Paying for the Fixed Costs of Roads

David Levinson

Address for correspondence: Professor David Levinson, Department of Civil Engineering, University of Minnesota, 500 Pillsbury Drive SE, Minneapolis, MN 55455, USA; levin031@umn.edu.

Abstract
This paper explores alternative financing mechanisms to pay for the fixed costs of roads, particularly in cases without rising marginal costs. Mechanisms considered include tolls, gasoline taxes, and developer payments. The problems with each are discussed. An example looking at problems of temporal and spatial free-riding is presented.
1.0 Introduction

Public finance economists often suggest pricing of public facilities in the form of a three-part tariff (Blewett and Nelson, 1988). The first part reflects a charge for the capital facility (for example major highways). The second part is the charge for bringing the facility (for example local streets) to the user, and is usually covered by the private sector in constructing a new subdivision. The third part is a charge based on actual use, and is analogous to congestion pricing.

In recent decades, the third part of the tariff, user charges, and especially dynamic charges, has seen a great deal of attention in the academic community (see de Palma and Lindsey, 1997; Viton, 1995; Walton and Euritt, 1990), but more importantly in practice, particularly with the development of High Occupancy/Toll (or HOT) lanes in the United States (Elliott, 1975; Fielding and Klein, 1993; Sullivan, 1998), cordon charges in Norway (Larsen, 1995), distance based charges for freight in Germany (Rothen-gatter, 2004), and area based charges in Singapore (Field, 1992; Holland and Watson, 1978) and London (Bell et al., 2004). Clearly there is resistance to user charges, which serve both to manage roadways by temporally or spatially redistributing traffic, and to raise revenue. They are significantly more complex and require a much larger bureaucracy to administer than more traditional road financing measures that rely on taxes. Reports from the London Congestion Charge indicate administrative costs upwards of 67 per cent of revenue raised (Hensher, 2003), and other toll based systems typically lose 20–25 per cent of their revenue to administration, which is significantly less efficient as a revenue-raising scheme than petrol taxes, which often lose less than 1 per cent of revenue to administrative processes (Wachs, 2003).

The first and second parts of the tariff, paying for the fixed costs of roads, especially when marginal costs are not rising, has not seen the same level of academic interest. Yet there are a number of complexities because of the variety of vehicles using roads, the spatial structure of networks, the varying lengths of trips, multiple levels of government, and long life of infrastructure.

This paper reviews the issues around paying for the fixed costs of roads. Over history, a number of different approaches have been tried; some more appropriate to their context than others, and some that may see new life as changes in the underlying environment of transport network finance takes place. The next section considers the classification of roads, both functionally, economically, and by ownership. This is followed by a review of the history surrounding the evolution of road financing in the Anglo-American context. The charging of developers, which can take a variety of forms, is
considered. Then an example, illustrating some of the many issues, is provided. The paper concludes with some thoughts on the future of funding the fixed costs of roads.

### 2.0 Classifying Roads

Roads serve many purposes, but broadly they can be organised along two dimensions, providing access to land and serving the function of movement, and of course some roads do more of one than the other. Roads that serve movement tend to be faster and have higher capacities (serve more vehicles per hour), while those that serve land access will be slower with lower capacities. This distinction gives rise to the hierarchy of roads found on any network and described in many transport engineering texts. Figure 1 illustrates the concept, where roads tend to fall along the diagonal.

This hierarchy, applicable with relabelling to other modes, explains a lot about transport financing and, when we discuss road financing, we have at least three distinct types of facilities that each require different means. The first and second parts of the tariff correspond to the top and bottom of the hierarchy respectively. Underlying that tariff is an assumption that roads are public facilities. However, the issue of financing cannot be isolated from the issue of ownership (public vs. private), nor, within the public sector, of which level of government manages the road, nor how private sector roads are regulated. Economists often define public and private goods using the dimensions of rivalry and excludability, as summarised in Table 1.

