Urban Transportation Policy: A Guide and Road Map

Kenneth A. Small
Department of Economic
University of California at Irvine
Irvine, CA 92697-5100

ksmall@uci.edu

First version: April 2007
This version: June 2007

Prepared for: *Unraveling the Urban Enigma: City Prospects, City Policies*, conference and book, Wharton School, University of Pennsylvania, conference May 3, 2007.

Abstract

The main transportation issues facing cities today fall into familiar categories– congestion and public transit. But the emerging needs in each area are quite different than those most widely understood and commonly analyzed.

For congestion, there is now a far richer menu of options that are understood, technically feasible, and perhaps politically feasible. This is accounted for by several factors:

- Product differentiation: One can now contemplate offering roads of different qualities and prices, and allowing users to choose.
- Privatization: Many selected road segments are now operated by the private sector. Transportation officials are keenly interested due to financial constraints.
- Attitudes toward pricing: Road pricing is routinely considered in planning exercises, and field experiments have made it more familiar to urban voters.
- Goods movement: Urban trucking has grown in its environmental effects and links to the urban economic base, especially in port cities. One result is serious interest in tolled truck-only express highways.

For public transit, several similar factors call for changes in policy:

- The dominance of large public transit agencies has led to an undesirable homogenization of service. There is a need for political mechanisms to allow each type of transit to specialize where it is strongest.
- The spread of "bus rapid transit" has opened new possibilities for providing the advantages of rail transit at lower cost.
- The prospect of pricing and privatizing highway facilities could reduce the amount of subsidy needed to maintain a healthy transit system.
- Privately operated public transit is making a comeback in other parts of the world. Lessons there may offer pointers for the US.

The single most positive step toward better urban transportation would be to encourage the spread of road pricing. A second step, more speculative because it has not been researched, would be to use more environmentally-friendly road designs that provide needed capacity but at modest speeds, and that would not necessarily serve all vehicles.

Acknowledgment

The author is grateful to Robert Inman, Charles Lave, Robert Poole, Jr., and Martin Wachs for helpful comments on earlier drafts. All responsibility for the contents lie with the author.

Urban Transportation Policy: A Guide and Road Map

Kenneth A. Small, University of California at Irvine

Introduction

Cities exist and thrive because they enable people to access each other. Thus they depend on a good transportation system, as confirmed by the strong impacts of transportation infrastructure on both economic growth (Gramlich 1994) and urban structure (Giuliano 2004). Furthermore, there is little doubt that, at least in a city of any size, a healthy economy requires a transportation system that includes both private and public modes, since neither alone can possibly accommodate the enormous variety of trips that such an economy generates. Each mode involves important policy decisions about the extent of capital investment, the level of service provided, and the financing and pricing of that service.

Given its importance, it is no surprise that transportation is often a top concern to urban residents. Periodically this concern rises in prominence as one or another part of the system appears to be near breakdown. The high labor intensity of public transportation, combined with a variety of pressures toward more dispersed trip patterns, subject it to severe cost pressures that occasionally erupt in service cutbacks or unsustainable fiscal drains. Meanwhile traffic congestion on highways, inherent in urban life but never really accepted, continues its steady march toward an apparently intolerable future. Adverse environmental effects of traffic, and of the activities that support it, just add to unease about the health of the underlying system which depends so strongly on motor vehicles.

Often the reaction to both of these problems is to propose infusion of funds into public transit. Whatever the wisdom of this approach for ameliorating the fiscal problems of transit, experience indicates that it can have at best a very small effect on traffic congestion. Yet highways carry the overwhelming majority of urban trips in virtually all metropolitan areas. Therefore, realistic planners typically look to highway capacity enhancements as the main weapon against congestion. If one wants a simple explanation for why congestion is growing, it is not hard to find: for many decades, road capacity has grown far more slowly than vehicle-miles of travel — especially in urban areas. Table 1 shows some representative figures for the US, covering 1980-2005.

Table 1. Demand factors and supply of road capacity: US 1980-2005

	1980	2005	Growth
All Areas:			
Adult population (age 15+) (millions)	175.3	235.7	34.5%
Registered motor vehicles (millions)	155.8	241.2	54.8%
Vehicle-miles of travel (billions)	1,527.3	2,989.8	95.8%
Lane-miles of major roads (thousands)[a]	552.2	724.6	31.2%
Urban areas:			
Vehicle-miles of travel (billions)	855.3	1,951.9	128.2%
Lane-miles of major roads (thousands)[a]	219.9	351.5	59.9%

[a] "Major roads" are defined as expressways and other principal arterials.
Sources: US FHWA (1997), Table MV-200, VM-201, HM-260; US FHWA (2006a), Tables MV-1, VM-2, HM-60; US Census Bureau (2006), Table 11.

Yet while planners are busy projecting capacity needs (often requiring hopelessly infeasible levels of funding) and detailing apocalyptic scenarios in case the capacity is not forthcoming, many analysts are pessimistic about what can be accomplished through more road capacity. Since World War II, enormous capacity investments have been made; they have accommodated an impressive growth in population, mobility, and motor-vehicle traffic in particular, but they have not stabilized congestion levels. Neither have transit initiatives, which have made minor if any inroads into the decline of transit mode share. Improved signal timing, freeway ramp metering, carpool lanes, transit reorganization, and land-use policies all have their effects on congestion, but they are small. Downs (2004) draws a stark conclusion: congestion is here to stay because there is only one policy — road pricing — that can stop it, and the public will not support road pricing. In other words, public policy is at an impasse.

There is a consolation: congestion is self-limiting. It can exist only because people — lots of them — are willing to put up with it. Discourage them enough, and they will remove themselves from the upward pressure of highway use. An equilibrium is thus reached in which the cost of travel, including users' own time, becomes high enough that demand is kept in check. The trouble with this solution is that it may be a quite unhappy equilibrium, disliked by those affected, politically potent, and economically inefficient. Moreover, the functioning of the urban system, broadly conceived as an economic and cultural system aimed to provide enjoyment of life, is compromised.

I offer here a slightly less pessimistic view than that of Downs, focusing on some newer developments in transportation policy and analysis. Some of these developments make pricing more likely, while others broaden the comfort zone within which pricing and non-pricing innovations might work. While many of these developments are genuinely emerging trends, some are gleams in the eye of this beholder, representing possibilities that I believe are unleashed by emerging trends but in non-obvious ways.

I first consider developments related to highway congestion, then those related to public transit. I conclude with observations about what transportation policy could look like if things go well.

Highway congestion and level of service

Product differentiation and heterogeneous preferences

Downs's thesis about the impossibility of serious congestion relief is formulated in the context of a homogeneous population. He postulates that people uniformly want to travel in dense areas at peak times, and have common political preferences against such possible ameliorative policies as regional land-use controls, elimination of housing subsidies, or road pricing. This is of course an abstraction, and I believe a useful one. But recent events have highlighted another aspect of urban life: the heterogeneity of types of users and of people's preferences concerning their travel conditions. Accounting for such variety changes the terms of the impasse described by Downs.

Decades ago, urban road managers realized that a highway system's efficiency might be increased by encouraging people to carpool, so that the same number of trips can be served in fewer cars. Thus began an extensive and still ongoing process of restricting certain expressway lanes to carpools. The hope was to stimulate more carpooling, which it does to a small extent. But the primary effect is to lower aggregate passenger travel times by offering faster service to those people, such as some long-distance commuters to large employment centers, who choose to carpool for other reasons. (To the dismay of planners, carpool lanes have also attracted family members, especially parents carrying children, who would not otherwise be using more than one motorized vehicle.)

The system of carpool lanes is not universally admired, especially where there is insufficient carpool traffic to keep an entire lane well utilized. In a few cases, carpool lanes have been decommissioned and returned to general use. Dahlgren (1998) notes the rather narrow range of parameters that lead to carpool lanes meeting commonly applied criteria for success such as high utilization, high level of service, and savings in aggregate travel time.

But interestingly, the system of carpool lanes has opened the door to a broader notion of how differing levels of service can be offered to people in different circumstances. In analytical terms, we could say that with carpool lanes, cars are differentiated by the total value of time they carry, on the presumption that each passenger values his or her time enough so that the aggregate value in a carpool is typically larger than that of a solo driver. This is a crude form of product differentiation, whereby two otherwise similar products (expressway travel) are offered with different quality levels.

In other realms of economic life, product differentiation is much more common and is usually accomplished by pricing. No one thinks twice about the option to pay more for better theater tickets, faster computers, nicer restaurants, or first-class seats on airplanes. But public services have typically not been differentiated; and the provision of infrastructure for highway travel, because it is typically provided by government agencies, has mistakenly been viewed as another public service — despite the fact that highway travel lacks the features that normally define public goods.

