The Public Enterprise Concept and Road Supply

George Docwra

INTRODUCTION

The purpose of this paper is to provide a fairly broad sketch of some of the implications of the application of the public enterprise concept to the road supply industry. Specifically, we will be concerned with the issue of guidelines for price, output and investment policy. As in the paper by Kolsen (1975) the assumption is made that the principal objective governing the behaviour of public enterprises is that of efficiency in resource use. Where other objectives are deemed to be important - such as the granting of subsidised services to particular groups of consumers - it is assumed that these should not be financed by internal cross subsidisation, but instead, by direct grants from the government to the public enterprise. By adopting this approach the community is placed in a position where it is able to ascertain with ease the financial costs involved in using public enterprises to achieve income redistribution objectives.

As far as the road supply industry is concerned it is fairly clear that this industry does not exhibit all of the characteristics of a public enterprise. While road space fits the 'natural monopoly' category; is supplied in the main by government (the proportion of total road output supplied by private firms is inconsequential), it is not sold in the market in the same sense as say, for example, electricity, gas and telephone services. Certainly the user of road space is confronted with charges imposed by government which affect either his decision to purchase a vehicle (or particular type of vehicle) and/or his decision as to how much road space to consume (i.e. number of trips) but these charges, which usually take the form of vehicle registration fees and petrol taxes etc. bear little relationship to the demand for road space at different time periods and cost of supplying road space of varying quality in various locations. In other words, the present methods of financing road space provide little (if any) guidance to the important and related problems of: (i) achieving optimal use of existing road capacity i.e. the short run problem and (ii) the long run problem of determining the optimal quantity/ quality of road space in various locations.

Traditionally, in Australia and elsewhere investment decisions affecting the road supply sector have been, and in most cases still are, based on a mixture of political, historical and technical criteria. It is only in recent years - in Australia's case, since 1969 - (that attempts have been made - mainly at the Australian Government level)[1] to evaluate alternative road investment programmes in terms of their economic benefit/cost characteristics. While such a move is clearly a step in the right direction there are nonetheless a number of changes, especially on the pricing side which could be made, the effect of which would be to greatly improve not only efficiency of resource use within the road sector but also between the road sector and the rest of the economy.

---

1. The 1969 Commonwealth Aid Roads Act represented a significant departure from previous Aid Roads Acts in that it provided, to a large degree, for the allocation of Commonwealth road funds to the States on the basis of economic benefits and costs. In previous years road grants were allocated to the States according to the formula 1/3 area; 1/3 population and 1/3 motor vehicles registered. The basis for the 1969 Act is to be found in the 1969 Report of the Commonwealth Bureau of Roads.

Given that the purpose of this paper is to treat the road supply industry as a public enterprise we will be concerned, as noted above with discussing (i) the principles on which prices for the use of road space should be determined, given the objective of efficiency in resource use, and (ii) the relationship between pricing policy and investment criteria.

All of those who are familiar with the literature dealing with optimal price, output and investment policy for public enterprises in general, and road space in particular, will be aware of the lack of complete agreement among economists as to the nature of 'appropriate' price and investment guidelines. At the general level the prescriptive advice tended by writers such as Dupuit (1844), Lewis (1949), Coase (1946), Little (1960) and Hazelwood (1950), for example, differs from that offered by writers such as Hotelling (1938), Vickrey (1948) and Williamson (1966) to name but a few. Likewise differences are found in the literature on road track pricing and investment policy. For instance the pricing recommendations contained in Walters' World Bank study (1968) are quite different from those contained in the UK Ministry of Transport Road Track Cost document (1968).

In part these differences in prescriptive advice are a result of differences in assumptions made with respect to constraints on pricing and/or investment policy. The Ministry of Transport document, for example, makes certain assumptions as to what is possible, from both a practical, and one suspect political point of view, while the rules which emerge from Walters' discussion of road price and investment strategies are intended as representing the theoretically correct (in the Classical tradition of the 'least constrained' model), rules against which actual and alternative policy proposals should be evaluated. Aside from differences in constraints there are, however, differences in rules which are of a more fundamental nature. Here we refer specifically to arguments as to whether the view that "the tub should stand on its own bottom" has anything to do with the efficiency in resource allocation criterion or whether it is essentially an equity and/or political issue. In addition there are other important differences in argument associated with the use of the consumer surplus criterion for investment decisions; the interpretation of the 'bygones are bygones' argument and the implications of cost complexities such as joint and common costs, 'lumpiness' and non renewable and specific assets, for the determination of optimal pricing policy.

No attempt will be made in this paper to critically review and classify the various approaches to the discussion of road track price and investment policy. Instead we will commence our discussion in the next section with a short outline of the relationship between optimal price and investment policy for the road supply industry, assuming that road space is supplied under competitive market conditions - or more "realistically" supplied by a government monopolist who is expected to behave as if road space were supplied under conditions of perfect competition. In other words, behave in the public enterprise tradition.[1] The purpose of this section is simply to make quite explicit the link between price, output, and investment decisions under competitive conditions since the competitive model forms the basis of much of the theoretical analysis of the'right' pricing policy for real world public enterprises. In the following section we consider, albeit briefly, Walters' model - since the

1. As asserted by Bonbright, for example, the view that the supply of particular commodities or servides should be treated as public utilities "... implies that the (outputs) should be offered for sale instead of being given away and that the sale prices should bear a fairly definite relationship to cost, or to cost plus a fair return typically well below the point of monopoly profits. In other words the so called 'theory of public utility rates' already starts with certain presumptions about the relevant principles of rate determination". Principles of Public Utility Rates, (Colombia University Press) 1966, p.26)

approach here is generally accepted as typifying the conventional wisdom on road track pricing and investment policy. The fourth section offers an alternative approach, still in the context of the 'least constrained' environment and which departs rom the conventional wisdom in at least three important respects, namely, (i) with respect to arguments relating to 'covering' the entire costs of road supply (ii) on the matter of the importance of indivisibilities in road supply and the use of the consumer surplus criterion as the appropriate criterion for evaluating alternative road investment options, and (iii) with regard to the treatment of joint costs. These issues are inter-related, and the discussion of them is intended as an extension of the arguments advanced in the paper by Kolsen (1975).

Following this the fifth section is devoted to a discussion of pricing rules for road space in the context of various constraints. In particular we consider some of the implications of the principles developed in the third section for the pricing of road space given that it is not possible, for whatever reason, to use sophisticated charging devices (in the third section it is assumed that such devices are available and are inexpensive), and that in effect we are forced to rely more or less on existing road user charges e.g. vehicle registration charges and petrol taxes as the means of financing road supply. In this "more practical" framework we ask specifically how a given road budget (exogenously determined) might be more efficiently raised from a resource allocation point of view, given existing road user taxes. While the argument of this section offers a solution which is a far cry from what is conceptually the 'first best' it is certainly a great improvement on current practice.

ROAD SUPPLY AND THE PERFECTLY COMPETITIVE SOLUTION

We begin our discussion of the public enterprise concept and road supply in the standard textbook fashion by assuming that road space is supplied (i) by a single public authority (ii) that the industry exhibits constant returns to scale (iii) that inputs are perfectly divisible (iv) that there exists one (homogeneous) class of road users (v) that the rest of the economy is made up of perfectly competitive industries (vi) that the road supply industry is required to simulate the behaviour of competitive industry[1] and (viii) that the road supply industry is a multi product firm (as indeed are most public enterprises) capable of supplying different quantities/qualities of road space in various locations. At one end of the spectrum it is capable of supplying low quality/low capacity road space while at the other it is able to supply high quality/high capacity road space. Between these extremes we suppose that there are a very large number of quantity/quality combinations.

Like any other firm the road authority will be confronted with two main problems. In the short run period (i.e. when not all inputs are variable) it will be concerned with the problem of how to make best use of existing road capacity, while in the long run it will be confronted with the task of determining optimal adjustments to capacity/quality. Before outlining the nature of the adjustment mechanism by our hypothetical road authority we briefly note the costs involved in the supply and use of road space. First, there are those costs which are incurred by the supplier. These, in turn, maybe subdivided into (i) road track costs and (ii) maintenance

1. See for example the discussion by Mohring, H. and Harwitz, M. (1962). <u>Highway Benefits: An Analytical Framework</u> (Northwestern University Press), pp. 80-87.

costs. Second, there are those costs which are imposed by users on other users as well as costs imposed by users on non users. As far as track costs are concerned these are fixed or 'sunk' costs which the road authority incurs when it builds roads. Specifically they relate to "any long term contractual (capital) commitments - such as the purchase of land, the laying down of track, etc. Once committed they are inescapable except in the very long run." (Walters (1968), p. 23)[2]

Regarding maintenance costs two categories are identified. There are those which are imposed by the user on the road authority and are a function of traffic volume and composition. Given the same traffic volume and composition (at precisely the same time) these costs will vary from one part of the road network to another according to variations in the technical quality of the road network. In addition there are those maintenance costs which are invariate with respect to traffic volume and composition. Instead they are influenced by such factors as time, and variations in climatic and weather conditions.

