ROAD INDIVISIBILITIES

Some Observations

By D. N. M. Starkie*

This paper critically examines the belief that the provision of road capacity is subject to pronounced lumpiness or indivisibilities, so that the road planner's choice of capacity is heavily constrained. This view, expressed in extreme form, will be found in Paul (1960) and Walters (1968).

The first part of the paper provides a brief theoretical treatment of road indivisibilities, illustrating their significance. The treatment differs from previous accounts by first redefining the production function for road services. It then shows, for example, that the presence of indivisibilities is a necessary but not a sufficient reason for efficient road user charges to diverge from long-run marginal costs. In order to focus on the indivisibility issue, it is postulated that there are no barriers to an efficient system of user charges. It is recognised, nevertheless, that there are difficulties, apart from political acceptability, in implementing such a system.

The second part of the paper is an analysis of the extent to which indivisibilities occur in practice. The data used for this purpose relate to rural main roads in Australia.

THE SIGNIFICANCE OF INDIVISIBILITIES

In the following analysis it is assumed that a road of any given size has a rigid capacity. This assumption, though unusual in the economic treatment of road problems, is consistent with the "service capacity" concept used in the planning of roads by highway engineers. This concept relates the output of a road to a specified quality of travel. Quality is expressed in terms of several factors, including speed, traffic interruptions, safety, freedom to manoeuvre and driving comfort. In practice, a number of different quality bundles, known as levels of service, are specified, and emphasis is placed on a particular limiting value of travel speed. The limiting values for one level of service differ for different basic types of roads.

From the economist's viewpoint, each level of service will have a separate demand

* Professorial Fellow, Department of Economics, University of Adelaide. I am grateful to an anonymous referee for valuable comment on an early draft of this paper.

function. The investment decision therefore becomes a matter of choosing the quantity-quality bundle of road supply that will maximise welfare (Docwra and Kolsen, 1982). In this context the level of output, beyond which it is impossible to sustain the quality of travel demanded (in practice the limiting speed), becomes the service capacity. The limit is a rigid one, and diagrammatically is indicated by a vertical curve corresponding to a given amount of road track. This can be regarded as the short-run marginal cost curve. The cost of person A using a road which is at service capacity is that person B is deprived of its use at that quality of service; the cost of use by the first person is the value forgone by B.

This is illustrated in Figure 1, which, to simplify the analysis, shows the demand and cost functions relating to one level of service (or quantity-quality bundle). It is further assumed in Figure 1 that there is an existing road with a service capacity fully adjusted to the level of demand D_i.[1] The user charge is at a level, therefore, which covers total use-related costs $OZLV_i$ (these costs of wear and tear and traffic supervision are assumed to be incurred at a constant rate per unit of output) and provides also additional revenues ZP_iAL sufficient to cover capital charges on, and non-use-related costs of, the existing road.

Now suppose that a once-and-for-all increase in demand (at the same level of service) is forecast, and that this demand forecast is shown by the demand curve D_e.[2] With service capacity limited to that of the existing road, user charges will rise to P_e and produce a surplus in excess of total current costs and capital charges. The efficient solution will be to expand the capacity of the existing road.

The line Q–R shows the path of long-run marginal costs if existing service capacity is expanded. In this context these costs are the additional use-related costs plus the capital costs of expanding output at the margin. Use-related costs per unit of output are assumed to be the same for the new capacity as for the old; capital costs are expressed as an annual equivalent charge. A further assumption is constant returns to scale.

Under these conditions, optimal service capacity is V_1 where the road user charge equals $SRMC_1$ and long-run marginal cost, and the optimal *addition* to capacity is given by $V_1 - V_i$. In this new equilibrium position the costs of expanding capacity will be covered by revenues from user charges (net of current costs). Total revenues are equal to the area OP_1XV_1; use-related costs equal the area $OZYV_1$.

Suppose, however, that there are *technological* constraints on expanding road capacity, so that capacity can be provided only in large, fixed amounts. In other words, road expansion is characterised by technical indivisibilities or lumpiness. In Figure 1 this is illustrated by supposing that the existing road with a capacity V_i *has* to be expanded to produce at least the level of potential output, V_2. In these circumstances Q–R ceases to be the long-run marginal cost curve. Instead Q–R denotes the costs of expansion when averaged over the increment of additional capacity.