Recent events, motivated by quite other considerations, have expanded the idea of product differentiation on roads beyond carpools to paying more for better quality, just as for other goods. I will describe in the following subsections how this came about, so here let me just describe the outcome. In several US metropolitan areas, beginning with Orange County and San Diego (California) and Houston (Texas), express lanes were constructed or converted from carpool lanes and offered not only to carpools but also to solo drivers willing to pay a toll. In most cases the toll varied quite steeply by time of day. In each case the concept was presented as a modification of carpool lanes, either as an explicit conversion for the purpose of making better use of spare capacity in an existing carpool lane, or (in Orange County) as a substitute for previous plans to add a carpool lane. (In Orange County this change was not welcomed by neighboring Riverside County, where most of the users live, although there was a sweetener: users of the very congested corridor got four new express lanes instead of two).

These early experiments were encouraged, if belatedly, by a federal program begun in 1991 to support innovative demonstrations of congestion pricing. Finding no takers for actual implementation of pricing on an entire corridor, the US Federal Highway Administration incorporated these "high occupancy/toll" (HOT) lanes into the program later in the 1990s. Indeed, HOT lanes were so much more popular than congestion pricing that the Congestion Pricing Pilot Program was renamed, changing "Congestion Pricing" to "Value Pricing" — a term originally coined to indicate product differentiation involving pricing, but one that has come to be used synonymously with (or euphemistically for) congestion pricing.

HOT lanes have spread to several other states and are prominent components of a number of local, state, and federal transportation plans. Poole and Orski (2003) develop a nationwide proposal for networks of such lanes, including considerable investments in connecting interchanges so that users could avoid mixing with regular traffic when changing from one HOT route to another. (This feature also is levered off plans for carpool lanes, which in some states involve similarly expensive interchanges but restricted to carpools.) Their proposal also incorporates bus transit vehicles.

HOT lanes have been joined by some additional instances of new toll roads that serve as congestion relievers. Toll roads have of course existed in many states, sometimes in urban areas, and have provided *de facto* product differentiation vis-à-vis non-express arterials. But recently, as urban areas have largely filled in their developed areas with expressway systems, new toll roads serve more often as alternatives not only to arterials but to other urban expressways. An early example just outside the US is Highway 407 in the Toronto metropolitan area, opened in 1997 and running parallel to and just a few miles from the highly congested Queen Elizabeth Way. A more recent example is E-470 in Colorado, a tolled half-beltway connecting heavily developed suburbs north and east of Denver to each other and to Denver International Airport. Public opinion shows an increasing if reluctant acceptance of such ventures as sources of funding for needed capacity, acceptable because there are reasonably close substitutes that remain free.

The upshot of this activity is that people are becoming used to the idea of tolled facilities offering premium service. Furthermore, dramatic improvements in pricing technology, supported by aggressive corporate development within a dynamic and quite competitive industry, have made it possible to implement far more sophisticated pricing systems. Market penetration of electronic toll payments has reached well over half on many toll roads, and HOT-lane operators increasingly offer toll schedules that assume some sophistication on the part of users. In a small number of cases, including HOT lanes in the San Diego and Minneapolis regions, pricing is “dynamic”: that is, the price is varied in real time, depending on congestion levels in the adjacent lanes, in order to keep the HOT lanes busy yet congestion-free. In what came as a surprise to many, users had little trouble adapting to dynamic pricing and it works smoothly.

Thus two prerequisites for road pricing — public familiarity and feasible charging technologies — have entered the public realm more or less by accident. In addition, transportation planners are by financial necessity becoming locked into systems that contain priced facilities. Thus it is likely that most current experiments will continue even though a priced facility is occasionally returned to free status, due for example to paying off a bond or to political problems with a private operator (an example of which is described later).

Privatization

Financial pressures have induced state and local authorities to seek arrangements with private investors to hasten the process of building capacity. There are many institutional forms of private involvement in road finance, ranging from financing a publicly designed road to building and managing the road privately. All of them involve some form of privately provided finance recouped by toll revenues. Usually the arrangement is specified as a franchise that spells out the rights to operate a toll road for a specified period of time with some limitations on tolls, toll increases, or the rate of return to capital.

Private participation in road operations is significant for policy toward congestion for several reasons. First, private investors, as well as financial institutions supporting them through loans, have a strong incentive to accurately forecast demand for the road. This raises the level of knowledge about impacts of a given capacity expansion on the road network, and it helps steer investment away from projects that benefit only a narrow interest group.

Second, private firms have experience with price setting and generally understand such important features as price sensitivity, public perceptions, marketing, and the roll of price differentiation. This is precisely the kind of knowledge needed to bring analytical models of price setting into a form that can be implemented practically.

Third, private road operators have a financial incentive to enhance economic efficiency in road use, which is also the goal of standard congestion-pricing theory. It is well known in the academic literature that a private road operator, even one with a monopoly, will choose to differentiate prices by time of day in a manner very similar to that called for by standard congestion pricing recommendations (Small and Verhoef 2007, Section 6.1.1). This is because the private operator can charge higher tolls if it can provide a high level of service by keeing congestion down. In fact, the pricing structure that provides the greatest revenue is basically the same as the structure that maximizes public benefits as normally defined, although the *level* of prices may be substantially higher in the former case. For this reason, as private operators propose, bid on, or negotiate franchise agreements, they will tend to encourage public authorities to consider differentiated toll schemes that might otherwise be ruled out for simplicity, but that in fact can both increase revenue and help manage congestion.

The ability of private operators to capture benefits of improved efficiency is so strong that even existing toll roads have begun to have their operating rights sold by state or local governments to private investors. Two very large such sales occurred in recent years: the Chicago Skyway, owned by the City of Chicago, was franchised in 2005 and the Indiana Turnpike in 2006. These two road segments, which connect to each other, form parts of Interstate Routes 80 and 90, two major east-west corridors in the northern United States. Long-term leases were granted in return for up-front payments of $1.8 billion and 3.8 billion, respectively. Although the agreements do not specify time-of-day pricing, they do give the franchisees incentives to adopt road-maintenance strategies that better match user preferences. It seems only a matter of time before the financial advantages of price differentiation by time of day lead to proposals that include such differentiation.

Leasing an existing asset introduces some interesting issues of politics and public finance. The effect is usually to shift control over net revenues from a dedicated toll authority to a political authority, and to shift the timing of these revenues from gradual receipt over many years to a lump sum at the time of the lease. Whether and how this shift is accompanied by controls over spending the revenues can greatly affect transportation funding in current and future years.[1]

[1] The Chicago Skyway lease provided up-front funds to the City of Chicago for general use. For the Indiana Turnpike, authorizing legislation places the proceeds in a ten-year highway improvement fund. The Governor of Pennsylvania proposed to place proceeds from a lease of the Pennsylvania Turnpike in a transportation endowment fund of indefinite lifetime, whose earnings would support transportation improvements.

Another interesting feature of private highways is that in several cases, private firms have submitted unsolicited proposals for new highways. One notable example is the proposal by a consortium of Fluor Enterprises and Transurban (USA) for new HOT lanes on parts of the Washington Beltway and Interstate 395 in Virginia. These notoriously congested roads serve hundreds of thousands of Washington-area users, many of whom have influential positions in national government. Thus direct consumer experience with private pricing proposals will to some extent also become experience by policy makers. What is especially relevant about unsolicited proposals is that the private firms are free to suggest pricing schemes that otherwise might never make it through a public bureaucracy.

Private participation in highway capacity, then, not only makes it possible to add capacity more quickly than could be done otherwise. It also brings with it a host of factors favoring the use of pricing for congestion management. Like the experiments in product differentiation described earlier, experiments with privately operated highways may break the impasse that makes congestion so intractable.

Public attitudes toward road pricing and privatization

Generally, the public does not like road pricing. Most people think that free travel on roads is a traditional and fundamental right (although in fact turnpikes played very important roles in earlier centuries). Furthermore, the argument for pricing is abstract and involves offsetting the welfare losses directly experienced by individual users with welfare gains in the form of toll revenues, which users may not trust will be spent wisely.