As far as the second main category of costs is concerned we refer to (i) road user costs (ii) congestion costs and (iii) community costs. In the first group we place those costs which the road user incurs as a result of his decision to use the road. These costs include fuel and tyre costs; wear and tear on the vehicle and so on. Congestion costs (an externality, but one which is internal to the industry) represent those costs which are imposed on road users as a result of additions to the traffic flow. Given the width of a road, together with other technical characteristics, such as road

2. A detailed classification of road track costs is to be found in: Haritos, A (1973). Rational Road Pricing Policies in Canada, (Ottawa, Canadian Transport Commission)

curvature, there is some volume of traffic for which it is usually referred to as the 'free speed' situation. As the volume of traffic is increased additional vehicles impede the movement of other vehicles and as a result cause an increase in time and operating costs. Finally, community costs represent those costs which are imposed by road users on the community in general. They take the form of noise costs, pollution of the atmosphere by motor vehicle exhaust fumes, loss of amentity and so on. From the point of view of what follows we assume that such costs are internalised, or don't exist, and that the only costs relevant to the analysis are track costs; maintenance costs and congestion costs.

We are now in a position to consider the behaviour of our "competitive" road authority in both the short run and long run period. Following other writers we will direct our attention to a particular section of the road network and assume, to begin with, that the short run situation is as shown in Figure 1.

Along the abscissa we measure traffic volume in terms of vehicles per hour (v.p.h.) while along the ordinate we show costs and price in terms of cents per vehicle kilometre. Short run variable maintenance costs are assumed to be constant and are depicted by the line BB'. Invariate maintenance costs do exist - although they are not shown in the diagram. The curve CC' represents variable maintenance and user private cost function. Beyond traffic volume OX_0 user operating costs increase as traffic volume increases and vehicles impeded one another. Traffic volume reaches a maximum at OX_3. If the density of vehicles is increased further the actual flow of vehicles will decrease. This is shown in the diagram i.e. once OX_3 is reached the CC' cost function "bends backwards". The marginal social cost function which incorporates the effects of congestion (as manifest by increased time and wear and tear costs etc.) is described by the curve CC". Finally, the line DD represents the demand function for use of the road per period of time.

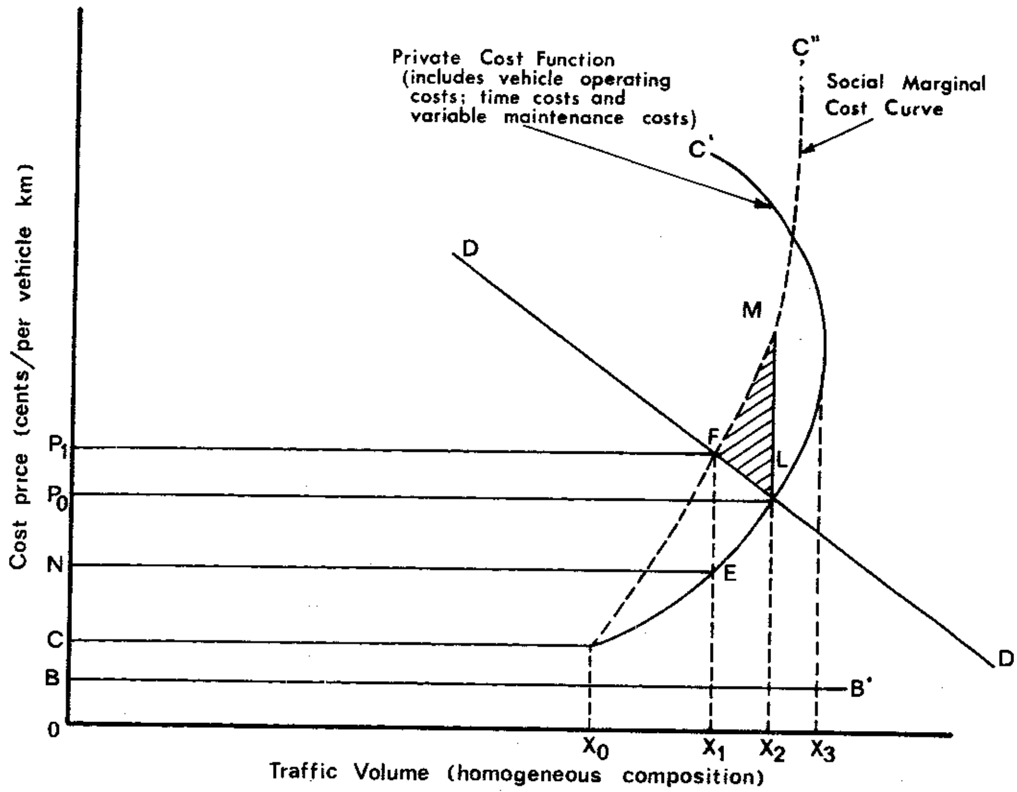

FIGURE 1

As can be seen from the diagram the benefit maximising output would occur at that point where the demand function intersects the marginal social cost curve i.e. where price = p_I = short run marginal social cost (srmc). The optimal traffic flow is thus OX_I. At a price output combination given by the intersection of the curve CC' with the demand curve DD, benefits to the marginal user as measured by P_o are less than the costs imposed by the marginal user on all other users as shown by LM. In this case output is too large. Similarly, at a price greater than P_1 benefits to the marginal user are greater than social marginal costs - indicating that there are some users excluded from using the road who value (in terms of their willingness to pay) the use of the road more highly than the costs which their use would impose on other users and the road authority. In this case output is too small. In other words, it is only where marginal private benefits equal marginal social costs that net benefits are maximised. Whether or not the road authority will retain, in the short run, the quality of the road network will depend on the relationship between revenues collected from road users and those short run maintenance costs which could be avoided by preventing use of the road. In terms of Figure I, revenues collected are obviously greater than short run variable maintenance costs as shown by the area $NEFP_I$. Whether such revenues are sufficient to meet short run invariate maintenance costs is another matter. Assuming they are, then the quality of the road will be retained. On the other hand, if revenues are less than total short run maintenance costs, but equal to short run variable maintenance costs then the road authority would be expected (in the absence of any compelling non economic factors) to allow the quality of the road to deteriorate. At the lower limit if revenues are less than total short run avoidable costs then no further expenditure would be incurred.

So much for the nature of the short run adjustment process. Let us now consider the process of adjustment in the long run period. This is illustrated by way of Figure 2. Here we show the short run marginal cost function (excluding short run time related maintenance costs) for a particular part of the road network as being constant until at some volume of traffic (e.g.X_2) the curve becomes vertical - depicting a rigid capacity constraint. The curve C' represents the short run marginal cost function (srmc) of road I which has a maximum capacity, X_2, while C" represents the srmc function associated with a road having a maximum capacity of X_3. The long run marginal cost function is described by the line b + c and represents the long run costs associated with producing a given level of output (i.e. traffic volume). For convenience it is assumed that there is only one demand period represented by any one of the demand functions D^I to D^{IV}. Suppose, that initially, capacity is given by X_2 and demand D^{III}. Optimal use of capacity is obtained at a price equal to b + c. At this price/output combination price = srmc = lrmc and a normal competitive rate of return on capital is achieved. The road

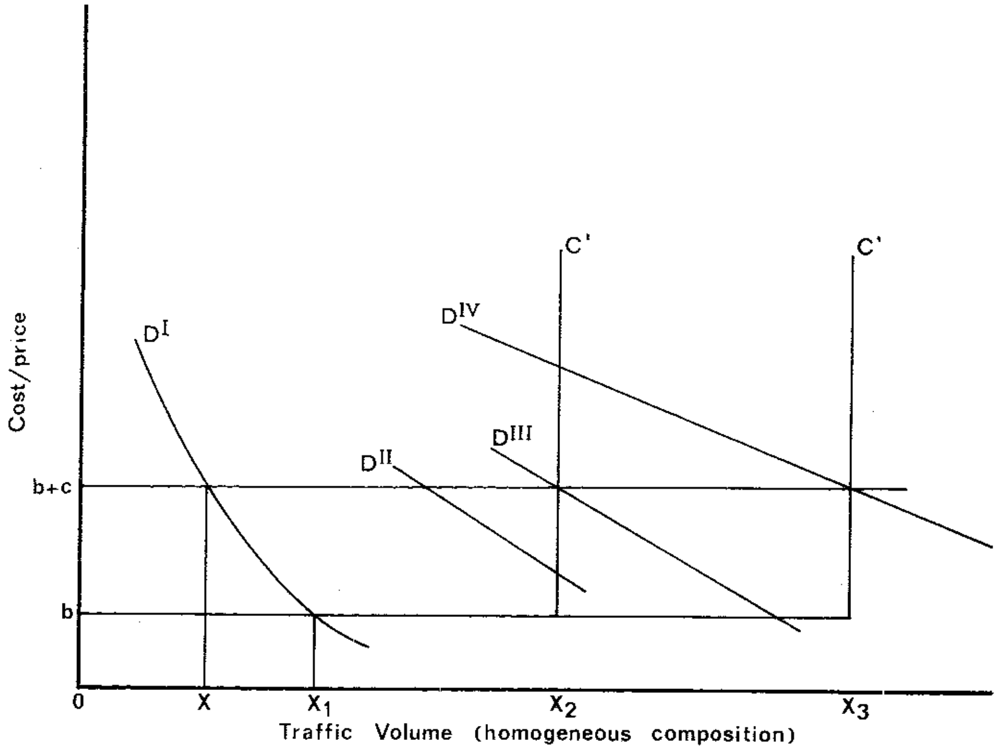

FIGURE 2

authority is in long run equilibrium - there being no incentive to either expand or contract capacity. It will also be observed that at this price/output combination the road is 'congested' - representing, however, a level of congestion which is optimal (in the long run) for the particular road in question.

