TRANSPORTATION INFRASTRUCTURE COSTS AND PRICING

Helen KRAMER)*

Transportation poses a distinct set of problems for economic theory and policy. This is so because of the technological nature of transportation infrastructure projects and the associated behavior of costs. These investments are characterized by high capital intensity, high durability, increasing returns to operation up to planned capacity in the short-run (which is typically 20—30 years), increasing returns to scale in the long-run and marked indivisibility, in that most projects have a minimum size. As a consequence the behavior of the marginal and average cost curves differs substantially from the typical case of traditional microeconomic theory, in which decreasing returns to a fixed factor of production, ultimately decreasing returns to large scale and perfect divisibility are assumed.

According to a study prepared for the Commission of the European Economic Community, »increasing average returns appear to be the general rule in transport infrastructures«.[1]) This holds for both infrastructure and those associated operation services which are independent of the volume of traffic. However, the same experts found that substantially increasing *marginal* returns to capacity can occur only in infrastructure investments, when new infrastructure is established and brought into service, and not in the operation of an existing infrastructure. Total real capital costs are less for constructing a larger infrastructure at the outset, which may be underutilized for some time, than those incurred for constructing a smaller project and expanding its capacity when it has become overcongested. This is another aspect of indivisibility.

INVESTMENT CRITERIA

Logically, the first problem that should be confronted is the choice of criteria which are to be the basis for deciding on construction of infrastructure projects. Only some form of coordinated planning and consistent application of investment criteria can make rational choice of investment pro-

*) Researcher at the Institute of Economic Studies, Belgrade.

[1]) European Economic Community, *Studies: Options in Transport Tariff Policy* (Brussels, 1965), p. 25.

jects possible. In non-centrally planned economies, cost-benefit analysis provides the criterion for choice among alternative projects most frequently advocated by economists, particularly in the fields of transportation and water resources. Those projects are adopted which promise the greatest net social benefits over social costs, when these magnitudes are discounted to take account of differing time periods.

Cost-benefit analysis provides no simple formula for investment decision-making. However, it is nonetheless a very useful planning tool in that it imposes the necessity of making careful studies to quantify costs and benefits as far as possible, instead of relying on vague qualitative judgments. Estimation of benefits also serves the useful function of providing evidence as to the charges that users of transportation infrastructure could reasonably be expected to pay. The requirement of justifying any proposed project by a cost-benefit analysis should reduce the danger of making incorrect investment decisions out of ignorance or in response to the political pressures of narrow interest groups. Imposition of user charges wherever technically and administratively feasible can be considered desirable not only as a financing device but also as a check to the latter tendency, provided it is consistent with a policy of economic development. It is to the thorny issue of pricing the use of transport infrastructure that we will now turn our attention.

PROBLEMS OF PRICING TRANSPORTATION INFRASTRUCTURE

1) Marginal Cost Pricing Systems

Variants of a marginal cost pricing system are derived from the theory of optimal resource allocation. Actually, none of the proposed variants bases price solely on marginal cost, and they may be called with more precision »MC plus« systems. In the case of the EEC experts cited above, the »optimum price« or »economic charge« consists of a cost charge which is equal to marginal cost, plus a rent called a »congestion charge«, which is imposed only when capacity is fully utilized, and which is designed to eliminate excess demand. The congestion charge so defined is not an economic cost, since it does not measure the value of resources consumed, but a pure rent for the use of durable and indivisible factors. The other variants, formulated in discussions of road pricing, employ the concept of marginal social costs, in which marginal cost is defined in such a way as to include congestion costs and external diseconomies such as air pollution and noise. The EEC pricing criteria are more general, being applicable to railroads and waterways as well as roads.

a) *The EEC Experts' Proposal: Economic Charges*

Since the use of the term »marginal cost pricing« in economic literature is apt to be misleading, it should be emphasized that the optimum price, as defined by the EEC experts, cannot be determined from cost factors alone, but must be set at a level that will equate demand with capacity. Hence the price setters must estimate the price elasticity of demand.

