Limits to Competition in Urban Bus Services in Developing Countries

ANTONIO ESTACHE* and ANDRÉS GÓMEZ-LOBO**

**The World Bank, Washington, DC, USA and ECARES, Université Libre de Bruxelles;*
***Department of Economics, University of Chile, Santiago, Chile*

ABSTRACT *During the past three decades, urban public transport policy has gone through several phases. From public ownership and monopoly provision, the 1980s and 1990s were characterized by a strong liberalization of the sector. This experience showed the limits of liberalization of the sector in terms of safety, prices and accountability. The paper discusses the market failures that justify this claim and the regulatory options available in this emerging new role of government. It illustrates how they are being used in practice in some countries.*

Introduction

Any review of the perception of best practice in terms of the market structure and organization of urban public transport over the last 30 years or so would reveal the popularity of a hybrid model, i.e. a model in which the public and the private sector share responsibilities for the delivery of the service. A more careful review would reveal an evolving consensus with respect to the optimal degree and form of government intervention in this hybrid model. This evolving consensus is particularly obvious in bus transport, which represents the largest mode of urban transport in the developing world since, according to the latest data published by the United Nations (2001), over 40% of all trips to work are done by buses in low- and middle-income countries. The high use of buses in these countries is a direct consequence of low private car ownership levels.

The new consensus on the organization of the bus services market has roughly emerged as follows. During most of the 1970s, public provision and self-regulation were the norm, but they ended with major fiscal difficulties and the consequences of the two major oil price shocks. Starting in the 1980s, liberalization and privatization of services became the new norm, but they ended with major safety and environmental problems in addition to some social issues resulting from tariff rebalancing in the sector. The result, toward the end of the 1980s, was a wave of policies introduced to mitigate some of the excesses of competition by restricting

Correspondence Address: Antonio Estache, The World Bank, 1818 H. Street, NW, Washington, DC 20433, USA. Email: aestache@worldbank.org

entry in the sector. Since the end of the 1990s, the state seems to be returning firmly, at least as a regulator and as a facilitator of modal integration.

The main purpose of the paper is to discuss why this last stage might be the most rational from an economic point of view, thus confirming the current intuition of many policy-makers in the sector. The general perception is that the liberalization experience has demonstrated that there are indeed limits to competition and that the industry is characterized by many market failures that require regulatory intervention.

The paper emphasizes two sets of regulatory issues. The first is related to the industrial organizational of a competitive bus market since it drives the efficiency potential of public transport services. They address the potential market failures that might plague the industry, including pollution and congestion externalities, network effects, the need to coordinate the large number of agents involved, and the interactions with other urban planning issues. The second set of regulatory issues flow from the first. Indeed, if regulatory intervention is warranted, the challenge is to intervene without compromising the efficiency benefits of competition among private providers.

The paper addresses these issues as follows. The second section relies on a presentation of the experience of Santiago, the capital of Chile, as an illustration of the typical ways in which the regulatory issues tend to emerge in practice. The third section provides a detailed review of the main sources of market failures in the bus industry that would justify a regulatory role for government in developing countries, including excessive entry, socially inefficient frequency decisions and agency issues resulting from various types of ownership structures in the sector. The fourth section presents the various possible regulatory responses to these problems. The fifth section shows how these solutions are converging into a new hybrid model and how they are being implemented in practice in the context of the recent reforms of urban transport in Bogotá, Colombia. Finally, it must be recognized that there are several preconditions that must be met for the hybrid to be successful. Otherwise, regulatory failures might more than compensate for the potential benefits of this model. Therefore, the paper ends with some comments on the risks associated with the new model of regulation in this sector.

An Illustration: Evolution of the Role of Government in the Bus Sector in Santiago, Chile

The recent history of urban public transport in Santiago, Chile, provides a good illustration of the typical evolution of urban transport policies and of the reasons behind the changes that have occurred regarding best practice in this field. This history can be divided into three distinct periods. The first period, ending in 1979, was characterized by heavy state intervention, both as a service provider (with the company Empresa de Transportes Colectivos) and as a regulator of prices, routes and permits for private operators. During this period, there was a chronic shortage and low quality of services. The social costs of this insufficiency were in the form of long waiting times for bus arrival and congested buses.

The second period began in 1979, when the sector was liberalized with the introduction of free entry, freedom to establish routes and, beginning in 1983, freedom for each operator to set tariffs. The rationality for these reforms laid in the conviction that a free market would generate an optimal level and quality of services. Competition, it was thought, would guarantee an efficient level of prices. Unfortunately, in practice, things did not turn out as expected.

During this second period, there was a significant increase in the number of buses and the geographical coverage of the system. Between 1979 and 1983, the number of buses increased by 40%, from 5185 to 7278. This increase and the problems it generated spurred a U-turn in the liberalization reforms. Between 1984 and 1988, entry was restricted, although illegal entry persisted and the number of buses reached close to 11 000 by 1988. In 1988, just before the transition to democracy, the sector was again completely liberalized and the number of buses reached its peak of 13 698 in 1990.

The effects of liberalization clearly generated benefits to users as waiting times were reduced and the distance to the nearest bus route was shortened. As the process continued, average capacity utilization dropped to 55% for buses and to 32% for the smaller 20–30-seat 'taxi' buses (Cruz, 2002). In spite of this fall and of the associated large number of operators in the system, users saw the benefits of improved service offset by a real fare increase of almost 100% between 1979 and 1990. This increase was not related to increases in fuel prices. In fact, with the exception of 1986, fuel prices throughout the period were lower than their original level in 1979. The behaviour of fares during the liberalization period implies that competition was not successful in curbing market power.

In addition to these price increases, the reforms were beginning to be associated with major quality problems. Indeed, the increase in the number of buses, their reliance on diesel fuel and the increase in the average age of buses—with lower technological standards—transformed the bus sector in one of the main generators of congestion and air pollution externalities in Santiago. By the late 1980s, Santiago's atmosphere was one of the most polluted in the world (CONAMA, 1998). The bus industry was not the only source of air pollution, but it was one of the leading contributors. Congestion was related to the fact that 80% of bus routes passed through the six main arteries of the city and clogged the main roads of the central urban area (Malbran, 2001).

