The efficacy of congestion pricing

Zachary J. Ridder
University of Tennessee at Chattanooga, pfc665@mocs.utc.edu

Follow this and additional works at: https://scholar.utc.edu/honors-theses

Part of the Economics Commons

Recommended Citation
The Efficacy of Congestion Pricing

By Zack Ridder

Departmental Honors Thesis
The University of Tennessee at Chattanooga
Economics Department

Project Direction: Dr. Leila Pratt

Examination Date: 3/30/16

Examination Committee:
   Dr. Leila Pratt
   Dr. Bruce Hutchinson
   Dr. James Guilfoyle
   Dr. David Giles
The Efficacy of Congestion Pricing

By

Zack Ridder
I. Introduction

As the number of vehicles continues to climb road congestion will become a more prominent feature of urban centers and will likely worsen over time. This thesis will show that the most simple and effective way to combat congestion is the implementation of road pricing schemes in and around major metropolitan areas. For the most part the benefits of such a system outweigh the potential costs. These tolling systems force commuters to internalize the marginal social cost of driving, as opposed to only paying the average social cost and incurring a deadweight loss.

This paper will first provide a review of the literature, tracing the theoretical development of congestion pricing from its early roots to modern studies. Following this review a section on economic principles that are integral to the study of road pricing, specifically the concept of marginal cost pricing and how to reach the efficient quantity of traffic will be discussed. This will be followed by a study of several cities where congestion pricing tried. The last section of this paper makes the case for using road pricing in major cities.

The economic foundations of road or congestion pricing were first put forward by William S. Vickrey, Alan A. Walters, and Carl S. Shoup in the 1960’s. There was, they argued, an imbalance between what a motorist was willing to pay and what they were made to pay for travelling by road. This was a result of roads being paid for with flat taxes rather than fees for travel. As a result roads seemed free to the average commuter, and thus there was a tendency to over consume. This overconsumption led to congestion, gridlock, and traffic accidents. These conditions all caused a loss in time to commuters, which was increased with every additional car added to the overburdened roadway.
The answer to this problem was theoretically simple, though politically onerous: place a charge on motorists equal to the marginal social cost their driving would impose on others. In doing so the imbalance and overuse would correct itself as drivers diverted to less well-travelled routes or changed the times they drove to avoid the new explicit cost. If the demand for roadways was inelastic and the supply limited, this system would result in a more even distribution of traffic and a reduction in congestion. The only problem that remained was the practical difficulty of ensuring payment of the toll, without causing slowdowns at tolling stations which would make the problem of gridlock even worse. Technology provided the answer to this issue and now tolling can be done at any speed. Singapore’s ERP lanes are the physical proof that variable rate road pricing not only has come of age, but that it can be used as an effective policy to curb congestion in and around urban areas.
II. Theory and Literature Review

The starting point for any study of congestion pricing is the works of William S. Vickrey and A.A. Walters, both economists, who first wrote about the concept during the 1950’s and 60’s. Facing urban growth following the Second World War, and a massive increase in the number of cars on the roads, both Vickrey and Walters identified congestion as a problem for commuters and government alike. Vickrey first wrote about the problem in regards to the New York subway system in 1952, postulated the idea of increasing the fares during peak times would ease congestion.\(^1\) He argued that such a price increase would not actually reduce the traffic flow, but would increase it by spreading traffic out over a longer period of time and distributing it more efficiently.\(^2\)

His basis was built on the cost in time to commuters as well as established economic principles for demand models. By imposing a higher cost to travel during the peak times, and a lower cost during off-peak hours, he showed that an average family would be made better off monetarily if they shifted their travel habits only slightly to take advantage of the lower cost during off peak times.\(^3\) Based on his experience with railroad and subway systems he suggested, that to avoid the bottleneck that might occur by collecting such tolls, that every person wishing to travel be issued a card that could be verified by a machine. They would then be billed based on the time and destination of their journey.\(^4\)


\(^2\) Ibid.


\(^4\) Ibid.
Congestion can, generally, be divided into six categories, each of which has its own set of causes, effects, and solutions. The first is single interaction, where two units of transportation, cars in this case, are in such proximity that at least one must slow down in order to avoid an accident or collision. Normally, single interaction congestion is a problem in areas of light traffic and the delay to each participant is roughly the same as the delay they are causing to another person.\(^5\) Multiple interaction simply refers to a higher density of traffic in which one might expect even longer delays.

Bottlenecking occurs when the capacity, or in this case the number of available lanes or road space, is insufficient to meet the demand of the users. This can vary with time, so an area that at 10pm might be completely congestion free might be subject to a bottleneck at 7:30am. Bottlenecks develop naturally around areas of limited space that are preceded by higher capacity, a two lane tunnel connected to a four lane highway as an example. A triggerneck is caused by a bottleneck when traffic that has no intention of using the area afflicted by a bottleneck is delayed by the congestion resulting from the bottleneck. Traffic that is backed up and preventing drivers from taking their desired exit from taking or merging is a good illustration of this phenomenon.\(^6\)

Network and control congestion is normally a municipal problem as it deals with a buildup of congestion that necessitates the use of some form of regulation to control or mitigate it. Traffic lights, stop signs, or routing limitations are the hallmark of this form of congestion. Finally, there is general density congestion which is based on the density of transportation as whole, regardless of what form it takes. General density also covers the argument for constructing more roads or lines of service. Even if the different types of transportation do not


\(^6\) Ibid, 452.
directly affect each other building more infrastructure for either will, due to the oversaturation and finite space, become increasingly expensive and time consuming.\textsuperscript{7}

Each of these forms of congestion can occur, either independently or in concert, in either a city or a highway situation. While road pricing is a potential solution for any of these it should also be noted that other methods, such as gas taxes, import duties on cars, subsidized rail or bus travel, or emissions controls on vehicles can also be employed, with varying degrees of success. The difference in approach between congestion pricing and a gasoline tax or an import duty is an important one, however, as the consumer has the ability to mitigate their costs by spreading out or taking alternative modes of transportation to minimize their marginal cost. Commuting will still take place, but with the addition of the toll the commuter faces more explicit choices on how much they will drive.

