INCENTIVE STRUCTURE IN TRANSIT CONCESSION CONTRACTS: THE CASE OF SANTIAGO, CHILE, AND LONDON, ENGLAND

Andrés Gómez-Lobo E.

Department of Economics

University of Chile

Julio Briones

Department of Transport Engineering and Logistics

Catholic University of Chile

The Clean Air Institute

INCENTIVE STRUCTURE IN TRANSIT CONCESSION CONTRACTS: THE CASE OF SANTIAGO, CHILE, AND LONDON, ENGLAND

Author
Andres Gomez-Lobo E.

Co- author
Julio Briones

CONTENTS

1 Introduction

Public transportation is an essential element in the daily lives of most people in Latin America as it is their main means of transport. In most large cities in the region public transport represents over 50% of motorized trips, representing the predominant transport mode for lower income households.[1] This service is therefore essential and impacts indirectly on a whole range of social and economic issues including access to education, employment, and health care, among others.

Despite public transport existing in most developing countries, these services are often inefficient, low quality and unsafe. There are disorganized route structures, poorly maintained vehicles and intense competition for passengers in the streets which leads to inefficient services.[2] These issues have motivated many cities to prioritize the need to reform public transport systems. In fact, many cities in Latin America are attempting to improve and formalize their bus operations; for example, by introducing mass transit solutions, developing integrated transport systems and optimizing bus routes. Such improvements are motivated by the need to reduce road traffic accidents, offer faster and higher quality transit services to users, reduce congestion, reduce CO_2 emissions, improve local air quality, and at its most general, curb rising private transport use by providing an attractive public transport alternative.

The major structural modification behind these reforms is the change from an economic model based on 'competition in the market' among informal transport suppliers to a model of ex-ante 'competition for the market' by formal transport firms. That is, most reforms are based on some type of concession for the supply of services.

In this context, one of the critical elements for efficient formalization and operation of public transport is the preparation of a robust concession contract for proper operation of the service. This should ensure that the final public transport service satisfies the needs of the population and provides a contractual regime which ensures financial sustainability, certainty, stability, and legal protection of the service providers.

The concession contract is an understanding between a company and a host government in which the government specifies the rules under which the company can operate locally. In the case of a public service concession, a private company enters into an agreement with the government to have the right to operate, maintain and carry out investment and operate services for a given number of years.

Many issues are critical when preparing an effective concessions contract. But key to the success of a reform are the economic incentives provided in the contract. These incentives will determine the performance of operators and will ultimately determine whether the new services meet users' and planners' expectations behind reform. Thus, the economic incentive structure provided is crucial to the acceptability and ultimate success of any public transit reform. As such, providing correct economic incentives in a concession

---

[1] See Urban Age (2008).

[2] See ITDP (2007) for practical explanation for these problems. See Estache and Gómez-Lobo (2005) and Gómez-Lobo (2007) for a more theoretical approach to this phenomenon.

contract is a pre-requisite for obtaining all the other benefits from the modernization of public transport systems in developing countries.

The objective of this report is to analyze the principal clauses of bus concession contracts to allow them to be adapted to other cities, and their concrete impact on the behavior of bus operators, the quality of the service provided, and transit systems finances, based mainly on the experience of *Transantiago* in Santiago, Chile, and in London, England . The ultimate aim is to derive lessons learned and provide recommendations for other cities that plan to embark on a contracting exercise or are faced with contract renegotiations with bus operators.

In Santiago, Chile, a full-scale reform of the public transport system was undertaken in 2007. *Transantiago,* as it is called, completely changed the route structure, the fare payment method, and the contractual relation with operators as well as many other dimensions of the city's public transport system. Unfortunately, the results were initially disastrous due to the design and contractual errors.[3] This lead to a series of contract renegotiations between 2007 and 2011 aimed at improving the incentive structure of contracts in order for the system to provide adequate services.

Extracting the lessons from the *Transantiago* experience regarding contractual design and incentive structure is important for several reasons. First, no other city in the developing world had previously attempted such an ambitious and all-encompassing modernization of its public transport system. Other notable reforms were limited to BRT type systems, most notably *Transmilenio* in Bogotá (and the pioneering experience of Curitiba before that).[4] BRT schemes, although extremely valuable, are limited in scope and ―as will be argued below― in such systems it is easier to regulate and enforce performance. As cities are contemplating more widespread and radical reforms, the case of Santiago provides a unique experience to analyze what works and, more importantly, what does not work, as far as incentive mechanisms are concerned in transit concession contracts.

Second, the continual contract renegotiation experience of *Transantiago* ―with at least three major renegotiation rounds between 2007 and 2011― provides a rich and varied experiment with different contractual mechanisms from which lessons regarding their application and results can be analyzed.

In this report, we also complement the above experience with an analysis of the incentive structure provided by the concession contracts for bus operators in London, England. This review provides a 'best-practice' benchmark from an integrated public transport system deemed to be successful and where services are provided by private operators under concession contract arrangements. Finally a brief review of other bus concession contracts in Latin America is presented in an annex to this study.

As stated above, the ultimate aim of this report is to provide policymakers with information and guidance regarding the different contractual options available in public transit

---

[3] A detailed review of this experience can be found in Gómez-Lobo (2012) and World Bank (2009). A brief recount summarizing the contractual issues is provided in Annex 1 to this report.

[4] See EMBARQ (2010) for a summary of different experiences around the world.

concessions and their possible benefits and pitfalls. As many cities in the region are now introducing ―as the SITP in Bogotá― or considering more widespread transit reforms that go beyond BRT schemes, the issue of incentives and performance in concession contracts will probably become an increasingly important topic determining the success of initiatives aimed at modernizing these critical services.

Section 2 of the report presents a conceptual discussion regarding incentives in concession contracts and their importance for operator´s performance and service quality. This discussion centers on the two main contractual mechanisms used to affect performance: (a) the payment mechanism (how are operators' revenue determined?) and (b) fines, penalties and bonuses linked to performance obligations (what are the sanctions for under-performance? What are the rewards for over-performance?).

Section 3 of the report discusses other important issues related to contractual design such as drivers' compensation (what restrictions or conditions are placed on drivers' payment and contractual relations?) and incentive mechanisms for fleet renovation and environmental performance.

Section 4 presents a detailed case study of the experience with several contractual mechanisms for the case of Santiago and London.

Finally, the report ends with the main conclusions and recommendations regarding contract design and incentive issues for public transport reform.

Annex 1 presents tables summarizing the key points of the Santiago and London experiences. A summary of other European experiences is presented in Annex 2, and a short review of other Latin American experiences, two BRT and one wider transit system reform experience which is in its infancy, are presented in Annex 3.

2   Incentive structure in concession contracts

In this section we discuss the incentive structure of concession contracts in urban transit. A concession contract establishes the service and quality obligations that operators must meet and the mechanisms and rules by which these operators are compensated or penalized financially. In what follows, we assume that operators are private for-profit companies and so respond to economic incentives. This is the usual case in public transit concession contracts ―either competitively tendered or otherwise― in developing countries, at least in Latin America.[5]

There are two broad categories of financial incentives in a contractual relationship. The first is determined by the way operator's net income (revenues minus costs) is determined. The **payment mechanism** scheme laid out in the contract will influence the operator's behavior in so far as operators try to increase their profitability, either by increasing income or

---

[5]   Some exceptions include the public operators in the city of Porto Alegre in Brazil and one operator in the BRT scheme in Mexico City. See NEA et al (2008) for a relevant discussion regarding incentives and risk sharing issues in public transport contracts and a comparison of experiences among European cities. See Annex 2 to this report for a summary table of this comparison.

decreasing operational costs. A misalignment between the variables and actions that determine the operator's profitability with the quality of service objectives of the authorities will produce unintended results. In the best of cases these will reduce the expected benefits of a transit reform and in the worst of cases may derail the reform altogether.

The second broad group for aligning an operator's performance with the authorities' service objectives is defining **fines, penalties and rewards** linked to performance standards. How these fines or rewards are established and how performance is monitored will also be a crucial component to the success of a transit reform.

It must be borne in mind that concession contracts will contain both kinds of clauses. Therefore, there is a continuum of possibilities as to the relative weight placed on each of these broad groups of incentives for operators' performance. In turn, there are cost and benefits to these different alternatives and trade-offs involved depending on the institutional capacity, the type of transit reform (BRT or other) and the main objectives pursued by the authorities in each particular case. The available options and the trade-offs involved in each case are discussed in the following section of this report.

Throughout the following discussion it will be assumed that the payment mechanism, fines and penalties and other contractual clauses in the concession contract are clear and do not imply discretion on the part of the authorities in the determination or interpretation of these clauses. Therefore, we abstract from negotiations and political economy issues between the authorities and operators. We also abstract in what follows from the possibility of renegotiating the concession contract in order to focus on the incentive properties of different contractual clauses.

2.1 Payment mechanisms

The way operators make a profit from offering transport services will determine their incentives. In turn, profits will depend on who bears the risk of cost overruns (or who receives the benefits of cost savings) and who bears the risk of revenue or demand changes.

In theory, risk should be assigned to whichever party is in a better position to control those risks or, if they are external to each party, then to whoever can absorb those risks at a lower cost. For example, it is reasonable to assign operational cost risk to operators since they are in a better position to take actions to lower these costs. The exceptions are cost risk due to input price variations that are not controllable by operators; for example fuel prices and other input price changes. Therefore, most concession contracts assign cost risks to operators although they are usually shielded from input price variation, usually by adjusting their payments with an input price index. In this last case, it is users (or the State) that bear the costs and benefits of input price variations.[6]

Since assigning operational cost risks to concessionaires and readjusting payments to compensate for input price variation is quite common, in what follows we will not consider

---

[6]   The way risks are assigned will affect the economic efficiency of the system and may have financial implications for the system as a whole. Ideally, risks should be assigned so as to minimize the economic costs of providing transit services for a given quality of service desired.

special contract provision as regards costs. Rather, the focus will be on the mechanisms that determine operator's income and the incentives provided.

Critical for the discussion that follows is the extent to which operator's income depends on the number of passengers transported. In transit services, it is often the case that contracts protect operators from demand risk. If the number of passengers transported decreases, operators' incomes do not fall (or do not fall in proportion to the demand decrease). Likewise, if the number of passengers transported increases, operators' incomes do not increase proportionally.

In order to compare alternative payment mechanisms it is useful to set out some criteria that can be used to evaluate each scheme. These criteria are:

The incentives provided to cater to demand. That is, the interest the operator will have to provide an efficient and high-quality service. Since quality of service is multi-dimensional and difficult to specify contractually in all its dimensions, the incentives provided by the payment mechanism may be crucial in determining outcomes. These include the effective stopping at stops to pick-up or let passengers alight, the cleanliness of buses, the treatment of passengers by drivers, the control of non-payment and informing the authorities when route changes or extensions may increase demand, among others.

Frequency and regularity of service. The payment mechanism will also determine the incentives operators have to provide a timely and regular service. Higher service frequency and regularity entail costs and will only be willingly provided (save for fines and penalties discussed below) by operators if revenues compensate these higher costs. Higher frequency and regularity reduce waiting times and make public transport more attractive to users, increasing demand for these services.

The safety of operation. The more revenues are linked to the effective number of passengers transported, the more operators will have the incentive to compete for passengers, often creating non-trivial traffic problems and safety concerns. Regarding this issue, the exclusivity of service determined in the contract will also be relevant. If operators have exclusivity of service in certain areas or streets, then competition among operators for passengers will be attenuated and more demand risk placed on operators may not generate high safety concerns.

The revenue risk and access to financial markets. The payment mechanism will also determine the revenue risk assumed by operators which in turn will affect the profit variability of the concession and the access to formal financial markets. The revenue risk may increase the financial costs of a reform, as operators will demand a higher expected return on their investment to compensate for the higher risk of the concession. Profit variability may be particularly important when the reform aims to renovate and modernize the fleet, requiring concessionaires to tap private sector financing. Usually, banks and other financial institutions will pressure for some kind of revenue guarantee in concession contracts in order to provide this funding. With respect to the risk premium charged by operators or their creditors for taking

on demand risk, the precision with which demand can be forecasted may be important. In cities where there is plenty of information regarding demand patterns and behavior and demand has been very stable in the past, then demand risk will be lower and making operators face demand risk may not be too expensive. On the other hand, in cities where a radically new reform is being implemented, where there is scant information regarding demand behavior and where it is difficult to predict future demand with certainty, then the risk premium will be large and it may more reasonable to shield operators from demand risk.

Planning and enforcement capacity required. Some payment mechanisms require a strong institutional capacity on the part of the authorities in order to define and specify services and frequencies. Some service quality dimensions may be difficult or costly to monitor.

Broadly, payment mechanisms can be grouped into three main categories:

- Fixed payments,
- Payments based on operational variables, and
- Payments based on passengers transported.

Naturally, there are intermediate options between these extremes as they are not mutually exclusive and most contracts will combine elements of each category. However, in order to discuss the incentive properties we will treat each one separately in what follows.

Fixed payments

Contracts can shield operators from demand risk by establishing fixed payments to operators. This can be accomplished directly by establishing a compensation to operators that is not dependant on the demand or operational variables. It can also be achieved indirectly, by establishing income guarantees in the contract or other formulas that attenuate, sometimes completely, the demand risk faced by an operator.

Fixed payments give maximum income security to operators and will probably reduce the risk premium demanded by them, lowering the financial cost of the reform. It will also increase access to financial markets, as financial institutions can predict concessionaires' revenue stream with relative certainty.

In addition, operators will not have incentives to compete aggressively in the streets for passengers, since their income will be independent of this effort. Therefore, this alternative ranks high if safety is a major concern for the authorities.

However, since income is fixed, operators have the incentive to increase profits by reducing costs which can impact on service quality such as service frequency and contracts will have to depend much more heavily on fines and penalties in order to achieve frequency and regularity of service, as well as other operational results —such as stopping at bus stops, pleasantness of drivers, cleanliness of buses, fleet maintenance, etc.— when payments to operators are fixed. Often reducing operating costs is based on restricting the activities "not so visible to the user or the authority" such as preventative and corrective maintenance of the buses. It is therefore necessary to incorporate performance clauses, quality

programming and program implementation of fleet maintenance and validation procedures. However, it is important to note that this activity increases the need for supervision and monitoring by the authority.

Finally, fixed payments do not generate any incentives for operators to cater to demand. They will not be responsive to user's needs, nor will they propose route changes or extensions that may improve services and increase demand. Route planning by transport authorities will be critical to guarantee good coverage and services, particularly as urban structure changes through time. Thus, fixed payment schemes require strong planning capacity (to determine the service required) as well as enforcement capacity.

Payments based on operational variables

It is quite common for bus concession contracts to establish compensations based on the effective supply of services. In the transport economics literature, contracts based on operational variables are often called Gross Cost contracts (Hensher and Brewer, 2001). An example is to pay operators a price based on the seat-kilometers (or seat-miles depending on the country) offered during a certain period of time.

This is by far the most common practice in developed countries (see NEA, et al (2008)) and is also quite common with BRT schemes in developing countries (ITDP, 2007). In fact, the review of concession contracts for three BRT cases in Latin America presented in Annex 3 (Bucaramanga, Transmilenio phase 1 and 2, and SITP trunk routes, both in Bogotá, Colombia) illustrates payment schemes linked directly or indirectly (and to varying degrees) to operational variables. This is also the case of the contracts used in London since the year 2000 and reviewed in Annex 2. As described in Section 4 of this report, the contracts renegotiated in 2008-2009 in the case of Transantiago can also be classified in this category.

This scheme has the advantage of reducing external revenue risk to operators but maintain incentives to provide an adequate level of service, at least in terms of frequency and regularity. This may be particularly important for low demand routes or non-peak hours, when demand is low and services may not be privately profitable.

Thus, this scheme ranks well in terms of risk and access to finance and frequency of service. Concrete examples for the case of Santiago and London will be discussed further below.

In spite of its advantages, operators do not have incentives to cater to demand since their revenue is unaffected by demand inducing efforts. Thus just like in the case of fixed payments, regulators will have to depend more heavily on fines and penalties to maintain service quality in dimensions that are costly to measure and include in a payment mechanism.

As an example, in Santiago, the contracts renegotiated in 2007 included a payment mechanism whereby operator's revenue depended on the number and capacity of buses in operation compared to the number and capacity established in the operational plan. This mechanism was augmented by including also the kilometers operated compared to the operational plan in the contracts signed in 2009. However, some operators could meet these

operational targets by supplying the correct number of kilometers, buses and capacity and still offer a bad service; for example, by not stopping at bus stops when requested.

Thus the payment mechanism based on operational variables induces operators to meet operational criteria that may not coincide perfectly with quality of service as expected by users and authorities.

In addition, this payment mechanism will also require strong planning capacity (to determine services required and frequency) as well as enforcement capacity. In this respect it is interesting to note that in the years 2008 and 2009 the authorities in London tested a variant of their Quality Incentive Contract by adding driving quality and vehicles' internal and external presentation as additional quality variables linking performance to operators' payments. However, they desisted from extending these contracts beyond a few pilot tests because of the high monitoring costs for these variables. Thus, even in cities with a very high institutional capacity, linking payments to operational variables may be costly and impractical.

Finally, under this type of payment scheme, operators will not have incentives to compete for passengers in the streets and thus this scheme ranks high in terms of safety.