Towards the end of the 1980s, the high tariffs, high average age of buses and low average capacity utilization of an oversized bus stock together with the environmental and congestion externalities served finally to put an end to the free market experience. In 1991, new regulations were introduced that ushered in a new hybrid model for the organization of the industry in Santiago. Under a new tendering system, the authorities established the route coverage of services, while tariffs were determined through the competitive bidding process (subject to periodic adjustments for changes in input costs). The competitive tendering process—although not perfect—served to stop and reverse the real tariff escalation of the 1980s, lower the number of buses and increase the average capacity utilization rate. The authorities also directly retired close to 2600 of the oldest buses in the system, paying close to US$14 million in compensation to bus owners (Cruz, 2002).

By 2001, there were only 8179 buses in operation, in spite of the fact that the average number of passengers during a working day increased from 3 575 942 to 4 275 913 between 1991 and 2001. Occupancy of buses doubled during this period. The average age of buses dropped to 6 years and over half of the current stock met EPA(Environmental Protection Agency)-91 or EPA-94 emission standards. Service quality—measured by network coverage and waiting times—was not affected by the reforms since the authorities did not modify the existing route design when tendering was first introduced. Waiting times at bus stops averaged only 4 min (Ministerio de Obras Públicas Transporte y Telecomunicaciones, 1997).

The new role of government was further fine-tuned during the 1990s when several other norms and regulations were introduced, including a limit of the maximum age of buses (10 years), emission standards, bus types (automatic transmission), among others, that helped to increase the quality of service and reduce the environmental externalities caused by the sector.

Although Chile's bus services are now firmly grounded in a hybrid regime with a strong regulatory commitment to address market failures, there are still several aspects of the system inherited from the free market experience that have not been corrected but which are being worked on. Among the remaining problems are the following.

First, when concessions were introduced, no effort was made to optimize the network configuration or to introduce some sort of tariff integration. Therefore, currently too many routes overlap, individual services are usually very long—the average length of a route is 63 km (EMG Consultores, 2002)—and tariff integration is non-existent except for some minor integration between some bus services and the underground metro system. These features generate aggregate economic inefficiencies in the transport sector, as discussed further below. Another consequence is that there is still an excessive number of buses in the system compared with an optimized network (SECTRA, 2002).

Second, since operators still earn their revenue from ticket collection, buses still have strong incentives to compete head to head on the road. Besides the difficulty, this creates for an orderly bus-stop design—buses stop anywhere generating more frequent stop-and-goes, increasing travel times and thus further undermining the economic efficiency of the transport system as a whole—safety is the greatest emerging problem. In 2001, there were 7392 accidents in the Santiago urban area where a bus was involved; of the 5699 injured people in those accidents, 112 were fatalities. On average, therefore, there is one death every 3 days in an accident involving a bus from the urban transport system. Although this involvement rate does not necessarily imply that the buses were at fault in all the accidents, if the experience of Bogotá—discussed further below—can be used as an analogy, it is highly probable that a large proportion of these accidents was due to the economic incentives that characterize the current bus system in Santiago. In addition, casual empiricism of the driving habits of bus drivers in Santiago suggests the same conclusion.

These pending issues illustrate some of the problems inherited from the liberalization experience. The weakness of regulation was large, but certainly not the only one. A related factor, possibly one of the main obstacles to further reform of the organization and regulation of the bus system, has been the pressure from the bus owner's lobby. Once the bus system had been liberalized, new strong interest groups were created by the new market conditions that made future reforms more difficult.

Market Failures in the Urban Transport Industry

More generally, what Santiago's story reveals is that market failures in the bus industry are not only the result of environmental and congestion externalities, but also the consequences of the economic characteristics of the business. The network characteristics of the industry, the peculiarities of demand for journeys, the specific organizational structure chosen for the delivery of the service and major governance issues can indeed result in many other forms of market failures. Policy-makers and the general public are typically concerned with the need to regulate and minimize congestion and environmental externalities related to public transport. Safety concerns are also present in the policy agenda. However, the failures related to the economic characteristics of the industry might also justify a regulatory intervention even in the absence of the more traditional congestion and environmental problems as reviewed below.

Lack of Curb Rights

The street curbs and bus stops where vehicles pick up their passengers are public property and thus a 'common property' resource for all transport companies, even informal ones. According to Klein *et al.* (1997), this is the main market failure in urban bus transport and explains many of the failures of this industry. Their argument is interesting, but its relevance depends on the characteristics of the demand for public transport. There are transport markets where demand depends on the prior existence of a regular service, with predetermined schedules and a high quality of service. In these markets, a company must invest in these characteristics, possibly by operating at a loss for a period, in order to induce demand for public transport. However, if a company undertakes this investment and generates the demand for the service, a competitor (formal or informal, or even private vehicles looking for passengers in order to use high-occupancy lanes in certain cities) can take away the original company's clientele at the points where passengers congregate. Since curbsides are public property and thus there is no way to avoid this 'business stealing', the company cannot recoup its original investment. Knowing this, the company does not invest in the first place, demand is not induced and the public transport service might disappear.

There are two solutions to the above problem. Klein *et al.* (1997) propose a system of transferable curb rights. Under this scheme, only companies that own rights over some curb space can pick up passengers at those points. These rights could be sold or rented, thus ensuring that the most efficient companies own these rights. Klein *et al.* suggest the use of video cameras to enforce these curb rights. The second solution is to give one operator exclusive rights over a route and then to enforce this exclusivity by keeping out potential informal operators. To avoid the abuse of market power by the exclusive operator, competitive tendering should be used to determine the franchise operator and fare levels.

Considering the country-specific conditions, the lack of curb rights might help to explain the decline of public transport in some cities in developed countries. When high quality and scheduled transport services are preyed upon by interlopers, the formal high-quality service declines, thus potential passengers, unsure of the time the next bus will pass by their stop, might prefer to use another mode of transport. As demand decreases, service supply decreases further, and a downward spiral of decline of the bus market might ensue. If the market does not disappear, it will be dominated by the generally lower-quality interlopers.

However, as a general explanation of the woes confronting bus markets in developing countries, the lack of curb rights argument is less convincing. In most developing countries, individuals might not have many alternative modes of transport available besides buses. In these markets, demand might be 'thick' in the sense that no prior investment in a regular scheduled service is required to induce demand. As recognized by Klein *et al.* (1997), the argument of a 'dissolving anchor' in public transport due to the lack of curb rights does not apply in the case of 'thick' markets. In many large cities, such as Santiago, Manila in the Philippines, Dakar in Senegal or Bangkok in Thailand, the problem is not one of disappearing public transport services but of excessive entry and frequency, especially during off-peak hours.

