal travel cost, as a function of vehicle flow, against travel demand (depicted in **Figure 2**). Here, the average cost (AC) of highway trips/hour is

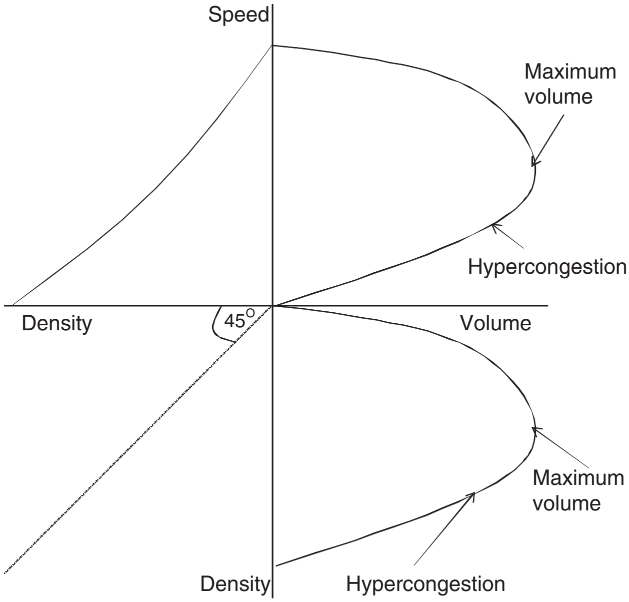

Figure 1

The fundamental diagram of traffic congestion.

Table 1 Typical density, speed, and flow relationships under optimal highway conditions[a]

Density (vehicles/mile)	Speed (mile/hour)	Volume (vehicles/hour)
<12	>60	<700
12–20	57–60	700–1100
20–30	54–57	1100–1550
30–42	46–54	1550–1850
42–67	30–46	1850–2000
>67	<30	Unstable

[a]Source: Homburger et al. (1992).

$$AC = c_m + VOT \cdot T(V), \tag{1}$$

where c_m denotes the money cost per trip before any toll, reflecting, for example, fuel costs, vehicle wear and tear, possible parking fees, etc.[4] VOT [value of (travel) time] represents the amount drivers would be willing to pay to save one hour of travel time (see below). $T(V)$ is the time per trip (the inverse of speed), which rises with increasing traffic/hour on the highway. The average cost curve in **Figure 2** bends backward after the

[4]I omit the possibility that congestion could lower fuel economy and thus raise fuel costs per trip. A typical assumption is that vehicle fuel consumption increases by 30% under heavily congested conditions, though there is considerable uncertainty over this figure (see, e.g., Greenwood & Bennett 1996; Small & Gómez-Ibáñez 1999, sect. 3.2).

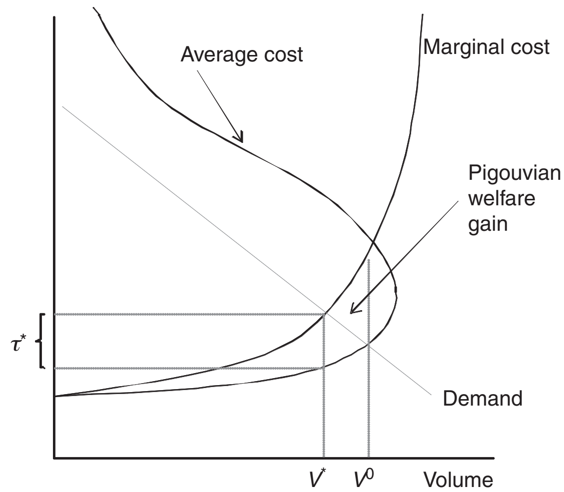

Figure 2

Traditional model of optimal congestion tolls.

maximum traffic flow is exceeded. This reflects the hypercongested portion of the speed-flow curve in **Figure 1**, where additional vehicles increase delays to other drivers by enough to reduce the overall traffic flow. However, as discussed below, a dynamic analysis is needed to study hypercongestion, which is a transitory phenomenon. For the demand curve in **Figure 2**, the equilibrium traffic flow with no toll is V^0, because motorists use the highway until the average cost per trip equals the benefit (or height of the demand curve).

Multiplying *AC* by *V* and differentiating gives the marginal social cost (*MC*):

$$MC = AC + MEC; \;\; MEC = VOT \cdot VT'(V). \tag{2}$$

The marginal cost exceeds the average cost by the marginal external cost of congestion (*MEC*), which equals the increase in travel time attributable to extra congestion from one more trip, times the number of trips/hour, times *VOT*.

The socially efficient traffic flow for the demand curve in **Figure 2** is V^*, where the marginal social cost of an extra trip equals the marginal benefit. This flow could be induced by levying the Pigouvian congestion toll, τ^* in **Figure 2**, equal to the gap between *MC* and *AC* at this point, or $wV^*T'(V^*)$. Note that this toll is less than *MEC* at the pretoll traffic flow because *MEC* declines with the reduction in traffic flow.[5] The welfare gain from the Pigouvian toll is the gap between the marginal social cost curve and the demand curve, integrated over the reduction in traffic flow.

Revenue raised by the Pigouvian toll is $\tau^* V^*$. Before any revenue recycling, all drivers are worse off under the toll (assuming they have the same VOT) which is a challenge for political feasibility (see below). For motorists who continue to use the highway, average costs increase because the toll exceeds the VOT savings, whereas drivers who are diverted from the road to their next best alternative are also worse off.

[5]According to Litman (2007b, pp. 5.5–5.10), optimal tolls would reduce traffic volume to roughly 1500 vehicles/lane/hour on highways and 800 on urban arterials.

2.2. Bottleneck Model

Although the basic speed-flow model is a useful starting point for analyzing situations when traffic conditions do not change quickly, or when the focus is on average traffic levels over extended periods, it has two main shortcomings (Small and Chu 2003). First, it cannot accommodate the possibility of hypercongestion. Yet, hypercongestion is a real phenomenon at various choke points in the road network (e.g., at stoplights, at highway entry or exit points) where large queues form and then clear during rush hour. In fact, for some of the world's most congested cities, such as Bangkok, Athens, Rome, and Jakarta, most of the road network is hypercongested most of the time. Second, the traditional model captures only one behavioral response, that is, whether to drive on the congested highway. In reality, within the peak period, drivers can change their departure time to avoid the point at which congestion is most severe. Thus, traffic inflow to the highway per unit of time is endogenous not fixed. The bottleneck model of congestion was developed to address these two shortcomings (e.g., Vickrey 1969; Arnott et al. 1991, 1993, 1994).

In bottleneck models, motorists have a preferred arrival time and incur a rising cost for early or late arrival (e.g., workers with fixed work schedules). Drivers choose their trip departure time to trade off these "schedule delay" costs against travel-time savings from leaving before, or after, the rush-hour peak. In the simplest setting, one bottleneck permits a maximum throughput of vehicles/hour. As the inflow rises above the maximum outflow, a queue forms, peaking and then progressively declining as the end of the peak period is approached.

The optimal toll in the basic model rises over the first part of the peak period and then falls, keeping traffic inflow equal to capacity outflow. This dynamic toll thus "flattens" the peak by inducing some people to depart earlier, if they leave before the peak, others to leave later, if they depart after the peak, and still others to use alternate routes or avoid driving all together. In contrast to the speed-flow model, the optimal toll eliminates congestion entirely because there is no queuing when traffic inflow is at or below choke-point capacity.

The inverted-U schedule for the optimal congestion toll during the peak period is an important policy insight from the bottleneck model. Numerical simulations in Arnott et al. (1993) suggest that more than half of the welfare gains from congestion pricing may come from trip rescheduling within the peak, rather than avoidance of peak travel altogether.

Another insight is that the costs of congestion are greater than the extra travel time alone. They also include schedule-delay costs, because people deviate from their preferred arrival times to save time during their trip (Arnott et al. 1994, Small & Chu 2003). However, the extent to which schedule-delay costs are picked up in empirical estimates of the VOT, and therefore are incorporated in estimates of congestion costs, is unclear.