Nevertheless, the various experiments and demonstrations undertaken during the last two decades have resulted in considerable changes in attitudes among affected users. Many people recognize that the private sector can deliver congestion relief sooner than the public sector and that paying tolls is the price of that accelerated schedule. In areas with HOT lanes, public acceptance has tended to rise over time, often reaching majority support. For example, before the HOT lanes opened in Orange County, California, about 65% of solo drivers approved of providing toll lanes to manage congestion; a year later, 69% to 82% of them approved, depending on whether or not they used the toll lanes. Before the project, a smaller proportion of solo drivers, 43%, approved of varying the tolls with the severity of congestion; a year later this rose to 60% of those continuing to use the free lanes and 73% of those using the toll lanes. (Carpoolers generally approved less of these concepts and did not show much change over time.) Approval of the concept of *private companies* operating toll roads rose among nearly all groups, to around 50%.[2]

Perhaps the most surprising political development in congestion pricing is the proposal by New York's Mayor Bloomberg for a cordon toll of $8 during daytime hours for entry into Manhattan south of 86th Street (except for circumferential travel on designated express arterials at the island's borders). Travel purely within the cordon would be priced at half that amount. This proposal is one of eight selected for further consideration under the US Department of Transportation's "Urban Partnerships" program, which offers substantial funding for innovative proposals specifically including congestion pricing. Bloomberg's proposal touched off a substantial political saga both in New York City, surrounding suburbs, and the New York State Legislature. Early press coverage suggests that two significant factors in obtaining support would be the availability of federal money, and the use of fees to obviate increases in existing bridge tolls and public transit fares.[3]

This is not to say that experience is always positive. The governor of Indiana faced considerable hostility over the sale of the Indiana Turnpike, which will lead to higher tolls, despite the fact that a high fraction of toll payers live in other states. Similar proposals for the New Jersey Turnpike and the Pennsylvania Turnpike have produced strong opposition. The extensive private investment program undertaken by Texas, with its first concession agreement in June 2006 for a new $1.35 billion road, has led to a backlash that seems certain to curtail the extent of planned privatization.[4]

---

[2] These numbers are estimated from Sullivan (1998), Figures 6-1, 6-3, and 6-12. Updates in Sullivan (2000) show that these approval ratings subsequently dropped substantially, probably due to some controversial actions of the private toll operator described later.

[3] See Schaller (2006), Bindrim (2007), "Congestion Pricing Deadline" (2007). For more details of the proposal, see Grynbaum (2007).

[4] The Texas legislature voted overwhelmingly in May 2007 for a series of limitations including a two-year moratorium on new franchise agreements and restrictions on the terms that can be offered. About the same time, a multi-billion-dollar agreement for extension of a toll road north of Dallas was reopened to allow a competing bid by the public toll-road authority in the region. See "Texas Toll Moratorium Hodgepodge" (2007) and " 'Price Check' Bid Upsets Cintra's Texas SH 121 Deal" (2007).

An example illustrates how even successful instances of privatization can be reversed, and yet still contribute to the growth in use of tolls to manage traffic congestion. The HOT lanes described earlier in Orange County, California, were originally constructed and operated under a long-term franchise by a private consortium. As noted earlier, the deal creating these lanes was already unpopular in inland Riverside County, where most of its users reside. Two public-relations snafus further eroded support for the private operation. First, the private operator made a clumsy attempt to reap a tax windfall by proposing to sell the lanes at a handsome profit to a newly created nonprofit organization, which would be eligible for tax-exempt bond financing. Press exposure revealed a less than fully arms-length relationship between the seller and the proposed purchaser. Second, severe congestion on the regular (free) lanes of the corridor returned more rapidly than expected following its drastic decline upon completion of the 91 Express Lanes (which expanded capacity in the corridor by about 50 percent). When the California Department of Transportation attempted to add some new capacity under the guise of safety improvements at a merger point, the private operator invoked the non-compete clause in its franchise, which turned out to be one of the most restrictive ever written for a private highway: namely, it prohibited any such expansion and lacked the more common provision for compensation in case of overriding public need. The upshot was that the express-lane franchise was bought out in 2003 by the Orange County Transportation Authority. The sale price gave the private operator a healthy profit on its nearly eight-year ownership of the road, and the terms of the loan underwriting the public takeover ensure that pricing will remain in place for many years. Indeed, as of April 2007, the price has been raised several more times, making it one of the most expensive roads per mile of travel in the US, with a peak rate of $9.50 for the 10-mile outbound trip between 4:00 and 5:00 p.m. on Fridays.[5]

Nevertheless, the public is getting used to road prices being among the innovations they are likely to see as policy makers grapple with intractable congestion. It will become increasingly hard to defeat pricing proposals on purely ideological grounds, forcing discussion into more objective consideration of actual effects. This enhances the possibility that pricing proposals with especially high congestion-relief benefits will get a hearing.

[5] The toll schedule is provided at http://www.91expresslanes.com/tollschedules.asp. Generally the inbound peak tolls, which occur during the morning, are only about half as large as the outbound peak tolls.

It is also important that analysts in federal, state, and local agencies are gaining experience with road pricing. Many such agencies have at least some staff with training in economics, and other staff members have grown in their ability to understand and assess analyses of economic efficiency. It is now quite common, even routine in some agencies, for a menu of proposals in urban transportation to include pricing. For example, the US Federal Highway Administration issues regular reports to Congress about the condition of the nation's highway infrastructure and needs for investment; its 2006 report includes the following statement in bold type:

> … congestion pricing has the potential to significantly improve the operational performance of the Nation's highway system, while significantly reducing the level of future capital investment that would be necessary to achieve any specific level of performance" (US FHWA 2006b, p. xi).

The report goes on to estimate that reduction in needed capital investment at about $21 billion per year. Two decades ago, such a statement by a highly visible public agency would have been considered impolitic.

Thus both citizens and technocrats are giving pricing solutions a hearing instead of dismissing them out of hand. This raises the chance that a successful package can be constructed, one that improves efficiency by using price to lower congestion, but still provides overall benefits that citizens understand and in which they have trust.

Goods movement

The movement of freight within and through urban areas has long been an important part of the economies of urban areas, and a significant requirement of transportation facilities. Freight is increasingly important to regional and national economies as trucking serves primary distributional roles for inter-regional and international shipments, often entering through water ports. The four largest US ports alone — New Orleans, Houston, New York, and Los Angeles/Long Beach — handled 710 million tons of traffic in 2004; one-fifth of this, valued at $442 billion, involved international trade.[6] Much freight traffic originating at ports travels by truck and/or rail via the urban infrastructure to inland destinations, some of it passing again

[6] US Census Bureau (2006), Tables 1043, 1061.

through urban hubs like Chicago. (The Chicago region generates an estimated 3,500 truck trips per day just connecting its own rail terminals, due to a shortage of connecting rail capacity.)[7]

Trucks impose considerable environmental and safety costs in the form of air pollution (especially particulates, which are the most clearly documented causes of severe health effects), noise, and collision damage to passenger vehicles. The explosion of port activity accompanying recent expansion of global trade has accentuated these problems and focused many residents and policy makers on finding alternatives to the large truck volumes found on some urban corridors, such as those serving the ports of Long Beach and Los Angeles, California.

One outcome of these factors is an interest in truck-only roads or lanes, often conceived as new capacity to be built for the dual purposes of congestion relief and channeling truck traffic to where it is less harmful (Poole 2007). Usually truck-only lanes are planned as toll facilities, either because they are proposed for accelerated investment or simply on grounds of equity. Typically the trucking industry has opposed special road charges aimed at them, but it has shown more flexibility toward truck-only toll lanes provided they are not made mandatory by prohibiting trucks from other highways. Thus, the outcome of attempts to deal with special and growing needs of freight transportation could be another type of differentiated highway service, offering premium service for those freight movements for whom the faster and more reliable travel are worth the payments.

Local distributional activity by trucks also creates significant problems for the urban transportation system, especially congestion resulting from loading and unloading. Shippers' desires for rapid and predictable deliveries, combined with carriers' attempt to minimize labor costs, may lead to larger vehicles on dense city streets than would be efficient from an overall system point of view. Local businesses depend on such deliveries and often strongly resist attempts to regulate them in the interests of traffic management. This can greatly complicate the politics of congestion.

Highway design standards

[7] McCarron and LaBelle (2002), p. 1.

The increased importance of product variety and differentiation casts doubt on some long-standing assumptions about design standards for highways. Furthermore, some standards that made sense when most travel was intended to be under free-flowing conditions are inappropriate to high-density urban settings where congestion must and should be common even under ideal policies.

Perhaps the clearest example of problematic design standards is the US interstate highway system. Interstate highways are expected to meet nationwide standards for lane width, sight distance, grade, shoulders, and other characteristics (AASHTO 2005). But simple economics suggests that where land costs, construction costs, and traffic volumes are high, one should trade off costly features like lane widths, shoulders, and long sight distances for more capacity whenever possible. In other words, those design features that provide better safety and ride quality at high speeds become less important relative to those that increase throughput at moderate (congested) speeds.