In the case of markets with thick demands, the lack of curb rights might still be a relevant explanation for some phenomena in developing countries. First, the curb rights argument might be relevant for the inexistence of private services in lightly travelled routes or during night or early morning hours. Second, it might explain certain distortions in the investment decision of operators that may degrade the quality of service. Since operators must compete for passengers at the curb, investments that allow operators to compete more effectively will be undertaken. One example might be the strong preference operators in Santiago have for vehicles with manual gear systems rather than for automatic gears. In 2002, the industry successfully lobbied for a change in the regulations that forced new buses to have automatic gears. The control and faster acceleration provided by manual gears might be an advantage for racing and positioning vehicles in the competition for passengers at curbs.[1] Another distortion might be the incentive for introducing faster and smaller buses that are more manoeuvrable compared with their larger and slower counterparts. Although from a private point of view the smaller buses might be preferable, the additional congestion, pollution and higher investment costs per seat capacity might not justify this decision from a social viewpoint.

Efficient Entry and Fares

Besides the lack of curb rights, there are some other reasons why a free market in urban transport might not provide the efficient number of services in developing countries. One of the main reasons might be the underestimation of the importance of fare regulation to achieve a socially efficient level of service in the sector. Fares tend to be controlled politically to ensure their affordability, but they also might need to be regulated or subsidized to induce the right entry/exit decisions and frequency decisions by private operators. Congestion, pollution and other externalities are clearly emerging as a growing problem as well for the design of the optimal organization of the bus industry. Less appreciated maybe is the importance of 'agency' problems in the design of the industry structure. To formalize the ideas of this section and to show the interactions between the various forms of market failure presented, a very simple transport model is first developed that gives basic insights into the forces behind excessive or insufficient entry and then each market failure is discussed in some detail.

A simple public transport model. Consider a transport system that operates exclusively on buses during a certain cyclical period (say 1 day, 1 week or 1 month). During this period, $N(p)$ passengers arrive uniformly to take a bus, where N is a function of the price p of one trip. Assume also that each firm owns only one bus that is operated only once per period. The marginal operating cost is assumed to be zero—again an assumption that does not change the results—but there is a fixed cost, F, per period. This fixed cost can be the capital value of the bus plus fuel and maintenance costs of undertaking the routes regardless of the number of passengers carried. If the arrival of these buses B is equally spaced, the demand for a particular bus will be N/B since it receives all passengers that arrive (uniformly) at the bus stop since the last bus passed. Under these assumptions, profits to a bus owner are:

$$\pi_i = \frac{p \cdot N(p)}{B} - F$$

When there is free entry, the private equilibrium condition (zero profit condition) is:

$$B^p = \frac{pN(p)}{F}$$

This implies that if there are B^p buses, there are no private financial incentives for additional entry and each firm obtains a normal rate of return on its investment.

However, this number of buses might not be socially optimal. For instance, a good planner would also take into account the effect of the number of buses on the average waiting time of passengers with a cost for a passenger of c per unit time. On average, a passenger will have to wait $T/2B$ time units before he/she can board a bus, where T is the time it takes a bus to complete the circuit (which is normalized to 1). Therefore, the social welfare function is given by the sum of the user and operator welfare associated with this service:

$$W = \frac{-cN(p)}{2B} + B \cdot \left(\frac{p \cdot N(p)}{B} - F \right)$$

Maximizing this function with respect to the number of buses gives:

$$\frac{\partial W}{\partial B} = \frac{cN(p)}{2B^2} - F = 0 \Rightarrow B^s = \sqrt{\frac{cN(p)}{2F}}$$

The main point of this derivation is that depending on the specific parameter values that characterize the market, the socially optimal number of buses could be smaller or larger that the private equilibrium computed above.

The ratio of the social optimal number of buses to the private equilibrium can be expressed as:

$$\frac{B^s}{B^p} = \frac{1}{\sqrt{2}} \cdot \sqrt{\frac{F}{p \cdot N(p)}} \cdot \sqrt{\frac{c}{p}}$$

The net result will depend on two effects. The ratio of fixed costs to revenues (the second term) represents the excess entry effect typical of differentiated goods models. The higher is the fare (assuming demand is inelastic) the higher are the revenues. This induces entry into the industry, which in turn might be inefficient due to excessive duplication of fixed costs. Thus, for high fares, the social optimal level of buses will be lower than the private equilibrium, independently of congestion or environmental externalities.

The last term represents the 'waiting time externality', which is well known to transport specialists. Mohring (1972) argues that private operators would supply too little service (in terms of frequency) since they do not take into account the social benefits of the reduction in waiting times of all passengers as additional buses are run on a network. If an extra bus is put into service, there will be an added private cost that the firm must recoup. However, the extra bus will decrease the interval between buses for the whole route. For example, if five buses pass uniformly each hour through a given point in the network, there will be a 12-min interval between buses. Passengers—if they arrive uniformly to take a bus at that spot—will wait on average 6 min before they can board. If an extra bus is put into service, the time interval between buses will fall to 10 min, and the average waiting time will be 5 min. Therefore, all passengers—independently of whether they actually use the new bus—will gain an extra 1 min in the form of less waiting time. This is the positive externality created by the additional bus. For a high waiting time cost parameter to fare levels, there will be an under provision of services in the private equilibrium compared with the social optimal level.

Too much or too little entry? Both effects point to an over provision of services in a private 'commercial' equilibrium when fares are high. Why would it be likely that in a competitive equilibrium, fares will be above efficient levels? The argument is familiar in the study of horizontal differentiated product markets and also applies to buses. Indeed, buses going to different destinations are not perfect substitutes. But even buses passing the same final destination of a passenger might not be perfect substitute because of the time lapse between the arrival of buses or the headway between buses. An individual might prefer taking the first bus that arrives at a stop rather than wait for the next one, even if there is a positive probability that the next bus is cheaper. Thus, the bus that just arrives at a stop has some 'captive' clients, which implies that it has some market power to raise tariffs (Fernandez and De Cea, 1990). Raising tariffs, however, creates excessive returns to investment and spurs the entry of new firms or buses beyond the socially efficient level. Evans (1987) uses a theoretical horizontal differentiation model, along the lines of Salop (1979), to show that a competitive equilibrium will imply high tariffs and excessive entry. Furthermore, some authors, such as Gómez-Ibañez (1999), have argued that in practice the 'waiting time externality' is likely to be small in quantitative terms. When buses pass at a high frequency, the gains from a reduction in the headway between buses are likely to be small.

