Chapter 1

INFRASTRUCTURE & THE RAILWAYS

1.1 The Economics of Infrastructure

Infrastructure development and the extended linkages this has with economic growth and development has come into close scrutiny in the post-recessionary world situation of the 1990s. Two major reasons contributing to a lack of soft alternatives in this regard, are the tail-off of high achievable LDC growth rates in capital goods with the incursion of consumer-oriented living standards and consequently increased production costs, and exhaustion of scale economies within the existing spatial distribution of resources. Although country-experience of infrastructure development has been wide and diverse, with impacts varying with size of the country involved, the importance of new thrusts in this direction derives from potential increases in market-size and productive capacities which would ensue. The rationale behind this, drawn diversely from the literature on Schumpeterian innovation cycles, Kondratieff long waves and Kuznets swings[1] postulates that long sustained expansionary waves leading to economic development would arise from concentrated focal investment in those economic sectors having vertical interdependency-linkages with the widest spectrum of economic activity. Development of infrastructure is also important from the point of view of 'closing the development gap' and 'catching-up', so as to leave the LDCs advantageously placed with respect to the new World Trade Order and its anticipated spin-offs.

The first question that has however to be confronted here relates to the origins and the modes of generation of infrastructure. The vast development canvas afforded by economic history has generally shown the generation of infrastructure to have been an act of governance, raising obvious conflicts with the *laissez faire* concepts of the state. Infrastructure therefore exists generally as the result of design rather than of chance, or of the `invisible hand', since the scale of coordination and promotion required for putting infrastructure into place is usually well beyond the capacity of disaggregated economic enterprise. Moreover in most cases, the large investments required for the development of infrastructure have either been directly sourced or else supported from government coffers.

The second question that confronts the issue relates to the impact of infrastructural bottlenecks in constraining the growth of the economy, or conversely to the multiplier impact of infrastructural investment in achieving the acceleration of development. Infrastructure is directly a *development* variable rather than a *growth* variable, since the qualitative levels of economic activity undergo a sea-change once infrastructure has been instituted. Thus the development options that governments face relate to the relative location of development investments *i.e.* whether the investment to be made will be at the economic base, or at downstream positions. Rules of complementarity between infrastructural and other investments dictate a sequencing of these investment flows so that capacities created by investment of one type are fully exhausted before the next expansion in capacity occurs.

The present study will look into the character and role of infrastructure in development through the medium of one of its most important historical constituents, namely transportation. Within the different modes of transport, the greater part of two centuries of economic development has borne close association with the development of railways. Although the degree of dependence of economic activity upon railways varies according to country and according to the point of time from which the development of transportation commenced, most advanced economies today can attribute their state of advancement to critical periods in the past, when the advent of railways improved communication linkages adding new muscle to their industrial sectors. India, in this respect, was one of the fortunate few where the railways were introduced around the same time as they were in the advanced countries. However the course of subsequent economic development differed significantly, the historical reasons for which have usually been identified in colonialism. The coming of national independence and the institution of planned economic development made a marked change in the placement of the Indian economy vis-a-vis the railway infrastructure that was already in

position, and nearly five decades of development experience are now available to assess the impact that transport development has had on the economy in its character of an infrastructural variable. The various facets of this impact will now be outlined and assessed.

1.1.1 Types of Infrastructure

Under the older name of *social overhead capital* [Rosenstein Rodan, Nurkse, Hirschman], infrastructure encompasses a gamut of *public utilities*, like power, telecom, water supply, sanitation, etc., and *public works* like roads, dams, irrigation-canals, and the entire range of transportation services. Since the process of economic development involves complex interplays between different constituent sectors of the economy, the provision of infrastructure, which defines the social envelope of this interdependence, determines the ultimate pace of and locational balances within the process. The extent of such provision depends on *infrastructural investment*, which may then alternately be represented as a capital variable within economic space, that determines the ultimate efficacy of capital outlays made on other technical inputs in raising output and productivity.

A first distinction might be made between public works and public utilities as components of infrastructure. The word-prefix 'public' attached to each indicates a commonality between them arising from their common characteristics as *public goods*, and on account of their collective or social provision. Beyond this, the two categories become dissimilar because while the former boosts the scope for the economics of value, production and distribution, the latter improves welfare. Thus despite their inclusion within the common genre of *infrastructure*, not all its constituents are common in physical character or in the location of their ultimate impact on the economy. Instead, each constituent is a separate entity, capable of exercising its influence on production and distribution processes either singly, or conjointly with other infrastructural constituents. Because of this ambivalence, the crucial factor deciding the efficacy of infrastructural outlays is the order of their precedence. Thus the provision of power without roadways would limit the infrastructural impact on production, because it remains delinked from the markets, and the provision of water supply without sanitation would not necessarily improve welfare, because of the absence of basic hygiene. In a long-term analysis of development, perhaps the most crucial infrastructural input would be transportation, because of its capacity to unify fragmented markets and consolidate production and the allocation of resources.

The present study is focused on this particular aspect of transportation as a part of development infrastructure. Although the overall infrastructural impact of transport services on the economy may be similar, distinctions are often drawn between alternative transportation modes in terminology such as *surface transport*; *land-based* and *water-based* transportation; *inland* and *international* waterways; global *seaways*; *air transport*; etc. The endowment of transportation facilities that any country is provided with at a given point of time is a joint-product of geographical location and economic history, with the chronology of development exercising a deterministic role. For instance, while inland waterways had once held an all-important position in transport scenarios of the past, their actual presence was restricted to countries which had adequate systems of navigable rivers. Later, with a revolution in other modes of transport, while the domestic importance of waterways declined considerably, the seaways still hold on to that importance by virtue of their primacy in international freight movements. Even within *surface* transport, the two important constituents i.e. the roadways and railways are not necessarily perfect alternates, since while the former have developed mainly to serve the need for moving low-volume freight over shorter distances, the latter have over time become increasingly specialised towards the movement of bulk freight. The newer development of airways in more recent times has centred around the need for unprecedented speed in the delivery of special freight over vast distances, but still remains little more than a fringe choice for the LDCs because of the huge amounts of capital investments required. Thus surface transport is still the first investment choice while developing the infrastructure for transportation in such countries, and the railways hold special importance to countries of larger size with vast internal distances.

1.1.2 Transportation & the Mobility of Resources

Economic development devolves upon the integration of isolated growth trends in economic activity into nationwide economic acceleration. One factor crucial to such consolidation is mobility of men and materials over the economy, since such mobility widens the ambit of markets for inputs and finished products and

demand and supply linkages therein. Infrastructure plays a role auxiliary to the development process. While infrastructural variables often have the character of public goods supplied by public utilities, they are critical to the determination of the overall productive capacity of the economy. There are, however, important distinctions between the different components in infrastructure. Infrastructural sectors such as energy or power produce direct outputs. The transport infrastructure on the other hand produces services that indirectly bear on gross output through their influence on factor and product mobility and hence on resource allocations. The transportation network serving the economy by interlinking its sectors and regions is an infrastructural variable. Commodity flows over this network and their growth indicate infrastructural demands and their realisation. The supply variable is the carrying capacity of the transportation network which determines the extent to which such flows can take place, and thus the future course of development.

The present study will examine transport infrastructure in the Indian economy as a determinant in economic and industrial development processes. However the exact nature and direction of causation in this respect is not universally established. Contemporary opinion on the role of infrastructure in economic development thus falls broadly into two divergent schools of thought. One school maintains that development of social overheads should be made with anticipatory demand in mind, the rationale being that creation of social overheads generates cost-reducing external economies. The other protagonist school argues that transport and other infrastructural facilities should be created in response to capacity shortages within the economy, since long-term investment carries risks and uncertainty when made in anticipation of long-term demand that might not even materialise. Whether the two sets of opinions are so absolute as to be mutually exclusive, or whether each individual opinion is a product of the economic circumstances and the times in which it arose will also need to be examined by reference to the long history of railway development across the world.

1.2 Global Availability of Transport Infrastructure

The growth of global transport infrastructure has been mainly concentrated over the period of 200 years since the closing years of the 18th century. It has thus kept pace with and is probably causally linked to phenomena such as the growth of world trade, colonialism and industry over the same period; however, where the 19th century might be called the 'Railway century' because of the assimilation of steam power in transportation, the present century is appropriately the 'Road century' because of the advent of the internal combustion engine. A difference would therefore be noticeable between the country-transport networks developed over the previous century as compared to those developed more recently. Most industrialised countries, especially in Europe, have extensive railway networks in addition to highways and superhighways, raising spatial incidence of transport infrastructure to very high levels. Countries in the LDC group generally have much lower incidence, and also show pronounced absence of railways relative to roads. However an obvious exception might exist in respect of the decolonised countries, especially those in South Asia. The development of railways in these countries proceeded almost alongside railway development in the colonial nation - and as will be seen, with capital being solely sourced from the latter. That the colonial power in this respect was almost invariably Great Britain is not merely coincidental, since the economic might that the British Empire acquired was largely the product of its colonial enterprise, including provision of transport as infrastructure for this. Other European colonial powers *e.g.* France, had much less to show by way of transport development in their colonies such as Algeria. Although the same colonised countries did, along with others, acquire an overlay of roads in the 20th century, the fact remains that the penetration of railways that already existed meant that fewer roads were constructed till their eventual independence. It also stands to reason that since foreign (British) capital in large amounts had been sunk into railway construction and operation, there was obvious reluctance to open the tranportation sector to competition from roadways.

Hence a cross-comparison of transport infrastructure across the countries of the world shows wide variability both in the relative extent of road and rail networks, and in terms of their spatial incidence. Cross-comparison nevertheless proves illuminating and establishes uniform principles for evaluating the incidence of transport infrastructure across alternative modes.

Geometric cross-comparisons of the mode-wise extent of transport infrastructure across 130 low, medium and high income countries in 1992-93 are presented in the figures. The two important incidence indicators that form the horizontal and vertical axes of *Fig.*1(*a* & *b*) and 2(a & *b*), are respectively, *availability of transport across space,* and *availability of transport across settlement,* denoted in the figures as ratios for

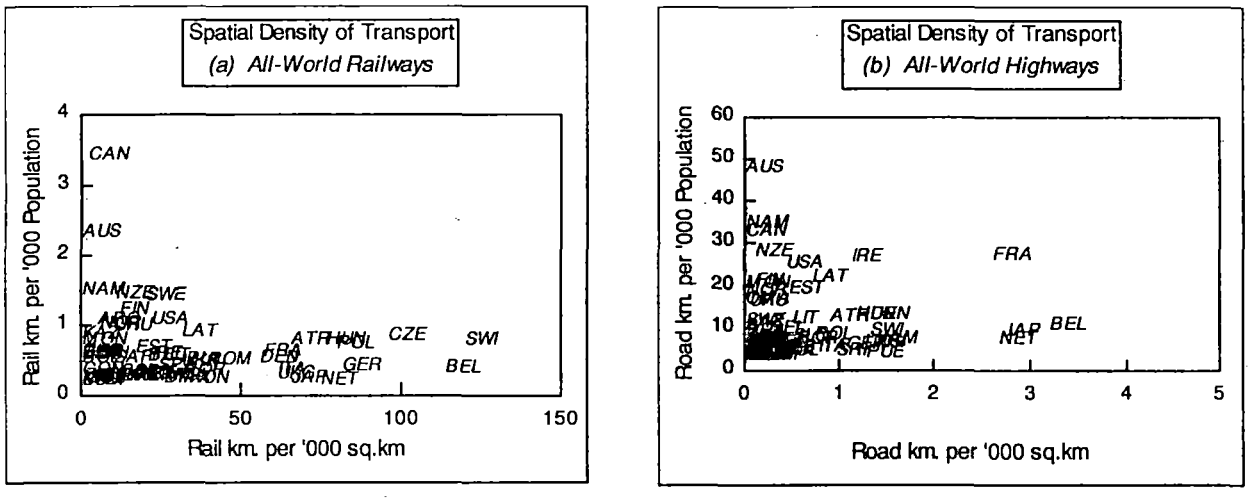
modal route-kilometres per thousand sq.km area and modal route-kilometres per thousand population. The ratios are computed separately for railways, highways - the two principal modes of surface transport. The former ratio indicates the spatial density of the modal network, while the latter reflects the relative access the population has to it.

The formulation of the indicators used here might be briefly explained. As such, the necessity for transportation arises both because of the distances that have to be traversed within a country, and because of the size of the population that needs to be provided transport services. In practice, either principle - or both, in exceptional circumstances - have been responsible for the peculiarities in evolution of modal networks in different countries. Since as background to the incidence concept, the idea of population density determines this choice, cross-comparison can be made on the incidence features of railway, highway and all-mode transport in the countries under reference, against foreknowledge of their spatial and demographic dimensions.

1.2.1 The Country Analysis

Comparing the relative position of railway and highway networks in all 130 countries, as presented in Fig 1.1*a* and 1.1*b*, pronounced differences are noticed between the general incidence characteristics of railways and highways. While railway incidence seems particularly influenced by the settlement density of a country, the incidence of highways is determined more by its spatial size. In Figure 1*a* for instance, countries with highest spatial densities for railways are all European and include densely-populated Switzerland, Belgium and Germany, among others. In contrast, and inspite of high spatial density of railways, the correspondingly high density of population ensures that relative access of the population to railway transport remains much smaller than in Canada and Australia, and even in Sweden, implying that countries with the highest spatial density of railway coverage are also countries with the highest settlement density of population. Railways are therefore *settlement*-oriented. Bearing in mind that Canada and Australia are rather exceptional in terms of their phenomenal size and nominal population, therefore acting as geometric outliers in the *space*-oriented cluster, the dominance of settlement density in determining the incidence of railways is even more apparent.

***Figure* 1.1: Spatial Density of Modal Transport by World Countries**



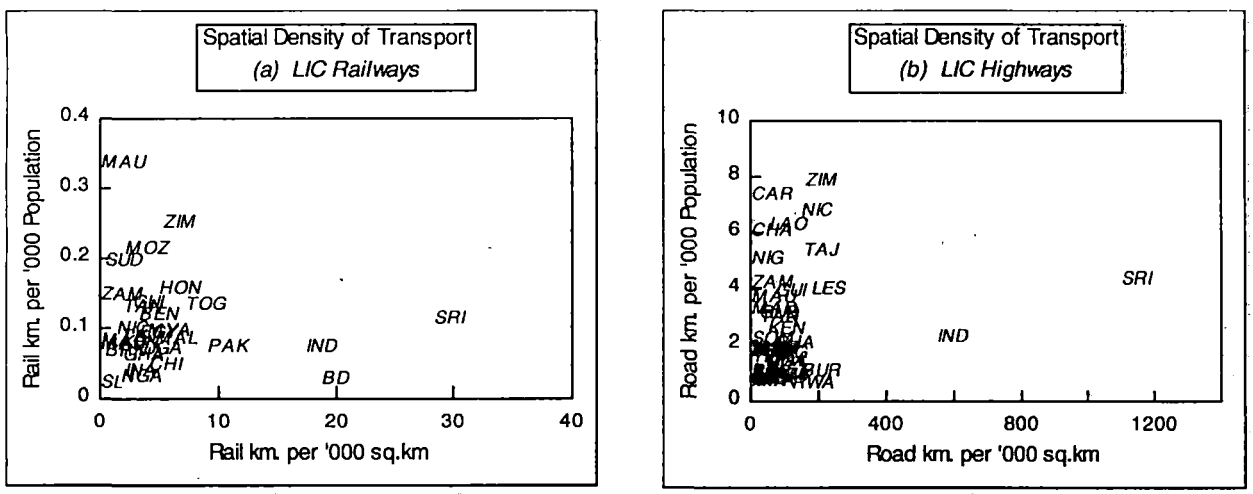
Sources: Plotted Inter-Country Data using standard 3-letter country identifiers, from *World Factbook 1993*, CD-ROM Database incorporating US State Department Data, Wayzata Technologies, 1993

The incidence of highways, on the other hand, seems largely complementary to the incidence of railways in the countries considered, implying that the development of their highway networks is *space*-oriented, and supplements rather than duplicates rail networks in most countries. However, a few countries maintain their relative cluster positioning in *Fig* 1.1*a* and 1.1*b* Thus the vast territories of Australia and Canada have low spatial coverage by both highways and railways, and are therefore *under*-served by transport even though their sparse populations may also not warrant higher density of coverage. In contrast, countries like France, Netherlands and Belgium in Europe, and Japan show high spatial densities for both railways and highways, indicating that while the high density of settlement in these countries has necessitated parallel development of both transport modes, the development may well have been competitive, especially over periods when the

total traffic generated by their economies proves insufficient for full utilisation of total transport capacity. Although the UK and the US lie closer to the general cluster of countries in the figures, and are therefore not individually distinguishable among them, they too show competitive development between railways and highways. However, where the order of this competitiveness in the UK is similar to that in France and shows settlement-orientedness, it acquires space-orientation in the US, which is a large country with a relatively sparse population.