A third insight is that the welfare gains from bottleneck pricing can be roughly the same as the toll revenue collected, unlike in the traditional model, where welfare gains are typically much smaller than revenue transfers. Because it is optimal to eliminate bottleneck queues completely rather than partly, welfare gains are first order, unlike in the static model where they are second order. In principle, this should reduce public opposition to tolls designed to alleviate specific bottlenecks in the road network.

Hybrid models combining elements of both traditional and bottleneck models have also been developed. For example, Mun (1994) developed a dynamic model of travel between two distant points with a queue in the middle that forms and eventually clears during the peak period. Travel time is determined by the standard speed-flow relation on either side of the bottleneck, but it also includes wait times from queuing. The model thus captures hypercongestion at a choke point, but over the entire peak period, the average travel cost curve always increases in traffic flow. In hybrid models, the optimal toll has both a static component analogous to the Pigouvian tax and a dynamic component to address the bottleneck (Arnott & Kraus 1998, equation 20).

2.3. Empirical Implementation of the Basic Model

Most empirical assessments of optimal congestion tolls use the speed-flow approach, which is the model I focus on here.[6] Implementing the basic speed-flow model of congestion charges requires three pieces of information: the VOT, the speed-flow functional relationship, and the demand response to tolling.

2.3.1. Value of travel time. If people value the pure disutility from an extra hour of work and an extra hour of travel time equally, then the VOT would reflect the net-of-tax wage. More generally, travel may be valued at less than the net wage if, for example, people prefer to be in a car rather than at work, or vice versa, if they prefer the work environment to being in a car.

There is a large empirical literature on the VOT. Some studies use revealed preference methods (e.g., estimating willingness to pay auto fuel and parking costs to save time over an alternative, slower travel mode), whereas others use stated preference methods (e.g., directly asking people what toll they would be willing to pay for a faster commute). Stated-preference studies (e.g., Calfee & Winston 1998) often yield a much lower VOT than do revealed-preference approaches. Brownstone & Small (2005) suggested that this may be due to survey respondents' overestimating the actual time savings from travel options with higher monetary costs.

Given the difficulty of controlling for schedule-delay costs, the extent to which people's estimated willingness to pay for shorter travel times also reflects the advantages of their being able to schedule trips closer to their preferred times is unclear. Another cost of travel delay is the added uncertainty over arrival times when trip times are stochastic, which matters when the cost of being late by a given margin exceeds the cost of being early by the same margin. In response to this uncertainty, people may leave earlier than they would otherwise prefer. However, it is also unclear to what extent the cost of this "buffer time" is reflected in existing estimates of the VOT (Small & Verhoef 2007, p. 54).

Most literature reviews recommend a VOT for personal auto travel equal to approximately half the gross market wage, or somewhat less than the net-of-tax wage (e.g., Waters 1996, U.S. Department of Transportation 1997, Mackie et al. 2003). The recommendation suggests, for example, that if the gross hourly wage for urban areas is \$20 (Bureau of Labor Statistic 2006, table 1), then the time cost for driving 1 mile when the road speed is 30 mph will be 33 cents. For comparison, if the vehicle fuel economy is 20 miles per gallon and the retail gasoline price is \$3.50 per gallon, then fuel costs per mile are 15 cents. Commuting trips are also valued somewhat more highly than leisure, shopping, and other trips, partly

[6]Sometimes, delays at bottlenecks, averaged over the peak period, are taken into account when calibrating speed-flow curves. Quantitative assessments of optimal tolls based on bottleneck models typically must postulate a distribution for schedule delay costs, because there are no direct data to measure them. Numerically solving these models is challenging, given that bottleneck congestion is inherently a disequilibrium phenomenon (for more discussion, see de Palma & Marchal 2001).

because of penalties for late arrival at work. Trips made in the course of work are typically valued at the gross wage, as they reflect the VOT to the employer (Department for Transport 2007). Moreover, as a possible result of stress and frustration, the VOT may be substantially higher under heavily congested, compared with free-flowing, conditions (e.g., MVA Consultancy et al. 1987, Wardman 2001, Steer Davies Gleave 2004). As regards the VOT/income elasticity, Mackie et al. (2003) recommend a value of 0.8.

2.3.2. Speed-flow curves. The most commonly used functional form relating travel time per mile (the inverse of speed) to traffic flow (where these are observed from time-lapse satellite data or ground-based traffic counts) is

$$T = T_f\{1 + \alpha V^{\beta}\}, \tag{3}$$

where α and β are parameters and T_f is time per mile when traffic is free flowing. A typical value for the exponent β is 2.5–5.0. With $\alpha = 0.15$ and $\beta = 4.0$, Equation 3 is the Bureau of Public Roads formula, which is widely used in traffic-engineering models (for a review of the literature on speed-flow curves, see Lindsey & Verhoef 2000).

Differentiating Equation 3 and substituting in Equation 2 give

$$MEC = VOT \cdot \beta \cdot AD, AD = T - T_f, \tag{4}$$

where AD (average delay) represents the excess of time per mile over that under free-flow traffic conditions. Using this formula, Parry & Small (2009) valued the marginal external cost of congestion at 28.0 cents/vehicle-mile for peak automobile travel in metropolitan Washington, D.C., in 2002. **Table 2** extrapolates this estimate to various other large cities, on the basis of travel delay per mile of travel in that city relative to that for Washington, D.C. (see Schrank & Lomax 2007). These rough estimates vary from 33.7 cents/mile in Chicago, Illinois, to 23.5 cents/mile in Detroit, Michigan. These figures may mask considerable variation in marginal external costs across individual links in the network as well as across points in time within the peak period. For example, for peak travel in Twin Cities, Minnesota, Mohring (1999) estimated marginal congestion costs that vary from less than 2.5 cents to more than 50 cents/vehicle-mile, across road classes.

Accurately estimating the parameters of speed-flow relations for specific roads can be tricky, however. For example, the relation will vary with highway gradient, bends, presence of hard shoulders, frequency of stop lights and intersections, etc., so estimates may not transfer across different roads. Furthermore, estimating the point of maximum flow can be challenging because of considerable scatter in observed speed and flow data (Small & Verhoef 2007, ch. 3), and estimates of the speed-flow curve for specific highway segments are sensitive to bottlenecks near that segment.

2.3.3. Demand responses. Calculating Pigouvian congestion tolls, as opposed to the marginal external cost at prevailing travel flows, requires simultaneously solving for the marginal external cost and the demand for travel as the highway is priced. Demand responses are also needed to assess the net benefits from congestion tolling.

There is a large empirical literature on the overall responsiveness of driving practices to price fluctuations, usually measured by fuel prices or fuel costs per mile. Aggregate cross-sectional studies compare travel behavior across metropolitan areas, or sometimes across different zones within an urban area. Other studies use time-series data, though the results

Table 2 Marginal external congestion costs for selected urban centers[a]

Urban area	Annual person hours of delay[b]	Marginal external congestion cost[c]
Los Angeles, California	490,552	32.4
New York, New York	384,046	31.7
Chicago, Illinois	202,835	33.7
Dallas, Texas	152,129	25.9
Miami, Florida	150,146	28.7
Atlanta, Georgia	132,296	24.7
San Francisco, California	129,919	28.1
Washington, D.C.	127,394	28.0
Houston, Texas	124,131	25.4
Detroit, Michigan	115,547	23.5
San Diego, California	90,711	25.9
San Jose, California	50,038	26.0
Orlando, Florida	40,595	24.4

[a]Sources: Schrank & Lomax (2007), Parry & Small (2009).
[b]Measured across all times of day (thousands of hours).
[c]Measured for peak auto travel (cents/vehicle mile).

can be sensitive to the specification for autocorrelation. A rough rule of thumb is that the elasticity of vehicle miles with respect to fuel prices is between -0.1 and -0.3 (e.g., Goodwin 1992, Goodwin et al. 2004).