The marginal cost component of the economic charge for use of transport infrastructure consists of those costs of infrastructure operation which

vary with the volume of traffic: variable costs of maintenance, renewal and safety measures. Most of the costs of infrastructure, however, are fixed: most damage to roads, for example, is caused by time and weather rather than traffic, and railroad tracks, bridges, tunnels, embankments, drainage systems, guard rails and signalization must be maintained irrespective of traffic. These fixed costs and amortization and interest on the capital invested can be covered only by the rent component of the economic charges, which will be positive only when congestion occurs. Thus, it can be anticipated that an infrastructure built according to the criteria of optimum resource allocation will incur a deficit over its economic lifetime, defined as the (negative) difference between the sum of the discounted value of investment and operating costs and the discounted value of revenue from all economic charges.[2])Since the investment rule for optimal allocation requires that the sum of marginal investment cost and discounted marginal operating costs (defined in relation to capacity and independent of the volume of traffic) equal the discounted value of the revenue expected from congestion charges, the size of the total discounted deficit is determined by the difference between the average and marginal investment and operating costs, or in other words on the extent to which infrastructure construction and operation show increasing returns.

A deficit will be realized when infrastructure investment has been adequate, coming at least close to the economically optimum level, while underinvestment will result in high congestion charges which might be sufficient to eliminate the deficit. »Whether the infrastructure is optimal or whether it is ill-adapted to present and future demand, the option of economic charges presupposes that in any case the best possible use must be made of the infrastructure as it stands. This objective cannot be attained unless the charges for its use are equal to the economic charges.«[3])

However, the system of economic charges can lead to an optimal allocation of resources among competing transport infrastructures only if it is assumed that all investment decisions have been taken correctly. Where cost-benefit analysis would show a negative result with respect to a particular project, but it is built anyway because of political interests, price policy cannot optimize factor combinations in competing infrastructures. The »political« project will divert some traffic from other infrastructures, decreasing utilization of capacity, which can render an originally correct investment decision non-optimal and increase the size of the deficit which has to be financed from public funds. It is the problem of the deficit which is the main weakness of the system of marginal cost pricing.

In the case of railroads, where infrastructure and transportation services are under the same management, it would be difficult in practice to separate the accounts in order to distinguish between a deficit attributable to infrastructure and a deficit that might arise in other aspects of railway operation. Public assumption of the deficit would therefore weaken incentives

---

[2]) This definition holds only at the time of entry into service of the project. After that date the definition must employ an economically arbitrary term — the non-amortized value of the capital investment. The deficit is then defined as the sum of all present and discounted future expenditures plus the non-amortized part of the initial investment minus the discounted value of all present and future revenues.

[3]) E.E.C., op. cit., p. 78.

to efficient management of railway transport services, and would be inconsistent with self-management and the principles of the economic reforms.

Another problem is that the relatively low level of the »optimal« charges may distort investment location decisions and evoke demand for transport that would not arise if users had to cover full investment costs of the infrastructure.

The question of financing the deficit raises a number of crucial issues. From the point of view of equity, it would seem desirable to tax the economic rents of infrastructure users, but this cannot be done without reducing utilization of the infrastructure below the optimum level. The theoretical solution proposed by advocates of marginal cost pricing systems is a neutral tax, so called because it does not alter the marginal conditions for economic decision-making. One is compelled to observe, however, that references in the literature to neutral taxation as a means of financing the deficit are suspiciously brief and vague. The only concrete example offered of such a tax is a poll tax or equal lump sum paid by everyone, infrastructure users and nonusers alike. Any suggestion, implied or explict, that such a tax has no effect on the allocation of resources and social welfare is ill-founded, in that it neglects the distribution problem. The transportation content of all goods and services is not equal, nor are the benefits derived from a given infrastructure equally distributed. At least some members of society would be paying more in poll tax than the value of extra benefits they receive from transportation improvements, while others paid less: »... even a poll tax imposed on an already optimally organised society is not neutral if its proceeds are redistributed anmong the community. It shifts the economy to a different optimal position — one in which some people are better off and others are worse off«.[4])

A related issue is that a certain allocation of resources is »optimal« only on the basis of a given distribution of income, which may not be regarded as equitable. Consequently, consideration of the implications of the deficit and the means for its financing must take into account the effects of any particular policy on the distribution of income. While a new lump sum tax would be »neutral« in its effect on short-run production decisions under the (unrealistic) assumption that all producers behave according to the rules of optimalization, it would nevertheless alter the pattern of consumption spending and saving, since its immediate effect is to reduce disposable incomes. As a consequence the composition of output would be altered, and given the diffused nature of benefits from improved transport, it would be extremely difficult, if at all possible, to demonstrate that the lowest paid workers whose consumption would be reduced in the short-run were not on balance worse off.