![Figure 1](image-url)  
*Functional Highway Classification and Type of Service Provided*
Table 1

<table>
<thead>
<tr>
<th>Excludability</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rivalry</td>
<td>Private</td>
<td>‘Congesting’</td>
</tr>
<tr>
<td></td>
<td>Limited Access Arterials</td>
<td>Linking Collectors</td>
</tr>
<tr>
<td></td>
<td>Club</td>
<td>Public</td>
</tr>
<tr>
<td></td>
<td>Local Streets</td>
<td></td>
</tr>
</tbody>
</table>

Table 1 suggests that limited access arterials, including fully restricted facilities such as freeways, as well as major signalised arterials with limited driveway access, exhibit, with toll technologies, the properties of excludability and rivalry (congestion), making them, in economic parlance, private goods. The recent experience with private roads in the United States has been mixed at best. Southern California’s privately built median toll lanes of SR91 were recently taken over by the Orange County Transportation Authority (which operates other toll roads in the county) for $207.5 million in cash and assumed debt (Level 3 Communications, 2003). This takeover was prompted by increasing congestion in the corridor, which the public was unable to remedy because the franchise agreement given to the private owners prohibited expansion of competing facilities. The privately built Dulles Greenway in Virginia has also had a mixed experience, though is now doing better than it was in its early years when it had to restructure its debt and lower tolls to attract traffic, having recently expanded capacity to accommodate demand and raised toll rates (Wooldridge et al., 2002). Like California’s experience with SR91, Spain, France, and Italy established private concessions to operate toll roads in the post World War II era, though many of these were either consolidated or taken over by the national government. Eastern European Countries have followed in the post Cold War Era. Many developing countries in Asia and Latin America use private toll roads for intercity transport, with a variety of different concession strategies and varying levels of success (Cobin, 1999; Gomez-Ibañez and Meyer, 1993; Hambros, 1999; McCormack and Rauch, 1997).

In contrast, local streets are, or can be, excludable — gated communities being the classic example — but because of their low levels of traffic, are not rivalrous, suggesting they are club goods (Buchanan, 1965; Cornes and Sandler, 1996). These club goods are generally publicly owned, but may also be private, as in St Louis (Hunter, 1988). Benton Place in St Louis is the oldest such private street in the United States (Beito and Smith, 1990; Lafayette Square Marquis, 1998), wherein the adjoining properties are
required to join the Benton Place Association, which assesses funds annually for maintenance. Laid out across private lots, Benton Place was designed to control traffic and, like the Dutch Woonerf, to allow use of the road space for public activity such as children’s play.

The middle level of roads may suffer congestion, but are not really excludable, as they connect local streets with each other and with limited access arterials, and typically serve many adjoining businesses, putting them in a third category.

Levinson and Yerra (2002) examine the issue of which level of government should pay for roads: local governments are closer to the problem, but may not be able to attain scale economies, while state governments, which can achieve some economies, may have problems of span of control. Levinson and Yerra find that there is a share of expenditures by each level of government on each highway class, which results in a minimum expenditure for each funding category (both capital and operating expenses).

3.0 Rights of Passage

The word ‘road’ from the Old English ‘rad’ shares its root with the modern word ‘ride’, and in that etymology is an important understanding of what roads were historically: rights of passage, places where one could ride across property held by another. In 1285, the Statute of Westminster, apparently the earliest English law concerning the subject of road maintenance required of manors ‘…that highways leading from one market town to another shall be enlarged where as bushes, woods, or dykes be, so that there be neither dyke nor bush whereby a man may lurk to do hurt within two hundred feet of the one side and two hundred feet of the other side of the way’ (Webb and Webb, 1913). The assumption is that the manors maintained roads (rights-of-way) across their property, and were responsible for ensuring no danger. There were in this same era more important streets and roads, which would be maintained by other organisations. For instance, in England, Catholic monasteries often maintained roadways along major pilgrimage routes, a practice that came to an end with Henry VIII’s break with Rome and the creation of the Anglican church.