Payments based on passengers transported

Although less common in developed countries, most transit services in developing countries are provided by private operators whose revenues depend directly on the number and type of passenger transported.[7] This is probably due to several factors, including: loosely regulated markets and the weak capacity of agencies to plan, monitor, and enforce complex concession contracts in these countries. In addition, these types of schemes are usually seen in non-integrated systems where it is not necessary to control revenue collection centrally since operators keep the fare revenue they collect.

As these operators ply the streets for passengers, adding to the growing number of private vehicles, they often create traffic chaos and generate important safety concerns and this issue is usually one of the motivations behind reform.[8] For instance, this was the case of Santiago's public transit system between 1979 and 2007. Should concession contracts then mimic these incentives by paying operators based on the number of passengers transported?[9]

Based on safety concerns the answer would seem to be negative. In addition, this alternative increases the revenue risk assumed by operators, particularly if some of the

---

[7] This type of contracts is also called Net Cost contracts in the literature (Hensher and Brewer, 2001) and the case of London between 1996 and 1998 is an example of the use of these contracts in the context of a developed country.

[8] This effect will be moderated if driver's salary is fixed or independent of passengers transported, an issue discussed further below.

[9] Another way to increase safety is to sign exclusive contract for certain routes/areas. In such cases, there will be no competition between different operators. However, there may still be some competition between driver of the same company if they are paid depending on the passenger carried (but probably only in services that exhibit very high frequencies).

demand volatility does not depend on actions and effort expended on their part. For example, demand for public transport will increase or decrease according to the general state of the economy and employment. Also, general economic growth and private car ownership growth will affect public transit demand in ways difficult to predict for an operator at the beginning of a long-term concession.

Making a concessionaire face these non-controllable demand risks will hamper attempts to tap financial markets for the needed fleet modernization or other financing needs of a reform. This will be particularly so if there is uncertainty regarding post-reform demand levels, making cash flow projections difficult. This is often the case in cities where there is scant information as to ridership levels in the informal pre-reform system.

To dispense completely of payments conditional on demand can create severe problems. Operators (and indirectly drivers) will have absolutely no revenue incentive to control non-payment (fare evasion), to propose route changes or service extensions that better serve users, or even to stop at bus stops to pick-up passengers. In this case, incentives can be provided by fines and penalties defined in the contract, but the authorities must have an effective monitoring technology to enforce such clauses, as will be discussed below. Infrastructure, most notably boarding stations in BRT type schemes where passengers pay before boarding, can be used to control non-payment in lieu of operators' control.[10] But in more extensive transit reforms, it will not be economical or practical to build pre-boarding station over the entire network and the way non-payment is controlled becomes an important issue. However, the infrastructure is a means to facilitate the operation and should address first the issue of demand and access, not just the issue of evasion of payment.

As for frequency and service regularity, payment conditional on effective demand will have differing effects depending on the characteristics of the route or service. For those services where demand is high, and revenues cover costs, operators will probably have incentives to provide services according to the operational plan.[11] However, if demand is low and revenues do not cover costs (as in night services), then operators will be better off financially by skimping on frequency. In this case, other control mechanisms must be used to guarantee service levels. This is one area where payment by seat-kilometer or a combination of seat-kilometer and passengers transported might be preferred.

Finally, making operators face some demand risk may provide some incentives for operators to propose route changes or new routes in order to tap new demand sources or improve existing services. In this respect, this payment option may reduce the requirement for planning and monitoring capacity on the part of the public agency in charge of public transport. This may be particularly useful in a fast growing city, where it may be difficult for a centralized authority to determine quickly the changing patterns of travel and new sources of demand. Operators on the other hand may have more local information regarding

---

[10] It must be stressed that station infrastructure should primarily be designed to speed up the operation of the system (faster boarding and alighting of passengers) and only as a secondary consideration the control of non-payment.

[11] The issue of whether a monopoly transit operator has incentives to over-provide or under-provide frequency is a debated issue. See Gómez-Lobo (2011) for a review.

transport demand as they are "in the street" everyday. Giving them incentives to turn this knowledge into network change proposals, by linking their payments to effective demand, may be an efficient way for the authorities to optimize services.

Table 1 provides a summary of the foregoing discussion. It can be seen that no mechanism dominates on all counts, although it can be seen that fixed payments will probably be dominated by a scheme where operators are paid according to operational variables. In practice, the best option will probably be a combination of these mechanisms. Indeed, the use of a single payment mechanism would be unlikely to succeed.

Table 1: Incentive properties of different general payment mechanisms

Payment mechanism	Fixed	Operational variables (seat-kilometer)	Passengers transported
Cater to demand	-	-	+
Frequency and regularity of service	-	+	+ (if demand is high) - (if demand is low)
Safety	+	+	-
Risk and access to finance	+	+	-
Planning and monitoring capacity requirements	-	-	+

Note: (-) indicates that the incentives generated by the payment mechanism generate a negative outcome, (+) indicates a positive outcome.

2.2   Fines, penalties and rewards linked to service obligations

An alternative way to provide performance incentives is to establish explicit service obligations in the contract and enforce them through fines, penalties and rewards.[12] It would seem natural to rely on both types of mechanisms in order to optimize performance, and real world concession contracts do include both.

However, it is important to note that providing performance incentives through payment mechanisms or through a system of fines, penalties and rewards may not be equivalent in terms of costs and effectiveness. Using the latter to guarantee good performance has several drawbacks. These include:

- *The need to explicitly define service obligations*: while all contracts define some set of service obligations, the correct definition of these obligations and their associated fines will be much more critical in the case where performance is dependent exclusively on the incentives provided by fines and penalties. This may be straightforward in some cases (frequencies, for example) but it may be much more difficult for others (for example, the way passengers are treated). In addition, as contracts are always incomplete, situations may arise that call for a particular behavior on the part of operators that was not even considered in the original contract. Finally, service obligations are often discrete, in the sense that a certain performance parameter is set and fines (rewards) are charged for not achieving (achieving) this parameter. But generally there are no incentives for over-complying with the performance standard. Payment mechanisms, on the other hand, are usually more continuous providing incentives for a range of performance values.

- *The results may be inefficient as compared to payment mechanism*: for example, as will be described in more detail below, in Santiago, as operators were paid according to seat-kilometers supplied, they did not always have incentives to stop for passengers at bus-stops. The service obligation answer to this problem would be to establish in the contract the requirement for buses to stop at all valid stops where passengers are waiting. However, this may be very difficult to enforce, requiring inspectors at stops to verify that buses are complying with this requirement. Another alternative would be to control this behavior through technological means.[13] However, this is clearly inefficient requiring more technology and higher travel times due to unnecessary vehicle stops (or higher spacing between bus stops to reduce this time cost), at least compared to a system where operators are given more demand risk and therefore have incentives to stop for passengers without requiring monitoring and enforcement.

---

[12]   In some systems, such as Transmilenio in Bogotá and Transantiago, a fraction of fines go to a special fund used to reward operators for good performance. This mechanism has the advantage that it may reduce the incentives for operators to avoid paying fines (since they may expect to receive back some of these resources as performance payments) through legal or non-legal (bribing inspectors) means. In addition, inspectors may have fewer incentives to pass excessive fines since no institution benefit financially from these fines.

[13]   For example by verifying that the bus stopped and doors were opened at each stop on a route using GPS and other electronic information.

- *Higher enforcement costs*: as the previous example illustrates explicit service obligations require more monitoring and enforcement effort on the part of the authorities. This may be particularly challenging where institutions are weak or badly funded —as is often the case in developing countries— or vehicles are not equipped with the required monitoring technological (possibly due to cost considerations).

- *Ineffective in certain contexts*: even when obligations can be defined in the contract and performance can be monitored, fines and penalties may still be ineffective. This will be the case when fines and penalties are so high that they risk the bankruptcy or continuation of the concession. This may not be a viable option for the authorities so these fines or penalties cannot be applied and these clauses provide no incentives for performance. We will give an example further below when discussing the case of Transantiago.

The point of this discussion is that in general a concession contract will have to balance the cost and benefits of the different options to provide incentives. This will usually require a mixed system where some incentives are provided by the payment mechanism and others through fines, penalties and rewards linked to performance indicators.

The particular weight put on each mechanism and design option will also depend on the type of reform undertake, the institutional capacity of the authorities, and the costs of the different options.

For example, in BRT type schemes, non-payment is easy to control since passengers usually pay before boarding in a small set of enclosed pre-payment stations. Guards and infrastructure (tourniquets) placed in these stations are usually sufficient to guarantee low levels of non-payment, as is usually the case with metro systems. In addition, operational performance (frequency, regularity, stopping at bus stops, etc.) will also be easier to monitor and enforce in such a system, since there are a limited number of stations and network kilometers. Frequency and regularity will also respond more closely to operators' efforts since the exclusive bus corridors in such a system preclude noisy interaction with private transport and congestion.

However, in more city-wide reforms —such as Transantiago or the SITP in Bogotá— that encompass many more network kilometers and where there are no exclusive bus corridors, monitoring and enforcement of service obligations may be substantially more costly. In this case, more weight should be placed on positive incentives through higher demand risk, rather than fines, penalties and rewards. This is particularly so for service dimensions that are difficult to define and express formally in a contract.

Figure 1: Optimal incentive mechanism and reform type
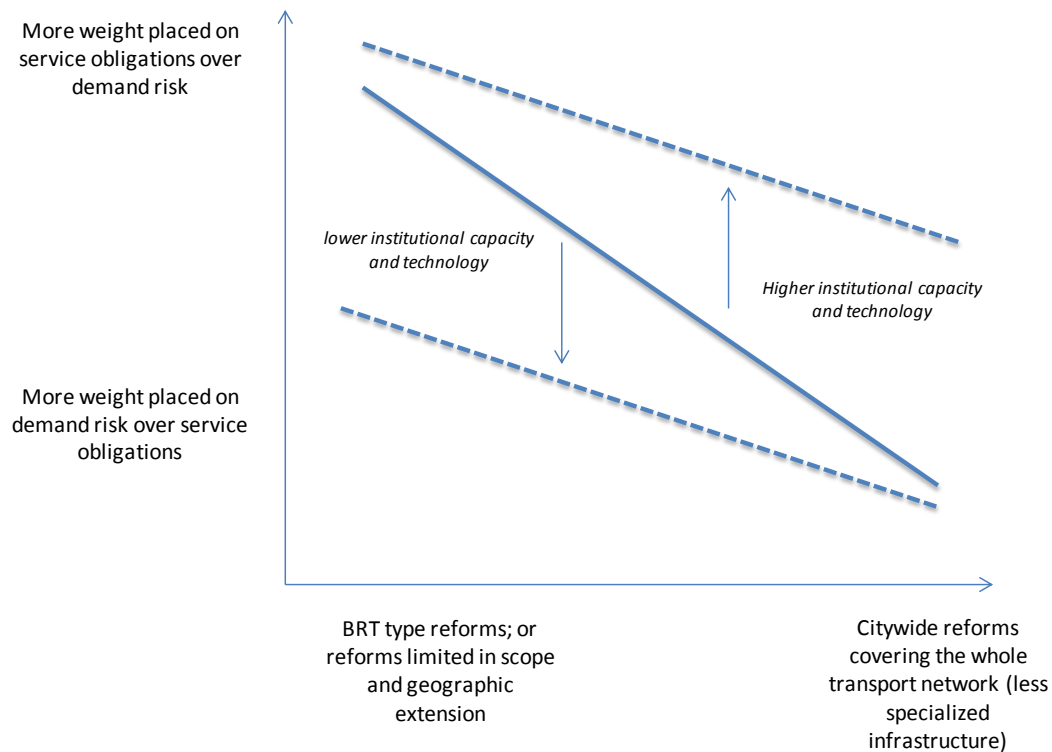
The foregoing discussion can be summarized by Figure 1, which indicates the optimal balance between demand risk (payment mechanism) and service obligations (fines and penalties) in the incentive structure of a concession contract according to the nature of the reform.

As the curve in this figure illustrates, the more ambitious and encompassing the reform, the more important will be to provide incentives through demand risk compared to service obligations. In addition, the institutional capacity and the available technological elements will shift this curve. Thus, for cities with a high institutional capacity and where there are technological elements (TICs) capable of monitoring the system at a low cost, the balance can be tilted more towards service obligations (as in developed country cities). However, as institutional capacity is low and there is less technology available to monitor the system (as in developing countries) it will be optimal to depend even more on demand risk rather than service obligations in order to provide adequate incentives.

The evidence provided by the Santiago case study and the cases reviewed in Annex 3 of this report, seems to confirm the above prediction.

In the case of Transantiago, a citywide reform without specialized infrastructure in most of the network, contracts have been renegotiated several times since the reform was introduced in 2007. In each round of renegotiations, the authorities have pushed for concessionaires to face more demand risk in an effort to provide better incentives for performance. Although these contracts also include a payment mechanism based on operational variables and a complete set of performance obligations linked to fines,

experience has shown that they are not enough to guarantee high quality services and it is important for these operators to face non-marginal demand risk.[14]

However, the case of the Integrated Public Transport System (SITP) reform recently launched in Bogotá, Colombia, is even more interesting.[15] In this case, the revenues for trunk services that operate in the exclusive corridors of the Transmilenio network depend exclusively on operational variables. However, the revenue of zonal services that operate outside this network ―and where monitoring is expected to be more difficult― also depend on the number of passengers transported. Thus, providing some incentives through demand risk (payment mechanisms) was deemed important for services operating outside the BRT network while providing incentives through service obligations (fines and penalties) was considered sufficient for services operating within the BRT network.

3　Additional issues in contract incentives

In transit reforms, there are additional issues that merit particular attention as regards the incentive structure. First, contracts may contain possible clauses that restrict the way drivers can be paid or establish particular labor relations conditions. Drivers are in a sense the operators' 'agent' in the street. The incentives that operators give drivers, and the way concession contracts can restrict these incentives, will alter drivers' performance and the service quality perceived by users of the public transit system. Therefore, in what follows we also review possible contractual clauses related to drivers' labor relations with operators.

Second, the incentive structure also determines the risk borne by operator's compared to users and the State. This risk may affect the investors' access to financial markets, which may be crucial to a reform's success in modernizing and renovating the fleet. This may be particularly important in order to reduce the environmental impacts of public transport systems, such as air and noise pollution. In this report, besides analyzing how different incentive structures affect the risk faced by operators we also review particular clauses that may impinge on the incentives operators may have to renovate the fleet and improve environmental standards.

3.1　Driver's compensation

If there are no contractual constraints it would be expected that the incentives provided to drivers through their contractual relationship, would mimic the incentives faced by operators. It is optimal for operators to align the incentives of their 'agents' (drivers) to their own. For example, if operator's revenue and profits depend on passengers transported, then it is to be expected for driver's to earn a fraction of their income conditional on the number of passengers transported.

In certain cases it may be desirable to limit the extent to which drivers' income depends on passengers transported. For example, in order to prevent these drivers from racing. Thus, in reforms such as Transantiago, where traffic safety was a major concern, clauses were

---

[14]　Details of the evolution of these contracts are presented in Section 4.
[15]　See Annex 3 for a brief review of the contracts in this case.

included in the original concession contracts to prevent operators paying drivers based on the number of passengers transported. This is also the case for the SITP reform in Bogotá.

However, it must be borne in mind that clauses that limit the type of labor relationship that operators can have with their drivers may prevent the correct alignment of incentives between the principal (operators) and their agents (drivers).

This is important because it can lead to two types of problems. First, it may blunt the incentives provided through the payment mechanism or performance standards. For example, a provision requiring an operator to control fare payment may not be effective if this operator cannot in turn link driver's income to effective demand. Second, it can also lead to renewed informal relationships between operators and drivers, a phenomenon that reform usually tries to overcome.

It is interesting to note that in the case of Transantiago, the latest contracts eliminated these clauses.

3.2 Incentive mechanisms for fleet renovation and environmental performance

As mentioned earlier the risk characteristics faced by operators may affect their access to private sector financing for fleet renovation which is an important consideration when concerned with the environmental impact of the transport system. Therefore, the incentive structure of the concession contract will determine in part the viability of the reform in this respect. However, contracts usually contain a series of other clauses directly related to fleet renovation and other environmental issues. In many instances these are quality standards that must be met by any potential concessionaire; for example, when the contract establishes a minimum fleet of certain types of buses.[16] However, in other cases, contracts are more flexible and operators are induced to generate environmental benefits in more indirect ways. For example, when there are monetary benefits or extensions of the horizon of the contract when certain fleet renovation or other environmental beneficial actions are taken. Why not impose these actions or fleet targets as service obligations and tender contracts with these obligations included? What potential benefits are there from generating incentives for these actions that may or may not be undertaken? The main reason for this indirect approach is when the authorities are uncertain as to the costs of these actions. If they are imposed as service obligations in the contract they may increase the costs of the reform beyond what was originally planned. Therefore, by giving operators incentives the potential financial costs for the reform are known and the actions will only be taken if the benefits are above the private (ex–ante unknown) costs for operators.[17]

---

[16] For example, in Colombia, buses used in BRT and integrated systems are required to meet minimum Euro IV standards. To encourage the use of cleaner technologies, the Environment Ministry reduced the import cost of heavy vehicles (buses and trucks) with clean technologies (electric, hybrid and CNG) from 15% to 5% (Decree 2658 of 2011).

[17] This argument has much in common with the well known benefits of quantitative targets as opposed to prices under uncertainty (see Weiztman, 1974).

4   Contractual clauses in practice and their effects: the Santiago and London experiences

In this section we review the experience of Santiago and London in more detail. We emphasize the effects of different contractual mechanisms and their evolution through time. The aim of this section is to illustrate the conceptual issues discussed above with concrete examples from these two experiences. In Annex 3 we briefly review three other experiences in Latin America.

