Introduction

William Vickrey, winner of the Nobel Prize for Economics, is considered the father of Congestion Pricing. He first proposed it in 1952, for the New York City subway system, recommending that fares be increased in peak times and in high-traffic sections and be lowered in others. Elected officials considered it risky, and the technology was not ready. Later, he made a similar proposal for road pricing.

Vickrey considered time-of-day pricing as a classic application of market forces to balance supply and demand. Those who are able can shift their schedules to cheaper hours, reducing congestion, air pollution and energy use – and increasing use of roads or other utilities. “You’re not reducing traffic flow, you're increasing it, because traffic is spread more evenly over time,” he has said. “Even some proponents of congestion pricing don't understand that.”

He has admitted that his ideas have sometimes not been well received by those who set public policy because, “People see it as a tax increase, which I think is a gut reaction. When motorists' time is considered, it's really a savings.”

He suggested, “One possible detection and billing method would use electronic identifier units carried in each vehicle, which would activate recording devices in or on the road. Computers would sort the information and determine charges; motorists would be billed monthly.” This is exactly how modern Road Pricing systems function.

Principles of Efficient Congestion Pricing

Here are some of William Vickrey’s thoughts on how to implement efficient congestion pricing for roads and parking.

1. Charges should reflect as closely as possible the marginal social cost of each trip in terms of the impacts on others.

Charges may be set to exceed such costs as a means of obtaining additional revenues for the government at slight excess burden, justified possibly in naive terms as imposing a surcharge tax on this activity comparable to tax surcharges on other activities and purchases. There is no excuse for charges below marginal social cost. If it be urged that such charges in a particular case would be desired on distributional grounds, there will always be more efficient and equitable ways of effecting such redistribution. If efficiency and revenue are both sacrificed, it will seldom if ever be possible to recoup the loss of efficiency elsewhere.

2. Charges should vary smoothly over time.

Only in this way can everyone be given an incentive to shift the time of travel, if only by small amounts, away from the peak. If charges vary discontinuously, excessive incentives are given to rush to get ahead of a jump in the charge, or to lag waiting for a drop in the charge. There is a likelihood of creating mini-peaks just before a scheduled increase or just after a scheduled decrease. Few of those moving at the height of the peak will be willing to make the substantial shift in travel time of an hour or more necessary to obtain a lower charge. It will often be easier and cause less disruption to get 12 drivers to shift the time of their trip by 10 minutes each than to get one person to shift his trip by two hours.

If staggering of working hours were to succeed initially in eliminating peak congestion, in the absence of road pricing it would then fail, as firms seeing that congestion has been abated then drift back to their preferred times and recreate the congestion. Reasonably smooth variation of the prices over time is needed to prevent this from happening.
3. Efficient charges cannot be determined solely by conditions at the time of the individual trip, but must take into account the impact of the trip on other traffic from the time the trip is made until the end of the congestion period.

Example: If queuing conditions are such that flow through a choke point is at capacity from 7am to 10am, a car going through the choke point at 7:10 may encounter a short queue and be delayed by only 5 minutes, but will be responsible for there being one more car delayed in the queue from 7:10 until 10:00, or 2.83 vehicle-hours of delay; conversely, a car going through the choke point at 9:55 may have been delayed by 15 minutes, but will impose only 5 minutes of delay on others. A motorist who starts his trip 30 minutes earlier may reduce his own waiting time from say 15 minutes to five minutes, but unless he gets entirely free of the queue, he will increase total waiting time by 30 vehicle-minutes and the aggregate waiting time of others by 40 vehicle-minutes. If charges were calculated in terms of his own delay, the incentive would be perverse. Efficient congestion charges would eliminate the bulk of the queuing delay without any decrease in peak flow and possibly an increase in shoulder flow. Marginal social cost is measured by the expected time from the passage of the particular car through the choke point until the next time at which the queue drops to zero.

Where hypercongestion exists, such that traffic density is above the point of maximum flow, which is generally the case whenever speeds are below about 60 percent of the low-density speed, impacts may be even more severe. Injecting an additional car into the area decreases the total flow, decreasing the rate at which cars succeed in getting off the congested street network, increasing the density and further decreasing flow in an exponential manner. I have estimated that, under the 1983 conditions in midtown Manhattan, an additional 3-mile trip at 11 a.m. can account for between 100 and 300 additional vehicle-hours of delay over the remainder of the day. By keeping the accumulation of traffic density below the point at which flow begins to decrease, efficient road pricing can actually increase peak hour flows, and in many cases increase total daily flows, while increasing speeds, reducing air pollution, and yielding needed public revenues.

4. Efficiency can be enhanced, for a given level of data collection, by charging on the basis of the trip segment from one observation point to the next, rather than by merely the passage of an observation point.