[1] We also assume here a single uniform demand period. For a similar analysis which covers fluctuations in the demand cycle, see in particular Williamson (1966).

[2] Again to simplify the analysis we assume that the forecast is not subject to a probability of error. For an extension of this analysis incorporating choice of optimum capacity under conditions of an uncertain demand, see Rees (1976, 143–165).

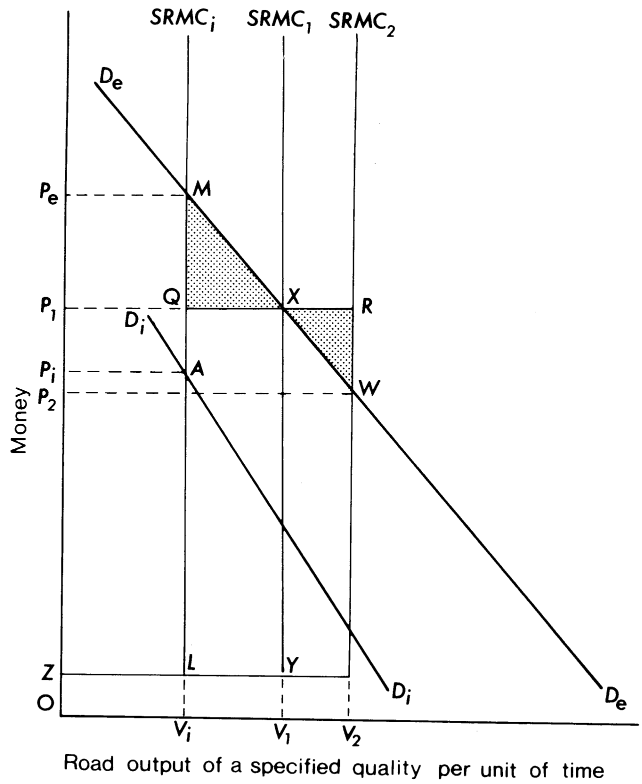

FIGURE 1

Demand and Cost Functions and Indivisibility

With an expanded road of capacity V_2 and level of demand D_e, user charges are given by the intersection of $SRMC_2$ and the demand curve. This level of charge, V_2-W, provides a contribution towards the costs of expanding capacity, but it fails to cover total expansion costs.

Nevertheless, expansion of capacity can be justified by surplus criteria. The additional benefits of expansion are given by the area P_eMWP_2, comprising gains in consumer surplus. These gains are offset by losses of producer surplus P_eMAP_i + $P_iAQXRWP_2$. The net change in welfare is $MXQ - XRW$ (the difference between the shaded areas in Figure 1), which in this case is a positive increase in consumer surplus.

Thus the key nature of an indivisibility is illustrated. The presence of indivisibilities is a necessary though not a sufficient reason for efficient user charges for a road to diverge from long-run marginal costs. If extra capacity is infinitely divisible, the efficient investment solution is always to expand the capacity of a road till the user charge, short-run marginal cost and long-run marginal cost are equal.

EVIDENCE ON LUMPINESS

The service capacity of a road is determined by a number of design features. The feature most evident, and most frequently spoken of, is the road lane; Walters (1968, p. 41) refers to this as the main quantitative unit of highway output. The number of lanes can have a very substantial effect on highway capacity, and it is probably the widespread knowledge of this that has enhanced the belief that road capacity is a very lumpy form of investment.

There is less appreciation of the many other road features that determine service capacity. These include the width of a lane, the lateral clearance from obstructions alongside lanes, the surface of road shoulders, horizontal and vertical alignments, and the provision of auxiliary lanes. The effect of these and other factors on capacities is still a subject of research. Nevertheless, enough is now known to enable the road engineer to select from a detailed range of service capacities without resorting to the more radical measure of increasing the number of lanes, and hence the basic type of road.

Studies have shown, for example, that with wider lanes vehicles have less effect on the traffic flow in adjacent lanes; this increases capacity. However, the ability of the road planner to vary capacity in this way and yet maintain a particular level of service is limited. Lanes that are too narrow or too wide will cause hazards inconsistent with the specified service level. Nevertheless, there remains a substantial opportunity to home-in on a particular capacity by adjustments of this type.

