1. Introduction

'To improve the British economy, I'd tax all foreigners living abroad' (Chapman et al. 1989).

Will toll roads ever be the norm and 'free' roads a distant memory? Today's electronic toll collection technology can assess tolls from vehicles traveling at full speed on toll roads, bridges, and tunnels, permitting us to take seriously the notion that direct financing of roads might again become widespread. Transportation economists advocate tolls that vary by time of day to finance highways, mitigate congestion, and internalize the external costs of vehicle emissions and pavement damage.[1] Yet others argue that 'the prospects for widespread adoption of congestion pricing are extremely limited' because only a small political constituency (principally transportation economists and planners) favors this kind of pricing (Wachs 1994, p.15).

While the use of tolls to manage the externalities of congestion and pollution is relatively new in the realm of highways, road pricing to build and maintain infrastructure has a long history. However, most roads are still financed through gas taxes and general revenue. Future trends and policies may again shift the balance.

Implementing congestion pricing on existing toll roads is straightforward, as the adoption of time-of-day variation on roads such as the New Jersey Turnpike and many toll bridges has shown. Many different services already have prices that vary by time-of-day, including telephones (cheaper evening and weekend rates), movie theaters (the matinee show) and restaurants (the 'early bird' special). Giving discounts to travelers during the uncongested off-peak should attract much less opposition than an extra (punitive) toll on peak period travelers. Setting the right tolls, so that time-of-day pricing is as efficient as possible without being too complex, is a challenging but surmountable problem. However, establishing toll roads in the first place is difficult. This problem has two components: constructing new toll roads and converting old 'free' roads.

New or widened roads can be financed either through tolls or from more broadly based revenue sources such as gas taxes. With the completion of the interstate highway system, American localities must bear a greater share of the cost of new highways. But along with the greater financial responsibility comes increased flexibility. While the federal government generally prohibits tolls on newly constructed interstate highways,[2] no such restriction exists on locally funded roads. Besides raising otherwise unavailable funds, toll financing ties use more closely to payment. A number of new toll roads have been started in the past decade, these include two private roads in the United States as well as many in other countries.

America's current toll roads have never had to face the issue of conversion from the 'free' road model (aside from limited experimentation with high-occupancy/toll lanes, which allow toll payers voluntarily to buy into excess capacity on high-occupancy vehicle lanes, as discussed in Chapter 11). Though there is a historical precedent for the conversion of 'free' roads to the toll model, as discussed in Chapter 2, it is nevertheless a politically contentious issue. Its success depends on how government spends the new toll revenue.

A popular saying insists 'There is no free lunch'. Similarly, there is no 'free way'. The real issues are the directness of the charge and who bears it. Directness varies by means of collection. Payments may be collected by time of day, on each road segment, for every trip, at every fill-up of the gas tank, or once a year as a tax. Depending on the mechanism, different individuals may pay more or less than their fair share. Those who pay nothing for the use of the road are *free riders*. Though the ride may be free to them, it costs someone else. Although there is no free way, some may take a free ride.

States in the east of the United States continue to finance many highways with tolls, but this is rare in the west. Clearly one can point to historical and political reasons explaining this fact, but underlying the history is a set of preferences that shape each state's decision. In brief, the preference can be summed up by the folk saying 'Don't tax you, don't tax me, tax the fella behind the tree'. Local governments typically rely on a mix of revenue sources, each of which is borne by a different set of people. For instance, taxes on car rentals, hotels, and entertainment are common in tourist areas. Speed traps on major highways through small towns are another example. While many conventional taxes cannot reach non-residents – who don't have to pay local income, property, or sales taxes – road tolls can. And the proportion of non-residents using the road in the physically smaller eastern states, or tourist areas like Florida, is greater than in the west.

If a state places a toll booth near the state line (referred to as a *boundary toll*), it expects that at least one-half of the tolls will be paid by non-residents. However the proportion of tolls paid by non-residents is higher than the share of total distance they travel.

This book argues that the issue of who gets to use roads without paying the full cost is critical to understanding the choice of highway finance mechanism. In larger localities employing boundary tolls, an increasing number of trips stay inside the boundary and thus do not incur tolls. However, the larger the community under tax financing, the more trip-makers there are who are subject to taxation. A tax-based financing system, particularly in a small jurisdiction, is inequitable to local residents and may not be politically stable. On the other hand, a toll-based system is inequitable to non-residents, which does not create the same political instabilities. A similar argument applies to placing the burden of new infrastructure on existing rather than future residents.