4.1   The experience of Santiago (Transantiago)

An ambitious reform of the public transport system, Transantiago, was introduced in Santiago, Chile, in February 2007[18]. This reform completely changed the route structure, the fare payment method, the contractual relationship with operators as well as many other dimensions of the city's public transport system. The results were immediate and disastrous.

After the reform was implemented on February 10th, 2007, a date known as the 'Big Bang', due to the complete overhaul of the old system from one day to the other, it was clear that public transport supply was insufficient for a city where almost 56% of motorized trips used public transport. Waiting times and total travel times increased substantially, congestion was notorious at bus stops, inside buses and in the metro system, and users were forced to make costly and unpopular transfers between transport modes and vehicles in order to complete their trips.

The consequence was a social and political upheaval not seen in the country since the return to democracy almost 20 years earlier. This led to several changes in the system between 2007 and 2011 particularly in the contracts with operators. During this period at least three major contractual renegotiation processes were undertake.

In this Section we present a brief description of the original reform, the main characteristics of the concession contracts and a general overview of the changes in these contracts between 2007 and 2011. We then analyze in more detail the contractual clauses related to operator's payment mechanism, labor relations with drivers, and fines and other penalties. We also review the experience with these clauses and how they evolved over time in response to the problems encountered.

4.1.1   General description

One of the most important characteristics of the Transantiago reform was the transformation of the route network from a point to point, non-integrated and overlapping route scheme to an integrated trunk and feeder system. The city was divided into 10 zones, where local and feeder services would operate under a franchise arrangement in each one, plus 5 trunk operators that would provide longer services that crossed the city.[19] In addition, the metro was to be integrated into the system, providing an additional "backbone" trunk

---

[18]   For a more detailed description of this reform, see Gómez-Lobo (2011).

[19]   After the tendering process of 2004 only 9 concession contracts were finally signed for the feeder zones since there were no operators interested in the central local area services (Zone A).

service for the new system. Route overlap in the new network was minimal and even when it occurred the overlapping services were usually operated by the same concessionaire to avoid competition in the streets

In order to integrate fares among different services and modes an electronic pre-payment card was introduced. For fare integration to work, revenues must accrue to a centralized agency that then distributes this income among operators according to the terms of their respective contracts. To this end, the Transantiago Financial Administrator (AFT) was created and tendered to a consortium of Banks and a technological firm in 2005.

Another institution, called SIAUT, was designed to provide customer information, including route maps and a web page application to help users plan their trips.

To curb air pollution emissions and noise, as well as to offer users the benefits of modern low floor buses, the fleet would be renovated.[20] New buses had to meet Euro III or IV Technical Specification standards. However, to keep costs down and to allow existing operators to participate in the new system, only a fraction of buses had to meet the new Technical Specification standards. In fact, only 58% of the total fleet met these standards in February 2007. In addition, to reduce costs even further, a large fraction of the new fleet was designed to be high capacity (160 passengers) articulated buses.

4.1.1.1   Contracts, risk-sharing and incentives

The key variable that determined the bi-monthly payment to operators was the *PPT* (Payment per Passenger Transported) and was the main bidding variable in the competitive tendering process held in 2004.

Although actual payment would be the result of multiplying the *PPT* by the number of passengers transported every two weeks, a complex mechanism was introduced to reduce the demand risk faced by operators. The result was that payment to operators would be based on a fixed pre-established demand estimation (called the "reference demand" and included as an Annex to each contract). In practice, operators faced negligible demand risks amounting to 10% of the <u>deviation</u> between the reference demand and effective demand. Thus, if demand fell 10% below the reference demand in a given month, the payment formula would increase the PPT in the next month to compensate for this drop and operators would only loose 1% of projected income. There were also other mechanisms in the contract to protect the cash flow of operators, particularly those that would be making investments in new buses. Although these mechanisms were meant to lower financial risks and thus enable operators to obtain funding for fleet renovation, they blunted incentives to cater to demand.

An operational plan would be established every three months determining the services and frequency that each operator had to meet each period of the day. Income did not depend directly on compliance with the operational plan nor, as mentioned above, on passengers transported, but rather on penalties defined in the contract. A long list of fines for different infringements was defined in the contracts. However, if an operator accumulated more than

---

[20]   The existing fleet was composed of high floor buses that required climbing several steps when boarding. However, once inside, the floor was level and there was ample seating capacity (80 seats).

6.000 UF[21] of fines in a twelve-month period, the authorities were obliged to terminate the concession contract.

Contracts also stipulated that drivers had to have formal labor contracts and could not be paid according to passengers transported. This was introduced to curb driver's incentives to compete for passengers in the streets with the accompanying safety risks. The downside is that it also eliminated all incentives for drivers to control non-payment and cater to demand.

4.1.2   Overview of the problems encountered in 2007 and its aftermath

After several postponements, the reform in its full form was finally implemented on February 10th 2007. The route network was changed overnight, fare integration was introduced and the new payment mechanism for each concessionaire came into operation.

Problems arose immediately and there was considerable chaos in the city. During the first few months after February 10th it was not possible to determine how many buses were operating nor their frequency or regularity. However, one thing was certain: supply was insufficient to cover demand. Bus stops were overcrowded, passengers fought their way into available buses, waiting and travel times increased and there was a large degree of dissatisfaction with the new system.

Faulty contract design was one of the major reasons for the problems encountered (Gómez-Lobo, 2012). As will be discussed in more detail below, operators had scant incentives to meet the operational plan. Payment to bus operators in the original contracts did not depend on seat-kilometers supplied or on passengers transported.

Although there were penalties for non-compliance these were ill defined, relatively low in monetary terms and difficult to enforce given the lack of a monitoring technology.[22] Furthermore, as discussed above, the contracts had a limit of 6.000 UF in penalties during a 12 month period before the authorities were forced to terminate the contract. This was useless in practice as an incentive mechanism. It was unclear who would operate the services before a new concessionaire could be found. This transition could last months and would imply leaving users without services during that period. Thus, terminating the concession contract was not a viable option for the authorities once the new system was in operation and thus paradoxically this clause restrained the authorities' capacity to pass fines and enforce the operational plan. It was a non-credible threat that eventually worked to the operator's benefit.

Thus, the system lacked the "carrot" incentives of a competitive system (whereby operators' income depends on passengers transported) and the "stick" incentives of penalties. Unsurprisingly, operators found it profitable to reduce costs by lowering supply since income was unaffected. Non-compliance with the operational plan was the norm during the first period of the reform.

---

[21] The UF is an indexed monetary unit commonly used in contracts in Chile in order to avoid the effects of inflation. One UF is worth approximately US$ 46 at the time of writing.

[22] Note, for example, that according to the contracts an operator could provide just 60% of the required frequency and not risk being fined.

The original contract design, particularly fines and penalties, assumed a strong institutional capacity to monitor and enforce these contracts. However, the institutional structure was quite weak in February 2007. There was a special unit within the Ministry of Transport and Telecommunications charged with these tasks. However, it initially lacked the financial and human resources to properly undertake these activities.

Unsurprisingly, the above problems led to a series of contract renegotiations. In the next sections we will present the details of the contract clauses, the problems encountered and the evolution of these clauses over time.

All contract documentation is available at the following website: http://www.coordinaciontransantiago.cl/corporativo/index.php?option=com_content&view=article&id=38&Itemid=23. From now on, the original tendering documents for operators (*Bases de Licitación de Vías*, 2003) will be referred to as the *Bases*.

4.1.3 Detailed review of clauses related to operator's payment, their effects and evolution[23]

As mentioned above, payment to operators was initially based on the tendering variable *PPT* (Payment per Passenger Transported) multiplied by the number of passengers transported in a concessionaires services (Clause 3.5.2.1.2 of the *Bases*). Formally,

$$PC_t = PPT_t \cdot Q_t \qquad\qquad (1)$$

Where PC_t is the payment due in period t, PPT_t is the *PPT* in period t, and Q_t is the number of passengers transported by the operator in period t. The PPT_t of each concessionaire was the *PPT* tendered by the company at the awarding stage in 2004, adjusted for input price variations between that date and the payment period t (clauses 3.5.2.4 and 3.5.2.5.3 of the *Bases*).

However, the PPT_t was also adjusted each period to shield operators from demand risk. To this end, the following formula for the evolution of this variable was included in clause 3.5.2.5.1:

$$PPT_t = PPT_{t-1} \cdot \frac{0.9 \cdot Q^r_{t-1} + 0.1 \cdot Q_{t-1}}{Q_{t-1}} \qquad\qquad (2)$$

where PPT_{t-1} is the *PPT* of the previous payment period, Q^r_{t-1} is the reference demand in period t-1 and Q_{t-1} is the effective demand in period t-1.[24] The reference demand was a demand projection estimated by the authorities and included as an Annex to the tendering documents. The reference demand was set for the twelve months of 2005 and would then grow by 1.7% a year.

---

[23] A complementary description of the mechanisms described in this section is provided by Beltrán, Gschwender and Palma (2012).

[24] This formula applied for the first year of operation. For the second year and beyond, the formula would be Q^r_{t-1}/Q_t for those specific services where commercial speeds fell by 1 km/hour from one year to the other, and the above formula for the rest. This would dampen even more the demand risk faced by operators.

When effective demand was different from reference demand, then the *PPT* would be adjusted. However, in order for this adjustment to compensate the demand drop, reference demand had to be updated for the following months according to the following formula:

$$Q^{r^t} = Q_j^{r^t} \cdot \frac{Q_{t-1}}{Q_{j-1}^{r^{t-1}}} \quad \forall \, t > j \qquad (3)$$

Where Q^{r^t} is the new reference demand for month t ($t > j$) updated in month j, $Q_j^{r^t}$ is the old reference demand for month t, Q_{t-1} is the effective demand in month j-1, and $Q_{j-1}^{r^{t-1}}$ is the reference demand for month t-1 in month j-1. In order to understand this formula, let's assume we are in month j. Last month effective demand was below the reference demand established for that month. Therefore, the *PPT* was adjusted according to formula (2). In addition, all the reference demands from month j onwards are also updated to internalize the drop in demand.

The impacts of the above dynamic formulas are very difficult to understand at an intuitive level and as such they may be rightly criticized for making the concession contract quite opaque. In order to grasp their implication we must use a simulated example. Let's assume that reference demand is initially 100 for each of twelve months and the initial PPT is $10 per passenger. Therefore projected income is $1,000 for each of the twelve months. Now, assume that in month one, demand falls by 10% permanently.

Table 2 shows the impact on all variables including effective income. It can be seen that except for the first month, income only falls by 1% a fraction of the fall in demand. Therefore, in essence the payment mechanism established in the original contracts shielded operators from most demand risk. They only faced a demand risk equivalent to 10% of the demand <u>shortfall</u> over projected demand. If demand fell by 50% over projected or reference demand, the effective income of operators would only fall by 5%, after the first month.

There were also two other mechanisms to protect concessionaires from demand risk. One was a preferential payment mechanism in case system wide revenues in a given period were insufficient to pay all operators (clause 3.5.2.1.2 of the *Bases*). These favored trunk operators with modern fleets. The other was a present value of income mechanisms, whereby if an operator at the end of the concession period had not earned (in present value terms) an amount equivalent to the original *PPT* tendered (adjusted only for input price variations) times the original reference demand, then the concession period would be extended to allow the operator to earn the revenue shortfall (clauses 3.5.5 and 3.5.5.1).

Table 2: Simulation of payment formulas

Month	Initial reference demand	Initial PPT	Projected income	Effective demand	Adjusted reference demand	Adjusted PPT	Effective income	% Change in Q	% Change income
1	100	10	1000	90	---	10	900	-10%	-10%
2	100	10	1000	90	90	11	990	-10%	-1%
3	100	10	1000	90	90	11	990	-10%	-1%
4	100	10	1000	90	90	11	990	-10%	-1%
5	100	10	1000	90	90	11	990	-10%	-1%
6	100	10	1000	90	90	11	990	-10%	-1%
7	100	10	1000	90	90	11	990	-10%	-1%
8	100	10	1000	90	90	11	990	-10%	-1%
9	100	10	1000	90	90	11	990	-10%	-1%
10	100	10	1000	90	90	11	990	-10%	-1%
11	100	10	1000	90	90	11	990	-10%	-1%
12	100	10	1000	90	90	11	990	-10%	-1%

In essence, the payment mechanism basically implied a fixed payment irrespective of operational variables (seat-kilometers supplied) or the number of passengers transported. Plus, there were other mechanisms to guarantee a minimum income in each payment period (preferential payment mechanism) or along the total duration of the concession (present value of income guarantee).

The incentive structure provided in the original contracts was catastrophic. Operators did not have incentives to comply with the operational plans. Absent an effective enforcement mechanism, the fines and penalties defined in the contract ―that will be discussed further below― were insufficient to guarantee the supply of services. Operators earned a fixed income irrespective of the number of passengers transported or the number of buses they supplied. In fact, during the first few months after the reform was introduced, not even the 10% demand shortfall risk described above was used. Since many buses were not yet equipped with the electronic payment card validation equipment and the clearing system (confirming that payment to each operator was in fact based on the number of passengers transported) was not yet audited, the authorities paid operators only based on reference demand.

The result was an undersupply of buses that contributed directly to the chaos and negative perception of the system during this initial period. Clearly, the payment mechanism needed to change in order to provide incentives to comply with the operational plan.

4.1.3.1 *First contract renegotiation 2007- increasing operator incentives*

The first contract renegotiation was undertaken several months after the reform began and the formal changes were signed in mid-2007.[25]

One of the main changes in the payment mechanism introduced during this period was an increase from 10% to 35% in the demand shortfall risk assumed by operators.[26] In other words, the evolution of the *PPT* would now follow the formula:

$$PPT_t = PPT_{t-1} \cdot \frac{0.65 \cdot Q_{t-1}^r + 0.35 \cdot Q_{t-1}}{Q_{t-1}} \tag{2´}$$

Now, a permanent demand shortfall of 10% of reference demand would imply a permanent reduction of 3.5% of revenues. There was a fixed amount added to the operator's payment to compensate for a permanent fall in income from this change, but in the margin it was expected that operators would have more incentives to cater for demand due to the new formula that affected variable income.

However, the above change did not seem to alter operators' behavior much. In part, this may have been due to the compensation through a fixed payment of the drop in revenues from this modification compared to the original contract. Alternatively, the revenue risk assumed may have been too small compared to the cost of supplying more services and may thus have been insufficient to change operators' incentives. The plausibility of this last conjecture is reinforced when one considers that operators had exclusivity of service (although this started to change by mid-2007) in their operational area or routes and service overlap was minimal. Therefore, operators had a somewhat captive demand that would remain fairly stable even if frequency of services were reduced.

The other major change introduced in mid-2007 provided more powerful incentives. Effective payment to operators was made conditional on an Index that measured the compliance with the operational plan. Payment to operators would now be:

$$PC_t = ICPH_t \cdot PPT_t \cdot Q_t \tag{1´}$$

where *ICPH* was the 'Capacity Hour Compliance Index'.[27] This index was calculated as follows:

> A day was broken down into half-hour intervals, denoted by *i*;
> During each half-hour interval, the capacity supplied (measured as bus capacity) required according to the operational plan was calculated, $PH_{plan\ op,i,t}$;[28]

---

[25] The exact date of formal contract modifications differed for each operator as each signed the documents on different dates. For Trunk 1 services, for example, the contract modification was signed on November 9th 2007, although it was operational since August of 2007. This modification is available from the web page cited at the beginning of this Annex.

[26] The contract modification documents are more complex than what the following discussion would suggest. There were other clauses introduced that in practice had no effect since the conditions for their application were not met. In what follows we distill the most salient features of these modifications.

[27] This is the author's translation of '*Indice de Cumplimiento de Plazas Hora*'.

[28] Bus capacity per hour was chosen as the capacity variables since there was a wide array of different buses in operation, each with a different capacity for transporting passengers.

The effective capacity supplied by an operator each half-hour was calculated using GPS signals for each bus, PH_{it};

Then, the *ICPH* was defined as:

$$ICPH_t = \frac{\sum_i IPCH_{i,t} \cdot PH_{Plan\ op,i,t}}{\sum_i PH_{Plan\ op,i,t}} \tag{4}$$

$$ICPH_{i,t} = \begin{cases} 1 & if\ \dfrac{PH_{i,t}}{PH_{Plan\ op,i,t}} \geq 0.94 \\[2em] \dfrac{PH_{i,t}}{PH_{Plan\ op,i,t}} & if\ \dfrac{PH_{i,t}}{PH_{Plan\ op,i,t}} < 0.94 \end{cases}$$

In other words, the *ICPH* was a weighted average of the compliance with the operational plan each half hour over the payment period. Full compliance in a half-hour period required an operator to supply at least 94% of the capacity required by the operational plan. In addition, operators had to compensate the shortfall (between 94% and 100%) by providing 120% of the capacity shortfall (if it the shortfall was in a peak-period) or 110% of the capacity shortfall (if it was in a non-peak period) during another half-hour over what was required by the operation plan during the other half-hour.

The introduction of the ICPH had an immediate effect in the number of buses in operation. Figure 2, taken from Beltrán, et al (2012), shows how the number of buses in operation increased sharply as the new ICPH index came into operation in August 2007.