Fig 1.2*a* and 1.2*b* confine country analysis solely to economies with lower income, including India and China. Although as would be expected, the transport coverage provided by rail and highways is lower than in the high income countries which are more visible in the earlier figure-set, it is seen however that railway development has been settlement-oriented only in India, Bangladesh and Sri Lanka, and to a lesser extent in Pakistan, and is otherwise well below 10 route-km/1000 sq.km in the 42 low-income countries [LIC] being considered. India, Pakistan, Bangladesh and Sri Lanka thus all act as outliers to the general pattern of railway development established by the LICs, and it is curious to note that they also have a shared history of railway development under the British, with the present railway systems of the first three countries having once been part of a common pre-Partition railway system. As will be seen presently, much of the early development of railways across the world drew finance from British capital markets and it is only natural that the former British colonies should have benefited by acquiring more extended railway systems than other LICs.

***Figure* 1.2: Spatial Density of Modal Transport in Low-Income Countries**



Sources: Plotted Inter-Country Data using standard 3-letter country identifiers, from *World Factbook 1993,* CD-ROM Database incorporating US State Department Data, Wayzata Technologies, 1993

However comparing the incidence of railways with that of highways, only two of the former four countries remain outliers - namely, India and Sri Lanka. Sri Lanka, because of its compact island nature, is better endowed in spatial terms by transport facilities. India, because of the strong thrust towards the building of road infrastructure in its Five-Year Plans, has expanded its spatial access to highways, although to some extent, this thrust has also arrested the further expansion of railways over the same period. Most LICs on the other hand form a dense cluster along the vertical axes, implying uniformly low coverage by transport infrastructure in general, and accessibility of transport services to the population being arrayed in terms of the ratio of country-population to country-size.

The first overview of global transport infrastructure that has been gained through the foregoing country analysis enables general principles to be laid for evaluation. On the strength of these principles, it is observed that the LICs are under-served by transportation in both spatial and population terms, and that a greater degree of complementarity characterises the development of railway and highway infrastructure in the high income countries [HIC] than in the LICs. Inspite of these broad features, relative similarity is noted between the development of transport modes in India vis-a-vis the Western European countries indicating that fruitful study might be made of the parallel development experience.

1.2.2 *Per Capita* Analysis

A feature that has remained invisible in the foregoing analysis is the direct role of population-size in determining transport access. This is now examined by plotting country positions for *per capita* spatial access to transport by individual modes against *per capita* spatial access to transport by all modes, in Fig. 1.3*a* & 1.3*b* for railways and *Fig.* 1.4*a* & 1.4*b* for highways. Comparing high-income and low-income economies, nearly perfect correlation is noted between plotting variables for highway transport when countries are ordered in terms of the relative sparseness of populations, although *per capita* road access is considerably lower for the LICs. This would establish, as a general principle, that the road access ratio for the population is positively related to sparseness of population, and that relative variation between countries in the length of highway networks does not materially change this order of relation.

***Figure* 1.3: Spatial Access of the Population to Railway Transportation**

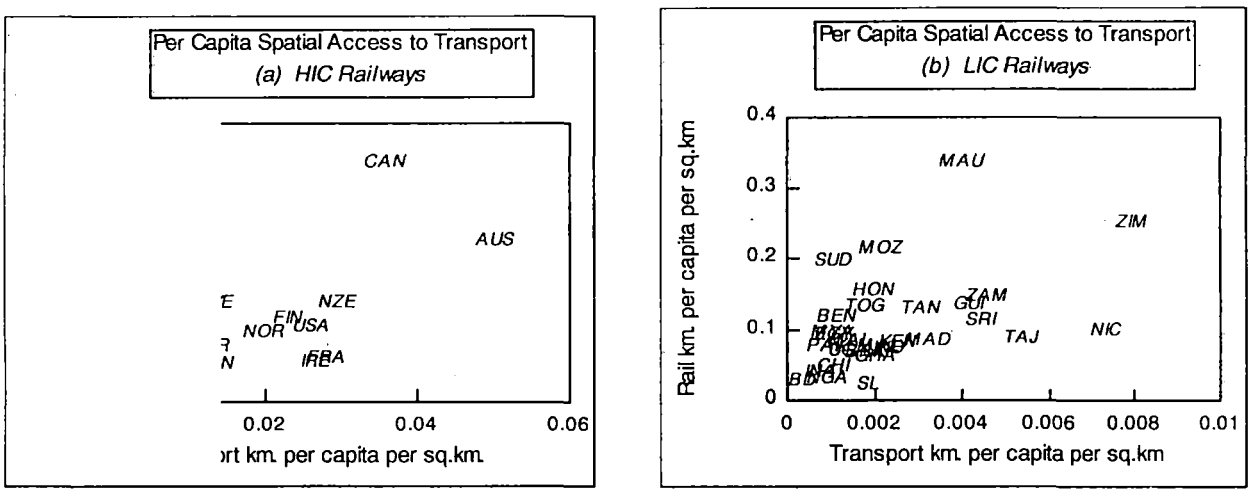
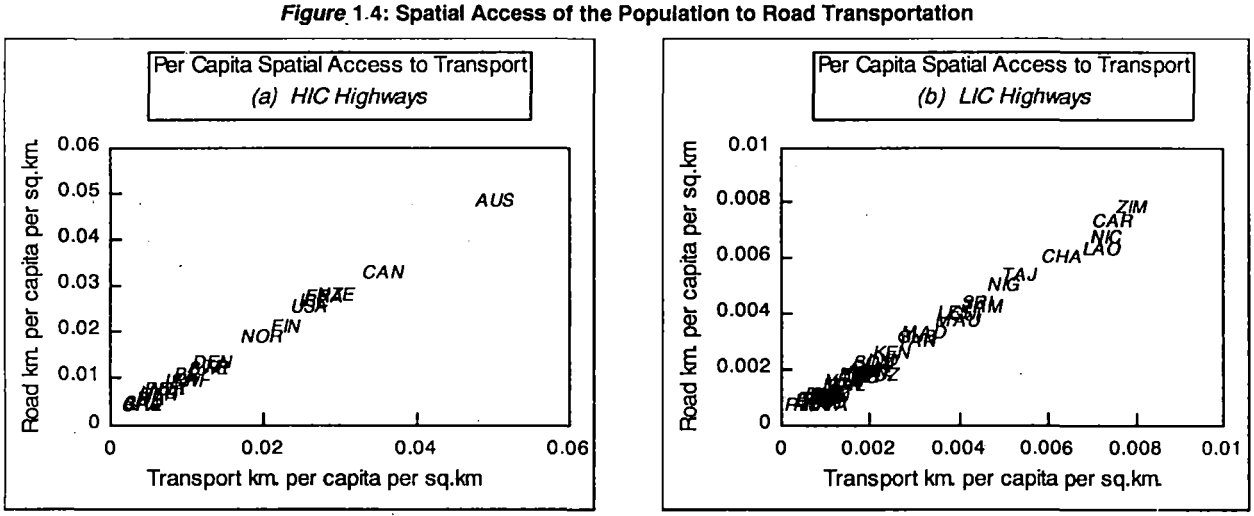


***Figure* 1.4: Spatial Access of the Population to Road Transportation**



Sources: Plotted Inter-Country Data using standard 3-letter country identifiers, from *World Factbook 1993*, CD-ROM Database incorporating US State Department Data, Wayzata Technologies, 1993

The same arraying of countries is retained for the corresponding railway plots. However the cluster plots for the railways show more obvious scatter, with the degree of scattering increasing for the LICs. Ignoring obvious outliers like Canada and Australia among the high income countries, and Mauritania and Zimbabwe among the LICs, all of whom share low population-to-land ratios, the principal difference between the former and latter country-groups appears in the directional spread of the cluster. While the higher income countries are more closely aligned along a SW-NE direction and would therefore show higher correlatedness between plotting variables, the position of the LDCs indicates more amorphousness in clustering behaviour. In general, the LICs are more deficient in railways as compared to highways in *per capita* terms, since railway transport inequalities visible in *Fig.* 1.3*b* - which in most cases are railway deficiencies - are made up when the highly correlated highway plot is also considered.

between 1850 and 1870 and its reverberations were felt in many countries of the world, including India. It may be noted however that the headiest expectations lasted during the phase of railway construction and had petered down by the time the railway companies entered the phase of railway operations and traffic competition. Another factor behind the reversal of expectations was the realisation by governments that railway enterprise could not remain unregulated, because of the transport monopolies that had been created. A third and more crucial factor was the drying up of foreign private capital flows after the end of speculation and economic downturn.

1.3.2 Private Funding for Railway Infrastructure

Although infrastructural development promises immense downstream benefits in terms of activity and productivity increases, the burden of financing this development is often daunting. The 'low-level equilibrium trap' identified in many LDCs springs therefore from both funding and infrastructural inadequacy, since the domestic funds flows can only be increased by putting infrastructure into place. In the absence of domestic solutions two options remain open of government subsidies *i.e.* 'merit goods' funding, and of foreign borrowing. In the context of railways, the problems are aggravated by the low returns to capital, compared to the scale of capital funding, which makes it difficult for railway enterprise to capture enough revenues to pay its own costs. As had been noted earlier, this situation argues for the provision of government subsidy so that infrastructure is not undersupplied. Similarly, the existence of high domestic or foreign interest rates which prejudices the public-borrowing option, argues for the same. In these circumstances, the past trend in most infrastructural systems had been to rely mainly on public finance and subsidised operation, even at the cost of efficiency losses leading to the sacrificing of future development in infrastructure.

Reviewing the history of infrastructural finance, the role of government guarantees and subsidies is well-evident and will be commented upon at length in the next chapter. The primary function behind these was to encourage the inflow of foreign capital, in an uncertain area where domestic investors feared to tread. Even the early American railroads, lavishly supported by land grants and federal construction subsidies, sought their finances from the money market in London.[6] While the railways constituted the major infrastructural innovation of the 19th century, vast investments in the same period also went into canals and dockyards, telegraph systems, roads and turnpikes, electricity utilities and the like. As a result, the success of the railway companies in competing for foreign capital lay in the magnitude of concessions they were able to secure from governments and the high returns they were consequently able to offer to their investors.

But, as remarked earlier, the period of railway capitalism might be subdivided into the phases of railway *construction* and railway *operation*. While the first of these phases were secured by various grants, subsidies and guarantees, the second was unsecured. The immediate consequence of the convergence of railway construction activity throughout the world financed by a single financial market was that with the eventual completion of the construction phase, the majority of the world's railways entered the unsecured operational phase together. The drop in financial returns that each underwent coalesced into a major market slump occurring in the 1870s, with no coordinated world revival of railway funding having occurred since then. This would hold that the success of privately-funded railway infrastructural development - for which the 19th century is held up as example - was in the Schumpeterian sense, a 'historic individual', and cannot argue for private infrastructural finance as necessarily being the superior alternative.

1.3.3 Infrastructure Finance Theory

An important consideration in treatments of the finance of infrastructure is the concept of *asymmetric information,*[7] whereby the financial promoter is more informed about the risks of enterprise than the investor. Where infrastructural capital is raised through private promotions within the stock market, the possible extent of asymmetry that would prevail is the obvious cause of high returns promised to the investors in risky capital issues. The resultant rise in interest rates outprices safe capital issues and increases the overall risk portfolio of the market, in the financial phenomenon technically known as *adverse selection*. It thus becomes easier and easier for the speculative promoter to promote high-risk infrastructural issues, since the investors are lured away from safe investments into the vortex of higher and higher investment risks by promises of higher interest rates. This is technically the problem of *moral hazard*, as a consequence of which higher interest rates actually lower the investor's expected returns. The more marked the information

asymmetry, the greater will this divergence be.

Although the presence of speculation with asymmetric information pushes up interest rates in anticipation of high rates of capital returns, it also goes to the detriment of those high risk projects which may yield high potential returns because of the crowding of purely speculative issues. In such cases, the promoter's liability to subscribe to the capital of his own issue rises, and will lead to reluctance on his part to enter the capital market unless he finds his position protected by some form of interest guarantee. The defect of this system is that the guarantee also promotes the diversion of resources to nonproductive purposes.[8] In extreme cases, the promoters may resort to 'bankruptcy for profit' in a phenomenon called 'looting' noted in Akerlof & Romer [1993].[9] Financial prudence advises closure of firms whose earnings from the liquidation of assets outweigh the returns from enterprise. One modality of looting involves the draining out of operating funds receivable under the guarantee from the firm well in advance of this so that only a shell operation remains at the time of closure. Another more reprehensible method falsifies the accounting returns of the shell company to show low cost-to-return ratios in order to acquire new capital inflows on other ventures. The crux for fair operation of government guarantees on infrastructure is close monitoring of performance, which would however also undo the information asymmetry that commenced the process.

The theoretical analysis made above adequately reflects the prevalent mood of the financial market during the 19th century railway boom. As long as high expectations were fuelled by the abundance of new railway construction projects, investor response remained high. Following the completion of most major construction projects by the 1870s, a financial turning point was reached. Although private capital did remain engaged in railway operation for a considerable time to follow, both credit needs as well as informational asymmetry were much less in this case and hence did not lead to the earlier response from the financial markets. The next development following the commencement of large-scale railway operations was the advent of extremely low returns to operating capital so much so that the anticipation of shared profits under which the government had originally extended guarantee support failed to materialise. Instead the governments found themselves making a constant outgo of funds to finance the guarantee. It was this unhappy experience - common among all governments who had financed railway construction through guarantees and subsidies - that led to the spate of railway buy-backs and takeovers once the promoters' sleight-of-hand became apparent.

Early transport infrastructure projects were usually characterised by extreme asymmetries of information because of their innovative nature, the lack of demand projection, and a dearth of respectability among their promoters. An added factor was the fact that the actual resource potential of newly-opened regions, on which potential traffic would depend, had not been properly assayed. In contrast, project promoters and their privileged friends resorted to vast speculations in the lands through which the railways would pass, because of insider information. Financial market reluctance to finance uncertain infrastructural issues in distant lands was overcome when the State acquiesced to the provision of an incentive structure of operating subsidies, interest guarantees and collateral land grants to underwrite the issues. Although the eventual impact of each was identified, the respective incentives operated on different individual sentiments in the market. The subsidy reduced the risk premium on the investment. The guarantee relaxed credit rationing by counteracting informational asymmetry. The land grant provided security and reduced collateral risks. Much of railway finance in the 19th century was raised in these circumstances, from foreign financial markets, and provided a major impetus to further development of these markets.[10] Even in the comparatively-rich US, railway securities constituted more than half of the outstanding foreign debt till 1914.[11] The alternative model of local finance was experimented with, *e.g.* in France's Freycinet, but proved a failure in mobilising adequate infrastructural finance. Since with time common knowledge has grown, the asymmetry has lessened, making it extremely difficult today to raise finances in the same degree and manner.

1.3.4 Promoting Infrastructural Finance

Investment in infrastructural facilities like railway transport is also characterised by the fact that a major part of the initial capital investment into the building of networks is either irrecoverable or recoverable only at some future date when the scale of operations has mounted to a point when credible operating profits begin to appear. Moreover, a large part of such costs need to be 'sunk' and ignored in future references to 'profit', leaving only incremental costs undertaken to keep the service running open for consideration.[12] The period of low returns to capital is justified by the gestation lag between the building and operation of an infrastructural facility. In the best of circumstances, the rate of returns remains low because of the social nature of the good

provided makes it necessary to keep user-prices low to the user to maximise the stream of social benefits. Most infrastructural services are thus priced at marginal cost so that only the incremental operational costs are actually recovered. The unattractiveness of investing at such low rates of return adds to the difficulty of mobilising funds for the development of infrastructure. Experience with private commercial investments undertaken towards developing the railways in Britain and India during their early phase as also with the more recent entry of private investment in India's power sector shows that such investment is forthcoming only at premiums where governments assure a guaranteed rate of interest irrespective of capital gains or losses. As just illustrated from the literature, the incentive can prove highly ineffective when the investing companies begin to find it more appropriate not to show any capital gain on their investment in their accounts lest the guaranteed interest be withdrawn. Governments bearing heavy interest burdens because of the guarantee eventually find it more worthwhile to invest on their own rather than carrying the burden any further.