Thus, it might be likely that the excessive entry effect may dominate, especially in developeing countries where the time costs might be low. One symptom that excessive entry has occurred is that a considerable number of buses will be running almost empty, something that is still observed in Santiago. In addition, the fact that price competition is not very strong in bus markets has been widely documented. For instance, Klein *et al.* (1997, p. 68), in a review of the British competitive experience, concluded that:

> it has been rare in the British experience for companies to compete by offering lower fares. Rather, real bus fares increased 17 percent between 1986 and 1994. Instead of lower fares, companies chose to offer more frequent service than their competitors.

There might also be another phenomenon generating an excessive number of buses in a competitive transport system when there are fixed schedules. If there is an incumbent provider or a cartel (in the case of Santiago, it could have been the operators' association), then they might have an incentive to pre-empt entry by competitors. This pre-emption strategy, as in Schmalensee (1978), implies 'filling' all the possible routes of the network in order to make entry by a small rival unprofitable and may also result in too many buses in operation.

Congestion and other externalities. Ultimately, whether there is excess supply in the bus industry is an empirical matter that needs to be analysed with reference to particular experiences and circumstances. However, there are additional influences on the optimal number of buses. In particular, if congestion and environmental costs are considered and these externalities are proportional to the number of buses on the network, then the socially optimal number of buses might still have to be lower. This implies that the possibility that a free market will result in excessive, rather than insufficient, entry is greater.

There is an interesting dichotomy here between developed and developing countries. Whereas in the former the modal split is usually heavily biased towards private cars, in the latter, public transport accounts for the majority of trips. For example, a recent report by NERA and TIS.PT (2001) reviewed the public transport system in nine cities of developed countries. It was found that public transport usually accounts for less than 25% of trips (the highest share of public transport was found in Zurich, Switzerland: 37%). In London, UK, to cite another example, public transport accounts for 34% of daily trips (Transport for London, 2003). Private car transport, on the other hand, accounts for well over 50% of trips (43% in the case of London). This is in stark contrast to developing countries, where in a city such as Santiago, 61% of trips are undertaken on public transport according to the 2001 Origin and Destination Survey.

The modal split will have an influence on the causes of congestion and pollution. In a developing country where buses are the main source of urban transport, these will have a greater responsibility for congestion and pollution. This argument is reinforced by considering that in developing countries bus technology is often older and more polluting than in developed countries. It is possible then that in developing countries, a free market might provide an excessive supply of services (due to these externalities), while in a developed country they are a solution to these same externalities. The policy recommendations must then be very different in each case. While in the latter subsidies to public transport might be justified to increase services, in developing countries, efforts might be needed to rationalize the public transport system.[2]

Agency Problems

The last source of market failure commonly encountered in the sector and that might entail regulatory intervention is associated with incentives problems built in the design of contracts given to drivers in a competitive market. Indeed, in a competitive bus market, the number of passengers using a particular bus is partly a function of the effort the driver makes to seek potential passengers on the road, stop to pick them up (usually not at a formal bus stop) and in general try to 'beat' other buses for the client. The market solution for this incentive problem is for the owner (principal) to design a contract for the driver (agent) that aligns the latter's incentives to his objectives. In this context, this means that drivers are paid on a variable salary that depends on the number of tickets sold.

Although paying drivers based on the number of passengers collected makes sense from a private point of view, from a social perspective it has several negative effects. The ensuing competition on the road between buses vying for passengers can create enormous safety problems, as described above in the case of Santiago. It also makes it very difficult to design a rational system of bus stops, since drivers will have a strong incentive to pick up passengers anywhere between stops. Frequency problems are also created by this structure as drivers use a strategy of 'head-running' (running just ahead of competitors) or wait until a competitor appears before undertaking the route.

Paying drivers a fixed rather than a variable wage is one way to eliminate the above problems. However, this will not arise naturally in a competitive bus industry, since a unilateral change of driver's contract by one company will usually entail an economic loss. Even imposing this by decree to all bus companies would be difficult. As long as owners' profits are related to the number of passengers carried, there will be a strong incentive to pay drivers, either formally or informally, based on the number of passengers picked up.

The strict enforcement of traffic laws and regulation can also be a solution to the problems noted above. However, these enforcement actions go 'against the grain' of strong economic incentives, and in developing countries, where institutions tend to be weak, they might not be effective. In these countries, it might be preferable to break the link that operators face between profits and passenger carried.

In many regulatory experiences around the world, including Bogotá (see below) and London, bus companies are paid according to quality variables or schedule completion rather than passengers carried. However, this in turn requires some system of revenue sharing between bus companies that might be difficult to achieve without heavy regulatory intervention. In addition, this policy option has the drawback that operators might have the incentive to reduce costs by lowering the number of frequencies served or not stopping at all bus stops. However, inexpensive modern geographical positioning system technology is now available to monitor the compliance with route scheduling of operators. Therefore, technological change has made it possible to pay operators according to variables other than the number of passengers carried (e.g. the compliance with route schedules and other quality measures) without a significant change in the service provided to passengers.

Revenue collection by drivers or other personnel on buses can also generate inefficiencies through a more subtle channel. When drivers or another employee on the bus must collect fares, there is the possibility that these employees will not report all revenue earned to the owner. Although this can be monitored by inspectors who check that passenger have been given tickets, fully effective policing is costly. In fact, in Santiago, owners and analyst agree that drivers supplement their incomes by around 20% through this type of fraud.

This possibility of fraud by drivers makes it harder for bus owners to delegate the operation of buses to hired employees. Thus, one observes that bus owners tend also to be drivers, or drivers are close relatives of bus owners. The monitoring problem described here is one reason why in many cities in the developing world private bus companies tend to be small, informal and family-oriented businesses.

In Santiago, the average number of buses per owner is 2.11. Over 70% of buses are owned by entrepreneurs who own five or less buses.[3]

Authors such as Nash (1988) argue that in a competitive bus market, economies of scope or scale will not be exploited because of the transaction costs involved in coordinating schedules, integrating services and fares, and agreeing on a revenue-sharing scheme between numerous private operators. This argument, however, is not fully convincing because there are many examples where a competitive market results in an efficient exploitation of these economies, specifically through mergers and vertical or horizontal integration. A good example is the airline industry, or interurban bus services, where private companies have adopted a scope-and-hub system to exploit economies of density (Brueckner and Spiller, 1994). However, the diseconomies of size that arise due to the monitoring costs of a large fleet of buses when drivers are responsible for revenue collection, and which generates an atomized ownership structure, may make it difficult to exploit potential economies of density or scale in an urban transport network. In fact, preliminary work undertaken in Santiago shows that there are large economies of density still to be exploited (SECTRA, 2002).