At the same time as suggesting potential solutions to congestion Vickrey also took aim at the root cause of the problem: The idea that transportation infrastructure should be free to use by motorists. This idea, which he lambasted as “perverse”, had its roots not in economics, but in political expediency.\textsuperscript{8} Vickrey, using the example of the New York bridges and tunnels, it made no sense to impose a toll on the newest bridges to assist in paying for them, but not levy at least an equal toll on the older bridges for their upkeep. The inevitable result, he argued, would be that any rational commuter would, naturally, take the free bridge over the toll bridge, valuing money more than their time and causing massive congestion around the free bridges.\textsuperscript{9}

\textsuperscript{7} Ibid, 452-453.

\textsuperscript{8} Ibid, 455.

\textsuperscript{9} Ibid.
The proximate problem of congestion, a lack of available routes and an overabundance of travelers, did, however, have a simple answer: constructing more and more infrastructure to compensate. Vickrey tackled this issue too, arguing that the investment in new infrastructure, based on the idea that it would, at some level, be profitable, usually resulted in an inaccurate charging scheme that did not represent the marginal cost to the consumer.10 While as discussed above there are at least six types of congestion that exist, but, with reference to the idea of expanding infrastructure Vickrey believed the bottleneck situation could be addressed.

A bottleneck situation exists where there is some sort of fixed capacity and, at a specific point, that capacity is less than the demand. As long as the demand remains lower than capacity there will be few or no delays, but as demand increases, so too do the wait times. Bridges, tunnels, lanes of traffic, and power lines are all subject to bottlenecks.11 Without tolls, there will be spikes of congestion during peak times, as the capacity of the bridge or traffic lanes is static. To expand the available capacity is an explicit expense, one that takes time and further disrupts traffic cycles and causes the very bottlenecking that the construction of new capacity is intended to prevent.12 As a result, Vickrey argued, it was more sensible to simply impose a toll, which represented an additional source of revenue for the government, and allow individuals to make their own decisions. Some might pay the toll, others would seek alternative routes, and even more might avail themselves of public transportation or share rides.13

---

11 Ibid, 252.
12 Ibid, 253-254.
13 Ibid, 258-259.
A traffic system can only handle so many cars before it reaches its maximum capacity and, while a government can build more infrastructure to attempt to ease the problem this has been shown to be less than cost effective. Using Washington, DC as an example in 1963 Vickrey estimated that for every $3000 dollar car added to traffic during peak times a $23,000 investment had to be made to offset the additional vehicle.\textsuperscript{14} Marginal cost pricing is charging a price equal to the extra expense involved in producing an extra unit of output. Commuters, through taxes, pay the average cost of driving on a road, but not the actual marginal cost which lowered congestion taken into account. To correctly apply road pricing the price of the toll must be set to equal to the marginal social cost to society that the driver's journey imposes on other drivers.\textsuperscript{15}

With systems like London this is done with a day by day charge, whereas in Singapore this is done by charging by the number of trips through tolling gantries, but each are trying, within limits, to adhere to this basic idea. These charges should, ideally, be small but fairly frequent and they should change to allow for different times or roads, incentivizing uses at off hours or taking alternate routes.\textsuperscript{16} Without some form of road pricing inefficiency develops and society as a whole suffers costs due to increased congestion, pollution, and even accidents.

This concept illustrated in the standard model of congestion in Figure 1\textsuperscript{17} with a uniform road, normal with and without junctions, All commuters are assumed to be the same except for their marginal willingness to pay, which is given by demand curve $D$. The marginal private and marginal social benefits (MPB and MSB) are set equal to each other as well as to $D$. The

\begin{itemize}
  \item \textsuperscript{14} Vickrey, “Pricing in Urban Transport,” \textit{The American Economic Review} Vol. 53 No. 2 (1963): 456
  \item \textsuperscript{15} “Principles of Efficient Congestion Pricing,” accessed March 18, 2016, \url{http://www.vtpi.org/vickrey.htm}
  \item \textsuperscript{16} Martin G. Richards \textit{Congestion Charging in London}, 28.
  \item \textsuperscript{17} Georgina Santos, “Urban Congestion Charging: A Comparison between London and Singapore,” \textit{Transport Reviews} 25: 512
\end{itemize}
efficient equilibrium is found at point $C$, where the marginal social cost and the demand curve, which is also the MSB, intersect. However, the actual equilibrium is found at point $B$ where average social cost (ASC) is equal to the marginal willingness to pay. This means that drivers are only paying the ASC, and as a result an externality develops in the form of a marginal congestion cost, represented by the distance between points $A$ and $B$. This results in the area of $ABC$ being a dead-weight loss.\textsuperscript{18}

![Diagram showing average and marginal costs, dead-weight loss, and efficient charge](image)

**Figure 1.** Average and marginal costs, dead-weight loss, efficient charge

Imposing a toll moves the consumer back along the demand curve towards $C$, which corresponds to the efficient quantity of traffic on the road at a given time. To be effective the toll

\textsuperscript{18} Ibid.
would have to equal line segment $CE$, which would force motorists to completely internalize the cost of their decision to drive. As a result of the toll the flow of traffic would become optimized at point $C$, while the rectangle $CEGF$ would represent revenue gained by collection of the toll.\(^{19}\) These proceeds could be used to cover the cost of implementing and maintaining a system capable of imposing this toll, and to subsidize public transportation, which would further lower road congestion.

A. A. Walters, a contemporary of Vickrey, applied similar ideas to the highway system during the 1960s. Arguing against other economists that there should be a decrease in taxes he suggested that the reverse would actually be more effective. Pointing to examples in the rail industry and the airline industry, he advanced the idea that marginal cost pricing had been misapplied by previous writers who had focused only on the decrease in demand that would result if a tax was leveled on gasoline or more tolls were levied. To counterbalance the loss of demand on the roads there would be, Walters predicted, an upswing in demand for the closest competitor, in this case the railroads.\(^{20}\)

Like Vickrey, Walter's plans for how to implement either the tax increase or the road tolls looked to technology. Cars would be equipped with a mileometer, which would indicate that a toll was being leveled by a raised flag much like a taxi, or else purchase a daily sticker indicating the payment of the toll. Those utilizing stickers would be charged more per unit than those utilizing the milometer to both encourage a shift to the milometer and also give a slight discount to those who made daily trips into and out of the controlled zone.\(^{21}\) To ensure compliance he

\(^{19}\) Ibid.


\(^{21}\) Ibid.
suggested the use of cameras to record anyone driving without having the toll.\textsuperscript{22} In particular he singled out large cities, London and New York specifically, as being ideal places to impose tolls due to the limited means of entrance or exit, which would make actually implementing road pricing much simpler.\textsuperscript{23} These limitations on entrance indicate another factor of road pricing, the inelasticity of demand of commuters for roads.