Another way to look at this is in terms of an equilibration of travel times across different types of highways. When the overall capacity of an urban traffic system is in heavy use, there will be a tendency for higher-quality roads to become congested more severely than others until their levels of service are equalized. This simple equilibrium concept is exposited in a highly stylized form by Pigou (1920) and more generally for urban traffic by Downs (1962). To the extent it is valid, extra expense incurred to improve design speeds on major roads has no payoff during congested periods, whereas anything to improve capacity has a huge payoff. In heavily congested urban areas, most people experience these roads under congested, not free-flow, conditions and so the need for capacity should dominate the design process.

A simple example is lane width. The standard 12-foot-wide lanes of US interstate highways provide safety margins for roads carrying mixed traffic of cars and trucks at high speeds and often under difficult conditions of weather and terrain. On most urban commuting corridors, trucks are fewer and it is practical to limit speeds to well below those of intercity travel. Indeed, urban expressway expansions are sometimes carried out by converting some shoulders to travel lanes and restriping all lanes to be narrower, sometimes to an 11-foot width. But if the road has an interstate highway designation, exemptions are required and these may be considered temporary until a fuller and more expensive reconstruction can be undertaken.

Even if pricing were in place to limit congestion, optimal speeds would be far from free-flow during much of the day. Keeler and Small (1977) analyze the tradeoff between capital expense and road capacity for congestion reduction, using construction costs and speed-flow relationships estimated for the San Francisco Bay Area. They find that in the two main central cities of that area, San Francisco and Oakland, optimal speeds during the four busiest hours would be about 50 miles per hour with congestion pricing in place. Without pricing, the optimal speed would surely be lower. In today's most congested cities, with land costs much higher than those observed by Keeler and Small, the optimal speed for an expressway is probably lower still.

Many of the aesthetic and environmental objections to urban expressways are related to their size and visibility, which are magnified by designs permitting safe travel at very high speeds. To the extent that aesthetics carry extra weight in urban areas because they affect more people, those considerations also argue for reducing the free-flow speed for which roads are designed. Even speed itself is an environmental factor due to tire and engine noise, which become so severe that large expenditures are sometimes undertaken for sound walls and extra sound insulation on nearby homes. These expenditures could be reduced by using lower-speed road design.

Samuel (2006) documents a wide variety of innovative ways that capacity can be added to urban road networks in a more environmentally and aesthetically friendly manner. These include advanced intersection designs and tunnels. Tunnels carry urban express traffic in Oslo, Sydney, and other cities and are now planned for the completion of missing links in the Long Beach Freeway near Los Angeles and in the A86 ring road around Paris. In the Paris case, where the missing section will pass under the historic palace of Versailles, the planned car-only design permits high capacity with 10-foot lanes, low clearances, and a 43 mi/hr speed limit (Samuel 2006, p. 19). Tunnels are an attractive option for many urban motor vehicle movements, but the high cost of a large cross-section makes it especially important to plan for low speeds and limited vehicle sizes.

While exotic highway designs offer promise, conventional designs already exist for carrying substantial volumes at moderate speeds, in the range of 40-50 miles per hour, while presenting a much less obtrusive public face. Lakeshore Drive in Chicago provides service at such speeds most of the time while preserving the Lake Michigan lakefront as a mark of the city's beauty. Storrow Drive in Boston similarly provides substantial capacity at moderate speeds without ruining the landscape. Neither road meets interstate highway standards, nor do they allow urbanites to cross vast developed areas in an hour. Some older parkways, such as the Arroyo Seco in Los Angeles and the Merritt Parkway in Connecticut, similarly fall below some modern interstate standards and operate in an uneasy compromise, having historic scenic value (the Arrooyo Seco Parkway is a National Scenic Byway) but being overloaded and thus under pressure for safety upgrades which make them more like conventional freeways.[8]

A few cities have moved to upgrade their major arterial streets to "superstreets," which provide a level of service closer to that of limited-access highways by means of turn lanes, traffic signal coordination, under- or over-passes at key intersections, and the like. The most important (and expensive) component is improved intersections, for which a number of innovations have been proposed including some that conserve on land consumption (Samuel 2006, pp. 48-61). An important consideration of such proposals will be to demonstrate they are safer than conventional arterials, which have substantially higher accident rates than freeways.

Many of the roads built to such intermediate standards do not accommodate large trucks. Indeed, evidence suggests that accommodating trucks adds 30 to 60 percent to the cost of an urban road.[9] For tunnels, the savings are potentially larger because a given cross-sectional area can handle two to three times as many lanes for cars as for mixed traffic (Poole and Sugimoto 1995, Fig. 1). A move to make lower-profile roads a larger portion of our high-capacity road network will undoubtedly raise objection from truckers, who like to maintain full routing flexibility. But trucks are a minority of traffic, especially during rush hours, and their needs add greatly to the aesthetic and environmental problems of roads that carry them. It simply does not make sense to build the entire network around trucks. Instead, it is better to apply the principle of differentiated products and provide some roads well suited to trucks and others well suited to handling massive peaked flows of passengers.

---

[8] The Arroyo Seco Parkway, opened in stages between 1938 and 1953, was originally considered a model of safety with its 11-foot lanes and 45 mi/hr speed limit. Renamed the Pasadena Freeway after various upgrades, it now carries nearly five times its original design volume and, with its original tight curves but a speed limit of 55 mi/hr, is one of the most accident-prone of Southern California freeways. A community task force in the 1990s spearheaded its designation as a National Scenic Byway and recommended a return to the original speed limit. See Loukaitou-Sideris and Gottlieb (2003).

[9] Meyer, Kain and Wohl (1965, pp. 204-206) present data suggesting that a cars-only expressway would cost 77 percent of a mixed-traffic expressway (as calculated by Keeler and Small 1977, p. 8). Dehnert and Prevedouros (2004) find that an underpass at an arterial intersection costs 61 percent of one that accommodates trucks, mainly due to shorter length and lower clearance.

Thus an important part of making future cities livable is to provide mobility through designs that are both aesthetically and economically sensible. An interesting byproduct would be that even when providing the same level of service, such roads would be perceived as only moderately rather than severely congested — simply because there would be a smaller gap between actual and potential speed. The most common measure of time lost in congestion is precisely this gap (Schrank and Lomax 2005); by this measure, the quickest way to reduce congestion would be to lower the speed limit on all expressways! That of course would be perceived as artificial, but road designs that carry current traffic volumes in a more aesthetic manner might indeed be perceived as providing better service (i.e. less deteriorated due to congestion) even if at the same speed as now.

Research on the comparative costs of high-capacity roads designed for different speeds and vehicle sizes is needed before any firm recommendation can be made. Equally pressing is a better understanding of their safety implications. But safety depends on many factors besides road designs, leading to a further question. If roads are designed to be safe only at moderate speeds, can we prevent people from choosing higher speeds and thereby compromising their own and others' safety? Attempts to retain the scenic character of the Arroyo Seco Parkway, mentioned earlier, have foundered partly on this problem. Rather than simply accept driver behavior as given, we should consider what public policies might be undertaken, in conjunction with road design changes, to encourage compatible driver behavior. One such policy is to introduce visual clues that cause drivers to slow down. Another is to regulate speed (or other behavior) differently from how it is done today. We now consider this latter approach in more detail.

Regulation of driver behavior

Imagine a visitor from another planet with a highly organized society possessing technologies similar to ours. Governments on this planet provide a variety of transportation services, facing similar tradeoffs between cost and quality as we do. They provide mass transportation in scheduled vehicles, and also more individualized transportation in vehicles carrying one person or a small number of people traveling together. The planet's residents undergo occasional tragic accidents, just as we do, but they affect only a small proportion of trips and are accepted as one of the costs of living, albeit one they try constantly to reduce.

You might think this visitor would feel quite at home in any of Earth's developed countries. But the visitor is struck by a disparity that seems incomprehensible. Although our rail and air transportation is carefully organized with elaborate attention to coordination among vehicles, our personal vehicles are subject to no such control other than some very rudimentary signal lights (roughly the same technology introduced on railroads in the early 1900s). Could it be, the visitor wonders, that Earth has substituted intensive driver training of all its citizens for the technological measures with which this other planet manages its many small vehicles? Inquiry reveals that no, we lack such training; indeed, the disparity of habits and methods used by our drivers is quite astounding. How could it be that with such advanced technology we have failed to apply it to reduce the rate of accidents and traffic interferences that differentiate our highways from the smoothly functioning system familiar to our visitor?