The alternative sweetener of offering free land grants to companies investing in infrastructure which was tried with some measure of success in several countries including the US and India is particularly attractive to overland transportation infrastructure such as railways and roadways, since these require large acquisitions of land to build route networks. However this option becomes progressively more limited as lands are occupied by settlements necessitating evictions for the land guarantee to operate, and is therefore less appropriate to present world circumstances.

It is seen then that the major hurdle every country has had to cross in the development of infrastructure is the limited access that such projects have to commercial sources of finance. The twin reasons behind this are the 'lumpiness' of the investment required for building infrastructural capacity, and the recurrent need for subsidiary investments thereafter in order maintain capacity at optimum levels. Financial investment in transportation infrastructure and in the railways in particular during their formative period, as experience widely shows, has generally been organised by private enterprise. Even so, intervention from the State has come sooner or later because the strong public-good characteristics of transportation services associate its development with the fulfillment of social welfare goals that a profit-maximising private firm may have scant regard for. Partly because of this, the huge scale of financial demands for infrastructural investment render it effectively impossible for private enterprise to develop more than a nominal component in any infrastructural project. Experience with private investment during the development of the early railways has therefore generally demonstrated that while private endeavours during the primary developmental stage have the merit of consolidating the capital made available from the partial contributions of several small investors, investments on that low scale only bring about piecemeal development of transportation and cannot ultimately provide an integrated railway transportation system.[13]

Another set of impediments to infrastructure development arises from the *opportunity costs* of investment, *i.e.* the scarcity-determined choice criteria by which an investor selects from among alternative investment avenues. Although accumulation of idle capital in Britain during the early 19th century from the fruitful consequences of British mercantilist and industrialisation policies of the earlier period provided the capital required for the development of domestic railways as well as for later commercial investment on the American and Indian railways, such investment outflows took place only after the limitation of opportunities for investment at home and the promise of phenomenal returns elsewhere. The question might arise whether similar circumstances are today being mimicked in the infrastructural development scenarios of the world. The major circumstance that is missing in today's scenarios is the colonial situation of the previous century, where it served the interest of the colonial power to build infrastructure in order to maximise the benefits derived by the home country from its colony. Thus where government policy in the home country may have once proved instrumental in determining the direction of international financial flows, in today's scenarios the borrowing countries have to compete equally or unequally for capital funds with the home country. Whereas the scarcity of capital is definitely more acute, such as in the LDCs, the ability to offer attractive rates of return is diminished both by the lack of riches and the lack of spending power. The infrastructure sector in such countries is therefore caught in a trap of government indecision on whether to go ahead in building infrastructure or to augment other spheres of economic activity. Similar dilemmas are faced by potential financiers in investment decisions on whether to invest in infrastructural sectors which offer low rates of return but stretch their yield for a considerably long period, or in other activities which offer more attractive returns but for shorter durations. While both dilemmas can be theoretically resolved through yield-

maximisation, the presence of greater levels of uncertainty and high risk premia in infrastructural investment eventually queers the pitch against it. A similar perception also guides the choice of investments between alternative transport modes, such as between the roadways and the railways.

Such weaknesses of the private financial process and the character of railway transportation as a public good has shifted the onus of developing railways almost entirely to governments, and even though ownership of the railways may differ between countries, the dominant presence has been of public ownership at some stage or the other, with the sole exception of the US. Circumstantial changes that may have taken place in this in context of the present moves towards privatisation would need to be reviewed from the context of the concerned countries, rather than comparing and emulating dissimilar situations.

It will be interesting to identify the considerations governing infrastructural investment decisions in the present global reference. While an LDC like India still has to cope with acute capital shortages, it is offered the option of either curtailing its investments and slowing down economic growth, or borrowing from surplus countries to fulfil production targets and social objectives. This is a major factor guiding the largest directed flows of project investment today into infrastructural areas, since such finance as is forthcoming within the country is too small to sustain investment levels. While the non-viability of total dependence on private investment for the development of railways needs no further elaboration at present, it is necessary to explore the ways and means open to the State in funding the infrastructure for future development. Fiscal policies of the government may well provide resources for public investment, but other imperative needs may require that the major portion of such funds be diverted to other important economic sectors such as agriculture or industry to maximise production and employment. This leaves governments with the option of public borrowing by attracting investments into bonds and equities. While the first of these imposes the greater burden in debt repayment - a feature that is plaguing the railways in India presently, the latter involves both risk factors and uncertainty, as well as dividend liabilities. The history of infrastructural finance shows in many instances that the breakthrough in such circumstances has been made by tapping foreign financial markets where idle capital exists, although private participation in infrastructural development has eventually to be bounded by government regulation. Although in the absence of direct equity participation for reasons of this kind, investments in infrastructural bond issues by foreign investors may be welcomed, it will be worthwhile to remember that such investments are forthcoming only because the opportunity cost of investing in such bond issues is presently low, and when better investment opportunities become available elsewhere the flow of money into the development of subsidiary infrastructure is liable to dry up, leaving the borrowing country in a deeper crisis than before. Thus caution needs to be exercised when initiating such capital flows.

The only recourse left ultimately to the railways and other infrastructural organisations is to augment their own resources through efficient utilisation of assets and minimisation of operating and other costs. While efficient cost structures therefore need to be devised, the important strategic initiative comes from the ability of the government or the organisation to market transport services, and to increase traffic elasticities by creating capacity in excess of demand. While pricing modifications such revisions of tariffs and fares can also help in building up a surplus in infrastructural revenues, it is important to channelise these surpluses into expansion and modernisation of operations in order to attract new traffic, rather than squandering them in other areas which will yield few developmental dividends.

1.4 Infrastructure as a Public Good

Infrastructure, also referred to sometimes as *social overhead capital*, is defined in the classic literature as any sector providing "services basic to any production capacity"[14] and as including "all public services".[15] Particularly in case of transport infrastructure, certain classical arguments have commonly been advanced for the public ownership of utilities. One of these drawn from the economics of information suggests that the state, being more informed than the market, is more qualified to make provisions in this respect. Other arguments draw on 'natural monopoly' attributes, negative social and environmental externalities, distributive equity of benefits, or the provision of merit goods, to advance the case. However, the generally low world levels of investment in expansion and renewals of facilities in a period when demand for transport and other infrastructural services has been rising have led to obvious bottlenecks over the 1970s and 1980s, and to mounting concerns in many countries about the adequacy of infrastructure to sustain economic growth and development into the future. Such concerns have related in part to the pricing of these services as public goods, in which case inadequacies and inefficiencies in infrastructure have been directly attributed to *market*

failure, rather than to failures in the underlying government processes.

Areas of conflict emerge on examining the literature on the role of transport infrastructure in promoting economic development. The approach of 'new economic history' applied in railway studies such as Fogel [1964],[16] Fishlow [1965][17] and Hawke [1970],[18] is sometimes cited for evidence of 'what would have happened if specific transport infrastructure had not come into place'. The approach however suffers from obvious limitations because of its heavy reliance on hypothetical econometric projection. Studies of long term macro trends, *e.g.* Biehl [1991][19] or Andersson & Strömquist [1988][20] indicate strongly positive association between the provision of transport infrastructure and economic development, without necessarily establishing the causality between the two. On the other hand, studies at micro level on the regional and local impact of transport services, *e.g.* Blum [1982],[21] Nijkamp [1986],[22] Rietveld [1989],[23] and Lakshmanan [1989][24] lead to ambiguous or even opposite conclusions on the terms of association between transport development and spatial disparity, since increases in production in specific regions can be deemed to have occurred either because of site-specific creation of new economic activity, or because of transfer of economic activity from other regions. In particular, issues of *public* versus *private* provision of transport infrastructure bring in comparisons between the relative transaction costs and efficiency levels of each, which are seldom explicitly addressed in the literature.

Canonical *public goods*, which in the terminology of public economics[25] involve the externalities of *non-rivalry* and *non-excludability* on the part of their users, can only really exist in market situations when supply is considerably ahead of demand. This creates obvious conceptual difficulties in the efficient pricing of such goods, since the imposition of user costs that realistically reflect the costs of production of public goods necessarily introduces some measure of economic excludability, even in unsaturated markets. More serious difficulties are created by underpricing of scarce services in transport markets which are usually congested, where overuse of facilities by one user physically excludes some other, and the canonical requirement in case of public goods, that all users consume the service, is not met. Another argument, again drawn from public economics, relating to the provision of transport infrastructure on noncommercial terms dwells on the conceptual apparatus of *merit* goods, satisfying "wants so meritorious that their satisfaction is provided for through the public budget"[26] The social obligations of railway undertakings that require the provision of certain services priced well below cost, or the continuance of specific services under conditions where they cannot break even because of the dearth of adequate traffic, are much cited instances, although it may also be noted that the provision of merit goods does not *ipso facto* enjoin that they must be provided under public ownership.

Another argument that becomes intuitively appealing particularly when advanced in relation to railway infrastructure is based on technical indivisibilities in sunk and fixed capital costs which generate increasing returns and decreasing unit costs as the scale of infrastructural services expands. The implication that the optimum scale for provision of infrastructure is too large to be efficiently provided by multiple enterprises, and therefore that a *natural monopoly* exists, argues for the activation of public ownership so that services are provided at reasonable and regulated costs, instead of being set at the high potential monopoly rent-seeking levels that market-based pricing would inevitably lead to. One strong counter-argument that militates against this derives from the spatial impact of centralisation and the distortions in location that can result from it. Other weaker arguments question the degree to which infrastructural costs might actually be deemed 'indivisible', and whether the regulation of monopoly rent-seeking actually necessitates public ownership or legislative protection.

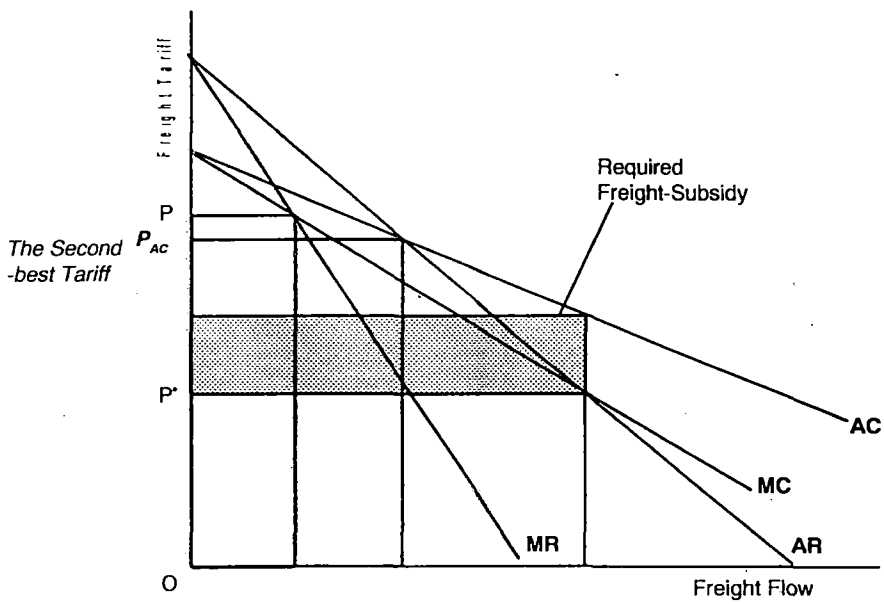
The more deep-rooted historical rationale behind public provision of transport infrastructure however derives from apprehensions that privatised provision would lead to suboptimal supply and inadequacy in services, both in terms of their density and spatial spread. Whether such fears are verified by fact requires examination of the comparative history of infrastructure development to see whether phases of private ownership led to squeezing of investment and services, and whether transition to public ownership reversed that trend. In the contemporary situation of infrastructure through much of the world, it is widely apparent that public ownership has not in fact solved the problem of resource inadequacy which lies at the heart of the inadequate provision of infrastructural services.

1.4.1 Pricing Infrastructural Services

As such, the issue at the core of both historical and theoretical approaches to the allocational efficiency of infrastructure and therefore to its ownership - whether private or public - relates to the pricing of its services. Were publicly-owned transport infrastructure to be treated as a non-marketable merit good, any consideration of demand and supply would be irrelevant to the determination of its nominal price. The problem of market-based pricing and allocation however becomes more and more tractable as one moves further and further from canonical definitions in determining the true character of infrastructural goods. Thus the questions now examined relate to whether nominal *i.e. ad hoc* pricing of the type usually resorted to is the only means of pricing infrastructural services at reasonable levels, or whether an exercise in market pricing can indeed be made.

The key feature that delineates infrastructure is in practice defined by the decreasing-cost/increasing-returns-to-scale envelope within which it is provided,[27] and any efficient pricing solution must be appropriate to this. Temporal considerations are also involved. The realisation of higher per unit prices in the short run can augment the resources invested on the provision of infrastructure in the long run. Conversely, the maintenance of equity pricing at levels oblivious to infrastructural costs in the short run can lead to inadequacies in investment and services in the long run. The question whether social and equity considerations will be served by market-determined prices would then rest on whether the prices so determined would in fact be so high as to apply a squeeze on the extent of services provided and to exclude a subset of users, or whether a tractable solution could otherwise be found. The intrinsic structure of the pricing decision involved appeals intuitively to application of second-best pricing which is examined geometrically below. Symbols in the diagram carry their usual meaning.

***Figure* 1.5: The Second-best Problem in Transportation Pricing**



The monotonically declining cost curves seen in the figure would allow a private monopolist to earn rent-seeking profit at a unit-tariff P. If instead, resort were made to Pareto-pricing at MC = AR = unit-price P* of the infrastructural service, the service-provider would sustain a loss equivalent to the shaded rectangle because of average price realisation at levels lower than average cost of the service, with the gap between AC and AR having to be made up by providing an equivalent state subsidy if the continuation of the service were sought. The general opinion in favour of public ownership in natural monopoly conditions argues that the state is best placed to monitor the efficient use of its subsidy, and any other form of organisation leads to waste of public money. But an alternative method that obviates the need for subsidies sets the price for the infrastructural service at a level that ensures acceptable returns. If the acceptable level of returns is defined at break-even levels *i.e.* at normal profit, the second-best price P_{AC} would be set at AC = AR, under the average-cost pricing principle also applicable under natural monopoly conditions to contestable markets for infrastructure outlined, for example, in Baumol *et al* [1982][28] In sufficiently open i.e. *contestable* markets, monopoly suppliers of infrastructural services would be obliged to price at average costs, rather than at rent-seeking

marginal cost price, to maintain their market share. However, the foregone cost of services, defined as the difference between average cost and marginal cost, would constitute an uncovered social loss attributable to the non-marginal character of pricing. The standard solution[29] under such conditions is to move from single pricing regimes to discriminatory pricing regimes which would recover the full costs of service provision without setting prices so high as to result in exclusion of users, so long as the relevant cost curves were not significantly higher than the disaggregated demand or AR curves. The system is usually followed for yield management and revenue maximisation in several instances of contestable markets such as airline cartels, motor syndicates, etc.

Even though inadequacies in infrastructure are often blamed on market failures rising out of the public good character of infrastructure, reasoning of this kind can often obscure the underlying failures of government policy. Government intervention of the wrong kind can worsen the overall impact of market imperfections or failures, in situations where market forces might have offered better solutions. One of the widely prevalent practices in partial charging for the use of infrastructure for recovery in addition to user rates is to resort indirectly to lumpsum charges, the yields of which are usually not earmarked for reinvestment in infrastructure. Thus, to the extent that infrastructural costs to the service-provider exceed direct recovery in the form of user-charges, the amount foregone represents a deadweight loss to the service provider. The case is usually descriptive of government failures in the provision of transport services, which lead also to overuse and congestion of infrastructure in the short run, and to a shortage of resources for investment in infrastructure in the long term.