If we now assume that demand is given by either demand functions D^I, or D^{II}, it is clear that a road of capacity X_2 is of more than optimal size. For each demand function the revenue obtained from setting price equal to srmc is less than lrmc. In the long run the road authority will reduce capacity until such time a capacity/quality combination is reached at which revenues from optimal charges equal long run marginal costs. Thus for demand function D^I capacity would be reduced from X_I to X. In contrast, if the demand function is depicted by D^{IV}, revenues derived from an optimal charging policy exceed the long run costs of producing a road having a capacity given by X_2 i.e. price = srmc > lrmc. Such a situation is clearly a signal for the road authority to expand capacity, which given the demand and supply conditions assumed, means adjusting the quantity/quality of road space to the maximum output level given by X_3. Given our assumptions of a competitive road supply industry, long run equilibrium for the entire part of the road network will be attained when price/lrmc ratios for each part of the network are equal to unity. Naturally, it is reasonable to expect that there will be a wide range of demand and cost conditions for road space in various locations. Setting the price/lrmc ratios equal to unity does not imply that all roads will have the same long run quantity/quality characteristics.

So far the analysis has assumed a single demand period for each part of the road network. The above argument, however, can be extended to take account of more than one demand period. The same principles apply. For our competitive road supplier the decision to expand or contract capacity is based on a consideration of revenues and costs. However, there is one modification which needs to be made to the argument. In the above discussion we had argued that for a single demand period an increase in the supply of road space would take place when demand (or price) is greater than lrmc. This is also a condition for expansion (under competitive conditions) when there is more than one demand period for the same piece of road space. What we have here is simply the familiar 'peak-off-peak' model. To take simple case of the two demand period model, demand in period I (off peak period) maybe such that the optimal price is equal to short run avoidable costs (no contribution to capital cost is extracted from users). Demand in period II (the peak period), on the other hand is such that the optimal charge results in revenues in excess of short run avoidable costs. Obviously, in this case, the quantity/quality characteristics of the road will be retained in the long run only if revenues from the sale of road space during the peak are at least equal to capacity plus operating costs. Likewise if expansion is justified then it will be so because revenues from the sale of peak services exceed capacity costs plus operating costs (lrmc).

Similarly, contraction will be justified if peak period revenues are less than capacity costs. (For a detailed discussion of the peak/off peak problem the reader is referred to the contributions of Lewis (1941), Steiner (1957), Hirshlielfer (1958) and Kolsen (1966).)

The main point which we have attempted to emphasise in this section (albeit in a somewhat laboured fashion) is that in the Classical model of "simple competition" the link between price, output and investment policy for the road supply industry, and indeed other producers is quite explicit. Given the objective of revenue maximisation (in this case being consistent with the objective of efficiency in resource use) no producer would set price at less than the costs which could be avoided by ceasing production of output (single product case), or of particular outputs, where more than one product is involved (the multi product firm case). Moreover, it is the relationship between existing and expected revenues and the cost of replacing capacity which determines whether too much or not enough capacity has been provided.[1] Specifically, new investment - in our case, road track investment - will take place in those locations when expected revenues indicate that existing plus additional capacity is able to earn its replacement cost.

So much for the 'competitive' solution. We now direct our attention to a model of the road supply industry which is generally viewed as providing a more realistic account of the economic and technical characteristics of road supply (in contrast to our competitive model) and as such as providing a more acceptable framework for the determination of optimal price and

---

1. See for example, Ponsonby, G.J. (1960) "Earnings on railway capital", Economic Journal, December.

investment policies. While many economists have argued along the lines we are about to consider, the most recent and complete statement of the argument, in the context of the road supply industry, is to be found in the study prepared by Walters (1968). For convenience we will refer to the model as the Walters' model.

WALTERS' MODEL - THE CONVENTIONAL WISDOM

The model of the road supply industry which is advocated by Walters (1968), and indeed most economists, involves an acceptance of the short run pricing rule of the competitive model, but not the prescriptive advice regarding investment decisions. More to the point, it is contended that the optimal pricing strategy for a road authority is to set price equal to srmc while investment decisions will need to be evaluated (in contrast to the competitive rule) on the basis of consumer surplus calculations. The reasons advanced by Walters for this departure from the competitive rule of basing investment decisions on revenue/cost relationships are, it is argued, to be found in the nature of the supply characteristics of the road supply industry. Whereas in the competitive model the assumption is made that road space is perfectly divisible, Walters' model takes a realistic stance by introducing such factors as indivisibilities; economies of scale and joint products - the latter referring, in this case, to the relationship between two dimensions of road space, capacity and quality. Let us begin then, by outlining the arguments relating to these factors.

<u>Indivisibilities, "lumpiness" and Economies of Scale</u>

Apart from reference to 'indivisibilities' associated with the purchase of the materials from which road space of various qualities is constructed (i.e. it is argued that it is cheaper to purchase, in large quantities, materials such as cement and gravel etc.)[1] the principal forms of indivisibilities or 'lumpiness' are seen to arise from technological factors. For instance it is suggested that there are significant indivisibilities on the input side of the road supply industry in the form of equipment and machinery. As a consequence the supply of road space is characterised - over a given range - by economies of scale. Further, there are also indivisibilities on the output side. Here reference is made to the fact that a road must meet 'minimum' technical specifications (i.e. in terms of width) if it is to be of any use to some, if not all vehicles. In other words a road must be at least as wide as the narrowest motor vehicle. To be of use to all types of vehicles it must obviously satisfy a different set of technical specifications. In addition mention is also made of indivisibilities associated with quality improvements to road space. In practice a road authority will not make minute changes to the nature of the surface of a particular section of the road network. Thus it is argued:

> with improvements such as surfacing it is technologically silly to put pavement down 1/100th of an inch at a time. The road authority can pave the whole road to a sensible minimum depth, or it can pave sections of the road one after another. Both involve discontinuous or lumpy improvements. (Walters (1968), p.41)

The upshot of the argument is simply that indivisibilities associated with the supply of road space result in a polarisation of both capacity and quality. Real world suppliers of road space are thus seen to view changes (in contrast to our supplier of our competitive model) in both quantity and quality of road space as having to be made in discrete units. For one thing there is a standard size lane. Roads are normally described as one lane, two lane, three lane and so on, rarely (if at all) is

1. This is really an example of economies of scale.

a 1.5 lane, 2.5 lane road etc. Likewise, insofar as the quality of the surface of the road is concerned the range of options accepted in practice are much fewer than those implied in our model of a perfectly competitive road supply industry. In general a road has either an earth surface, a gravel surface, or a sealed surface. It is of course recognised that it is possible to vary the quality of a road's surface at different sections along its path, in accordance with variations in demand but even so the range of quality variations fall a long way short of those implied in the competitive model.

Accepting - at least in general terms - the indivisibilities are a characteristic of the road supply industry it seems reasonable to ask why this should be a cause for special concern. After all indivisibilities/economies or scale are not uniquely related to the road supply industry - they are features of many productive activities within both the public and private sectors of the economy. The view taken here is that the literature, generally, has not been very clear in its discussion of this matter.

Insofar as Walters' analysis is concerned the claim is made that indivisibilities in the supply of road space are "of far more moment than in other areas of economic activity." (Walters (1968), p. 40) This view is advanced on the ground that while indivisibilities exist elsewhere they may frequently be ignored because of the size of the market.

> In manufacturing industry there are many striking examples of indivisible inputs - blast furnaces, giant presses, the entreprenuer himself - but normally these indivisibilities may be considered negligible in relation to the size of the market. Even large absolute indivisibilities may be ignored. But given the limited local market of roads, even small indivisibilities may be large in relation to the size of the market. (Walters (1968), p.40).

There are at least four aspects of this claim which require some comment. On the one hand to assert that the existence of 'large' indivisibilities in many areas of economic activity other than the road supply industry are 'negligible' in relation to the size of the market, and as such maybe 'ignored' presumably means that over the long run, revenues from the sale of the outputs associated with such highly indivisible plant are at least (assuming that demand has been correctly estimated) equal to or greater than the opportunity costs incurred in undertaking investment of that size. It is argued, for example, that in the case of manufactured goods that "there is a world market which, when goods can be stored may be extended over a lengthy period of time." (Walters(1968), p. 40). In contrast, in the case of road space, "the market is local and momentary". (Walters (1968) p. 40) And further, when we compare road space with other types of public enterprises such as those supplying electricity or water, the existence of excess supply in a particular market "may be sold at some cost, to other markets". (Walters (1968), p. 40). For road space, however, this cannot be done. To reiterate, if it is argued that a manufacturer or supplier of, say, electricity, makes a decision to increase investment, and that such an increment in capacity, by virtue of technological and economic constraints, results in a 'substantial' increase in ability to provide additional units of output; and further, if this increase in ability to supply additional outputs poses no problem, then it does so for the reason that expected revenues from the sales to each of the various markets, must at least be equal to the lrmc of increasing capacity. There seems to be no other acceptable interpretation of Walters' statement that in particular areas of economic activity it is possible to ignore the existence of 'large' indivisibilities.

Our second point concerns the implications of the claim that in the road supply industry indivisibilities are of much more importance than elsewhere. As it stands the meaning of this claim is far from clear. In one sense the argument that road capacity created in one locality cannot be utilised elsewhere is really an argument about specificity rather than an argument about indivisibilities as such. But this is not peculiar to road space. The same is true of railway track capacity, hotel accommodation and so on. The fact that capacity is not fully utilised during all periods (or for that matter during any period) should not necessarily be viewed as a matter for concern or as a factor which is unique to the supply of road space. In some cases the existence of excess capacity will simply be a manifestation of the fact that expectations regarding demand have not been realised i.e. a mistake has been made. In other cases the story will be different: the existence of spare capacity will be consistent (in the long run) with the ability of the enterprise (public or otherwise) to at least earn a normal return on capital. Thirdly, the point also needs to be made that there is a greater degree of substitutability between roads - especially in the urban context - than is recognised by the indivisibility argument. In other words, there are usually more ways than one of getting from A to B and thus a number of options as to the size and location of capacity adjustments.

