that the increases could be attributed for the most part to improved customer and employee accessibility.

APPLICATIONS

Possible applications of these results are many. The most important use is in connection with impact statements and public involvement programs. This application provided the original motivation for the study. The results of this study have quantified the property value effects of a limited-access highway. This information can be used for generally assessing property value effects in similar locations when a highway is constructed. The effects on property value are a great source of public concern. This evidence will provide facts for detailed discussions on this topic.

There has been interest in partly financing highway construction by capturing part of the accessibility benefits through property taxes. The property value effects are caused by the user benefits from the highway and do not represent an additional benefit. If existing taxes on highway users are at an appropriate level, then an additional tax on property is not called for. If additional taxes are indicated, they could take either form with similar long-run effects. A related point is that care must be used in applying the results of the benefit side of this study to benefit/cost analyses. Double-counting would result if user benefits were fully evaluated and property value effects were added.

These same considerations do not apply to the adverse property value effects of noise. Noise represents an externality that must be considered in benefit/cost analysis in order to make efficient decisions. The distributional effects of these externalities might also provide a basis for the payment of compensation to the residents affected. Such compensation should be paid to the house owners at the time of the highway effects origination but not to those who purchase the house after the effects take place. Currently, the Federal Highway Administration requires that controls such as noise barriers be used to reduce highway noise to 70 dB(A) in residential areas unless exceptions are granted. If additional taxes are indicated, they could take either form with similar long-run effects. A related point is that care must be used in applying the results of the benefit side of this study to benefit/cost analyses. Double-counting would result if user benefits were fully evaluated and property value effects were added.

Finally, this study might prove useful in making decisions between various transportation modes. Such a choice between modes must be based on all of the effects of the construction of each mode.

ACKNOWLEDGMENT

This paper is based on a research project undertaken by the staff of the Social and Economic Planning Section of the Department of Transportation. The study was conducted the study assisted by Brian Limotti, community affairs consultant; Carrol DeSonia, Keith Greetham, and Gary Bruner of the Land Management Division assisted in data collection and interviewing. Evan A. Iverson, supervisor, Social and Economic Planning Section, served as project director. Guidance and suggestions were received from Keith Densley, chief right-of-way agent of the Department of Transportation; G. Murray Duncan, acquisition supervisor; F. Dale Curfman, real estate specialist, Washington Division Office of the Federal Highway Administration; and Willa Mylroie, Research and Special Assignments Engineer.

REFERENCES


Publication of this paper sponsored by Committee on Application of Economic Analysis to Transportation Problems.

Some Conventional and Not-So-Conventional Views of Congestion

A.O. PEARMAN AND K.J. BUTTON

The purpose of this paper is to explore the extent to which the conventional treatment of highway congestion, as developed in the economic analysis of road pricing, provides an acceptable theoretical or practical foundation for policy. The conventional theory is first outlined, and it is emphasised that, although it is probably technically sound, it relates to highly abstract circumstances. The main body of the paper then develops two themes. First, a number of arguments are put forward that imply that, in quantitative if not qualitative terms, the conventional analysis of congestion seems unlikely to provide an adequate basis for the proper formulation of policy. Second, some reasons for regarding congestion as an effective allocative mechanism in its own right are given. Although the arguments in the paper are not developed sufficiently far to reach firm conclusions of an operational kind, there are clear indications that traffic management and related policies aimed at securing efficient use of existing highway facilities should proceed with care when valuing congestion savings and when assessing optimal congestion levels.
A great deal of money continues to be spent to relieve traffic congestion. A great deal of time is spent in traveling in congested conditions. Where does the appropriate balance lie? It is all too easy, certainly for the layperson, and maybe even for the professional, to accept the conventional wisdom that congestion is bad, and that it is unquestionably the task of transport planners to eliminate it. Recent concern on both sides of the Atlantic about securing better use of existing transport infrastructure through traffic management and related policies has highlighted the issue, but perhaps not adequately.

We queue in congested conditions in shops and banks. We wait for a library book to be returned rather than buy it. We wait for state hospital treatment as an alternative to paying the fees associated with the private sector. Has society got it wrong everywhere? Clearly not—and nearly all professional transport planners will possess a degree of familiarity with the main arguments that underlie the determination of the correct level of congestion that should be permitted on a highway facility.

However, what we hope to demonstrate is that the conventional wisdom of urban traffic congestion is nothing like as firmly founded as a textbook familiarity with the topic might imply. We express some doubts about the theoretical basis of the conventional argument, and even more about the value of its implications as a guide to policymaking in the real world. Congestion, indeed, may even have positive advantages as a device for helping to distribute scarce resources among different sections of society.