The impact of the long term problem might be separately examined. Firstly, the inability to levy adequate direct rates on infrastructure use sends out wrong signals on investment needs and priorities. Secondly, long term provision of infrastructural services deteriorates both in technological quality and extent. The Public Choice School's[30] reasoning therefore argues for earmarking of all revenues realised from public infrastructure.

The new interest that has arisen around privatisation of infrastructure as a potential solution for the problems of cost-inefficiency and capacity creation in publicly-owned infrastructure is now examined. Although privatisation is supposed to internalise congestion costs and optimise long run capacity, in practice such assumptions too would require idealised conditions unlikely to be fulfilled by real world situations. Arnott *et al* [1993][31] demonstrates that while a private monopolist might set capacity at cost-efficient levels to minimise total costs, including those he transfers to the user, the service charge levied does not necessarily yield socially optimal usage, because of the markup imposed, and because the privatising monopolist's markup price would turn out to be rather sticky even when cost-efficiency ultimately reduces marginal cost.

1.4.2 The Theory of Contestable Markets

Infrastructural pricing problems of the kind alluded to above - also known more illuminatingly as problems of 'bottleneck' pricing - spring mainly from the nature of the infrastructural goods concerned and the restrictions to entry that apply to the domain of the infrastructural activity because of the presence of sunk costs. This is of particular relevance to the railways, where the infrastructure required involves a high percentage of sunk costs and where, because of social objectives, railway services assume the form of a public good. Public ownership of infrastructure has generally followed the apprehension that bottlenecks become inevitable if the alternative of market pricing is resorted to, because of the diminished productive and allocative efficiency of the market mechanism in services that have pronounced merit goods characteristics. While providing sufficient instance of this, the railways also constitute an infrastructural sector where effective competition is not possible immediately owing to the vast spectrum of services offered. The various modes of operations that have been followed historically by railway systems across the world have therefore ranged until very recently, from outright government ownership to strict control of railway pricing under statutory regulation of rates. Developments since the 1980s which reveal the inefficiencies of regulated systems have however seen a consolidated movement towards more open railway pricing systems, as the underlying theory of infrastructural pricing has become more and more robust.

A theoretical transition to deregulation theory was accomplished in the early 1980s following developments around the theory of contestable markets. Most concisely stated in Baumol *et al* [1982][32], the theory holds that in *contestable* markets *i.e.* markets where no significant sunk costs are involved and no barriers prevail on entry into the market, only internally-efficient firms offering an optimal multiproduct mix and operating

at optimum scales and cost-structures that involved no cross-subsidisation, would eventually survive. However the operation of differential time-lags between the instantaneous demand response to the foray of any new contestant into the market, and the reaction of existing incumbents to his entry would endow the former with the 'hit-and-run' ability to enter and contest the market, to charge short-term prices above competitive rates, and to exit the market without having to bear any exit costs well before rivals could retaliate. Since the theory assumes additionally that all new contestants stand on level footing, the initial market structure prevailing in the contestable market is irrelevant. In the absence of cost-barriers to entry and exit, incumbents in the market are left no alternative but to acquiesce to the contest, and would, in the interests of survival, provide goods and services at most reasonable prices in attempts to foreclose the contest. Thus with *perfect* contestability, the social objective of welfare maximisation is ensured, as the industry under deregulation - regardless of the number of participating firms - performs under perfectly competitive conditions, where neither excess profits exist, nor can inefficient firms survive. The task of government in this context is no longer to determine and protect the 'optimal structure' of production, but to ensure instead that there are no restrictions on free entry that would limit the credibility of the threat of competition. If conditions of contestability obtain therefore, even monopoly firms will behave as if subject to factual competition, in order to avoid attracting new contestants who would destroy the monopoly.

In the case of infrastructural multiproduct monopolies such as State Railways or otherwise regulated systems with restriction on entry, imperfect contestability prevails. Static welfare could then conceivably be maximised through discriminatory Ramsey pricing[33] of infrastructural goods above their marginal cost to recover full costs. In practice however, the threat of hit-and-run entry would always impose certain restrictions on the set of sustainable multiproduct prices from which the optimal Ramsey price would be charged, and thereby hinders fulfilment of the welfare objective. In a departure from perfectly contestable markets, a multiproduct monopoly which can force barriers to entry can actually charge very high Ramsey rates for exclusive services, against less-than-competitive rates for other services. Contestability theory thus bridges the entire price spectrum from marginal cost pricing to Ramsey pricing, by introducing the concept of a contestable monopoly which could theoretically maximise its profits through average cost pricing at least-cost production, but will resort to Ramsey pricing under the threat of entry. Ramsey pricing in such cases would thus fulfil welfare objectives even though the underlying market situation would be different from the theoretical formulation of perfect contestability.

The attack that contestability theory has however faced on several fronts proves the concept inappropriate to contexts which demand the achievement of firm viability in the long run. In the railways, for example, the requirement of railway plant and capital equipment is fixed and highly specialised, and the irreconcilability of fixed costs with the assumption of zero sunk cost made under the theory renders the market uncontestable. With reference to the objective of long term viability, the promotion of contestability through deregulation of industry can in fact lead to welfare losses by depriving the incumbent firm of its ability to attain revenue adequacy and establish a viable natural monopoly by utilising the existence of sunk costs in its favour through discriminatory Ramsey pricing.

The paradoxical character of this situation now becomes apparent. Enhancement of direct competition among incumbents and promotion of contestability in the market has adverse impact upon economic welfare because it drives prices to marginal cost levels and deprives firms from the benefits of Ramsey pricing. This however is inconsistent with the theoretical underpinning of contestable markets based on their welfare-maximising structure and contestant behaviour. If vigorous intramodal competition in transport infrastructure produces revenue inadequacy, with losses in pricing freedom[34], it logically follows that sunk costs, cross-subsidisation, etc. can materially improve welfare if they reduce price competition among existing incumbents. The encouragement of contestability via enhanced competitive access is rejected therefore, because it erodes the market power required for the practice of Ramsey pricing in industries undergoing a transition to deregulation.[35]

1.4.3 Ramsey Pricing

The discussion on the Ramsey pricing system[36] extends into methods of demand-based second-best pricing, as alluded to earlier. It allows the consistent choice of welfare-maximising prices that ensure adequate revenue for a regulated multiproduct firm to cover its costs, against recurrent losses due to marginal cost pricing under conditions of decreasing marginal costs and increasing returns to scale. Since fixed costs alone

would not lead to Ramsey pricing, the system has often been applied to studies of 19th century railway value-of-service pricing[37], where owing to the presence of scale economies, railways were unable to generate sufficient revenues to simultaneously cover investment and operating costs. The advantages of the Ramsey pricing system are most apparent for natural monopolies which deter entry of other firms to maintain market power.

A point to be specially noted with respect to the Ramsey rule is its assumption of zero cross-elasticities of demand, which in turn imply the avoidance of cross-subsidisation. Ramsey pricing in effect postulates stand-alone cost tests as a means of escaping cross-subsidies.[38] However the possibility of cross-subsidy cannot be ruled out in practical applications of Ramsey pricing to multiproduct firms, since the presence of strong complementarities can lower prices of some of the services below marginal cost levels, while the prices of the remainder exceed marginal costs by more than would have been necessary to maintain revenue adequacy. But cross-subsidies of this kind also provoke the entry of new competition, thereby undermining the Ramsey pricing scheme, which can then only be tenable under an entry barrier.[39] On the other hand, the prohibition of cross-subsidies to obtain fair pricing may sometimes reduce total welfare, since it would be inconsistent with the Ramsey pricing principle for infrastructural and other services that show strong complementarity.[40]

Most theoretical treatments of Ramsey pricing are made under the assumption that the regulated enterprise sells directly to final consumers. In the alternate situation where firms purchase their inputs from a regulated firm while optimising their own costs, adjustments to the Ramsey pricing principle would ordinarily be required,[41] and only if the downstream enterprises are perfectly competitive will the Ramsey rule apply. Adjustments would invariably be needed wherever the downstream industries under reference are neither price-takers nor cost minimisers. This limits the applicability of Ramsey pricing to railway freight services which provide inputs to regulated industry. In case of the freightage of coal to rate-of-return regulated utilities such as electric power plants, the prices chargeable by railways under the Ramsey principle would need to be reduced, firstly because the utilities would tend to produce less than optimal output, and would therefore require reduction in their coal input-prices in order to augment output, and secondly because such utilities tend to overcapitalise, the price of their coal inputs would also have to be lowered in order to constrain capital investment.[42]

In cases where both competing road and rail lines are operated by regulated firms and are under complementary use, the welfare-maximising prices of individual transport services would lie below the Ramsey prices that would have been set for each transport mode separately. Such a situation might occur when the costs incurred on each service differ in magnitude, reflecting differences in routes and lengths of carriage. But revenue-pooling Ramsey pricing which maximises the cumulative revenue yields of both transport modes would ensure equalisation of marginal cost pricing at identical utility levels, which makes the Ramsey principle at least the second-best pricing method. The alternative to maximising pooled revenue would be to let one modal firm break even, whilst the other profited behind an entry barrier, corresponding to the method of 'totally regulated second-best' pricing mentioned in Brauetigam [1979].[43]

The Ramsey pricing method was devised to ensure minimum revenue yields that fully cover the costs of providing infrastructural goods. Critical evaluations of Ramsey pricing in railway applications argue that the principle proves untenable in face of the inability of railways to generate increasing returns to scale in situations of excess capacity that are marked by a high cost-to-revenue ratio. Excess capacity becomes unavoidable in the case of railway freight transport, which aims at maintaining the interregional balance in outbound and inbound payments, while not necessarily implying a balance in the respective tonnages. Empty haulage on one journey-leg thus becomes inevitable.[44] Nevertheless the debate continues, and proponents of the Ramsey pricing principle would see no hindrance in applying the method to the determination of railway rates under conditions of excess capacity.

1.5 A Review of Literature

A survey of the literature relevant to the present study has been undertaken, in order to strengthen focus on the issues and problems that need to be dealt in evaluating the development of railway transportation infrastructure in India. Areas in the literature considered to be pertinent relate to existing studies on economic and industrial development, on transport economics, and on railway transportation in particular. These are reviewed briefly in separate sections below.

1.5.1 Literature on Economic & Industrial Development

Development literature identifies one of the major challenges in economic development as being the provision of infrastructure services to match sectoral demands of the economy. World Bank national accounts, reported in the World Development Report [1994], estimate the value added by infrastructural services - of which transport is the largest - as 7 to 11 percent of GDP. It has been revealed in Ratter [1973] that World Bank lending is largely focused on building infrastructure in the LDCs, and that one-third of total lending by the Bank since its inception in 1946 has been to the transport sector. Bennathan & Johnson [1987] is a recent study of transport in general economic contexts, which explores its input-output linkages. Another study on transport infrastructure is Galenson [1987], which addresses labour absorption in the transport sector. It may be noted that transport alone commonly absorbs 5 to 8 percent of total paid employment.[45] A point of interest in infrastructure studies relates to assessments of the productivity impact of infrastructural investments. Canning & Fay [1993] is a seminal study of this nature, whose scope is extended into the empirical realm by the exploration of fiscal policy - economic growth relationships in Easterly & Rebello [1993].

It is a moot point whether infrastructure leads economic growth, or the reverse applies. This causality has been explored by a number of studies which include Duffy-Deno & Eberts [1991], Holtz-Eakin [1988] & [1992], Mera [1973], and Uchimura & Gao [1993]. The gamut of research results cover both one-way causation in either direction, as well as dual causation. It is seen in these studies, however, that veracity of the causal hypothesis is conditioned by the econometric methodology adopted, so that the studies under reference do not always concur, but may instead have differing conclusions; an example is the opposition between the two Holtz-Eakin studies where the use of more sophisticated econometric modelling in the later study reverses the conclusion. Mera[46] makes an interesting spatial analysis of the Japanese case while exploring relations between social overhead capital and production functions of regional character.

An alternative to the measurement of infrastructural impacts measures their influence in reducing production costs. A survey of studies of this genre is provided in Aschauer [1993]. Although no consensus exists on the magnitude or exact nature of the impact of infrastructure on growth, most studies agree on the high potential payoff in terms of economic growth. The role infrastructural investment plays is substantial, significant and often much greater than investments in other forms of capital, although actual realisation of payoffs depends also on the economic policy-frame. This aspect, namely the differential impact that private-sector or public-sector orientations in the policy-frame have upon the productivity of infrastructural investments, is investigated in studies such as Ford & Poret [1991], or Hulten & Schwab [1993]. Munnell [1992] is a useful perspective on infrastructural investment & growth.

Literature sources on the general economic and sectoral impacts of infrastructure shed considerable light on processes of assessment. Kessides [1993], in reviewing general developmental contributions made by infrastructure, makes an entry into associated policy issues. Peters [1990] is more revealing on the developmental impact of transport in India: analysing data from the farm sector from 85 districts over 13 states, the study shows among other things, that increased market access concomitant on lowered transport costs has resulted in an agricultural spurt that is further bolstered by the enhancements in bank support occasioned by easier communication. The role played by infrastructural status as a determinant in the location of international investment-flows is explored in Wheeler & Mody [1992]. Access to at least minimal infrastructure has important welfare connotations in poverty contexts. The urban aspect of this is investigated in terms of transport and mobility needs in Kranton [1991].

An important determinant in the impact of an infrastructural investment is of course the quality of services provided - inefficient and poorly maintained infrastructural utilities lead to unreliable services. This problem, particularly acute in developing countries, is explored in context of the electric power sector, in Besant-Jones [1993]. Another glaring feature of infrastructural services in LDCs is the unequal access provided to the poor; this is studied in terms of spatial inequalities in transport facilities, in Camara & Banister [1993]. It is pointed out that although transport services are often subsidised in order to maintain low tariffs, the attendant benefits are cornered by the more affluent, while further expansion in services is curtailed by their nonprofitability.

Infrastructural services are often public goods provided by public organisations. Arguments exist that endemic organisational failures and generally poor performances could be better tackled by privatisation rather than by public-sector reform. The alternative, in the face of sociopolitical compulsions, is to give public infrastructural organisations some measure of autonomy and financial independence. General issues and

experiences in public enterprise reforms are summarised in Shirley & Nellis [1991]. Nellis [1988] investigates the alternative of private-sector tieups through contractual arrangements, which maintain social objectives while improving performance. Trivedi [1990] is a study of such performance agreements in Indian contexts.

The explicitly public mode of organisation prevalent over most of South Asia is explored in Seabright [1993], which also addresses reform issues. Corporatisation as an alternative to official control, is one of them. The more radical solutions involve partial or total privatisation to allow free play of market forces. Reforms in the infrastructural sector thus have many ramifications. A concept central to public organisation is the existence of a 'natural monopoly' in such services because of a market situation where a single provider can serve the market at lower cost than multiple organisations. This argument has come under challenge in modern contexts where diversification in organisational services has meant that such a provider is often performing in areas outside the natural monopoly.

As such, some of the literature under reference explores the potential for sectoral unbundling - the dissolution of monolithic organisations and services into separate enterprises, very often independent of each other. Moyer & Thompson [1992] reviews the scope for this in the case of railways. Nilsson [1993] analyses regulatory reforms on the Swedish Railways. A study of railway privatisation and its effects in Japan is provided in Fukui [1992]. Ramamurti & Vernon [1991] lays out general principles for privatisation and control of state-owned enterprise.

Where infrastructure is privately-owned, quality and quantum of services provided may be maintained voluntarily through self-regulation. Gwilliam [1993] investigates the possibility in urban transport services. Voluntary maintenance of product quality in public utilities is studied in Rovizzi & Thompson [1992]. Participative means of regulation also obtain, although these entail consumer information and more competition. Sectoral reorganisation of public utilities in India are investigated in this light in Paul [1993]. The economics of regulation are outlined in Viscusi *et al* [1992], and Winston [1993].

The last major area of relevance in the literature on infrastructure development surveyed in this review relates to the sources of finance for infrastructural investment. Although public funding has been the major conduit for such finance, the issues that the literature identifies are created by the growing inability of governments to bear ever-escalating infrastructural costs. In the face of the general resource crunch that governments today face, capital spending on infrastructure is the first to be discarded; operations and maintenance closely follow. Reductions in infrastructural spending inevitably impose long-run economic penalties, but governments find this to be politically more expedient than reducing public employment or wages. Hicks [1991] elaborates on this problem in developing-country contexts.

