**DRAFT**
**COMMENTS WELCOME**
**March 2003**

# The limits to competition in urban bus services in developing countries[1]

**Antonio Estache**
**World Bank Institute and**
**ECARES, Universite Libre de Bruxelles**

**and**

**Andrés Gómez-Lobo**
**Department of Economics**
**University of Chile**

**Abstract**

This paper makes the case for the return of regulation in the organization of urban bus services in developing countries. During the past three decades urban public transport policy has transversed several phases. From public ownership and monopoly provision, the eighties and nineties was characterized by liberalization of the sector. The experience of several countries, in particular Chile, indicates that liberalizing the sector may not be the welfare maximizing option. This paper discusses the market failures that justify this claim and presents the regulatory options available in this emerging new role of government. Throughout the paper we illustrate ideas with examples from Chile, Colombia and a few other countries.

Keywords: transport, efficiency, regulation

JEL Classification:

[1] The usual disclaimers apply.

## 1. Introduction

Gomez-Ibañez and Meyer (1997) identify three main options for providing and regulating urban bus services: public monopoly, unregulated private firms and a hybrid model in terms of the relative role of the public and the private sector: regulated private firms. 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 adoption of the hybrid model.[2] A more careful review of the evolution would reveal an evolving consensus with respect to the optimal degree and form of government intervention in this hybrid model.[3] This evolving consensus has roughly emerged as follows.[4] During most of the 70s, public provision and self-regulation were the norm but 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 ended with a number of major safety and environmental problems in addition to some social issues resulting from tariff rebalancing in the sector. The result, towards the end of the 1980s, was a wave of policies introduced to mitigate some of the excesses of competition by restricting entry in the sector. Since the end of the 1990s, the state seems to be returning to the sector, at least as a regulator and also as a facilitator of modal integration.[5]

The main purpose of this paper is to discuss why this last stage may be the most rational from an economic point of view, thus confirming the current intuition of many policymakers in the field that there are indeed limits to competition in urban transport. The liberalization experience of the last 20 years has clearly demonstrated that the industry is characterized by many market failures and as such urban transport sector policymakers face

[2] See for instance, World Bank (2002) and NERA and TIS.PT (2001).

[3] In this paper when we refer to public transport we are thinking mainly of bus services. Although public transport in an urban area is also serviced by metro and light rail, it is fair to say that the most pressing regulatory issues are related to bus services, at least in developing countries. However, we will touch upon the relation between bus and other forms of public transport when discussing intermodal and tariff integration.

[4] See Gwilliam, Nash and Mackie (1985a; 1985b). Gómez-Ibañez and Meyer (1997), although not refuting the above consensus, argue that the empirical case against deregulation is not so strong.

[5] Although environmental regulations in this sector are important they will not be discusses here. They usually take the form of noise and emission limits on vehicles, together with age restrictions or other forms of technological restrictions.

many of the problems common to public utility regulation in general. Just like the regulation of natural monopoly network industries such as sanitation, electricity and telecommunications, the urban transport sector also faces particular problems that considerably complicate the regulatory task. These include pollution and congestion externalities, network effects, the need to coordinate the large number of agents involved and the interactions with other urban planning issues. Perhaps the only factor which makes urban transport regulation comparatively easier is the possibility of organizing competition for routes and the existence of competition from other modes of transport such as private cars and rail. However, even this adds an additional complication to the regulation of this sector: the need to oversee potential anti trust issues.

In this paper we provide an overview of some of the issues involved and present some practical experiences that could be considered as best practice, as well as some other less successful but equally illuminating cases. We believe the discussion in this paper is important given that most of the academic literature in transport has a developed country bias, were the problems and issues are different from those of the developing world.[6] In developed countries private transport is prevalent and one of the main policy preoccupations has been whether it is convenient or not to promote public transport (for environmental and congestion reasons) and how to do it, including the granting of subsidies. Given that subsidized bus services are ubiquitous in the developed world, and public provision common, a second concern has been the promotion of productive efficiency. Direct competition in urban bus markets is rare in the developed world[7].

In contrast, in developing countries most people use public transport, particularly buses, in urban areas. Subsidies are much less common and the main policy issues relate to investment, network coverage and affordability of tariffs and other social considerations. Full-scale competition in large urban bus markets has only applied in developing countries

---

[6] For a good compilation of the literature see Oum et-al. (1997).

[7] There are some expereriences in urban areas outside London in the United Kingdom, for example Birmingham, Manchester, Liverpool and other three cities (Gómez-Ibáñez and Meyer, 1997).

and the benefits and costs of these experiences, together with their regulatory implications, have not been well documented in the literature.

In this paper the emphasis is placed on two sets of regulatory issues. First, those related to the organizational and network aspects of the system and which determine the overall effectiveness and efficiency potential of the public transport services. These have to do with economies of density or scale, intermodal integration, insufficient or excessive entry, principal agent problems between owners and drivers and some other issues related to the industrial organization of this industry. Second, there are issues related to the productive efficiency of individual transport services. In other words, for a fixed network configuration, how can regulators promote least cost services?

The paper is organized as follows. Section 2 discusses the experience of Santiago, capital of Chile, as a typical example of how the new regulatory consensus seems to be emerging. Santiago's experience is also interesting because it was a pioneer in bus market liberalization, with radical and long lasting reforms implemented during the eighties. Thus it provides a clear example of the problems that can arise in a competitive bus market. Section 3 reviews the main sources of market failures in the bus industry that would justify a regulatory role for government. It covers economies of density and network efficiency. It also offers a discussion of the various dimensions of efficient entry and frequency decisions, including the fact that waiting time represents an externality, that there are economic costs associated with the excessive duplication of fixed costs and the consequences of congestion and environmental externalities. This section concludes with a discussion of agency issues resulting from various types of ownership structures in the sector and the problems caused by linking operator's revenue with the number of passengers transported. Section 4 presents the various possible regulatory responses to these problems. Section 5 presents the recent experience of Bogotá, Colombia, where successful reforms were implanted along the lines suggested in Section 5. The last section concludes.

## 2. An illustration: evolution of the role of government in the bus sector in Santiago

The recent history of urban public transport in Santiago, Chile, provides an extremely effective illustration of the various visions regarding urban transport policies and 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, ending in 1979, was characterized by heavy state intervention, both as a service provider (with the company Empresa de Transportes Colectivos) and as 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 5,185 to 7,278 (see Table 1). The increase in the number of buses 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. However, as the process continued average capacity utilization dropped significantly. Average occupancy in buses and taxi buses dropped 55% and 32% during this period, respectively.[8] In spite of this drop in capacity utilization and the large number of operators in the system, real tariffs increased by almost 100% between 1979 and 1990 (see Table 1)[9]. This increase was not related to increases in fuel prices. Figure 1 shows that, with the exception of 1986, domestic fuel prices were lower than its original level in 1979.

**Table 1. Prices and number of buses in Santiago**

| Year | Average tariff (Ch$ of 2001) | Number of buses |
|---|---|---|
| 1979 | 136 | 5,185 |
| 1980 | 133 | 6,043 |
| 1981 | 135 | 6,081 |
| 1982 | 164 | 6,579 |
| 1983 | 194 | 7,278 |
| 1984 | 223 | 8,240 |
| 1985 | 261 | 8,653 |
| 1986 | 257 | 8,904 |
| 1987 | 240 | 10,680 |
| 1988 | 233 | 10,900 |
| 1989 | 237 | 11,841 |
| 1990 | 267 | 12,698 |
| 1991 | 236 | 13,353 |
| 1992 | 203 | 11,891 |
| 1993 | 184 | 11,034 |
| 1994 | 179 | 11,562 |
| 1995 | 182 | 10,228 |
| 1996 | 186 | 9,255 |
| 1997 | 181 | 8,711 |
| 1998 | 186 | 8,623 |
| 1999 | 204 | 8,407 |
| 2000 | 253 | 8,272 |
| 2001 | 281 | 8,179 |

*Sources:* Tariffs: Instituto Nacional de Estadísticas. Buses: 1997-2001: Ministerio de Obras Públicas, Telecomunicaciones y Transporte; 1990-1997: Sanhueza and Castro (1999); 1979-1989: Cruz (2002).