The aim of the final section of the paper is to assess the implications of the previous sections. It cannot offer a neat conclusion, but it does suggest that, although we understand a little about the theoretical basis of the conventional argument, and even more about the value of its implications as a guide to policymaking in the real world. Congestion, indeed, may even have positive advantages as a device for helping to distribute scarce resources among different sections of society.

CONVENTIONAL WISDOM

Economic analysis of traffic congestion has been undertaken in a number of ways. One approach, as for example by Mohring (1), Kraus and others (2), and Wheaton (3), has its foundations in conventional economic theory. It is a largely algebraic application of the theory of consumer behavior to circumstances that are broadly consistent with those found on congested urban roads. A second approach, which differs more in emphasis than final form from the first, is characteristically graphical in its mode of thought, starts from a physical description of the development of congestion as traffic volumes increase, and only becomes an economic analysis when the time and other resource commitments consequent on tripmaking are aggregated into a measure of trip-generated cost. For the most part, the second of the two approaches has received the greater exposure in the professional literature, and certainly in student textbooks, and it is this that we take as our starting point.

Early thinking on the economics of road traffic congestion is described by Pigou (4) and Knight (5), but the principal impetus to more recent work stems from the contributions of Walters (6-8) and Beckman and others (9). The account that follows draws particularly on the descriptions given by Else (10) and Johnson (11). Conventionally, the analysis is based on a number of rather restrictive assumptions, the result of which is a tractable problem, but one whose solution, it will be seen later, is not necessarily very valuable. The assumptions are as follows:

1. A homogeneous traffic flow moving in a uniform direction.
2. A single road with only one entry and exit point.
3. Resource allocation considered only with reference to the road itself (i.e., ignoring interactions with other sectors of the economy), and
4. A demand for the use of road space expressed as a demand for a number of vehicles to emerge per time period, with instantaneous adjustment of density over the whole road so that density is uniform at every point.

If we let

\[ F = \text{flow (vehicles emerging from the road/min)} \]
\[ D = \text{density (average vehicles/mile at a given time)} \]
\[ S = \text{speed (miles/min attained over the road)} \]
\[ T = \text{time (journey time for any vehicle)} \]
\[ L = \text{length of the road (miles)} \]

Then the following relations hold,

\[ T = (D) \] (1)
\[ F = DS \] (2)
\[ S = L/T \] (3)
\[ F = [D/(D)] L \] (4)
\[ T = g(F) \] (5)

Equation 1 states that \( T \) will depend on \( D \) \([T = f(D)]\), as illustrated in Quadrant I of Figure 1, with \( f' (D) > 0, f'' (D) > 0 \), and \( f(D) \) asymptotic to \( D_{\text{max}} \), reflecting a maximum density, beyond which movement along the road sizes up entirely. At low levels of density, travel time is almost independent of density.

In equilibrium, the flow of traffic that emerges from the road \( F \) is equal to the product of traffic density and speed (Equation 2), speed in turn being defined by Equation 3. A combining of Equations 1, 2, and 3 enables flow to be expressed as a constant (L) multiplied by a ratio \( D/f(D) \), which in graphical terms is simply the slope, measured relative to the vertical axis, of a ray from the origin to any point on the \( f(D) \) curve in Quadrant I. This ratio is clearly at a maximum when the ray is tangential to \( f(D) (D = D_{1}) \), and is zero when either 0 or as \( D \) tends to \( D_{\text{max}} \). Graphically, Quadrant II in Figure 1 constructs the relation between \([D/(D)] \) and \( D \) and Quadrant III shows the linear relation (Equation 4) between the density/flow ratio and flow itself.

The final step in developing the conventional, backward-bending relation between flow and travel time involves, algebraically, substitution of Equation 1 and its inverse into Equation 4. Graphically this can be achieved by noting that any technically feasible flow \( (F_{2}) \) in Quadrant III can be associated with two separate travel times \( (T_{2}' \) and \( T_{2}'' \)) in Quadrant I, via the relation depicted in Quadrant II. Thus, the backward-bending relation \( T = g(F) \) in Quadrant IV can be constructed.

Thus far, the relations discussed are noneconomic, based on the physical characteristics of the road in question and the observed behavior of travel times as traffic density increases. However, to de-
Determine what flow of traffic will in fact be the output of the road, economic concepts are introduced. First, it is posited that travel cost will be a continuous, monotonically increasing function of journey time. This is meant to reflect the fact that travel time itself has an opportunity cost (a value) and that costs of travel such as fuel consumption and engine wear will also generally be greater as densities increase and speeds fall.