The alternative to government funding is private funding. Besides meeting the resource-gap, private funding is regarded in some studies to be more efficient. Kay [1993] reviews the veracity of this argument. It is pointed out in Lane [1992] that government's natural advantages in raising large funding for infrastructural projects are squandered if there is an accumulation of excessive debt. Project financing has certain established principles, outlined in standard references such as Nevitt [1989]. The case for private financing of infrastructure is enunciated in Pyle [1994], Gomez-Ibanez & Meyer [1993], and Martinand [1993]. The latter two studies have transport contexts. Since modes, operations and organisational character of infrastructure undergo evolutionary transitions over time, no one paradigm might be said to exist. Grubler [1990], for instance, makes an exhaustive survey of the underlying process.

The economic envelope for the present study is provided by the economic and industrial perspectives that have determined the direction of development policy in India over the planning era. The study is undertaken in reference to the infrastructural linkages that associate these planning perspectives with the process of economic development. To this purpose, a reading has also been made of several studies which explore the Indian development experience. Among these, studies which focus on the changing patterns of industrial growth in the country over different periods prove most relevant to the investigation of transport policy linkages. In studies of this nature, the alternating high-growth and low-growth phases experienced by the Indian economy are explored in Ahluwalia [1992], Sandesara [1992], Kulkarni [1991] and Alagh [1995]. Alagh[47], particularly, addresses the infrastructural context as a rationale for recent economic reforms in India, exploring the strategies and models that underlie Indian planning while showing how past experience could help determine future development strategy. The conclusions drawn in the studies under reference are relevant for the analytical exploration of infrastructural issues in the development of transportation to be carried out by the present study.

1.5.2 Literature on Transportation

The literature sources on transportation are diverse, covering matters of policy, theory and evaluative study. Distinctions between transport as an element of infrastructure, and transport economics as a discipline are dealt with in Winston [1985], among others. Transportation is here described as a service vital to economic development, while the economics of transportation addresses ultimate 'real world' policy implications of providing this service. The developmental linkage of transportation is further explored in Sharp [1973], and is historically reviewed in various studies surveyed in the subsequent subsection.

The principal exercise in many transport studies is transport demand-analysis. Houthakker & Taylor [1970], for example, is a particularly wide-ranging econometric projection of U.S. demands over a vast range of goods and services, including various items of transport. In the case of transport, income elasticities are found to be more important to demand-determination than price elasticities. It has been pointed out in Stubbs *et al* [1980], that there is no 'correct' unit of transport output that can be used universally in studies pertaining to transport, because the unit of supply is often larger than the unit of demand. Although the problem is not acute in railway freight studies, where wagonload consignments are common, it still persists in the case of railway parcel traffic where consignments accepted are less than a trainload or wagonload.

The main theoretical interest in the character of infrastructure originates from the fact that it is the archetypal 'public good'. However as later refinements of analysis have shown, its nature as a public good varies from pure to impure, depending on mode of provision and pricing. Publicly provided infrastructure contributes positively to the growth of private sector activity and output, indicating some complementarity between the two, rather than subsitutability. Changes in the mode of provision, such as the privatisation of railways, transform the character of infrastructural goods. The growth model of Barrow & Martin [1992] therefore distinguishes between the cases of publicly-provided private goods, and impure public goods, or publicly provided goods subjected to congestion where latecomers are excluded from the user-market by saturation of services. This line of analysis is also reflected in Aschauer [1988] and Lynde & Richmond [1992].

Another area of concern in the case of publicly provided infrastructural capital relates to the degree to which this is optimal. Berndt & Hansson [1991] thus provides a measure of optimality under assumptions that infrastructure is a pure public good. In such a case the total marginal benefits of public capital are a summation of shadow values over all private sector firms, plus a summation of all marginal benefits accruing to final users. In a uncongested user-markets, total marginal benefits can then be the largest benefit that accrues to a subset of users, rather than the sum of benefits to all potential users, and must then equal the social cost of providing public infrastructural goods for pricing decisions to be reached. This raises the important question as to whether the sum of marginal benefits to all potential users can actually be estimated, in the case of public infrastructure? However, with user-market congestion, the above result would become qualified.

The traditional analysis of public infrastructure assumes typically that infrastructure is an unpaid factor in production. In practice, this does not hold since various rates and charges do apply to infrastructure, along with taxes, license fees and levies. Exclusion of these from consideration in fact overestimates infrastructural impact. Barro & Martin[48] shows that recovery of the costs of providing infrastructure as a pure public good is better accomplished through lumpsum user taxation, since social returns to investment can thereby be maximised. However when infrastructure is provided as an impure public good, the task is best accomplished by general taxation since this internalises excessive use of infrastructure. The possibility of financing infrastructure through public borrowing rather than public taxation throws up new questions. In general, the literature holds that both modes of financing have the same impact on the economy in real terms. However this would imply that private consumption and expenditure remain unchanged when short term levies and cesses are imposed as tax-means to finance infrastructure, as implied in Ricardian debt-neutrality. In capital market situations where the cost of borrowing is high, the tax-means of finance would probably be more efficient than the creation of new debt. It is only when government funds are earmarked for the development of specific infrastructure rather than for general expenditure that the optimal mix of tax and debt for infrastructural finance can be defined.

One of the other questions explored in the literature concerning infrastructural investment, and investments in transport in particular, pertains to the process of project selection. Social benefit-cost analyses which are often applied to the purpose involve evaluation of future streams of expected benefits. Theoretical issues in Cost-Benefit Analysis receive thorough attention in Dreze & Stern [1987]. New directions in project appraisal

and planning are deliberated upon in Little & Mirlees [1990], and Squire [1990]. A general consideration of problems of project appraisal is found in Henderson [1965], which stresses the need for government to evolve a common discount rate for both private and public investment, to obviate inconsistencies that might otherwise arise in project preferences. Peacock [1973] summarises the debate on the choice of an appropriate discount rate. Discount rates based on Social Opportunity Cost [SOC] and Social Time Preference (STP) are examined in Feldstein [1972], as criteria for choosing between investments in alternative transport modes. The preferred discount rate varies: Hirshleifer [1961], for example, advocates the use of the SOC rate, where opportunity cost on every investment would be measured by returns foregone on projects which might have alternatively been undertaken. Henderson,[49] on the other hand, states that since investments in transport infrastructure involve risks and uncertainties and deferred benefits, and are subject to depreciation over time, the STP discount rate which establishes present values for future consumption would be more appropriate. Sensitivity analysis is identified in van Horne [1977] and Bromwich [1976], as the commonest method for assessing and measuring risk and uncertainty associated with durable transport asset-investment; the technique is often applied to transport demand-projection studies.

The economics of transport occupies itself with concerns of the pricing of transport services. Dupuit [1844] had established the classic principle that 'the best tariff would be one which makes all users of communication pay a toll proportionate to the utility they derive from the passage', extending the Marshallian concept of consumers' surplus into transport pricing applications. Pricing considerations are important in determining the viability and long-term sustainability of transport services. Millward [1971] is a thorough analysis of the marginal cost pricing principle, which also delves into the presence of externalities in transport as implicit transport costs. The study makes the observation that social cost pricing in transport cannot be applied conjointly with marginal cost pricing elsewhere. Turvey [1968] finds long-run marginal costs to be generally more appropriate as a basis for transport prices because benefits from durable transport assets with an extended lifespan will then at least equal or exceed the value of resources expended in acquiring them. This of course has special relevance when the replacement of an asset or an expansion of the transport network is being considered. For comparison, an assessment of marginal cost pricing problems and spot pricing in the electric power sector is found in Bernstein [1988].

Problems do arise in applying such general costing principles to specific transport contexts. The marginal cost pricing principle is often at variance with governmental or developmental priority; various tariff commissions such as the Rail Tariff Enquiry Commissions in India have explored this rationale. It has been reflected in Howe & Mills [1960], that opportunity costs of railway assets, which have alternative uses, may be less than the accounting costs of interest and depreciation allocated to them. The principal problems in railway pricing, i.e. joint costs and indivisibilities, can however lead to significant and persisting divergences between short-run and long-run marginal costs, as pointed out in Joy [1971]. For such reasons, price discrimination is normally resorted to in order to maximise revenue from services provided. Continuance or closure of an existing service is then defined by whether this revenue exceeds long-run marginal costs, with this principle being particularly applicable in the case of pricing of freight services.

Another aspect of costs in transport services relates to the role of time, since the transport service essentially involves transit over time between specific spatial points. Alternative theoretical approaches have been devised to handle time and also transit-delays. Lee & Dalvi [1969], for example, use a notional 'diversion charge' in their model which subsumes the role of travel time in determining that change in period of transit or transport tariff which would induce a shift to an alternative transport mode. Else & Howe [1969], using the cost-benefit formulation to compare road and rail transport modes over two distinct traffic-zones, shows that the use of this approach does not always lead to consistent appraisal over all sectors, and that the all important factor in the incompatibility of conclusions stems from the time element.

A number of studies in the transport literature relate to government policy. Gwilliam & Mackie [1975] makes observations on how regulations and often state ownership have now pervaded all sectors of transport. Four main spheres that government transport policy controls are identified in Stubbs *et al* [1980]; these are the quality, quantity, and organisation of the transport sector, and the efficacy of the resource allocation process. Railway services have a non-marketable character which is conducive to their operation under government. Intervention in the transport market is also necessary to ensure an optimal traffic flow that utilises full capacity, and an optimal modal split in transport services which avoids needless competition between alternative modes. Other factors providing the rationale for government participation in the transport

sector include the need to exploit the scale economies available to large undertakings in situations where normal economic forces would not lead to mergers, and the need for policy controls in sectors where government has a large fiscal involvement through subsidy to operations.

The origins of transport subsidies in the railway sector are identified in Joy [1964] with government grants to cover losses on unremunerative but socially beneficial railway services. Joy [1973] details the consequences of exposure of railways to free competition in both passenger and freight sectors, which would bring pricing down to equivalence with short-run marginal costs, and would permit utilisation of railway assets (e.g. wagons) over their full life, or to the point of technological obsolescence, with the effect of reducing capital expenditure on asset-replacements. Lee [1977] introduces a discussion of objectives into the policy debate. The upshot of this discussion is that while subsidies to railway freight operations should be terminated, those to railway infrastructure (assets) should continue. The study comments on the effectiveness of subsidies in altering modal choices in transportation via resource allocation policy.

A problem that has particularly stimulated research is the optimal modal-mix in transport services, where questions of competitiveness or complementarities between roadways and railways enter. Using international cross-comparisons, Silbertson [1970] constructs a simple model which shows that rail services may not be close substitutes for road transport services. These competitive dimensions are also explored in Stubbs *et al*[50], which shows, however, that in the case of Britain, road traffic has been able to make great inroads into rail traffic. This competitive element is of course of great relevance to a study of the transportation sector in India.

1.5.3 Literature on Railways

A large segment in the literature on railways relates to the historical development of railway transportation, commencing in Britain. Aldcroft [1976], for instance, observes that the transport revolution which ushered in the 20th century was accompanied by unprecedented growth in the use of transportation services as well as the changing importance of different transportation modes. While recounting the vicissitudes that the American and British economies have been facing from the increasing demands for personal transportation, stress is laid on the pressing need for policymakers to deal with such demands expeditiously. Because of imperial imperatives, railway development in India did not lag far behind railway development in Britain. The early history of the Indian railways is comprehensively surveyed in Tiwari [1937] & [1946], Sahni [1953], MacPherson [1955], Hughes [1964], and, more recently, in Christensen [1981]. The MacPherson study, particularly, looks at railway financing in India during the formative period, while Tiwari [1937] examines the evolution of railway rate-structures in the light of Indian industrial development. The role of government in the railway development process receives comment in Christensen,[51] in respect of the efficiency and quality of railway transportation services.

With the advent of the Railway Age in the middle and late 19th century, financing of transportation infrastructure assumed new dimensions. Although the initial development came at the hands of enthusiastic private investors, this did not preclude the need for government intervention, particularly in newly-developing countries. The rationale offered for government intervention through controls, assistance and subsidies has already been covered above. The character of railway services as a public good to society and the lessons learnt about the inadequacies of private handling of this public utility have been studied in the context of US railways in Eichengreen [1995], and in Cain [1980]. The former study, while providing an overview of the transition from private funding to public funding of railways, also has the importance of identifying various negative implications of public funding in developing countries. A policy prescription of 'back-to-the-future' is suggested as a means of ruling out inefficiencies in public funding while tapping local finances, and also in order to reduce excessive dependence on external borrowing for completion of projects. Cain,[52] in extension, points to negative aspects of the 'public service' image that railways were increasingly forced to adopt, which ignored managerial and operative inefficiencies that the system might be suffering from.

Railways by nature have technical advantages over other modes of transportation: they can transport heavy loads over long distances with minimum energy consumption, and at a speed competitive to other means of surface transportation. Railway investments however entail bulk capital expenditures on a considerable inventory of immobile assets like tracks, bridges, marshalling yards, etc. Irrespective of the ownership of railway capital, the fact thus remains that huge capital assets of permanent nature have to be created for

railways to serve transportation needs. Meenakshisundaram [1972] observes that as demand for transportation grows faster than GNP, suitable policy-criteria for guiding investment have to be recommended. Returns from investment on permanent capital assets can only be realised as deferred benefits; thus this choice of present investment for future consumption assumes importance in the railway literature. Although much has already been stated contextually to the pricing of transport services, specific issues regarding railway freight rates remain that need review. In view of the fact that the railways as a transport utility are characterised by joint supply of both passenger and freight services, the joint cost principle has often been referenced in many studies. The fixation of freight rates and of different principles that can be applied to their determinations are explored in Srivastava [1971], Pulley [1969] and Chandra [1969].

In developing economies like India, where the root problems are to maintain demand-supply equlibria over all sectors of the economy, and for the provision of freight services to be made a step ahead of demand, it is quite likely that pricing considerations would be important in maintaining sectoral balances in the economy. Peculiarities inherent in investment, services provided and the monolithic structure of the Indian Railways [IR] and other such factors also impinge upon any exercise in price fixation, as observed in Srivastava[53]. The pricing principles experimented with by railways over an extended period of time include the 'cost of service', 'value of service', 'joint-costs', 'ability to pay' and 'what the traffic can bear' principles, along with flat rates and telescopic rates. While Pulley[54] suggests five major variants of pricing policy, including marginal-cost pricing, and marginal cost pricing with uniform-increment price discrimination, as viable alternatives for the railways, studies with reference to the Indian railways consider the telescopic-rate principle to be the most scientific method of assigning freight rates. The recommendations given by various rail-tariff commissions such as the RTEC are explored in this light elsewhere in this study.

Another important operational sphere of railway transportation concerns freight flows within the network. Demand-supply relationships in freight services assume formidable significance in developing economies, where any disequilibrium perpetuates instability and has deleterious effects on economic growth. Studies on this theme in an Indian context form core references for the present study. An overall picture of railway freight operations in India, including projections of anticipated railway freight demand in the future, is found in the Report of the National Transport Policy Committee [1980]. Other reports of relevance include those of nongovernmental organisations like Tata Economic Consultancy Services [TECS], reports of the Railway Board for the Ministry of Railways and various Corporate Plans of the Indian Railways, published by the Government of India. Another important source for annual information on IR comprises the Indian Railways Yearbooks [IRYB] of different years. These public documents all provide data and analyses on various issues that are also explored by the present study.

Srivastava[55] studies different aspects of the Indian transportation system with attention on its development over time, and maintains a larger focus on IR. Problems commonly encountered in railway operations in India are duly identified and noted. More recent growth trends in IR operations are analysed in a regional framework in Rao [1984], where regression methodologies are incorporated. This study offers a holistic idea of the micro-level impacts of transportation, since it is framed for a particular Indian district. Freight-structure trends and the associated modifications of transportation planning these call for in India are well covered in the advisory report of the NTPC [1980]. The Committee recommends an intermodal mix that would optimise and ensure complementarity between road and rail transportation. While energy-use considerations are also examined by the NTPC study, the lagged responses within the infrastructure-economy linkage which affect freight flows and their commodity-character, are investigated through a distributed lag exercise. A theoretical formulation and empirical testing of the consequent supply-demand imbalances in Indian railway freight transportation is made using econometric procedures in Rao & Sriraman [1985].