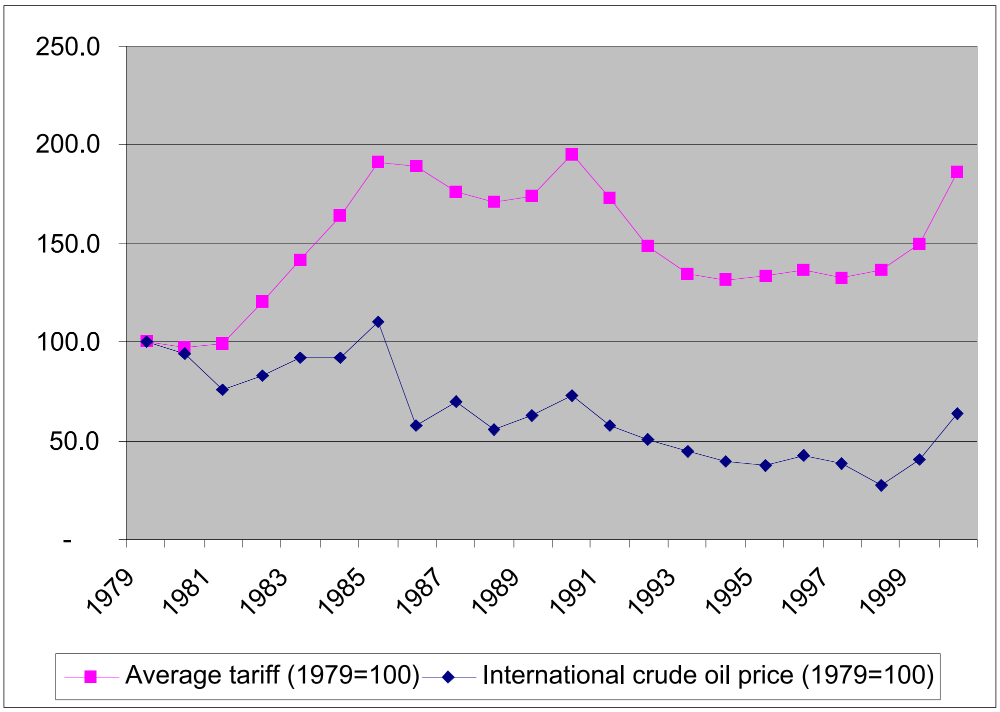
---

[8] Cruz (2002). Taxi buses were smaller buses of 20 to 30 seats compared to the larger 40 seat buses.

[9] The flip coin of this statement is that because tariffs were increasing, there were strong incentives to enter the sector in spite of the fact that low capacity utilization implied operating on a point high on the average cost curve. Therefore the true economic inference is that high prices induced excessive entry. However, once there was excess capacity in the sector it is curious that this did not induce price wars among operators. Possible reasons for this behavior will be discussed in the next section.

The observed behavior of tariffs during the liberalization period is very perplexing, implying that competition was not successful in curbing market power. That this may be a generic structural phenomenon in bus industries across the world is an issue we discuss further below.

**Figure 1: The evolution of tariffs and fuel prices**




Note: International crude oil prices were transformed to domestic currency using yearly average exchange rates. During this period international fuel price changes were passed directly to domestic prices. Therefore, the international price of oil is a good indicator of domestic fuel prices during those years.

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 eighties, Santiago's atmosphere was one of the most polluted in the world.[10] The bus industry was not the only source of air pollution but it was one of the leading contributors.

[10] Together with Mexico City, Santiago was the most polluted city in the region. In 1990, there were 95 days with particulate matter concentration levels above 150 _g/m$^3$, 30 days above 240 _g/m$^3$, and 6 days above 330 _g/m$^3$ (CONAMA, 1998). For reference, in the United States the norm for daily concentrations of PM10 is 150 _g/m$^3$, the same as in Chile. Thus for more than one third of the year, Santiago's atmosphere did not meet national or international concentration norms in 1990.

As a source of congestion, the fact that 80% of bus routes pass through the 7 main arteries of the city has clogged the main roads of the central urban area (Malbran, 2001).

Towards the end of the eighties, the high tariffs, the high average age of buses, the low average capacity utilization of an oversized bus stock, together with the environmental and congestion externalities served to finally 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 revert the real tariff escalation of the eighties, lower the number of buses and increase the average capacity utilization rate (see Table 1).[11] By 2001 there were only 8,179 buses in operation in spite of the fact that the average number of passengers during a working day increase from 3,575,942 million to 4,275,913 million between 1991 and 2001. Occupancy of buses doubled during this period.[12] The average age of buses dropped to 6 years and over half of the current stock meets EPA-91 or EPA-94 emission standards. Service quality—measured by network coverage and waiting times—were not affected by the reforms. The authorities did not modify the existing route design when tendering was introduced so the network coverage is the same as before. Waiting times at bus stops are on average only 4 minutes (Ministerio de Obras Públicas, Transporte y Telecomunicaciones, 1997).

The new role of government was further fine-tuned during the nineties 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, which helped to increase the quality of service and a reduction in the environmental externalities caused by the sector.

[11] The authorities also directly retired close to 2,600 of the oldest buses in the system, paying close to US$14 in compensation to bus owners (Cruz, 2002).
[12] The passenger numbers are taken from the origin destiny surveys undertaken by the government in 1991and 2001.

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. The authorities are currently attempting to introduce further reforms along the lines of those implemented in Bogotá, which will be discussed further below.

Among the remaining problems are the following. First, when concessions were introduced no effort was made to optimize the network configuration nor 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 kilometers (EMG (2002))—and tariff integration is nonexistent except 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 are still an excessive number of buses in the system compared to an optimized network (SECTRA, 2002; Vivanco, 2002). Pollution and congestion could be lowered further still if more reforms are implemented.

Second, operators still earn their revenue from ticket collection. This generates strong incentives for buses 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—the greatest problem relates to safety. In 2001 there were 7,392 accidents in the Santiago urban area where a bus was involved; of the 5,699 injured people in those accidents 112 were fatalities. On average then there is one death every three days in an accident involving a bus of the transport system. Although comparative international data are hard to obtain, these figures seem dramatically high. If the experience of Bogotá—discussed in section 5 further below—can be used as an analogy, it is highly probable that a high number of these accidents are due to the economic characteristics of the current bus system in Santiago.

These pending issues serve to illustrate some of the problems inherited from the liberalizing experience. But they also point to another interesting issue. One of the main obstacles to further reform of the bus system has been the pressure from the bus owner's lobby[13]. Once the bus system is liberalized, strong interest groups form that make future reforms more difficult. From a political economy perspective it may be easier to follow the example of London Transport, where routes previously operated by a public company were tendered without the intermediate liberalization experience.[14]

## 3. Market failures in the urban transport industry

What Chile's recent history reveals is that market failures in the bus industry are not only the result of environmental and congestion externalities but that 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 result in many other forms of market failures.[15] Policymakers and the general public are typically concerned with the need to regulate to minimize congestion and environmental externalities related to public transport[16]. However, the failures related to the economic and network characteristics of the industry may also justify a regulatory intervention even in the absence of the more traditional congestion and environmental problems. In this section we review some of the economic problems encountered in bus transport industries. This discussion is organized at three levels: the aggregate network level, the individual bus lines and the problems related to the relationship between owners and drivers.