The answer presumably lies in political attitudes, social norms, and the history of our highway transportation system. We in fact have the technology to carry out many forms of driver regulation that would reduce accidents and improve traffic flow. Speed "governors" have been used on certain rental vehicles for decades. Controlled braking has become standard equipment on many vehicles, and gap control between vehicles is now offered on many models as an option to reduce rear-end collisions. Mobile communications, used for toll collection and driver information and guidance, would make it possible to activate such devices according to a centralized traffic-management plan.

Such centralized control would assuredly come at considerable cost. But not necessarily more so than current systems, such as route guidance, that are rapidly gaining in open markets. Furthermore, the cost of a centralized system of vehicle management is unknown because no one has seriously proposed it. Drivers' choice of speed, acceleration, lane movements, and expressive gestures (up to a point) have traditionally been viewed as part of freedom of movement, and controls on them would likely be derided as "social engineering".

Yet it seems quite possible that such controls would permit just the kind of more favorable tradeoffs discussed in the previous subsection, by which capacity could be increased while limiting the accompanying land consumption and aesthetic impact, not to mention the pollution and fuel consumption resulting from thousands of vehicles all trying to gain an edge on each other in the competition for road space. Specifically, it seems the main factor preventing more widespread acceptance of curved, narrow-lane highways is that they are considered unsafe. Enforceable limits on speed and lane-change maneuvers would reduce or eliminate this disadvantage while increasing the maximum throughput possible on a given pavement.

Clearly, any such proposal today would face a skeptical reaction at best. But might we see such a change in attitudes in the future, even in a nation devoted to individual freedom? I think it is possible. People accept considerable intrusion into their freedom of movement in the interests of security (airport check-in procedures), safety (drunk-driver laws), and, as already argued, optional premium service (tolled express lanes). They also accept data collection on their movements when it suits their convenience (mobile phones, credit-card records, electronic toll collection). What is needed is a demonstration of direct, perceptible benefits from accepting moderate limitations. One way to accomplish this would be to construct optional, premium-service roadways available only to vehicles equipped so as to be part of a traffic management system including speed control.

An important byproduct of such an innovation would be improvements in our record of motor-vehicle crashes. Although large gains have been made in terms of reducing crash rates (per vehicle-mile traveled), motor-vehicle accidents remain one of the largest categories of costs of driving: $0.12 per vehicle-mile by one estimate, about one-seventh of total short-run average variable cost for a typical US urban rush-hour work trip including the cost of travel time, unreliability, and vehicle capital.[10] Indeed, it appears that driving is the most dangerous activity undertaken regularly by most people. A public-relations campaign that demonstrates both safety and congestion-relief benefits from a program of voluntary limitations on driver behavior could have considerable appeal.

Public Transit

Public transit in the US plays a key role in supporting large employment clusters, serving lower-income populations, facilitating tourism, and helping selected markets. But it carries far

[10] The cost estimates are from Small and Verhoef (2007, table 3.3). The rate of motor vehicle accidents per million vehicle-miles declined from 11.7 in 1980 to 3.7 in 2004, during which time the rate of traffic fatalities (within 30 days of an accident) fell from 0.033 to 0.014: US Census Bureau (2006, Tables 1080, 1082).

fewer people than automobiles, and its share of trips has declined inexorably since the years soon after World War II. The main reasons for this decline appear to be rising incomes and the widespread decentralization of employment and residences; secondary but still important reasons include strong federal support for the Interstate Highway System, tax and housing policies that favor single-family residences, zoning restrictions on high-density housing, tax policies favoring free or highly subsidized parking at workplaces, low gasoline taxes compared to most developed nations, and little explicit pricing of highways. Kain (1999) and Small and Gómez-Ibáñez (1999) describe these factors in more detail.

A lot of money has been put into transit service in the last half century. Much of it is motivated by a desire to use transit to lure drivers off congested highways. In the US, these efforts have met with very limited success. But even where successful in diverting auto users, expansion of transit has not been the hoped-for solution to congestion. A primary reason for this is the existence of "latent demand" for peak-period road use in large and highly congested urban areas. Many people have been deterred by congestion itself from traveling when and where they would most prefer. Whenever new road space is opened up by a successful diversion to transit, some of the latent demand becomes again realized, tending to fill up the road space. The result is only small, if any, improvement in peak conditions, although there may be substantial benefits to the individuals involved.

There is little prospect that public transit will ever return to being a dominant force in urban transportation in the US. But there are several ways in which its use can be expanded and its value can be raised. These changes in turn would raise the attractiveness of cities as places to live and visit, and would help marginally to ease the pressure of highway congestion.

Specialization

The same forces of differentiation discussed in connection with highway transportation affect public transit as well. The markets served by transit are very different from each other. Probably its two most important markets are affluent suburbanites traveling to downtown business destinations and poor residents traveling within inner cities (Meyer and Gómez-Ibáñez 1981). This differentiation in markets creates a need for more differentiation of products.

That need is accentuated by the strong scale economies that characterize scheduled services in large vehicles, as demonstrated by Mohring (1972). These economies arise from two sources: savings in operator costs if vehicles become more fully utilized, and savings in user costs associated with accessing transit vehicles if service is increased. These user costs include walking or driving to transit stops, waiting for a vehicle, and making transfers between transit lines — all of which can decreased, if more vehicles are in service within a given area, by increasing the spatial density of transit lines and the frequency of service on each line.

These scale economies arise from higher rider density within a given area. As a result, standard transit service is well suited for offering frequent and densely packed service in areas with high rider density, but is poorly suited to offer service where rider density is low.

Thus product differentiation and scale economies lead to a common conclusion: transit operators should specialize. They should seek their strongest markets and pour intensive resources into them, including marketing resources and supportive political actions such as high-density zoning near large transit stations. Weaker markets should not be served at all, or should be served mostly using some model other than regularly scheduled vehicles.

Unfortunately, trends since the 1950s have been in the opposite direction (Wachs 1989, Garrett and Taylor 1999). Many transit agencies today are large, multi-jurisdictional conglomerates subject to bureaucratic and political pressures to homogenize their service. To a large extent this is a byproduct of the public takeovers of financially failing private firms following the precipitous decline in ridership in the 1950s and 1960s, as ownership of single-family homes and automobiles surged. Such takeovers were often accompanied by consolidation of several transit systems into a single large one. The result has been metropolitan-wide transit authorities which, in order to achieve the necessary political support in widely dispersed jurisdictions, have tried to offer at least rudimentary service everywhere. This is just the opposite of specialization.

Of course, abandoning low-density markets will raise strong protests, some of them pointing to inequities and harm to particular disadvantaged groups. Indeed, these markets often contain groups — including poor, physically handicapped, and elderly residents who cannot feasibly use private automobile — that society apparently deems worth supporting even at considerable cost. Finding a type of transit appropriate to such groups has proven a challenge. Demand-responsive transit, consisting of small buses or vans dispatched according to pre-arranged requests, is one way, but it has been very expensive and in the US has largely become a service solely for elderly and handicapped. New dispatching algorithms may be helping these services, and has been demonstrated even for general use in a few sites in Europe.[11] In many situations, shared-ride taxi service would be a relatively economical approach if institutional barriers can be overcome.

The point is that serving low-volume markets with standard scheduled service is very costly and not an economical way to take care of special needs. Even offering free or highly subsidized taxi service to target populations would often be less expensive. Meanwhile, by focusing service on those markets where demand is strong, transit operators can take advantage of scale economies and create the kind of frequent service and dense route coverage that can entice people to view public transit as a real alternative to private automobiles, rather than as an occasional convenience.

Role of buses

A focus on high-density markets raises the question of the appropriate type of transit vehicle. Large cities worldwide have, of course, found rail transit to be a vital part of their transportation system. But does rail transit make sense for the second tier of city sizes or, in the largest cities, for service in lower-density suburbs?