Figure 2: Number of buses in operation: July 2007 to December 2007.
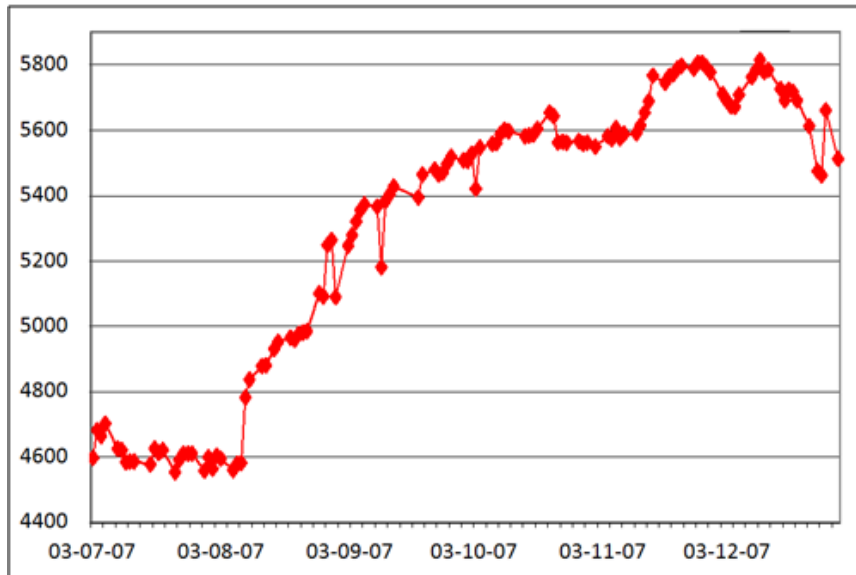
Source: Beltrán, Gschwender and Palma (2012)

The performance of the system increased after the *ICPH* was introduced. Users' waiting and travel times started to fall and the surveys showed an improvement of passengers' perception of the system. However, by mid 2008 the authorities realized that the *ICPH* was not enough to guarantee firms' compliance with the operational plan.

As operators gained experience with the new index, they started gaming the system. They could do this through several ways.

First, to count towards the *ICPH* index, a bus had to be moving (at a speed over 2 km/hr) and had to have at least 10 GPS signals in the half-hour interval. However, the ten signals could be achieved in 5 minutes. A bus arriving at a terminal during the first 15 minutes of a half-hour period could be replaced by another bus during the rest of the half-hour in the same service. The GPS signals would indicate two buses in operation on that route in the half-hour while the effective capacity was only equivalent to one bus.

Second, the system did not provide incentives for a quick turnaround of buses at the terminal ends of routes. Once the bus counted in the half-hour period, drivers could over-rest at the terminals without a financial penalty.

Third, there was no assurance that a bus whose GPS signal indicated it was moving, was really on the service route.[29] Furthermore, with the monitoring technology in place, the index could only be calculated for the entire fleet of each concessionaire at the aggregate level. It was not possible to calculate this index at the route or service level. This implied that operators could over provide capacity in low cost services and under-provide capacity in high cost services during a given period without being penalized.

Fourth, an operator had the incentive to provide capacity with larger buses at a lower cost but implying a lower effective frequency of services.

Finally, the 94% rule could be used to 'smooth' the operational plan. Less services could be provided during peak hours, where costs are higher, and the shortfall compensated by offering 120% of the shortfall during non-peak hours. Thus, effective supply compared to required supply started to look as in Figure 3.

---

[29] This possibility was somewhat improved by a change in the calculations which required a bus to have effective validations (i.e. passengers boarding) in order to count for the index.

Figure 3: Illustration of effective and required capacity supply by mid 2008
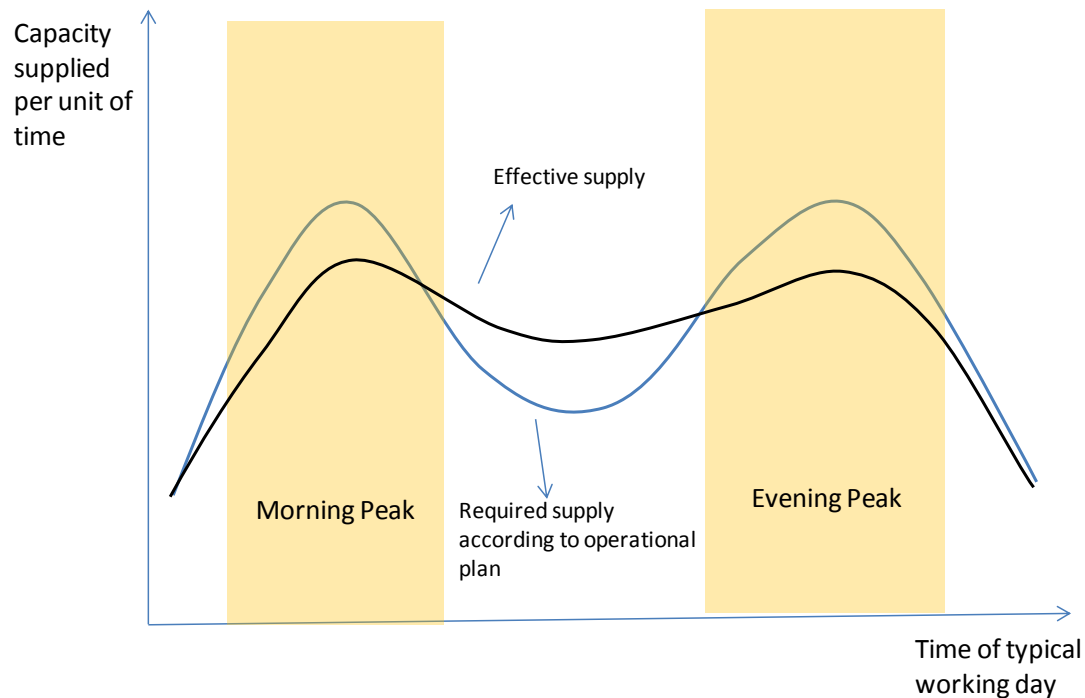
It was clear that some operators were able to maintain a high *ICPH* while at the same time exhibiting unsatisfactory frequency levels, particularly during peak periods.

4.1.3.2   *Improvements in the incentive structure in 2008*

Two events happened in mid-2008 that allowed for an improvement in the incentive structure. First, in July of that year a new fleet monitoring system became operational.[30] This system allowed the authorities to track buses much more closely. Now, the number of buses on route, their headway and effective frequencies could now be monitored for each service in real time.

A second event that year gave the authorities the ability to apply new indicators and thus use the new fleet tracking system to improve the systems operational performance. Since its inception in February 2007, the transit system had been running a financial deficit.[31] During 2008 the deficit was being funded by a private credit to the system provided by the Inter-American Development Bank.[32] This credit was deemed unconstitutional by the constitutional court in September 2008. From that date the system had to be funded by a

---

[30]   See Beltrán, et al (2012) for more details.

[31]   The financial deficit was close to US$825 million in 2008, representing close to 50% of operational costs of the system. By 2011 this deficit had declined to US$ 722 million representing close to 40% of operational costs.

[32]   It had to be a private credit since Congress had not approved this credit or other subsidies.

special provision in the Constitution that allowed the Executive Branch to spend up to 2% of the yearly budget as an 'Emergency Fund' without Congressional approval.[33]

As these were emergency funds reserved for very special circumstances it was deemed that they could not be used to pay operators for 'services that had not been provided'. Thus, two new indices were unilaterally introduced by the authorities and discounts to operators' payment were made for non-compliance with these indices.[34]

The new two indices were the Frequency Compliance Index (*ICF*) and the Regularity Compliance Index (*ICR*). The first measured the effective frequency on a route-direction over the required frequency according to the operational plan. The second measured the regularity with which services were rendered and varied from 0 to 1 according to the coefficient of variation of bus dispatched on each route.[35]

These indices were measured on all routes and periods. However, the payment discounts —which varied according to the level of performance of each index— were applied sparingly on a certain percentage of operators' routes in order to avoid bankrupting operators.

These new indices and their accompanying income penalties gave the authorities new teeth to control the system. During the second semester of 2008 there was a marked increase in the performance of the system, reduction in waiting and travel times, and positive perception on the part of users. The systems average *ICF* during morning peak hours increased from close to 75% in August 2008 to over 95% in August 2009 (Beltrán, et al, 2012). During the evening peak this index passed from less than 70% to close to 90% during the same period. The *ICR* evidenced a similar evolution as the *ICF* during that period.

4.1.3.3 *Contract renegotiations in 2009*

The success of the *ICF* and *ICR* indices and the shortcoming of the *ICPH* index gave way to a new round of contract renegotiations in 2009. The aim was to change the contracts in several ways, including the formal introduction of these indices. During this period the authorities had realized that the difference between the *ICPH* and the *ICF* index could be explained by the lack of a kilometer variable in the *ICPH*. This gave rise to a new index called the Compliance with Capacity Hour Kilometer Index (*ICPHK*). This index was incorporated in the new contracts signed in the second semester of 2009 and was formally introduced as follows:

$$DSC_t = PC_t \cdot \left(1 - \frac{\sum_i \alpha_{i,t} \cdot PH_{Plan\ op,i,t}}{\sum_i PH_{Plan\ op,i,t}}\right) \tag{4´}$$

---

[33] The Constitution states that the funds could be used in an emergency to maintain the provision of public services, which was interpreted to include public transit services.

[34] Since this was not a renegotiated change, the remnant of payment due to operators according to the contracts was considered a debt of the system with operators that would be paid when the system had sufficient funds. However, everyone was aware that these payment shortfalls would most probably never be paid. Naturally, the concessionaires challenged the discounts applied but the courts upheld the authorities' interpretation.

[35] Detailed definitions and measurement issues can be found in Beltrán, et al (2012).

where DSC_t is now the discount applied to the payment of an operator in period t and

$$\alpha_{i,t} = \begin{cases} 1 & if \ \dfrac{PH_{i,t}}{PH_{Plan\ op,i,t}} \geq 1 \\[2ex] \dfrac{PH_{i,t}}{PH_{Plan\ op,i,t}} & if \ \dfrac{PH_{i,t}}{PH_{Plan\ op,i,t}} < 1 \end{cases}$$

where $PH_{i,t}$ and $PH_{plan\ op,i,t}$ are now defined in terms of <u>capacity-kilometers</u> offered and required according to the operational plan. Notice also, that the 6% buffer in the *ICPH* index was eliminated. Finally, it is also important to note that the *ICF* and *ICR* index were also included in the new contracts in order for the authorities to control and enforce the operational plan in specific routes.

The contracts signed in 2009 made the payment mechanism similar to a system of seat-kilometer or capacity-kilometer that is found in many transit systems around the world (as gross cost contracts).

However, by late 2009 and early 2010 a new problem started to arise. Under the pressure to meet the more demanding index defined in terms of capacity-kilometers and where any shortfall from 100% compliance from the operational plan would penalize payments, operators started giving drivers orders to meet certain time constraint for completing their services irrespective of the quality of service provided. Thus, drivers began concentrating on arriving at the terminal end of their route by a certain time and not stopping at bus stops to pick up passengers. This became an increasing problem through 2010.

Another issue that was not resolved by the contract change in 2009 was non-payment. Non-payment on buses fluctuated between 11% and 18% of passengers during 2007-2009.[36] The authorities had tackled this problem by creating guarded boarding stations at key bus-stops where users had to pay before boarding buses. However, the number of such stations was limited (155 by the end of 2009). The other alternative was to enforce payment through inspectors and fines. However, there were several limits to this strategy. First, inspectors required special equipment to read the electronic payment cards in order to ascertain whether a particular passenger paid or not. This equipment was not available until mid-2008 and so payment enforcement only started that year. Second, by law inspectors could not fine users unless they were accompanied by policemen. This limited the scope of the enforcement strategy as the number of available police officers for this task was limited.

4.1.3.4 *Contract renegotiations in 2011- controlling operator behavior with enhanced incentives*

When the transport tariff started rising in 2010, non-payment became rife increasing to over 30% on some services. By this time it became clear that the most effective way to limit non-payment was to give operators some incentives to control this behavior.

Therefore, the new authorities took charge of the system when the government changed in March 2010 decided to undertake a new contract renegotiation process. This process

---

[36] In the metro, non-payment was very low since there were guards in each station.

culminated with the signing of new contracts in late 2011 that became operational earlier this year.

As far as the payment mechanism is concerned, the new contracts increased the demand risk faced by operators. It was deemed that this was the only way to give operators incentives to cater to passengers (and stop at bus stops) and also control non-payment.

The new payment mechanism was structures as follows (Clause 5.4.1 of the new contracts):

$$Y_t = PPT_T \cdot q_t + PK_T \cdot \left(km_t + 0.33 \cdot (kme_t + kma_t) \right) \cdot ICT_t - D_t + O_t \qquad (4'')$$

where Y_t is the payment in period t[37], PPT_T is the payment per passenger in month T, q_t is the number of passengers transported in payment period t, PK_T is payment per kilometer in month T, km_t is the number of kilometers required in period t, kme_t and kma_t are additional and special kilometers required in period t[38], ICT_t is the Capacity Compliance Index in period t, D_t are additional discounts in period t, and O_t are other payments in period t.

The *ICT* is very similar to the *ICPHK* index defined earlier and calculated every half hour:

$$ICT_t = \frac{\sum_{i=1}^{n} \left[Min\left\{1, \frac{PKH_{i,t}}{PKH_{op,i,t}}\right\} \cdot PKH_{op,i,t} \right]}{\sum_{i=1}^{n} PKH_{op,i,t}}$$

where $PKH_{op,i,t}$ is the capacity-kilometers established in the operational plan in the half-hour i in period t and $PKH_{i,t}$ is the actual capacity-kilometers undertaken.

Several comments are in order with respect to this new payment mechanism. First, the payment was split into a fixed part and a variable part. The fixed part is made conditional on the number of kilometers offered and is thus akin to a payment scheme based on operational variables.[39] The variable part, which represents around 70% of payment, is just the *PPT* times the demand.

Although this payment mechanism increased the demand risk faced by operators another mechanism in the contract was introduced to limit this risk (clause 5.4.1.7 of the new contracts). Once a year, the average number of passenger per kilometer offered are to be calculated as:[40]

---

[37] In this equation the sub index T denotes de month, while t denotes the payment period. There are two payment periods per month.

[38] Due, for example, to special public events. These kilometers are also included in km_t which implies that they are remunerated at 33% more than the PK_T.

[39] The new contracts in 2009 had also introduced a fixed payment in the mechanism.

[40] In order to save space in what follows we do not present the exact formulas in the contract, but rather a more compact representation of them.

$$IPK_t = \frac{\sum_{i=t-24}^{t-1} q_i}{\sum_{i=t-24}^{t-1}(km_i \cdot ICT_i)} = \frac{Q_t}{KM_t}$$

Considering that there are two payments per month, the above formula effectively calculates the number of passengers per kilometer offered during the last 12 months.

Then the following quantities are calculated:

$$IPK_{inf} = \overline{IPK} \cdot (1 - \alpha)$$

$$IPK_{sup} = \overline{IPK} \cdot (1 + \alpha)$$

$$Q_{inf} = IPK_{inf} \cdot KM_t$$

$$Q_{sup} = IPK_{sup} \cdot KM_t$$

where $\overline{IPK}$ is the passenger per kilometer parameter set in the contract for the operator. Then once a year the following payment is made the operator:

$$AIPK_t = \begin{cases} PPT_{t-1} \cdot (1 - \beta_1) \cdot (Q_{inf} - Q_t) & \text{if } Q_t < Q_{inf} \\ 0 & \text{if } Q_{inf} < Q_t < Q_{sup} \\ PPT_{t-1} \cdot (1 - \beta_2) \cdot (Q_{sup} - Q_t) & \text{if } Q_{sup} < Q_t \end{cases}$$

Although, each operator may have different parameters α, β_1, β_2, typical values are 0.03, 0.35 and 0.85, respectively. In effect what this mechanism does is to reduce demand risk. For example, for the parameters just presented, the operator assumes a demand risk equal to the *PPT*, for variations of plus and minus 3% of passengers per kilometer offered. If demand falls below that 3% threshold, he is reimbursed once a year for 65% of the shortfall in demand from this threshold. If demand is above the 3% threshold, then the operator must return to the system 15% of the excess demand (valued at the last *PPT*).

Thus, the new contracts increase demand risk but not completely. In fact, following the example presented in Table 2 above —and assuming that initially both mechanism are calibrated so that for demand equal to reference demand or $\overline{IPK}$, both mechanisms pay the same amount to the operator— a permanent drop of 10% in demand (assuming kilometers offered remain constant) implies a variable income reduction of around 5.4%. However, if the fixed part of the payment formula is taken into account then the variability in income is even less. There is of course the financial cost of being paid only once a year for a demand shortfall in the new mechanism, while the old mechanism compensated operators with a one month lag.

For a demand increase operators get to keep a much higher percentage of the marginal income than what they have to return from demand increases. Therefore the mechanism is not symmetric.

Finally, it is important to note that other income guarantee mechanisms of the original contract, such as the minimum net present value of income guarantee over the life of the contract described above, were eliminated.

Unfortunately, the last contract change came into effect in the middle of 2012 so it is too early to evaluate its impact. However, from the description presented in this section it can be concluded that there has been a marked evolution towards more demand risk in the payment mechanism of contracts. In spite of this, operational variables are still crucial in the payment mechanism. Also, some type of demand risk protection has also been unavoidable in the evolution of these contracts.