---

[13] This interest group has used the threat of strikes repeatedly. The last strike, which drove Santiago to a standstill, occurred in August 2002. However, the event helped to show the importance of this sector as a source of air pollution. Particulate Matter concentration levels dropped by half that day compared to a day with similar atmospheric conditions (Garreud and Aceituno, 2002).

[14] However, in this case there is another important interest group that could pressure the authority: the public bus company workers.

[15] While we picked Santiago de Chile to illustrate the point, we could have picked any medium to large city in the developing world since the evolution of the organization of the sector during the last third of the 20th century seems to have followed a very similar pattern across countries.

[16] Sometimes this sector is responsible for creating these externalities, at other times it is part of the solution to the externalities created by other modes of transport such as private cars. The latter is the case in developed countries.

### 3.1 Economies of density and network efficiency

A transport network can be conceptualized as a series of flows of passengers to and from different nodes. Figure 2 illustrates a very simple network. There are 4 nodes, labeled A to D, from which and to which people wish to travel. Let's assume that in a given period there are 100 people per hour who wish to go from point A to D, from B to D and from C to D[17]. Also, because of the structure of the road network, buses from point A and B have to pass point C on their way to D. One possibility to satisfy this travel demand is to offer three services, each carrying 100 passengers between the three nodes and point D. Assume that the cost is $1 per passenger for each segment along these routes. The total cost of satisfying this demand would then be $500 per period.

However, it may be that for technological reasons the cost of carrying a higher load of passengers does not increase proportionally to the number carried. For example, carrying 300 passengers may cost less than $1 per person. This may happen because bigger buses can be used (reducing congestion also), because a specialized transport network already exists (such as a subway system) or for some other technological reason. Let's assume that the cost of carrying 300 passengers drops to $0.5 per individual. With this assumption it would be cheaper to satisfy the transport demand by integrating the system as in Figure 2. In this case, passengers from point A and B travel to point C, where they then have to change to another transport mode (or type of bus) which will take 300 passengers from C to D. Ignoring transfer costs this transport alternative would cost $350 in the aggregate[18].

[17] For ease of exposition we will ignore the flows in the reverse direction

[18] In practice, the costs of building the necessary transfer infrastructure (for example for modal changes during trips) and the costs that such transfers impose on users must also be considered when evaluating the efficiency of an integrated network system.

**Figure 2: A simple transport network**

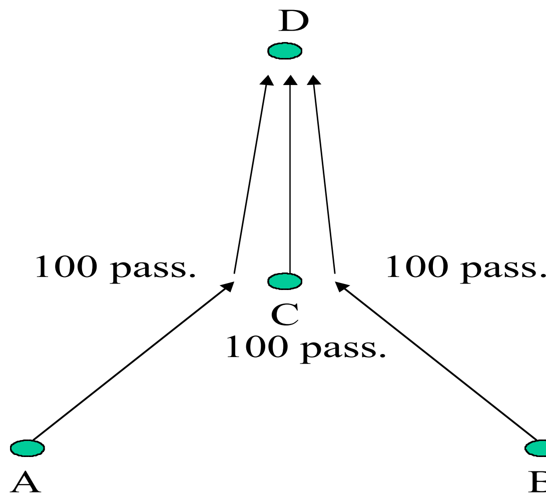
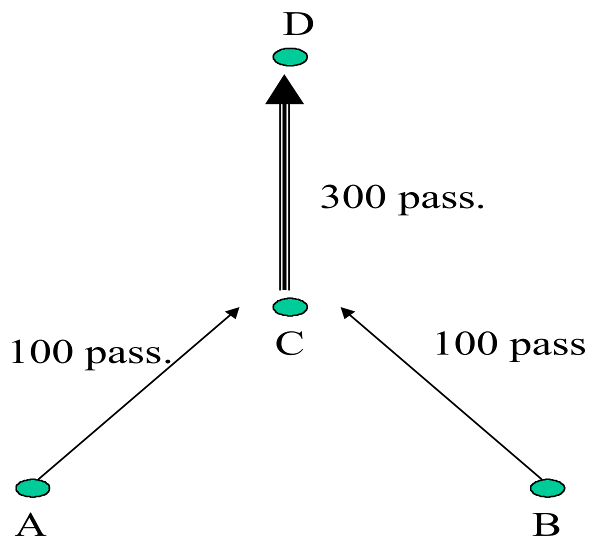



**Figure 3: An integrated transport network**




The above example rests on the assumption that there are economies of density in the transport network. That is, average transport costs decrease as the density of the passengers increases in certain parts of the network. In such situations, it may be more efficient to have services with shorter routes, more transfers between routes or modes by passengers (integrated transport system) and transport capacity designed to meet the requirements of each segment of the network. This in turn requires some kind of fare integration whereby a passenger can switch routes or modes by paying a single ticket or a fraction of the sum of individual fares.

That these economies of density exist—at least in the case of Santiago—has been shown in several studies (SECTRA, 2001; Vivanco, 2002). These economies of density are also presumed to exist in the airline industry and give rise to the 'scope and hub' configuration that this industry has adopted in the United States since deregulation[19]. However, in the airline industry competitive firms have exploited these economies of density efficiently. Why would regulatory intervention be required to achieve the optimal system configuration in the bus industry when in the airline industry it arises naturally from competitive dynamics? There are several reasons.

First, there may be high fixed costs to building the required transfer and ticketing infrastructure to implement an integrated transport system. However, something more is needed in this explanation, otherwise one could expect one company to grow, take advantage of these economies of scale, charge lower tariffs and become the dominating company in a city. In competitive bus industries where drivers collect fares, bus owners face difficult monitoring problems that generate strong diseconomies of size[20]. Thus taking advantage of economies of density would require reforms to eliminate this barrier to firm growth, such as independent revenue collecting agency and a revenue sharing scheme. But it is unlikely that this can be implemented without government regulation and coordination.

Second, the demand for air travel has different characteristics from the demand for bus journeys. The price of a flight is much higher compared to the cost of waiting time, and travel is less frequent. Thus it may be convenient for a passenger to wait a couple of hours at an airport for a connecting flight when this saves a few hundred dollars. Thus airlines are able to attract passengers by implementing a scope and hub network, lowering costs and then offering passengers lower prices, even thought this may imply more transfers and

---

[19] See Brueckner and Spiller (1994).
[20] This issue will be further analyzed below in the section regarding principal-agent problems.

longer travel times. The potential income saved by the passenger more than compensates for these additional nuisances.

In the case of buses the value of a journey is much lower and more frequent. If a company implements an integrated system in a city and offers a lower overall tariff provided passengers use only its buses, it is not at all clear that they will increase their patronage dramatically. When a person is standing at a bus stop he may prefer to take the first bus that comes along, even if its more expensive, rather than wait for the lower prices 'integrated' bus. The extra waiting costs may not compensate the lower price, especially since it is uncertain when the cheaper bus will appear. Thus, even though passengers may be better off if the whole system is transformed into an integrated network, an individual company that unilaterally tries to implement such a system may not grow. In our example, it may be that a proportion of passengers at node C (Figure 3) will take a competitor bus that charges a higher price but arrives first at the bus stop. If this non-integrated service from C to D is profitable it will take away demand from the integrated company that then may not be able to recoup the investment in the higher capacity vehicles (or transfer infrastructure)[21].