This can be handled by dividing the congested area into zones of suitable size, large enough so that checkpoints can be installed at every crossing of a zone boundary by a thoroughfare. In processing, each entry of a zone by a vehicle can be matched with its exit and a charge assessed according to the indicated trajectory, rather than separately for merely passing a control point.

5. There is much to be said for charging on an ex post, strict liability basis in terms of the actual impact that a trip can be calculated to have had on the traffic as actually experienced, over the balance of the day, rather than according to some schedule fixed in advance.

A present schedule will be unable to allow for increased congestion caused by weather, sporting events and the like, whereas motorists would have the opportunity to inform themselves of these situations and adjust their trips accordingly. In the case of adventitious events such as fires, accidents and the like, one could perhaps allow a grace period of 15 or 30 minutes from the time the occurrence has been broadcast before increased charges become effective, to avoid unduly charging motorists who would have had no opportunity to alter their plans. The question is whether motorist expectations can be relied upon to make better estimates of impending conditions than can be represented to them in schedules and broadcasts. The problem is that where congestion threatens, traffic conditions seem often to vary very widely from day to day even when there is no broadly recognizable cause for the variation.

6. All vehicles should be charged without exception, including trucks, doctors’ cars, press cars, and cars of public officials and diplomats, among others.

Even where these charges are borne by third parties, it is useful to maintain the integrity of the market pricing principle and avoid disputes as to qualification for exemption. Public vehicles such as police cars and fire trucks should also be charged, even if this is only an accounting transfer, in order to give a true picture of the cost of these operations and provide an incentive for performing their functions more efficiently, and possibly induce a better budgeting of public funds with due regard for true costs. One could even argue in principle for the application of surcharges where sirens or other priority signals are used, though many would consider this to be carrying principle too far.
7. Taxicabs present a special problem of ascertaining the charge at the time of incurrence, so that it can be charged to the customer.

Where other vehicles are being billed monthly, on a segment by segment basis, it may nevertheless be possible to adopt for taxicabs a special regime in which they are charge on a modified point basis. The beam emitted by beacons or scanners can be modulated to indicated a level of charge per mile to be added to the meter charge until the next checkpoint is encountered, or until a set mileage increment has been run up, the latter to avoid excessive charges for portions of trips outside the zoned area. While not as accurate as the segment charging method, it should be sufficiently accurate to serve the purpose so that the cost of a more accurate system might not be justified.

8. Curb parking, where permitted at all, should be charged on the basis of clearing the market.

In principle the price should be made to reflect as closely as possible the marginal social cost of the occupancy of a space in terms of the cost to other would-be parkers of added difficulty in finding a space, or of having to resort to other modes. In practice, this could be approximated by a rule saying that if, over a suitable number of weeks, fewer than, say 5% of the spaces are typically vacant during a particular time slot, the charge should be increased, and if vacancies are consistently more than, say 20%, the charge should be reduced, or eventually eliminated. Charges may appropriately be made to vary with the size of the vehicle.

The coin-operated parking meter controlling a single space is an inappropriate device for collecting such charges. Existing meters operate at only a single rate when specified to be on, and adjusting the rate to adapt to changing conditions is expensive. Collecting and auditing the proceeds are expensive, and theft and vandalism are often problems.

9. One simple and inexpensive method of collection would be by means of parking cards.

Parking cards can be sold in convenient denominations through service stations, or along with lottery tickets, or by mail. Cards would be imprinted prior to use with the license number of the vehicle with which they are to be used. One or more of these cards having a total value equal to the anticipated charge would then be punched or mutilated by the user to specify the time at which parking begins, and displayed on the vehicle. The time paid for would be determined by comparing the value of cards displayed to a schedule posted at suitable intervals. The main problems are the sophistication required of the driver unless the schedule of charges is simplified to the point of failing to obtain the most efficient use of the space, and the longer time it would take the wardens to ascertain whether or not a car is in violation, as compared to merely observing the flag on a mechanical meter. On the other hand, capital, maintenance, collection and auditing costs should be much lower than for the mechanical meter. Charges can be adjusted from time to reflect current circumstances by merely changing the posted schedules. Some difficulty may be encountered in getting habitual users to adjust promptly to comply with a new schedule.

10. Another method would use parking ticket vending machines.

Vending machines would be set up at convenient intervals where a button can be pressed to indicate the class of vehicle, and money inserted until a display indicating the time up to which parking has been paid for reaches a satisfactory value. Especially where the appropriate level of charges would involve an inconvenient number of coins, payment can be by inserting a credit card or a stored value card and using a button until the desired time limit has been reached. Pressing another button then causes a ticket to be issued showing the type of vehicle and expiration time, and possibly the serial number of the card used, which might include the license plate number of the vehicle with which it is to be used. This ticket would then be displayed on the vehicle. The data on the ticket can be indicated in a large-scale coded form that can be checked by the warden from a moderate distance, expediting the checking process. A major advantage of this scheme is that the schedule of charges can be fairly closely graduated to balance demand and supply without causing difficulty for the parker or the warden, and can be updated fairly easily as conditions change. A minor disadvantage is the need for the parker to walk from his parked car to the ticket machine and back. Capital and maintenance costs are higher than for the card method; The cheaper card method
might be considered an interim measure pending a decision as to how large an area to cover with the ticket machine method.