Many cost studies over the last 40 years have compared the costs of rail transit with those of other modes providing similar service (Small and Gómez-Ibáñez 1999). They have found that for nearly all situations in the US where rail transit does not already exist, buses can provide essentially the same amount of service as rail at far less cost. These arguments have been instrumental in a few cases in causing city governments to scale back ambitious rail plans, for example Honolulu in the 1970s and Houston in the 1980s. Yet these decisions tend to be reversed later (Houston opened a light rail system in 2004). Meanwhile many cities large and

[11] Computer-aided demand-responsive bus scheduling through a travel dispatch center was implemented through the European Union's SAMPLUS research program in several areas, including at least one that is open to all residents of a moderate-density urban site: Porto Romana (population density 3,600/sq km) in the Florence metropolitan area. The Porto Romana service uses flexible routes with predefined stop points and reported an operating cost of €1.69/trip, not including the dispatch center (Mageean and Nelson 2003). Other demonstration sites had much higher costs per ride, possibly because of lower density of users.

small, serving areas dense and not so dense, have opted for rail systems. There are many reasons: pressures from private interest groups such as downtown landowners, desire to enhance a city's public image, advantages of rail in terms of comfort and convenience to users, and support from federal grants.

Recent innovations in the design of bus transit have to some extent overcome each of these reasons for preferring a more expensive rail system. These innovations are loosely known as "bus rapid transit," designating a variety of initiatives giving bus transit some of the characteristics usually associated with rail. These characteristics include specialized boarding stations, off-vehicle fare payment, fixed and well publicized routes, real-time information at stations, restricted rights of way, preferential signal timing, low-emitting vehicles powered by electricity or natural gas, and marketing.

The prototype bus rapid transit system is that of Curitiba, Brazil, which carries close to 2 million passengers daily. Larger South American cities, including Sao Paulo (Brazil) and Bogotá (Columbia), have also built major systems — Sao Paulo runs 400 buses per hour during the peak and carries over 30,000 passengers per hour (International Energy Agency 2002). There are also substantial long-standing exclusive busways operating in Ottawa and Pittsburgh, and a new line in Los Angeles with grade-level intersections that opened in 2005.

Several cities in the US and Canada have recently designated new bus rapid transit lines as upgrades to existing bus routes. Early results on two Los Angeles lines showed increases in average travel speed of roughly 25 percent, to about 14 mi/hr on the busy Wilshire corridor (Levinson *et al.* 2003, Section III.B); express service (stopping only at selected locations) was subsequently introduced on Wilshire to provide a faster option. In Vancouver, three regular bus routes were upgraded to "rapid bus transit" status between 1996 and 2002, also achieving a reported overall speed of 14 mi/hr in two cases ("BRT at TRB" 2005) . The move to bus rapid transit connects also to privatization: recent federal legislation has created the "Public-Private Partnership Pilot Program," affectionately known as Penta-P, which will consider proposals for bus rapid transit from Houston and Atlanta.[12]

[12] "Federal transit gets back on the PPP track" (2007). Interest in bus rapid transit is now worldwide; see for example the thirteen papers in *Journal of Public Transportation Special Edition: BRT* (2006). As just one example, Sydney, Adelaide, and Brisbane in Australia have reported travel-time savings of 37% and higher, substantial patronage growth, and perhaps even a positive effect on adjacent property values from introducing bus rapid transit on specific corridors (Currie 2006).

What these experiments have in common is an attempt to focus resources on a small number of bus lines in order to dramatically improve their level of service. This focus mimics what happens when a new rail line is built. It is also exactly what is called for by the specialized markets and the scale economies that characterize urban transit. Bus rapid transit therefore offers a promising prospect for making public transit an economical yet vital part of urban life in an advanced highway-oriented economy like that of the US.

Subsidies

Two features of urban transit create a strong case for subsidizing it (Kerin 1992). The first is the scale economies characterizing transit costs when user costs are included, as already described. When average cost declines as a function of usage, each new user costs less than the average to the overall system (including other users), so it makes sense to encourage such users with fares set below average cost. The other feature is the underpricing of peak-hour automobile travel, which is a substitute mode for public transit. This underpricing creates an inefficient amount of traffic congestion; thus setting transit fares below cost, especially during peak periods, may be desirable to entice drivers off congested highways. Small and Verhoef (2007, section 4.5.1) provide a more formal treatment of these two arguments for transit subsidies.

Working against the case for subsidizing transit is the strong evidence that a large portion of transit subsidies has been absorbed in higher costs to transit agencies (Pickrell 1983, Lave 1991, De Borger and Kerstens 2000). Much of the higher costs have been in the form of higher wages, mainly for drivers through union negotiations. One might regard this type of higher cost as a transfer payment and so not strictly a social cost, although it is socially wasteful if it results in using higher-skilled workers than needed for the job. But to the extent that subsidy payments are funded by taxes with adverse effects on economic efficiency, even transfer payments exact a cost. Furthermore, some of the higher costs are the in the form of less efficient operations, more administrative overhead, and a bias toward a higher-than-efficient ratio of capital to other inputs in producing transit services.

Thus transit subsidies pose a dilemma: there are sound economic reasons for subsidies, yet too much of them may be wasted in unnecessary expenditures.

A possible resolution to this dilemma is offered by the tentative moves toward more pricing of highways, discussed earlier. Quantitative evidence suggests that of the two rationales for transit subsidies just described, diverting drivers from congested roads justifies much larger subsidies than scale economies (Van Dender and Proost 2004, Parry and Small 2007). If congested roads are priced at marginal cost, the rationale based on congested roads disappears. In a world with road pricing, it therefore may be possible to offer efficient transit service even while greatly reducing current transit subsidies.

Small (2004) provides an example of how a “virtuous circle” can magnify the positive effect that road pricing has on transit use, using congestion charging in central London as an example. An immediate impact of road pricing is, of course, to divert drivers to public transit as a substitute mode. This then creates a cycle of new impacts. First, the new riders provide new fare revenue for the transit operator. Second, the transit operator can now economically expand service offerings to handle the new ridership, which in turn reduces user costs for existing riders. Third, new riders are attracted by better service offerings, creating additional rounds of service improvements and diverting yet more drivers from congested roads. (This latter diversion is not necessarily a net social benefit, since now those roads are priced and so the diversion entails lost revenue; but it does support the goal of reducing road congestion.) At the same time, where transit vehicles share street space with cars, reduced congestion improves service quality and substantially reduces labor costs. Over a longer period, the changes in transit use can also encourage land-use changes in the form of transit-friendly development, thereby extending the number of people for whom transit is an attractive service.

Thus, congestion pricing on roads supports not only goals of congestion relief but also of transit finance. Furthermore, road pricing makes it economical to increase transit service; this stands in contrast to attempts to increase transit service as a mechanism to divert auto drivers, which tends to falter by inducing a realization of latent demand, as described earlier. One could say that rather than looking to transit to solve problems of highway congestion, a task to which it is inadequate, we should look to road pricing as a way of resolving the financial dilemmas of public transit! Happily, it does this as a byproduct of its main function, which is to efficiently reduce road congestion.

A more prosaic solution is to structure subsidies to minimize adverse effects on operating costs. One way to do this is “user-side subsidies”: calculate subsidies based on the number of users rather than on costs of serving them. This is equivalent to giving the subsidies in the form of payment to users, for example a stated amount per ride. The operator then has an incentive to provide service valued by users in order to attract them, and no incentive to increase costs.

Privatization

Transit operations have recently been subjected to numerous experiments aimed at using the private sector to improve performance. Here, I discuss the main categories and a few examples; Nash (2005) and Karlaftis (2007) offer more thorough reviews.

A simple form of privatization is contracting out. The public authority retains control over what services are offered, but some of those services are delegated to private firms under specific contractual terms. For example, London Transport contracted with private bus operators as part of British bus deregulation during the 1980s. This approach has also been used in Santiago, Chile (Gwilliam 2005), and in Bogotá, Columbia (Estache and Gómez-Lobo 2005). Contracting out has been used extensively in the US since the late 1970s (Frick, Taylor, and Wachs 2007), Australia since the early 19990s, and many other nations. Most such contracts have proceeded relatively smoothly, although there are some notable exceptions such as Melbourne, Australia, where financial failures and poor service have been attributed by at least one observer to poorly structured incentives in the contracts (Mees 2005).

Going somewhat further, the public authority can franchise some of these services by licensing private firms to operate them under less specific guidelines. Performance goals may be mandated or encouraged through incentives, including the prospect for favorable consideration for later renewal of the franchise. Such franchises are usually controversial; they allow more room for private initiative than contracting out, but they also are subject to contract disputes, strategic renegotiations, and abandonment of obligations by a financially failing firm.

Most examples of franchising in public transportation have been for regional rail service, notably in the UK and Sweden. However, starting in 2003 the London Underground has franchised the maintenance and upgrading of its infrastructure and vehicles to two private companies — over the strenuous objections of the Mayor of London and his Commissioner for Transport and also leading to some safety problems.[13] As with many franchising arrangements, a

[13] See "Tubes Untangled" (2004) and Small and Verhoef (2007, sect 6.3.4).

big motivation for the London Underground private initiative was to accelerate a capital improvement program that, by all accounts, was urgently needed due to years of under-funding under public operation.