4.1.4   Clauses related to labor relations with drivers, their effects and evolution

In the original contracts there was a clause (4.3.1.1 of the *Bases*) that explicitly prohibited operators from paying drivers as a function of the number of passengers transported. The intention was to reduce competition for passengers in the streets, which, as mentioned above, was one of the main drivers for reform.

However, this clause also had the adverse effect of stifling any incentives drivers may have had to cater for demand and control non-payment. In the latter part of 2009 and 2010 drivers –consistent with the operators' need to meet operational variables such as the *ICPHK*- did not even have strong incentives to pick-up passengers in the stops unless other passengers were alighting.[41]

Therefore, as contracts evolved and placed more demand risk on operators, it was necessary to change this clause so that driver's payment structure could be accommodated to induce the right behavior from the principal's point of view (operators). Thus, in the new contracts signed in 2011, this restriction was eliminated.

The possible problem related to this change is that individual drivers may start racing in the streets, repeating old practices, if operators structure their incomes on the number of passengers transported. The lower overlap of services from different companies may improve this, since operators may give drivers less incentives to race for passengers when their services do not compete with those of other companies.[42] Unfortunately, at the time of writing it is too early to tell whether the new contracts have worsened driving behavior or not, since the new contracts became operational only a few months ago.

4.1.5   Clauses related to fines and penalties, their effects and evolution

Like in all concessions, the original Transatiago contracts had a long list of fines and penalties related to service quality. In clause 3.6.1.1 of the *Bases* fines are defined according to the seriousness of the fault. Thus, there was a 10 UF[43] fine for every time an

---

[41]   Alighting passengers would complain (sometimes loudly) inside the bus if drivers did not stop at their requested stop so skipping stops when there were alighting passengers was rarer.

[42]   One possible measure to ameliorate this problem is to force operators to rotate drivers among services and schedules. However, this requires additional monitoring capacity on the part of the authorities.

[43]   The U.F. (*Unidad de Fomento*) is a monetary unit used in most contracts in Chile that varies daily according to last month's inflation rate. Thus, it maintains its real value through time, making all

operator commits minor offences (such as a driver not wearing a uniform; unclean buses; etc.). There were also events fined with 20 UF, 30 UF, 40 UF, 100 UF and 200 UF and other specific fines for other concepts. For the purpose of this review, we will describe in more detail the fines related to operational underperformance, since these fines were meant to control service quality in lieu of a direct payment mechanism linked to performance.

Fines for operational underperformance included the following:

- A 10 UF fine for every time a bus unjustifiably denied transporting passengers even though it had capacity to do so (3.6.1.1. (1º b)).

- A 10 UF fine for each service where the transport capacity was below that established in the operational plan (3.6.1.1. (1º g)).

- A 30 UF fine for every time an operator accumulated 5 fines of 10 or 20 UF in a six month period or every time an operator accumulated 10 of these fines throughout the concession period (3.6.1.1. (3º c)).

- A 100 UF fine for every bus found, unjustifiably, operating out of the route considered in the operational plan (3.6.1.1. (5º b)).

- A 100 UF fine for every time an operator accumulated 5 fines of 30 UF or 40 UF in a year (3.6.1.1. (5º d)).

- A 100 UF fine every time the average occupancy rate was above a certain threshold, except if the operator was providing the maximum transport capacity defined for the route (3.6.1.1. (5º g)).

- A 200 UF fine for every day that an operator dispatched less than 60% of the required frequency in a two hour period in a certain route (3.6.1.1. (6º a)).

There were also fines for lack of service regularity (3.6.1.1. 7º) measured according to a formula defined in Annex 8 of the *Bases*. There was an additional mechanism establishing fines for non-payment, using a benchmarking technique among companies with over 5% of non-payment (3.6.1.1. 9º).

Finally there was a provision in clause 3.6.1.2. (d) of the *Bases*, whereby if a concessionaire accumulated more than 6,000 UF in fines in a twelve month period the authorities were obliged to terminate the concession.

As we can see all major operational problems that were encountered at some stage with Transantiago (low capacity in operation, low frequency, high non-payment, not stopping at bus stops, etc.) were covered by the penalties established in the contract. In addition, the

payments defined in this monetary unit constant in real terms. At the time of writing one U.F. was worth around US$ 46.

risk of losing the concession should have been a powerful incentive to avoid operational underperformance.

However, in practice the system of fines and penalties defined in the contract did not work. Why did was this? There are at least four complementary explanations.

- First, the contract established a three month period after the implementation of the reform (last paragraph of clause 3.6.1.1. of the *Bases*) where the fines defined in (2), (6), (7) above and the fines for irregular service (3.6.1.1. 7º) did not apply. Thus, from February 2007 until 10th May 2007 the authorities could not apply fines for low frequency and low capacity supply.

- Second, some of the operational faults were ill-defined. Take for example (7) above. In a route with 20 buses an hour (typical in certain trunk routes) an operator could supply 12 buses (perhaps with higher capacity) and still not be subject to a fine.

- Third, and much more important than the above two points, is that in order to apply a fine, the authorities needed to verify the fault using inspectors in the street since there was no other monitoring technology available. There was also a somewhat lengthy bureaucratic procedure for applying fines, including an appeals mechanism for operators.[44] Needless to say, considering the scale of the crisis after February 2010, the authorities did not have the required manpower to enforce the operational performance through fines.

- Finally, the 6,000 UF limit on fines proved to be non-credible and it eventually benefited operators. Notice that the authorities could have easily terminated all concession contracts within weeks or months of the reform. All it would need to do is to find 30 services perhaps in different days where frequency was below 60% of that established in the operational plan for a given concessionaire. In mid-2007 this would have been an easy task with respect to any operator.

However, the authorities could not terminate all concession. Even the termination of a small area concession (zone G) in September 2008 proved to be extremely difficult. The transition from one concessionaire to the other required several months, and during the interim period the quality of service fell incessantly. Terminating other concession contracts was just not an option for the authorities. Thus, the above 6,000 UF fine limit actually worked in operators' favor since the authorities had their hands tied as to how many fines they could apply to a given concessionaire.

The original concession contracts also established a mechanism to reward operators for good quality performance (clause 3.5.6 of the *Bases*) based on passenger surveys, regularity and other quality measures. The financial rewards would be funded from the fines applied to operators. However, this mechanism would first apply twelve months after the initial introduction and was never applied.

---

[44] It is interesting to note that in some systems, such as in Bogotá, contracts stipulate a reduction (30%) in the amount of a fine due when fines are paid early in order to reduce incentives for operators to use administrative and legal instruments in challenging or postponing fines.

As was described above, from mid-2007 the emphasis was placed on improving operational performance through the payment mechanism. However, fines could still be applied and they became increasingly useful to guarantee quality of service in specific routes and services. Thus, the payment mechanism guaranteed an overall or aggregate level of services, while fines became a useful tool to enforce performance in specific underperforming routes.

The 6,000 UF fines limit was eliminated in the 2009 contract renegotiation and, as explained above, the *ICF* and *ICR* index and their associated payment discounts were formally introduced in the contract to be applied for specific services while overall service performance was enforced through the *ICPHK* index.

The new contracts introduced in 2012 mimic this last approach. As can be seen by equation (4´´) above, the payment formula includes a capacity Compliance Index (*ICT*) to control overall capacity supply but also includes specific discounts. The performance variables giving rise to these last discounts are defined in Annex 6 of the new contracts. These are reformulations of the *ICR* and *ICF* index. In the case of the ICR, it was redefined to avoid the pitfalls encountered with the old index (see Beltran, et al, 2012). In the case of the *ICF*, several indexes were defined for each route-period-day, route-period month, for all routes in a certain period of the day and for the aggregate company level. Underperformance of each of these indices gives rise to a discount depending on the route type (high or low frequency) and the shortfall between actual performance and required performance. In all cases, these new indices are measured through electronic means rather that inspectors in the street.

One possibly critical aspect of the new contracts is that total discounts cannot surpass 2% of the total payments due to an operator in a given payment period, which effectively limits the use of these discounts as an incentive device if performance deteriorates significantly. In addition, underperformance of certain indices up to some threshold does not generate fines. Although this may be reasonable if performance is now measured electronically and at all times, it may give rise to gaming on the part of operators as happened with a similar mechanism, the *ICPH* index, after the 2007 renegotiations. It is too early to tell what the effects of these new provisions will be on behavior.

Finally, Annex 7 of the new contracts includes a long list of faults giving rise to fines as in the original contracts.

4.1.6   *Clauses related to fleet renovation, driving practices and other environmental performance issues*

The evolving Transatiago contracts also had several clauses related to fleet renovation and other environmental performance issues that we describe in this section. First, we must first define what was understood at the time by fleet renovation. Buses in the old system were either EPA 91, EURO I, EPA 94 or EURO II diesel standards. The new Technical Specification standard buses had to be EPA 98 or EURO III or above if they were diesel.

The original *Bases* established a duration of 13 years for two out of the five trunk services (T2 and T4) where new buses had to be introduced from the beginning of the concession.[45] However, for trunk services T1 and T5 (*Clause 3.4.4.2. of the Bases*) there was flexibility as to the fleet characteristics and duration of the contract. In the case of these two services, if the concessionaire offered at the bidding stage to have a completely renovated fleet, then the duration would be 13 years. If on the contrary, the concessionaire offered to use buses from the old system, the concession would only last four years. However, if by the end of the third year into the concession the operator renovated the fleet completely, then the concession period would be extended 9 years so that the total duration was 13 years.

In the case of trunk service 3, the duration of the contract was always set at two years and it was expected that these services would operate with old buses. Also, in the case of feeder services, the concession duration was 6 years (beginning in August 2005 and later postponed to October 2005) and no incentive mechanism was considered to renovate the fleet.[46]

In the case of trunk services, the 13 years concession period could be further extended (up to a maximum of 6 additional years) if the operator renovated the fleet with buses with lower Particulate Matter (PM_{10}) and Nitrogen Oxide (NOx) emissions than the EURO III or EPA 98 diesel technology. This mechanism is presented in Clause 3.4.5 of the Trunk *Bases*.

The mechanism worked as follows. At least one year before the concession ends, the concessionaire must apply for the extension. Subsequently, the authorities estimate the emissions that would have occurred during the last 11 years of the concession assuming that the fleet was comprised of Euro III (or EPA 98) diesel technology:

$$EPB_T = \sum_{l=1}^{3} EP_{lT}^{DEIII} \cdot KR_{lT}^{DEIII}$$

$$EPB = \sum_{T=13}^{144} EPB_T$$

---

[45] It must be noted that this period was counted from the end of August 2005 (later changed to October 2005), when a one-year transitional phase was considered in the contract where operators used the new buses but with the network, payment method and route design of the old system. The definite implementation of the new system was programmed for the end of October 2006 (later moved to February 2007). Therefore, the duration of the concession under the new system was 12 years originally.

[46] There was one incentive to offer new buses in the feeder services. Clause 5.4 of the feeder *Bases* established that in case of a tie in the evaluation of bids (that was based exclusively on the PPT offered) then winner would be selected according to the highest payment to the "*Reserve Técnica*" fund. If there was also a tie in this last variable, then the winner would be chosen as the one that offered the highest proportion of TS standard buses in the fleet. It was highly unlikely that this last criterion would eventually have any impact on the probability of winning the concession so it probably had no effect on the incentives for renovation.

where T is the month, KR_{lT}^{DEIII} is the number of kilometers offered during month T of the concession by buses type l assuming they are of EURO III or EPA 98 diesel technology.[47] EP_{lT}^{DEIII} is the weighted average emissions of MP10 and NOx for the bus class l in month T. EPB then are the total (weighted emissions) of the concessionaire from year 2 to year 12 of the concession assuming his fleet was composed entirely of EURO III (EPA 98) technology.

The actual emissions are calculated with the real fleet, as:

$$EPE_T = \sum_{l=1}^{3} EP_{lT}^{t} \cdot KR_{lT}^{t}$$

$$EPE = \sum_{T=13}^{144} EPE_T$$

where EP^t and KR^t represent the weighted emissions and kilometers offered of buses type l but with technology t.

The extension period is then estimated to be (in months):

$$PE = Min\left[72; \left(1 - \frac{EPE}{EPB}\right) * 132\right]$$

where the second value of this formula is truncated to the lowest integer.

The weighted average emission parameter (EP) for each type of bus was set in the contract according to the values shown in Table 3.

Except for one inconsequential clause (*Bases* clause 4.3.1.2) related to driver training, no provisions were considered in the original contracts relating to driving practices or incentive mechanisms for eco-driving practices.

What were the effects of these incentives in practice? The concession extension incentive seemed to have worked. The winning bid for T1 offered a new fleet from the start. The authorities did not expect this to happen so quickly in the case of T1 and this even generated some operational problems.[48] In the case of T5, although the concession started

---

[47] Three bus types were defined in the Transantiago standard for trunk services (Clause 4.1.2 of the *Bases*): *B2* (between 12 and 14 meters in length), *C1* (between 14 and 16.5 meters in length) and *C2* (articulated buses longer than 16.5 meters).

[48] The new articulated buses of this concession would not fit under a bridge in some routes in the central part of the city (*Calle Bandera*). The authorities had expected to undertake the needed infrastructure investments required during the following year expecting this concession to operate with old buses for a few years.

Clean Air Institute

with old buses, within a few years the operator had renovated the whole fleet in order to extend the concession duration.

Table 3: Weighted average PM$_{10}$ and NOx parameter (EP)

Technology	Bus Type		
	B2 (under 12 meters)	C1 (12 to 14 meters)	C2 (articulated over 16 meters)
Euro III or EPA 98 diesel	1.00	1.12	1.15
Euro III or EPA 98 diesel with original particle filter from manufacturer	0.72	0.81	0.84
Euro IV diesel	0.50	0.56	0.58
EURO III or EPA 98 Natural Gas	0.35	0.39	0.41
EURO III or EPA 98 Hybrid (diesel – electric)	0.36	0.41	0.43
Electric	0.00	0.00	0.00

Source: Table Nº3 of the trunk services *Bases*.

Therefore, extending the concession from 4 to 13 years at the beginning did induce fleet renovation. The mechanism for extending the concession period in exchange for emissions reduction also had an effect as will be discussed further below.

The concession for T3 ended in 2008 and a new tendering process was undertaken during 2009 for these services. The new contracts stipulated a complete renovation of the fleet. In addition, the new contracts for feeder services negotiated in 2009 considered the complete renovation of the fleet for these services for EURO III or EPA 98 technology with filters.

Evidently, this renovation had an effect on the system's cost, but by the end of 2010 almost the complete fleet of 6.100 buses was Transantiago standard.[49] However, in order to make this renovation feasible, two new features were included in the contracts. The first was a fixed monthly payment (Renovation Payment) that would be index to the US dollar exchange rate only. The purpose of this payment was to guarantee the payment of the debt for the renovation of buses and that would be independent of demand or passengers transported. However, feeder services had a concession horizon up to October 2011 and would not be able to pay the full cost of the new buses in such a short period of time. Therefore, the new contracts established that in the new tendering process the authorities

---

[49] This decision had to do with the political consequences of having most of the poorer areas of the city served by old buses. This was deemed to be unfair to most users in these areas. It is also interesting to note that the original contracts stipulated (Clause 4.1.3.2 of the trunk *Bases* and 4.2.3.2 in the case of the feeder *Bases*) that filters had to be installed in old buses to reduce particulate matter emissions. In the mayhem following implementation this clauses were never enforced and with the renovation of the fleet became obsolete.

would oblige the new concessionaire to buy the buses and pay for the remaining renovation quotas to the previous concessionaire. In order for this to work, some provisions were included to guarantee the fleet maintenance standards of the outgoing concessionaire.

This last mechanism was never tested since the new authorities that took power in March 2010, decided not to tender the feeder services when these contracts expired in late 2011. Rather, they started renegotiation talks with trunk operators for the signing of a completely new contract. As part of the deal, the trunk operators would take over the feeder services once these contracts expired.

As mentioned earlier the new contracts negotiated during 2011 became operational during 2012. These contracts include the same mechanism to extend the concession period in exchange for emissions reduction as the original contract described above (Annex 4.H of the new contracts). In fact, some operators had already taken advantage of this system.

This was the case of the new operator U2 (previous T2 with feeder service G now included) that in 2009 had installed 564 filters in its fleet of 982 Euro III (EPA 98) diesel technology buses. The authorities estimated that over the life of the concession until 2017 (one year before the end of the concession) these filters would imply an extension of the concession for 22 months, until August 2020. Therefore, the new contract recognized that the concession would last until this last date.

Beyond August 2012, the U2 concessionaire could extend the contract for another three years if it continued to improve its fleet technology under the new contract. To this end, a new emissions table was established in the new contracts, reproduced as Table 4 below. This table includes parameters for the bus types used in feeder services that are now operated by trunk concessionaires.

Two other concessionaries have also used this mechanism to increase the duration of the concession. Therefore, it seems that the mechanism is having an effect in fleet renovation and emissions reduction. This is an interesting mechanism because users receive the environmental benefits of fleet renovation in the present but the costs are borne in the future in the form of a longer horizon before the contract can be retendered. There is also evidence that particulate matter pollution levels in Santiago (MP10) have decreased after the implementation of the Transantiago reform (Figueroa, Gómez-Lobo, Jorquera and Labrín, 2012). If this is the case, then there will be further benefits as operators take advantage of this mechanism and further reduce emissions.