However, estimating the potential demand response to peak-period pricing of a link in a road network, as opposed to a uniform increase in the price of all driving, is problematic. Typically, this can be done only after the response to pricing that link has been observed. *Ex ante*, studies may extrapolate estimates of the degree of substitution in demand between priced and unpriced routes from studies of other, previously implemented, congestion pricing policies in other cities, making some adjustment for differences in the proximity of other roads to the priced road across cities.

3. COMPLICATING FACTORS IN THE DESIGN OF CONGESTION TOLLS

Although the speed-flow model is the basis for most empirical assessments of optimal congestion tolls, other assumptions—besides the absence of bottlenecks—are often unrealistic. I now discuss literature that relaxes these other assumptions.

3.1. Limited Pricing Across Freeway Lanes

To date, in the United States, part of the political deal-making needed to implement congestion pricing on freeways has involved leaving motorists with the option of an unpriced, though more congested, alternative lane on the freeway. Suppose, first, that all drivers are homogeneous and a toll is applied to one of two parallel freeway lanes. Because the lanes are perfect substitutes, traffic from one lane will move to the other until the average cost of the unpriced (but more congested) lane equals the average cost of its priced (but less congested) counterpart. In an extreme case, where the demand for travel on the freeway is perfectly inelastic, all the reduction in traffic on the priced lane is shifted onto the other lane and the toll reduces welfare.[7] More generally, with some elasticity in demand for freeway travel, some drivers diverted from the priced lane will give up using the freeway altogether. Nonetheless, accounting for the partial shifting of drivers onto the congested alternative greatly reduces the optimal level, as well as welfare gains, from the single-lane toll. Welfare gains from tolling one lane (initially carrying half of the freeway traffic) are well below half of the potential welfare gains from pricing both lanes. For example, Verhoef et al. (1996) estimated the potential welfare gains from a single-lane toll at only 10% of those from first-best tolls applied to both routes (see also Braid 1996, Liu & MacDonald 1998).

3.2. Driver Heterogeneity

In reality, the VOT differs greatly among drivers, which affects the optimal set of freeway tolls and welfare gains from single-lane tolls. With driver heterogeneity, the best pricing scheme is not a uniform toll across all lanes, but rather differentiated tolls that allow drivers to sort themselves into lanes that are more or less congested, depending on their VOT. Surprisingly however, studies allowing for differences in the VOT find the efficiency gains from differentiated tolling may not be large. For example, Verhoef & Small (2004) and Parry (2002) estimated that uniform tolls may generate more than 90% of the potential welfare gains from the first-best, differentiated set of lane tolls. This is because in the first-best outcome the difference between the tolls, or among marginal external congestion costs across different lanes, is modest. Although people in the high-toll lane have a higher VOT, which raises the marginal external cost of congestion for that lane, the fact that there are fewer drivers in that lane—thus, there is less congestion—partly counteracts the higher cost.

In contrast, the welfare effects of single-lane tolls are substantially enhanced when account is taken of the possibilities for drivers with a high and low VOT to sort into priced and unpriced lanes. Small & Yan (2001) estimated the efficiency gains from single-lane tolls could be three times as large, when driver heterogeneity is taken into account (though the welfare gains are still below half of those from first-best pricing of both freeway lanes). This reflects greater gains from reducing congestion in the priced lane (where drivers have a high VOT) and smaller losses from extra congestion in the unpriced lane (where drivers have a low VOT).[8]

[7]In effect, the reduction in congestion from the first vehicle diverted off the priced lane will be exactly offset by added congestion on the unpriced lane. Because marginal congestion costs are now (slightly) higher on the unpriced lane, any further diversion of traffic between lanes will increase total congestion costs and lower welfare.

[8]Heterogeneity in the size of passenger vehicles is less important than driver heterogeneity because differences in the amount of road space taken up by cars versus light-duty trucks (pickups, sport utility vehicles, minivans) are modest relative to average headways between vehicles on the road. Estimated differences in the marginal external costs of congestion for different types of passenger vehicles are therefore not large (Federal Highway Administration 1997, table 5-23). Heavy-duty trucks, however, take up more than twice the road space of passenger cars, implying they should have a separate, and higher, toll.

3.3. Network Effects

Generally, a congested freeway segment is just one link in a road network covering an urban center. By diverting drivers from the freeway, congestion tolls may exacerbate congestion on substitute roads elsewhere in the system or may reduce it on complementary roads feeding into the priced segment. Ideally, congestion on other roads would also be internalized through tolling; in which case, (small) changes in traffic on those roads have no efficiency effects. More realistically, there will not be comprehensive pricing of all other congested roads because of political constraints and perhaps because of high monitoring costs associated with pricing crisscrossing city streets. Under these conditions, the second-best toll differs from the Pigouvian toll, the greater the marginal congestion costs on other roads (net of any toll on those roads), and the greater the portion of drivers diverted off the priced freeway that moves to other roads, as opposed to those who cancel their trip or substitute to the off-peak period (MacDonald 1995, Verhoef 2002).

General statements about the sign of an adjustment to the second-best toll—let alone the magnitude—to account for network effects are difficult, however. This is because the availability of substitute routes and extent of complementary feeder roads are both highly case specific. What is needed is a carefully calibrated computational model of the particular road network under study that realistically captures the main substitution possibilities.

Typically, economically based network models disaggregate an urban system into travel zones where each zone consists of stylized links (such as inbound, outbound, and circumferential) representing an aggregation of arterials and side streets within the zone. Other links, such as freeway segments and bridges, may be represented separately. On the demand side, households are aggregated into groups, perhaps by income class. A decision tree involving the choice of whether to take a trip and, if so, then including which destination, mode, time of day, and route may be used to determine travel demand.

However, few economic-network models exist, given the daunting amount of researcher time and data collection required to develop, calibrate, run, and update them.[9] One such model, developed by Safirova and coauthors, has been applied to metropolitan Washington, D.C. (e.g., Houde et al. 2007). Proost and colleagues have also developed other models for various European cities (e.g., de Borger & Proost 2001). The SATURN model has been used to examine alternative cordon tolls in certain U.K. cities (e.g., May & Milne 2000, May et al. 2002, Santos & Newbery 2002). METROPOLIS, a disequilibrium, dynamic model of bottleneck congestion developed by de Palma and coauthors, has also been applied to European cities (e.g., de Palma & Marchal 2001). The limited results available suggest the empirical significance of network effects and, hence, the potential usefulness of such models (e.g., Safirova & Gillingham 2003).

But even within sophisticated network models, neither the link- (or zone-)specific speed-flow curves, nor the own- and cross-price elasticities of travel demand by road and time of day, can be known precisely. Policy makers may therefore need to rely on trial-and-error approaches where tolls are initially set on the basis of existing models and then

[9]Traffic-engineering models of road networks are far more common than economic models and are widely used in forecasting future traffic flows and the traffic impacts of policies such as infrastructure upgrades. Engineering models do not provide welfare-based measures of congestion costs, however, and often do not integrate demand-side behavioral responses to tolls and changes in congestion. Thus, they cannot be used to estimate the welfare effects, and economically efficient levels, of congestion tolls.

revised as models are updated in response to observed, policy-induced changes in travel patterns (e.g., Yang et al. 2004).

3.4. Tolling to Open Up Underutilized Road Capacity

So far, I have compared congestion tolls against a baseline with no policies. However, in some metropolitan areas, certain freeway lanes are restricted to high-occupancy vehicles (HOVs). Converting these lanes to high-occupancy toll (HOT) lanes, where drivers of single-occupant vehicles can pay to use the lane, has several beneficial effects. Unlike in the basic speed-flow model, those paying the toll are better off (they would not pay the toll unless they value the travel-time savings more than the toll). Drivers remaining on the unpriced alternative lane benefit from reduced congestion because single-occupant vehicles switch to the HOT lane. The government also gains tax revenues. The only losers are passengers of HOVs, who suffer from a decline in speed on the HOT lane, though this slowdown may be limited if the toll is high enough to retain reasonable traffic flow. For Washington, D.C., Safirova et al. (2004) estimated the welfare gains from HOV to HOT lane conversion are almost as large as those from more comprehensive pricing covering all lanes of all freeways that currently have HOV lanes.