Finally, the indivisibility argument seems: (i) to unduly neglect, or underate the possibilities of making better use of existing capacity by means of traffic management techniques and (ii) ignore the fact that the minor changes to the capacity of roads (resulting in substantial improvements in quality) can, and are in fact, frequently made. For example, extra lanes can be provided for relatively short lengths of particular sections of the road, such as steep inclines, to enable vehicles to overtake slower vehicles. That there might be substantial economies to be achieved by making 'large' adjustments to capacity is not a necessary and sufficient condition for making such changes. The decision to do so (given the efficiency criterion) must be based on a comparison of benefits and costs.

Our point of contention with the argument of the conventional model is in part a disagreement about the importance of indivisibilities in the road supply sector vis-a-vis other productive activities in the economy, and in part a disagreement about the nature of the ingredients of an efficiency pricing policy. These differences will be made clearer in a moment.

Joint Cost Characteristics

Associated with the argument about the significance of the indivisibilities/economies of scale argument for the problems of the road supplier, is the argument that the provision of road space exhibits joint cost characteristics. For Walters this characteristic is manifest in terms of the relationship between capacity and quality i.e. it is argued that as capacity is increased, so too is quality and that the trade offs between capacity and quality which are implicit in the competitive model of section 2 are much greater than those which are possible in practice. (Of course if pure jointness existed between capacity and quality, then no trade off would exist). To a large degree the discussion by Walters of the joint product characteristics of road supply is comprehended by the discussion of lumpiness and discontinuity. However, there are at least two points which need to be made; and they are: (i) that joint products characteristics occur in a number of ways and (ii) that joint costs are common to both a wide range of publicly and privately supplied outputs. Regarding the former point 'jointness' occurs, first, in the sense that some minimum size plant (road track) is necessary if a number of vehicles (e.g. passenger cars; trucks and heavy trucks) are to be able to make use of the road. By allowing for jointness in this sense we introduce an additional factor, namely, heterogeneous user classes, which is generally overlooked, or incorrectly dealt with in discussion of road price and investment policies.[1] Secondly, there is jointness in the sense that capacity provided for peak demand use is available for off peak use. Services of the road provided during the peak demand periods are clearly not the same as those provided during the off peak i.e. they are different products. Thirdly, jointness occurs because capacity provided to enable journeys from A to B is available for journeys in the opposite direction.

While the road pricing solution advanced by Walters takes account of the second mentioned form of 'jointness' it ignores problems posed by the existence of different user classes.

The road supply industry is not examined as a multi product enterprise - providing road outputs to different groups of users. In order to pursue this and other points raised so far we now bring together the various strands of the conventional model.

The main assumptions are as follows: (i) that the non road supply sector of the economy is made up of perfectly competitive industries and that resources required by the road supply industry are derived from a wide range of industries rather than from one or a few groups in particular. This assumption is a standard one and is made on the grounds that it enables us to ignore the effect of changes in policy in the sector under examination, on other sectors. Problems of 'second-best' are ignored, or alternatively, assumed to be unimportant

---

1. See for example the treatment of joint costs by Meyer et.al., The Economics of Competition in the Transportation Industries. (Harvard University Press, 1964) Chapter IV, especially pp. 69 - 73.

(ii) that functions take the form of those dipicted in Figure 2 (iii) in implicit assumption that roads are used by a homogeneous group of users (iv) another implicit assumption that there are no institutional, political, pricing or budget constraints i.e. the model is in the Classical tradition of 'least constraints' and (v) that road space is characterised (as discussed above) by indivisibilities, economies of scale and joint costs.

The optimal pricing rule which emerges from the model, as noted at the beginning of this section, is that - for the homogeneous group of users - price should always be set on the basis of srmc, which means that if particular roads are congested then price should be equal to congestion costs, whereas if there are no congestion costs then price should be equated with variable maintenance costs.[1] The point of departure with the Classical solution is to be found on the investment side. Here it is argued that whether or not the quantity/quality characteristics of a road should be retained should be determined on the basis of estimates of revenues plus consumer surplus. This is explained in the following way.

Suppose we have short run and long run cost functions (given indivisibilities) for a section of the road network, and alternative demand functions as depicted in Figure 3. If initially, road 1 already exists, and the relevant demand function is D^1D^1, then given that it is argued that price should equal srmc, the question of whether the quality of the road should be retained (supposing that short run invariate maintenance

---

1. It should be noted that the disagreement between this definition of srmc and the definition in the theory of the firm is to be found in the treatment of non-variable maintenance costs. In the theory of the firm these would be included, since if price equals only variable maintenance costs, it is possible to close down the road and save more on costs than is lost in benefits.

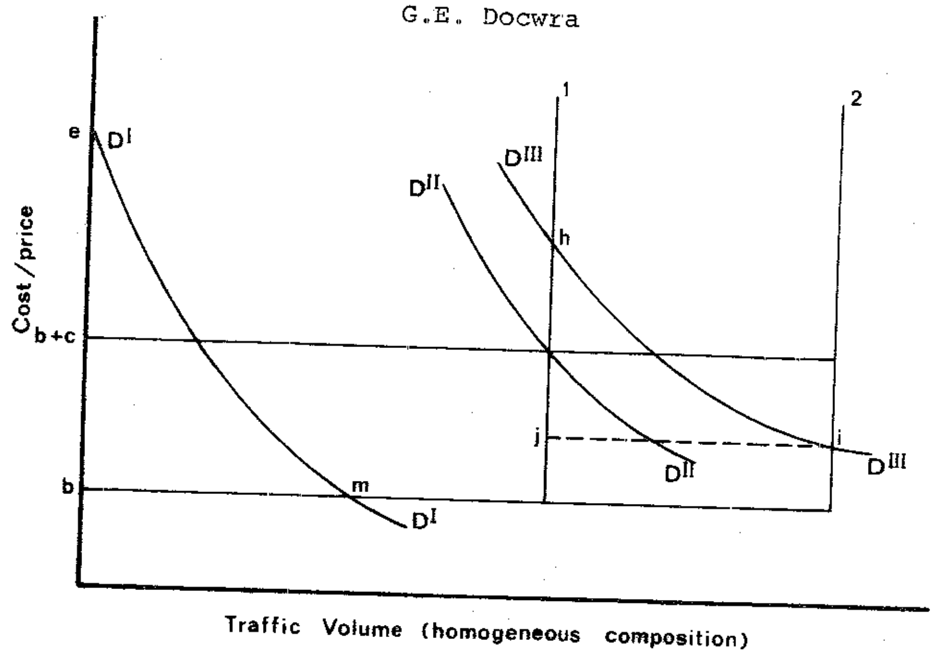

FIGURE 3

costs exist) is seen to depend on whether benefits to users as measured by consumer surplus (emb) are equal to or greater than invariate maintenance costs.[1] Similarly if we start denovo and assume that the choice is between road I or no road at all - the "all or nothing case" - then the question of whether the road should be built, given demand function D^1D^1 and p = srmc = short run variable maintenance costs, will depend on whether consumer surplus estimates are equal to or greater than long run costs. On the other hand if demand conditions are given by $D^{11}D^{11}$ and road 1 is constructed the equilibrium conditions of the

---

1. This is really an example of an extreme "all or nothing case". Walters concedes that "it may be possible to adjust maintenance expenditure to keep the road in various degrees of repair, or to maintain a variable width of road". However, it is also argued that "the consequences of a skimped maintanance program will be manifest in increasing operating expenses and perhaps also in increased variable maintenance expenses" and further, that "these margins need to be balanced one against the other. This reduces but does not eliminate, the need for a consumer surplus criterion when deciding to keep the road open". (Walters(1968), p. 46).

competitive model are satisfied i.e. p = srmc = lrmc and the road authority achieves a normal return on capital for this part of the road network. What happens if demand shifts to $D^{111}D^{111}$? Should an expansion in capacity be undertaken? Again according to conventional arguments[1] the answer is in the affirmative if incremental benefits as measured by the area hij + revenues are greater than incremental costs. If this is the case, an expansion in capacity to road size II will result in the road authority making a loss on this part of the road network. Conceivably, therefore, in long run equilibrium the road authority would be earning less than a normal return on capital. While urban roads are likely to generate surpluses i.e. revenues over and above long run costs this will not be the case for most rural and intercity roads. Specifically, given that optimal pricing policy is defined as setting price equal to srmc there is no reason to suppose that total revenues generated from the system, given correct demand forecasting, will equal total system costs. Moreover, it is argued that there are no a priori reasons why the road authority should attempt to recover the long run costs of providing and maintaining road space.

> whether the surpluses collected from the urban highways will counterbalance the deficits (or, strictly, negative surpluses or rents), on the intercity and rural highways is an arithmetical matter of considerable administrative and political importance. Whether the state raises taxes to finance the net deficit or whether it enjoys a surplus so that it may remit other taxes - these are matters of much economic interest....

---

1. For other statements of this argument the reader is referred to Neutze, G.M., "Investment Criteria and Road Pricing", The Manchester School of Economic and Social Studies, Vol. 34, 1966. Williamson, O.E., "Peak Load Pricing and Optimal Capacity under Indivisibility Constraints", American Economic Review, Vol. 56, 1966. Millward, R., Public Expenditure Economics: An Introductory Application of Welfare Economics, (McGraw-Hill, 1971) Chapter 8.