Table 4: Weighted average PM$_{10}$ and NOx parameter (EP) in the new 2012 contracts

Technology	Bus Type					
	A1/L	A2	B1/M	B2	C1	C2
Euro II or EPA 94 diesel	1.000	1.000	1.000	1.000	---	---
Euro III or EPA 98 diesel	1.000	1.000	1.000	1.000	1.131	1.250
Euro III or EPA 98 diesel with original particle filter from manufacturer	0.449	0.569	0.663	0.742	0.839	0.927
Euro IV diesel	0.265	0.342	0.402	0.453	0.515	0.572
Euro V diesel	0.192	0.243	0.283	0.316	0.357	0.394
Euro IV Gas	0.238	0.307	0.362	0.407	0.464	0.515
Euro V Gas	0.173	0.218	0.255	0.284	0.321	0.355
Euro V Hybrid	0.154	0.194	0.226	0.253	0.286	0.316
Electric	0.000	0.000	0.000	0.000	0.000	0.000

Source: Annex 4.H. of the new 2012 contracts. *Notes*: buses A1/L correspond to vehicles between 8 and 9 meters in length with no more than 26 seats (including the driver) and weighing less than 10 tons, A2 to buses between 9 and 11 meters in length, and B1/M corresponds to buses between 11 and 14 meters in length with more than 26 seats (including the driver) and weighing between 10 and 14 tons, B2/P buses correspond to buses between 12 and 14 meters and weighing more than 14 tons. See D.S. 122 (1991).

4.1.7 *Lessons from the Santiago experience*

There are many lessons to be distilled from the Transantiago experience. However, as far as contracts and incentives are concerned the main lesson is that designing contracts for such a complex reform is not an easy task. Also, there will always be unexpected behavior on the part of operators or users that will induce the authorities to attempt to change the incentive scheme of contracts. In this respect, a gradual approach, with shorter contracts and in only some areas of a city where contracts can be tested first and then modified before extending them to the whole system may be ideal. In the conclusions to this report we present more ideas regarding the optimal sequence of a transit reform.

As far as the detailed incentive structure of contracts is concerned, the Transantiago experience indicates that operators should face some demand risk, particularly in services that do not operate in exclusive lanes or infrastructure where monitoring costs may be expected to be lower.[50] Also, good performance incentives based on operational variables are required in order to guarantee optimal performance standards.

In the case of fines and penalties, the Transantiago experience is also instructive in that sometimes, draconian penalties, such as the unilateral termination of a contract by the

---

[50] This lesson seems to have been adopted in the recent SITP reform in Bogotá. In this case, trunk operators do not face demand risk but zonal and inter-zonal operators that provide services in the general street network face demand risk equivalent to around 20% of their income. See Annex 3 for more details.

authorities, may be counterproductive. Fines and penalties must be enforceable and their application must not generate costs to the authorities or users if they are to be credible and thus useful as an incentive mechanism.

Restricting drivers' payment method was probably not a wise idea and those clauses have recently been removed from the latest contracts renegotiated in this city.

On the positive side, it would seem that voluntary incentives for fleet renewal and environmental improvements based on concession duration extensions have worked reasonably well in the case of Transantiago. These flexible mechanisms may provide an interesting experience for other reforms in the region.

Transantiago was a very ambitious reform and problems related to contractual design were only part of the problems faced by the new system. Space precluded a more complete analysis of this experience and the wider lessons to be learned for transit reform in Latin America. The interesting reader is referred to Gómez-Lobo (2012) for such a review.

4.2 Review of the London bus concession contracts

4.2.1 Introduction

This Section complements Santiago's experience with an analysis of the incentive structure provided by concession contracts for bus operators in London, England, providing a best practice benchmark, at least in citywide non-BRT type systems. Analyzing London's experience is useful due to the numerous changes in the contractual design during the last thirty years and the relative success shown by the last generation of contracts.

Reviewing London's experience of contractual reforms of bus companies poses many questions for those interested in designing contracts for public transport services in other cities. How have London bus contracts evolved during the last thirty years? What were the main drivers of these changes?

In this section we argue that there were two important drivers for contractual reforms: the need to reduce subsidies due to government budget constraints; and the need of improving quality of service.[51] Hence, setting the right incentives to deliver a less expensive, more efficient and high quality service became a central issue for designing contracts.

This review is structured in chronological order to show the evolution of contracts. This historical analysis allows the identification of the main drivers for change and how different contracts induced different economic behavior by companies.

---

[51] In 2007/08, the average fare for a journey stage was around USD 1 and the average subsidy for a journey stage was around USD 0.60. These values are computed based upon the data for the 2007/08 year provided in TfL (2009): Subsidy GBP 653 million, traffic Revenue GBP 1,053 million, journey stages 2,176 million. Also, an average exchange rate between July 2007 and June 2008 of GBP 1 = USD 2 was used.

4.2.2   Overview of the London bus system[52]

London has one of the largest bus systems in the world. During this year (2012), every weekday, more than six million people are transported on 7,500 buses. The bus network has more than 700 routes and 19,500 stops and stations, covering all areas of Greater London. More than 90% of Londoners live less than within 400 meters of a bus stop.

London's bus system stands out not only because of its size but also for the quality of service provided. Historical indicators of quality, such as reliability and punctuality, show that progress is evident, especially during the last ten years.[53] The current contractual scheme, Quality Incentive Contracts, can account for this improvement. Based upon economic rewards and penalties, Quality Incentive Contracts are designed to induce bus companies to act in such a way as to produce the level of quality expected by the Greater London Authority and its transport agency, Transport for London.

Developing Quality Incentive Contracts has not been an easy task for London´s transport authorities, having evolved from a complex process of trial and error during the last thirty years. In this period, bus companies evolved from just one public firm to several private organizations. Contracts with bus companies in different periods alternatively allocated production and revenue risks to transport authorities or private companies. The transport authorities changed many times during these years as well, being controlled by different levels of government, metropolitan or central.

4.2.3   The past

During the last forty years, London's bus contract scheme experienced significant changes as bus company ownership evolved from public to private companies. During some periods, contracts allocated cost and demand risks to government, while in other periods to companies. Finally, the transport authority responsible for managing contracts was sometimes under control of the metropolitan government and sometimes under control of central government.

This complex process of change can be analyzed considering five periods in which different types of contracts were used. Also, these periods can be characterized in terms of ownership of bus companies (public and/or private), level of government managing contracts and allocation of production and demand risk. Table 1 summarizes this analytical structure.

---

[52]   This overview is based on TfL (2012a).
[53]   See Figures 4 and 5.

Table 5: Main characteristics of different periods of London bus contracts[54]

Period	1970-1985	1985-1996	1996-1998	1998-2000	2000-present
Type of contract	Contract with one subsidiary public firm	Gross cost contract	Net cost contract	Gross cost contract	Quality incentive contracts
Ownership of bus companies	Public (1 firm)	Public (13 firms) and private	Private	Private	Private
Transport authority under control of:	Metropolitan government	Central government	Central government	Central government	Metropolitan government
Production risk allocation	Government	Operator	Operator	Operator	Operator
Demand risk allocation	Government	Government	Operator	Government	Government

During the first stage, from 1970 to 1985, there was just one public company providing bus services. This company was a subsidiary of London Transport, the transport authority at the time. London Transport was under the control of the Greater London Council, which was the metropolitan government. Thus, metropolitan government internalized both the costs and demand risks of bus services. The operation of London Transport's subsidiary was inefficient. There were no incentives for improving efficiency because market competition with other firms did not exist. Thus, running bus service in London during these years demanded growing amounts of public funds every year. Also, quality of service was poor. There was significant number of scheduled services that did not operate because of staff absence or lack of buses due to mechanical problems.

In the second stage, between 1985 and 1996, this situation changed radically. In 1984, the London Regional Transport Act was passed, establishing that London Transport's subsidiary would be privatized; in a first stage, competition would be introduced through the tendering of services and later through deregulation; and the Greater London authority would be abolished. London Transport would be controlled by central government. The aim of this reform was to start a gradual process of privatization of the public bus company. The reform was expected to reduce public subsidies and to transfer risks out of government. Also, this reform aimed at improving the operational efficiency and the quality of bus service.

In 1985, new contracts were introduced. London Transport' subsidiary was split into thirteen smaller subsidiaries and a route tendering process was conducted. In this process

---

[54] Years are approximate because there were transitions between periods.

the thirteen new public firms had to compete against private firms. The tendering variable was the revenue required to operate the specified services. Routes were awarded to the operator that could run the services at the most cost-effective price. These contracts were denominated as 'Gross Cost' contracts and implied paying operators a certain amount for supplying services. Thus, these contracts transferred operating cost risks to firms although demand risk remained with London Transport.

Around 40% of the initial contracts were awarded to private companies rather than London Transport's subsidiary companies. These new public and private companies conducted their own wage negotiations, took appropriate steps to reduce their overhead costs, and competed against each other for the contracts to run LT bus routes. Finally, in terms of quality of service, gross cost contracts introduced higher standards for safety and reliability. Contracts could be terminated for poor performance. Under the new contracts, operating companies were not paid for cancelled services when it was their responsibility.

Thus, in terms of the conceptual review provided in this report, during this stage London bus contracts can be classifies as having a payment mechanism based exclusively on operational variables, with no demand risk transferred to operators.

During 1994 and 1995, the thirteen public subsidiary firms were privatized. London Transport retained the function of planning routes and setting the fare structures.

Between 1996 and 1998, the reform process went one step further. Authorities at the time were interested in transferring both production and demand risks to private operators. Gross Cost contracts were replaced by a new generation of contracts called Net Cost contracts. In this new scheme, the bus operator gets to keep the fare revenue. This shifted most demand risk to operators who would now have the incentive to generate more revenue by increasing the quality of the service provided. These net cost contracts were initially not subject to competition as the routes were already allocated to each subsidiary and the terms of the new contracts were agreed by negotiation. The duration of these contracts varied in order to give each company a reasonable time horizon to recoup investments and to allow the network to be tendered over a reasonable timescale.

However, net cost contracts did not produce the expected results. On the contrary, quality of service actually decreased as operators focused on lowering costs rather than improving quality to attract new users. Hence, it became clear for the transport authorities that contracts had to be reformulated again.

Between 1998 and 2000, gross cost contracts were applied again as a transition to a new generation of contracts called Quality Incentive Contracts.

4.2.4   *Present practice*[55]

From 2000 to the present, Quality Incentive Contracts were designed and applied in order to induce bus companies to act in such a way as to produce the level of quality expected by

---

[55]   We would like to thank Clare Kavanagh, Director of Performance, TfL London Buses, for providing very useful information for this section. Any errors or omissions are the authors' sole responsibility.

Londoners and its authorities. After more than ten years, progress in quality is evident as graphs at the end of this section show.

Quality Incentive Contracts work with a payment scheme that has two components: one that covers all costs required to finance a specified transport service, like in Gross Cost Contracts;[56] and the other one is an incentive provision that adds or discounts resources according to the quality performance of transport operators. Transport for London measures bus performance and the results are contrasted against operational conditions defined in contracts as Minimum Performance Standards. In the case of meeting or exceeding these standards, financial bonuses are paid to operators. On the contrary, deductions are applied to payments for underperformance. Contract extensions are also used as an incentive mechanism for performance. If an operator goes beyond Minimum Performance Standards, then it has the possibility of extending the duration of the contract in two years.

There are three types of incentive provisions in Quality Incentive Contracts:

- Operated Mileage

- Reliability Performance Payments

- Contract Extensions

First, Operated Mileage is the incentive provision to operators to deliver the scheduled mileage established for each route in its contract. Miles which are not provided are divided into two categories:

- Lost mileage under the control of bus companies, such as staff absences and sickness, or mechanical problems, are not paid. This deduction is applied in proportion to annual contract price and scheduled mileage.

- Lost mileage beyond the control of bus companies due, for example, to adverse traffic conditions, is not deducted.

Second, Reliability Performance Payments are calculated on an annual basis by comparing the Operator's annual reliability performance on each route against the contracted Minimum Performance Standards. These are generally set for the life of a contract based upon specific routes characteristics like length and average journey time in the route, types of areas served and whether these are congested or not. Reliability measurements are different for high and low frequency routes.

For high frequency routes (more than 4 buses per hour, for example, those operating during the day) waiting time reliability is the most important attribute. This depends on how evenly spaced buses are on each route. Hence, Transport for London measures intervals between buses, expressing this as "Excess Waiting Time", which is defined as the extra time that passengers waited above the expected waiting time had services been perfectly reliable.

---

[56] According to TfL (2009), costs include: driver wages, fuel, bus depreciation, insurances, maintenance materials, and other labor and staff costs.

The quality incentive scheme for high frequency routes imply an additional payment of 1.5% of the base contracted price for every whole 0.10 minute in Excess Waiting Time below the minimum standard set in the contract. On the other hand, the operator's income is reduced by 1% of the contracted price for every 0.10 minute in Excess Waiting Time above the minimum standard established in the contract.

For low frequency routes (four buses per hour or less) punctuality is the most relevant attribute. In this case, the transport authority contrasts the departure time from the stop to the advertised timetable, computing the percentage of departures that are on time. If significant delays are detected, then this also gives rise to a payment deduction to the operator. The measure is expressed as a percentage of departures that are "On Time", being this a window from 2½ minutes earlier than expected to five minutes later than expected. The objective is to increase On Time departures to 100%. Additionally, the percentage of journeys running "Early" is also monitored, where early is defined as a bus departing between 2½ and eight minutes ahead of the schedule. It is normally considered that there is little excuse for early running, as passengers may not have arrived at the stop.

Bonus payments are paid at a rate of 1.5% of the contract price for each full 2% "on-time" performance that exceeds the minimum standard set in the contract and deductions are made at a rate of 1% of the contract price for each full 2% below the standard.

All bonus and deduction for quality performance are capped at 15% and 10% of the contract price for each payment period, respectively.

Finally, contracts, that have an initial period of five years, can be automatically extended for two years if an operator meets or exceeds the reliability "Extension Threshold" criteria. This reliability threshold is slightly higher than the reliability Minimum Performance Standards, being based upon Excess Waiting Time and On Time scores. If a route qualifies for an extension, then the operator is free to choose to work or not this route in the same current contractual conditions. If the operator does not accept the extension, the route is tendered in the usual manner. If the operator accepts the extension, then the route is withdrawn from that year's tendering program, and is tendered two years later.

There are some routes that work with just some of the economic incentives because the cost of monitoring exceeds its benefit. This is the case of night routes, school services and other low frequency services. For instance, some of these routes only have performance payments but not contract extensions.

It is important to note that Transport for London worked on a second generation of payments based upon the following quality aspects:

- Driving quality, including customer service and the professionalism, passenger interaction, smoothness of ride, serving the stop;

- Vehicles internal and external presentation, including cleanliness, damage, defacing and wears to interior and exterior features.

There was a pilot experience during 2008 and 2009 with this second generation of quality contracts. Although the pilot was successful, these new quality aspects were not finally

included in contracts because its monitoring cost was significantly expensive. This is interesting because it illustrates the trade-off discussed in this report between performance incentives and monitoring costs of different contractual mechanisms.

There are a few other features of the contracts that deserve mention.

First, contractual accounting periods are based upon four weeks, having 13 periods each year running from April to March. 75% of the contract price is paid during each period. The balance, less deductible lost mileage, is paid at the end of the following period. Payments (and deductions) under the reliability incentive scheme are made annually, and for the more recently introduced driving and vehicle quality incentive schemes, quarterly.

Second, among the responsibilities of Transport for London is the determination of bus routes and frequency, setting fares and collect revenues and provide revenue protection (on-bus revenue protection inspectors), among others. On the other side, operators must develop the timetables (that have to be approved by Transport for London), manage all aspects of the day-to-day operation of routes and control the use of passes and collect any cash revenue on buses.

Sanctions are applied to aspects related to security conditions for passengers and also satisfaction. Transport for London examines a range of indicators including the technical ability of the driver, maintenance procedures and mechanical condition of the vehicles, Customer Satisfaction Surveys, focusing on Bus Services, Night Buses and Bus Stations, and an assessment of an operator's ability to provide a safe service.

Unsatisfactory performance is discussed with individual operators, and if necessary operators may be required to produce and implement action plans to resolve performance issues. Current and past performance is taken into account in the evaluation of tenders and recommendation for award of new contracts. Transport for London usually resolves any performance issues through normal contract management. However if performance continues to be poor and it is considered that it cannot be resolved by other means, as an ultimate sanction, London Buses retains the right to terminate any contract.

Quality incentive contracts have dramatically improved the level of service of buses in London. Companies have effectively responded to its three incentive provisions. Figure 4 shows that Operated Mileage has had a significant improvement. Figure 5 shows that since 2001, when quality incentive contracts were introduced, Excess Waiting Time decreased from 2.2 to 1.0 minutes. Before 2001, no clear pattern can be seen. One interpretation for this would be that before 2001, contracts were not much concerned with quality of service but rather focused in cost and revenue incentives. For instance, while gross cost contracts were applied, between 1985 and 1996 and later between 1997 and 1998, operators were more focused on supplying contracted capacity than paying much attention on having regular intervals between buses along routes. Also, under net cost contracts, applied between 1995 and 1998, bus operators had incentive to transport people at the lowest cost possible, reducing quality of service. Finally, Figure 6 shows that the percentage of On Time Departures has increased since 68% in 2001 to 83% in 2011.