3.5. Tolling by Private Operators

Private toll roads have been around for some time in Europe and the Pacific Rim and are emerging in the United States (Gómez-Ibáñez & Meyer 1993). If monopoly operators are free to maximize profits, however, the toll will be set above the socially efficient level (Small & Verhoef 2007, ch. 6). Although the operator internalizes congestion by accounting for the increased willingness of drivers to pay for highway use as congestion falls, an additional markup is imposed to exploit monopoly power. The more inelastic the demand for highway travel, that is, the more limited the availability of alternative roads, the greater the markup is. In addition, the divergence between the monopoly toll and the second-best optimal toll is greater if drivers diverted by the toll add to congestion on parallel unpriced lanes of the freeway or on other roads. Under some conditions, the monopoly toll may actually reduce welfare relative to the case of no tolls (Verhoef & Small 2004). Therefore, without a competitive bidding process for toll rates, which would undermine the monopoly markup, there is a case for imposing maximum toll regulations, though the *ex ante* measurement of the efficient toll ceiling may be challenging to obtain before observing behavioral responses to the toll.

3.6. Interactions with Other Externalities

Congestion tolls affect other highway externalities, and to the extent that these externalities are not internalized through other policies, they should be factored into assessments of the welfare effects and, arguably, the optimal levels, of congestion tolls. But do these adjustments make much practical difference? Averaged over urban and rural areas, and over time of day, these other externalities, though not as large as congestion, still appear to be significant. Nationwide, marginal congestion costs have been put at the equivalent of approximately 5–7 cents per vehicle mile (e.g., Federal Highway Administration 1997, 2000). Estimated traffic accident externalities for the United States are almost as large, around 2–7 cents (e.g., Federal Highway Administration 1997, Miller et al. 1998, Parry 2004). The Federal Highway Administration (2000) estimated nationwide local pollution damages at 1.7 cents per mile for 2000, though emission rates are declining over time with more stringent emissions standards for new vehicles.[10] These other externalities are partly counteracted by federal and state fuel taxes, which amount to approximately 40 cents per gallon (or 2 cents per vehicle mile).

However, the key point here is that, to a far greater extent than other externalities, congestion is highly specific to region and time of day. The marginal external congestion costs for peak travel shown in **Table 2** are large relative to the above figures for pollution and accidents, net of fuel taxes. Thus, accounting for other externalities should have a modest effect on the welfare effects of urban, peak-period tolls.

3.7. Interactions with the Broader Fiscal System

Congestion taxes can interact with distortions the broader fiscal system creates elsewhere in the economy. Most importantly, federal and state income taxes and payroll taxes combine to drive a substantial wedge between the effective gross wage firms pay (which, in a flexible market, reflects the value marginal product of labor) and the net wage households receive (which reflects the marginal value of forgone nonmarket time). Therefore, to the extent that a new policy causes an increase or decrease in labor supply, it will induce an efficiency gain or loss in the labor market. In fact, the welfare effects of even tiny changes in labor supply can be empirically important relative to those of reducing congestion, given the large labor tax wedge and the huge size of the labor market in the overall economy.

Congestion tolls can affect the labor market two ways. First, revenues may be used to lower the burden of labor taxes, producing an efficiency gain. Second, tolls levied on heavily used commuter roads reduce the returns to work effort—net of commuting costs—and may deter labor force participation at the margin. This deterrence effect, however, is partly dampened because the reduction in congestion lowers the time cost of commuting. According to Parry & Bento (2001), the net impact of a Pigouvian congestion tax, with revenues used to reduce labor taxes, is to increase labor supply, and welfare gains in the labor market are roughly the same size as those from correcting the externality.

However, the critical issue here is the importance of using congestion tax revenues in a socially productive way, either to offset reductions in distortionary taxes or, more generally, to finance public-spending projects that yield comparable efficiency gains. In fact, if revenue recycling does not lead to significant efficiency gains, the externality-correcting tax may lower overall social welfare, as the gains from correcting the externality may be outweighed by the efficiency losses in the labor market (Parry & Bento 2001). This point needs to be heeded when political feasibility may necessitate either some use of toll revenues to finance compensation schemes for motorists or earmarking of revenues for transportation enhancements.

[10] Mainstream estimates of global-warming damages, though highly contentious, are modest relative to these other externalities. Most estimates of these damages are in the order of approximately \$5 to \$25 per ton of CO_2 (e.g., Aldy et al. 2009). Burning a gallon of gasoline produces 0.0024 tons of carbon, or 0.0088 tons of CO_2, and approximately 1/23 gallons are consumed per vehicle mile driven (National Research Council 2002; Federal Highway Administration 2006, table VM-1). Therefore, a \$10 price on CO_2 amounts to approximately 0.4 cents per mile. An additional highway externality is road damage. However, it is standard to attribute this largely to heavy-duty trucks rather than light-duty vehicles. This is because pavement wear and tear is a rapidly rising function of a vehicle's axle weight.

The importance of judicious revenue use is suggested by **Figure 2**. Given that travel demand is inelastic, the amount of revenue can easily be several times larger, and conceivably an order of magnitude larger, than the Pigouvian welfare gains from correcting the externality. Using the central values for labor supply elasticities from the empirical labor economics literature, Parry & Bento (2001) noted the efficiency gain from using $1 of revenue to cut labor taxes, as opposed to financing lump-sum transfers, is $0.25. Thus, the potential gain from revenue recycling (0.25 times the amount of revenue) can easily be as large as, or perhaps much larger than, the welfare gains from congestion reduction.

3.8. Freeway Tolls Versus Cordon Tolls and Area Licenses

An alternative to pricing individual freeways is to implement a cordon toll where drivers pay as they pass points in the road network, where these points connect to form a cordon around a city or city center. Another version of this policy is the area license in which case drivers must pay even if the trip starts and terminates within the area, without crossing the border. These schemes have potential appeal for old European cities where the downtown areas are a mass of higgledy-piggledy streets that would be impractical to price individually. They have also been proposed for some U.S. cities, most notably New York City.

However, cordon-pricing and area-license schemes are inefficient in that they impose the same fee regardless of trip distance. They can also exacerbate congestion elsewhere in the road network, as people change their routes to bypass the pricing region. Despite this, well-designed cordon tolls and area licenses may still capture a large portion of the efficiency gains from more comprehensive pricing (e.g., Akiyama et al. 2002, May et al. 2004, Safirova et al. 2004, Santos 2004b, Verhoef 2002). In particular, efficiency can be improved by varying the toll with driving direction and time of day, by appropriate placement of the pricing boundary, and possibly by using multiple pricing rings.

3.9. Summary

The conceptual framework for designing congestion taxes is well developed, in that we have a reasonable grasp of the potentially important factors to consider when assessing the optimal levels, and welfare effects, of pricing schemes. The importance of network effects, bottlenecks, existing HOV lanes, etc., varies considerably across cities, however. Optimal policies therefore need to be assessed on a case-by-case basis, requiring individually calibrated models on local traffic flows, speed-flow relations, and behavioral responses to tolling.

In fact, more work is needed on empirical models for policy analysis of different urban centers. This includes developing network models that realistically capture changes in congestion throughout the entire road system. In addition, more aggregated simulation models can also play a valuable role in interpreting numerical results from computational models and in roughly gauging the empirical importance of other factors difficult to capture in a detailed network model, such as toll variation within the peak period and impacts on distorted labor markets. Research is also needed on schemes that may help overcome political opposition to congestion pricing and on the efficiency or feasibility trade-offs involved in creating broad coalitions of net beneficiaries from the policy. I now turn to these issues.

4. PRACTICAL OBSTACLES TO CONGESTION PRICING

In the past, opposition to congestion tolls in the United States from the public and elected officials has been strong. However, the development of electric metering technologies has, to some degree, addressed two of the traditional concerns: implementation difficulties and abuse of information collected on individuals' driving habits.