> What must be emphasised here is that there is no economic rationale for "balancing the road budget"... Each policy must stand or fall by the consequences and not by any abstract obiter dicta. If it be thought that the roads should bear higher taxes than those which emerge from economic analysis - then let the case be aruged in terms of alternatives, such as running a budget deficit, or reducing government expenditure. The balanced road budget is, to the economist, merely a graven image (Walters (1968) pp. 59-60).

In our view there are a number of good reasons why some aspects of the prescriptive advice of the conventional wisdo should be rejected. To begin with we have already expressed our concern with the view that indivisibilities in the supply of road space are of far greater importance than they are elsewhere in the economy; and even if they are, it is still not clear why this should matter all that much. A second point of difference, however, concerns the interpretation of the 'right' relationship between prices and costs and, in respect of costs, the issue of which costs are deemed to be relevant. In Walters' model the 'right' pricing rule is judged as setting price = srmc; the rationale for such a rule being, as argued earlier, the benefits from the use of existing capacity will be maximised. In other words, 'relevant' costs are those costs (given that road capacity already exists) which can be avoided by not allowing anyone to use the road. While there can be no disagreement with the notion that, in the short run, price should reflect short run costs it is far from obvious why, on efficiency grounds, (i.e. as distinct from political and/or administrative grounds) no attempt should be made by the road authority to recover capital costs. To argue that once capital costs have been incurred that "bygones are bygones" is in our view a misinterpretation of the "bygones are bygones" argument. The argument should only apply if the demand for road space has been overestimated - and even then one would argue that an attempt should be made to retrieve some part of the capital costs i.e. if it is possible to do so without significantl distorting resource use. From an efficiency point of view what matters is the manner in which the charging system attemps to recover costs. If too much investment has been made, then setting price = srmc represents the lower limit to an efficient pricing policy. In the absence of congestion this means that price should = short run avoidable maintenance costs (including short run time related maintenance costs).

Support for this argument goes at least as far back as Dupuit's classical work 'On the measurement of the utility of public works' published in 1844. Most writers who have referred to this work (e.g. Hotelling (1938) or, who have, in turn drawn on the Hotelling's exposition) have placed sole emphasis on Dupiut's discussion of the short run, consumer maximisation issue, to the neglect of other aspects of his argument. Thus, for example, in discussing the issue of an optimal pricing policy for a public project, such as a bridge, it is argued

> As the toll increases, so does the utility of the bridge diminish in proportion; it becomes zero when the toll equals 0.fr. 15 at which price no one crosses the bridge; it is therefore possible for the loss of utility to rise to as much as 102,000 francs. Does this mean that there should only be very low tolls or even that there should be none at all? That will not be our conclusion when come to speak of tariffs; but we hope to show that their height needs to be studied and operated according to rational principles, in order to produce the greatest possible utility and at the same time a revenue sufficient to cover the cost and upkeep and interest on capital (Dupuit (1844), p. 40).

And, more recently, in Munby's discussion of the "public enterprise basis" of pricing road track costs:

> The logic of a replacement costs measure of capital is that it represents over time the costs of keeping the system going, which

> consumers should be ready to pay if over time they want this use of resources rather than some other - a choice which becomes a real choice when investment decisions have to be made. If assets are not to be replace, or if technical change requires a totally new kind of asset, then the charge on past sunk capital ceases to have economic relevance except as an indication of what consumers may be persuaded to pay, i.e. the benefit which they obtain from existing assets. Thus, in arriving at a charge performing the function of allocating resources, one seems to be left with the dilemma that either replacement cost... is the right answer, or any charge on capital is arbitrary and meaningless, as bygones are bygones (Munby (1968), p. 165-166).

And, further in evidence submitted to the United Kingdom Select Committee on Nationalised Industries.

> There are two main considerations in relation to pricing policy. Firstly, prices should as far as possible, reflect marginal costs. Secondly, prices as a whole should be such as to cover total costs. These two objectives may well be in conflict; they are not of exactly the same nature. The first objective is concerned with the allocation of resources, where as the second is partly concerned with this, but also involves questions of financial control from a management point of view.
>
> The relative importance of these two criteria depends on the circumstances of the case (a). In so far as there is substitution between different purchases, and there is a degree of elasticity of demand, it is important that prices reflect marginal costs in the nationalised industries, (b) the second criterion (of covering total costs) is important partly as a check on investment, and partly as an incentive to efficiency. As a check on investment, the criterion involves both a backward look on previous decision to see whether they have been rightly taken, and a forward look in relation to present investment decisions. It is important that those who have to make decisions should know whether the public is ready to pay for all the resources involved in producing the commodities in question, but this check is much more indirect than in the case of currently

> committed resources which are repeatedly being used in production. (Munby (1972), p. 588).

Our second point concerns the investment check argument - referred to in the above statement by Munby and by other economists as well, in particular, Coase (1970) and Little (1960). In essence the argument stems from a concern about the theoretical validity of the consumer surplus concept as a basis for evaluating public enterprise investment decisions (again, in the context of the least constrained environment), in addition to problems of a purely practical nature i.e. problems of measurement. While one is not prepared to accept Little's dictum that the consumer surplus concept "is a totally useless theoretical toy" the case for its use in the public enterprise sector of the economy is not as obvious as the theoretical models of the conventional wisdom would seem to suggest. For one thing investment decisions in the private sector of the economy are not based on consumer surplus calculations. Here the revenue cost calculus is the determining factor. If the Walters' approach is adopted then there is more than a strong possibility that (on efficiency grounds) too large a proportion of the community's resources will be allocated to the road sector. Secondly, there is the problem that the consumer surplus criterion does not enable us to determine, after the event, whether or not investment decisions were more or less correct (unless of course one resorts to an expost consumer surplus calculation, assuming that this is easier to do than the former!) In contrast, profitability does yield an objective test - even though profitability might not be an 'ideal' measure of the efficiency consequences of investment decisions.

> ...Within very wide limits, it would be anyone's guess as to whether the consumers' - surplus criterion would be satisfied or not. It will, of course, be objected at this point that guesswork is, in any case, impossible to avoid. Before the event, profitability is only guesswork. Surely, it will be said, if

> consumers' surplus provides a correct criterion (which, of course it does not, except in very special cases) it is better to guess at it, rather than at profitability.... This defence is plausible, but beside the point. Profitability is not, of course 'ideal' (except, again in very special cases), but, at least, one knows after the event whether one guessed right or not. The great trouble with any consumers' - surplus criterion is that one does not know, even after the event, whether the criterion was satisfied. In fact, the plain truth is that it does not yield us a criterion at all - or if it can be said to yield a criterion, then it is one which is open to anyone's interpretation within very wide limits. (Little (1960) p. 179).

In addition to these considerations two other points warrant mention. First, it is somewhat paradoxical to say the least, as Coase (1970) for example, has argued, that having estimated by use of the consumer surplus criterion that consumers are in fact willing to pay a sum of money equal to or greater than total resource costs, that no attempt should be made to extract such a sum from consumers by means of user charges. Since in the private sector of the economy consumers are expected to at least meet the total resource costs involved in the production of various outputs, where such outputs are produced under conditions of decreasing costs or otherwise, and/or whether lumpiness in investment is involved, the failure to confront the consumers of public enterprise outputs (including road space) with total resource costs will more than likely result in a distortion in resource use. The failure to recognise this stems in the main from an excessive pre-occupation with the short run objective of maximising benefits from the use of existing facilities. Even Vickrey one of the staunchest advocates of the short run marginal cost rule would appear to concede this point. Thus, it is argued:

> Marginal cost pricing must be regarded not as a mere proposal to lower rates generally below the average cost level but rather as an approach which implies a drastic re-arrangement of the pattern and structure of rates. Indeed it is this

> restructuring of rates that is likely to be the greatest contribution of marginal cost pricing of the overall efficiency of our economy, while the further gains that might be obtainable from the reduction of rates from a self sustaining level to a marginal cost level are, once the pattern of rates has been made to conform as closely as possible to marginal cost, are likely to be relatively small (Vickrey (1955) pp.618-619).

Finally, there is the matter of how deficits are to be financed. If users of road space are not required to meet the total costs involved (assuming of course that investment decisions are optimal), then this will result in a redistribution of income in favour of consumers of road space and also lead to distortions in resource use elsewhere in the economy. As to the redistribution issue there do not appear to be any compelling reasons why consumers of public enterprise outputs should in general be subsidised; while regarding the other point, as a practical matter, there are unlikely to be neutral taxes. In other words, financing large deficits by direct taxes, or by other methods, will result in distortion in resource use in the other parts of the economy, the total effect of which might well be to outweigh whatever gains are achieved by adhering, in the public enterprise sector, to a srmc policy pricing in a situation where either economies of scale exist or indivisibilities are important.

AN ALTERNATIVE LEAST CONSTRAINED MODEL

From what we have said so far it is clear that an alternative model of the road supply industry having the objective of determining the optimal quantity/quality of road space in various locations would differ from the model of Section 3 to the extent that investment decisions would, in general, be judged on the basis of revenue cost considerations (subject to the proviso that externalities - apart from congestion costs, are not important), and further, require recognition of the joint cost factor in the supply of road space. It is this characteristic, which in addition to location factors, accounts for the multi-product nature of the road supply industry. While brief mention has already been made of joint costs some additional comment would seem worthwhile.