However, this improvement on quality of service was not free. Subsidies have increased significantly between 1999/00 and 2007/08, from GBP 41 million (USD 82 million) to

GBP 653 million (USD 1,306 million).[57] This growth can be attributed to both costs and revenue reasons. On the cost side, the prices of relevant production inputs have increased. Also, route network size has grown significantly requiring the supply of more mileage. Finally, bus speeds in London have decreased significantly, demanding more vehicles in order to meet frequencies set by TfL. On the revenue side, real average fare per passenger journey stage at 2007/08 prices have decreased from 61 pence (USD 1.22) in 1999/00 to 48 pence (USD 0.96) in 2007/08. However, the number of passenger journey stages increased from 1,296 million in 1999/00 to 2,176 million in 2007/08 (TfL, 2009). As a result traffic revenue grew from GBP 789 million (USD 1,578 million) in 1999/00 to GBP 1,053 million (USD 2,106 million) in 2007/08 (figures at 2007/08 prices).[58]

---

[57] Considering an average exchange rate between July 2007 and June 2008 of GBP 1 = USD 2.
[58] As a reference, according to the World Bank, GDP per capita in the United Kingdom is USD 38,818 and in Chile its USD 14,394.

Figure 4: Mileage operated by the London bus system
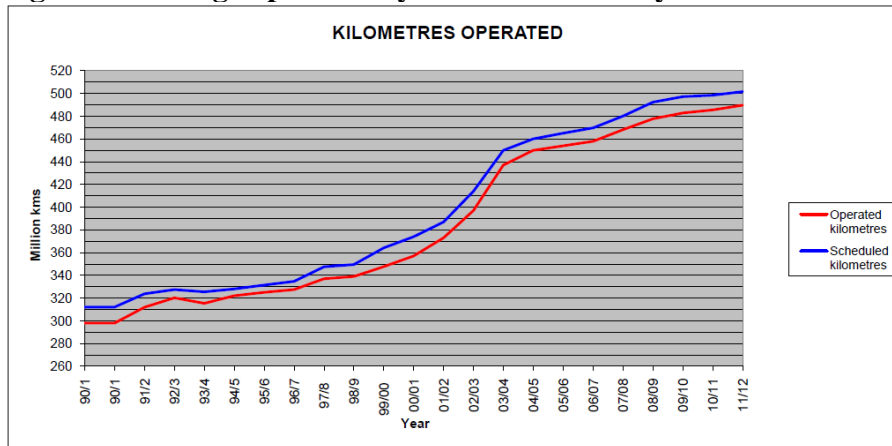
Figure 5: Excess waiting time
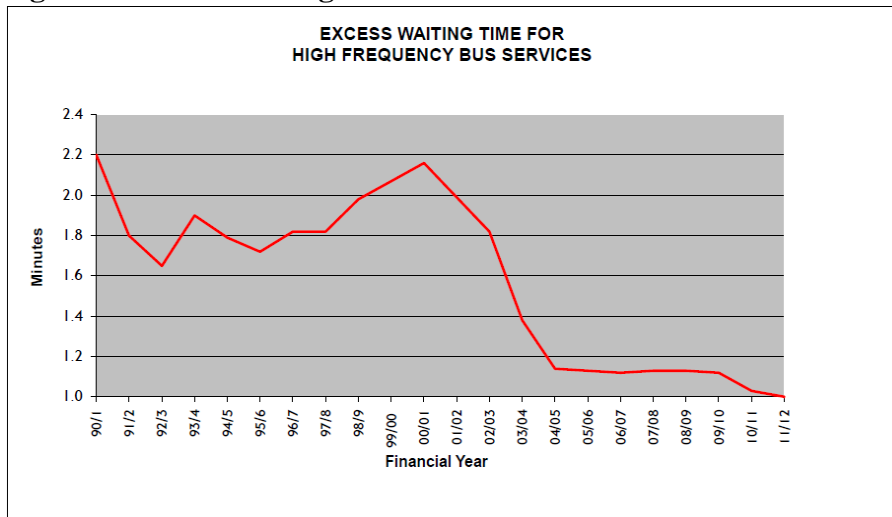
Figure 6: Percentage of "on time" buses
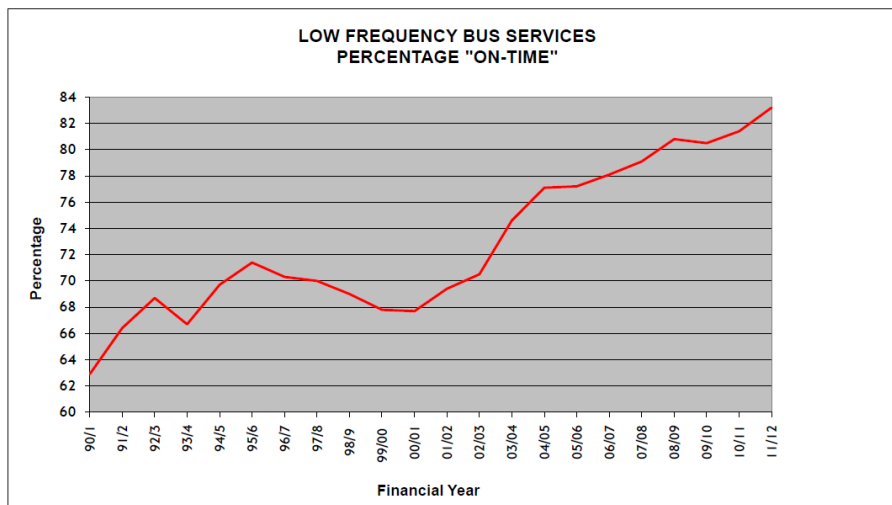
Sources: TfL (2012b)

4.2.5   Future developments

Although there is clear evidence of positive results from the introduction of quality incentive contracts, these achievements have required growing amounts of public resources. In 2009, Transport for London conducted a study (TfL 2009) analyzing the possibility of returning to net cost contracts (as in 1996-1998 period) in order to reduce subsidy requirements. Net cost contracts would transfer revenue risk back to bus companies.

The net cost contract approach has pros and cons which have to be carefully assessed. On the positive side, operators would have incentive to increase demand. This could increase their revenue and, eventually, operators would need fewer subsidies. Operators should be interested in improving quality standards in order to attract more users. However, when net cost contracts were applied between 1996 and 1998, quality actually decreased so care must be taken if these types of contracts are reintroduced.

On the negative side, operator would experience more risk, thus increasing required returns and so subsidies could even increase with the new contracts. Also, this type of contract could adversely affect the financial stability of companies. This is clearly an area of public concern since failure by one of these companies could have important social impacts.

TfL (2009) concludes that in the current situation it is not clear whether moving to a contractual scheme based upon net cost would have more benefits than costs. This change would require further analysis.

4.2.6   Lessons from the London experience

Figure 7 presents a diagram contrasting incentives of bus contracts during last 30 years in London. The arrows indicate the actions that each type of contract promoted. Under Gross Cost contracts companies' revenue were fixed and could only increase profits (π) by decreasing costs (TC). However, firms only focused on reducing cost without regards to quality. Also, these contracts did not provide any incentives to increase demand.

Hence, it was thought that under Net Cost contracts bus operators would have more incentive to increase passengers in order to raise their total revenue and thus profits. It was thought that firms would improve their quality of service in order to attract more passengers. Also, it was expected that higher revenues for companies would mean that less subsidy would be required. However, companies would need to be financially compensated for the additional risk they were assuming and thus subsidies could actually increase.

Finally, Quality Incentive Contracts were applied. These were based upon Gross Cost contracts but considering provisions for quality. Incentive for reducing costs remained but not enough to compromise the supply of a reasonable level of quality since this would also affect revenues through quality performance bonuses and deductions. Operators would not require a risk premium under these contacts (unlike Net Cost contracts) and so subsidies would not be affected by this change.

Figure 7: Incentives of bus contracts in London over the last 30 years
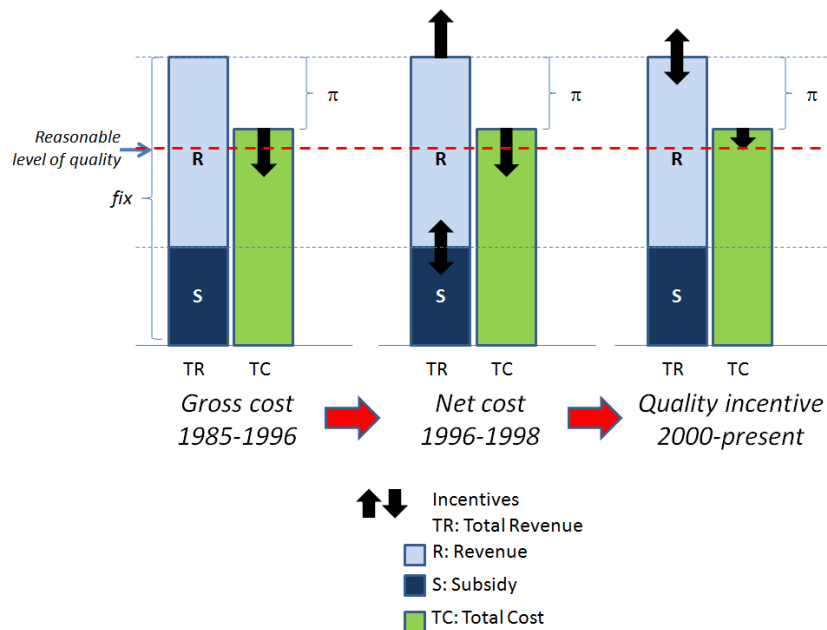
This analysis would seem to indicate that the Quality Incentive model provides a good balance between the objectives of users, the authorities and companies. It takes advantage of the best features of a Gross Cost model; production risks are under the control of operators so they have incentives to be operational efficient and control cost, but do not have to bare excessive demand risk which would increase costs through the risk premium charged by operators. Also, these types of contracts focus directly on the quality provided to users, one of the main problems of Gross Cost contracts. However, the experience in London shows that quality performance is not cheap and requires increasing levels of subsidies and bonuses.

5   Conclusions and recommendations

One of the main conclusions from this study that comes out clearly from the two case studies reviewed in this report is that designing effective bus concession contracts —that provide the right incentives for quality provision— is extremely difficult. This is particularly so for extensive city-wide bus systems as opposed to narrower BRT type schemes. This in turn reflects the difficult trade-offs involved between the economic and financial interest of the public sector, private operators and users in the design of any major transit reform.

In London it took more than three decades of trial and error (and substantial public subsidies) before they developed a contract (Quality Incentive Contracts) with good performance characteristics. In Santiago, contracts have undergone three major renegotiations in four years since the reform was introduced.

The importance of the above point is that policymakers must be aware that this is a crucial design issue that must be treated with great care in future reforms. Mistakes in the design of the incentive structure of contracts may derail a public transport modernization reform jeopardizing the expected benefits of such policies, including the associated energy and environmental benefits.

Thus, policymakers should learn from past experiences and pitfalls in order to avoid making the same mistakes in other reforms. We hope this report makes a valuable contribution in this respect, at least as far as detailing some of the problems faced in the case of Santiago and London.

However, no matter how carefully contracts are designed, problems are bound to arise. Therefore, besides learning from past experience, policymakers should opt for reform processes that are gradual in time and space and leave some room for trial and error, rather than all-encompassing, "big-bang" approach as used in Santiago. A gradual approach leaves some room for correcting problems as they arise without subjecting users to the costs and frustration of a major and generalized crisis in the public transport system.

One possible approach would be to concession some parts of the city, perhaps the least critical, before embarking on a wider reform. Another approach is to experiment with different contract designs in pilot cases, as was done recently in London (2008 and 2009) with a new generation of Quality Incentive Contracts. Still another approach is a gradual process of reform on a functional basis. For example, introduce a centralized revenue collection system with a pre-payment or electronic charge card should arguably be a first step and only tender operating concessions when this payment system is consolidated. The exact details of such a gradual approach will depend on the specific characteristics of the city and the type of reform pursued. But the central point remains that undertaking a major reform of a critical public service such a transit cannot and should not be done at one stroke.

Another major conclusion of this study is that there is no 'perfect' bus concession contract since these will always have to trade-off conflicting objectives. For example, how much demand risk one wishes to transfer to operators will depend on the trade-off between

providing incentives for operators to cater to users with the objectives of traffic safety and financial affordability of reform.

Achieving the perfect balance between these conflicting objectives will always be elusive and perhaps controversial. In fact, this is why there is still an on-going discussion in London as to the right contract type. In the case of Santiago, it is too early to know whether the new contracts will provide the right incentives, but it is unlike that current contracts will prove to be perfect. It is highly probable that in the future, discussions will again arise as to the convenience or not of certain contractual arrangements.

In spite of the above, from the case studies presented in this report some general guidelines for contract design can be established.

First, it seems that shielding operators from demand risk to some extent is unavoidable. This makes the costs of the system lower, since companies will charge a lower risk premium, and will also contribute to operators' access to financial markets in order to fund fleet renovations and other investments.

However, it is also recommended that operators face some demand risk. This will induce companies to cater to demand and provide an adequate quality of service for users in a decentralized manner. This is perhaps even more important in developing countries where the alternative of using other mechanisms to provide quality is hampered by monitoring and enforcement costs. However, it is interesting to note that even in the case of London there is an on-going discussion as to whether operators should face more demand risk (use a net cost contract instead of a gross cost contract with quality incentives as is done now).

Second, it seems recommendable to condition payments to operators based on performance with respect to some key operational variables. The experience with the first Transantiago contracts, or the London experience with net cost contracts during the 1996-98 period, indicates that operators must be incentivized directly (through bonus or fines for performance). Counting on indirect incentives (through the effect of quality of service on ridership when companies face demand risks) has proved to be insufficient at least in the case of Santiago and London.

Thus, from the above review, it would seem that optimal bus contracts would subject operators to some demand risk but would not make payments completely dependent on ridership. They would also include important performance based bonuses or discounts in order to incentivize quality of service provision directly. Finally, it would seem that including restrictions on the type of labor relation between operators and drivers in the contract —beyond of course requiring this relation to be formal and within the legal framework— is not recommendable. This will probably distort the functioning of the incentive structure that the authorities are trying to implement with the contract.

Finally, it must be borne in mind that according to past experience most concession contracts will be renegotiated during their life-time.[59] This highlights the need to design flexible contracts that can be adapted to future circumstances. However, introducing flexibility to change contracts explicitly is difficult since operators will naturally want some stability in the "rules of the game" and financial security in order to invest and commit to the reform. There is the risk of opportunistic behavior by one of the parties if contracts are not clear or complete.

Therefore, it is probably not possible to design contracts with much flexibility, unless clauses are introduced that specify how the contract can be changed, when and with what compensation to operators. Alternatively, shorter contracts could be tendered in order to give the authorities the possibility of modifying the contracts in the future through a new tendered process. Another option would be to include a mechanism in the contract that allows the authority to buy back the concession under certain circumstance. Unfortunately, none of these recommendations are easy to implement and all have drawbacks but it may be important to recognize at the design stage of a reform that the probability of contract renegotiation is probably very high.

---

[59]   As pointed out by one reviewer of this report, the authorities in Bogotá are attempting to renegotiate the Transmilenio contracts. Therefore, even in the case of this successful reform contracts are being subject to a renegotiation process.

6   References

Beltrán, P., A. Gschwender and C. Palma (2012), 'The impact of compliance measures on the operation of a bus system: The case of Transantiago', *Research in Transportation Economics*, forthcoming.

EMBARQ (2010), Lessons Learned from major bus improvements in Latin America and Asia, World Resources Institute, Washington D.C.

Estache, A. y A. Gómez-Lobo (2005), 'The Limits to Competition in Urban Bus Services in Developing Countries', *Transport Reviews*, vol. 25(2), March, 139.158.

Figueroa, E., A. Gómez-Lobo, P. Jorquera and F. Labrín (2012), 'Develando Econométricamente los Impactos del Transantiago en la Contaminación Atmosférica de Santiago de Chile', Documento de Trabajo Nº355, Departamento de Economía, Universidad de Chile.

Gómez-Lobo, A. (2007), 'Why Competition Does Not Work in Urban Bus Markets: Some New Wheels for Some Old Ideas', *Journal of Transport Economics and Policy*, Vol. 41, part 2, May, pp. 283-308.

Gómez-Lobo, A (2011), 'Monopoly, subsidies, and the Mohring effect: a synthesis and an extension', Working Paper Nº 336, Department of Economics, University of Chile, April, available at:
http://www.econ.uchile.cl/uploads/publicacion/715bc4d96f1f147cb2d851c28e33c20d9a083
77f.pdf.

Gómez-Lobo, A (2012), 'The ups and downs of a public transport reform: the case of Transantiago', Working Paper Nº 354, Department of Economics, University of Chile, available at:
http://www.econ.uchile.cl/uploads/publicacion/479bc827d4ea165274125814b5889d03b035
3b8f.pdf

Hensher, D. and A.M. Brewer (2001), *Transport: An Economics and Management Perspective*, Oxford University Press.

ITDP (2007), 'Bus Rapid Transit Planning Guide', Institute for Transportation & Development Policy, available at: http://www.itdp.org/microsites/bus-rapid-transit-planning-guide/brt-planning-guide-in-english/.

NEA, inno-V, KCW, RebelGroup, TØI, SDG, TIS.PT (2008), 'Contracting in Urban Public Transport', report produced for DG Energy and Transport, European Commission, Amsterdam, January 14th.

TfL (2008), London's Bus Contracting and Tendering Process, Transport for London.

TfL (2009), *Independent Strategic Review of Bus Services in London*, Transport for London Board, Agenda Item 5, 29 July 2009. Available at:
 http://www.tfl.gov.uk/assets/downloads/Item05-Independent-Bus-Review-July09.pdf

TfL (2012a), http://www.tfl.gov.uk/corporate/modesoftransport/1548.aspx, Transport for London.