There are two types of elasticity—inelastic, where the demand or supply of a product does not rise or fall at a corresponding rate relative to changes in price, and elastic, where the supply or demand is sensitive to changes in prices. The demand for road lanes is inelastic, as individuals must drive to and from their jobs daily. This inelasticity of demand is most clearly seen in the peaks where congestion is generally at its worst: rush hour traffic as commuters try to reach their jobs in the morning and return from them in the evening. With an individual's livelihood relying on them reaching their place of employment motorists are faced with an inelastic demand for routes to their destinations. During such times the more direct or desirable routes are overburdened and quickly reach their capacity. Congestion pricing addresses this inelasticity by imposing a toll that encourages people to stagger their arrival and departure times, take a different route, or switching to a different form of transport. These changes shift the demand away from a specific route and on to less used roads or public transport, spreading the capacity out more uniformly.\textsuperscript{24}

This inelasticity of demand can vary based upon times, as off peak travellers are more sensitive to changes in price than those travelling at peak times.\textsuperscript{25} In the case of Milan an

\textsuperscript{22} Ibid.
\textsuperscript{23} Ibid.
\textsuperscript{25} Ibid, 452-453.
increase of 1 percent in the toll is estimated to have caused a 0.3 percent decrease in trips into the controlled zone.\textsuperscript{26} These decreases are similar to the results found in London and Singapore, and usually come with a corresponding increase in the use of public transportation as the drivers take the lower cost option and switch from one form of transport to another in response to the toll.

In a different paper for the World Bank in published 1968, Walters directly discussed the idea of roads as public goods and how this was misleading. The salient features of a public good are that it is non-exclusionary and non-rivalrous. In other words a public good cannot be easily denied and the use of it does not detract from anyone else utilizing it. In the case of roads the addition of a vehicle can lower the enjoyment of another person driving on it. On small private roads this might not be the case, but with regard to heavily travelled routes it most certainly is and, as a result, these roads do not function as pure public goods.\textsuperscript{27}

The massive expense of public roads makes it difficult to allocate to any one individual the cost of expanding the road or building the bridge, as a result the cost must be divided over the population as a whole. This, however, is inequitable as there will be people charged for a benefit that they do not utilize. Further, while infrastructure represents a large upfront cost and then periodic upkeep, road pricing is a source of revenue. This income can itself be allocated to help maintain the existing infrastructure, build new lanes for travel, or supplement public transit. This leads us to the question of equity, one of the main arguments against road pricing that has been cited in virtually every attempt to either implement it or to repeal it. The potential for inequality does indeed exist within any road pricing scheme and can be broadly lumped into


three categories: who pays, who benefits, and how are the externalities distributed? When considering who is paying the toll most road pricing is done by time and location, thus commuters entering and exiting the Central Business District during peak hours in the morning and afternoon are the users most likely charged

The second category of equity effect revolves around, who benefits. No matter how congestion pricing is set up, to be effective it has to be capable of charging everyone a price that reflects the marginal social cost of a trip. By imposing the cost consumers are shifted to different times and routes, lowering congestion and increasing travel times for all motorists. As each commuter gains in time by not spending it in traffic, the monetary charge is, according to Vickrey, less than the potential gain.

Another benefit of congestion pricing is reduced pollution. Pollution is usually measured in terms of nitrogen oxides (NOx), various kinds of particulate matter 10 micrometers in size or less (PM10), and carbon dioxide (CO2). The first two are associated with negative effects on health and the last is a contributing factor in global warming. With these health risks mitigated or reduced there is a benefit to society both in terms of a reduction of healthcare costs for related illnesses and also as an increase in life expectancy per capita. All congestion pricing schemes have some associated reduction in pollution, even if that is not their avowed goal, because


30 Ibid.


number of cars driving at peak times is diminished either by drivers taking different routes, which lower the concentration of vehicle emissions at congested points, or by taking lower emitting public transportation.

In the case of Milan the benefit in welfare has been estimated at 3 billion dollars. A referendum to continue the congestion pricing scheme resulted in 80 percent voting in favor of the Ecopass system.33 This vote, which flies in the face of the normally negative views the public has for road pricing schemes, indicates that the residents, both inside and outside of the controlled zone, felt that they received some benefit from the policy.34 This added benefit is also present in London, which saw an increase in life expectancy for both the population within the charging zone and those outside of it.

From the outset, roads are undervalued and underpriced. The benefits that society receives from roads is greater than the cost to the average commuter for building or maintaining them.35 But how are these benefits measured and who do they affect? There are several methods for valuing the social benefits of congestion pricing, and taken as a whole they go a long way to allaying some of the arguments against such a policy. One of the earliest and easiest is the increase in speed for traffic within a congested area, as an increase in speed equates to a decrease in the time a commuter takes to get to their destination.36 Another deceptively simple model for estimating the benefit to society is in the reduction of congestion, measured by the number of

33 Ibid.
34 Ibid.
cars on the road at peak and off peak times. If the premise that any congestion represents a cost to society, any reduction of it would therefore constitute an increase in social benefits.
III. Case Studies

Introduction

While tolls, bridge fees, and wheel taxes have been a part of transportation history for many hundreds of years the modern history of congestion pricing as a policy only truly began in the latter half of the twentieth century. The increasing number of automobiles on the road created two new dilemmas for regulators: traffic congestion and pollution. The idea of road pricing was a cheap and easy way to control both of these factors and a number of cities in the last sixty years have implemented it, with varying levels of success. This section will examine several examples of congestion pricing in cities around the world.

Singapore

In response to rising urban congestion Singapore enacted the Area License Scheme (ALS) in 1975 and updated it to an electronic system, the Electronic Road Pricing Scheme (ERP) in 1998.\textsuperscript{37} The city-state of Singapore is unusually well suited for congestion pricing as it is located on an island at the tip of the Malay Peninsula. The congestion pricing measures were intended to supplement an earlier attempt in 1972 to limit the number of vehicles by imposing an import duty on vehicles equal to 45 percent of their market value and a registration fee of 25 percent (in 1995 this fee was increased to 55 percent).\textsuperscript{38} All of these schemes have common roots in two major studies, one in 1967 and another in 1974. Both studies concluded that the only way


to lower congestion and also accommodate Singapore’s growth would be to limit car ownership and usage.  