With respect to the regional distribution of income, however, the system of marginal cost pricing has a distinct advantage over the alternative systems that will be considered below. That is, it is the only system truly consistent with a development policy and unlike others can be applied without modification to the insuficiently developed regions and purely local networks, which should not be made to bear the burden of self-financing.

---

[4]) E.J. Mishan, »A Survey of Welfare Economics, »*Surveys of Economic Theory*«, Vol. 1 (New York, 1965), p. 171.

b) *Marginal Social Cost Pricing and Application to Roads*

One of the purposes of charging for use of infrastructure is to control overall demand for transportation and to distribute it among alternative infrastructures. The essence of the argument in favor of economic charges or marginal social cost pricing is that when charges reflect costs, resources will be efficiently distributed between transportation and other goods and services and also between one type of infrastructure and another. For these charges to be effective in accomplishing their allocative goal, they must be imposed on users in such a way as to make the cost of their journeys clear to them in advance. Otherwise economically unjustified journeys will be undertaken which add to congestion and raise social costs.

On the other hand, short-run demand for transportation is highly inelastic, and is influenced by non-monetary factors such as speed, reliability and comfort. Infrastructure charges, which sometimes represent only a small part of transportation costs, would have to vary widely in order to affect demand.

Since a survey of the literature has failed to uncover any attempt to apply this pricing system to railroads, the following discussion will be confined to road pricing.

PROBLEMS OF MEASURING MARGINAL COSTS

The *social* costs of a vehicle journey consist of the *marginal* costs of:

1) maintenance (to repair damage caused by road use)

2) periodic reinforcement of the road

3) policing

4) safety

5) congestion or loss of time imposed by the marginal vehicle on all other road users.

The first four categories of variable costs are borne by the administrative unit that manages the road. It is these costs plus the congestion cost that must be charged for, since the vehicle owner already bears the average *private* costs of the journey, which consist of wear and tear of the vehicle, fuel and the value of the driver's time and risk. When the social costs are recovered on the basis of user charges per vehicle-kilometer, the private costs of vehicle owners will then reflect all the consequences of their use of the road.

The main difference between the system of marginal social cost pricing and the system of economic charges is that the latter imposes a charge in order to achieve optimal demand for infrastructure capacity, while the former treats the congestion charge as the real cost of an additional journey. Congestion increases vehicle running costs, vehicle repair costs, rate of wear of tires and parts, and the time required to complete a journey. It also increases air pollution, and some authorities advocate charging for this as well as noise, but offer no suggestions how to measure such external diseconomies.

One of the major unsolved problems in road pricing is how to allocate costs among different categories of vehicles. The cost of road reinforcement

varies considerably according to category of vehicle, increasing more than in proportion to weight, so that the corresponding part of marginal cost is very high for heavy vehicles and very low for light.[5]) A difficulty in calculating these costs is that expenditures for reinforcement or renewal vary discontinuously, so that in most cases it is impossible to determine directly the expenditure variation caused by an additional unit of traffic. Consequently, correlations must be sought between variations that go beyond the marginal scale in any strict sense.

Marginal costs of maintenance and policing are difficult to allocate by vehicle category. However, it is well known that overloaded vehicles cause great damage to road surfaces, although no reliable estimates of costs are available. This imposes the need to establish a system of inspection and heavy fines for violation of loading regulations.

Another factor to be taken into consideration is that maintenance costs vary considerably with the quality of the road, being very low per vehicle-kilometer for paved roads, somewhat higher for gravel and several times higher for earth roads. Regression equations relating maintenance costs to volume of traffic must be interpreted with caution, however, because the causal relationship is not one-way. Traffic will tend to be higher on well-maintained than on poorly-maintained roads.[6])

Maintenance costs also may vary according to the composition of traffic, or proportion of light to heavy vehicles. However, an American study found virtually no difference in effects on maintenance costs between heavy and light vehicles on roads designed for heavy traffic, although damage caused by weather may be compounded more by heavy than by light vehicles.[7]) The problem in less developed countries such as Yugoslavia is that heavy vehicles often use roads not designed for such purposes, and in such cases it is certain that they cause higher maintenance costs per vehicle than passenger cars and light vans.