Table 1.1 illustrates this idea: taxes, particularly property or income taxes, but even gas taxes when drivers buy their gas near home, fall disproportionately or entirely on local residents, while non-residents ride for free. Many tolls, particularly boundary tolls, fall hardest on non-residents, while residents get off easy.

Table 1.1 Incidence of revenue mechanisms on user groups

	Residents	Non-residents
Tax	Hard ride: Payment exceeds fair share	Free ride: No payment for use of road
Toll	Easy ride: Payment less than fair share	Hard ride: Payment exceeds fair share

This book explores many of the underlying reasons that localities choose to use taxes or tolls of various kinds. One hypothesis is that smaller political units have more motivation to impose tolls than large ones. First, the smaller the community, the greater the share of toll revenue from non-residents. Second, for larger regions, tolls collected at the state or county line may prove insufficient to recover costs. However, under the right circumstances, boundary tolls enable a jurisdiction to achieve the locally ideal policy of 'taxing foreigners living abroad' as suggested in the opening quote.

While tolls are common for certain expensive facilities such as tunnels and bridges, they are less common on streets and highways, which are more typically funded through user taxes or general revenue. This research identifies critical technological, economic, geographic, and political factors associated with a government jurisdiction's choice of revenue mechanism (for instance, taxes or tolls) for its network. In contrast to the large thread of research which focuses on optimal financing decisions, this book frequently uses game theory to analyse the political and economic implications of alternative revenue mechanisms and organizational structures for the road network.

This book analyses roadway network financing, constructing models that include the basic features of the economic structure of transportation networks. It examines the demand and supply interaction, the choices available to actors (consumers and producers), and the linkage between the two when the residents of a jurisdiction own the local network. The idea of decentralized, local control and multiple jurisdictions (for example, different states) and user groups distinguishes this analysis from one where a central authority (such as the federal government) tries to maximize global welfare. The model's theoretical results should be consistent with what is empirically known about network financing. It should thus explain what network financing choices are made under various circumstances. Policies that alter circumstances to affect the desired choice of revenue mechanism can then be drafted.

Four main rationales for road pricing motivate this research: financing infrastructure and relieving congestion through capacity expansion; changes in the vehicle fleet and tax base for highways; social costs; and allocating existing infrastructure more efficiently.

RATIONALES FOR ROAD PRICING

Financing Infrastructure

Historically, the primary rationale for road pricing was to finance both the capital and operating costs of infrastructure. Construction of new highways can result in cross-subsidies when financed out of general revenue, or even gas taxes. Typically, a new highway only serves a portion of the population, those using the origin–destination pairs that it connects, but is funded by a broader population. The political impact of the cross-subsidies can be reduced if a large 'package deal' is assembled (for instance, with new highways in every political district). This was the case with the interstate highway program of the 1950s to 1980s, and with many recent highway bills. However, the ability to form package deals of new roads becomes more difficult as the conventional highway system reaches maturity. There are strong arguments on both equity and efficiency grounds for users (those who benefit directly) paying for its use. Road pricing, unlike the gas tax, much less property or income taxes, ties revenue for the use of roads to the users of that specific facility.

Measuring the monetary flows into and out of the highway system is a complex task. Table 1.2 shows a simple balance sheet analysis from the 1999 highway statistics. The vast majority of highway revenue comes from

Table 1.2 Highway revenues and expenditures: 1999

Revenues and Expenditures	$ millions
Highway revenues	
Motor-fuel and vehicle taxes	76,937
Tolls	4,978
Bond issue proceeds	11,276
Intergovernmental transfers	1,762
Total revenues	94,953
Highway expenditures	
Capital outlay	59,499
Maintenance and traffic services	29,212
Administration and research	8,714
Highway law enforcement and safety	9,946
Interest on debt	4,584
Bond retirements	5,471
Net funds placed in reserves	12,473
Total expenditures	129,899
Revenues minus expenditures	(34,946)
Ratio of revenues to expenditures	0.73
Non-highway user revenues	
Property taxes and assessments	6,066
General fund appropriations	14,693
Other taxes and fees	5,519
General investment income and other receipts	6,715
Intergovernmental transfers	1,952
Net non-highway revenues	34,945
Non-highway expenditures from additional highway use taxes	
Non-highway purposes	8,873
Mass transportation	8,951
Collection expenses	3,199
Territories	212
Net non-highway expenditures	21,235
Revenues minus expenditures	13,711
Ratio of revenues to expenditures (adjusted)	0.89

Source: Federal Highway Administration (2000), Highway Statistics 1999, Table HF-10.

user fees in the form of gas taxes, licensing fees, and related vehicle taxes. Tolls represent under 5% of total revenue. Furthermore, general (non-highway-user) revenues have been limited since the early 1980s, though growing in recent years. While highway user revenues cover 73% of expenditures, highway users pay additional taxes that go to non-highway purposes, such as mass transit. Counting those payments (adding the amount to highway revenue), highway users cover 89% of their costs. This compares with less than 40% of operating costs that transit users pay, which does not include transit capital costs. Highway finance is principally a pay-as-you-go system, with a relatively small role for bonds, even for capital expenses with a long lifespan. Net bond usage (bond proceeds minus bond retirements minus interest on debt) is just $1.2 billion, compared with nearly $60 billion in capital expenditures. Funds placed in reserve (in a sense, the opposite of bonds in that they are savings for future expenses rather than borrowing for them) exceed net bond payments approximately ten-fold.