As outlined in section two, one necessary criterion for an effective implementation of congestion pricing is the ability to limit access to and from areas. In Singapore, this is relatively easy to do because access to the Central Business District (CBD) was initially going through twenty-two specific entrance points (though over time this number increased to thirty-three).

It is worth noting that Singapore was a city of over four million people when this policy was under consideration and that an application of congestion on this scale had not, to this point, been attempted.

Although limiting the number of private vehicles entering the CBD, the government of Singapore wished to exempt vehicles that affected business, public transport such as buses, and low profile conveyances like motorcycles. In addition, public service and military vehicles were also exempted. The ALS utilized physical licenses that were purchased and displayed on vehicles entering the Restricted Zone (RZ). The policy was very successful. Authorities set a congestion price of S$3 a day or S$60 a month. Many thought that this policy would not work because it would be unenforceable during the morning commute. However, the policy resulted in

---


a staggering 43 percent decrease in the total number of motor vehicles travelling within the Restricted Zone (RZ) within only four weeks.\textsuperscript{44}

This system lasted for almost thirty-three years with only minor changes, such as extending the time that the congestion charge was in effect in the mornings by an hour. In 1995 the Singaporean government committed almost S$200 million to spur the adoption of an electronic system that would update the ALS.\textsuperscript{45} By September of 1998 the ERP was installed. This system is based on radio frequency identification, optical detection and has taken advantage of various advances in both imaging and smart card technology. Each car was fitted with a transponder In-vehicle Unit (IU) at a cost of less than S$300 per vehicle, and the government further subsidized the transition by paying almost half of the cost for the upgrade.\textsuperscript{46}

The IU requires the insertion of a debit card equipped with a smart chip before making any journeys to through the ERP area. Visitors to Singapore from Malaysia or elsewhere are able to rent an IU, though most have found it more efficient to simply purchase and install one outright if they enter the city with any regularity.\textsuperscript{47} As the vehicles enters the city it must pass through a number of gantries where a double layer of sensors first charges the debit card inserted into the IU and then records the vehicle type and location. The transaction takes place instantaneously, confirmed by the IU in the car with a beep, and the commuter can continue unmolested upon their way. In the event of insufficient funds the ERP forwards the offending vehicle’s information to a central computer and a fine is issued. A network of cameras record the

\textsuperscript{44} Ibid.


\textsuperscript{46} Ibid.

\textsuperscript{47} Ibid.
license plate to ensure that a charge is leveled even upon vehicles that are not equipped with an IU.\textsuperscript{48}

While the stated intention of the ALS and the ERP was to reduce the number of cars on the road at peak times, there were associated effects on the health of residents due to a lowering of air pollution. While experiencing a massive economic growth spurt during from the early 1980s through the late 1990s, Singapore managed to maintain a moderate to good level on the pollutant standard index while also increasing the number of vehicles on the road by 26 percent.\textsuperscript{49} Even with this increase in commuters the peak morning and evening times travel at speeds ranged from 25 to 30 kilometers per hour, in contrast with a city like New York where the average is only 10 kph.\textsuperscript{50}

Because Singapore was the first location where congestion pricing was implemented, it has been central to several studies on the effects that policy changes have on the civilian population. Two of the areas affected were retail prices and welfare. An increase in the cost of the ERP had ramifications beyond the morning commute. Different real estate sectors saw increases and decreases in value when the ERP was raised. The retail sector saw a 19% difference between properties within the Restricted Zone (RZ) versus those outside the RZ following a S$1 increase in the ERP in November 2010.\textsuperscript{51} Evidence suggests that a fluctuation

\textsuperscript{48} Ibid.


\textsuperscript{50} Ibid, 358.

in value might be related to the loss of foot traffic caused by the cost associated with traveling.\textsuperscript{52}

Such a dramatic variance was not reflected in either office or residential real estate prices.

Another consideration in the case of Singapore is the effect that the ALS and ERP had on overall welfare. While there were exemptions allowed for those carpooling and motorcycles the fact remains that, for some, the cost to enter the CBD was simply too high.\textsuperscript{53} Those who could not pay the toll incurred significant scheduling costs, making them worse off than they were before either the ALS or ERP was imposed.\textsuperscript{54}

The issue of welfare has, at least in Singapore, been partially assuaged by a robust public transport system that consists of buses and rail lines that make millions of trips into and out of the city center daily. Buses are the most numerous and most used used of all public transportation with over 3300 of them in service, running 240 routes, for a total of almost 3 million trips during a normal business day. Interestingly, these “public” transpots are owned by private corporations, but they are licensed by the Land Transport Authority (LTA) and must conform to their strict standards. In addition they are not allowed to set their own lines of service, the LTA mandates both companies’ running times and routes.\textsuperscript{55}

This extensive bus system is further supplemented by the Mass Rapid Transport (MRT) rail system and the Light Rapid Transport (LRT) both of which consist of approximately 141 km of line and carry upwards of a million passengers a day. Despite the massive costs associated with rail, there are plans to expand this service over the next 30 years.\textsuperscript{56}

\textsuperscript{52} Ibid.


\textsuperscript{56} Ibid, 96-97.
transport in this case has not been in its ability to move people quickly, but the poor integration of their stations into the development of the city as a whole and its impracticality when compared to simply hopping onto a bus or taxi. In an attempt to correct these deficiencies older stations are being refurbished and new LRT lines are being built to better tie the rail system to the needs of the city.

Last of all is the taxi service, which boasts some of the lowest prices in the world. Over 18,000 taxis make their way through the streets of Singapore, carrying approximately 1 million passengers with them. In contrast to the rail and bus lines these taxis are not regulated by the LTA, however stiff competition has kept the price of travelling by taxi very low. These three services, taken together, provide options to those wishing to enter the CBD and thus directly address the question of welfare. Prices on these alternatives are low, either due to a mandate by the LTA, in the case of rail or buses, or due to the brisk competition within the taxi system.