Marginal costs of congestion and safety are generally higher than for variable maintenance and policing and are also difficult to allocate. For congestion costs only crude approximations are feasible, such as applying coefficients to a type of vehicle taken as a standard unit to calculate occupation of road space and averaging costs per standard vehicle unit on this basis. As in cost-benefit analysis, the problem is encountered of imputing a value to time, in this case time lost rather than saved.

Marginal costs of safety, which reflect the accident risk imposed by an additional vehicle, could be reflected in vehicle insurance costs, if these varied more in relation to the marginal safety risk factor attributable to each user. Thus, different insurance premiums might be charged on the basis of accident rates associated with various categories of driver, classified according to age, sex, driving experience, place of residence and average distance driven annually.

---

[5]) European Conference of Ministers of Transport. Economic Research, Third Seminar, »Pricing the Use of Infrastructures«, mimeo (Paris, Sept. 1970), p. 14.

[6]) A.A. Walters, *The Economics of Road User Charges*, World Bank Staff Occasional Papers No. 5 (Washington, D.C., 1968), p. 170.

[7]) Ibid., p. 171.

Conclusions

At best, the marginal costs of traffic on roads can be only roughly approximated. Considerable averaging among roads and by categories of vehicles is of necessity involved. For railroads the problem of measuring marginal cost is even more complicated because of the technical interrelatedness of infrastructure and transport services and the even greater relative weight of joint costs and common costs.

The fatal weakness of all marginal cost pricing systems, however, is the deficit incurred. Objections to financing the deficit from general taxation have been discussed above.

Finally, it is not possible to demonstrate convincingly that a marginal cost pricing system in any form for transport infrastructure will achieve a better allocation of resources in the economy and a higher growth rate of real social product than alternative systems.

2) Systems of Budgetary Equilibrium

Of all the possible alternatives to the system of marginal cost pricing, a version of the system of budgetary equilibrium without borrowing (called pay-as-you-go financing) is the least arbitrary and simplest to apply. Each year, users are charged the full costs of that year's expenditures on infrastructure investment and operation (including maintenance and renewal). Thus, there is no deficit, no subsidization and no problem of determining amortization schedules.

There are two circumstances under which the user charges imposed under this system would be similar to those corresponding to the system of economic charges: when an infrastructure is congested and congestion charges under the latter system would be high enough to eliminate the deficit, and when infrastructure is expanding slowly so that annual charges have to cover only current operating expenses. In all other cases budgetary equilibrium would impose higher charges, and to the extent that demand is price-elastic would reduce infrastructure utilization to below the optimal level.

Another difference is that investment might be less under the system of budgetary equilibrium. The investment rule must be redefined, substituting discounted total revenues for total social benefits. Since in practice it is not possible to recover all consumer and producer surpluses (or economic rents), total revenues will be less than total benefits and investment will be less than optimal. However, this may be true as well under the system of ecnomic charges, where investment is constrained by the availability of funds from the public budget.

A more serious problem is that mistaken investment decisions, particularly overoptimistic ones, under this system can cause substantial deviations from optimal utilization. Charges aimed at budget balancing will be much higher than marginal costs, and the already underutilized infrastructure will have even more excess capacity as a result. This danger can be reduced by equalizing charges over a large enough region. Regions should be demarcated so as to include complementary parts of a network and both congested and uncongested routes. The more homogeneous the subset of routes (road or rail) for which budget balancing is required, the greater is the

likelihood of inefficient pricing, with attempts to cover costs leading to progressively increased user charges and reduced traffic.

It should be clear from the foregoing that under no circumstances should the system of budgetary equilibrium be applied to large underdeveloped areas (large relative to the size of the region over which charges are equalized) *nor to purely local networks.* Consistent with a policy of development, these are best financed out of the Fund for Development of the Insufficiently Developed Republics and Province or the republican budget, users paying only the marginal costs.

The degree of geographical equalization may also be inadequate in the case of very large indivisible investment projects, such as electrification of the railroads, construction of an entirely new infrastructure (such as Belgrade — Bar railway), or when an existing network is enlarged so much that it practically amounts to construction of a new network. Widening the area within which charges are equalized would have the disadvantage that if infrastructure costs differ substantially from one network to another, equalization between them would distort the distribution of traffic away from the optimal allocation.