Finally, even if a private operator was able to exploit the economies of density in a network, the optimal private network configuration may not coincide with the socially optimal one. This occurs because it the firm would be unable to appropriate the value of the reduction in travel time associated with lower congestion in an integrated system, which will also accrue to passengers of other operators and private cars. Thus the optimal integration from a private point of may fall short of the socially optimal one.

[21] Notice that in essence this argument is the familiar one related to an industry with large scale economies: it is socially efficient to have one 'natural' monopoly provider. The difference here is that unlike the assumption usually made in a natural monopoly setting, competition would be *sustainable* if the larger firm were unable to drive out the other firms by lowering prices. Also notice that a 'natural monopoly' provider in this setting does not imply having just one transport company. This is because the cost economy does not come from the joint production of the services in the different segments. Different segments of the network can be operated by different firms, as will be discussed further below.

### 3.2 Efficient entry and tariff regulation

At the level of individual lines there are some other reasons why a free market in urban transport may not provide the efficient number of services. Lack of regulation may result in under or overprovision of the socially efficient level of service in the sector. Tariffs may also have to be regulated in an urban transport system and in some cases it makes sense to subsidize these tariffs.

#### 3.2.1 The waiting time externality

In a very influential paper, Mohring (1972) argued 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[22]. 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 5 buses pass uniformly each hour through a given point in the network, there will be a 12-minute interval between buses. Passengers—if they arrive uniformly to take a bus at that spot—will wait on average 6 minutes before they can board. If an extra bus is put into service, the time interval between buses will fall to 10 minutes, and the average waiting time to 5 minutes. Therefore, all passengers—independently of whether they actually use the new bus—will gain an extra minute in the form of less waiting time. This is an externality created by the additional bus.

Morhing (1972) argued that since there is an added social benefit in increasing service frequency, the private equilibrium where the number of buses is determined by the extra private cost and the extra revenue obtained by the marginal bus will be socially inefficient[23]. Thus, bus companies should be subsidized in order for them to offer the socially optimal level of service. Note that even in a market where tariffs are determined

[22] See Box 1 for a more formal derivation of this result.
[23] This result is more general. In product differentiation models if firms cannot appropriate all of the social surplus from entry then there will tend to be insufficient entry.

freely, the private equilibrium will not supply the correct number of services. An individual supplier cannot appropriate the extra surplus—in the form of charging higher tariffs—created by an increase in frequencies, since most of the benefits of waiting time reduction will accrue to passengers that are not his client. A monopoly provider who increases frequencies and fares on all buses could resolve this problem. However, this industrial structure would require regulation in order to curb the monopoly power of the company.

### 3.2.2 Excessive duplication of fixed costs

While Morhing (1972) showed that there is a clear risk of under provision, there are also factors in the way the sector is often organized that could result in excessive supply. Private firms could indeed put too many buses on the road compared to the socially optima level and this can happen independently of congestion or environmental externalities. The argument is the familiar one from differentiated product markets with fixed costs, where the business stealing effect of competition produces too much entry when there are fixed costs.[24] An additional entry reduces all firm's demand forcing them to operate on a high point on their average cost curve.

Why are buses differentiated products? Obviously, buses going to different destinations are not perfect substitutes. However, even buses passing the same final destination of a passenger may not be perfect substitute. This occurs because of the time lapse between the arrival of buses. An individual may 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 this next bus is cheaper. Thus, the bus that just arrives at the stop has some 'captive' clients, which implies that it has some market power to raise tariffs[25] [26].

[24] See also Mackie, Preston and Nash (1995).
[25] See Fernandez and De Cea (1990) for a rigorous derivation of the demand curve faced by a bus.
[26] It is interesting to note that this effect will probably not occur in other transport industries, such as airlines. In that case, the infrequency of use and the characteristics and value of the product justifies the use of scheduling. In turn, this implies that passengers can plan their trips without incurring unexpected waiting costs. In the bus industry scheduling is much more difficult to apply because due to traffic and other influences it is difficult to predict the exact arrival time of a bus at a stop.

Raising tariffs, however, creates excessive returns to investment in this industry that spurs the entry of new firms or buses. Since there is a fixed cost to entry (the capital costs of the buses), in equilibrium too much entry may ensue. The inefficiency of this situation is due to the excessive duplication of fixed costs.[27] Obviously the excess entry problem will be compounded if there are environmental and congestion externalities associated with the number of buses in operation.

There may also be another phenomenon generating an excessive number of buses in a competitive transport system. If there is an incumbent provider or a cartel (in the case of Santiago it could have been the operators association) then they may have an incentive to preempt entry by competitors in order to avoid competition from independent operators. This preemption strategy 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.[28]

**Box 1: A simple model of supply of public transport**

The following model is very simple, but it offers some important intuitions with respect to the potential market failures in a competitive public urban transport system. Let's assume a transport system that operates exclusively on buses during a certain cyclical period of time (say one day, one week, or one month). During this period, $N(p)$ passengers arrive uniformly to take a bus, where $p$ is the price of one trip. Assuming that the regulatory authorities fix the price, the total number of passengers per period is exogenous to the operators. $B$ is the number of buses in circulation. It is also assumed that each firm owns only one bus that is operated only once per period[29]. There are no operation costs—again an assumption that does not change the results—but there is a fixed cost of $F$ per period. This fixed cost can be the capital value of the bus or terminal cost.

Assuming that buses are 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. Therefore, profits to a bus owner are:

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

Assuming free entry, the private equilibrium condition (cero profit condition) is:

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

When there are $B^p$ buses, there are no private financial incentives for additional entry. Each firm is obtaining a normal rate of return on its investment.

---

[27] The duplication of fixed cost effect still applies if prices are fixed (see Box 1).
[28] This argument is related to Schmalensee (1978).
[29] The model can be made more realistic by separating the number of buses and the rotation of theses same buses in the system. However, this would not add any particular insights compared to our simple model.

However, the above equilibrium may not be socially optimal. A central planner would also take into account the effect of the number of buses on the average waiting time of passengers. Assuming that the average waiting cost for a passenger is *c* per unit of time. On average, a passenger will have to wait T/*2B* time units before he can board a bus, where T is the time it takes a bus to complete the circuit. Therefore, the social welfare function is given by:

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

Maximizing this function with respect to the number of buses one obtains:

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

Depending on the parameter values, the socially optimal number of buses could be smaller or larger that the private equilibrium. The result will depend on two countervailing forces:

a) In the private equilibrium, bus owners and operators produce a negative externality on other operators. The marginal bus owner does not take into account the fact that his entry into the market will take away some of the clients of other operators and there is an excessive entry of buses. The social cost of this externality is that fixed costs, *F*, are being duplicated without need. Another way to look at this is to recognize that the fixed costs generate a decreasing average costs curve for firms (economies of scale). When too many buses enter, each firm will be underutilizing its capacity and operating on a high portion of the average cost curve. We could label this effect the "common property externality". The higher is the price for a journey the more this excess entry effect will be present.

b) There is another effect that goes in the opposite direction. A private operator will not take into account the reduction in waiting times of passengers in his decision to introduce an extra bus into the network. We could label this the "waiting time externality" that generates insufficient entry and is originally due to Mohring (1972).

### 3.2.3 Will competition guarantee efficient prices?

In the previous sections we have hinted that competition may not drive prices down to their efficient level in a bus industry. If there is a natural tendency for prices to be above costs then the excess entry problem noted above will be fostered. If lowering prices does not increase demand for an operator, then private operators would not be able to successfully exploit the economies of density of the network as noted in section 3.1. Therefore, how prices behave in a competitive market is crucial for many of the arguments raised above.