Highway infrastructure here can be divided into two basic types, conventional and advanced, though the logic for the two is largely the same. Advanced highway infrastructure can be distinguished from more conventional infrastructure by its use of intelligent transportation systems. Automated highways, while now on the back-burner, have been suggested to address roadway congestion. While it is unclear at this early date what shape advanced highway systems will take, certain forms may require complete separation of traffic into groups equipped with the appropriate technology and groups which are not so equipped. Without the appropriate financing mechanism such as road pricing, constructing this infrastructure would entail cross-subsidies from the existing (non-equipped) fleet users to those with the new technology, a transfer from poor to rich.

Change in Tax Base with Alternative Fuels

In the absence of a property rights framework and a market solution for air pollution, policy makers have focused on regulation. California and the northeastern United States have mandated that a certain fraction of the new car fleet be either low emission or zero emission vehicles (ZEVs). ZEVs are intended to reduce air pollution, or at least relocate it to the place where electricity is generated, hopefully an area where fewer people will be affected. While the precise rules are subject to change and delay (and have been altered several times to date), the trend toward switching to non-gasoline based fuels is at least nascent. To the extent that roads are financed through gasoline-based excise taxes, a shift to alternative fuels would result in a decrease in revenues for roads.

Some (albeit limited) evidence for this trend are incentives provided for ZEVs (California Air Resources Board 1997). The federal government offers a tax credit of $4,000 toward the purchase of a ZEV. Several California air-quality management districts offer 'buy-downs' to a manufacturer for each ZEV sold for a limited number of sales. Some utilities offer discounted electricity rates for off-peak electric vehicle recharging. A memorandum of agreement reached in 1996 between the California Air Resources Board and seven auto manufacturers postponed the previously announced ZEV mandates. In its place there was agreement to produce (in total) 3,750 advanced battery-powered ZEVs in 1998, 1999, and 2000 (California Air Resources Board 1996) and some measures to provide an equivalent emission reduction through production of cleaner light-duty vehicles. While this level of non-gasoline based vehicles is unlikely to cause major financing shortfalls, future targets of 10% ZEVs as in California, or even higher levels if technological breakthroughs in fuel cells manifest themselves, suggest that it may become a more significant issue in the future.

Social Costs

At the center of social cost debate is the question of whether various modes of transportation are implicitly subsidized because they generate unpriced externalities, and to what extent this biases investment and usage decisions. A proposed solution is to price travel based on the amount of externalities generated. To the extent that externalities vary in space, it may be appropriate to charge for them through road tolls rather than more general sources such as gas taxes. The main externalities include congestion, air pollution, carbon emissions, noise pollution, pavement damage, and increased accident damage. Chapter 3 discusses social costs.

Congestion Pricing

The Bureau of Transportation Statistics (1999) reports that between 1980 and 1997, highway lane-miles increased by 4%, registered motor vehicles increased by 31%, and vehicle-kilometers traveled by 67%. This was despite roughly stable journey to work times (Levinson 1998). At some point, rising congestion costs outweigh construction costs (after accounting for other social and private costs), indicating a net social benefit for the construction of new roads. Road pricing has effects which bear on this issue, including reduced demand (and improved short-run efficiency) in the absence of additional infrastructure and increased revenue for constructing new infrastructure.

Congestion reduction and more efficient allocation of resources are often cited as some of the main benefits of road pricing, particularly peak period pricing. Clearly road pricing is a necessary prerequisite to congestion (or time-differentiated) pricing. Qualitatively, the idea behind congestion pricing is that different people have different values of time. Without pricing, everyone travels slowly. But if roads are priced, individuals with a high value of time will be able to pay money and travel faster, while others will not pay the money and not travel at that time (or travel on more congested and slower alternative free roads).