We have already mentioned that the peak/off-peak demand pattern for road space is one example of 'jointness'. This is because the quality of the product provided during the peak is entirely different from that provided during the off peak. Put differently, road outputs provided at 6p.m. on a weekday are not perfect substitutes for outputs provided at 2a.m. All of this is in accord with the Marshallian definition of a product class. Moreover, since capacity which is provided to meet peak demands is automatically available for use at other times of the day - an example of what is known as 'time jointness' - there is no economic reason why the price of road space during the off peak should be the same as that during the peak. As in the case of the electricity supply industry, for example, where the same demand characteristics are evident, what matters is that price should not be less than those costs which could be avoided by not producing at all; while for peak period demand the congestion (opportunity) cost factor is all important. At the widest price range capital costs, invariate maintenance and variable maintenance costs are financed entirely from charges imposed on peak users, while off peak users contribute no more than costs which vary directly with use.

On the other hand, if demand in some off peak periods were to increase, then as soon as congestion becomes evident the price paid by off peak users will need to reflect this. In these circumstances off peak users will also be contributing to joint capital costs; all invariate maintenance costs and all variable maintenance costs. This of course, is precisely what one would expect to happen under conditions of perfect competition. If pure jointness exists then whether or not a producer will regard some outputs as "waste" or as products will depend on whether the price which consumers are prepared to pay, is at least equal, in the short run, to short run separable cost. From the long run perspective the relevant issue is whether total revenues from the sale of the joint products are equal to or greater than total long run separable + joint i.e. non separable capacity costs and invatiate maintenance costs. Of particular importance, however, is the fact that under conditions of perfect competition joint costs are allocated according to elasticities of demand, or according to "what the traffic will bear". The same principle would of course apply under other market conditions - the main difference in policy being with respect to the quantity of each of the various outputs sold. Whereas under competitive conditions the respective joint outputs of the firm will be treated as products rather than as a "waste", as long as price is equal to separable costs, under monopoly the supplier has more discretion. What is signficant, from the point of view of this discussion is that in either case relative prices for the joint outputs are determined by demand elasticities. Such differences, however should not be viewed as price discrimination. To avoid confusion it is perhaps better, in the joint product case, to refer to differences in relative prices as representating price "differentiation". This is obviously an important distinction since price discrimination as such i.e. charging different prices for the same product, when such prices do not reflect differences in costs of production is unlikely to be condoned as a legitimate practice for public enterprises. For a detailed discussion of this issue the reader is referred to the early debate between Pigou (1912, 1913) and Taussig (1891, 1913) on the matter of railway pricing policy, and in more recent discussions by Kolsen (1968), Ferguson (1972) and Kahn (1970).

In addition to 'time jointness' the other major element of jointness arises because different permanent ways are not normally constructed for the use of distinct user groups. Instead, road space is provided for the use of heterogeneous groups of user classes. The reason for this is obvious, namely, that it is cheaper in terms of resource costs. However, having said that it is clear that some parts of the capital costs can be unambiguously attributed to particular user classes - these are separable capital costs and assume importance when the long run problem comes up for consideration. Once separable capital costs (for each user class) are determined what remains is joint to all user groups. Apart from time related maintenance costs these costs include such elements as the right-of-way costs and the capital costs of the minimum quality road which must be incurred before any traffic can be allowed to use it. While such joint costs have been recognised in various road track cost allocation studies, they have, as observed earlier, been dealt with in essentially arbitrary ways. One of the most popular "methods" being the so called incremental cost method - a method employed in such studies as the U.S.A. Highway Cost Allocation Study (1961); the inquiry into Land Transport in Victoria (1971) and the study of road pricing policies in Canada prepared by Haritos (1973) on behalf of the Canadian Transport Commission. No attempt will be made to detail the errors of this approach - except to say that not only is it "backward looking" in perspective, but is also capable of producing a variety of allocations, each of which are equally arbitrary. Again, in terms of the above arguments the appropriate basis for attempting to recover such costs is that of charging according to what the traffic will bear.

In brief, road track costs may be classified as:-

(i) maintenance costs which are a function of use - referred to as separable maintenance costs.

(ii) time related - non separable maintenance costs.

(iii) capital costs which can be unambiguously imputed to particular user groups i.e. separable capital costs.

(iv) capital costs which cannot be unambiguously imputed to any particular user group. These costs are joint to all user classes and are therefore referred to as non separable capital costs.

The manner in which these costs are recovered, and the output combinations produced, will obviously depend, as already noted, on the behavioural rules governing the road supplied. For a public enterprise approach the objective may simply be viewed as that of maximising revenue subject to the constraint, dictated by statute, that the monopoly position of the road supplier is not exploited. And, as argued elsewhere, this is usually interpreted to mean that not only should price cost relationships for each part of the road network (given different supply and cost conditions) be the same in long run equilibrium, but also that profits from either the entire road system, or any part, should not exceed competitive levels. Whether such an outcome is 'right' on second best grounds is another issue. Even so it would seem to be at least a signficant move in the 'right' direction when compared with current road price output policy - which is characterised by an absence of a systematic relationship between prices and costs.

In what remains of this section we sketch the broad outlines of one model, discussed by Kolsen (1968) which attempts to examine the relationship between prices and costs in an 'optimal' road system. While the model makes no attempt to consider each of the joint cost elements cited above, it nevertheless, highlights a number of the issues involved in setting out the ingredients of an alternative least constrained model.

The system of charges suggested is one which is based on the quantities and qualities of road space consumed. Thus:

> If road space is to be provided in the quantities (qualities) for which users are prepared to pay, it is necessary to know whether and where users value additional quantities (qualities) more highly than other goods and services on which their incomes could be spent. The supply of road space which is optimal in this sense would permit users to relate the value of additional road space to the cost of providing it... The optimal degree of congestion would be achieved when users are indifferent between present payments for existing quantity of road space together with the costs to users caused by congestion, and the costs of greater quantity (quality) of road space together with the lower congestion costs. (Kolsen (1968), p. 84).

Specifically, what is required for the model (since it is assumed that quality and quantity are directly related) is an indicator of quality which is both measurable and 'reasonably' unambiguous, thereby making it possible to determine the willingness of users to pay for additional quality. The indicator which is chosen as 'best' meeting these requirements, is travel speed. To simplify the argument it is assumed, to begin with that: (i) all users have vehicles capable of similar speeds; (ii) increases in maintainable speeds between zero, and say 40 mph are viewed as increases in quality of road space, for which users are prepared to pay. However, the value placed on increases in speed is not the same for each user; (iii) vehicles are identical, the occupancy factor being determined by the number of vehicles; (iv) attainable speed is a function of road capacity and the presence or absense of other vehicles; (v) there are no economies of scale in road supply, and finally; (vi) an on vehicle metering system is available which records mileage and the speed at which mileage was undertaken.[1]

The meter would record different prices for different speeds, and as argued by Kolsen the initial level of prices would not be very important "because these prices will be adjusted so as to result in a maximum total revenue from, or similar cost/revenue relationships for, each road category". (Kolsen, (1968), p. 86). Clearly, price would need to be higher when speeds are lowest because low speeds would, in most cases, imply that the particular piece of road space being used is in high demand. Thus, for example, one might envisage the price for road space being x cents per mile when average speeds are less than say, 5 mph, 0.8x cents per mile for 5 to less than 10 mph, and 0.5x cents per mile for ten to less than 20 mph, and so on.

At the same time that users are charged different prices for using road space at different times of the day etc. and in different locations it is possible to record traffic volumes (and speeds) for the various parts of the road network. The effect of this is to enable the road supplier to relate revenues and speeds to location. In other words it is not essential to "know where each particular vehicle was at all times". (Kolsen (1968), p. 87) Instead all that is required is that "at certain prices with certain traffic conditions, certain quantities (qualities) of services were demanded". (Kolsen (1968), p. 87) In addition origin and destination studies would also be necessary in order to determine the optimal location for increases in capacity.

1. For a detailed discussion of road pricing methods see Beesley, M.E., (1973). Urban Transport: Studies in Economic Policy, (London: Butterworths) Chapter 9.

So far the model is much the same as that discussed in the second section. Consider, however, the effects of removing assumptions (i), (iii) and (iv). Clearly the removal of (i) and (iii) is to introduce a heterogeneous vehicle mix. Consequently the relationship between speed and vehicle occupancy is less direct than initially assumed. But this is not a cause for concern. So long as there are perceptible levels, congestion, all that is required is that we are able to express the effects of particular vehicles (e.g. ten ton trucks) on vehicle flow in terms of a 'displacement'factor. Thus, if our truck reduces traffic flow (ceteris paribus), by as much as, say, four 'standard vehicles' (e.g. "mini minors") then the price paid by the truck is four times that of the "mini minor". Similarly, if other factors such as differences in acceleration rates and degree of manoeuvrability are important, then these can be accounted for in a similar fashion.

What about the effects of removing assumption (iv)? Suppose for example that economies of scale exist. Again, as suggested earlier, this should not necessarily be a cause for great concern. While it is true that the sum of (prices) revenues equal to the cost of producing the marginal bundle (i.e. bundle of joint products - peak and off peak) will result in a loss if economies of scale exist, such a price/cost ratio is merely a lower limit. For reasons already noted it is by no means obvious that the 'right' pricing policy for the road authority is one which sets out to deliberately make losses.