Changing social and economic relationships led over a period of centuries to the replacement of the statute labour system of building roads to one of turnpike authorities. The Highways Act of 1555 established a process whereby surveyors, reporting to Quarter Sessions (quarterly meetings equivalent to a modern County Council) would be elected to supervise four (and later six) days per year of statute labour from adult
male residents without property (those with property would effectively buy their way out of this service by providing capital, including in-kind capital such as horses and tools, as well as the more generic money), a procedure that lasted to 1835 (Albert, 1972; Pawson, 1977; Webb and Webb, 1913).

The problem with statute labour rather than money as a means for financing road construction is the resulting productivity problem. While workers can be forced to report to duty, they cannot be made to work hard, and perhaps equally importantly, because of the temporary nature of their servitude, they are not likely to work skilfully or efficiently, and are unable to obtain the efficiencies associated with an appropriate division of labour.

The solution to this problem is a paid, and somewhat better skilled, labour force, which then raises the question of how to raise revenue to pay for labour. The answer, once it began emerging in the late 17th century, was obvious: place tolls on the roads. The nature of tolls varied from place to place; the first to be tolled were the heavy vehicles (especially those on narrow wheels) that rutted the unpaved roads of the day. As the economy continued to grow during the 18th century, heavier and heavier freight was being transported longer and longer distances, damaging more and more road surfaces. The English Parliament passed enabling legislation allowing communities to establish tolls, something especially important for communities with a lot of through traffic that did not add to the local economy but did damage local roads. Bogart (2003, 2004) provides evidence that these non-profit turnpike trusts did provide better quality roads, lowering the costs of carrier services (carriage), illustrating the trade-off between investing in capital (the roads) and investing in labour and services. Bogart (2005) provides statistical evidence that communities in the United Kingdom that had more through traffic were more likely to toll their roads.

In the US, toll roads were generally profit-seeking local companies, whose stock in part was owned by local governments and in part by local investors. While profit-seeking, they were often not profit-making, a fact that was well-known; Klein (1990) suggests that investment by local elites was viewed as a social obligation to help the community develop as much as, if not more than, an attempt to earn a high rate of return.

The technology for enforcing tolls was quite simple: the turnpike, a spear placed across the road that would lift when the toll was paid. While traffic levels were relatively low, the only inconvenience of such technology is the requirement of having change, and the stopping to pay someone in a tollbooth or tollhouse. Toll road (as well as canal and river navigation and railroad) operators strove to maximise revenue, and this meant developing means to price-discriminate between different customer classes, which became increasingly complex as the industry matured.
(Odlyzko, 2004). As traffic levels grew, especially in the 20th century incarnation of toll roads, delay at these tollbooths became another inconvenience. Toll farming, the leasing of the right to collect tolls in exchange for a fixed sum, became a widespread (if unpopular) practice throughout Europe, much like tax farming (Albert, 1972; Turvey, 2003).

This first toll era lasted until the second half of the 19th century. Competition from railroads took away long-distance business, and thus revenue, from toll roads (Baer et al., 1993). Toll roads, which operated at the sufferance of government, were not especially popular with users (Albert, 1979), and when the opportunities arose, the roads were disturn-piked and taken over by local, county, or state governments. The new government management, accompanied by the good roads movement of the late 19th century, led a further improvement in roads, and the paving of roads in the early 20th century was concomitant with the rise of the bicycle and then motor car as modes of transport. Until the advent of the gasoline tax, special assessments and general government revenue were used to pay for road improvements.

Figure 2 shows the history of the federal gasoline tax in the United States and in one state, Minnesota. The federal tax followed by a decade

![Figure 2](United States and Minnesota Gasoline Tax and Allocation)

Sources: Buechner, William (n.d.) History of the Gasoline Tax. American Road and Transportation Builders Association, [http://www.artba.org/economics_research/reports/gas_tax_history.htm](http://www.artba.org/economics_research/reports/gas_tax_history.htm)

Ryan, Barry, Presentation at Oberstar Forum, University of Minnesota, 17 April 2005
the development of gasoline taxes in the individual states, and has been used primarily in the transportation sector for the funding of capital projects rather than operations or maintenance. In the US, unlike other countries (Newbery and Santos, 1999), the tax has been largely earmarked (or hypothecated) to roads. Within both the authorising legislation passed every six years or so, and the annual appropriations bills, there is further earmarking, so that funds are dedicated to specific projects (Utt, 1999). However, the gasoline tax has not kept pace with inflation, indicating a real drop in funds available for new road construction since peaking in the state in 1940 (as rural roads were paved) and for federal taxes in 1960 (with the onset of the interstate era).