1.6 Railway Tariff Formation

In any sector that extends services to the economy, the equilibrating role in matching service demand to the supply of services is played by the charge applied to the users of the service. In static contexts, the role of service charges in the tertiary sector of the economy is thus very similar to the role of prices in clearing markets in microeconomic theory. However, when applied to infrastructural services within a development context, this analogy must be modified because of the dynamics of the development process.

1.6.1 The Nature of Railway Costs

Railway plant requires huge investments in fixed and highly specialised capital suited only to the provision of transportation. Although rolling stock and even equipment can be physically relocated, a very large part of this investment that goes into grading and earthworks during line construction is not switchable to alternative uses. Hence as long as railway capital invested continues to yield some manner of return, railway services continue to be provided, since the abandonment of operations means a total write-off of immobile investments. However no returns or low returns reduce the inducement to maintain railway plant and to replace worn-out assets. New investments in railway services may nevertheless be summoned if they increase the possibility of returns being yielded by the original investments, although, because of both lumpiness in initial capital costs and regulative barriers, large profits accruing to one railway enterprise do not immediately induce the competition to build rival lines. Barriers to fast capital-inflow and outflow in railway enterprise also prevent these from acting as regulators to the level of railway tariffs and rates.

The large volumes of investment represented in existing physical railway plant thus account for peculiarities in the character of fixed costs and variable costs (or 'prime' costs) involved in railway operation.[56] For a given level of railway plant and property, railway fixed costs, along with the interest outgo on bonds with which these are financed and rental values of railway lands tend to remain constant regardless of the volume of traffic. Many of the railway costs classified as variable costs or operational expenses too, such as expenditure on maintenance of lines and structures are independent of the volume of traffic carried, even though expenditure of this type is usually undertaken or deferred depending on the volume of business and profits and fund availability. Maintenance expenditure on railway equipment such as traction and rolling stock, although more variable, also does not vary in exact proportion to the volume of traffic carried. The same is true in excess-capacity situations for personnel wage bills, fuel expenditure and miscellaneous other costs included in railway budgets as *transportation expenses*, since the variable costs involved in transporting partially-loaded trains is nearly the same for trains with full load.

Transportation enterprise is generically subject to certain economic laws apart from the regulation imposed by legislation. The principle of increasing returns asserts, for instance, that expenditures do not increase to the same degree as revenues when the volume of business increases. Once a transportation system has been established on the foundations of fixed capital investment, an expansion in the volume of shipments causes operating expenses to rise, but has little effect on constant expenditures and results in decreased expense per unit. This holds true as long as unused plant capacity is available, that is until, for instance, double tracking on a railway route is necessary, or increased equipment and terminal facilities are required along a highway. In each mode of transportation the relationship of constant to variable expenses depends on its physical equipment and the nature of its operation.

Although estimates of the proportion of fixed costs in railway enterprise have now been scaled downwards from the two-third proportion found in early literature, the fact nevertheless remains that a large mass of the short-run costs of railway enterprise is invariant in respect of the volume of traffic. In circumstances such as these, with increasing volume of traffic, unit railway traffic costs tend to decline uptil the point of optimal utilisation of railway plant, and rise thereafter because of crowding of existing railway traffic handling facilities. Thus under situations of substantial excess capacity, railway operations are subject to increasing returns from economies of scale for two inherent reasons. The first of these arising from the efficiencies of more complete plant utilisation., spreads fixed railway costs over a larger number of traffic units. The second arising from the efficiencies of holding larger and technologically-upgraded plant, significantly lowers the unit fixed costs of operations.

The economics of increasing returns under which railways tend to operate inevitably predicate large swings in the net revenues of railway enterprise because of increases and decreases in traffic, the exact extent of which would depend on the operating cost elasticities involved. However, the common practice of deferring maintenance expenditure when railway funding is short renders the magnitude of these swings opaque, because net revenue surpluses may seem to occur in bad traffic years and net deficits in good traffic years relative to whether maintenance has been deferred or undertaken.

1.6.2 Railway Costing Principles

One of the principal economic laws that guides the process of railway tariff formation is the Law of Joint

Costs, which is always of particular relevance when output processes involve the production of two or more products from a single operation. Particularly in a railway situation, haulage of various categories of passenger and freight traffic over an identical set of tracks precludes the imputation of costs specific to any given traffic category on a scientific basis. In this distinctive economic milieu, problems of railway rate-making rank assume no less importance than the investment decisions that augment transportation infrastructure. Even though analysis of the existing structure of railway costs assists the evaluation of existing rate structure and the reformulation of rate policy, the railway rates defined cannot be set exclusively on considerations of costs but necessitate accounting of the social benefits derived from facilitation of transport.

Thus from the early days of their advent, it has been usual for railways to assume the role of 'price takers' and adopt rates determined by the railway regulatory authority.[57] While their success as an enterprise then depends on the volume of traffic and freight forthcoming at the rates offered, the railways do not wield much control over their business, since in their role as 'common carriers' they are obligated to carry the traffic offered.[58] It is hence often a matter of pious hope that the regulated tariff will exactly cover average operational costs without overstating or understating certain cost elements in the midst of the rigours of accounting procedure. With the emergence of competition and the gradual withdrawal of the state and its social concerns, many railways have sought the appropriation of rights to negotiate their own rates and choose their own traffic. However this also necessitates then that the railways critically evaluate their existing cost structure before setting 'meaningful' tariff charges.

The Micawberish concerns of railway administrators in the past to meet total expenses entirely out of revenues generated had often led to oversight of the basic structural characteristics in a railway system. When traffic costing commenced in Britain in 1949[59] the railways thus had to realise afresh that they could no longer remain a monopoly where cost accounting did not matter, and that they were specifically a multiproduct firm which had operations much larger and more complicated than the largest of private enterprises. Better accounting for railway costs therefore demands proper identification of the structure and sources of railway costs.

Two variants of costing procedures that are commonly encountered are *particular costing* and *generalised costing*. Particular costing is more appropriate to the charging of traffic regularly available in full trainloads or else in several wagonloads, where separate charges can be computed for marshalling and shunting services and terminal haulage and so on. Generalised costing may be invoked in the costing of 'smalls' traffic where particular costing usually leads to low realisation of charges or cumbersome costing procedure. The principle followed in this case assumes lowering in the tonnage costs of haulage as the ratio of consignment size to wagon capacity increases, but proportionately higher haulage costs for lighter consignments.[60] The generalised costing procedure is subject to a considerable amount of averaging, but proves more worthwhile when the extraction of cost information specific to separate items in the railway service imposes an additional computational burden on the railways.

Allocation of *joint* or *indirect costs* has always posed a nearly insurmountable problem for the railways. The characteristic of jointness in true joint costs can be attributed to technology and persists through all points in time. A common example drawn from the railways relates to the provision of technological inputs in the form of the terminals and marshalling yards that serve both freight and passenger trains simultaneously. In such cases, the charging of identical services at different costs depending on the user cannot be justified, especially since input depreciation is more because of age and obsolescence than from use. Segregability of specific service costs is however more appropriate to the *common costs* (more usually found than joint costs in railway references) which arise in investments that prove supportive to multiple freight shipments or passenger services.[61]

It had already been stated that while economics accords the *marginal cost* pricing principle a major role in the efficient allocation of resources, the contexts of infrastructure pricing raise the problem of whether to equate prices to *long-run* or *short-run* marginal costs. In the particular context of railway pricing, this involves evaluation of whether marginal cost pricing is at all efficient for a public utility services that involves a considerable margin of 'sunk' costs. A reformulation of marginal railway costs as *social marginal costs* that include *congestion costs* then becomes imperative, since expansion of railway user volumes reorder user-costs of fixed facilities that have gradually become overcrowded.

1.6.3 Rate-Setting Principles & the Differential Tariff

The demand for commodities and for transportation of commodities, although distinct, are related at absolute levels by market prices and market demands for those commodities. Even so, transportation demands can be considerably low depending on the freight tariff rates borne affecting the location of production facilities and distribution channels for commodities, without however affecting their levels of absolute demand. Immobility of commodities occasioned by high freight tariffs on them reduce market prices and increase market demands for them in contiguous production areas, while acting conversely in more distant distribution areas. Thus a threshold tariff exists beyond which commodity movements cease.

Transportation rates in all countries are based on the fundamental economic laws mentioned earlier in this section. When a freight rate is high, it is normally a small proportion of the selling cost. Under the law of increasing returns, revenues to the carrier increase disproportionately to costs, especially when constant costs are a large part of the total costs. On the other hand, a commodity with a low margin or profit per unit may be charged a low freight rate to facilitate a wider market and bring the carrier a greater volume of traffic. The increased volume compensates for the lower rates only when the return pays the variable expenses and contributes something toward the constant costs. Some dissatisfaction with this method of rate making by the railroads has been expressed by domestic water carriers in the U.S. Inland shippers believe that their industry has suffered because of this method, and they favor their system of fully distributed costs as being more just.

In a practice known in the literature as *differential charging* or *discrimination,*[62] tariffs charged by railways on commodity freight tend to vary in reflection of relative traffic demand., because of this huge underlying mass of fixed costs. Thus low tariffs just covering railway prime costs and not covering full costs of carriage might be applied to commodities for which traffic volumes are highly cost-elastic, since realisation of traffic spreads out fixed costs and minor upswings occurring in their traffic would lead to large revenue gains. Levy of higher tariffs on the other hand might shut off this traffic altogether. Differential freight tariff-setting thus refers to all situations in which differences in tariffs charged between commodities cannot be explained by differences in costs of carriage. The role played by the magnitude of fixed costs in determining the practice is made abundantly clear from the theoretical circumstance of a transport enterprise where fixed costs are nil, and all costs incurred are variable, in which case the levying of tariffs not covering full costs would always lead to loss unless compensated by higher-than-full-cost tariffs on certain categories of traffic.

Differential freight tariff-setting usually takes the form of discrimination between commodities and is made operable through the classification of freight into a limited number of commodity-groups for each of which common tariffs are applicable, and by the granting of special commodity-rates on articles not addressed by regular classified tariffs. Differences in rates that reflect differences in freightage and handling costs arising from bulk-character of certain commodities whose freightage costs per unit-weight are greater that costs per unit weight-density, or from requirements of specialised equipment and handling, expedited delivery or differences in handling risks and liabilities do not qualify as discrimination. Discrimination on the other hand refers to tariff-setting that is reflective of differences in the "ability of the traffic to stand high transportation charges.[63] and draws its rationale from the existence of fixed railway overheads and excess freightage capacity.

Railway freight tariff-setting in India incorporates these typical pricing principles within a structure of tapering telescopic rates defined on the commodity-classes for which rates are being set. Separate rates are set under the system for the three categories of smalls, wagon-load and train-load traffic with minimum chargeable weights and distances being set for each category. Differential rate-setting is incorporated within the process of commodity-classification in which the basic or Class 100 rate-scale is set, and from which all other class rates can be computed as multiples. Theoretically, the Class 100 rate is set so as to recover the fully distributed average costs of freightage, with the tapering rate structure affording concessional haulage to freight that is booked for longer distances. With rate classes on IR ranging from Class 80A to Class 300X, freight classes rated below Class 100 offer subsidised haulage while those above Class 100 recover more than the fully distributed average costs of freightage including terminal charges. Additional rate exemptions announced from time to time in the Railway Budget for certain commodities are accommodated by introducing new letter-suffixed rates derived from the appropriate class rate. Over time, the exemptions granted have seen the number of class rates go up to 60 from a basic rate structure involving 34 commodity classes. The telescopic element in each rate schedule also varies between commodity classes and has the effect of attracting traffic

of desired specifications to the railways by offering cost concessions.[64] The practice of setting railway tariffs for certain commodities at below the prime costs of haulage and making up the losses incurred by realisation of tariffs on certain other commodities at rates higher than full freightage costs amounts to the extreme form of discrimination known as *cross-subsidisation.*

A form of differential tariff-setting that qualifies as discrimination in the true sense of adjusting tariffs to traffic demand involves setting of commodity rates that bear no relation to traffic leads through devices such as setting varying commodity tariffs for equal freightage distances, setting equal tariffs for unequal distances, framing progressive tariff rates where the rates of progression per unit of distance are less than the rates at which carriage costs increase, and the extreme case of discriminating between long-lead traffic in preference to short-lead traffic travelling on identical line and direction through levy of higher tariff rates on the latter. Freight tariff equalisation which was adopted in India for almost three decades uptil the late 1990s to compensate for the locational disadvantages of power plants and industries to which coal and mineral traffic had to be carried over very long leads provides an instance of this tariff-setting principle. Tariff equalisation however ultimately undermines the economies of scale and the economies of location and has often been criticised on these grounds. Another less frequent mode of tariff discrimination concerns the setting of special rates personal to the shipper to secure particular classes of traffic. However the economic justifications for following this practice are less convincing, even though it will be seen that it is in vogue in France and has also been recommended in India as a means for railways to hold on to major freight clients, in the face of freight competition from the roadways.

Discriminatory tariff-setting is fostered under the monopolistic structure under which railway enterprise tends to operate. The principal reason for absence of competition in railway enterprise is the fact that struggles for traffic shares between rivals would drive tariff rates below service cost levels to the point that returns on capital invested would be too low to justify renewals of railway plant. This again reflects the presence of the high mass of fixed overhead costs involved, which justify continued operation of the freight service so long as it makes some contribution towards the cost of overheads, particularly when excess capacity prevails in the short-term. This aspect of railway tariff-setting has been a point of classic focus in the literature (*cf.* the Taussig-Pigou controversy on whether the determinants of discriminatory tariffs for transportation are *joint* costs or *constant* costs, to which Pigou's reply is paraphrased in Pigou [1920],[65] which postulates that a discriminating railway enterprise in a competitive setting will face immediate competition for the traffic for which it has set the highest tariff, which in a discriminatory situation will be that traffic which makes more than a *pro rata* contribution to defraying overhead costs. The resultant struggle will force down rates on this traffic, while at the same time pushing down tariffs on other lower-rated traffic if excess capacity exists to levels below prime-costs.

Except for differences in route-coverage and in the qualitative nature of service, such factors also influence competition between railways and roadways on the evidence of countries like the US and France, leading to measures by railways to protect traffic shares by tariff incentives such as special rates for freight delivered in full trainloads or wagonloads, or fixed container charges without tariff discrimination on the basis of their commodity contents. At another level, roadways are unable to enter tariff competition for low-grade traffic which moves at rates well below those defined by the cost economics of road transportation define. In India, this has led over time to the surrender of low-bulk high-rated short-haul commodity traffic by railways to the roadways, while low-rated bulk traffic with long leads has remained with the railways because of the cost advantages enjoyed by that mode of freight. Although rising competition from roadways remains the bane of the railways in almost all countries around the world, possible long-term solutions dictated by similar duopoly situations that arise in industrial economics point towards eventual tariff coordination between railways and roadways in order to permit both to survive. Of course, the more disaggregated the road freight operation, the likelihood of subcontracting arrangements being used to break down rate agreements with the railways multiplies.

Although tariff discrimination in railway freight is also sustained because of apparent independence of individual transportation demands for most commodities from those of others, a relation may well exist between certain commodity-pairs whereby transportation demand for one commodity is increased by the decline in transportation demands for others on whom higher tariffs are being charged. Similarly, low rates on raw materials relative to high rates being charged for finished products can increase mobility of the former, eventually destroying the economics of locational advantage.

In sum, the three special characteristics that have distinguished the setting of railway freight rates are prevalence of discrimination and lower-than-cost tariffs on certain commodities because of the high proportion of operational overheads, prevalence of monopolistic structure in railway operations and the existence of transportation-demand independence between many pairs of commodities.

1.7 Elements & Scope of the Present Investigation

The conceptual framework that relates infrastructure to development theory is founded on the role of infrastructural capital in forming and servicing the expanding economic base. Railway transportation is accorded considerable prominence in the economics of public infrastructure because of its ability to interlink wide economic spaces and facilitate the economical movement of goods and resources between them. Both in India and elsewhere, the historic developmental role of railways has thus rested on the spatial economics of transportation through which economies of scale and scope created through vast investments sunk into the railway infrastructure are translated into economies of production in the form of reduced transportation costs. By carrying balancing movements of freight between surplus and deficit areas, the railways are also able to mitigate the locational diseconomies of production, aiding the outward dispersal of urban settlements which catalyse industrial activity.