Delineation of regions for equalization must also take into account the effect on conditions of competition between roads and railroads. When one mode of transportation (road) is expanding more rapidly than the other (railroad), in years when investment takes place user charges will be much higher on the expanding network. This would cause a temporary and unsustainable shift of traffic from roads to railroads, and increase the financial burden on remaining road users.

This points to a major drawback of the system of budgetary equilibrium without borrowing. There is no economically based argument why users in one year should pay all the investment costs of a transportation infrastructure that will last a quarter of a century or more. Certainly this conflicts with notions of equity. This holds all the more, the more discontinuous the investment, the smaller the number of those affected by the charges and the larger the investment in relation to the existing transportation network. An addition to the charges imposed on an individual user is justifiable economically only insofar as these charges are equal to the costs the user imposes on the infrastructure or other users. But the method of budgetary equilibrium without borrowing treats costs as if the infrastructure were used up in the year it is built. Thus charges are imposed on the user which are not related to the value of the resources he consumes in the respective period and which may exceed user benefits, while for future users charges are not added which correspond to benefits received. Moreover, this system would imply sudden changes in charges, which would create severe problems of adjustment in transportation enterprises.

All these considerations lead to the conclusion that it is desirable, even at the expense of some simplicity, to modify the system of budgetary equilibrium to permit borrowing.

Once borowing is permitted, the room for subjective decisions increases in defining total costs and in determining annual schedules of amortization and user charges.

Only one definition of total costs is consistent with the investment criteria for optimal resource allocation: total cost is equal to initial investment

cost plus total discounted future operating costs independent of traffic minus the discounted residual value of the infrastructure at the end of its economic life. Budgetary equilibrium could then be defined as being achieved when the total discounted value of revenues obtained over the infrastructure's economic life at least equals total costs, as defined. This definition is precise and objective only in application to new infrastructure projects, posing a problem of transition in application to existing infrastructures. Any attempt to define the capital value of the latter would be arbitrary and impose a heavy burden of additional computations. Anyway, there is no economic argument for charging present and future users for infrastructure that has already been paid for in the past.

There is one legitimate issue with respect to the transition period, and that is that imposition of the budgetary equilibrium requirement on new infrastructures may distort conditions of competition between new and old infrastructures if the latter are freed of capital charges. This problem can be dealt with by appropriate demarcations of the regions over which charges are equalized, and is much less serious when borrowing is permitted to finance investments or major renewal works which are spread out over several years.

Under this system annual user charges will equal the sum of interest and amortization on debt incurred in the past plus current expenditures not financed by borrowing. We see immediately that subjective judgments are involved in determining the annual charges because all amortization schedules are conventional in nature. The effect of inflation on capital values also poses a problem. The older an infrastructure, the greater the extent to which the real value of its capital debt has been wiped out by inflation and the less the real value of amortization and interest charges. Newer infrastructures would be thus placed at a competitive disadvantage, unless this factor were taken into account in determining the boundaries of the regions over which user charges were to be equalized. This is, of course, a problem only when transportation networks are sufficiently developed to offer users alternative routes between two points. It is unlikely to cause serious difficulties in Yugoslavia, where networks are not dense.

What may pose the crucial obstacle to application of this system in Yugoslavia is the absence of a capital market. International loans can be used only for very large projects or those related to development of tourism, and have the disadvantage of placing an additional burden on the international balance of payments during the period of repayment. National loans likewise are a device suitable only for major projects, and as the campaign to finance the Belgrade — Bar railway clearly demonstrates, entail relatively high costs of floating the loan.

Aside from this, the main problem in all systems of budgetary equilibrium is that they do not imply any particular method of apportioning charges among various categories of users. Only the portion of annual charges corresponding to marginal costs can be allocated on an objective basis. The remainder, corresponding to the annual deficit, can be apportioned only on the basis of some convention. The theory of optimal resource allocation requires only that the method of apportionment avoid distorting conditions of competition.

There is general agreement among experts that budgetary equilibrium should be applied to each mode of transport separately. However, if the de-

ficit that would be incurred under the rules of optimal resource allocation is very high and substantially differs in size between competing modes, this system would severely distort the conditions of competition and the high charges would result in inefficient use of infrastructure capacities. In this case the system of budgetary equilibrium in any form should be rejected.