The administrative costs of electronic debiting from smart cards, such as E-Z passes, are minimal and vehicles may not even need to slow down as they pass transponder points. Under a global-positioning system, motorists' driving behavior is monitored by satellite and bills may be periodically mailed to households on the basis of their mileage on congested roads, again at low administrative cost.[11] Under this system, privacy is more of a concern and would need to be addressed through strict legal requirements on information-collection agencies. With electronic debiting from prepaid smart cards, privacy concerns are largely redundant because there is no need to record the vehicle's tag number.

Two especially challenging obstacles to congestion pricing remain, however. First, motorist opposition may be intense if the new charges outweigh their VOT savings. Second, congestion pricing may be unfair from the perspective of vertical equity, given that everyone faces the same tax rate regardless of income. These issues are intertwined because they both depend on the incidence of congestion tolls.

4.1. Congestion-Toll Incidence

Conceptually, leaving aside network effects, the incidence of congestion tolls is straightforward. Consider the highly simplified setting represented in **Figure 3**, where D_i denotes the demand for mileage (per unit of time) on an isolated, congested freeway by income group i, where $i = L$ (low-income) or H (high-income). $AC_i^0 = c_m + VOT_i{\cdot}T$ denotes the average cost per mile of travel for income group i that, following Equation 1, consists of the monetary cost (before any toll), plus the product of the VOT for income group i, and the time per mile T (assume $VOT_H > VOT_L$). Suppose a toll of τ per mile is introduced. The cost per mile to income group i is now

$$AC_i^0 + \tau - VOT_i{\cdot}\Delta T, \qquad (5)$$

where ΔT is the reduction in travel time per mile due to reduced congestion. The burden of the policy (or consumer surplus loss) to income group i, denoted B_i, is the shaded trapezoid in **Figure 3** and can be expressed as

$$B_i = (\tau - VOT_i{\cdot}\Delta T)M_i^0 - \frac{1}{2}(M_i^0 - M_i^1)(\tau - VOT_i{\cdot}\Delta T), \qquad (6)$$

where M_i^0 and M_i^1 denote mileage for income group i before and after the toll, respectively. The two terms in Equation 6 are the first-order burden of the toll and the second-order reduction in the burden as people reduce freeway driving to avoid the toll, respectively.[12]

[11]Although technologies are constantly improving, flaws still need to be addressed. For example, under a global-positioning system, the signal from a vehicle is sometimes lost in the presence of tall buildings or other obstructions.

[12]Note that, for the low-income group, the more elastic their demand, the smaller the burden of the toll, because the elasticity reflects more ways to avoid the toll. Thus, the popular argument that tolls are unfair because they may push low-income drivers off the priced highway is somewhat misleading. The above analysis takes drivers' VOT as given. As noted above, however, the VOT may fall as tolls reduce the severity of congestion. This enhances the possibility that some drivers may be better off prior to recycling of the toll revenues (e.g., Santos & Bhakar 2006).

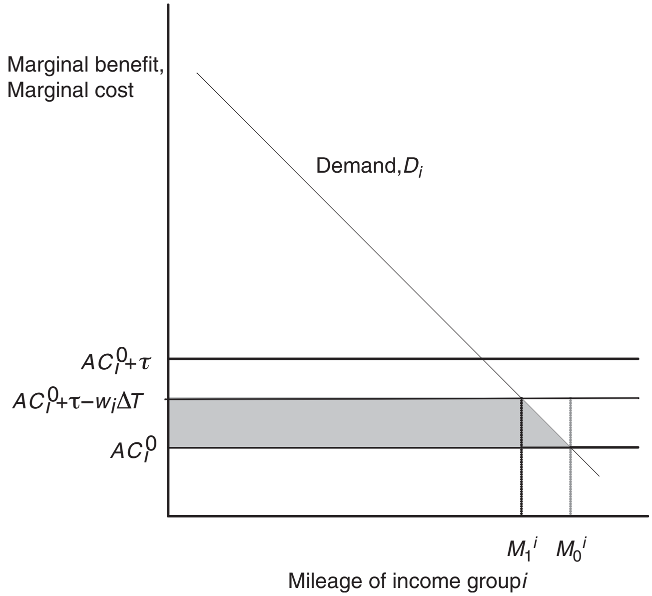

Figure 3

Incidence of congestion tolls.

Suppose that the proportionate reduction in driving is relatively small; then the burden, relative to income I_i, is approximately

$$\frac{B_i}{I_i} \approx \frac{(\tau - VOT_i \cdot \Delta T)M_i^0}{I_i}. \tag{7}$$

Conceivably, the high-income group could be better off under the toll, even before any revenue recycling, if they value the travel-time savings more than the toll payment. In contrast, the low-income group must be worse off because the toll must reduce the aggregate demand for travel on the freeway. The congestion toll is highly likely to be regressive, in that the burden-to-income ratio is greater for low-income households, or, $B_L/I_L > B_H/I_H$ (e.g., Cohen 1987, Glazer 1981, Small 1983). For the policy to be progressive, $VOT_H \cdot \Delta T < \tau$ and M_H^0/I_H would have to exceed M_L^0/I_L by a large enough margin to outweigh the smaller net-cost increase for the high-income group. That is, the income elasticity of mileage would need to be well above unity. Evidence suggests, however, that the income elasticity is around unity or below (e.g., Pickrell & Schimek 1997).[13]

In general, a more comprehensive incidence analysis should account for the distributional burden of changes in congestion elsewhere in the road network in response to the pricing of one link. The possible long-term impacts on homeowners and workers from induced changes in property values and wages should also be included. Ideally, burdens would be measured against some measure of lifetime income, rather than of annual income, because the former is a better measure of individuals' long-term consumption

[13]Note that I am defining a "regressive" and "progressive" policy in a different way from the usual use of these terms in discussions of tax policy. This is because I am defining the burden net of a time-savings benefit that is valued differently by different groups.

possibilities. Measuring against lifetime income tends to weaken the degree of regressivity (e.g., Poterba 1989). Despite these complications, congestion tolls are still likely to impose a disproportionately large burden on lower-income drivers. On the grounds of distributional equity and political feasibility, particular attention needs to be paid to compensating these groups.

4.2. Recycling Possibilities

As noted above, recycling congestion tax revenues in income tax reductions can yield relatively large welfare gains. If congestion tolls were implemented nationwide at the federal level, or (far more likely) at the local level, this could be achieved through accompanying legislation requiring automatic reductions in other taxes to keep total revenue constant. The drawback would be a large disconnect between relatively large numbers of people benefitting slightly from broad tax cuts and relatively few motorists bearing the brunt of the new toll. However, as evident from **Figure 3**, some of the toll revenue can be retained for broader tax reductions, given that compensation needed to prevent motorists being made worse off is less than the toll revenue collected.

These considerations suggest the attraction of proposals targeting some of the toll revenue at local transportation projects to help to compensate motorists, such as expanding other travel options like transit and bike paths (e.g., Small 1992). This approach need not entail significant loss of economic efficiency if these projects generate comparable welfare gains to those from cutting distortionary taxes. Another possibility for compensation is some toll rebate for low-income drivers, though this would partly undermine the overall effectiveness of the toll. More research is badly needed on the distributional and efficiency impacts of alternative packages of revenue uses for prospective congestion pricing schemes.

4.3. Policy Experience to Date

Introduced in 1975, the first attempt to use road pricing for congestion reduction was Singapore's area license (day pass). The scheme dramatically reduced congestion and raised travel speeds within the restricted zone, but congestion initially increased substantially outside of the zone, suggesting that the license price may have been excessive from a second-best perspective (Small & Gómez-Ibáñez 1998). In part, this problem was later addressed through supplementary tolls on major roads leading up to the restricted zone. Additionally, in 1998, Singapore replaced the area licensing with a toll debited electronically from smart cards on certain links, with the objective of maintaining an average speed of 30–40 mph on expressways and 12–18 mph on major roads (Santos 2005). Charges rise and fall in 30-min steps during peak periods, based on congestion levels observed in the previous quarter.