Although the present investigation shall mainly focus on the economics of railway freight operations in India, the infrastructural role of Indian railways shall be explored in the background from the twin perspectives of economic history and national economic planning, where the Indian railways have played a remarkable role. In development contexts, efficient freight transportation services that raise the productivity of producing sectors also expand the monetised economy because the resulting increases in income widen domestic markets by augmenting effective demand. In developing countries like India, where capital constraints are an impediment to economic growth, a burgeoning service sector will set the pace for parallel industrial growth. This is particularly true for infrastructural development and transportation, which lead to both creation and efficient deployment of investible and other resources, thus catalysing the process of capital formation.

1.7.1 The Railways in Indian Transportation Infrastructure

Among various transportation modes, surface transport has been the most accessible form of transportation for centuries with waterways being a popular mode of surface transport since the early years of transportation history. But waterways had their limitations in the sense that they could be adopted as a principal mode of transportation only in those countries which are well connected by rivers and canals. With the advent of the railways, waterways gradually lost their old importance, and over the period since, railway transportation has assumed a prime position in the hierarchy of transportation modes wherever railway infrastructure exists. However, the motorised roadways which made their advent long after the railways, progressed considerably over the 20th century, and have since begun to offer vigorous competition to the railways. Nevertheless, the railways with their advantage of being efficient in the transportation of goods and men at higher speeds and in greater quantity than the other modes of surface transport, and in the present scenario of increasing fuel prices, they are still a favoured mode of transport. The present study is an attempt to make an assessment of this popular mode of transport with specific reference to Indian Railways [IR] freight operations.

In commodity-terms, the output of the railways transportation sector is a service for which demand can be both direct and indirect. Direct consumption of railway transport service takes place for instance, when a family travels for recreational purposes, or when consignments of consumable commodities are transported to their final consumers. Indirect consumption of railway services arises as a result of direct consumption, e.g. when factors of production have to be moved to manufacturing points to facilitate production *i.e.* in consequence of derived demand. Although both direct and indirect flows of men and materials that take place along transportation networks are important to economic contexts, it is the latter that account for bulk utilisation of transport capacity in commodity shipments. Passenger flow along the transport network is however also capable of increasing the mobility of human resources and leads, over time, to greater degrees of urbanisation as a result of the association between industrialisation and the growth of towns. But transportation also involves freight flows, the volumes in which are good indicators as well as determinants of development.

1.7.2 Freight Operations on the Indian Railways

Freight movements within an economy take place through surface transportation over land and water, and also by air. Within the surface transport category, it is the land-based means that are predominant. Although land-based transportation includes both roads and railways, the two are not perfect substitutes, because of variations in their cost-efficiency and in the freight-mix they service. In India, despite developments in other modes of transport over the last quarter-century, the Railways have been able to retain their unique position as freight hauliers. Dominance of bulk commodities like foodgrains, metal ores, coal, cement, fertilisers and steel, etc., in the freight-mix largely explain their pre-eminence in internal goods transport.

A study of the history of IR shows that freight operations has been the prime consideration towards the development of railways in preindependent India. Indian railway history originates not from the construction of the first railway lines, but from the necessity to prevent famines that had plagued the country recurrently and were devastating enough for the colonial power to devise means to distribute foodgrains efficiently to the famine prone areas. The need to transport raw materials from the hinterland to ports and the reverse traffic movement of manufactured goods from ports to the domestic consuming centers also encouraged the British to invest in this rather expensive venture. Along with these economic motivations political considerations towards movement of military provisions and troops also contributed to the development of railways in India.

Freight movements in the post-Independence era had to traverse a long and winding path. Impending needs of the country to industrialise necessitated the movement of raw materials to production centres and as the economy grew, manufactured items found its way to the various consumption nodes. Retaining its character as a public good the IR also has to give preference to essential items like foodgrains and mineral ores for the core public sector industries. Thus although freight traffic has borne the characteristic of transporting broad based categories, there has been a gradual tendency to concentrate on bulk traffic over longer journey leads. As a consequence commodities that were previously carried by IR in the initial years of post-Independence, many categories have found their way to the fast growing roadways.

The National Transport Policy Committee in its report in 1980 had projected a 72:28 share of freight by the railways and the roadways. However recent trends have identified the share to be 40:60 for the respective sectors. While the projected freight tonnages for the 1990s have been achieved as will be seen in the study, the target has been attained with only 40 percent share in the intermodal freight movement as against the projected 72 percent. In order to increase this share from the present 40 percent it is necessary to bring back traffic from the roadways, a problem that IR has not attempted to solve until recently. For the general improvement of the degrading financial health of IR this alternative has become necessary along with the imperative of minimising the costs incurred to maximise its resources. The study while dealing with the freight movements on IR attempts to shed some light on these problems and possible solutions to overcome them.

1.7.3 Railway Freight-Capacity Planning & Transport Policy

Although passenger transport has an obvious productive dimension when it refers to movements of migrant and commuting labour, by far the most important productive flow along transportation networks is the freight flow in raw, semi-finished and finished goods.

Freight transportation has two pronounced aspects: it is a 'production facilitator' insofar as it is a bulk carrier of primary and intermediate inputs over economic space; it is also a 'market facilitator' in determining the size of demand brought to bear on basic, intermediate and final outputs produced within the economy. Inadequacy in freight transportation facilities is inhibitory to the actual process of development. Such inadequacies affect the economy in totality, but have a differential impact in spatial terms because of the regional imbalances they generate. Whether freightage capacity and the freight flows within it have commensurately met trends in the demands for freight services is worth investigating. This problem provides the formulation for the present study. The problem is studied with a focus on freight-haulage capacity and movement.

Several studies devoted to the estimation of transportation demand are found in the literature. Sriraman and Rao [1985], in an empirical study of railway freight demand in India identifies factors required for freight movements such as the length of haul and the quantum of shipment and factors affecting the relationship

between them. Effect of the changes in the relationship will be manifested in the change in the demand for wagons and this necessitates the operational decisions leadiing to an observed supply of freight services. Bottlenecks arise when the increased demand is not met by an increased supply of wagons. Thus alternative policies are required to augment the carrying capacity of the railway network and or reducing the demand for freight services that is unlikely to occur in a growing economy.

For an evaluation of freight services of the Indian Railways, different indices like commodity-tonnages originating and net tonne-kilometre haulages have been widely used in studies of the growth of freight traffic operations. Viewed in the long-term infrastructural context, time-series in these, have consistently pointed at emerging capacity-constraints. However, it is not enough merely to identify infrastructural lacunae in terms of transportation bottlenecks existing at a point of time. A longer time-frame of study is necessary, since infrastructural development in any case is not a short-period variable. The period under observation commences with the inception of railways in India, *i.e.* 1853 and analysis has been undertaken till 1994-95, while the recent developments on IR are also presented to assist in the future policy formulations.

Limitations in empirical study of long-range infrastructural development have been the need for extended time-horizons for analysis and the complexities of time-series modelling of infrastructural impact, in the face of which short-run planning mechanisms are particularly inadequate. Most infrastructural analyses in economic-history frameworks lead to heuristic conclusions that are poor in respect of operational rigour and therefore country-specific and often not incontrovertible.

While the shift in the planners' perceptions regarding the quantum of investment necessary and its mode of application within the freight infrastructure has been conscious and can be grouped into three distinct time-phases of the Indian planning experience - namely expansion, recession and post-recession - the economic consequences of such shifts have hitherto been glossed over in the absence of analyses that embrace the long time-horizon. Keeping in mind that the two schools of thought on social overhead capital alluded to the time-horizon conflict, no reconciliation of the shift from the long-run perception of transport development as a precursor to economic development to short-run exercises in capacity-shortage management can be achieved without distortion in the attendant development processes. This will be the problem highlighted in each of the empirical exercises that the study undertakes to reveal that the development of freight facilities in the Indian economy is characterised by a distinct hiatus, in consequence of shifting plan priorities. By all indications, such shifts in plan foci are deleterious and inhibitory to the pace of economic development that has been realised.

The basis that transport provides to development emanates from its infrastructural role. Transport augments the size of demand which will pull all other economic resources into economic activity through general market interlinkage; this in turn brings significant pressures to bear on transport networks. Thus transport infrastructure generates development, which generates the need for additional transport facilities. Serious bottlenecks may however be present in this process causing shortages of capacity and affecting the flow of resources and outputs important to economic development.

A priori, one would anticipate that as Indian planners had emphasized infrastructural augmentation sincerely in the early Plans, development - in terms of any economic indicator - would show up eventually. However, the feedback signals then appearing would require progressive investments in transport infrastructure. For progress to continue unhindered, effective coordination with the growing economy and appropriate phasing out of lumpy investments on transport capacity expansion would have had to be implemented. But public investment in the transport sector as a whole instead showed a declining proportionate trend, after the initial thrust witnessed over the first three Plans. Evidence in the chapters suggests also that actual expenditures too fell short of plan-outlays in real terms over almost every Plan, resulting in significant spillover of capital projects between Plans.

It will be worthwhile to mention that the Plan allocations that decide the fate of the capacity augmentation on IR is also responsible for the growth of the other production sectors in the economy. The targets set for the various industries and agriculture are dependent on the Planning Commisssion's decision of providing investment support through the Plan allocations. Thus shortfalls in one sector, *e.g.* the railways, can affect the other sectors of the economy as supply of freight services will be restricted due to lower capacity augmentation. On the other hand, declining Plan allocations for the agriculture, mining and industrial sectors would imply a shortfall in demand for freight services due to fall in production levels and would adversely

affect the transport sector through revenue shortfalls. The Railways were the worst affected by this resource crunch. A slowdown of the economy thus affected the IR because of freight target shortfalls and in the face of rising railway costs that are necessary for the maintenance of the capital stock and to sustain operations, led to increased dependence on the budgetary support from the Government. Consequently the subsequent Plans failed to sustain their earlier impact on the economy, and slackened the pace of infrastructural development.

Infrastructural thrust evident in the early FYP in India sought to provide the economy with a strong industrial base by facilitating the movement of raw materials and finished products without any constraints. The commodity freight-mix comprised of both light-weighted and bulk, short and long lead, low and high rated ensuring an even distribution of freight-flows from surplus to deficit regions. However the nonfulfillment of the production targets in the 3FYP and the consequent excess capacity existing within the railway transport network created an avenue for the planners' to bring about a shift in policy with a diminished infrastructural thrust and resort to demand management. This shift in the attitude of the planners failed to generate momentum in the railway transportation sector in the years that followed the lowering of investment in plan allocations and consequently had a deleterious effect on the development of the economy. Extraneous factors like famine, wars and oil shocks aggravated the crisis and the transport policies adapted themselves to combat such unforeseen circumstances. To overcome the uncertainties that affected the generation of traffic, freight flows on the Indian Railways (IR) were shifted deliberately towards bulk traffic mainly originating from other public sector units (PSU) ensuring some traffic as a better alternative to none or uncertain traffic. While essentials like foodgrains were mandatorily accomodated by the IR in consonance with the public utility nature of the railways, many short lead, low bulk and high-rated traffic were dispensed with to accomodate the PSU bulk traffic. For efficiency reasons freight policy stressed on long-haul, bulk traffic while the 'smalls' were forced to find alternative mode for transportation. Whether the policy shift from capacity management to demand management or from a mixed-bag of commodities to a specific commodity set has been for the betterment of the railway organisation and the economy as such is one of the concerns of the present study.

As a result of this deliberate policy to concentrate on long lead bulk freight, the IR freight-mix became unfavourably poised towards freight smalls and thus diverted towards other modes of transport like the roadways. The roadways sector partly under rate-competition and partly as a deliberate policy of altering the freight-mix as observed in the thesis could expand at the cost of the railways. Besides the fact that traffic in this category often involved highly-rated commodities, the switchover in transport mode also induced a loss of efficiency since several of these commodities still involved long leads. Loss of high-rated traffic resulted in revenue shortfalls and railway finances came to be heavily dependent on Budgetary provision from the Government. This along with the resource crunch affected the IR investment decisions adversely and led to formulation of policies that had a strong impact on infrastructure. The thesis examines the impact concentrating on the production sectors affected by the constraints in the supply of freight services and continued specialisation of the freight composition on the IR.

Another important aspect present in the infrastructural issues and the transport policy of the IR is the role of transport infrastructure in achieving an organised spatial distribution with respect to both production and consumption flows in order to achieve balanced regional development. Any bottleneck arising owing to capacity constraints or specialisation of freight traffic is bound to create a distortion in the regional growth pattern within a developing economy, where inadequacy in freight flows can jeopardise both the production and consumption activities.

Because of the importance of the issues that need to be investigated for development prospects in the future, for infrastructural expansion, and for the evaluation of Government policy in this respect, it is hoped that the study will break new ground in the study of infrastructural linkages and will provide insights useful to the reformulation of national transport policies.

1.8 Structure & Methodology of the Study

Considering the critical significance of the Indian Railways in meeting the country's infrastructural needs for passeneger and freight transportation over long distances, a scientific review of the entire transport planning process in prospect and retrospect is necessary. Identification of the lacunae present would aid the

formulation of new coordinated policies for effective development of infrastructure.

Fairly comprehensive published data covering diverse aspects of freight operations are available from Railway and other Government sources, from the Planning Commission and from National Transport Policy documents. The study has been mainly carried out on timeseries data assembled from this secondary database and on material drawn from documentary sources in the literature.

The present study is essentially an exploration of transportation economics and railway freight operations in India over an extended period of time. Chapter 1 (the present chapter) lays the foundation for the study by formulating an approach that can address the inherent infrastructural issues in transportation and railway development. An overview for this is developed through a cross-country analysis that illustrates the importance of surface transportation, and specifically its two principal modes namely the railways and roadways. It is seen that the preference for the dominant mode of transportation is determined by the geographical space occupied by a country and its population density, revealing that the choice of transportation infrastructure and the scale of capital investment committed to it is guided by spatial principles. In India, the geographical space is large enough to demand dual modes of transportation. Consequently, the capital requirements for developing infrastructure are accordingly large. An exploration of the theoretical literature and of development history reveals that although the provision of infrastructure has a social purpose, the method of infrastructural finance has varied between private and public investment. In the case of the railways an important method of financing infrastructure has been the partnership of private capital and the state. By granting leases, franchises and concessions, the state is placed in a powerful position to mobilise and direct private capital investment. Yet, because of its public utility character, the sunk costs of railway infrastructure cannot be recovered. Hence, the need for periodic refinancing arise.

In order to develop a comparative understanding of the development of railways in India as a background to study of the current problems faced by the Indian Railways, an exploration is made in Chapter 2 of the current productivity and efficiency of the railways in Britain, France, China, Japan and India. The comparative sizes of these networks are seen to have an important bearing on the relative passenger and freight intensities of the operations. The operational history of these railway systems is briefly explored to obtain a common understanding of the infrastructural issues that presently confront the Indian Railways. Although this choice of countries may seem arbitrary, some thought has gone into their selection. The railways in Britain are the oldest railway system in the world. Unlike the Indian railways the British system is passenger oriented because of the limited size of the country. Since it was British capital which brought the railways to India, similar patterns of financing guided the building of railway infrastructure in the two countries. The French railway system developed under the strong leadership of the state. Hence, social imperatives have always guided its operations. After nationalisation, the enhanced role of the state has made SNCF one of the most advanced state-owned railway systems in the world. Many of the technological initiatives made by Indian Railways under the Plans have sourced their knowhow and consultancy from SNCF.

The railways of Japan and China constitute two non-identical Asian railway systems. Like the British Railways, the railways in Japan are technologically advanced in passenger operations but carry limited freight because of their island character. Both the British and Japanese railways have undergone recent privatisation. The Chinese railway system is a large state-owned enterprise which has seen among the highest rates of network expansion in recent times. While the importance of the railways to China and India are similar because of their relative size and because of the undeveloped character of many parts of the two countries, their modes of railway operation and their technological choices differ considerably. The American railway system receives a special mention both as one of the largest railway systems in the world and also as one that has remained under disaggregated private ownership from the start. It also has among the most technologically advanced freight operations with high levels of running efficiency. Thus although the Indian Railways resemble many of these systems in some respects, they differ in many others. However, the historical study does not seek to reproduce the well-documented histories of these individual railway systems. Rather, it attempts to identify the common ground between the questions of public and private infrastructural funding, state and company ownership and the response to transport competition that have characterised the development of railways throughout most of the world. The study of the alternative solutions found by these railway systems provides a relevant perspective on many of the issues that currently confront the Indian Railways.