Within each mode of transport, networks must be classified according to the purposes they serve. A minimum classification would distinguish among (1) urban and suburban networks, (2) main (interurban) networks, and (3) local networks. The last may in the case of roads be subdivided further into rural secondary feeder roads (connecting to main networks) and access roads (serving only the local population). Budgetary equilibrium should be achieved for each of the first two categories separately, while the last category must be exempted from the budgetary equilibrium requirement *if it is desired to have rural roads corresponding to modern standards and to development needs.*

Users must also be classified. A minimum classification for roads must distinguish among passenger cars, buses, light trucks, heavy trucks, motorcycles, non-motor vehicles and pedestrians. Railroad users must be classified as passengers, types of freight and traffic in different directions.

3) System of Calculated Total Cost

This system attempts to eliminate the distorting effects of budgetary equilibrium on the conditions of competition between infrastructures of different ages by using replacement costs instead of historic costs as a base for calculating user charges. It thus takes account of price changes, and in a more complicated version can take account of technological changes in transport infrastructure. Each year and for each infrastructure replacement costs must be calculated anew, and the capital burden determined on the basis of a schedule of amortization and interest charges. This has a certain appeal, since it makes financial provision for maintaining infrastructure in the future, but it has the overwhelming drawback of imposing a tremendous annual burden of calculation on infrastructure managers.

Another objection is that in a period of inflation user charges on existing transport networks would be increased. This would raise costs of distribution and put additional upward pressure on costs of production, thereby contributing to still more inflation, higher reproduction costs, and another round of increases in transport user charges.

4) Short-Run Average Cost Pricing

An English economist has recommended SRAC pricing as a pragmatic solution to transport problems. Each category of user would be charged the annual average cost based on conventional accounting methods and including provisions for amortization and interest charges and reserves. Railroad and road managers would be required to just break even and to pass any revenues above total cost on to users in the form of lower prices. The author calls this a consumer surplus pricing criterion.[8])

[8]) C.D. Foster, *The Transport Problem* (London, 1963).

The only argument that can be made on this proposal's behalf is financial expediency. Its adoption would make it impossible to use prices as a means for attaining the best possible use of transport infrastructure. Where an infrastructure is incurring increasing costs owing to congestion, under this system the price charged would be less than marginal cost; congestion would be greater and user satisfaction less than under alternative pricing systems which take marginal cost into account. For new infrastructures, which are underutilized, setting price equal to AC will discourage worthwhile traffic which is willing to pay its costs (i.e., marginal costs), resulting in even more excess capacity.

SRAC pricing would also distort the conditions of competition between railroads and roads. Because railroads have a higher proportion of fixed costs, their marginal cost is generally lower and average cost generally higher than corresponding costs of road transport. SRAC pricing would make it impossible for underutilized rail networks to achieve better use of capacity and realize increasing returns (lower AC).

CONCLUSIONS ON PRICING SYSTEMS FOR TRANSPORT INFRASTRUCTURE

None of the pricing systems devised to date for transport infrastructure is completely satisfactory with respect to criteria of optimal resource allocation, distribution of income, and economic development. A marginal cost pricing system is desirable from the standpoint of sub-optimizing the allocation of resources within the land transport sector, but financing the deficit that results out of the government's budget is unacceptable because of its problematical income redistributional effects; in Yugoslavia it is also politically unacceptable because it runs counter to the principle of self-management and the goal of deétatization, as currently interpreted.

Of all the alternatives, the system of budgetary equilibrium with borrowing appears to offer the greatest advantages with the least drawbacks. It is possible to a significant extent in road pricing, and to a lesser extent in pricing railroad services, to base user charges on marginal costs, while recovering joint and common costs of infrastructure by supplementary charges.

However, purely local road and railroad networks (having no through traffic), and all networks in large underdeveloped regions must be exempt from the budgetary equilibrium requirement, consistent with a policy of promoting more rapid economic development. In Yugoslavia, therefore, a system of marginal cost pricing should be applied in Macedonia, Montenegro, Bosnia-Hercegovina and Kosovo, and the deficit should be financed out of development funds. Local networks in the rest of the country should be financed out of a special fund created in each Republic for this purpose. In no case should individual communes be responsible for financing construction and maintenance.

(Rad primljen decembra 1971)