Finally, given heterogeneous groups of users it will be appropriate to vary charges for user classes (within the variable charging scheme) in such a way as to take correct account of the other elements of joint costs. The aim being to ensure that prices will not exceed, for each user class, "what the traffic will bear". By proceeding in this way the road authority will be doing no more than what is currently common practice for other public enterprises, and indeed, what one would expect to occur under conditions of perfect competition. Thus on "uncongested" rural roads one would expect the price paid by heavy vehicles to differ from that paid by owners of sedan vehicles, and to differ for reasons other than differences in maintenance costs. In theory, if the quality of the road is worth retaining then the sum of revenues from all the user classes should be sufficient to meet separable maintenance costs (with no user class contributing less than separable costs) plus invariate (joint) maintenance costs - the contribution of each class to this element being determined according to differences in elasticities of demand. Moreover, if the "uncongested" road was worthwhile building, then the return on capital should indicate this to be so. As to the investment problem the road authority, would,given information on traffic flows and traffic composition, be able to determine whether it is worthwhile expanding capacity, and in addition, to provide quality characteristics for particular user classes. Thus if it is worthwhile providing specific quality characteristics for heavy trucks, then this would be determined by means of a comparison of revenues currently contributed by that class (given the cost of existing road space) and expected to be contributed by that class, with the additional costs involved in providing such facilities. Again, in long run equilibrium price/cost relationships for each part of the network would be the same - the precise relationship being determined by the objectives of the enterprise.

ALTERNATIVE (CONSTRAINED) MODELS

In this section we consider the problem of 'optimal' pricing policy for the road supply industry in a 'more realistic' or practical context. The discussion will be fairly short, attempting no more than to summarise some of the arguments developed in a recent study by Kolsen, Ferguson and Docwra (1974).

Previously we argued that to achieve an ideal pricing system for roads, it is necessary to have a charging method which accurately reflects demand and cost factors. Such a pricing regime could be achieved if a highly sophisticated meter system were available at relatively low cost. Since that apparently is not the case we will have to be content with something less than the 'ideal'. What we need to ask then, is whether it is possible to devise a charging system which not only reflects demand and cost factors, but is also capable of implementation. Not only is the answer in the affirmative, but there are clearly a number of such pricing methods some of which are obviously better than others and each in turn being better than current methods.

As far as the issue of variable charge versus fixed charge is concerned the former is preferable to the latter if it is possible to utilise the variable charge for the purpose of 'differentiating' between user classes according to differences in elasticities of demand.

The objections to existing variable charges are well known. For example, neither fuel taxes nor tyre taxes can adequately reflect demand elasticities or cost components. The fact that a particular type of vehicle (A) might consume 3 times as much fuel per mile as another vehicle (B) does not necessarily mean that 'A' imposes on the road authority and other users 3 times the costs which are imposed by 'B'; or for that matter can the differences be rationalised in terms of differences in elasticities of demand. Added to that, it is not practicable to 'differentiate' between user classes by charging different prices for fuel consumed by each class. Finally, such taxes can lead to distortions in engine or wheel design, as well as produce distortions in capital/output ratios. Other taxes, (e.g. the ton mile tax) while being somewhat better suffer similar defects.

In the Kolsen et.al. study (1974) it is suggested that is is feasible to implement a variable charge system. In essence, the suggestion is for the use of a simple inexpensive meter (such as the odometer) which cannot be tampered with, can be read on an annual basis and whenever changes in vehicle ownership take place.

As in the case of other multi-product public enterprises it is essential to be able to define a user (or consumer) class in terms of demand and cost factors, in order to be able to determine the 'right' charge per mile per vehicle type. What is suggested is that relevant cost components would include such factors as: (i) road occupancy characteristics of the vehicle; (ii) the area where the vehicle is used - this being important since the existing road system is characterised by substantial under and over capacity in various locations; and (iii) the time at which the vehicle is used. Relating the charge to these factors is necessary in order to take account of the separable cost elements, while for the joint cost component we need some rule-of-thumb for demand elasticities. For this the study cites the following: (i) the use to which a vehicle is put, i.e. whether the vehicle is used for private use, commercial use, ancillary and so on; (ii) the value of the vehicle; and (iii) the performance characteristics of the vehicle. Given the cost determined charge per mile for a particular vehicle, this is then multiplied by a demand determined coefficient (e.g. 1.0 for private use, 1.02 for commercial use, 0.98 for farmers and so on.)

While the use of the odometer is unsatisfactory for taking account of the use of road space during high demand periods (e.g. in large city areas) this can be overcome to a considerable degree by the use of special licences or stickers. The price paid for such licences would be determined by the road occupancy factor and could be sold for various time periods.

Obviously, the most expensive licence per vehicle type would be for a licence valid at all times during the peak for any location within the city.

Again, this combination of mileage and licence charges would enable the road authority (given the use of various devices to record traffic flows, speed and composition) to relate road user revenues to the cost of existing road space and enable more informed decisions as to what quantity/quality changes should be made, and in which locations.

Other Models

Apart from the adoption of a system of charging along the lines just considered the Kolsen et.al. study attempts to examine the problem of devising a more efficient system of road user charges in the context of constraints relating both to the size and means of financing the road budget. For instance, in Australia the principal sources of revenue for road expenditure are motor vehicle registration fees (a State government charge) and the petrol tax (a Commonwealth charge) - although with respect to the latter the relationship is certainly not explicit. What the study attempts, is to focus attention on the problem of rationalising user charges given the existing methods of financing road budgets. Clearly a number of models are possible. Three will be mentioned here.

The first assumes that petrol tax charges are taken as given and that the task for the road authority is to ascertain, according to the principles emphasised so far, the most efficient structure of motor vehicle registration fees. More specifically, since the existing structure of motor vehicle registration fees makes little economic sense (being determined on the basis of arbitrary technical criteria), one problem from a resource allocation point of view is to ask how the motor vehicle registration component of a State's road budget might be more efficiently raised, given that is has to be raised by means of a fixed charge.

The second model goes a step further by not only examining the fixed charge problem, but also asks how, given the use of the odometer, the petrol tax component might be more efficiently raised, while the third extends the argument by supposing that it is possible (and desirable) to separate the "tax" and the road budget component of the petrol charge. In each case the principles employed are similar to those suggested by public enterprise price theory.

Model 1

Since the objective is to recover the sum of revenues derived from motor vehicle registration revenues in a more efficient manner, it is necessary, as argued previously, to emphasise opportunity cost and demand elasticity factors. The model makes use of four data categories, namely, vehicle use, area of use, vehicle performance and vehicle value. For exposition purposes the study used data applicable to the State of Queensland. The total number of vehicles on register at June, 1972, were classified according to the above categories, and a total 'points' system calculated for the entire State. Given total motor vehicle registration revenues the value of a point is determined by the ratio of total motor vehicle registration revenues to the total number of points.

Area of use was deemed to be important because previous road investment decisions in the State have resulted in a substantial excess of quantity/quality of road space in country areas and shortages in cities and urban areas. The State was thus divided into three main areas, viz., the Brisbane urban area, Provincial City areas, and Country areas. The other categories (i.e. use, performance and value) were chosen as proxies for elasticities of demand for road space by the various user classes. Naturally, each of these variables differ in importance as proxies for elasticities of demand. In this model, value is assumed to be the most important and as such is given the highest absolute weight. A set of relative weights (based on "a reasoned judgement") were determined for value classes by vehicle types. Thus, for example, (for the year 1972) there were 1,458 cars and wagons in the value class $4,501 to $5,000. Since this class was assigned a weight of 19 (value class $500 or less was assigned a weight of 1) it contributed 27,702 "points" to the State total.

Regarding the performance factor, high performance vehicles were given a value of 1, and normal performance vehicles a value of zero. Weights for the remaining variables were assigned as in Table 1 below.

TABLE I[1]

Use/Area of Use	Relative Weight
Government/Local Government	1
Private	1
Ancillary	2
Commercial	3
Brisbane	3
Provincial City	2
Country	1

On the basis of the data collected for each of the four categories,

1. See Kolsen et.al. (1974), p.50.

together with the relative weights, it was possible to determine a total points score for the State. This amounted to approximately 7.362 million, and given (for 1972) a total motor vehicle registration sum of approximately $28 million, each point has a value of $3.80. Thus the charge for any vehicle depends on how many points it collects. For example, for a very expensive vehicle such as a new Jaguar, or new Mercedes Benz, used primarily in the Brisbane area for private purposes, the charge, given the weights employed in the study, would amount to about $228.00 in the first year (the weights being: value 55, performance 1, private use 1, Brisbane 3, giving a total weight of 60 points). Under the existing scheme the charge is about $60.50. If the same vehicle was used primarily for private use in the country the charge would amount to about $220.40. On the other hand, for a new mini minor used primarily in the Brisbane metropolitan area for private purposes, the charge would amount to about $41.80 compared with a present charge of $17.25, while for a vehicle of, say, ten years of age of normal performance, used privately for use in the city, the charge would be about $15.00 (in this case the weights would be: value 0, performance 0, Brisbane 3, private use 1, giving a total score of 4).

Model II

In this model the objective is to collect not only motor vehicle registration revenues in a more efficient manner, but also the petrol tax component. In 1972 petrol taxes generated in Queensland amounted to approximately 59 million dollars. The model assumes that the variable charge (i.e. the charge per vehicle mile) is put into effect by use of the odometer.