The interstate highway system beginning in the 1950s increased the federal role in the financing of US highways. First, new interstates were to be untolled (though toll roads that had been built from 1940 on were grandfathered into the system). Second, the federal government would pay 90 per cent of the cost, recovered primarily through federal gasoline taxes, leaving the states to match only 10 per cent to build the freeway. Lower level roads remained state responsibilities, although with the 1991 passage of the Intermodal Surface Transportation Efficiency Act, a much larger National Highway System became eligible for federal funds, with a reduced 80 per cent federal, 20 per cent state matching plan.

4.0 Developers

In the past several decades, especially with the decline in real gasoline tax revenues, local governments have begun to connect the need for new infrastructure with new development (Altshuler and Gomez-Ibáñez, 1993; Bauman and Ethier, 1987; Heath et al., 1989; Lee, 1989; Levinson, 1998; Nelson, 1989). Historically developers had been required to provide infrastructure on the site they were developing, but more recently, off-site improvements have been required. There are several legal tools that are useful for extracting resources from developers.

Developers may voluntarily provide infrastructure (or cash for infrastructure) to a jurisdiction, a system called proffers. In exchange, developers may obtain some sort of approval, often a rezoning or other permission to develop. When the required (or desired) infrastructure is expensive and beyond the scope of a single developer, multiple developers may join together and form a road club, signing an agreement with the relevant jurisdiction, in exchange for collective permission. The voluntary nature of these agreements may be considered suspect by members of the
development community; on the other hand, the developer has the right to develop at the legally allowed limit without a proffer. That limit may be too low a density to be considered profitable, however.

Communities may do away with the veneer of voluntarism and impose some form of *exaction*. An exaction is a requirement that developments pay money (or build infrastructure in kind) in exchange for approval. These, in the form of impact fees, became especially popular in California following the adoption of Proposition 13 in 1978, which greatly limited the rate of property taxes that could be collected by local governments. In *Nollan v. California Coastal Commission* (483 US 825 (1987)) the United States Supreme Court required there to be a nexus between the conditions (financial or otherwise) imposed on the development and ‘legitimate state interests’ the government is trying to extend. In a later case, *Dolan v. City of Tigard* (512 US 319, 114 S.CT. 2309 (1994)), the Supreme Court ordered ‘rough proportionality’ between the exactions and the public impact of the project. That said, most jurisdictions in the United States at present do not exact conditions from developers that at all compensate for the development’s impact on infrastructure.

Another way to achieve this end is the formation of *development districts* wherein developers would be permitted to proceed with development upon payment into a fund that prospectively covers the cost of all needed (planned) infrastructure.

A final alternative is the use of *benefit districts* (Stopher, 1993). In a benefits district, the value of property goes up as a result of the infrastructure having been built, due to the additional accessibility or reduction in congestion that the facility provides. Charging a premium (above and beyond the already increased assessment that should arise when property values increase) can be used to defray retrospectively the cost of the infrastructure or pay back the bonds. This is more widely used with rail projects, but can be associated with new highways as well.

We can organise these developer-payment tools according to their timing (prospective — before the infrastructure is built, or retrospective — after the infrastructure is built), whether they are voluntary or mandatory, and whether the geography applied is a particular development site, or a larger area, as shown in Table 2. One could further consider the

<table>
<thead>
<tr>
<th></th>
<th>Site-Based</th>
<th>Area-Based</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voluntary</td>
<td>Proffers</td>
<td>Road Club</td>
</tr>
<tr>
<td>Mandatory</td>
<td>Exaction</td>
<td>Development District (prospective)</td>
</tr>
</tbody>
</table>

Table 2
*Types of Developer Payments*
timing, whether the financing is pay-as-you-go or bond-based, but either can be used in these scenarios. One can imagine that it would be difficult to impose retrospective financing in anything but a mandatory, area-based way.