In either case it would be part of the duty of parking wardens to make a daily record of the usage observed, to serve as the basis for adjusting the charges to an efficient pattern.

11. Delivery vehicles and other vehicles making frequent short stops need special treatment, such as by using on-vehicle meters.

Delivery and similar vehicles could be provided with a meter capable of being set to run at varying rates that would be displayed for checking by enforcement agents. The meter would be periodically reset much like a postage meter. Double parking or parking in space subject to restrictions could be charged for at a suitably high rate rather than being erratically penalized or tolerated. It is ridiculous to see trucks double-parked just beyond bus stops when parking illegally in the bus stop would cause much less of an impediment to bus operations. Such charges would provide incentives not only for shifting deliveries to off-peak hours, but for shortening the time traffic is obstructed by employing additional helpers to speed the pick up and delivery operations. In the longer run, there can also be the replacement of congestion generating activities by activities creating less congestion, and for arranging loading bays that do not obstruct traffic. The main problem here is likely to be that of finding an objective and administrable basis for setting the level of charges at various times and places.

12. Political interference and bureaucratic bungling can spoil the game.

Unfortunately experience with the pricing of public services is not such as to give confidence that in practice a close approach to an efficient optimum can be achieved. In many instances, efficiency requires a substantial disruption of the patterns of activity of many, although especially in the case of persistent queuing and hypercongested networks the change may not be as great as many fear. Of perhaps greater political impact is the fact that, in many cases, individuals will be asked to pay fairly high market-clearing prices for what they have been getting at no direct money cost, though in many cases at a very high cost in terms of loss of time. It is an unfortunate fact of political life that those who expect to be injured by a change feel the threat very vividly and make a correspondingly great amount of political noise, whereas the potential beneficiaries are often not so vividly aware of the goodies that await them and thus tend to throw relatively less weight into the political balance.

At the bureaucratic level there is also the tendency to be satisfied with the status quo, especially where tenure of office is relatively secure. Even where change is obviously called for, it tends to come in homeopathic doses and to follow lines of proportional adjustment along traditional patterns rather than break into innovative territory. There is also a "not invented here" syndrome that is especially prevalent among large organizations that regard themselves as sui generis.

It is thus going to take a lot of pushing and earnest education to get progress made along the lines laid out above. But the potential gains are so great that it may just be possible to keep the meddling of special interests within bounds and overcome the drag of traditionalists to come out with something really worthwhile. And once a start has been achieved, who knows how much further one will be able to go.

Biography

William Vickrey was born in Victoria, British Columbia, in 1914, he received a bachelor of science in mathematics from Yale in 1935. He went to Columbia University for graduate work in economics from 1935 to 1937, when he received the M.A. degree. His doctoral thesis, "Agenda for Progressive Taxation," written for Robert Murray Haig for a 1948 Ph.D., was reprinted in 1964 as part of a series of economic classics.

His first study of efficient pricing of public utilities in 1939 and 1940 was of the electric power industry for The Twentieth Century Fund. In 1951, he studied transit fares for the Mayor's Committee on Management Survey in New York and in 1959 he presented to Congress a proposal to control the District of Columbia's
traffic congestion with electronically assessed user fees. He has addressed urban planning problems in Calcutta with the Ford Foundation and in Buenos Aires and New Delhi for the World Bank.

A conscientious objector during World War II, he spent part of his alternate service designing a new inheritance tax for Puerto Rico. After the war, he joined Columbia economist Carl Shoup on a team of economists who toured Japan in 1949 and 1950 to recommend reforms of the country's tax system.

Vickrey began his Columbia career as a lecturer in economics in 1946. He joined the faculty as assistant professor in 1948 and was named associate professor in 1950, professor in 1958 and McVickar Professor of Political Economy in 1971. He was chairman of the department of economics from 1964 to 1967 and retired as McVickar Professor Emeritus in 1982. He maintains an active schedule on campus and continues his research.

He was elected to the National Academy of Sciences in 1996 and served as president of the American Economics Association in 1992. He was elected a Fellow of the Econometric Society in 1967 and received an honorary doctorate from the University of Chicago in 1979 for work in game theory and social choice theory. He was awarded the Nobel Prize for economics in October, 1996. He died at age 82 of natural causes, two days after the Nobel award was announced.

Bibliography

During the course of his 60+ year academic career William Vickrey published hundreds of books and papers covering a broad range of economic subjects. Two excellent summaries of his work are listed below.