To increase the welfare of travelers (or potential travelers), the money collected needs to be redistributed in some fashion, either through lowering other taxes, through direct payments, or by reinvesting it in transportation. If a driver's value of time saved plus the amount returned is greater than the amount paid, that driver is better off. If the amount of money returned to another commuter is greater than the cost of deferring the trip (or traveling at a slower speed), then the commuter is better off. Road pricing will inevitably create both winners and losers (usually losers) without redistribution of the toll revenue. However, under the right redistribution policy, most people can be made better off. Chapter 10 discusses congestion pricing and Chapter 11 considers compensation policies.

OUTLINE OF THE BOOK

Chapter 2 provides a positive explanation for the historical rise and fall of turnpikes, as well as speculation about some of the necessary conditions for a significant re-emergence of turnpikes. The historical evidence is compared with specific analytical hypotheses about the effects of jurisdiction size and trip length on the choice of financing mechanism. Road tolls, present since ancient times, were deployed widely in the eighteenth century. By the middle of the nineteenth century, most intercity land travel in Britain and the United States used turnpikes. Yet, at the onset of the twentieth century almost all tollgates had been dismantled and the turnpikes converted to publicly owned, operated, and free highways. Disturnpiking occurred simultaneously with the centralization of control over roads – management moved from small local agencies, companies, and authorities to larger regions or states. Longer-distance travel was viewed as a responsibility of a higher level of government, which saw more users as local residents. What is non-local to a county may be local to a state. By the 1940s, the desire for limited-access highways serving long-distance auto trips led to another upsurge in toll road construction by states. The 1956 Federal Aid Highway Act arrested this trend by guaranteeing federal funding for a designated interstate highway system free of tolls. Just as before, what is non-local to a state is local to the nation. Thus, the vast majority of intercity roads in the United States constructed in the interstate era are not tolled because of centralized national policy-making. Recent interest in tolls in the United States has picked up with the completion of the interstate highway program. New road financing has largely become a state and local problem again. Because of the reduction in the transaction costs to government and travelers with electronic toll collection, tolls are more widely viewed as a feasible option. New public and even private toll roads are being constructed with electronic toll collection, while existing toll plazas are being converted.

Chapter 3 reviews the theoretical and empirical literature on the cost structure of highways and specifies and estimates cost functions. It develops a full cost model which identifies the key cost components, and then estimates costs component by component. It estimates a user's costs of owning and operating a vehicle. It builds a model of long-run total infrastructure expenditures on infrastructure by states. Data on Californian bridges are used to estimate a model of manual toll collection costs. The operating costs of electronic toll collection systems throughout the United States are also modeled. Measures of each externality: noise, air pollution, accidents, and congestion are constructed.

Chapter 4 examines the question of why some states impose tolls while others rely more heavily on gas and other taxes. A model to predict the share of street and highway revenue from tolls is estimated as a function of the share of non-resident workers, the policies of neighboring states, historical factors, and population. The more non-resident workers, the greater the likelihood of tolling, after controlling for toll roads planned or constructed before the 1956 Interstate Act. Similarly, if a state exports a number of residents to work out-of-state and those neighboring states toll, it will be more likely to retaliate by imposing its own tolls than if those states do not toll. Decentralization of finance and control of the road network from the federal to the state, metropolitan and city and county levels of government will increase the incentives for the highway-managing jurisdiction to impose tolls.

Chapter 5 considers questions of financing and the hierarchy of roads and investigates the various relationships between governmental and network hierarchies. Both infrastructure networks and government are typically hierarchically organized. The hierarchy of roads separates access from movement. There are both advantages and disadvantages to managing roads with a higher jurisdiction. While larger jurisdictions may be able to exercise scale economies, they also have a larger span of control, which implies increased management costs and slower decision times.

Chapter 6 investigates the problem of financing infrastructure over time when the number of users changes. The problem is confronted in many fast-growing communities desiring to coordinate the timing of infrastructure and development, yet still achieve economies of scale where they exist. The temporal free-rider problem is defined, whereby the group that finances the construction at a given time is not identical with the group that uses it. The continuous recovery method, which effectively establishes a property rights framework for infrastructure, is described. Continuous recovery enables existing residents to be appropriately compensated by new residents, independent of the number of new residents who ultimately arrive. The system is illustrated and compared with practice in a case that uses a non-continuous cost recovery system.

Chapter 7 considers the factors affecting the choice of revenue mechanism on a beltway in a cost recovery framework. The size of jurisdictions and the length of trips on the network dictate the proportion of trips which pass through each jurisdiction and the proportion which remain entirely within that jurisdiction. The spatial free-rider problem depends on the nature of the financing system; either through trips or local trips can be free riders. Free riding distorts equity and efficiency and produces potential political problems. The fixed and variable cost of collecting tolls and taxes may favor one method over the other and influence the spacing of tolls and thus the number of free riders.