While the fixed charge is collected as in the previous model the variable charge is collected on the basis that the charge per mile for the various user classes is such that substitution effects are minimised. The 'tentative' analysis suggested that this range should be no greater than about 4 to 1 over the entire vehicle population. For the model this meant that upper and lower limits were set at point scores of 60 and 15 respectively, i.e. all vehicles having a point score greater than 60 were rated at 60 while those having a point score less than 15 were rated at 15. This contraction in the value range required a recalculation of total points. Total points generated by characteristics such as performance, area and type of use amounted to 3.345 million, while the points generated by value, given the restricted range, amounted to 7.278 million, giving a total point score for Queensland of 10.623 million points. Since these points are then used to collect the 59 million dollars in petrol taxes the charge per annum amounts to 5.5 dollars. Assuming an average mileage of 10,000 miles per annum, the charge is equivalent to .056 cents per mile. Under this system an 'expensive' vehicle used for private use mainly in the city area, averaging 10,000 miles per annum would incur the following charges. First a fixed charge of 228 dollars (i.e. 60 x 3.80) and second, a variable charge of 3.36 cents per mile (i.e. 60 x .056) giving a total charge of 564 dollars. How does this compare with present charges for the same vehicle type? As observed earlier such a vehicle would be charged approximately 60 dollars for registration. Assuming an excise charge of 17.3 cents per gallon and an average petrol consumption rate of 14 m.p.g. the variable charge amounts to 123 dollars - a difference of 381 dollars (this of course is still only a small proportion of the purchase price of the vehicle). For vehicles in the 'normal' performance category the model generates a charge of 188 dollars compared with a charge of 141 dollars under present arrangements.

Similar calculations can be made for trucks. Given the upper and lower limits specified earlier, it is argued that there is little likelihood for the differences in variable charges to result in substitution at either the upper or lower value ranges (the upper range refers to values in excess of 8000 dollars while the lower range refers to values less than 4000 dollars). Whether substitution will take place between these limits will depend on whether the differences in the variable charge more than compensate for the differences in quality of the various vehicle types. The general conclusion is that:

> Judicious adjustments of the upper and lower stop rates will effectively eliminate changes in the truck mix before and after the variable charge. Additionally, a substantial proportion of the truck market falls in the region above the stop rate (i.e. upper limit) and so is uninfluenced by the variable charge. (Kolsen et.al. (1974) p. 63)

Much the same sort of remarks apply to the issue of whether there will be an inducement to hold vehicles for a longer period than would be the case if the variable charge were not introduced. Since running costs are an increasing function of vehicle age, while the variable charge is a decreasing function of age, and thus value, the answer to this problem will depend on the relationship between the two functions. Again, the model allows for some comment. First, and as already noted, there is little inducement, because of the variable charge, to retain a vehicle for a longer period once its value is below the 4,000 dollar limit. Second, as far as the upper limit is concerned, benefits from holding on to the vehicle will not occur until the value of the vehicle falls below the 8000 dollar range. But even here the modal suggests that the incentive effects are not likely to be all that important. This claim can be explained by way of the following example:

> Consider a 20,000 dollar rig depreciated at 20 per cent (straight line) ... Thus in the fourth year of operation the variable charge would fall from 3.36 cents per mile to 2.58 cents per mile - a difference of 0.78 cents per mile ... In the fourth year a truck has probably moved into the 120 - 160,000 mile range and is probably nearing major overhauls, thus in this year it is likely that running costs have risen by more than 0.78 cents per mile. (Kolsen et.al,(1974) pp. 63-64).

Model III

The third model to which we wish to draw attention argues for the use of a three part tariff structure - to take explicit account of the fact that, on average, Australian government expenditure on road space by way of grants to the States has, over the years amounted to about sixty-six per cent of petrol tax receipts.[1] In view of this there would be considerable advantages in having a charging system which: (i) collected motor vehicle registration fees as for Model I; (ii) collected the Australian road grant petrol tax equivalent (by not less than half of petrol tax receipts, or whatever proportion is deemed by policy makers as desirable), as a variable charge along the lines suggested in Model II and finally; (iii) collected the revenue component of the petrol tax by an appropriate method, for example, by means of a sales tax or the existing excise tax on petrol. For reasons of convenience and cost of collection, retention of the existing tax would seem preferable.

Such a scheme would have two distinct advantages. First, there is the advantage of separating the revenue tax element from the road component. This is of particular importance from the point of view of efficiency in resource allocation within the transport sector as a whole. If government

1. It is now considerably less than this.

policy favours the imposition of increased fuel taxes on the various transport modes, for revenue raising purposes, then this can be done in such a way that each mode is taxed on the same basis. Secondly, there is the advantage that the reduced variable charge means that there is even less likelihood of substitution (because of differences in the variable charge by vehicle types) between vehicle value classes.

CONCLUSION

In essence the message contained in this paper is that a multi-product public enterprise approach to the road supply industry provides a sounder theoretical basis for examining price and investment policy. In particular, emphasis has been given to the problem of 'jointness' in supply and the manner in which joint outputs should (on theoretical grounds) be priced. More than that, a first step approach to the pricing problem - which is both practical and likely to yield substantial benefits to the community has been suggested. Thus, for example, by implementing a three part tariff structure along the lines indicated in section 5 (supplemented by the use of special licences in high demand areas), together with information about traffic flows (and speed) by regions and by composition, road authorities would obtain higher quality information regarding revenue cost relationships for various parts of the road network than is presently the case - the effect of which would be to enable a more efficient use of the road budget.

REFERENCES

Coase, R.H. (1946). "The marginal cost controversy", Economica, Vol. 13.

Coase, R.H. (1970). "The theory of public utility pricing and its application", Bell Journal of Economics and Management Science, Spring.

Dupuit, J. (1968). "On the measurement of the utility of public works", reprinted in Munby, D. (ed.) Transport. (Penguin Modern Economics).

Ferguson, D.C. (1972). "Joint products and road transport rates in transport models", Journal of Transport Economics and Policy, Vol. VI, No. 1, January.

Haritos, Z. (1973). Rational Road Pricing Policies in Canada (Ottawa: Canadian Transport Commission).

Hazelwood, A. (1968). "Optimum pricing as applied to telephone service", reprinted in Turvey R. (ed.), Public Enterprise. (Penguin Modern Economics).

Hirshleifer, J. (1958). "Peak loads and efficient pricing: Comment", Quarterly Journal of Economics, Vol. 72, No. 3, August.

Hotelling, H. (1938). "The general welfare in relation to problems of taxation and of railway and utility rates", Econometrica, 6, Reprinted in Arrow, K.J. and Scitovsky, T. (eds.) Readings in Welfare Economics, A.E.A. series 1969.

Kahn, A.E. (1970). The Economics of Regulation: Principles and Institutions, Vol. I, (John Wiley and Sons Inc.) Chapter 3.

Kolsen, H.M. (1975). "The Public Enterprise Concept and its Application in Transport Economics", paper presented to the Australian Transport Research Forum, Sydney 29 and 30 April.

Kolsen, H.M. (1966). "The economics of electricity pricing in N.S.W.", The Economic Record, Vol. 42, No. 100, December.

Kolsen, H.M. (1968). The Economics and Control of Road-Rail Competition (Sydney University Press).

Kolsen, H.M., Ferguson, D.C., and Docwra, G.E. (1974). Road User Charges: Theories and Possibilities, Study prepared for the Bureau of Transport Economics, June.

Lewis, A. (1941). "The two-part tariff", Economica, Vol. 8.

Lewis, A. (1949). Fixed Costs, Overhead Costs, (Allen and Unwin).

Little, I.M.D. (1960). A Critique of Welfare Economics, Second edition. (London: Oxford University Press).

Munby, D. (1968). "Mrs Castles transport policy", Journal of Transport Economics and Policy, Vol. 11, No. 2, May.

Munby, D. (1972). First Report from the Committee on Nationalised Industries, Session 1967-68, Minutes of Evidence (London: H.M.S.O.)

Ministry of Transport (1968). Road Track Costs: A Report by The Ministry of Transport, (London: H.M.S.O.)

Pigou, A.C. (1913a). "Railway rates and joint cost", Quarterly Journal of Economics, Vol. 27 (3) May.

Pigou, A.C. (1913b). "Railway rates and joint costs", Quarterly Journal of Economics, Vol.27 (4) August.

Pigou, A.C. (1912). Wealth and Welfare, (London and New York: Macmillan), Chapter XII.

Victoria (1971). Review of Report of the Board of Inquiry into The Victorian Land Transport System (Melbourne: Government Printer).

Steiner, P.O. (1957). "Peak Loads and efficient pricing", Quarterly Journal of Economics, Vol. 71, No. 4, November, pp. 585-610.

Taussig, F.W. (1891). "A contribution to the theory of railway rates", Quarterly Journal of Economics, Vol. 5, July.

Taussig, F.W. (1913). "Railway rates and joint costs once more", Quarterly Journal of Economics, Vol. 27 (2), February.

U.S. Department of Commerce, Bureau of Public Raods, (1961). Final Report of the Highway Cost Allocation Study, House Doc. No. 54, 87th Congress, 1st Session (Washington, D.C.)

Vickrey, W.S. (1948). "Some objections to marginal cost pricing", Journal of Policital Economy, Vol. 56, July.

Vickrey, W.S. (1955). "Some implications of marginal cost pricing for public utilities", American Economic Review, Papers and Proceedings, Vol.45, May.

Walters, A.A. (1968). The Economics of Road User Charges (Baltimore: Johns Hopkins University Press). Also, World Bank Staff Occasional Papers No. 5.

Williamson, O.E. (1966). "Peak-load pricing and optimal capacity under indivisibility constraints", American Economic Review, Vol. 56, pp. 810-27. Reprinted in Turvey, R. (ed.) Public Enterprise, (Penguin Modern Economics).