New Regulatory Instruments to Cut the Risks of Market Failure

The previous section implies there are four main regulatory challenges in the regulation of the bus industry that form the backbone of the new emerging hybrid model.

The first is to design and integrate the transport network to exploit economies of density and scale without compromising the system's coverage. The necessary transfer infrastructure must be built and some system of tariff integration must be introduced (unless there is one operator for all services and relevant transport modes). It will usually also be convenient from a social point of view to create exclusive bus lanes on roads. This will guarantee that the lower travel times of an optimized system are not eroded away by an increase in private automobile journeys.

Second, some regulatory control must be exercised on entry decisions and frequencies, and tariffs must be regulated.

Third, operators' revenue must be decoupled from the number of passengers carried to avoid the negative effects that result from competition for passengers on the streets. This requires some integrated revenue-collection system, independent of operators, which can then distribute these resources among firms. In addition, the authorities would have to set up some system to monitor service compliance.

Finally, bus quality and technological specifications will need to be imposed in order to reduce environmental externalities and raise service quality.

It would seem that all of the above recommendations point to a return to the old model of monopoly provision of bus services, possibly by a public operator. However, the emerging hybrid model takes into account one of the important lessons of past experience: the need to guarantee the productive efficiency of the provision of transport services. Public provision as well as subsidies is not currently favoured by policy-makers due to their negative incentives for efficiency (De Borger and Kerstens, 2000).

Naturally, direct competition and free entry into the industry is one mechanism to foster productive efficiency. However, it raises the problems identified in the third section. There are, however, alternatives to direct competition that would promote efficiency while at the same time preserving the benefits of a centrally integrated and coordinated bus industry. When entry is restricted, a regulator wanting to set fares faces an asymmetry of information problem. The companies to be regulated have better information concerning the characteristics of the routes, the actual and potential costs of serving them, the costs of providing different service qualities, as well as many other economic and technological variables that determine the efficient cost structure. Therefore, in many ways, there is a direct analogy in the promotion of efficiency in this setting and the regulation of natural monopolies in utility industries.

Combining Restrictions with Competition for the Market

Restricting head-to-head competition does not mean that the market cannot be competitive. It has been long recognized that in some circumstances, competition *for* the market is a good substitute for competition *in* the market. Tendering bus routes can be a powerful regulatory instrument to address the asymmetry of information problems that arise from the need to pick among potential providers of services.

In utility industries such as water and electricity, the use of auctions to set prices is limited by the fact that assets are sunk and long lived, thus requiring periodic tariff revisions. However, in the urban transport sector where assets are much shorter lived, it is possible to set price conditions for a similar period as the life of assets. In addition, since there is usually a secondary market for used buses, these investments are not sunk, which means that incumbents do not need to be compensated for their undepreciated investment if they lose a contract. Therefore, tendering bus routes is much more common in this sector than in other regulated utility sectors.

Besides Santiago, Bogotá and London, many other transit authorities use tendering as a way to assign operators to serve previously defined network routes. There is enough experience in the tendering of bus routes to state that this regulatory scheme is indeed quite feasible. The potential gains from tendering can be substantial. In Santiago, the tendering system reversed the tariff escalation phenomenon that characterized the deregulation period. In London, it has been estimated that tendering reduced operating costs per bus-km significantly. Cost reduction estimates found in the literature include 20% of gross (Gaister, 1997), 14% of net administrative costs (White, 1995) and 35% of operating costs (Nash, 1995). In this last case, some of the gains are associated with renegotiation of working conditions within the existing publicly owned operators, probably due to the pressure introduced by tendering.

There are several ways to tender a route contract. In the case of Santiago, contracts are tendered for 5 years based on a multivariable selection criterion, which includes the fare offered by a bidder in addition to certain quality variables. In England, two basic systems were used to tender 3-year contracts. One is to award the contract to the bidder that offers the minimum subsidy. In this case, the firm receives revenue directly from passengers and the subsidy covers the estimate revenue shortfall compared with costs. The second system is based on a gross cost basis in which bids are received for the gross amount of transfers that the firm wants to undertake the service. In this case, the company does not retain any revenues from fares and funds its operation entirely from transfers.[4] To avoid drivers pocketing fares or not passing on revenues to the central revenue collection authorities, operators under this type of contract should have no responsibility for collecting fares. Rather, an independent revenue collection agency should be set up, as in the case of Bogotá.

There is some evidence that the last alternative lowers overall costs and transfers more (White and Tough, 1995). This is because with the minimum subsidy scheme, operators still face revenue risk from fluctuations in demand. This revenue risk decreases the interest of small operators for contracts, thus lowering competition at the tendering stage. Bidders will also include a risk premium in bids to compensate for the added risk, increasing transfers compared with the gross cost type contract.

In addition, the gross cost contract also has the advantage that a company's revenue is not directly related to the number of passengers transported, thus competition on the road is avoided. However, it lowers the incentives for companies to seek passengers or otherwise stimulate demand by providing a high-quality service. However, as noted above, there is now relatively inexpensive technology to monitor bus traffic and position centrally in order to regulate service quality.

In sum, if different routes do not overlap too much, there is effective traffic control enforcement and individual drivers within a firm coordinate their scheduling, then the risks of competition on the road may be low and a system that leaves some revenue risk to firms (minimum fare or subsidy contracts) may be preferable. However, in developing countries where traffic safety is an issue, a gross cost contract that rewards a firm according to their compliance with pre-established service provisions may be preferable.

Combining Entry Restriction with Yardstick Competition

While tendering is clearly an attractive instrument, it is not always efficient or feasible. The risk that operators might pressure the authorities to renegotiate a contract once it has been awarded undermines all the incentive properties of a tendering system. In addition, there might be political obstacles to introducing tendering, especially if other types of contracts have traditionally been used and these have performed relatively well. Another possible reason why tendering might not be effective is that there might not be enough firms to guarantee sufficient competition during the process, perhaps because firms have an ability to collude. This seems to have been the case in the tendering of contracts in Santiago in 1998, when there was only one bidder for 76% of route contracts and 97% of bids coincided with the maximum allowed tariff according to the bidding documents (Sanhueza and Castro, 1999). Therefore, it is relevant to examine the performance of alternative regulatory contracts on productive efficiency of transport firms.