Going further still, the public authority can simply turn over certain transit-service markets to private industry. The industry may be regulated like other public utilities, or it may be opened to free entry on the assumption that competition will produce a desirable result just as for other goods in a largely market economy. Just as with other businesses serving the public, regulatory oversight over such things as safety and financial disclosure can be maintained. Urban bus transit in Great Britain outside of London was deregulated in the 1980s, as was transit service in New Zealand in the early 1990s.

Because of scale economies, it is likely that service in a given area will be provided by just one or a few firms, even in a privatized system. Will the mere threat of entry cause incumbent firms to act competitively? The limited empirical evidence suggests not. For example, Evans (1988) describes the experience in Hereford, England, where transit service was deregulated beginning in 1981. There was brief period of intense competition, but then the dominant firm drove out all its rivals except in one small segment of its market. Fares ultimately returned nearly to the levels that prevailed prior to the experiment, although service levels remained substantially higher — perhaps a remnant of effectiveness of threat of entry.

Many researchers have investigated whether private or public firms operate more efficiently. The results are ambiguous (Karlaftis 2007). It appears that the key to efficiency is well-structured management incentives, and that these can be achieved under either public or private ownership at least in some circumstances. The US is somewhat behind the rest of the world in undertaking the more thorough-going types of privatization of public transit, which may be a problem because some observers have argued strenuously that US institutions cause its public transit operations to be especially inefficient (Winston and Shirley 1998). As experience accumulates, it should be possible to use private firms in appropriate ways to improve the efficiency of transit.

Conclusion: A road map for urban transportation policy

Policy makers can take advantage of the shifting terrain on which urban transportation operates through a number of steps. Taking any of them will help. Taking all of them would inaugurate a revolutionary change that would greatly improve urban life.

1. *Encourage highway pricing innovations*. In some cases, the direct benefits of a specific measure are not very large. For example, models of priced express lanes suggest that when the express lanes are kept operating at free-flow speeds, the net benefits are small compared to letting the same lanes operate as general-purpose lanes (Small and Yan 2001, Verhoef and Small 2004). But such express-lane innovations are typically improvements over the actual situation preceding them, which usually involved less capacity or express lanes restricted to carpools. Furthermore, these innovations are leading to more thorough-going proposals, including proposals to price all lanes in a corridor.

2. *Expand highway pricing to an entire corridor or area*. The same evidence just mentioned suggests that a fully priced corridor offers much greater benefits than a partially priced corridor. This remains true even if the price is set lower than optimal to meet political goals (Small, Winston and Yan 2006). Furthermore, cordon pricing of an entire downtown business district, along the lines of London and Stockholm, may be possible in a few US cities with dense downtowns, such as Manhattan, San Francisco, and Washington.

3. *Seek better tradeoffs between efficiency and public appeal for pricing schemes*. Small, Winston and Yan (2006) study a series of pricing policies for a freeway corridor with express lanes. The pure express-lane policies tend to have small net benefits, whereas optimal congestion pricing has large benefits but high prices that inhibit public support. They offer a policy that prices all lanes, at differential rates, with the lower-priced lanes at a low rate designed to strictly limit the direct costs incurred by users. This type of policy may be the compromise needed to enable pricing to extend beyond just a few selected express lanes.

4. *Encourage private participation in highways with good franchise terms*. Private highway finance has entered the US in a big way, but most of it is for conventional toll roads. If franchise terms are made flexible, bidders will find ways to use flexible pricing to everyone's advantage.

Innovative private operations will involve firms taking on demand risk, so it is important to give them enough pricing flexibility to have some control over this risk. Regulation is still needed, but with a soft touch — allowing latitude for price differentiation while regulating overall revenues (or profits) to avoid abuse of monopoly power.

5. *Pursue highway designs that emphasize high capacity at moderate speeds and with an environmentally friendly footprint*. Urban road designers should be allowed to sacrifice free-flow speed and ability to handle large vehicles in favor of high throughput of passenger vehicles. If curves or lane widths present safety hazards at high speeds, they should be encouraged to offer options in which electronic speed control on these corridors is used in return for fast and reliable service.

6. *Encourage niche transit*. Private entrepreneurs have proven very adept at finding profitable transit markets, even when it is illegal for them to do so (Cervero 1997). Public transit authorities should be forced to encourage such competition with their systems, rather than to outlaw them or to drive them out of business using predatory tactics, as happened in Los Angeles in the early 1980s (Teal and Nemer 1986).

7. *Break up large metro-wide transit providers by spinning off those serving lower-density areas*. Low-density transit service is a drag on the finances of big-city transit operators. There may be reasons to subsidize such service, but any such decision needs to be taken on its merits and not as part of a *quid pro quo* for keeping the larger operator afloat. Services for low densities and for populations with special needs should look quite different from regular transit service, typically involving small vehicles with flexible scheduling, and therefore need not be provided by the same agency.

8. *Configure federal capital-grant programs to encourage bus rapid transit*. Currently the US Department of Transportation is doing just this, but it is a political decision that can easily change. Legislators tend to like big visible projects to showcase their accomplishments, and rail has served this purpose but at great expense. Meanwhile many US bus operations are starved for

funds. Bus rapid transit offers a solution by making visible and attractive improvements at modest cost.

9. *Use open-ended user-side subsidies to improve incentives for public providers to control costs*. Transit policy is caught on the horns of a dilemma, by which needed subsidies are hijacked through union wage increases and/or operator inefficiencies. One way to lessen this tendency is to make the subsidies proportional to ridership, thereby forcing the agency to do everything it can to keep riders happy.

References

AASHTO (2005) *A Policy on Design Standards – Interstate System*, 5th Edition, Washington, D.C.: American Association of State Highway and Transportation Officials.

Brindrim, Kira (2007) "Congestion pricing OK if no fare hikes: poll," *NewYorkBusiness.com*, June 19.

Cervero, Robert (1997) *Paratransit in America: Redefining Mass Transportation*. Westport, CT: Praeger.

Currie, Graham (2006) "Bus Rapid Transit in Australasia: Performance, Lessons Learned and Futures," *Journal of Public Transportation, Special Edition: BRT*, 9(3): 1-22.

Dahlgren, Joy (1998) "High occupancy vehicle lanes: Not always more effective than general purpose lanes," *Transportation Research* 32B: 99-114.

De Borger, Bruno and Kristiaan Kerstens (2000) "The performance of bus-transit operators," in David A.Hensher and Kenneth J. Button (eds.) *Handbook of Transport Modelling,* Amsterdam and New York: Pergamon, pp. 577-595.

Dehnert, Gregory and Panos Prevedouros (2004) "Urban Intersection Congestion Reduction with Low Clearance Underpasses: Investigation and Case Study," *ITE Journal*, 74: 36-47.

Downs, Anthony (1962) "The law of peak-hour expressway congestion," *Traffic Quarterly*, 6: 393-409.

Downs, Anthony (2004) *Still Stuck in Traffic: Coping with Peak-Hour Traffic Congestion,* Washington, DC: Brookings Institution Press.

Estache, Antonio and Andrés Gómez-Lobo (2005) "Limits to competition in urban bus services in developing countries," *Transport Reviews*, 25: 139-158.

Evans, Andrew (1988) "Hereford: A case study of bus deregulation," *Journal of Transport Economics and Policy*, 22: 283-306.

Frick, Karen Trappenberg, Brian Taylor and Martin Wachs (2007) "Contracting for public transit services in the US: Evaluating the tradeoffs," in Kopp, Andreas (ed.), *Privatisation and Regulation of Urban Transit Systems*, Proceedings of Roundtable 138, European Conference of Ministers of Transport, Paris: Organisation for Economic Co-operation and Development, forthcoming.

Garrett, Mark, and Brian Taylor (1999) "Reconsidering Social Equity in Public Transit," *Berkeley Planning Journal*, 13: 6-27.

Giuliano, Genevieve (2004) "Land Use Impacts of Transportation Investments," in Hanson, Susan, and Genevieve Giuliano (eds.), *The Geography of Urban Transportation*, New York: Guilford Press, pp. 237-273.

Gramlich, Edward (1994) "Infrastructure Investment: A Review Essay," *Journal of Economic Literature*, 32: 1176-1196.

Grynbaum, Michael M. (2007) "New York pitches 'congestion pricing' to federal officials," *New York Times*, June 26.