Norway experimented with cordon tolling, though with little impact on congestion because the stated objective of the policy was to raise transportation revenue rather than deter congestion (e.g., Tretvick 2003, Ramjerdi et al. 2004). In London, congestion pricing has been given a large boost following its relatively successful implementation. An area-licensing scheme was introduced in 2003; initially, it covered 8 square miles of central London and was later expanded westward to incorporate Kensington and Chelsea. The fee for entering the charging area was first set at £5 ($9) and later raised to £8. Collection is by video cameras at checkpoints into and within the priced area that record each vehicle's license plate—drivers who have not prepaid are mailed a penalty amounting to £60 or more. In the first two years, the policy reduced congestion by 30% within the priced zone, without causing excessive congestion elsewhere in the network (Transport for London 2004). This is in large part because at least half of the diverted auto trips reflected people switching to mass transit, and only approximately one-quarter were diverted to other roads in the network (Small & Verhoef 2007, p. 151). However, by 2008, average speeds had fallen back to precharging levels as a result of a high number of road works and an increase in traffic from vehicles exempt from the charge (Santos 2008, Transport for London 2008).

Congestion pricing is gaining some, albeit limited, momentum in the United States, with federal funding for pilot schemes under the Value Pricing Program and the reduction of regulatory obstacles to freeway pricing (De Corla-Souza 2004). One type of scheme is the conversion of HOV to HOT lanes, for example, on I-15 in San Diego, California (Brownstone et al. 2003). Another scheme uses tolls to fund new infrastructure, such as the lanes opened on SR-91 in Orange County, California, in 1995.

5. CONCLUSION

Congestion pricing schemes implemented to date have demonstrated their potential to improve urban travel speeds, though appropriate design features can be critical (Santos 2004a, part II). Such features include sizeable and time-varying fees as well as pricing other parts of the road network if congestion-displacement effects are important. Another lesson is the possible need for price ceilings on fees private operators charge. For example, the SR-91 toll lanes reverted to public ownership in 2003 because excessive pricing by the private operator caused unexpectedly severe congestion on parallel, unpriced lanes.

In terms of feasibility, a number of factors, besides forceful political leadership, favored the introduction of congestion pricing in London (Leape 2006, Santos & Fraser 2006). One was the high level of public and business concern about traffic jams. Before the charge, for instance, the average driving speed in Central London was less than 10 mph. Opposition to the scheme was also weakened by a range of exemptions. Taxis are exempt and residents in the charging zone pay only 10% of the fee. Public support, particularly among commuters least able to afford the charge, was garnered by requiring that toll revenues be used to improve public transport. London motorists were also more receptive to video-camera surveillance, because this had previously helped reduce street crime.

Other urban centers in the United Kingdom and United States have yet to follow London's lead, presumably because favorable factors for the introduction of radical congestion pricing schemes have not yet come to a head. These circumstances could easily change down the road, however, as urban travel speeds continue to deteriorate. In fact, at a national level in the United Kingdom, there is serious debate about replacing fuel taxes with a nationwide charge on vehicle miles that would vary across regions and time of day (Department for Transport 2004).

In short, it is an exciting time to be a transportation economist, with political and public opinion beginning, albeit perhaps only gradually, to come around to the idea of congestion pricing. The pricing schemes that eventually emerge may deviate substantially from an economist's ideal—for example, charges may vary little across time of day, many vehicles and drivers may be exempt, and some toll revenues may be dissipated in wasteful spending. But we can envision policy refinement over time—for example, variable fees may be introduced in stages, exemptions may be "bought out" over time through one-off compensation payments, and requirements for efficient revenue uses may be phased in (e.g., revenue-neutrality provisions or requirements that funded projects pass a cost-benefit assessment).

Economists have their work cut out in empirically assessing the optimal design of, and efficiency gains from, congestion pricing. At the same time, they need to better reconcile efficiency and feasibility, particularly in the design of compensation schemes that avoid large burdens on politically influential motorist groups, at minimum cost in terms of forgone economic efficiency.

DISCLOSURE STATEMENT

The author is not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENT

I am very grateful to Todd Litman, Georgina Santos, and Kenneth Small for very helpful comments on an earlier draft.

LITERATURE CITED

Akiyama T, Mun S-I, Okushima M. 2002. Second-best congestion pricing in urban space: cordon pricing and its alternatives. *Rev. Netw. Econ.* 3:401–14

Aldy J, Krupnick AJ, Newell RG, Parry IWH, Pizer WA. 2009. *Designing climate mitigation policy.* Discuss. Pap., Resour. Future

Arnott R, Kraus M. 1998. When are anonymous congestion charges consistent with marginal cost pricing? *J. Public Econ.* 67:418–42

Arnott R, de Palma A, Lindsey R. 1991. Does providing information to drivers reduce traffic congestion? *Transp. Res. A* 25:309–18

Arnott R, de Palma A, Lindsey R. 1993. A structural model of peak-period congestion: a traffic bottleneck with elastic demand. *Am. Econ. Rev.* 83:161–79

Arnott R, de Palma A, Lindsey R. 1994. The welfare effects of congestion tolls with heterogeneous commuters. *J. Transp. Econ. Policy* 18:139–61

Arnott R, Rave T, Schöb R. 2005. *Alleviating Urban Traffic Congestion*. Cambridge, MA: MIT Press

Beckmann MJ, McGuire CB, Winsten CB. 1956. *Studies in the Economics of Transportation*. New Haven, CT: Yale Univ. Press

Braid RM. 1996. Peak-load pricing of a transportation route with an unpriced substitute. *J. Urban Econ.* 40:179–97

Brownstone D, Ghosh A, Golob TF, Kazimi C, Van Amelsfort D. 2003. Drivers' willingness-to-pay to reduce travel time: evidence from the San Diego I-15 congestion pricing pilot. *Transp. Res. A* 37:373–87

Brownstone D, Small KA. 2005. Valuing time and reliability: assessing the evidence from road demonstrations. *Transp. Res. A* 39:279–93

Bureau of Labor Statistics. 2006. *National Compensation Survey: Occupational Wages in the United States, June 2005. Bulletin 2581*. Washington, DC: US Dep. Labor

Bureau of Transportation Statistics. 2006. *National Transportation Statistics*. Washington, DC: US Dep. Transp. http://www.bts.gov/publications/national_transportation_statistics/2006/index.html

Calfee J, Winston C. 1998. The value of automobile travel time: implications for congestion policy. *J. Public Econ.* 69:83–102

Cohen Y. 1987. Commuter welfare under peak period congestion: Who gains and who loses? *Int. J. Transp. Econ.* 14:239–66

de Borger B, Proost S. 2001. *Reforming Transport Pricing in the European Union*. Northampton, MA: Edward Elgar

De Corla-Souza P. 2004. Recent U.S. experience: pilot projects. In *Road Pricing: Theory and Practice*, Vol. 9, *Research in Transportation Economics*, ed. G Santos, pp. 179–206. San Diego, CA: Elsevier

de Palma A, Marchal F. 2001. *Real-cases applications of the fully dynamic METROPOLIS tool-box: an advocacy for large-scale mesoscopic transportation systems*. Work. Pap. 2001-18, THEMA, Univ. de Cergy-Pontoise. http://www.u-cergy.fr/IMG/2001-18DePalma.pdf

Department for Transport. 2004. *Feasibility Study of Road Pricing in the UK—Report*. London, UK: Dep. Transp.

Department for Transport. 2007. *Transport Analysis Guidance: Values of Time and Operating Costs*. London, UK: TAG Unit 3.5.6. http://www.webtag.org.uk/webdocuments/3_Expert/5_Economy_Objective/3.5.6.htm.

Downs A. 1992. *Stuck in Traffic: Coping with Peak-Hour Traffic Congestion*. Washington, DC: Brookings Inst. Press

Federal Highway Administration. 1997. *1997 Federal Highway Cost Allocation Study*. Washington, DC: Dep. Transp.