The lessons learned in Singapore echo very strongly Wilson and Vickrey’s initial ideas for congestion pricing. Tolls are dynamic and attempt to approximate the marginal social cost of the journey, electronic toll verification eases any chance of bottlenecking developing at tolling areas, and the shift to public transportation has been accompanied by further investment in rail development to compensate. Allowing private companies to handle the bus system, while mandating standards and routes, as well as the taxis has kept the prices on public transportation low, making this policy less burdensome for the ordinary worker. While there are explicit costs and downsides associated with ERP, it has achieved its initial aim of lowering congestion with

57 Ibid, 97.
58 Ibid.
59 Ibid.
60 Ibid.
startling success. At the same time there have been measureable positives associated with the policy, ranging from increased speeds to better air quality, that represent a net social benefit to the citizens of Singapore.

**Milan**

With over half of its population using cars or motorcycles Milan has a car ownership rate that is one of the highest in Europe, surpassed only by Rome and, as a result, it also has a high level of atmospheric pollution. Its levels of airborne particulates regularly exceed the European Union mandated 50 micrograms per cubic meter.\(^{61}\) After trying several measures, such as banning older vehicles that could no longer pass the increasingly rigorous emissions standards, the city implemented a congestion pricing scheme in order to reduce pollution within an 8.2 square kilometer area.\(^{62}\)

This scheme, called Ecopass, started at the beginning of 2008 and imposed fees from 2 to 10 Euros with the cost being determined by the vehicle’s engine type.\(^{63}\) No tolls were assessed on hybrids and electric cars while progressively higher tolls were leveled on gasoline engines. Diesels and heavy goods vehicles (any vehicle with a mass of over 3500 kg), had the highest cost levied on them, but were still allowed within the zone. In addition there were discounts for vehicles making repeated entries into the city, a fifty percent rebate for the first fifty entries followed by a forty percent rebate for a further fifty entries. After one hundred visits the rebates were capped for the year. Residents of Milan could purchase a yearly pass that ranged from 50 to

---


\(^{62}\) Ibid.

\(^{63}\) Ibid.
250 Euros. Motorcycles and scooters were completely exempt from this toll, as were police, public transportation, and ambulances.

The charge had a dramatic initial result with the level of particulates in the atmosphere falling by as much as 23%, though these estimates do not take into account any other factors that might have influenced the drop. This drop in pollution was accompanied by a reduction of cars entering the area of approximately 23% during the controlled times and 17% during the rest of the day. There was also a change in when trips occurred after the introduction of the toll, with an increase in traffic occurring within an hour of the end of tolling for the day. As expected there was a corresponding increase in the use of public transportation within the controlled zone with the number of passengers increasing by 9.2%.

Milan experienced a social cost savings of 10.4 million Euros, a large portion of which came from the significant drop in the number of traffic accidents. The impact of Ecopass on government costs and revenues was a little more difficult to quantify due to the exclusion of the penalty payments in the total revenue to the city. It was suggested, however, that these penalty rates might be as high as three times the total toll revenues reported, netting the city of Milan a total of approximately 50 million Euros.

---


65 Ibid.

66 Ibid, 363.

67 Ibid.

68 Ibid, 367.

69 Ibid, 367-368.
The increase of public revenue and the decrease of pollution, the avowed goal of the program in the first place, indicated that Milan’s Ecopass system was a great success. It was extended for another year. However, after the initial drop in environmental pollution following the implementation of Ecopass levels of particulate matter started to slowly creep back up to their original highs.\(^{70}\) This recursion to high averages was likely due to the specific provisions of exempting motorbikes, which in turn incentivized drivers to switch to these higher polluting vehicles.\(^{71}\) While black carbon was decreased in the Ecopass area when compared to outside of the controlled zone, other particulate matter measurements showed no discernable difference when compared to levels before the implementation of Ecopass.\(^{72}\)

There were additional effects on the city as a whole. Accounting for other factors, studies showed rents increased by 0.75 percent after the Ecopass policy started, which indicated that some residents valued the increase of welfare that they associated with the potential reduction in pollution highly enough to pay slightly higher rents.\(^{73}\) At the same time there was between a 1.2 and 1.8 percent decrease in the market pricing of houses.\(^{74}\) The effect on housing nearly mirrored the cost to a commuter for an annual Ecopass, suggesting that the cost of the toll was being passed on from the residents to those selling houses.\(^{75}\)


\(^{71}\) Ibid.


The mixed success of Ecopass to achieve its goals can be traced back to the way in which the scheme was implemented in the first place. The exemption of motorcycles and scooters incentivized their use, which potentially offset the reduction of vehicle traffic. Compounding this problem was the narrow times that the charge was in effect, which made it easy to shift travel to other times.\textsuperscript{76} This example highlights the inherent difficulties in any congestion pricing scheme. To be effective in reducing pollution the policy must be targeted, and uniformly enforced across all generally used vehicles that are the major causes of pollution. Further, the times and locations tolled cannot be too specific, unless these routes are the only means in and out of the controlled zone. Otherwise, commuters will make the economically rational decision to travel during those times or through the avenues where the tolls are not leveled.

\textbf{London}

As early as the 1960’s the city of London was experiencing a boom in the number of vehicles travelling in and out of the city on a daily basis. Car ownership in the United Kingdom had doubled from half a million in 1958 to just over a million in 1963 and there was considerable public outcry about the high levels of traffic within the city.\textsuperscript{77} This criticism led the government to consider congestion pricing. However, this idea was rejected due to a perceived surplus capacity of the public transportation network, which opposing politicians argued commuters would use instead over time without the need to set up congestion pricing to push them in that direction.\textsuperscript{78} This extra capacity, however, was soon overtaken by urban population growth. By

\textsuperscript{76} Percoco, “Is Road Pricing Effective?” \textit{Transportation Research Part D} 25 (2013): 118.