As a consequence, it is possible to reconstruct the relation shown in Quadrant IV as a similarly shaped relation between travel cost and flow (Figure 2). This may now be regarded as an average private cost curve. For any given value of flow, it shows the costs that accrue to the individual of being one component of the flow of traffic. In fact, for any given flow, one of two possible cost figures will arise, either one on the lower part of the curve (AB) or one on the upper part (BC). This rather counterintuitive possibility is explained as follows. Normally, as more traffic seeks to use the road, flow increases and so does cost (points in the range AB correspond to such situations). However, the strength of demand for use of the road space can be sufficiently great that density exceeds the density corresponding to maximum possible flow (D1). In this case densities increase such that speeds decrease more than proportionately, and a flow less than the maximum results on the high cost portion (BC), of the cost-flow curve, the higher costs following from the greater time and resource costs associated with congested travel.

In formal terms, demand can be represented by a downward-sloping curve (JK), which implies that as the costs of travel fall, more people seek to travel. Thus, on which part of the cost-flow curve equilibrium is established depends on the precise position of the demand curve.

Given this formalization of the interaction between the demand for use of a road and its cost-flow characteristics, a number of questions conventionally follow. The first asks whether the volume of traffic that will choose to use the road is the correct one. The conventional answer is no. Assuming that the road can be viewed in isolation, economic efficiency in the use of resources requires that it is the intersection of JK with the marginal social cost curve and not the average private cost curve that should determine the appropriate level of road use. Some controversy exists over the definition of a marginal cost curve, as will be discussed later, but whatever attitude is taken in this latter respect, it can be concluded that, without the imposition of some structure of taxes, an incorrect level of traffic flow will in general result. This, combined with the personal frustrations of time spent waiting in traffic queues, leads to the conventional wisdom that congestion is a bad thing, which at least in some crude evaluations of the problem ought to be eliminated. This is another question that will be raised again later.

Finally, there is the problem of investment. Distortions in the market for the use of a service are likely to induce distorted responses in investment decisionmaking. Borins (12) and Wheaton (2) explore this topic, which is also one considered later.

DOUBTS CONCERNING CONVENTIONAL WISDOM

The conventional exposition of urban traffic congestion has been questioned both as a theoretical construct and, less directly, as a consequence of doubts about the validity of policy implications drawn from analysis based on that theory. This section will explore both aspects, starting with the empirical. Later some arguments will be put forward concerning possible advantages that stem from the presence of urban traffic congestion.

In considering the policy arguments that have developed from society's awareness of the disbenefits of increasing congestion over time, it is important to be clear about the scope of debate. In the previous section the assumptions adopted ensured that we were dealing with just a single type of road user, and that the costs considered were only those imposed by and on the road users themselves. Thus, no account was taken, for example, of the distinction between private and public transport nor that between the transport of passengers and goods. Also, the pollution effects of congestion were ignored, insofar as they impinged on anybody other than road users at the time of their travel. None of the wider effects that congestion might impose on society through influence on industrial location or urban structure were brought within the purview of
the analysis either. To follow through, all the possible arguments on such issues must remain outside the scope of a short paper such as this.

There are problems enough merely within the range of consideration implied by the implementation of the policy implications of the conventional arguments. The remainder of this section will be developed largely on that basis. It concentrates on the points concerned with making use of an urban road network of roughly the same scale as is already available for passenger transport. This amounts to the implementation of policies to shift cars from the places and times where congestion is currently excessive.

Under congested conditions, any tendency on the part of motorists to underestimate their own real costs of motoring or the costs they impose on others will encourage an excess of car use. In theoretical terms it is then very straightforward to argue that a tax given by MN in Figure 2, which will reduce traffic to the socially optimal level, should be imposed. In practice, however, such textbook conventional wisdom soon begins to lose some of its appeal. Even ignoring all the points raised in the previous paragraph, there are still difficulties. First, traffic conditions in urban areas vary widely according to time and location. The introduction of a flexible taxation system seems at present if not impossible, unlikely. Ignoring licensing and other rather coarse schemes such as that imposed in Singapore (13), the only flexible alternative appears to be some sort of vehicle-based metering scheme, presumably backed by modern computer technology, of the kind discussed, for example by Both (14, Chapter 5).

However, a scheme on those lines immediately highlights that a divergence between social and private costs is only a necessary and not a sufficient condition for the introduction of traffic limitation policies of the type under discussion. The running costs of such a scheme would be not inconsiderable and must, of course, be set against any savings in resource costs as a result of changes in patterns of car use (15). Economic costs are also likely to arise from trips previously made by car that are now suppressed or diverted to other modes. A commuter who arrives at work early may lose the time between arrival and the conventional starting time. Also, it may very well be that the costs imposed by extra use of public transport in the peak hours alone would be considerably in excess of fares charged and so diminish the net benefits of any road pricing scheme.