In Chapter 8, a similar but more sophisticated model is constructed to examine a long road representing an intercity highway crossing many states, such as the I-95 on the East Coast of the United States. Here, it is assumed that each locale acts to maximize the benefits to its residents from traveling plus the profits accruing to the local road, or local welfare. This measure explicitly excludes any benefits to non-residents. Each community selects a revenue instrument (such as taxes or tolls) and sets a rate of tax or toll to achieve its goal. The interaction between multiple political units and their residents complicates the analysis. Each jurisdiction's residents use both local and non-local streets, while residents and non-residents alike use its roads. The proportion of trips on a community's roads made by residents and by non-residents directly shapes the local welfare resulting from a particular revenue mechanism. This proportion depends on the size of the relevant city, county, or state. The choice between tax and toll must trade off the number of system users who don't pay their full cost because of where they live and travel, and the costs of collection. The sensitivity of travelers to tolls limits the revenue recovered. The decision of whether to impose taxes, tolls, or some combination of the two therefore depends on jurisdiction size.

Chapter 9 considers the tax toll problem in a repeated game context along a state border. Frontiers provide an opportunity for one jurisdiction to remedy inequities (and even exploit them) in highway finance by employing toll booths, and thereby ensure the highest possible share of revenue from non-residents. If one jurisdiction sets policy in a vacuum, it is clearly advantageous to impose as high a toll on non-residents as can be supported. However, the neighboring jurisdiction can set policy in response. This establishes the potential for a classical prisoners' dilemma consideration: in this case to tax (cooperate) or to toll (defect). Even if both jurisdictions would together raise as much revenue from taxes as from tolls (and perhaps more, since taxes may have lower collection costs), the equilibrium solution in game theory, under a one-shot game, is for both parties to toll. However, in the case of a repeated game, cooperation (taxes and possibly revenue sharing), which has lower collection costs, is stable.

In Chapter 10, a new graphical approach to congestion pricing is suggested to disentangle revealed demand at the given level of service (recognizing that level of service and demand are jointly determined) and underlying demand for travel at a given level of service. It is this underlying (or implicit) demand which should be used for welfare calculations, and which can suggest new approaches to differentiate the road network by level of service. This chapter then develops a disaggregate game theory approach to understanding congestion and congestion pricing, examining in depth the simplest case, that of two individuals choosing when to go, when the payoffs are interdependent.

Chapter 11 develops frameworks for evaluating the effects of financing decisions which will be considered in this study. Efficiency, our usual criterion, loosely speaking, says that no one can be made better off without worsening the condition of another. However, there are also equity considerations: while it may theoretically be possible to compensate individuals for losses from changes designed to make the system more efficient, unless that compensation is undertaken, winners and losers emerge. This chapter considers several compensation schemes associated with congestion pricing to try to achieve the efficiency goals while maintaining equity in the system.

Chapter 12 examines the deployment of electronic toll collection (ETC) and develops a model to maximize social welfare associated with the toll plaza. A payment choice model estimates the share of traffic using ETC as a function of delay, price, and a fixed cost of acquiring the in-vehicle transponder. Delay in turn depends on the relative number of ETC and manual collection lanes. Price depends on the discount given to users of the ETC lanes. The fixed cost of acquiring the transponder (not simply a monetary cost, but also the effort involved in signing up for the program) is a key factor in the model. Once a traveler acquires the transponder, the cost of choosing ETC in the future declines significantly. Welfare, which depends on the market share of ETC, includes delay and gasoline consumption incurred by travelers, costs to the toll agency, and social costs such as air pollution accruing to society. Finding the best combination of

ETC lanes and toll discount maximizes welfare. Too many ETC lanes causes excessive delay to non-equipped users. Too high a discount costs the highway agency revenue needed to operate the facility. The model is applied to California's Carquinez Bridge.

The conclusion summarizes the text and considers possible deployment scenarios leading to the widespread adoption of road pricing.

NOTES

1. See for instance: Bernstein and Muller 1993; de Palma and Lindsey 1998; Downs 1994; Dupuit 1849; Gittings 1987; Keeler and Small 1977; Mohring 1970; and Sugimoto 1994; Roth 1996; Small 1983; Small, Winston, and Evans 1989; TRB 1994; Verhoef, Nijkamp, and Rietveld 1996; Vickery 1963, 1969; Viton 1981, 1990.
2. Tolls are allowed on interstate highways in a few circumstances. Toll bridges, grand-fathered toll roads begun before the interstate system and extensions of those roads, and a few value pricing experiments permitted in the recent surface transportation bill (TEA-21) are the main exceptions.