Why would competition not lower prices? Moreover, why did the large number of new entrants in the city of Santiago during the eighties not compete through prices?[30] There are several reasons. First, in the case of Santiago many accusations of collusion were raised and the anti-trust authorities heard several cases related to the sector. It was claimed that the operators association was able to successfully impose its policy of equal (and high) tariffs on all of its members. However, if the property of the sector was so atomized, as noted above, how were the associations able to discipline its members? After all, cartels are often quite unstable.

One hypothesis to explain how the cartel was maintain is that the associations were able to offer insurance and legal services at cheaper rates than what the market would offer. Thus an independent operator that did not follow the associations' tariff policy would face significantly higher insurance and legal costs. Second, if the entrant was small the association members could practice predatory behavior on the entrant by guaranteeing that a member's bus would always run just ahead of the entrant's bus. Thus taking away a substantial number of potential clients. Finally, in view of the accusations presented to the anti-trust authorities it seems that plain criminal activity (physical aggression to owners, drivers and buses from independent operators) was also a method used to discipline members.

The second, complementary, hypothesis is more interesting because it points to structural problems in the transport market that may be generic and thus relevant to all countries. The basic argument—briefly exposed earlier—is that the demand faced by an individual bus on the street may be insensitive to the price charged within a certain range (Fernandez y De Cea, 1990). A passenger that arrives at a bus stop will take the first bus that appears if it is the low priced one. However, if it is a high priced bus, what will he do? He can either take the bus or wait for another bus. The first alternative has the cost of paying a higher price than necessary for the journey. However, the second alternative has the cost of waiting for the next bus to arrive plus the added inconvenience that the second bus may also be a high priced bus. Given this choice some individuals will prefer to take the first bus that arrives at a stop irrespective of whether it is a high or low priced operator, assuming the price

[30] All operators charged the same price irrespective of the particular characteristics of the route. This homogeneity in prices was noted by many analysts as an indication of collusion, since costs in different routes differed.

difference is not too high compared to the waiting time cost of individuals. This in turn implies that an operator that charges above average tariffs may not loose all his clients or that lowering tariff will not substantially increase demand. Every operator faces an inelastic demand curve on the road and therefore will exploit this 'market power' by raising prices. Thus, in equilibrium tariffs will be above costs.

### 3.2.4 Congestion and other externalities

Ultimately, whether there is excess supply in the bus industry is an empirical matter that needs to be analyzed with reference to particular experiences. However, from a conceptual point of view, there are additional influences that come to bear on this issue. 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 level of service will be lower still. This implies that the probability 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[31] reviewed the public transport system in 9 cities of developed countries. They found that public transport usually accounts for less than 25% of trips (the highest share of public transport was found in Zurich, 37%) in these cities. Private car transport, on the other hand, accounts for well over 50% of trips[32]. This is in stark contrast to developing countries, where in a city such as Santiago, Chile, 70% of trips are undertaken on public transport.

The modal split will have an influence on the causes of congestion and pollution. In a developing country, where the main source of urban transport is buses, these will have a

---

[31] NERA and TIS.PT (2001).
[32] The rest include walking, bicycle and others.

greater responsibility for congestion and pollution.[33] This argument is reinforced by considering that in developing countries bus technology will probably be older an more polluting than in developed countries.

It is possible then that in developing countries a free market may provide an excessive supply of services (due to these externalities) while in a developed country it may provide insufficient services. The policy recommendations are very different in each case. While in the latter subsidies to public transport may be justified in order to increase services, in developing countries efforts may be needed to rationalize the public transport system. The liberalization experiment during the 80's in Santiago discussed in section 2 would seem to be a case where the excessive entry phenomenon plus the negative congestion and environmental externalities was stronger than the positive waiting time externality.

### 3.3 Ownership structure

A third area of regulatory concern deals with the ownership structure and the contractual relationship between owners and drivers. In many developing countries with private provision of transport services in urban areas, a very atomized ownership structure arises. 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[34].

This ownership structure may be due to several factors. First, tax treatment of bus companies may be one explanation. For example, in Chile, bus companies pay on a presumed income basis and are not required to keep full income and cost accounts. Therefore, there are no institutional incentives to create formal companies with sophisticated management.

[33] This is not to deny that private cars are also important sources of congestion in cities of the developing world, just that their relative influence is probably smaller than in developed countries.
[34] Data provided by the Ministry of Public Works, Transport and Telecommunications.

Second, liability issues may be present. For example, in Chile bus owners are legally liable (in case of accidents an other legal claims) only up to the value of their property. This generates incentives to atomize the property of buses. In Santiago, it is not uncommon for an entrepreneur to divide his buses among family members.[35]

Finally, there are also diseconomies of size that arise due to the monitoring costs of a large fleet of buses when drivers are responsible for revenue collection and have to compete for passengers on the street. 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, full effective policing is costly. In fact, in Santiago owners and analyst agree that drivers can supplement their incomes by around 20% through this type of fraud[36].

This possibility of fraud by drivers makes it harder for bus owners to delegate the operation of buses to hired employees. Thus, one tends to observe that bus owners tend to be drivers also, or drivers are close relatives of the 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.

The problem with the above ownership structure is that it makes it harder to establish an integrated transport system, with tariff integration, since this requires some type of revenue sharing between firms. Small informal firms do not have the working capital or the administrative capacity to function in a sophisticated transport system. Thus, the ownership structure that arises when revenue in collected by drivers is an obstacle that makes it difficult to exploit potential economies of density in a network.

---

[35] Although this implies that the data on the fragmentation of bus ownership shown earlier for Santiago may be misleading (since the head of the family will the *de-facto* owner of the buses, taking this into account still shows that ownership is quite dispersed.

[36] One popular way that this is done is that some passengers return the ticket to the driver before getting off the bus, giving the driver the opportunity to reutilize the ticket and thus keep the extra fare without the risk of being caught by an inspector. Straight fraud is also common. Some passengers ask the bus driver if they will take them for less than the fare. The driver keeps the money and does not give a ticket to the passenger. Clearly, this last strategy is only possible due to the scarcity of inspectors or the low level of penalties.

### 3.4 Principal agent problems

In a competitive bus market, the number of passengers using a particular bus will be 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 this client. The natural private solution for this moral hazard (or hidden action problem) is for the principal (owner) to design a contract for the agent (driver) 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 as a private solution to the moral hazard problem faced by bus owners, 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 earlier 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. It is interesting to note that in a competitive bus system, this head-running phenomena generates a niche market for workers_called "sapos" (toads)_who stand in the main road arteries and sell passing drivers information of the time elapsed since competitive buses have passed that point on the route. It is questionable whether these services are socially beneficial since the impact on frequencies is not clear.[37]

Paying drivers a fixed wage 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

[37] Even if they do have a socially beneficial effect regulating service frequencies, it is not clear either whether this is the most efficient way to meet this objective compared to modern electronic technology.

loss[38]. Imposing this by decree to all bus companies will also be difficult, since the logic of a competitive bus industry is for companies to compete for passengers. As long as owner's 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.

Save very strict traffic monitoring, the only way to avoid the safety and other problems related to competition on the road is to break the link between profits and passenger carried. This is why 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 may be difficult to achieve without heavy regulatory intervention. In addition, this policy option has the draw-back that operators may 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 (for example the compliance with route schedules and other quality measures) without a significant change in the service provided to passengers.