Gwilliam, Ken (2005) "Bus franchising in developing countries: Some recent World Bank experience," Revised keynote paper, 8th International Conference on Ownership and Regulation of Land Passenger Transport, Rio de Janeiro, Brazil, June 2003. http://siteresources.worldbank.org/INTURBANTRANSPORT/Resources/bus_franch_gwilliam.pdf.

International Energy Agency (2002) *Bus Systems for the Future: Achieving Sustainable Transport Worldwide,* Paris: International Energy Agency and Organisation for Economic Co-operation and Development.

Journal of Public Transportation Special Edition: BRT (2006), 9(3).

Kain, John F. (1999) "The urban transportation problem: A reexamination and update," in Gómez-Ibáñez, José A., William B. Tye and Clifford Winston (eds.) *Essays in Transportation Economics and Policy: A Handbook in Honor of John R. Meyer*, Washington, DC: Brookings Institution Press, pp. 359-401.

Karlaftis, Matthew G. (2007) "Privatisation, regulation and competition: A thirty-year retrospective on transit efficiency," in Kopp, Andreas (ed.) *Privatisation and Regulation of Urban Transit Systems*, Proceedings of Roundtable 138, European Conference of Ministers of Transport, Paris: Organisation for Economic Co-operation and Development, forthcoming.

Keeler, Theodore E. and Kenneth A. Small (1977) "Optimal peak-load pricing, investment, and service levels on urban expressways," *Journal of Political Economy*, 85: 1-25.

Kerin, Paul D. (1992) "Efficient bus fares," *Transport Reviews*, 12: 33-47.

Lave, Charles A. (1991) "Measuring the decline in transit productivity in the U.S," *Transportation Planning and Technology*, 15: 115-124.

Levinson, Herbert *et al.* (2003) *Bus Rapid Transit*, Transit Cooperative Research Program Report 90 (Volumes 1 and 2). Washington, DC: Transportation Research Board.

Loukaitou-Sideris, Anastasia, and Robert Gottlieb (2003) "Putting Pleasure Back in the Drive: Reclaiming Urban Parkways for the 21st Century," *Access*, 22: 2-8. http://www.uctc.net/access/access.asp

Mageean, Jenny, and John D. Nelson (2003) "The evaluation of demand responsive transport services in Europe," *Journal of Transport Geography*, 11: 255-270.

McCarron, John, and Jim LaBelle (2002) *Critical Cargo: A Regional Freight Action Agenda for Jobs, Economic Growth and Quality of Life in Metropolitan Chicago*, Chicago: Metropolitan Planning Council. http://www.metroplanning.org/resources/images/criticalCargo.pdf

Mees, Paul (2005) "Privatization of Rail and Tram Services in Melbourne: What Went Wrong?" *Transport Reviews*, 25: 433-449.

Meyer, John R. and José A. Gómez-Ibáñez (1981) *Autos, Transit, and Cities,* Cambridge, MA: Harvard University Press.

Meyer, John R., John F. Kain and Martin Wohl (1965) *The Urban Transportation Problem*, Cambridge, Mass.: Harvard University Press.

Mohring, Herbert (1972) "Optimization and scale economies in urban bus transportation," *American Economic Review*, 62: 591-604.

Nash, Christopher A. (2005) "Privatization in transport," in Button, Kenneth J., and David A. Hensher (eds.) *Handbook of Transport Strategy, Policy and Institutions,* Amsterdam and Boston: Elsevier, pp. 97-113.

Parry, Ian W.H. and Kenneth A. Small (2007) "Should Urban Transit Subsidies Be Reduced?" Working paper 06-07-23, Department of Economics, University of California at Irvine, May. http://www.economics.uci.edu/docs/2006-07/Small-23.pdf

Pickrell, Don H. (1983) "Sources of rising operating deficits in urban bus transit," *Transportation Research Record*, 915: 18-24.

Pigou, Arthur C. (1920) *The Economics of Welfare,* London: Macmillan.

Poole, Robert W., Jr. (2007) "The Case for Truck-Only Toll Lanes," *Public Works Management and Policy*, 11: 244-249.

Poole, Robert W., Jr., and C. Kenneth Orski (2003) "HOT networks: A new plan for congestion relief and better transit," Policy Study 305, Los Angeles: Reason Foundation. http://www.reason.org/ps305.pdf

Poole, Robert W., Jr., and Yuzo Sugimoto (1995) "Congestion Relief Toll Tunnels," *Transportation*, 22: 327-351.

Samuel, Peter (2006) "Innovative roadway design: Making highways more likeable," Policy Study 348, Los Angeles: Reason Foundation. http://www.reason.org/ps348.pdf

Schaller, Bruce (2006) *Battling traffic: What New Yorkers think about road pricing*, Report No. 3, Center for Rethinking Development, New York: Manhattan Institute, December. http://www.manhattan-institute.org/html/rdr_03.htm

Small, Kenneth A. (2004) "Road pricing and public transport," in Santos (2004), pp. 133-158.

Small, Kenneth A. and José A. Gómez-Ibáñez (1999) "Urban transportation," in Paul Cheshire and Edwin S. Mills (eds.) *Handbook of Regional and Urban Economics, Vol. 3. Applied Urban Economics*, Amsterdam and New York: North-Holland, pp. 1937-1999.

Small, Kenneth A. and Erik T. Verhoef (2007) *The Economics of Urban Transportation*, Routledge, forthcoming.

Small, Kenneth A., Clifford Winston and Jia Yan (2006) "Differentiated road pricing, express lanes, and carpools: Exploiting heterogeneous preferences in policy design," *Brookings-Wharton Papers on Urban Affairs*, 7: 53-96.

Small, Kenneth A. and Jia Yan (2001) "The value of 'value pricing' of roads: Second-best pricing and product differentiation," *Journal of Urban Economics* 49: 310-336.

Sullivan, Edward (1998), *Evaluating the impacts of the SR 91 variable-toll express lane facility: Final report*, San Luis Obispo, California: California Polytechnic State University. http://ceenve.calpoly.edu/sullivan/sr91/sr91.htm

Sullivan, Edward (2000), *Continuation study to evaluate the impacts of the SR 91 value-priced express lanes: Final report*, San Luis Obispo, California: California Polytechnic State University. http://ceenve.calpoly.edu/sullivan/sr91/sr91.htm

Teal, Roger F. and Terry Nemer (1986) "Privatization of urban transit: The Los Angeles jitney experience," *Transportation*, 13: 5-22.

US Census Bureau (2006) *Statistical Abstract of the United States: 2007*, Washington, D.C.: US Census Bureau.

US FHWA (1997) *Highway Statistics Summary to 1995*, Washington, D.C.: Federal Highway Administration.

US FHWA (2006a) *Highway Statistics 2005*, Washington, D.C.: Federal Highway Administration. http://www.fhwa.dot.gov/policy/ohpi/hss/index.htm

US FHWA (2006b) *Status of the Nation's Highways, Bridges, and Transit: Conditions and Performance*, Washington, D.C.: Federal Highway Administration. http://www.fhwa.dot.gov/policy/2006cpr/

Van Dender, Kurt and Stef Proost (2004). "Optimal urban transport pricing in the presence of congestion, economies of density and costly public funds," working paper, Dept. of Economics, University of California at Irvine.

Van der Veer, Jan Peter (2002) "Entry deterrence and quality provision in the local bus market," *Transport Reviews* 22: 247-265.

Verhoef, Erik T. and Kenneth A. Small (2004) "Product differentiation on roads: Constrained congestion pricing with heterogeneous users," *Journal of Transport Economics and Policy* 38: 127-156.

Wachs, Martin (1989) "U.S. Transit Subsidy Policy: In Need of Reform," *Science* 244: 1545-1549.

Winston, Clifford and Chad Shirley (1998) *Alternate Route: Toward Efficient Urban Transportation,* Washington, DC: Brookings Institution Press.

"BRT at TRB!: Bus Rapid Transit presentations at the 84th Annual TRB Conference" (2005) *BRT Quarterly*, Tampa, Florida: Center for Urban Transportation Research, University of South Florida, Vol. 3.01.05: 1-4.

"Congestion Pricing Deadline" (2007), *New York Times*, Editorial, June 25.

"Federal transit gets back on PPP track" (2007) *Public Works Financing*, 213: 7, February.

" 'Price Check' Bid Upsets Cintra's Texas SH 121 Deal" (2007) *Public Works Financing*, 216: 4-6, May.

"Texas Toll Moratorium Hodgepodge" (2007) *Public Works Financing*, 216: 6-7, May.

"Tubes Untangled" (2004), *The Economist*, March 27, p. 57.