Federal Highway Administration. 2000. *Addendum to the 1997 Federal Highway Cost Allocation Study Final Report*. Washington, DC: Dep. Transp.

Federal Highway Administration. 2006. *Highway Statistics 2005*. Washington, DC: Dep. Transp.

Glazer A. 1981. Congestion tolls and consumer welfare. *Public Financ.* 36:77–83

Gómez-Ibáñez JA, Meyer JR. 1993. *Going Private: The International Experience with Transport Privatization*. Washington, DC: Brookings Inst.

Goodwin PB. 1992. A review of new demand elasticities with special reference to short- and long-run effects of price changes. *J. Transp. Econ. Policy* 26:155–69

Goodwin PB. 1996. Empirical evidence on induced traffic: a review and synthesis. *Transportation* 23:35–54

Goodwin PB, Dargay J, Hanly M. 2004. Elasticities of road traffic and fuel consumption with respect to price and income: a review. *Transp. Rev.* 24:275–92

Greenwood ID, Bennett CR. 1996. The effects of traffic congestion on fuel consumption. *Road Transp. Res.* 5:18–31

Hau T. 2005a. Economic fundamentals of road pricing: a diagrammatic analysis, part I—fundamentals. *Transportmetrica* 1:81–115

Hau T. 2005b. Economic fundamentals of road pricing: a diagrammatic analysis, part II—relation of assumptions. *Transportmetrica* 1:119–49

Homburger WS, Kell JH, Perkins DD. 1992. *Fundamentals of Traffic Engineering*, Berkeley, CA: Inst. Transp. Stud.

Houde S, Safirova EA, Harrington W. 2007. *Washington START transportation model*. Discuss. Pap. 07–43, Resourc. Future

Leape J. 2006. The London congestion charge. *J. Econ. Perspect.* 20:157–76

Lindsey R. 2006. Do economists reach a conclusion on road pricing? The intellectual history of an idea. *Econ. J. Watch* 3(2):292–379

Lindsey R. 2007. *Congestion relief: assessing the case for road tolls in Canada*. Issue 248, CD Howe Inst. Comment.

Lindsey R, Verhoef ET. 2000. Congestion modeling. In *Handbook of Transport Modeling*, ed. KJ Button, DA Hensher, pp. 353–73. Amsterdam: Pergamon

Litman T. 2006. *Smart Congestion Reductions: Reevaluating the Role of Highway Expansion for Improving Urban Transportation*. Victoria, BC: Vic. Transp. Policy Inst. http://www.vtpi.org/cong_relief.pdf

Litman T. 2007a. *Smart Transportation Investments II: Reevaluating the Role of Public Transit for Improving Urban Transportation*. Victoria, BC: Vic. Transp. Policy Inst. http://www.vtpi.org/cong_reliefII.pdf

Litman T. 2007b. *Transportation Cost and Benefit Analysis—Congestion Costs*. Victoria, BC: Vic. Transp. Policy Inst. http://www.vtpi.org

Litman T. 2008. *TDM Encyclopedia*. Victoria, BC: Vic. Transp. Policy Inst. http://www.vtpi.org/tdm/index.php

Liu LN, MacDonald JF. 1998. Efficient congestion tolls in the presence of unpriced congestion: a peak and off-peak simulation model. *J. Urban Econ.* 44:352–66

MacDonald J. 1995. Urban highway congestion: an analysis of second-best tolls. *Transportation* 22:353–69

Mackie P. 1996. Induced traffic and economic appraisal. *Transportation* 23:103–19

Mackie PJ, Wardman M, Fowkes AS, Whelan G, Nellthorp J, Bates J. 2003. *Value of Travel Time Savings in the UK: Summary Report*. Leeds, UK: Inst. Transp. Stud./Univ. Leeds

May AD, Liu R, Shepherd SP, Sumalee A. 2002. The impact of cordon design on the performance of road pricing schemes. *Transp. Policy* 9:209–20

May AD, Milne DS. 2000. Effects of alternative road pricing systems on network performance. *Transp. Res. A* 34:407–36

May AD, Shepherd S, Sumalee A. 2004. Optimal locations and charges for cordon schemes. See Santos 2004a, pp. 61–86

Miller TR, Levy DT, Spicer RS, Letina DC. 1998. Allocating the costs of motor vehicle crashes between vehicle types. *Transp. Res. Rec.* 1635:81–7

Mohring H. 1999. Congestion. In *Essays in Transportation Economics and Policy: A Handbook in Honor of John R. Meyer*, ed. JA Gómez-Ibáñez, WB Tye, C Winston, pp. 181–222. Washington, DC: Brookings Inst.

Mun S-I. 1994. Traffic jams and the congestion toll. *Transp. Res. B* 28:365–75

MVA Consultancy, ITS Univ. Leeds, TSU Univ. Oxford. 1987. *Value of Travel Time Savings*. Newbury, Berks.: Policy J.

National Research Council. 2002. *Effectiveness and Impact of Corporate Average Fuel Economy (CAFE) Standards*. Washington, DC: Natl. Acad. Press

Noland R. 2001. Relationships between highway capacity and induced vehicle travel. *Transp. Res. A* 35:47–72

National Surface Transportation Policy. 2007. *Transportation for Tomorrow: Report of the National Surface Transportation Policy and Revenue Study Commission*. Washington, DC: Natl. Surf. Transp. Policy Revenue Study Comm. http://www.transportationfortomorrow.org/final_report

Parry IWH. 2002. Comparing the efficiency of alternative policies for reducing traffic congestion. *J. Public Econ.* 85:333–62

Parry IWH. 2004. Comparing alternative policies to reduce traffic accidents. *J. Urban Econ.* 56:346–68

Parry IWH, Bento AM. 2001. Revenue recycling and the welfare effects of road pricing. *Scand. J. Econ.* 103:645–71

Parry IWH, Small KA. 2009. Should urban transit subsidies be reduced? *Am. Econ. Rev.* Forthcoming

Pickrell D, Schimek P. 1997. Trends in personal motor vehicle ownership and use: evidence from the nationwide personal transportation survey. *Proc. Nationwide Pers. Transp. Surv. Symp. Search. Solut. Policy Discuss. Ser. No. 17*, pp. 85–127. Washington, DC: US Fed. Highw. Admin.

Poterba JM. 1989. Lifetime incidence and the distributional burden of excise taxes. *Am. Econ. Rev. Pap. Proc.* 79:325–30

Pratt RH, Texas Transp. Inst., Cambr. Syst., Parsons Brinkerhoff Quade Douglas, SG Assoc., et al. 2000. *Traveler Response to Transportation System Changes: Interim Handbook. TCR Web Doc. 12*. Washington, DC: Natl. Res. Counc. Transp. Res. Board

Ramjerdi F, Minken H, Østmoe K. 2004. Norwegian urban tolls. See Santos 2004a, pp. 237–49

Safirova E, Gillingham K. 2003. *Measuring marginal congestion costs of urban transportation: Do networks matter*? Discuss. Pap. 03–56, Resourc. Future

Safirova EA, Gillingham K, Parry I, Nelson P, Harrington W, Mason D. 2004. Welfare and distributional effects of HOT lanes and other road pricing policies in metropolitan Washington, DC. See Santos 2004a, pp. 179–206

Santos G, ed. 2004a. *Road Pricing: Theory and Evidence,* Vol. 9*, Research in Transportation Economics*. San Diego, CA: Elsevier

Santos G. 2004b. Urban congestion charging: a second-best alternative. *J. Transp. Econ. Policy* 38:345–69

Santos G. 2005. Urban congestion charging: a comparison between London and Singapore. *Transp. Rev.* 25:511–34

Santos G. 2008. The London congestion charging scheme. In *Brookings Wharton Papers on Urban Affairs: 2008*, ed. G Burtless, J Rothenberg Pack, pp. 177–234. Washington, DC: Brookings Inst.