The ability to use yardstick competition or benchmarking is clearly an attractive alternative to tendering. It is common—especially in the transport industry—that regulators simultaneously regulate several services in contiguous spatial markets. Regulators then have a powerful tool at their disposal if they can compare or benchmark firms operating in the different markets. The use of benchmarking or yardstick competition if well applied can overcome the informational disadvantage of the regulator and in the limit can be used to reach a first best outcome (symmetric information).

The principal difficulty in applying yardstick or benchmarking-type regulations is that firms might not be directly comparable. The regulator must then strip-out the variation of costs across companies that are due to legitimate differences among companies (in the urban transport setting, this could be different route types, congestion levels, peak demand characteristics and other exogenous influences on costs) before comparing or benchmarking them. Naturally, companies have incentives to convince regulators that they are unique and thus 'non-comparable' with other companies.

In spite of the above difficulties, there is at least one experience of the use of yardstick-type competition in the bus industry. In the late 1980s, some counties in Norway started adopting a standard-cost model to determine annual transfers. In such a system, the county and the companies agree upon a set of criteria for calculating costs of operating a bus network. It is a linear model that links driver costs, fuel costs and maintenance costs to the number of bus-km produced for different categories of routes (from inner-city, low-speed to long-distance, high-speed routes). Given fares and route schedules, the standard-cost model determines the level of transfers granted by the regulator. Once the parameters of the model are set, realized costs by one company that happen to deviate from the standard-cost figures will not affect the level of its next annual lump-sum transfer. This gives the standard-cost model a flavour of yardstick competition (Shleifer, 1985). The main characteristic of yardstick competition is that transfers be based on a benchmark estimated based on cost performance of a larger set of companies. Dalen and Gómez-Lobo (2002) apply an econometric stochastic cost frontier approach to an 11-year panel of Norwegian bus company data and show that the yardstick-type contract (standard-cost model) increases the cost efficiency of firms. In addition, there are dynamic effects to the use of yardstick contracts. Costs are lower the longer the contract has been in effect. However, the measured impacts tend to be small.

Offering Menus of Contracts to the Operators

There might be occasions when neither tendering nor yardstick competition is feasible. In this case, the problem faced by the regulator is identical to a natural monopoly situation where the regulator must regulate a single natural monopoly under asymmetric information. The new theory of regulation, as in Laffont and Tirole (1993), offers a normative framework for regulatory policies in such contexts. What are the implications of the theory for policy-makers? The first is that offering just one type of contract is not optimal. Regulators should try to offer a menu of options and allow firms to self-select according to their private information. This has seldom been the case in practice. There are very few experiences where regulators formally offer a menu of options to firms. However, it could be that informally, during a negotiation stage, the regulator offers a menu of such contracts.

Note that several authors have tried to apply the theoretical results of the optimal second-best contract under asymmetric information to the urban transport sector. From a theoretical point of view, there is the work of Pedersen (1994). On the other hand, Wunsch (1994) actually attempts to derive the optimal menu of contracts for transit firms through a mix of econometrics and calibration. Wunsch uses data on 177 mass transit firms in Europe to estimate the distribution of the asymmetry of information cost parameters of firms. "[T]he asymmetry of information between the regulator and the agent is assumed to be limited to the unexplained variance of a cost estimation based on a cross-section of 177 transit firms"

(Wunsch, 1994, p.2). Therefore, Wunsch estimates a cross-section cost function conditioned on the characteristics of each transit system and obtains that the confidence interval around fitted values has a standard error of about 15% of costs. Wunsch uses this information, plus some calibrated parameters for other functions, to calculate the optimal menu of contracts to offer firms.

Gagnepain and Ivaldi (2002), in a study of the public transit system in France, use a structural approach (previously explored by Dalen and Gómez-Lobo, 1995) to recover a firm's underlying cost efficiency distribution. They then model the effects of the introduction of an optimal regulatory contract, including route tendering.

Towards a Solution in the Real World: The Experience of Bogotá, Colombia[5]

The case of Bogotá, the capital of Colombia, illustrates how with political will and a well-structured project, it is possible to improve radically the transport system in a short period. The reforms in Bogotá took 36 months to implement.

During the 1990s, Bogotá struggled with many of the problems alluded to above. Buses competed vigorously for passengers in the streets (called the 'the war of the penny' by Colombians) generating unnecessary traffic risks and a chaotic system of stops and goes. High congestion generated velocities as low as 10 km/h during peak periods. Even short trips would sometimes take considerable time. In 1998, it was estimated that the average trip took 70 min. The average age of buses was 14 years in 1998 and the service quality they could offer was low. The average occupancy rate was 45%. It has been estimated that 70% of particulate matter emissions from mobile sources could be attributed to the bus system.

Towards the end of the last decade, the Mayor of Bogotá undertook a radical reform of the transport system in the centre of the city. The reform was organized around what came to be called the TransMilenio Project that became operational in December 2000, only 2 years after it was first proposed. This project is based on a system of exclusive bus lanes—along the busiest corridors of the city—to be used by bus operators. Private concessionaires, chosen through a competitive tendering process, operate these central corridors. An extensive network of feeder routes, also operated by private concessionaires, complement these corridors.

To date, three exclusive corridors totalling 35 km have been built. In addition, 22 feeder routes are in operation with a total extension of 66.7 km covering over 40 neighbourhoods of Bogotá. Infrastructure investments during this first stage totalled US$213 million. The plan is to build up the system—over 15 years—to 22 central corridors (for a total extension of 388 km) and with the capacity to transport 5 million people daily. The total cost of the reform is estimated to be US$1970 million.

Institutionally, the project is organized around a publicly owned company, TransMilenio S.A., that designs the network, writes the contracts later tendered to private operators and administers the system. It is responsible for tendering the operation of the central corridors, the feeder routes, and the ticketing and payments system. Tickets are based on a system of prepaid cards, also administered by a private concessionaire. The lanes, bus stops, terminals, pedestrian bridges and transfer stations were built and owned by the public sector.

Among the most salient features of the TransMilenio Project is a network design with enclosed bus stops every 800 m, with pedestrian bridges and other services. Modern vehicles especially designed for passenger service were introduced through the conditions stipulated in operators' contracts. There are now 411 large articulated buses with automatic transmission systems, wheel suspension and modern natural gas or diesel motors operating in the central corridors, in addition to 147 standard buses on the feeder routes. A dual system of regular services (buses that alight at every stop with a frequency of 3 min) and express services (stopping only at a subset of destinations and with a frequency of 4 min) was introduced. Schedules and routes are monitored by an electronic surveillance system based on a Satellite Positioning System and controlled by a specially created traffic control agency.