2000 the Mayor of London, Ken Livingstone, promised in his campaign for office to implement a road pricing scheme if elected.\textsuperscript{79}

Faced with the dual problem of promulgating this policy as a law and implementing it effectively on the ground. Fortunately for him the 1999 Greater London Authority Act provided him exactly the powers needed, as its provisions explicitly addressed the Mayor’s authority to enact a road pricing plan.\textsuperscript{80} With almost a million commuters entering and exiting the city daily, and spending almost 30 percent of their transit in gridlock, the public had identified congestion and an increase in public transportation as the city's two most important issues.\textsuperscript{81} With a mandate to action Livingstone first applied a £5 charge for privately owned cars, higher for commercial vehicles, to enter the central London area. The pricing was based on the Road Charging Options for London, an independent study supported by the government in 2000. It suggested that that the £5 charge would achieve the goals of lowering congestion during peak times, shifting travellers onto public transportation, reducing car related emissions, and also provide a revenue stream that could be used to further buttress the public transportation of the city.\textsuperscript{82}

The practical application of the London Congestion Charge was based on the work done in Singapore and followed the same model of area licensing that was established there. Vehicles paid a daily charge of £5, that was increased to £8 in the summer of 2005, to enter the zone between 7:00 am and 6:30 pm. There were the normal provisions for emergency and government vehicles, motorcycles and mopeds, as well as sizable discounts or exemptions for residents and

\textsuperscript{79} Ibid.
\textsuperscript{80} Ibid
disabled drivers. Those who fail to pay upon entry have until 10 pm before they face late payment fees, and in the event that the commuter does not pay by midnight there is a fine of £100, , which can be cut in half if paid within 14 days. This flat fee system is worth commenting on as it flew in the face of the traditional pricing plans for congestion charges. In the past it was believed that to be truly effective congestion pricing had to vary by time, but one of Mayor Livingstone’s avowed goals was that the pricing feel “fair” in order to create a lasting system.

The effects of the London congestion charge were seen relatively quickly with increased speeds for commuters increased by 10 or 20 percent between 2002 and 2004 and traffic delays reduced by almost 26%. Additionally there was a reduction in PM10, particulate matter, of almost 12 percent and NOx, nitrogen oxides, levels fell by approximately 12 percent. These declines, specifically the reduction in NO2, in emissions were shown to have added between 18 to 183 years of life per 100,000 population to residents in the Greater London area.

The lowering of congestion within the city corresponded with the expected uptick in the use of public transportation as commuters shifted away from driving into the city. The bus

---


84 Ibid.


services in particular saw massive changes, both in speed, with a 21 percent increase within the controlled zone, and in the number of passengers serviced an increase of 38 percent.\textsuperscript{91} Even with this increase the bus and underground system maintained reliable and consistent schedules of service.\textsuperscript{92}

Retailers in the congestion zone during this period saw lower year end earnings in 2002, the year after the charge went into effect, and a majority pinned the blame on the congestion charge. However other factors such as fear of terrorism and the economic slowdown could also explain this dip.\textsuperscript{93} In fact, some research showed that the congestion charge had statistically insignificant impact on the retail sector.\textsuperscript{94}

To consider the effects on equity we must understand one of the facets of the argument in transportation economics-namely the problem of social exclusion. There are several definitions, but the underlying concept is that social exclusion is any process which prevents residents, or those in close geographic proximity, from participation in any normal activities for that society.\textsuperscript{95} With regards to London, it is clear that there is indeed less traffic within the central control zone than there was previous to the congestion charge being implemented, but there has not been a correspondingly large drop in people entering or exiting the city. The charge has merely changed the means by which people commute, transferring drivers over to public transportation,\textsuperscript{96} and as

\begin{flushleft}


\textsuperscript{93} Ibid, 114.

\textsuperscript{94} Ibid, 124.


\end{flushleft}
a result there was little social exclusion experienced in London’s road pricing scheme. There are tangible gains to society, with one study putting the total net benefit to London at 67 million pounds a year for the first 30 months of the congestion charge.97

The exemptions and reductions in the toll, for the disabled and residents, also helped to make this policy an equitable one since they addressed at risk social groups that might see an undue penalty otherwise.98 As with any policy there will be a cost to some group and, though congestion pricing did indeed represent a burden on lower income households, the overall benefit to society, in terms of increased traffic flow, health, and lower accident rates, outweighs the inconvenience to these smaller groups.99

It is certain that London’s vast public transportation network has aided, if not outright enabled, this policy to be so successful while remaining relatively equitable. One of the main objectives of the program from the start was to shift people from individual vehicles to public transit and it is clear that this goal was met. Public opinion has shifted with the introduction of the toll as well, from an initial split of 40 percent for and 40 percent against the idea to 55 percent for and only 30 percent against.100 Here again, the starting goal, to market this policy in a way that the public would perceive as both fair and reasonable, has proven to be successful. Using a flat rate rather than a dynamic one was unorthodox, but has contributed to this perception.


100 Ibid,495.
The method may be slightly different, but the idea and the implementation could well be lifted straight from the pages of Vickrey or Wilson’s early work on the subject. The toll is paid in advance, cars are monitored by the omnipresent CCTV’s, the tolls is set as close as possible to marginal cost, and, while there are exceptions and discounts, the toll has been applied as broadly as possible. As might be expected, the issues with this policy were mainly political and technical in nature. It is clear that technological advances and innovation could make road pricing schemes even more efficient, and many such changes have already been suggested for London and other cities.\(^\text{101}\)

As two of the largest cities where road pricing has been initiated a comparison between the effectiveness of the London Congestion Charge and Singapore’s ERP is unavoidable. Both use electronic tolling and enforcement to curtail any potential gridlock at access points, though the ERP remains the more effective system. This is based on two factors: accessibility and difficulty to set up. While Singapore has 45 gantries that need to be maintained for tolls, London would require no less than 174 in order to set up an equivalent system.\(^\text{102}\) Another difference between the two systems is the number of exemptions given to vehicles. London has opted for a wider range of exclusions and discounts, while Singapore has been much more stringent. The result of this is a greater effect on traffic patterns within Singapore.\(^\text{103}\)

The pricing and payment of the two systems is also very different, with the London Congestion Charge being applied as a flat amount for a day, while the ERP drivers in Singapore


\(^{102}\) Santos, “A Comparison between London and Singapore,” *Transport Reviews* 25: 521

\(^{103}\) Ibid, 519-520.
are charged by the number of they gantries the pass through.\textsuperscript{104} While these individual charges are lower, they more effectively capture the concept of congestion pricing by tying the decision to use a gantry to an immediate cost, rather than London’s daily or monthly payment. In addition, the ERP rates are graduated, further deterring bad driving habits by having incremental increases over a period of minutes to penalize either speeding up or slowing down at a gantry to get a better rate.\textsuperscript{105} Added to the distinction by time there is a greater attempt to distinguish types of vehicles under the Singapore system, where vehicles are differentiated by type.\textsuperscript{106}

These differences, distinct as they are, do not necessarily reflect better or worse systems as both road pricing schemes were tailored to a specific city, rather than a theoretical generality. Singapore and London cover vastly different physical areas, with the London Congestion Charge covering almost 22 square kilometers as compared to Singapore’s ERP which only encompasses 7 square kilometers.\textsuperscript{107} Congestion times differed as well, with Singapore seeing traditional rush hour congestion while London experienced a steady rate of high congestion throughout the day.\textsuperscript{108} The ERP system is certainly closer to the ideal road pricing system outlined by Vickrey, but it is equally clear that the London model has unique difficulties to overcome and lessons to offer.