Chapter 3 thus commences the analysis of the trends of various indicators of railway operational performance

Chapter 3 thus commences the analysis of the trends of various indicators of railway operational performance which have bearing on the movement of freight by railways in India. A long-period analysis is first made of the capital situation of the railways with reference to the returns to capital-at-charge, and the changes in the patterns of capital funding are identified. These are related to trends in asset acquisition and the operational trends of the Indian railways. While both originating tonnages and journey leads are seen to increase continuously over the post-Independence period, a tendency to shift the focus of railway operations to certain bulk commodities is also noticed in the face of declining support from budget allocations. Since the augmentation of freighting capacity has not kept pace with the increasing demands from traffic, the railways have had to devise their own means of carrying additional bulk-freight through the substitution of other commodities.

Chapter 4 makes a study of the railway wagonfleet in India. The patterns of wagon acquisition by the railways and the technical changes in the wagonfleet composition that have been made to accommodate the needs of specialised streams of freight are identified from the data. It is seen that irregular acquisition of wagons has imposed freighting constraints in the railways and has also driven the Indian wagon-fabricating industry into industrial sickness. The character and evolution of the industry is also studied contextually, and the impact of wagon shortages on the efficiency of railway freight operations is identified.

Chapter 5 studies the patterns of fund allocation to the Indian Railways in successive Five Year Plans and also the internal allocations of these funds by the Railways for different infrastructural uses. The policy reasons behind the slowdown in wagon acquisition and the changes in wagonfleet composition are examined through a comparative assessment of important Railway reports and freight projection studies. It is seen that wagonfleet inadequacies have led to the emergence of transportation bottlenecks in the country. In the technical part of the chapter the distributed-lag model is applied to identify the determinants of freighting capacity of the railways. The econometric analysis studies the changes in wagon acquisition and the railway wagonfleet in relation to changes in the haulage of freight and in other operational and economic variables. The disequilibrium between the various freight projections and the rates of wagon acquisition are identified and analysed in terms of its impact on the economy. It is also seen that the shifting transport priorities in the course of the Plans lie at the core of the problem.

The technical study in Chapter 6 seeks to devise an econometric basis for the exploration of the economic and distributive consequences of transportation bottlenecks in India, on secondary data reflecting the longterm trends of railway freight operations and the cross-dependent patterns of commodity movement. An extended timeframe of 35 years, commencing at the end of the First Five-Year Plan and extending uptil the end of the Eighth Five-Year Plan, allows the technical study to bring out the essential freight demand and freight supply characteristics in the operational management of the Indian Railways. Wagon loadings of 15 major commodities on the two principal gauges *i.e. broad gauge and metre gauge,* are used as surrogates for railway freight productivity and a comparison across gauges is undertaken to identify the impact of freight policy on the sectors associated with the production of these commodities. While production of agricultural commodities as well as light industries are located in dispersed corners of the country, the freight in certain commodities had a metre gauge (MG) dominance earlier since it fed the feeder lines of the railways which also served as connecting links to the backward areas of the country. The infrastructural impact of the intermodal transportation policy of phasing out MG railway operations and concentrating on the broad gauge (BG) that was part of the National Transport Policy report in 1980, on commodities carried on the MG network is thus analysed in order to identify the displacement of commodities from the MG to the roadways. The chapter undertakes a multiple regression analysis on the wagon-loading data for commodities that belong to different sectoral groupings, and seeks to define the basic sectoral flow and composition of railway freight traffic in India. To understand the patterns of serial correlation which are present in the data, the longterm cross-variability of freight flows is studied through a phase-analysis of the systematic component of the residual errors which identifies the patterns of freight shortfalls and subsequent substitution of commodities that has led to increasing railway freight cyclicity over successive plan periods. The commodity coefficients are then estimated using the Cochrane-Orcutt adjustment procedure and interpreted in the order of their importance to railway freight operations in India.

The spatial and sectoral impact of longterm shifts in transport policy are studied in Chapter 7. The profitability positions of India's zonal railways are analysed with respect to their gauge-networks and the intensity of their freight and passenger operations. The impact of the changes in railway freight-flows that have been

observed previously are found to fall disproportionately on regions served by MG-dominated railway zones. Multicollinearity in the data is studied through use of the factor analysis approach to identify the patterns of commodity dominance and displacement in railway freight operations over the extended timeframe. The patterns captured by the factor analysis are reformulated into an econometric freight-adjustment model which identifies the most common freight demand and supply situations encountered in railway freight operations in India. It is also shown that these commodity displacements are essentially caused by the inadequacy of railway infrastructure and freighting capacity, which have led to increasing competition in the Indian freight market from the roadways.

Chapter 8, which is the final chapter, draws together the various strands of analysis from the preceding chapters. While the focus of the present study remains on the freight operations of the Indian Railways, several larger issues concerning railway planning and the financing of infrastructure are found to have entered the analysis. By its very nature, railway infrastructure or any other form of infrastructure offers limited scope for profit-taking but extends the developmental benefits of economies of scale for a very long period of time. This is seen to have occurred through the construction of railways in India. However, railway infrastructure continues to require periodic support for renewing its physical plant and technology which demand huge capital investments. Since the present study has partially explored the consequences of inadequate capital renewal on railway freight operations in India, the ultimate question asked relates to the current capital needs and future sustainability of the Indian Railways.

The basic need to identify such displaced commodities arises from the need to observe the infrastructural impact of the transport policies on the downstream industries. While a certain freight-mix has been present in the early plan periods there has been a noticeable shift towards bulk traffic especially catering to various PSUs. As a policy objective to carry essentials and PSU traffic at a lower rate, the IR have failed to generate resources to finance their investment projects. Unlike France, it has not tried to bring back the commodities that has diverted to roadways because of the NTPC policy to maintain a higher average journey lead to keep the IR as a viable organisation. In addition to this the IR has to provide a high capital-at-charge for the maintenance of the present capital stock as well as contribute to the general finances in the form of dividends. However it is also required to increase capacity to attract more traffic from the roadways. With dwindling plan allocations the avenues of an expansion through renewed investment are almost absent and the railways are gradually losing grip over their finances. Thus on the one hand to generate more resources through increased traffic it is imperative that IR expands its capital base and on the other hand to enhance its capital base the IR either has to increase its traffic to earn more revenues or the government has to fully undertake the responsibility of rejuvenating the IR.

Thus in the present study of railway development in India, the contrasts between the colonial and the post-colonial pace of infrastructural development and the auguries these may have had on the economy of the country are present within the econometric results of the study. The point of interest for the study relates to how changes in developmental perceptions and priorities over along period of time can de-emphasise the earlier goals of infrastructural history and whether these can be reestablished by the reorientation of development policy into longer timeframes, as has been accomplished by other modernising countries. Thus, the conclusions from the study extend outside the Indian situation and point towards plausible reformulation of development studies in general, to capture the role of infrastructure.

Chapter References

1 *cf.* Solomos Solomou [1990]: *Phases of Economic Growth, 1850-1973:Kondratieff Waves and Kuznets Swings,* Cambridge University Press, Cambridge

2 *cf.* Daniel Defoe: *The Anatomy of the Stock Exchange Alley*, quoted D.N.Ghosh [1997]: 'Gullibility of the Investor and Pastimes of the Marketmen', *Economic & Political Weekly*, 32(42):2631

3 D.N.Ghosh [1997] *op.cit., Economic & Political Weekly*, 32(42)2631-32

4 Gillian Harris, Andrew Pierce [1998] : "The Heroic Failure that Sank Scotland', *The Times*, London, syndicated reprint, *The Statesman* daily, Calcutta edn., 26 April 1998

5 *cf.* Reminiscences of Edward Callow, in Charles Duguid: *The Story of the Stock Exchange,* quoted D.N.Ghosh [1997] *op.cit.* p2631

6 Barry Eichengreen [1995]: 'Financing Infrastructure in Developing Countries: Lessons from the Railway Age, *The World Bank Research Observer* 10(1):75-91, p76

7 Joseph Stiglitz, Andrew Weiss [1981]: 'Credit Rationing in Markets with Imperfect Information', *American Economic Review,* 71:393-410

8 Michael C.Jenson, William H.Meckling [1976]: 'Theory of the Firm: Managerial Behaviour, Agency Costs and Ownership Structure', *Journal of Financial Economics,* 3:305-360

9 George Akerlof, Paul Romer [1993]: 'Looting: The Economic Underworld of Bankruptcy for Profit', *Brookings Papers on Economic Activity,* 2:1-73, quoted Barry Eichengreen [1995] *op.cit.* pp76-78

10 *ibid.*

11 Barry Eichengreen [1970] *op.cit.* p81

12 John B.Lansing [1966]: *Transportation and Economic Policy,* The Free Press, New York, p89

13 *ibid.* p31

14 Ragnar Nurkse [1952]: *Problems of Capital Formation in Developing Countries,* Basil Blackwell, Oxford

15 A.O.Hirschman [1958]: *The Strategy of Economic Development,* Yale University Press, New Haven, Connecticut

16 R.W.Fogel [1964]: *Railroads and American Economic Growth: Essays in Econometric History*, W.W.Norton, New York

17 A.Fishlow [1965]: *American Railroads and the Transformation of the Ante-Bellum Economy*, Cambridge, Massachusetts

18. G.R.Hawke [1970]: *Railways and Economic Growth in England and Wales, 1840-1870,* OUP, London

19. D.Biehl [1991]: 'The Role of Infrastructure in Regional Development', in R.W.Vickerman *(ed)*: *Infrastructure and Rural Development*, Pion, London

20. A.E.Andersson, U.Strömquist [1988]: 'The Emerging C-Society', in D.F.Batten & R.Thord *(eds)*: *Transportation for the Future,* Springer-Verlag, Berlin

21. U.Blum [1982]: 'Effects of Transportation Investments on Regional Growth: A Theoretical and Empirical Analysis', *Papers of the Regional Science Association,* 49:151-168

22. P.Nijkamp [1986]: 'Infrastructure and Regional Development: A Multi-Dimensional Policy Analysis', *Empirical Economics,* 11:1-21

23. P.Rietveld [1989]: 'Infrastructure and Regional Development: A Survey of Multiregional Economic Models", *Annals of Regional Science,* 23:255-274

24. T.R.Lakshmanan [1989]: 'Infrastructure and Economic Transformation', in A.E.Andersson, D.F.Batten, B.Johansson, P.Nijkamp *(eds)* [1989]: *Advances in Spatial Theory and Dynamics,* North Holland, Amsterdam

25. R.Musgrave [1959]: *The Theory of Public Finance*, McGraw Hill, New York

26. *ibid.* p13

27. *cf.*K.Train [1991]: *Optimal Regulation,* MIT Press, Cambridge, Massachusetts.

28. William J.Baumol, John C.Panzar, Robert D.Willig [1982]: *Contestable Markets and the Theory of Industry Structure,* Harcourt Brace Jovanovich, New York

29. *cf.*R.Turvey [1975]: 'A Simple Analysis of Optimal Fares on Scheduled Transport Services', *Economic Journal,* 85:1-9

30. *cf.*J.M.Buchanan [1963]: 'The Economics of Earmarking' *Journal of Political Economy,* 71:457-459

31. R.Arnott, A.de Palma, R.Lindsey [1993]: 'A Structural Model of Peak-Period Congestion: A Traffic Bottleneck with Elastic Demand', *American Economic Review,* 83:161-179

32. William J.Baumol *et al* [1982] *op.cit.*

33. Sylvester Damus [1981]: 'Two-Part Tariffs and Optimum Taxation: The Case of Railway Rates', *American Economic Review,* 71(1):65-79

34. Almarin Phillips [1980]: 'Ramsey Pricing and Sustainability with Interdependent Demands', in Bridget M.Mitchell, Paul R.Kleindorfer *(eds)* [1980]: *Regulated Industries and Public Enterprise,* Heath, Lexington, Massachusetts

35. William B.Tye [1989]: 'Infrastructure Costs and Contestability Theory', in *Transport Policy, Management and Technology towards 2001,* Selected Proceedings of the 5th World Conference on Transport Research, WCTR, Yokohama, vol.1: pp25-38

36. William J.Baumol , D.F.Bradford [1970]: 'Optimal Departures from Marginal Cost Pricing', *American Economic Review,* 60:265-283

37. *ibid.*

38. Sylvester Damus [1984]: 'Ramsey Pricing by U.S. Railroads', *Journal of Transport Economics and Policy,*

January, 18:51-61

39. Gerald L.Faulhaber [1975]: 'Cross-Subsidization: Pricing in Public Enterprises', *American Economic Review,* 5(5):966-77
40. *ibid.* p972
41. B.J.Spencer, J.A.Brander [1983]: 'Second Best Pricing of Publicly Produced Inputs', *Journal of Public Economics*, February, 20:113-119
42. Henry McFarland [1986]: 'Ramsey Pricing of Inputs with Downstream Monopoly Power and Regulation', *Journal of Transport Economics &Policy,* January, pp89-90
43. Ronald B.Braeutigam [1979]: 'Optimal Pricing with Intermodal Competition', *American Economic Review,* 69(1):38-49
44. Sylvester Damus [1981]: 'An Agenda for Research in Transport Pricing in an Era of Deregulation', *ICC Practitioners Journal,* 49(1):23-32
45. World Bank [1994]: *World Development Report 1994: Infrastructure for Development*, OUP, New York, p13
46. K.Mera [1973]: 'Regional Production Functions and Social Overhead Capital: An Analysis of the Japanese Case', *Regional and Urban Economics,* 2(May):157-85
47. Y.K.Alagh [1995]: *Indian Development Planning & Policy: A Re-Evaluation*, Vikas, New Delhi
48. R.J.Barro, Xavier Sala-i-Martin [1992]: 'Public Finance in Models of Economic Growth', *Review of Economic Studies*, 59:645-661
49. P.D.Henderson [1965]: 'Notes on Public Investment Criteria in the United Kingdom', *Bulletin of the Oxford University Institute of Economics & Statistics,* 27:55-89
50. P.C.Stubbs, W.J.Tyson, M.Q.Dalvi [1980]: *Transport Economics*, Allen & Unwin, London
51. R.O.Christensen [1981]: 'The State and Indian Railway Performance: 1870-1920', part 1, *Journal of Transport History*, 2(September):1-14
52. P.J.Cain [1980]: 'Private Enterprise or Public Utility ? Output, Pricing and Investment on English and Welsh Railways', *Journal of Transport History* ,1(September):9-28
53. S.K.Srivastava [1971]: *Economics of Transport*, S.Chand, New Delhi
54. K.M.Pulley [1969]: *Rail-Road Pricing and the Auto-Parts Incentive Rates*, Ph.D. thesis, University of Pennsylvania, Philadelphia
55. S.K.Srivastava [1971] *op.cit.*
56. *cf. e.g.* J.Dupuit [1844]: 'De la mesure de l'utilité des travaux publics', *Annales des ponts et chausées*, 2d Ser.(8): p332 & J.Dupuit [1849]: 'De l'influence des péages sur l'utilité des voies de communication', *Annales des ponts et chausées*, 2d Ser.[17]:170
57. Stewart Joy [1989]: 'Railway Costs and Planning', *Jounal of Transport Economics and Policy*, January
58. *cf.*A.M.Milne and Laing [1956]: *The Obligation to Carry,* Institute of Transport, London
59. C.D.Foster [1963]: *The Transport Problem,* Blackie, London, p74
60. *ibid.*p77
61. John B.Lansing [1966]: *Tranportation and Economic Policy,* The Free Press & Collier-Macmillan, New York & London, p50
62. D.Philip Locklin [1969]: *Economics of Transportation,* Richard D.Irwin, Homewood-Illinois, 6e, p134
63. *ibid.* p135
64. RFFC *op.cit.* vol.1, pp543-555
65. A.C.Pigou [1920]: *Economics of Welfare,* Macmillan London, p257n