Capacities are also affected by lateral clearance between the edge of the running lane and obstructions, such as poles or the sides of bridges or culverts. Any obstruction within two metres of the carriageway encourages vehicles to move towards adjacent lanes, thus reducing opportunities for overtaking. The effect increases as the clearance is reduced, but is less for continuous obstructions than for those that occur intermittently along the roadside.

Normally the provision of shoulders alongside the carriageway will have the effect of removing obstructions lateral to the lane edge. Shoulders alone will not as a rule affect capacity, except where the provision of a shoulder has a bearing on the surface adjoining the carriageway. Paved shoulders allow for marginally higher capacities than do gravel shoulders. That is because with gravel shoulders vehicles are more inclined to avoid the edge of the carriageway, thus reducing the effective lane width.

Horizontal and vertical alignment is of substantial importance to road capacity. Although the specification of a particular level of service means that a road is designed so that a vehicle can travel at a speed (the design speed) without being exposed to hazards from impaired sight distance or severe gradients, specification of a design speed is not enough to allow for the full effects of alignment. So the highway engineer again has the option of selecting various service capacities by designing the road to different standards of sight distance and gradient. Gradient is particularly important where in hilly terrain heavy vehicles form a significant component of the traffic flow. Heavy vehicles have a disproportionate effect on capacities where gradients are severe.

A course sometimes adopted to meet the last point is to introduce auxiliary lanes on the steeper sections of a road. Recent research (Hoban, 1982) has shown that short stretches of auxiliary lanes can have a substantial impact even along level stretches of two-laned highway. Hoban's results suggest that, where there is a systematic provision of auxiliary lanes, "the two and a half lane road" (to use his phrase) may be a practical and effective alternative to both realignment and duplication.

These points can be illustrated by reference to data in the *Highway Capacity Manual* of the Highway Research Board (1965) and its derivatives, such as the *Guide to Traffic Engineering Practice* of the National Association of Australian State Road Authorities (1976). These data allow for the effect of different features on various service capacities of a hypothetical road section operating under ideal conditions. The drawback of using them is that without reference to actual project costs there is no information to show that all design options lie on the production frontier.

To circumvent this problem, Figure 2 has been compiled from three case studies for which comprehensive data were available on the costs of different designs of roads. All cases relate to existing two-lane main roads in rural areas. Roads of this type form the backbone of the Australian network of rural arterial roads. Between 1974–75 and 1978–79 that network absorbed 30 per cent of total expenditure on construction and reconstruction of Australian highways.

The first study (Hoban, 1980) shows the choices of capacity available by constructing one, two or three lengths of auxiliary lane (each 2 kilometres long) on both sides of an existing 21-kilometre highway. In this instance, capacity is defined by reference to the Highway Research Board (1965) level of Service "A". The existing road has a high design speed, wide lanes and shoulders, and generally wide lateral clearances alongside the road. The terrain is flat, and sight distances are generally good.

The second study (Hoban and Fawcett, 1981) considers the choice of capacity provided by auxiliary lanes (designed to different speed standards) and short sections of duplicated carriageway. The original study included 14 options which avoided duplicating the total length of highway, but examination of cost data showed that nine of these were not on the production frontier. The road section studied was 9 kilometres in length and, in contrast to the previous case, included some severe gradients and generally poor sight distances. Capacity is measured with reference to service level "B".

The third study (Finch, 1980) introduces additional design options in the form of different pavement and shoulder widths as well as realignment. The original study included 13 options, but seven of these were not on the production frontier. The horizontal and vertical alignment of the existing highway was poor, and the pavement width only five metres. The choices shown in Figure 2 for this highway are for capacity measured at service level "C".