Santos G, Bhakar J. 2006. The impact of the London congestion charging scheme on the generalised cost of car commuters to the city of London. *Transp. Policy* 13:22–33

Santos G, Fraser G. 2006. Road pricing: lessons from London. *Econ. Policy* 21:264–310

Santos G, Newbery D. 2002. *Estimating urban congestion charges*. Discuss. Pap. 3176, CEPR

Schrank D, Lomax T. 2007. *The 2007 Urban Mobility Report*. College Station, TX: Texas Transp. Inst./Texas A&M Univ.

Shoup DC. 2005. *The High Cost of Free Parking*. Chicago, IL: Am. Plan. Assoc. Plan. Press

Small KA. 1983. The incidence of congestion tolls on urban highways. *J. Urban Econ.* 13:90–111

Small KA. 1992. Using the revenues from congestion pricing. *Transportation* 19:359–81

Small KA, Chu X. 2003. Hypercongestion. *J. Transp. Econ. Policy* 37:319–52

Small KA, Gómez-Ibáñez JA. 1998. Road pricing for congestion management: the transition from theory to policy. In *Road Pricing, Traffic Congestion and the Environment: Issue of Efficiency and Social Feasibility*, ed. KJ Button, ET Verhoef, pp. 213–46. Cheltenham, UK: Edward Elgar

Small KA, Gómez-Ibáñez JA. 1999. Urban transportation. In *Handbook of Regional and Urban Economics*, Vol. 3, ed. P Cheshire, ES Mills, pp. 1937–99. Amsterdam: North-Holland

Small KA, Verhoef E. 2007. *The Economics of Urban Transportation*. New York: Routledge

Small KA, Yan J. 2001. The value of "value pricing" roads: second-best pricing and product differentiation. *J. Urban Econ.* 49:310–36

Standing Advisory Committee on Trunk Road Assessment. 1994. *Trunk Roads and the Generation of Traffic. Department of Transport*. London: HMSO

Steer Davies Gleave. 2004. The effect of road congestion on rail demand. *Rep. Passeng. Demand Forecast. Counc.*, Assoc. Train Oper. Co., London, UK

Transport for London. 2004. *Congestion Charging Impacts Monitoring: Second Annual Report*. London, UK: Transp. Lond.

Transport for London. 2008. *Central London Congestion Charging: Impacts Monitoring. Sixth Annual Report*. London, UK: Transp. Lond. www.tfl.gov.uk/roadusers/congestioncharging/6722.aspx

Transportation Research Board. 2006. *The Fuel Tax and Alternatives for Transportation Funding*. Washington, DC: Natl. Res. Counc.

Tretvick T. 2003. Urban road pricing in Norway: public acceptability and travel behaviour. In *Acceptability of Transport Pricing Strategies*, ed. J Schade, B Schlag, pp. 77–92. London, UK: Pergamon

US Department of Transportation. 1997. *The Value of Travel Time: Departmental Guidance for Conducting Economic Evaluations*. Washington, DC: Dep. Transp.

US Department of Transportation. 2006. *National Strategy to Reduce Congestion on America's Transportation Network*. Washington, DC: Dep. Transp.

US Department of Transportation. 2008. *Transportation Vision for 2030: Ensuring Personal Freedom and Economic Vitality for a Nation on the Move*. Washington, DC: Dep. Transp. Res. Innovative Technol. Admin.

Verhoef ET. 2002. Second-best congestion charging in general networks. Heuristic algorithms of finding second-best optimal toll levels and toll point. *Transp. Res. B* 36:707–29

Verhoef ET, Nijkamp P, Rietveld P. 1996. Second-best congestion pricing: the case of an untolled alternative. *J. Urban Econ.* 40:279–302

Verhoef ET, Small KA. 2004. Product differentiation on roads: constrained congestion pricing with heterogeneous users. *J. Transp. Econ. Policy* 38:127–56

Vickrey W. 1963. Pricing in urban and suburban transport. *Am. Econ. Rev.* 53:452–65

Vickrey W. 1969. Congestion theory and transport investment. *Am. Econ. Rev. Pap. Proc.* 59:251–61

Walters A. 1961. The theory and measurement of private and social costs of highway congestion. *Econometrica* 29:676–97

Wardman M. 2001. A review of British evidence on time and service quality valuations. *Transp. Res. E* 37:107–28

Waters WGII. 1996. Values of time savings in road transport project evaluation. In *World Transport Research: Proceedings of 7th World Conference on Transport Research*, Vol. 3., ed. D Hensher, J King, T Oum, pp. 213–33. Oxford, UK: Pergamon

Yang H, Meng Q, Lee D-H. 2004. Trial-and-error implementation of marginal-cost pricing on networks in the absence of demand functions. *Transp. Res. B* 38:477–93

Annual Review of
Resource Economics

Volume 1, 2009

Contents

Prefatory Article

An Amateur Among Professionals
Robert M. Solow 1

Policy Analysis and Design

Agriculture for Development: Toward a New Paradigm
Derek Byerlee, Alain de Janvry, and Elisabeth Sadoulet 15

Governance Structures and Resource Policy Reform: Insights from Agricultural Transition
Johan F.M. Swinnen and Scott Rozelle 33

Distortions to Agricultural Versus Nonagricultural Producer Incentives
Kym Anderson 55

Public-Private Partnerships: Goods and the Structure of Contracts
Gordon Rausser and Reid Stevens 75

Environmental Regulations and Economic Activity: Influence on Market Structure
Daniel L. Millimet, Santanu Roy, and Aditi Sengupta 99

The Development of New Catastrophe Risk Markets
Howard C. Kunreuther and Erwann O. Michel–Kerjan 119

The Curse of Natural Resources
Katharina Wick and Erwin Bulte 139

Experiments in Environment and Development
Juan Camilo Cárdenas 157

Behavior, Environment, and Health in Developing Countries: Evaluation and Valuation
Subhrendu K. Pattanayak and Alexander Pfaff 183

Resource Dynamics

Irreversibility in Economics
Charles Perrings and William Brock 219

Whither Hotelling: Tests of the Theory of Exhaustible Resources
Margaret E. Slade and Henry Thille 239

Recent Developments in the Intertemporal Modeling of Uncertainty
Christian P. Traeger 261

Rent Taxation for Nonrenewable Resources
Diderik Lund 287

Land Use and Climate Change Interactions
Robert Mendelsohn and Ariel Dinar 309

Urban Growth and Climate Change
Matthew E. Kahn 333

Reduced-Form Versus Structural Modeling in Environmental and Resource Economics
Christopher Timmins and Wolfram Schlenker 351

Ecology and Space

Integrated Ecological-Economic Models
John Tschirhart 381

Integrating Ecology and Economics in the Study of Ecosystem Services: Some Lessons Learned
Stephen Polasky and Kathleen Segerson 409

The Economics of Urban-Rural Space
Elena G. Irwin, Kathleen P. Bell, Nancy E. Bockstael, David A. Newburn, Mark D. Partridge, and JunJie Wu 435

Pricing Urban Congestion
Ian W.H. Parry 461

The Economics of Endangered Species
Robert Innes and George Frisvold 485

On the Economics of Water Allocation and Pricing
Yacov Tsur 513

Technology and Innovation

The Economics of Agricultural R&D
Julian M. Alston, Philip G. Pardey, Jennifer S. James, and Matthew A. Anderson 537

Supply and Demand of Electricity in the Developing World
Madhu Khanna and Narasimha D. Rao 567

Energy Efficiency Economics and Policy
Kenneth Gillingham, Richard G. Newell, and Karen Palmer 597

Recent Developments in Renewable Technologies: R&D Investment in Advanced Biofuels
Deepak Rajagopal, Steve Sexton, Gal Hochman, and David Zilberman . . . 621

Fuel Versus Food
Ujjayant Chakravorty, Marie-Hélène Hubert, and Linda Nøstbakken . . . 645

The Economics of Genetically Modified Crops
Matin Qaim 665

Errata

An online log of corrections to *Annual Review of Resource Economics* articles may be found at http://resource.AnnualReviews.org