Finally, when dealing with such a policy it is imperative to address political difficulty as well as the social equity in the successful creation of a congestion pricing plan. As noted above it is very clear that there is a cost associated with a road pricing scheme, either in time or money,

\textsuperscript{104} Ibid,

\textsuperscript{105} Ibid, 520.

\textsuperscript{106} Ibid, 522.

\textsuperscript{107} Ibid, 521.

\textsuperscript{108} Ibid, 522.
and that such a cost might well be a burden upon at-risk groups. Here is where the example of London can, and should, be invoked however. A pricing scheme can be undertaken if there is a sufficient mandate from the people, strong political leadership, and care is taken to tailor the plan around the existing needs of the city and population. There is no one size fits all solution in road pricing and the goals in each implementation must be clear and realistic, while the effects on public transportation or the poor must be thoroughly considered and, as much as possible, accounted for if a successful program is to be set up.

---


110 Ibid.
IV. Analysis and Conclusion

The idea of road pricing is not a new one, but its successful implementation in major metropolitan areas across the world in the last forty years has given considerable support to its application as a useful urban transport policy. From the seminal example of Singapore to the recent developments in London, Milan, Stockholm, and dozens of other cities and towns the trend has been slowly moving towards forcing motorists to internalize the full cost of their decision to drive, and to do so in an equitable manner. The basis of the policy, marginal cost pricing and a congestion charge to move commuters back along the demand line until they reach their willingness to pay, has been more and more rigorously applied, first with area licensing, then with vehicle type, and finally with sophisticated electronic pricing that varies by time and place.

These developments, naturally, have also raised new questions about its efficacy as a policy. Can it truly be applied equitably? What are the costs and benefits of implementing it? Is it effective? How should such a scheme be set up? It is worth looking at several of the points that one of the most active proponents of road pricing, William S. Vickrey, laid out during the 90’s and apply them to the systems set up in London, Singapore, and Milan. The first and most important principle for any road pricing scheme is price. The charge for each trip must reflect the marginal social cost of the trip in order to be truly effective. At worst it must at least equal this cost, though it can be higher in order to affect some larger social or revenue goal.\textsuperscript{111} Marginal cost pricing can be seen in every city with congestion pricing, some using dynamic charges which reflect the size of the vehicle or the amount of pollution it emits, while others, such as

\textsuperscript{111} “Principles of Efficient Congestion Pricing.”
London, use a flat charge. Either system has its own advantages, but both in theory and practice attempt to cleave to this principle.

As illustrated above in the theory section the problem of congestion, essentially, is in an over consumption of road by drivers who do not pay the real price of driving. To a large extent the public believes that roads are, or should be, free to travel on. At best there is lip service paid to the idea of roads being paid for with tax dollars. But such funds are inefficient and inequitable in equal measure. In the case of public works, the citizen who does not drive, but has still paid their taxes, has been forced to pay for a benefit that they do not receive, while those who do drive pay too little for the benefit. Those that drive are, therefore, encouraged to drive more than they would if they were paying the actual cost of such travel.

Congestion is the end result of this overconsumption, and here again there is a cost to society as each car added to the road only exacerbates the problem and imposes a cost on every other driver. In the words of Walters: “The nth driver does not consider the effect his travel has on the nth plus one driver.” The logical response to this imbalance is the leveling of a toll that is equal to this cost. Theoretically, this is a simple matter, but it raises practical questions about when these charges should be leveled, how they should be leveled, and who should be charged? The largest question, how to effectively set up an efficient system of road pricing, has been solved with the advent of new technology. Electronic Road Pricing systems, such as the one in Singapore, offer a quick and easy method of ensuring payment by travelers without increasing their travel time.

Such an electronic system also addresses two additional points made by Vickrey, namely that tolls should vary smoothly over time and that every vehicle must be charged. With manual tolling changes in toll rates can lead to gridlock, as drivers cluster in an attempt to make it
through while the charge is still low. Singapore’s ERP system has taken this into account, and now has graduated tolling rates that rise and fall based on the time. The use of gantries which all cars must pass through as the trigger for a charge allows for charging without a car changing its speed. Cameras automatically record and ticket drivers who pass through without paying the toll. Even Singapore, however, falls short of the ideal system as there are exemptions for public and governmental vehicles using the road. These should, according to Vickrey, also be charged, even if it is only as a means of accounting and budgeting to allow for more accurate pricing.  

As the historical examples demonstrate, the benefits of congestion pricing increase the more the basic principles are adhered to, but no two policies address the precisely the same issues. Each policy must be tailored to fit the needs and geography, both physical and political, of the city. In Singapore, an island city with an authoritarian government, the ability to make quick and decisive changes was helped by its relatively small area. Its initial experiment with road pricing, an area licensing scheme, was intended simply to reduce the number of private cars on the road and followed earlier heavy taxes levied on the importation and ownership of vehicles. While the taxes failed to achieve this goal, area licensing succeeded as congestion fell rapidly and traffic speeds increased. 

Such a system, however, was still not as efficient as it possible because it was based on static rates rather than the more desirable varied rate toll propounded by Vickrey or Walters. This parallels the London Congestion Charge, though for a different reason. While Singapore initially used flat rates because there was not a better system at the time, London’s road pricing scheme opted for flat rates even though a better system was available. there were two reasons for this decision. The first was geography. To make the congestion charge fully automated in the greater

112 "Principles of Efficient Congestion Pricing."
London area, which covers some 22 square kilometers and it would take approximately 200 electronic tolling gantries. The second was the political situation at the time. While the problem of congestion consistently ranked as one of the most pressing issues in London according to public polls, any road pricing scheme had to address the issue in order to be accepted by the public.

In London’s case the way to deal with both of these issues was to make the congestion tax so that (barring heavy goods vehicle, who pay more and those with exemptions or discounts) the public would perceive the toll as “fair”. This ploy has worked as more people support the toll now than when it was introduced. However, this also means that the London Congestion Charge is somewhat inefficient. While the average motorist pays more of the cost of their trip, they are likely paying too much for travel during off peak times and too little during rush hour. While there is some justification for this decision because London has constant congestion there remains room for improvement in the system.