5.0 Vehicle Type

The United States federal government has paid particular attention to cost allocation. Federal Highway Cost Allocation Studies (HCAS) were completed in 1982 and again in 1997, and aim to determine the responsibility of different vehicle classes (cars, trucks by type) to the cost of roads, including both infrastructure and social costs. In the US, cars, pickups, and vans, generally owned by individuals, comprised 92.4 per cent of vehicle distance travelled (FHWA, 1997). However, the costs imposed by trucks are more significant than then the 7.4 per cent of distance travelled may suggest. First, paving thickness is determined by truck weight per axle (an equivalent standard axle loading or ESAL is often used). So all roads that wish to accommodate trucks must be thicker even though cars do not require it. The truck-highway-paving system could perhaps be more optimal if the country had a different mix of vehicles, and more axles on vehicles (Small, Winston and Evans, 1989), which a different tax/toll structure might be able to induce. The disjoint ownership of vehicles and roadways may make some inefficiencies inevitable. Even if the paving problem was addressed, there are a variety of other costs that trucks impose on roads. Bridges must be stronger to bear the entire weight of the truck, a cost that is also captured in various cost allocation efforts. Overall the federal highway cost responsibility for trucks estimated by FHWA, considering both paving and bridges, is just over 40 per cent, while they are estimated to pay in user fees just under 36 per cent (FHWA, 1997). The cost allocation system according to the FHWA is least fair to pickups and vans, small single-unit trucks, and small combination trucks, which overpay by 40, 50, and 60 per cent respectively, while large single-unit trucks and combination trucks underpay by 50 per cent.

This analysis does not consider a number of other costs trucks impose. Trucks also determine grade and banking (or superelevation) on roads. In short, nearly 100 per cent of roads today are designed based on constraints imposed by fewer than 3.3 per cent of all vehicles and 8 per cent of all vehicle distance travelled. The cost of decreasing the grade of roads systematically has never been calculated, and is at this point in history in the US largely a sunk cost.
On the other hand, the width of the road (in terms of number of lanes or capacity), is determined by peak period travel, primarily passenger cars and individually owned trucks. Excess capacity cannot be rolled up at night.

### 6.0 An Example

Consider the case illustrated in Figure 3. We can use it to illustrate the spatial and temporal issues with a variety of fixed cost financing schemes. Assume that equal-sized areas $a$ and $b$ are the only two residential subdivisions of jurisdiction $j$. It is generally accepted that residents (or developers) of area $a$ pay for the on-site improvement that is cul-de-sac $u$, and area $b$ pays for cul-de-sac $v$. (See for example Levinson, 2002, Chapter 5). How to pay for links $z$ and $y$ remains to be resolved. Assume publicly provided, pay-as-you-go roadways, which is typical in the United States.

First, assume that development $a$ comes before development $b$. In that case, it is likely that the residents or developers of $a$, who comprise all of the residents of $j$, pay for $z$. This is fair in that the users of $z$ have paid for $z$. Problems (or opportunities) arise because of indivisibilities in the road network. Assume link $z$ is one lane in each direction, which provides excess capacity beyond what is required by the residents of $a$. A second developer proposes to build subdivision $b$. As part of that development, the on-site link $v$ is constructed. Who pays for link $y$? The developer of $b$ may suggest that since it is off-site, the whole community (both residents of $a$ and $b$) should share the costs. If this occurs, the residents of $a$ will have paid for all of link $z$ and half of link $y$, while the residents of $b$ will have only paid for half of link $y$. Clearly the residents of $b$ are at an advantage. The existing residents of $a$ may on the other hand try to exact link $y$ as an off-site improvement from developer $b$, in which case, the residents of $a$ will have paid for link $z$ and the residents of $b$ will have paid for link $y$, which sounds fairer, but the residents of $b$ still use link $z$ (while few residents...
of a use link y because of the spatial configuration), and so they still have been subsidised.