The balance of costs and benefits in a road pricing scheme determines its desirability. The danger in the conventional textbook presentation of the argument is that, in the interests of clarity, it tends to relegate the very difficult task of performing the balancing calculations to a secondary role, which leaves the case for pricing as all too easy to accept as a basis for policy, no matter how actually implemented. For example, the financial support of public transport as a second-best approach to proper pricing is by no means self-evidently desirable, either as an alternative to the do-nothing alternative or to any other traffic-reducing mechanism.

Evidence on the magnitude of potential benefits from traffic limitation is variable. Some early studies suggested considerable potential but made over-optimistic assumptions about potential traffic speeds in even a congestion-free urban environment (15, pp. 47-60). Recent thinking, at least in the United Kingdom, is that only benefits of moderate size (say tens of millions of pounds per year, rather than the hundreds of millions considered earlier) might be available. In such circumstances, the costs side of the cost/benefit analysis needs detailed consideration—a blanket case for road pricing does not appear to exist.

Furthermore, one of the principal components of the benefit element in congestion pricing is the time saved by allowing faster journeys. But, such calculations assume that time savings, independent of their size, may be evaluated at the same unit rate. That this is the case is by no means clear. A recent study on the value of time (17) considered in some detail the question of the evaluation of small time savings. Some empirical evidence (18-20) appears to suggest that constant marginal valuations of travel time savings cannot be supported, but this has been disputed (21). The main support of nonconstant marginal valuation are that (a) small time savings may not even be perceived by the beneficiaries and (b) even if perceived, they may not be of as much use as larger time savings. Certainly both of these arguments are intuitively plausible, and, on balance, the conclusion of Voorhees and others (17) is that there is no theoretical reason to assume constant marginal values of time. Whether over a reasonable range valuation is approximately constant is an empirical question that does not appear to have been tackled. The correct evaluation is a matter for the scheme beyond the topic of urban highway congestion. Nevertheless, the doubts that surround it again serve to undermine the conventional argument that the removal of congestion is clearly a good thing.

In addition to the empirical doubts about the support for policies aimed at removing congestion, questions have also been aimed at the theoretical basis of the conventional argument. In a recent paper, Else (10) suggested that the conventional analysis incorrectly defines the marginal social cost curve. Instead of considering the marginal social cost of an extra vehicle per unit of time to the traffic flow, what should be considered is the marginal cost of adding an extra vehicle to the road or, equivalently, the cost of adding to traffic density. Adoption of this approach suggests that the optimal flow will be greater than that yielded by the conventional analysis. It also follows from Else's approach that an optimum position on the backward-sloping part of the cost curve can be attained.

How well founded Else's criticisms are is at present a matter of controversy. Nash (22) has argued that they are for the most part unjustified, and, in particular, that the redefinition of marginal cost relative to numbers of vehicles rather than flow cannot be acceptable within the conventional theoretical framework of economics since all demands relate to flows, not stocks. Following this argument through supports the conventional position that a social optimum with flow at greater than the maximum is unattainable, and thus it seems that Nash has re-established the authority of the conventional wisdom in this case.

One point, however, where Else's paper does make a valid criticism of established modeling procedures relates to the costs imposed by congestion on following traffic, outside the time period for which the basic analysis is undertaken. The main argument of conventional analysis on a steady state with no recognition of variability in levels of congestion is clearly highly abstract. Moreover, it incorrectly suggests a single optimal level of taxation. By adopting a more dynamically oriented approach, Else is able to show, in some simple cases, that the pattern of optimal congestion taxes should vary through the congested period, because a rise in peak traffic densities, through increasing the length of the congested period, causes delays to off-peak traffic.
Thus, the desirable tax structure is one that has a high level of taxation at the start of the heavy traffic period, which gradually declines as it proceeds.

Brosie is not the first person to criticize the conventional analysis of congestion for too easily suggesting solutions on the basis of too narrow a definition of the problems. Apart from the question of overflows into other time periods, there are also spatial overflows. The conventional analysis is always presented in terms of a single link but, from very early on, this was recognized as dangerous. Walters [7] presented a model that considered highway networks, although he recognized that, at that time, empirical progress with the model was impossible. This was followed by a later paper [23]. More recently, and more practically, Wigan and Bamford [24] have investigated the effects of network structure on the benefits derivable from road pricing. The recent United Kingdom government paper [16] on Transport Policy warned,

The relationship between traffic speeds and volumes and the extent to which other traffic mitigation measures respond to improved travel conditions, will depend on the particular traffic situation in individual towns. Generalisation on the proportionate reduction in traffic volumes required to produce any given congestion savings is therefore unrealistic. Wherever possible project and policy appraisals should be based on a transport study model which simulates the complex interactions between traffic volumes and traffic conditions.