In contrast, Milan set up its Ecopass system to curb pollution associated with vehicle traffic. While the system has been successful in reducing vehicle traffic, the air quality changes have not been as profound as expected. If judged by its stated goal, Milan has only partially achieved what it set out to do when it first implemented Ecopass. At the same time, however, it has also seen a reduction in congestion and an increase in speeds which is the goal of most road pricing schemes. Milan case did not apply the toll uniformly, which violates one of the main points made by Vickrey for a truly efficient system. By allowing exemptions to motorcycles and not tolling every route into and out of the control zone the Ecopass system backhandedly incentivized behaviors that were counter to their original intentions.
In addition to the cost to commuters in terms of time there is also a cost to cities who attempt to address congestion and bottlenecks with the construction of additional infrastructure. While some building is necessary, the problem of congestion is usually not a lack of lanes but rather a clustering of motorists attempting to over utilize during a given period. The demand for roads can usually be met by existing supply if trips were staggered and less desirable paths were used. The construction of new infrastructure can actually increase demand instead. When you consider the cost of building new infrastructure to accommodate congestion the idea of roads as pure public good is even untenable.

If heavily travelled roads are not considered pure public goods then the question of equity shifts from the propriety of a commuter paying for the use of it to the more practical concern of who should pay for it. The sensible answer is that those who make greater use of the road should pay more than those that do not use it with any frequency. In contrast to taxes that are leveled across the board by a state or federal government, road pricing directly touches this first group while avoiding charges on the second. Whether driving a totally electric car or a diesel engine truck, both vehicles occupy the finite amount of roadway and both contribute equally to congestion. But, where a gas tax would recoup some of the cost for maintaining the road from the truck, the electric car is, at least partially, a free rider.

This hypothetical situation illustrates the point that, rather than being inequitable, road pricing can actually be a fairer system, as well as a more effective policy. With gas mileage increasing and a reluctance to levy new taxes, or raise old ones, congestion pricing represents a new source of revenue. If this revenue, less the cost of building and maintaining the road pricing system, is used to subsidize and expand public transportation then the argument about inequity becomes weaker. This was demonstrated in London, as Mayor Livingstone’s goal was to cause a
shift from cars to the underground and bus systems. In his own words the buses, “are now more reliable than at any time since systematic records began in the late 1970’s”.113

This increase in public transport to offset the increase caused by the congestion charge also goes to the heart of the dilemma of a road pricing scheme and its effects. While there is a gain in traffic speed and a lowering of urban congestion it is only achieved by forcing people to alter their habits by the application of a significant toll. In areas where the is no significant public transportation service the cost on working individuals is, to a large extent, unavoidable. They must simply pay the toll or else risk forfeiting their income. However, going back to the original concept as outlined by Vickrey, the commuter is saving in time by the reduction of congestion and this does counterbalance some of the cost of the toll, as well as the additional benefit of lowered atmospheric pollution.

The argument can be seen in two ways. The enactment of a toll imposes an undue hardship on lower income travellers, alternately the imposition of a toll is merely correcting for the over consumption caused by commuters paying only the average social cost. There is no easy answer to the first argument, except that individuals are priced out of markets regularly and which is nothing more than the nature of the supply and demand. There is a demand for road use, an inelastic one, and as a result people will pay the price to gain the benefit, whether it be the marginal social cost or the average social cost. At the same time, the revenue realized from congestion pricing should be used to incentivizes economically efficient behaviors. This change could be as simple as taking a different route to the CBD or driving on off peak hours, either one of which would result in a lowering of pollution and congestion in their own right.

The purpose of congestion pricing is not to force an undue penalty upon those that cannot afford to pay; instead it is to accurately represent the real cost of driving and force commuters to confront and internalize this cost. In the absence of tolls there are no immediate prices to indicate this real cost to consumers, and the over consumption that results from this imposes more costs in terms of time to each motorist. The efficient, and equitable, policy action to combat this is road pricing. The example of Singapore’s ERP shows that it is possible to create a system that is effective, but not onerous. London shows that even a flat rate can have a strong effect upon the habits of drivers, encouraging the shift away from peak times and returning revenue to support an expansion of public transit. Finally, while its end result is still in debate, Milan’s Ecopass system has seen similar reductions in congestion and the corresponding increases in speed that are predicted by every road pricing model.

The advances in technology used in Singapore's ERP system show how future congestion pricing might look like. Every vehicle would be charged, if only to better approximate the marginal cost associated with vehicles entering traffic. The machine in the car would ding as a motorist passed through a gantry making them aware that you have been charged, which reinforces the idea that they are trading money for time. The ability to change times gradually at tolling stations approximately models the changes in marginal cost over time and limits the attempts to take advantage of the system by slowing down or speeding up. The availability of other options allows consumers a choice of when and how to travel. All of these examples are directly in line with the initial theory of congestion pricing.

There is a distinction that must also be made between what is done with congestion pricing and what should be done. No congestion pricing scheme has yet reached the level of economic efficiency envisioned by Vickrey and Walters. The technology exists, the theory is
sound, but it is in the political application of road pricing that in all examples fall short. Even in an authoritarian government such as Singapore there are still improvements that could be made. In more democratic nations congestion pricing is only as strong as the political establishment that implements it. London compromised on a dynamic rate for the sake of "fairness" and Milan failed to uniformly charge all vehicles. These two cities are excellent examples of the difficulty in gathering sufficient political approval to utilize road pricing. Compromises are made and the result is a less efficient system. However, even if these road pricing schemes are not the ideal system of Walters or Vickrey, they do have the expected results. Congestion is lowered, pollution diminished, and travel times decreased in these cities. A slightly inefficient policy is better than a completely inefficient one.

In the light of these examples, and others taking place from the United States to Sweden, the implementation of road pricing is becoming a more appealing policy. While concerns remain it has proven to be a tool that works in reducing both congestion and pollution, if applied rigorously and fairly. The benefits, and revenues, of such a policy if applied correctly can enhance the overall welfare of a society by the increases in time as well as the added life years gained by lower emissions. While there are still practical issues to be addressed and improvements to be made congestion pricing remains an effective policy that should be utilized more fully.
V. Bibliography


Hensher, David A. “Exploring the relationship between perceived acceptability and referendum
voting support for alternative road pricing schemes.” *Transportation* 40 (2013): 935-959.