Figure 2 shows, for the different design options, plots of service capacity (read off the vertical axis) against the proportion achieved by each design of the maximum service capacity of a two-laned road operating under ideal conditions. This estimated maximum capacity is derived from National Association of Australian State Road Authorities (1976), which in turn is based on Highway Research Board (1965). Service capacities of all the existing highways are less than half the ideal capacity. The designs offer useful ranges of optional capacities. For example, the nine-kilometre highway has an existing capacity at service level B of just over 200 vehicles per hour, or 24 per cent of the ideal for a two-laned road. The designs for this road offer options of 36, 47, 51, 67 and 117 per cent of the ideal. The last is achieved by duplicating a

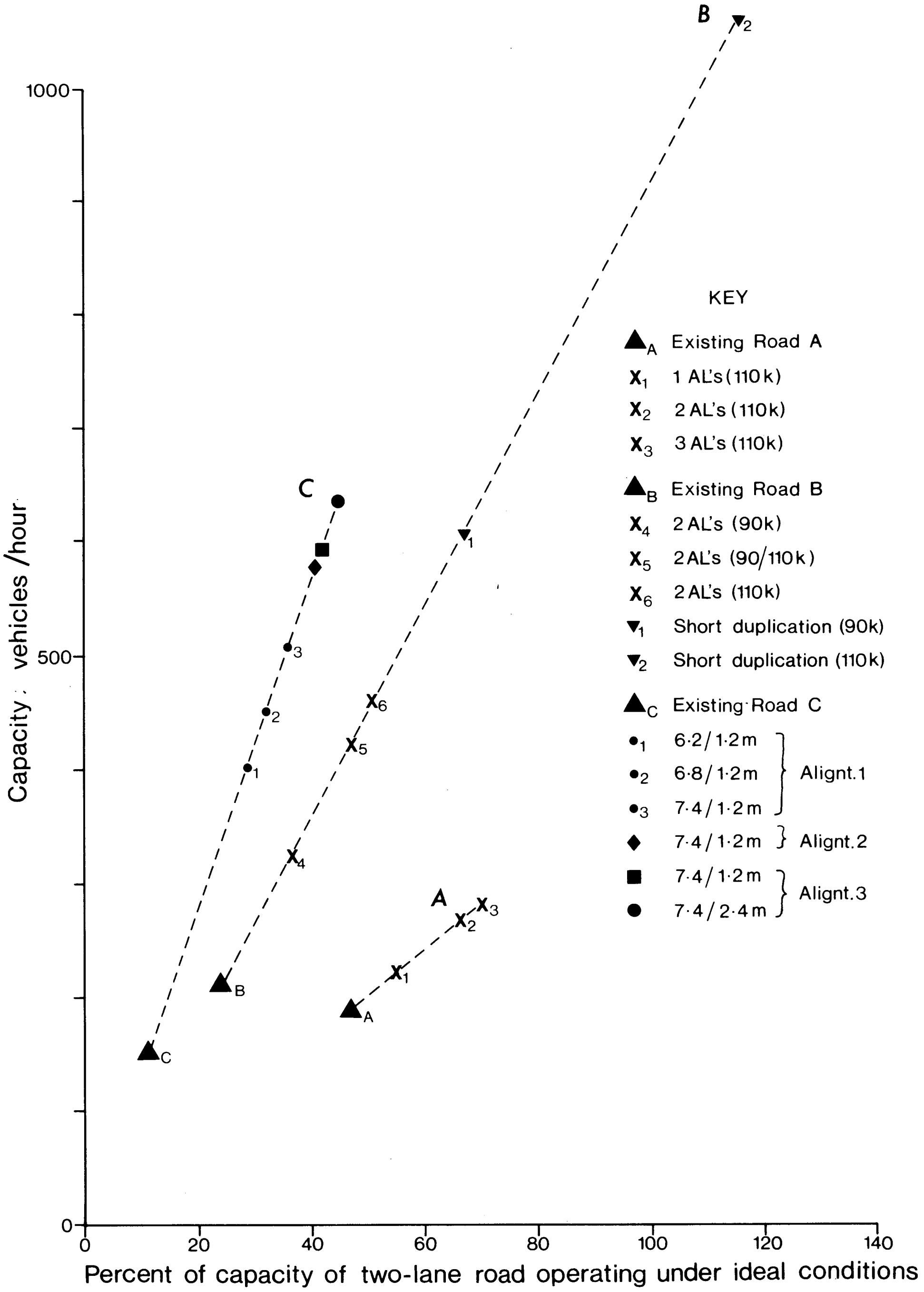

FIGURE 2

Capacity Increased by Auxiliary Lanes and Realignment third of the existing highway to a design speed of 110 kph.