Perhaps one of the most important changes introduced through the reform was the compensation regime for operators. Operators are now paid according to the number of kilometres travelled and the quality of service. This facilitated a radical change in driver's incentives; they are now under contract, work regular shifts and are not paid a bonus for passengers transported. The separation of the operation of buses and the collection of revenues was made possible by the introduction of the prepaid ticketing system operated by the special revenue collection company. This eliminated overnight the 'war of the penny', radically improving traffic safety and the quality of service. The electronic prepaid card system allows for tariff integration throughout the network.

In 2001, a year of operation, the evaluation of the experience has been very positive. Average velocity in the main corridors rose from 12.0 and 18.0 km/h (Calle 80 and Avenida Caracas, respectively) to 26.7 km/h after the project was in operation. Consequently, average travel times fell by 32%, equivalent to a 1-h saving daily for the average passenger.

Safety and service quality have improved dramatically. Figure 1 shows the number of accidents, injuries and fatalities on the roads corresponding to the TransMilenio network in 1999 — before the reform — and in 2001 after a year in operation. Note the significant reduction in the number of accidents, people injured or killed, and assaults.

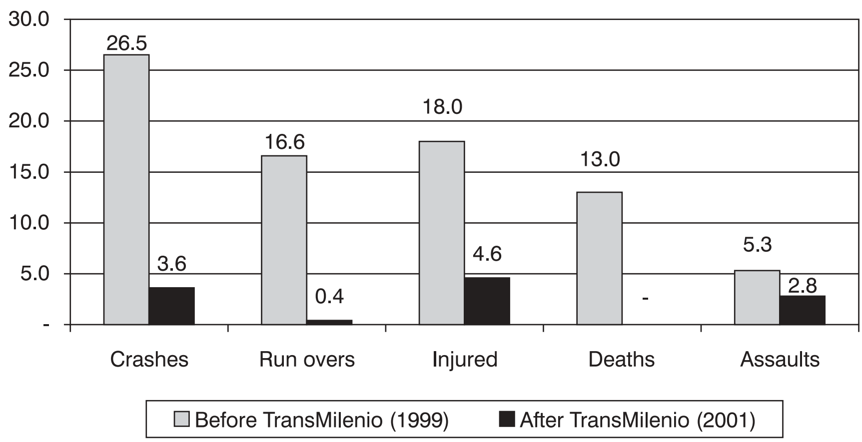

Figure 1. Weekly traffic accidents, injuries, mortality and assaults in the central bus network in Bogotá, Colombia. *Source:* Hidalgo (2001).

Pollution levels also dropped since the new system came into operation. Sulphur dioxide average daily concentration levels dropped by 43% when comparing January and March 2000 with the same months in 2001. Maximum daily concentration levels dropped by 54% (January) and 39% (March) when comparing both years. Nitrogen dioxide average daily concentrations dropped by 13% (January) and 41% (March), while maximum daily concentrations fell by 10 and 46%, respectively. For particulate matter smaller than 10 μm, the corresponding fall in average daily concentration levels was 31% (January) and 17% (March), and 16 and 13% in peak daily concentrations. Although it is possible that other phenomena (particularly climate differences) might account for these measured improvements, it is likely that the introduction of the TransMilenio plan was also responsible.

As for the funding side of the reform, tariffs increased only 6% and most of the infrastructure investment was funded through petrol taxes, multilateral loans and other domestic sources. Note that the TransMilenio system currently accounts for only around 15% of bus trips in Bogotá. The vast majority of bus trips are still undertaken in the chaotic bus system outside TransMileno's network.

Conclusions

The importance of correcting the 'developed economy bias' typically found in the literature on urban transport has been shown. The problems are indeed different in the two country groups. Direct competition in urban bus markets is rare in the developed world. The share of people using public transport, particularly buses, in urban areas decreases with development. According to the UN-habitat, 40.19% of trips to work are done by bus in the lowest income countries versus only 18.15% in Organization for Economic Co-operation and Development (OECD) countries. Finally, the ranking of the main policy issues relating to investment, network coverage, and the affordability of tariffs and other social considerations tend to be very different between developing and developed economies.

Considering these differences, the paper shows that under an emerging new hybrid model, the public authorities should consider more systematically a clear role in the determination of the network structure, service quality and frequencies, and force a separation between revenue collection and operating activities. However, unlike the traditional regulatory model, it is important for this emerging new model to continue recognizing the importance of efficiency concerns and of the role that private operators and modern regulatory instruments can have in meeting this objective.

Although this hybrid scheme is a way to combine the benefits of a public monopoly with the benefits of increased competition allowed by private provision, there are several risks and prerequisites in developing such a transport system. First, the scheme requires the planning authority to be capable of defining the network configuration and service levels adequately. Otherwise, shortages might appear and the supply of transport services will not be able to meet demand levels or expected quality characteristics. Therefore, some institutional capacity is required to define these variables. In addition, it is important to leave some flexibility mechanism in place in order to change route design, or other service quality levels, after a particular contract has been awarded to a private operator. The required institutional capacity should also include the ability to manage the tendering system and to monitor the contracts afterwards.

Second, by restricting entry into the market, direct competition is loosened. Thus, companies—although private—might not have sufficient incentives to increase productivity and control costs. The tendering of contracts avoids the problem provided agents do not perceive that renegotiation of contracts is possible. Tendering in general is to be preferred to negotiated contracts in this case both in terms of efficiency and of equity outcomes. Additional safeguards can be provided through several regulatory instruments, including yardstick competition and offering a menu of contracts to potential operators.

It is important to conclude by warning that reformers should never discount the risks of regulatory failure as a limitation to this hybrid model. If the middle road is not feasible, which of the two extreme cases (public monopoly or unregulated private provision) should be preferred? This will probably depend on a case-by-case basis, but the experience outside London and in Santiago would tend to indicate that a competitive market is probably better for users. Insufficient institutional capacity implies that the efficient operation of a public monopoly will also probably not be feasible. For small urban areas where there are not many economies of density and services are not complex, the competitive model might, however, require the introduction of selective public subsidies for the operation of unprofitable routes to maintain tariffs at levels consistent with the ability to pay of its users.

Acknowledgements

The authors are particularly grateful to José Gómez-Ibañez for very extensive comments on an earlier version of the paper and to Andrés Velasco for several productive discussions. They also thank Jose-Luis Irigoyen, Emile Quinet, Tomas Serebrisky and two anonymous referees.