Finally, note that doubts about the practical value of the straightforward conventional theory as a guide to controlling the use of congested highway facilities must inevitably spill over into doubts about investment decisions designed to alleviate congestion. A number of authors [3, 12, 25] have developed models that show that, in circumstances broadly similar to those analyzed previously, the quantity and timing of investment will be incorrect when based on the responses of motorists paying the average private cost rather than the marginal social cost of their road use in congested conditions.

If cost/benefit analysis is used as the investment criterion, in conjunction with average cost pricing, Wheaton [2] argues that investment in roads will be greater than it should be, because a second-best investment criterion should be used with a non-optimal pricing policy. Without this, what results is that less of the cost of road use is paid through congestion (time) and more is paid in money terms (investment). This conclusion is based on the standard steady-state model.

Henderson [25], in a paper that concentrated more on the timing of journeys, and so moved outside the conventional framework, suggested that the optimal level of investment was lower in the taxed than in the untaxed case, more as a result of peak-spreading. In similar vein, Borin’s paper [17] shows that the absence of marginal cost pricing will tend to bring forward the timing of investment programs, all other things being equal. Although the general qualitative thrust of all these papers is perhaps not undermined by some of the worries expressed earlier about the circumscribed range of models conventionally used, clearly their importance from a policy point of view is more open to doubt once the real effectiveness of introducing marginal cost pricing is questioned.

SOME ADVANTAGES OF CONGESTION AS A METHOD OF RESOURCE ALLOCATION

The tenor of the two previous sections has been such as to question the theoretical and particularly the empirical basis for the conventional view that the elimination of traffic congestion is a goal unquestionably worthy of pursuit. There is, however, a school of thought that goes further and puts forward the view that congestion is a positively helpful way of dealing with a resource-allocation problem. A latter set of arguments that states that the elimination of congestion is a useful allocative mechanism, suggesting solutions on the basis of too narrow a definition of the problems.

Transport Policy warned, when based on the responses of motorists paying the average private cost rather than the marginal social cost of their road use in congested conditions.

A second set of arguments that states that the regressive effects of road pricing is that pricing through congestion is likely to be progressive. Time, it is suggested, is distributed more equally than earned income. Thus, the opportunity cost of time (and thus willingness to queue) may be lower for those with lower wages, or no wages. Consequently, time prices act similarly to a tax that is proportional to wages. Thus, queueing can serve to vary the total subsidy involved in the provision of a public good by income class. Provided that the
loss that results from queueing is less than the cost of administering an equivalent means test, congestion provides an efficient means of attaining an equity objective. Matters are not quite so straightforward, however. Barzel (30) demonstrates that, if income elasticity is high and price elasticity is low, it becomes less likely that the poor will be prepared to pay by waiting. At the extreme, this suggests that subsidizing, say, opera, is highly unlikely to have progressive results. For roads the position is less clear—we again have an essentially empirical issue that hinges on the relative magnitudes involved in each case.

SUMMARY

The intention of this paper has been to raise some doubts about the extent to which the control of congestion is understood, not as an engineering problem, but as a socioeconomic one concerned with making the most appropriate use of scarce resources.

Despite some doubts, the basic theory, as applied to a highly simplified situation, seems to be technically correct. What is much less clear, however, is the extent to which the acceptance of this analysis as a basis for policymaking in the real world is justified. Even if, qualitatively, its implications are correct, there are significant quantitative uncertainties. Given that governments, local and national, are still pouring considerable sums of money both through subsidies and investment, into the relief of congestion, it is desirable to change this state of affairs. There seems to be ample scope for the transport economist, the transport planner, and the transport engineer to contribute to a debate that has a long and, in places, distinctly complicated as a basis for policymaking in the real world is justified. Even if, qualitatively, its implications are correct, there are significant quantitative uncertainties. Given that governments, local and national, are still pouring considerable sums of money both through subsidies and investment, into the relief of congestion, it is desirable to change this state of affairs. There seems to be ample scope for the transport economist, the transport planner, and the transport engineer to contribute to a debate that has a long and, in places, distinguished pedigree, but where the outcome is as yet considerably outstripped by the importance and complexity of the problems that must be solved.

REFERENCES