Occasional evidence of this nature has been used to argue a case for staged construction of highways. It is not the purpose of this paper to adopt a view on this question. In contrast, its chief purpose is to consider whether, at a particular point in time, the road planner does have a large number of capacity options he can choose from, or whether he is unduly constrained by technical lumpiness. The data in Figure 2 suggest that with the Australian rural network the planner is not unduly constrained in this way.

CONCLUSIONS

Most paved highways have a basic configuration of two lanes, a few of four or more; in the rural parts of Australia, some have a single bitumenised lane. The potential output of roads with these fundamentally different characteristics *can* vary considerably, and this has encouraged the view that roads are an example of marked indivisibilities. The inclination in the economics literature has been to emphasise options for road capacity in terms of lanes.

Sections of road, however, need not have a constant number of lanes, and lanes need not be constructed to produce only one level of output. Road output is determined by many characteristics of the lane (its width, gradient, curvature) and their interplay with the general road and traffic environment.

If large indivisibilities do occur they will be concentrated at the lowest and highest ranges of output. The difference between no road and the most basic form of road may constitute a large technical indivisibility. At the other end of the scale, with freeways, it may not be technically possible to select from a range of capacities. The choice is more likely to be between, say, four and six lanes, giving two very different levels of output. But most *practical* choices come in the context of expanding the capacity of an existing two-lane road. As the preceding data show, the choices here are varied and need not be constrained to choosing between large differences of potential output. Consequently, for the Australian rural network at least, it is difficult to accept the view of Walters (1968, p. 40) that "indivisibilities in roads [are] of far more moment than in other areas of economic activity".

This conclusion, that rural road indivisibilities are generally small, has implications for strategies of road pricing and investment. In these circumstances Millward (1973) has suggested choosing capacity in such a way that it is fully used, so that revenues cover long-run costs and pricing signals alone inform investment decisions. An offset against the small loss in efficiency from adopting this policy (caused in effect by the small indivisibilities) will be the less evident gains in efficiency from avoiding the detailed and difficult analysis required to measure changes in consumer surpluses. Given an efficient mechanism for pricing roads, this approach appears worthy of further consideration.

REFERENCES

Docwra, G., and H. M. Kolsen (1982): "Track Costs, Road Pricing and Cost Recovery". In Grenning M. R., D. N. M. Starkie, and M. M. Starrs, (eds.): *Pricing and Cost Recovery in Long Distance Transport.* Martinus Nijhoff, The Hague.

Finch, C. H. (1980): "Capital Cost Implications of Alternative Pavement Widths". In *Proceedings, Workshop on Economics of Road Design Standards*. AGPS, Canberra.

Highway Research Board (1965): *Highway Capacity Manual*. Special Report 87. Washington, D. C.

Hoban, C. (1980): *Overtaking Lanes: A Case Study of the South Gippsland Highway*. Civil Engineering Working Paper, 80/17. Monash University.

Hoban, C. (1982): "The Two and a Half Lane Rural Road". *Proceedings, Annual Conference*. Australian Road Research Board, Melbourne.

Hoban, C. J., and G. J. Fawcett (1981): *A Simulation Study of Alternative Road Improvement Strategies*. Australian Road Research Internal Report AIR 840–1, August, 1981.

Millward, R. (1971): *Public Expenditure Economics*. McGraw-Hill, London.

National Association of Australian State Road Authorities (1976): *Guide to Traffic Engineering Practice*. Sydney.

Paul, M. E. (1960): "Covering Costs by Receipts". *Bulletin of the Oxford University Institute of Statistics and Economics*, 22: 300–311.

Rees, R. (1976): *Public Enterprise Economics*. Weidenfeld and Nicolson, London.

Walters, A. A. (1968): *The Economics of Road User Charges*. Occasional Paper 5. International Bank for Reconstruction and Development, Washington, D.C.

Williamson, O. E. (1966): "Peak-Load Pricing and Optimal Capacity under Indivisibility Constraints". *American Economics Review*, 56, 810–827.