Finally, revenue collection by drivers or other personnel on buses generates incentives for armed and violent robberies.[39]

## 4. New Regulatory instruments to cut the risks of market failure.

The previous section implies that there are four main regulatory challenges in the regulation of the bus industry and which form the backbone of the new emerging hybrid model. The first is to design and integrate the transport network in order to exploit economies of

[38] In fact, there is one experience in company in Santiago that started paying drivers a fixed wage. Revenue for this company dropped by 40%, thus ending the experience.

density and scale without compromising the system's coverage.[40] 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 in order to avoid the negative effects that competition for passengers on the street entails. This requires some integrated revenue collection system independent of operators which can then distributes these resources among firms. This has an interesting implication for the ownership structure of the bus companies. They will have to become larger and more formal enterprises, capable of having some amount of working capital in order to operate in the short run since they wil not receive revenues from fares on a continual daily basis.[41] In addition, the authorities would have to setup 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 favored by policy makers due to their negative incentives for efficiency.[42]

---

[39] In Chile this has become such an important problem—with drivers being killed in order to steal the buses revenue—that proposals have been made to install special cabins on buses for drivers or revenue collectors.

[40] Transport engineers have developed sophisticated models that are used in order to find the optimal system configuration.

[41] In many countries with informal or small family oriented bus firms, operators do not have access to credit and usually require daily revenues to pay drivers, fuel and other expenses. Notice that decoupling revenue from passengers will by itself remove one of the obstacles to firm growth—the owners monitoring of drivers problem.

[42] For a summary of the literature on the impacts of ownership and subsidies on bus firm performance see De Borger and Kerstens (1999).

Naturally direct competition and free entry into the industry is one mechanism to foster productive efficiency. However, it raises the problems identified in Section 3. The purpose of this section is to discuss some 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 tariffs and transfers to transport firms faces an asymmetry of information problem. The companies that are 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.

### 4.1 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 usually last many decades and are sunk, while it is not convenient to set rates for that long. If periodic auctions were used to set prices some rule would have to be established to transfer assets in the case where an incumbent losses the competition. The complications and possible inefficiencies that may arise due to imperfect asset valuation and transfer rules limit the use of tendering in these industries. 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 must not 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[43].

Besides Santiago, Bogotá, and London, many other transit authorities use tendering as a way to assign operators to serve previously defined network routes[44]. 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 reverted the tariff escalation phenomenon that characterized the deregulation period. In London, it has been estimated that tendering reduced operating costs per bus kilometer significantly. Estimates range from 20% gross of administrative costs or 14% net of administrative costs (Glaister, 1998) to 35% (White, 1995).

There are several ways to tender a route contract. In the case of Santiago, contracts are tendered for a 5-year period based on a multi-variable 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 three-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 to costs. The second system is based on a gross cost basis. That is, bids are received for the gross amount of transfers that the firm is willing to receive in order to undertake the service. In this case, the company does not retain any revenues from fares (or may not even be responsible for collecting fares as in the case of Bogotá) and funds its operation entirely from the transfers.

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

[43] Another sector where tendering is common is garbage collection. The reason is the same as in the case of urban transport.
[44] For more cases see NERA and TIS.PT (2001).

will also include a risk premium in bids to compensate for the added risk, increasing transfers compared to the gross cost type contract.

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 also lowers the incentives for companies to seek passengers or otherwise stimulate demand by providing a high quality service. Therefore, if different routes do not overlap too much, there is effective traffic control policing and individual drivers within a firm coordinate their scheduling, then traffic safety and other risks of competition on the road may be low. In this case, a system that leaves some revenue risk to firms (minimum fare or subsidy contracts) may be preferable in order to give firms some incentive to meet its timetable and provide a high quality service to customers. However, in developing countries where traffic safety and assaults on bus drivers are common a gross cost contract that rewards a firm according to their compliance with reestablished service provisions may be preferable. As noted above, there is now relatively inexpensive technology to centrally monitor bus traffic and position.

### 4.2 Combining entry restriction with yardstick competition

While tendering is clearly an attractive instrument it is not always feasible. For example, there may 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 may not be effective is that there may not be enough firms to guarantee sufficient competition during the process, perhaps because firms have an ability to collude[45]. Therefore, it is relevant to examine the performance of alternative regulatory contracts on productive efficiency of transport firms.

[45] This seems to have been the case in the tendering of contracts in Santiago in 1998, where 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).

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 can, if well applied, 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 may 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 very unique and thus 'non-comparable' to 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.[46] This experience is reviewed in Box 3.

**Box 3: the use of yardstick competition in Norway**[47]

In Norway, responsibility for local transport is decentralized to the regional governments (counties). They define the network route, schedules, fares and the subsidies given to companies. Each county is free to choose its own regulatory policy. As a consequence, regulatory practice has evolved differently across the various regions of Norway. Two basic approaches have characterized regulatory practice in Norway. What both approaches have in common, though, is that they involve lump sum grants to companies (in 1989 the share of government subsidies amounted to about 40% of the total revenues of Norwegian transport companies). The majority of these companies were privately owned). Most counties originally bargained annually and individually with each company over both costs and transfers. This scheme exposed companies to a ratchet type effect. If costs are reduced it becomes harder to argue that the next lump sum grant must be kept at the previous level. On the other hand, if costs rise it may not be too difficult to convince the authorities that subsidies must increase. Therefore, it is reasonable to assume that the individual-bargaining scheme is of low power in terms of the cost reducing incentives that it provides.

[46] Different researchers have tried to analyze the impact of the yardstick regulatory contract in Norway. See for example, Jorgensen, Pedersen and Volden (1997) and Dalen and Gómez-Lobo (2002a). They all find a positive impact of the yardstick contract on firm performance. However, the effect is surprisingly small (around 2-3% of operating costs) and may not even be significant (see Dalen and Gómez-Lobo (2002b).
[47] This case study is taken from the work of Dalen and Gómez-Lobo (2002a).

In the late eighties some counties 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 kilometers 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 that is granted by the regulator.

The important aspect of the scheme is that the same standard cost model applies to all companies within a county. This makes the regulatory scheme more high powered: Once the criteria are set, realized costs by one company that happens 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 flavor of yardstick competition (see Shleifer, 1985). The main characteristic of yardstick competition is that transfers be based on a benchmark estimated on the basis of cost performance of a larger set of companies.

The standard-cost model is developed and maintained by a consultancy firm (ASPLAN VIAK) and is used by regional regulators to determine the level of subsidy granted to the companies. The main input to the model is a 4x5 output-vector Q that consists of the number of kilometers produced in four different types of routes by five different types of buses. $q_i^j$ denotes the number of kilometers in route type i with bus type j. Route types differ according to average speed of traffic and bus types differ according to size. The regulator controls this output-vector.

The benchmark costs of producing the output-vector is derived from a linear model that determines the drivers, fuel, maintenance and bus capital inputs:

$$c_k = p_k \sum_i \sum_j a_i^j q_i^j$$

where $p_k$ is the market price for input $k$ and $a_i^j$ is the input coefficient for output $q_i^j$. $a_i^j$ is derived from observed cost figures from companies within several regions and applies to all companies controlled by the regulator. The number of bus companies in each region is large enough to make $a_i^j$ almost insensitive to efficiency improvements by the individual company. Administration costs, $c_{adm}$, are set proportional to $\Sigma c_k$ (excluding fuel costs, however, that tend to vary considerably over time).

The level of subsidy, $S$, is now determined by the difference between expected traffic revenue, given fares and kilometer produced, and standard-costs:

$$S = P \cdot y^e \cdot \sum_i \sum_j q_i^j - \sum_k c_k - c_{adm} \ ,$$

where $y^e$ is the expected number of passengers per kilometer produced and $P$ is the average fare level set by the regulator. It is important to note that $S$ is not adjusted to the actual number of passengers per kilometer effectively transported.

More recently, both of the previous contracts are being replaced by a third type of contract, a subsidy cap. The companies and the county agree upon a reduction in the level of governmental transfers by X% per year, over a five years period. This last contract was introduced as a compromise after counties threatened to tender network services.