Notes

1. Another motivation for preferring manual gears might be that they consume less fuel. However, there is no evidence for either hypothesis, although one referee stated that this had also been suggested in the British case.
2. In the case of congestion, road pricing might be an alternative policy option that does not require the authorities to intervene in the design and management of the public transport system.
3. Data provided by the Ministry of Public Works, Transport and Telecommunications. Liability issues might also be responsible for this structure. In Chile, bus owners are legally liable only up to the value of their property. This generates incentives to atomize the property of buses. It is common for an entrepreneur formally to spread ownership of his buses among family members.
4. However, one referee has pointed out that in London, while drivers collect fares on gross contract routes, the proportion of on-bus fare collection is now zero in the central (yellow) zone and only around 0–2% elsewhere.
5. The sources for this case study are http://www.idu.gov.co, http://www.transmilenio.gov.co and Hidalgo (2001).

References

Brueckner, J. K. and Spiller, P. T. (1994) Economies of traffic density in the deregulated airline industry, *Journal of Law and Economics*, 37,(2), pp. 379–415.

CONAMA (1998) *Plan de Prevención y Descontaminación Atmosférica de la Región Metropolitana, 1997* (Santiago: Comisión Nacional del Medio Ambiente).

Cruz, C. (2002) *Transporte urbano para un Nuevo Santiago* (Santiago: Ministerio de Obras Públicas, Transporte y Telecomunicaciones).

Dalen, D. M. and Gómez-Lobo, A. (1995) *Regulation and Incentive Contracts: An Empirical Investigation of the Norwegian Bus Transport Industry.* Working Paper W96/8 (London: Institute for Fiscal Studies).

Dalen, D. M. and Gómez-Lobo, A. (2002) Yardsticks on the road: an analysis of the Norwegian bus industry, *Transportation,* 30(4), pp. 371–386.

De Borger, B. and Kerstens, K. (2000) The performance of bus transit operators, in: D. A. Hensher and K. J. Button (Eds) *Handbook of Transport Modelling,* ch. 36, pp. 577–595 (New York: Pergamon).

EMG Consultores (2002) *Análisis de la Estructura Tributaria del Transporte Público y de Carga.* Informe Final (Santiago: Ministerio de Obras Públicas, Transporte y Telecomunicaciones).

Evans, A. (1987) A theoretical comparison of competition with other economic regimes for bus services, *Journal of Transport Economics and Policy,* 21,(1), pp. 7–36.

Fernandez, J. E. and De Cea, J. (1990) Caracteristicas de equilibrio en un sistema de transporte publico urbano bajo condiciones de libre competencia, *Revista de Ingeniería de Sistemas,* 7(2), pp. 27–43.

Gagnepain, P. and Ivaldi, M. (2002) Incentive regulatory policies: the case of public transit systems in France, *RAND Journal of Economics,* 33(4), pp. 605–629.

Glaister, S. (1997) Deregulation and privatization: British experience, in: G. de Rus and C. Nash (Eds) *Recent Developments in Transport Economics,* pp. 135–197 (Aldershot: Ashgate).

Gómez-Ibañez, J. A. (1999) Pricing, in: J. A. Gómez-Ibañez, W. B. Tye and C. Winston (Eds) *Essays in Transportation Economics and Policy: A Handbook in Honor of John R. Meyer,* pp. 99–136. (Washington, DC: Brookings Institution Press).

Gómez-Ibañez, J. A. and Meyer, J. R. (1997) Alternatives for urban bus services: an international perspective on the British Reform, *Transport Reviews,* 17(1), pp. 17–29.

Hidalgo, G. D. (2001) *TransMilenio: el Sistema de Transporte Masivo en Bogotá.* Mimeo (TransMilenio S.A.).

Klein, D., Moore, A. T. and Reja, B. (1997) *Curb Rights: A Foundation for Free Enterprise in Urban Transit* (Washington, DC: Brookings Institution Press).

Laffont, J. J. and Tirole, J. (1993) *A Theory of Incentives in Procurement and Regulation* (Cambridge, MA: MIT Press).

Malbran, H. (2001) *El programa de medidas inmediatas y la política de prioridad del transporte público* (Santiago: SECTRA).

Ministerio de Obras Públicas Transporte y Telecomunicaciones (1997) *Estudio de Demanda del Sistema de Transporte Público de Superficie de Santiago 1997* (Santiago: Secretaría Regional Ministerial de la Región Metropolitana).

Mohring, H. (1972) Optimization and scale economies in urban bus transportation, *American Economic Review,* 62(4), pp. 591–604.

Nash, J. (1988) Integration of public transport: an economic assessment, in: J. S. Dodgson and N. Topham (Eds) *Bus Deregulation and Privatization: An International Perspective,* pp. 97–118. (Aldershot: Avebury).

NERA and TIS.PT (2001) *Models for the Provision, Regulation and Integration of Public Transport Services.* Final Report (London: Public Transport Partnership Forum of Ireland).

Pedersen, P. A. (1994) Regulating a transport company with private information about costs, *Journal of Transport Economics and Policy,* 28, pp. 307–318.

Salop, S. C. (1979) Monopolistic competition with outside goods, *Bell Journal of Economics,* 10(1), pp. 141–156.

Sanhueza, R. and Castro, R. (1999) Conduciendo el Transporte Público: la licitación de recorridos en Santiago, *Perspectivas en Política, Economía y Gestión,* 3(1).

Schmalensee, R. (1978) Entry deterrence in the ready-to-eat breakfast cereal industry, *Bell Journal of Economics,* 9,(2), pp. 305–327.

SECTRA (2002) *Licitación de Transporte Público 2003: Una propuesta conceptual de Sectra* (Santiago: Secretaria Interministerial de Planificación de Transporte).

Shleifer, A. (1985) A theory of yardstick competition, *RAND Journal of Economics,* 16,(3), pp. 319–327.

Transport for London (2003) *London Travel Report 2003* (London: TFL).

United Nations (2001) *Cities in a Globalizing World — Global Report on Human Settlements — 2001* (New York: UN Centre for Human Development).

White, P. and Tough, S. (1995) Alternative tendering systems and deregulation in Britain, *Journal of Transport Economics and Policy*, 29,(3), pp. 275–289.
White, P. R. (1995) Deregulation of local bus services in Great Britain: an introductory review, *Transport Reviews*, 15, pp. 185–209.
Wunsch, P. (1994) *Estimating Menus of Linear Contracts for Mass Transit Firms.* Mimeo (CORE).