The developer of b may note that the residents of a had use of the road before b even existed, and would have had to pay for z had b never happened, so that it is fair. If the developer is economically astute, he may also note that traffic from link b does not add to the marginal costs of constructing link z, and so the residents of a are really no worse off sharing the capacity.

A solution to the temporal problem would be to fund link z using bonds rather than using a pay-as-you-go system, bonds that are paid back by the whole community in jurisdiction j over the life of the facility (Levinson, 2001). When residents move into b, they pay taxes that help repay the bonds, ensuring that facility z is there when it is needed for residents in a, and paid for in proportion to use by residents of both a and b. However, this does not resolve the problem of paying for link y. If the same mechanism is used, then residents of a pay for a facility they do not use, thereby subsidising the residents of development b. Bonds paid by all residents, which works like a club, is only fair if all facilities are used approximately equally by all users. If one set of residents (for example those in b) place an undue burden, they should pay extra.

A solution to the spatial problem would be to fund each link based on which development caused its construction; this is the approach taken by impact fees. So link z would be charged to subdivision a and link y would be charged to subdivision b, again resulting in an effective subsidy for subdivision b.

A solution is required to address both the spatial and temporal free-rider problems, one that would result in the costs of link z being shared by a and b, while link y is paid for solely by subdivision b. It is here that the question of sector of ownership and the tying of financing to payment arises. Were the road private, or managed by an independent turnpike authority, the financing would clearly switch from the public to the user. If the annual cost to build links z and y were denoted Cz and Cy, and the annual number of trips out of subdivisions a and b were Ta and Tb respectively, average cost financing would charge a toll of

\[ \tau_z = \frac{C_z}{T_a + T_b}, \]

\[ \tau_y = \frac{C_y}{T_b}. \]

On the other hand, were the links to be separately privately owned by organisations with a profit-maximising objective, the issue of serial
monopolies arises (Chamberlin, 1933), which would lead to higher prices for consumers than if they were jointly owned, which remains higher than a welfare-maximising toll (zero if there are no marginal costs) or an average cost recovery toll noted above.

These problems would not arise were there no indivisibilities or economies of scale, but of course there are in real networks.

Now assume that subdivision $b$ comes before subdivision $a$. In that case, $b$ must build links $y$ and $z$ in order to have access, and subdivision $a$ gets to ride free. Society as a whole must pay more up-front costs (the same number of people could have lived in jurisdiction $j$ without link $y$ if subdivision $a$ had gone first), illustrating some disadvantages of leapfrog development. The other set of issues remains however.

**7.0 Conclusions**

Over history, a variety of techniques have been employed to finance the fixed costs of roadway infrastructure: collecting revenues from the general public, users, and developers, differentiating users by type, discriminating based on willingness to pay, to either cover total costs, or to earn profit. The examination of history and investigation of a specific example may give us insight into the future.

The costs of collecting revenues from tolls should drop with the advent of electronic toll collection (despite the experience in London), while rising congestion creates additional incentive to employ user-based financing systems. Mohring and Harwitz (1962) show that revenue from optimal congestion pricing covers the capital costs of the optimal level of infrastructure under certain conditions, including constant returns to scale in congestion and capacity expansion. However, as Verhoef (2000) notes, conditions are seldom optimal, leading to what is called ‘second-best’ pricing, such as the HOT lanes, cordon tolls, and area based tolls that are now being deployed, which do not have such economically convenient properties.

The absence of pricing surely requires additional funding for the fixed costs of infrastructure, while inefficient pricing may lead to shortfalls as well. While roads are considered a public service, average cost pricing, with careful attribution and allocation of costs temporally, spatially, and by user class will still be required for both efficiency and equity. Roads that are privatised have different objectives, though they may either be regulated much like public utilities, which again returns to some form of average cost pricing, or allowed to be truly profit maximising, using price-discrimination to differentiate users by their willingness to pay.
References