The following table shows how the use of the different types of contracts has evolved over the sample period 1987-1997. It shows the number of counties using each type of contract.

| Contract type | 1987 | 1988 | 1989 | 1990 | 1991 | 1992 | 1993 | 1994 | 1995 | 1996 | 1997 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Individual negotiation | 15 | 13 | 12 | 11 | 11 | 9 | 8 | 7 | 2 | 1 | 0 |
| Standard Cost | 4 | 6 | 7 | 8 | 8 | 9 | 9 | 9 | 10 | 7 | 3 |
| Subsidy cap | 0 | 0 | 0 | 0 | 0 | 1 | 2 | 3 | 7 | 11 | 16 |

During the first part of the sample period, most counties used the relatively low-powered individual negotiation type of contract. By 1992-1993, the yardstick model had become the most popular contract with 9 out of 19 counties using this type of regulation. From 1992 onwards, counties that previously used the low-powered contract switched to a subsidy cap. In the last two years of our sample period, many counties using yardstick contracts have also switched to the subsidy cap contract.

Dalen and Gómez-Lobo (2002a) 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 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.

### 4.3 Offering menus of contracts to the operators.

There may 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 he must regulate a single natural monopoly under asymmetric information. The new theory of regulation offers a normative framework for regulatory policies in such contexts.[48] What are the implications of the theory for policymakers? The first is that offering just one type of contract is not optimal (such as a Price-Cap in some counties in Norway). 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.[49] However, it could be that informally, during a negotiation stage, the regulator offers a menu of such contracts.[50]

It is interesting to 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 contract for transit firms through a

[48] The New Theory of Regulation began with the publication of Baron and Myerson (1982). An extensive treatment of the theory can be found in Laffont and Tirole (1993).

[49] The only documented experiences in this respect seem to be the in the US telecommunications market, where both state and federal regulators have offered choices of regulatory contracts to firms.

[50] This was the idea used by Wolak (1994) in his study of the California water supply sector. He estimated two models econometrically, one assuming the regulator offered the companies an optimal menu of regulatory contracts and another where the regulator observes the efficiency parameter of firms. His statistical tests show that the first model is more consistent with the data, implying that the regulator is effectively screening companies according to the theory presented above.

mix of econometrics and calibration. He uses data on 177 mass transit firms in Europe to estimate the distribution of the asymmetry of information cost parameters of firms. In his words "...the 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". Therefore, he 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. He the 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 their study of the public transit system in France, use a structural approach—previously explored by Dalen and Gómez-Lobo (1995)—to recover firm's underlying cost efficiency distribution. They then model the effects of the introduction of an optimal regulatory contract, including route tendering.

## 5. Towards a solution in the real world: the experience of Bogotá, Colombia[51]

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

During the nineties Bogotá, the capital of Colombia, presented many of the problems alluded to in the previous section. Buses competed vigorously for passengers in the streets (called the "the war of the cents" by Colombians) generating unnecessary traffic risks and a chaotic system of stops and goes. High congestion generated velocities as low as 10 kmh during peak periods. Even short trips would sometimes take considerable time. In 1998, it was estimated that the average trip took 1 hour and 10 minutes. The average age of buses was 14 years in 1998 and the service quality they could offer low. Average occupancy rate

[51] The source for the following case study comes from www.idu.gov.co, www.transmilenio.gov.co and Hidalgo (2001).

was 45%. It has been estimated that 70% of particular 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 center of the city. The reform was organized around what came to be called the TransMillenium Project that became operational in December 2000, only two 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 totaling 35 kilometers in length have been built. Also, 22 feeder routes are in operation with a total extension of 66,7 kilometers covering over 40 neighborhoods of Bogotá. Infrastructure investments during this first stage totaled US$213 million. The plan is to build up the system—in a period of 15 years—to 22 central corridors (for a total extension of 388 kilometers) and with the capacity to transport 5 million people daily. The total cost of the reform is estimated to be US$1,970 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.

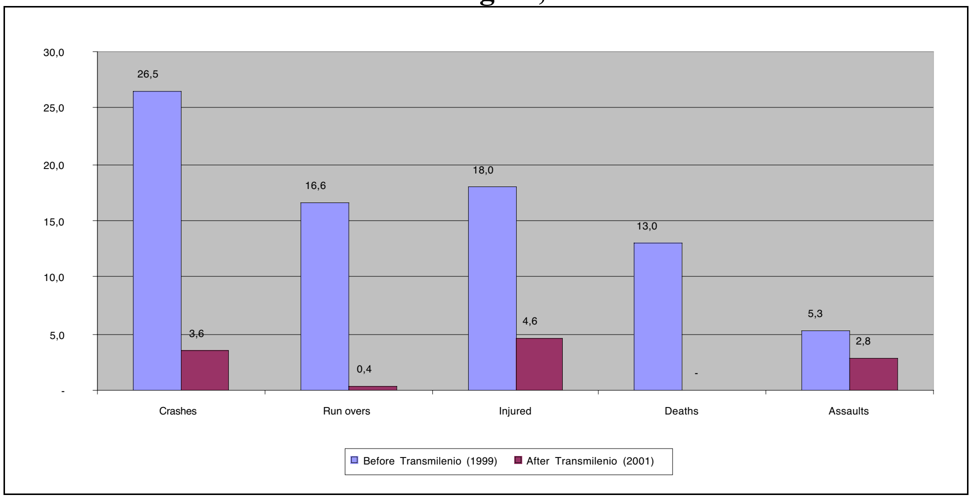
Among the most salient features of the TransMillenium project are a network design with bus enclosed stops every 800 meters, with pedestrian bridges and other services. Modern vehicles especially designed for passenger service were introduced through the conditions stipulated in operators' contracts. There are now 411 large articulated buses with automatic transmission systems, wheel suspension and modern natural gas o 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 minutes) and express services (stopping only at a subset of destinations and with a frequency of 4 minutes) 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 kilometers traveled 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 passenger transported. The separation of the operation of buses and the collection of revenues was made possible by the introduction of the pre-paid ticketing system operated by the special revenue collection company. This eliminated overnight the "war of the cents" radically improving traffic safety and quality of service. The electronic prepaid card system allows for tariff integration throughout the network

In 2001, after one year of operation, the evaluation of the experience has been very positive. Average velocity in the main corridors rose from 12 Km/h and 18 Km/h (Calle 80 and Avenida Caracas, respectively) to 26,7 Km/h after the project was in operation. As a consequence, average travel times fell by 32%, equivalent to a one hour saving daily for the average passenger.

Safety and service quality have improved dramatically. Figure 4 shows the number of accidents, injuries and fatalities on the roads corresponding to the TransMillenium network in 1999_before the reform_and in 2001 after one year in operation. A significant reduction in the number of accidents, people injured or killed and assaults can be noted.

**Figure 4: Weekly traffic accidents, injuries, mortality and assaults in the central bus network in Bogotá, Colombia**




*Source*: Hidalgo (2001)

Pollution levels have also dropped since the new system came into operation. Sulfur dioxide, $SO_2$, average daily concentration levels dropped by 43% between January and March 2000 with the same months in 2001. Maximum daily concentration levels dropped by 54% (January) and 39% (March) comparing both years. Nitrogen Dioxide, $NO_2$, average daily concentrations dropped by 13% (January) and 41% (March) , while maximum daily concentrations fell by 10% and 46%, respectively. For particular matter smaller than 10 microns, the corresponding fall in average daily concentration levels were 31% (January) and 17% (March) and 16% and 13% in peak daily concentrations. Although its possible that other phenomenon (particularly climate differences) may account for these measured improvements, it is likely that the introduction of the TransMillenium plan was also responsible.

As for the funding side of the reform, tariffs increased only 6% although most of the infrastructure investment was funded through petrol taxes, multilateral loans and other domestic sources.

## 6. Conclusions

During the last three decades the dominant model for the supply of urban bus services in developing countries has evolved from a highly regulated monopoly and public ownership structure towards a more liberal model with private ownership, liberalization and competition in the market. In this paper we argue that this last model has several shortcomings which has provoked a return to a more regulated industrial structure. In this new hybrid model, the authorities determine the network structure, service quality and frequencies, and they force a separation between revenue collection and operating activities. However, unlike the traditional regulatory model, this new hybrid system recognizes the importance of productive efficiency and 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 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 may 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 monitor the contracts afterwards.

Second, by restricting entry into the market, direct competition is loosened. Thus, companies—although private—may not have sufficient incentives to increase productivity and control costs. The tendering of contracts is a way to avoid this problem. However, there are many experiences around the world were a central authority determines the network configuration and services, the operation is undertaken by private firms (sometimes public-private joint ventures or even a mix of private and public companies) but the contracts are not tendered. Rather, they are either negotiated periodically or based on some kind of incentive contract. In this case, the issue of efficiency incentives arises and we noted several regulatory instruments that may help, including yardstick competition and offering a menu of contracts.

What happens when there is insufficient institutional capacity to implement a sophisticated system such as that applied in London or Bogotá? Moreover, one should never discount the risks of regulatory failure as a limitation to the 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 probably not be feasible either. In addition, for small urban areas where there are not many economies of density and services are not complex, the competitive model is probably better. A common criticism of the competitive model for small urban areas is that route coverage quickly diminishes as loss making routes are closed. However, this can be easily overcome through the introduction of selective public subsidies for the operation of unprofitable routes as has been done in the United Kingdom.

## References

Baron, D. and R. Myerson (1982), ‘Regulating a monopolist with unknown costs’, *Econometrica*, vol. 50, pp. 911-929.

Brueckner, J.K. and P.T. Spiller (1994), ‘Economies of Traffic Density in the Deregulated Airline Industry’, Journal of Law and Economics, XXXVII, pp. 379-415.

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

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

Dalen, D.M and A. Gómez-Lobo (1995), ‘Regulation and incentive contracts: an empirical investigation of the Norwegian bus transport industry’, IFS Working Paper W96/8, The Institute for Fiscal Studies, London.

------------------------------------- (2002a), ‘Yardsticks on the road: an analysis of the Norwegian bus industry’, *Transportation* (forthcoming).

------------------------------------- (2002b), ‘Do high powered contracts really work: an investigation of the Norwegian bus industry’, mimeo, Department of Economics, Univesity of Chile.

De Borger, B. and K. Kerstens (1999), ‘The Performance of Bus Transit Operators’, paper prepared for the *Handbook in Transport Economics*, Volume 1, Pergamon Press, NY.

Díaz, G., A. Gómez-Lobo and A. Velasco (2002), “El transporte público en Santiago: de enemigo público a servicio público”, Centro de Estudios Públicos, mimeo, Santiago.

Dourthé, A., H. Malbrán and M. Wityk (2000), ‘Regulación del Transporte Público de Superficie: la experiencia de Santiago de Chile’, mimeo, SCETRA.

EMG Consultores(2002), ‘Análisis de la Estructura Tributaria del Transporte Público y de Carga’, Informe Final, Ministerio de Obras Públicas, Transporte y Telecomunicaciones, agosto.

Fernandez, J.E. and J. De Cea (1990), ‘Caracteristicas de equilibrio en un sistema de transporte publico urbano bajo condiciones de libre competencia’, *Revista de Ingeniería de Sistemas*, VII(2), 27-43, diciembre.

Gagnepain, P. And M. Ivaldi (2002), ‘Incentive Regulatory Policies: The Case of Public Transit Systems in France’, The RAND Journal of Economics, vol. 33, no. 4, pp. 605-629.

Garreud, R. and P. Aceituno (2002), Las concencuencias ambientales del paro de micros del 12 de agosto, nota, Universidad de Chile.

Glaister, S. (1998), 'Desregulación y Privatización', in de Rus, G. and C. Nash (editors), *Desarrollos Recientes en Economía de Transportes*, Editorial Civitas and World Bank Institute.

Gomez-Ibañez, J.A. and J.R. Meyer (1993), *Going Private: the international experience with transport privatization*, The Brookings Institution, Washington D.C.

------------------------------------------ (1997), 'Alternatives for urban bus services: an international perspective on the British Reform', *Transport Reviews*, vol. 17, no. 1, 17-29.

Gwilliam, K.M., C. Nash and P.J. Mackie (1985a), 'Deregulating the bus industry in Britain: the case against', *Transport Reviews*, vol. 5(2).

Gwilliam, K.M., C. Nash and P.J. Mackie (1985b), 'Deregulating the bus industry in Britain: a rejoinder', *Transport Reviews*, vol. 5(3).

Hidalgo Guerrero, D. (2001), "TransMilenio: el Sistema de Transporte Masivo en Bogotá", mimeo, TransMilenium S.A.

Jorgensen, F., P. Andreas and R. Volden (1997), 'Estimating the inefficiency in the Norwegian bus industry from stochastic cost frontier models', *Transportation*, 24: 421-433.

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

Laffont, J.J. and J. Tirole (1987), 'Auctioning Incentive Contracts', *Journal of Political Economy*, vol. 94, pp. 614-41.

Mackie, P., J. Preston and C. Nash (1995), 'Bus deregulation: ten years on', *Transport Reviews*,

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

Mohring, H. (1972), 'Optimization and Scale Economies in Urban Bus Transportation', American economic Review, Volume 62(4), pp. 591-604.

NERA and TIS.PT (2001), Models for the Provision, Regulation and Integration of Public Transport Services, A final report for the Public Transport Partnership Forum of Ireland, London.

Paredes and Baytelman (1997), ‘Urban Public Transport Deregulation: the Chilean experience’, *Estudios de Economía*, vol. 23, pp. 195-210.

Pedersen, P.A. (1994), ‘Regulating a Transport Company with Private Information about Costs’, *Journal of Transport Economics & Policy*, vol. 28, pp. 307-318.

Oum,Tae-Hoon, et-al. (1997), Transport economics: Selected readings. Amsterdam: Harwood Academic in cooperation with the Korea Research Foundation for the 21st Century.

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

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

Shleifer, A. (1985), ‘A Theory of Yardstick Competition’, *Rand Journal of Economic*s, vol. 16, pp. 319-27.

Vivanco, J.L. (2002), *Estimación de los Ahorros Financieros Mínimos de la Optimización del Sistema de Transporte Urbano*, seminario para optar al Título de Ingeniero Comercial Mención Economía, Facultad de Ciencias Económicas y Administrativas, Universidad de Chile.

White, P.R. (1995), ‘Deregulation of local bus services in Great Britain: an introductory review’, *Transport Reviews*, 15, pp. 185-209.

White, P and S. Tough (1995), ‘Alternative Tendering Systems and Deregulation in Britain’, *Journal of Transport Economics & Policy*, vol. 29.

Wolak, F.A. (1994), “An Econometric Analysis of the Asymmetric Information, Regulator-Utility Interaction”, *Annales D’Économie et de Statistique*, Nº34.

World Bank (2002), Public Transport Policy Paper.

Wunsch, P. (1994), Estimating Menus of Linear Contracts for Mass Transit Firms’, mimeo, CORE, Belgium.