TfL (2012b), http://www.tfl.gov.uk/assets/downloads/long-term-trends.pdf, Transport for London.

Urban Age (2008), 'South American cities: Securing an Urban Future', Cities Programme, The London School of Economics and Political Science, available at: http://lsecities.net/publications/conference-newspapers/south-american-cities-securing-an-urban-future/#publication-section-South American cities.

Weitzman, M.L. (1974), 'Prices vs. Quantities', *The Review of Economic Studies*, 41(4), pp. 477-491.

World Bank (2009), 'Implementation completion and results report on a loan in the amount of 30.16 million to the Republic of Chile for a Santiago Urban Transport Programmatic Development Policy Loan', Document of The World Bank, Report No: ICR0000631, IBRD-73150, August 11, 2009, Sustainable Development Department, Argentina, Chile, Paraguay and Uruguay Country Management Unit, Latin America and Caribbean Region.

Annex 1: Summary of the Transantiago and London contractual experience

1. Operators' payment mechanisms

Table A1.1: Summary of payment mechanisms in Santiago and London

	Period used	Positive effects	Negative effects
Santiago			
Payment to operators mostly fixed (very small demand risk faced by operators) and payment not linked to operational variables.	February 2007 to July 2007.	Traffic safety improved considerably, with reduced accidents, injured and deaths involving buses.	No incentive to cater to passengers. No incentives to comply with operational plan. Low supply of buses, bad regularity of service and no incentives to control non-payment.
Payment mechanism supplemented with a capacity per hour supply index (*ICPH*). The *ICPH* measured the proportion of bus-seats in operation over the required bus-seats required by the operational plan. This index varied between 0 and 1 and operator's income was the payment defined in the original contract (see above row) multiplied by this index. Payment to operators was the payment due under the previous mechanism multiplied by the *ICPH* index, but with certain flexibility rules. Demand risk faced by operators was increased slightly.	July 2007 to September 2008.	Significant increase in the number of buses operating in the system. Did not raise any traffic safety concerns compared to the previous system.	Operators started taking advantages of some of the rules defining the *ICPH*. They learned to increase this index without providing the frequency required for each route in the operational plan. Index was at the aggregate level for each operator and did not help enforce the operational plan on a route basis. Introduction of index did nothing to provide incentives for operators (or drivers) to control non-payment.
The above payment mechanism (including the	October 2008 to first semester	Improved service regularity and	Drivers started skipping stops (even

![Clean Air Institute logo]

ICPH index) was complemented by discounts linked to a Regularity Compliance Index (ICR) and a Frequency Compliance Index (ICF) at the route level. Non-compliance with these indices were penalized through revenue discounts. In 2009, the *ICPH* index was changed to the *ICPHK* index. This index now measured the bus-seat-kilometers in operation compared to that required by the operational plan for aggregate performance enforcement. This index is a capacity-kilometer supply index. The *ICF* and *ICR* remained in use but for the control of specific routes and services.	2012.	frequency at the route level. Average waiting times and travel times reduced.	with passengers waiting) in order to meet ICPHK index requirements. Introduction of index did nothing to provide incentives for operators (or drivers) to control non-payment.
Contracts completely overhauled, more demand risk introduced (but not completely), payment still based on performance indices, ICR and ICF fine-tuned.	End of first semester 2012 to present.	Too early to know, although some evidence that fare evasion is decreasing	Too early to know.
London			
Privatization and tendered gross cost contracts. Contracts introduced higher standards for safety and reliability. Contracts could be terminated for poor performance. Operating companies were not paid for cancelled services when it was their responsibility.	1985-1996	Higher standards for safety and reliability. Reduced costs and subsidy requirements.	Quality of service not optimal according to authorities.
Net Cost contracts transferred demand risk to operators (i.e. payment based on passengers transported).	1996-1998	More incentives to be responsive to demand.	Operators concentrated more on reducing costs (with subsequent quality of service reductions)

			rather than increasing ridership through service improvements.
Gross Cost contracts.	1998-2000.	Stopped the vicious circle of cost reductions and service quality deterioration.	Quality of service not optimal according to authorities.
Quality Incentive contracts (gross cost contracts with bonus and discounts linked to service performance).	2000 to the present.	Improved service quality, frequency and regularity. Waiting times for users reduced.	Costs of the system increased (although cannot be all attributed to the contractual change since kilometers supplied increased significantly).

2. Drivers' payment mechanisms and other restrictions

Table A1.2: Summary of driver's payment restrictions in Santiago[60]

	Period used	Positive effects	Negative effects
Drivers' compensation could not be linked to the number of passengers transported.	February 2007 to first semester 2012.	Reduced competition in the streets for passengers with subsequent increase in public safety. Although this may have resulted from several factors, including the payment mechanism and the exclusivity of services in most zones, it is probable that driver's compensation scheme was a significant factor.	Drivers had no incentives to cater to demand or control non-payment. Muted any incentive effects that the authorities wanted to provide by making operators face more demand risk, since operators could not pass these incentives on to drivers.
No restrictions placed on drivers' compensation mechanism besides the requirement that labor laws be respected.	First semester 2012 to the present.	Too early to know, although some evidence that fare evasion in decreasing	Too early to know.

---

[60] For the case of London there are no clauses related to this issue.

Clean Air Institute

3. Fines and other penalties

Table A1.3: Summary of fines and penalties in contracts in Santiago

	Period used	Positive effects	Negative effects
Long list of operational requirements with associated fines Threshold level of fines an operator could accumulate in a twelve month period before concession had to be terminated by the authorities. This threshold was quite low in absolute terms	February 2007 to end 2009	Provided some incentives for performance.	Many operational requirements were not applied because they were: a) ill-defined, b) required monitoring and enforcement personal not available to the authorities, c) or did not apply during the first three months of the reform. The threshold level of fines before contract termination proved to be a non-credible threat since the authorities could not terminate contracts without generating critical disruptions to the transport system. In the end, it limited the enforcement capacity of the authorities since they could not pass more fines than the upper limit defined in the contracts.
The requirement of terminating a concession when the upper limit of fines was surpassed was no longer mandatory Overall less emphasis placed on fines and penalties to enforce quality provision compared to the use of the payment mechanism for this objective	End 2009 to first semester 2012	Fines could now be used in an unlimited way to enforce performance variables, although the authorities used the payment mechanisms for this purpose rather than fines and penalties	
	First semester 2012 to present	Too early to know.	Too early to know.

4. Fleet renovation and environmental performance incentive mechanisms

Table A1.4: Summary of fleet renovation and environmental performance mechanisms in contracts in Santiago and London

	Period used	Positive effects	Negative effects
Santiago			
Two trunk services concessions had a duration of 4 years. This period could be extended to 13 years if the operator renovated the entire fleet. These concessions, and two other trunk concession that had to start out with a renovated fleet could extend the contract beyond 13 years for another six years if they over complied with emission standards of their fleet or if they installed filters on their existing fleet.	2007 to the present	The fleets of the two original 4 years trunk concessions were renovated before the time limit established in the contract and these concessions were extended to 13 years. Some operators (particularly trunk service T2) have installed filters on their existing fleets and have already extended their contract beyond the 13 years limit.	Current contracts will last until at least 2020, making further adjustments to contracts and service obligations difficult. These contracts will not be tendered for another 8 years at least.
London			
Quality incentive mechanism for driving performance and presentation of buses	Pilot experience 2008 and 2009	Had desired effect	Too expensive to monitor and thus these quality variables are now not included in incentive contracts.

Annex 2: Other cases in Europe

In this Annex we briefly report the Awarding mechanism and risk allocation characteristics of contracts for a set of Europe cities (Table A2.1).

Table A2.1: Risk allocation and awarding scheme in bus contracts in some European cities

Awarding	Risk allocation to the operator		
	Limited risks	Production risk	Production and Demand risk
Directly award to publicly-owned operators	Prague	Krakow Rome	Amsterdam Barcelona Brussels Budapest
Competitive award to independent operators **According to multi-criteria procedure**		Dublin London	Lyon
Competitive award to independent operators **According to negotiated procedure**		Frankfurt Stockholm Warsaw	Manchester

Source: based upon NEA, et al (2008).

Annex 3: Review of other Latin American experiences[61]

In this Annex we briefly review the contracts from three other bus concessions in Latin America. These include the Metrolínea reform in Bucaramanga, Colombia, the Transmilenio system in Bogotá and the recent citywide reform in the same city called SITP.

1. Metrolínea, Integrated Massive Transport System in Bucaramanga, Colombia[62]

General description

This reform is an integrated transport system comprising 20.1 kilometers of exclusive bus corridor operated with high-capacity buses and with pre-boarding payment stations, plus 35 kilometers of trunk routes in mixed traffic roads operated by middle capacity buses with doors on the right (for regular roads) and left (for integration with the Metrolínea corridor boarding stations). The system also includes feeder services operated by small to medium sized buses with right side doors. All services have an integrated fare. Two 15 year concessions were tendered for the Metrolínea system and the first phase was introduced in 2009.

Payment mechanism

Operators are paid according to the following formula:[63]

$$IngOper_{j,i} = \left(\sum_{s=1}^{3} P_{s,j,i} \right) * FNS_{j,i}$$

where,

$IngOper_{j,i}$:   Income of operator j in period i.

$P_{s,j,i}$:   Payment of operator j, in period i, for services s, where s indexes feeder, normal ('*padrón*') and articulated bus services.

$FNS_{j,i}$:   Service factor of operator j in period i.

In turn, the payment for each service type is determined according to the following:

$$P_{s,j,i} = T_{s,j,i} * KM_{s,j,i}$$

where $T_{s,j,i}$ is the fare for service s tendered by the operator (adjusted for cost inflation to period i) and $KM_{s,j,i}$ are the kilometers of service type s undertaken by operator j in period i.

The period i corresponds to each two week period.

---

[61] I appreciate the grateful help of Ana María Perez (MPP student, Department of Economics, University of Chile) for the background search for this appendix.

[62] The description of this case is based on the documentation of Sistema Integrado de Transporte Masivo del Área Metropolitana de Bucaramanga, Licitación Pública M-LP-003-2007.

[63] This is a simplified version of the formula in the bidding documents.

Clean Air Institute

The *FNS* is parameter for operator j in period i and depends on the value of an Index of Performance Quality (*ICD*). The *ICD* is a weighted average of eight performance indicators:[64]

- *Regularity* (10% weight): this is an index measuring frequency compliance;
- *Punctuality* (10% weight): refers to time schedule compliance of services;
- *Operational* (15% weight): measures the compliance with several driving, parking and other operational behavior of buses;
- *Accidents* (20% weight): related to the number and seriousness of accident per kilometer of service;
- *Bus cleanliness and reliability* (15% weight): measures the number of mechanical problems and cleanliness of buses;
- *Environmental compliance* (20%): this is a weighted average of the compliance with the operator's Environmental Management Plan, the emission level of gases of the fleet, the level of noise pollution of the fleet and other environmental variables.
- *Infrastructure* (5%): based on the state and quality of the infrastructure operated by the concessionaire.
- *Attention to users* (5%): measures the compliance with a series of performance targets related to customer service.

Once the *ICD* index is calculated, the *FNS* measure is defined as follows:

Table A4.1: Service levels, Metrolínea

Service level	FNS	ICD
A	100%	ICD ≥ 97%
B	98%	95% ≤ ICD < 97%
C	95%	93% ≤ ICD < 95%
D	90%	91% ≤ ICD < 93%
E	75%	ICD < 91%

It is interesting to note that given the discrete definition of the *FNS,* operators may have an incentive to keep the *ICD* index close to 97%.

Penalties and sanctions

The contract stipulates a long-list of penalties for administrative, operational, bus quality, environmental and other performance issues. Most important from the point of view of this report are those related to operating performance. These include:

- Operating with service level E for three consecutive periods
- Operating with service level E during five periods during the last 12 months
- Departure delay of more than 2 minutes over scheduled departure time
- Not stopping at bus stops and other drivers' anomalies.

---

[64] The definition of each of these indices is contained in Annex 2 of the concession contract.

The contract rests mostly on the payment mechanism (*FNS*) to guarantee quality of service.

Also it must be noted that the concession contract terminates if the operator operates with service level E for 15 periods in a 12 month period. Considering the experience in Transantiago with similar draconian measures, it is probable that such drastic sanctions are not very effective in providing incentives since they may not be credible.

Driver contractual restrictions

The contract does not contain any clause related to drivers' compensation.

Summary of experience

Consistent with the BRT nature of the Bucaramanga system, contracts use mostly operational variables to determine payment (mainly kilometers of service and performance variables). There is no short-run risk demand for operators. However, if demand decreases significantly, the kilometers of service in the operational plan may be reduced affecting an operator's income indirectly.

In fact, as noted by one reviewer of this report, the above-mentioned risk has already occurred. The demand estimation for the system was too optimistic and operational revenue is below planned levels. Since subsidies were not considered in the design of this reform, the authorities, which have control of dispatch of the system, have lowered the frequency of service in order to reduce the number of kilometers of service and this way reduce operational costs.

2. Transmilenio, Bogotá, Colombia[65]

General description

Transmilenio is the BRT system in Bogotá, the capital of Colombia. It was first inaugurated in 2000, with 41 kilometers of exclusive bus corridors for trunk services and a number of feeder services. It was gradually been expanded to 87 kilometers and 115 stations of trunk routes and 90 feeder services (Phase I and II) and with the current opening of Phase III it has 109.2 kilometers. The Transmilenio system operates alongside the traditional chaotic and informal public transport system that is being upgraded with the SITP reform discussed in the next section.

An interesting feature of the Transmilenio contracts is that the duration of the concession is not fixed. Rather it is a variable length concession that ends when the average use of the fleet reaches 850.000 kilometers.

---

[65] This review is based on the 'Contrato de concesión operación troncal del sistema transmilenio' available from http://www.worldbank.org/pppiresource and only covers trunk services.

Payment mechanism

The payment mechanism for trunk services in the Transmilenio system is unusual for these types of contracts in that demand risk is shared among all operators. The payment of operator k is determined according to the following formula (*Clause 77*):[66]

$$ROT_k = \frac{C_k \cdot Km_k \cdot Av_k}{\sum_{i=1}^{n} C_i \cdot Km_i \cdot Av_i} \cdot Ingresos$$

where,

C_k :   tariff bid by the concessionaire during the tendering stage

km_k :   kilometers of service provided by operator k

Av_k :   an adjustment for the average service velocity of operator k

Ingresos: are the systems available funds from passenger income after deducting payment to feeder operators, revenue operator and other service providers in the system

i: indexes each trunk concessionaire

The AV_k parameter is a convex function of velocity that increases with the average speed of services.

Thus, the payment formula can be summarized as follows. First all of the systems income (net of payment for other services) is pooled together, and payments are discounted for BRT Agency, Trust Fund, FCS and feeders. Then it is partitioned among operators according to their share in the total value of kilometers of service rendered among all concessionaires adjusted for speed.

There are several things to note from this formula. First, concessionaires do face demand risk, but at the aggregate system level. That is, they suffer financial losses (less income) if trips decline in the aggregate not just in the number of transported by each firm. It is unclear how strong the incentives are in this scheme to provide demand-enhancing quality of service.[67] Free-riding on other operators' efforts to raise demand may be attractive. On the other hand, if there are few trunk operators in the system they may be able to coordinate and provide quality improvements that increase demand.

---

[66] This is not the exact notation contained in the contract but it is completely equivalent except for one term that was not included relating to payments to and from the Contingency Fund.

[67] The idea is that if an individual operator exerts costly effort to enhance some quality variable the benefits will be shared among all operators. Therefore, each operator has the incentive to free-ride on the effort of the other operators. At the aggregate level, the equilibrium level of effort will be lower than what would be exerted by each operator in the alternative scenario where they kept all the fruits of enhanced performance.

Penalties and fines

The contract stipulates a long-list of penalties for administrative, operational, bus quality, environmental and other performance issues. Most important from the point of view of this report are those related to operating performance. These include frequency, regularity and customer service.

There is also a system of bonuses for each concessionaire based on an index of frequency, punctuality and customer service ratings (Clause 78, 79 and 80). However, these bonuses are paid from the fines funds, so are probably not very significant in monetary terms. However, they probably do generate incentives for performance at the margin.

Driver compensation restrictions

There are no clauses related to drivers' compensation, only that drivers must be hired under formal conditions and must be trained.[68]

Summary of experience

The Transmilenio trunk services are mainly paid according to operational variables. However, there is some demand risk faced by concessionaires, but only at the aggregate level. In the case of these contracts, performance issues (besides kilometers of service) are enforced through fines.

3. SITP, Bogotá, Colombia[69]

General Description

Following the success of Transmilenio a decade earlier, the authorities in Bogotá decided to reform the complete public transit system. Up to this point, the modern Transmilenio operated alongside the informal and chaotic traditional public transport system that accounted for the majority of public transport trips.

This reform, called Integrated Public Transport System (SITP) began operations in mid 2012 and will be phased-in in 18 months. The reform includes trunk services and feeder services. As the previous Transmilenio feeder concession contracts end, these services will be taken over by the new zone and inter-zone operators.[70]

At the time of writing, 13 of the 614 routes operated under the SITP. According to news reports, the system was beginning to have a financial deficit due to competition from buses

---

[68] In some reforms, notably SITP in Bogotá, operators are obliged to hire at least 50% of drivers from the old system.

[69] This review is based on the document "Anexo 4A: Minuta del Contrato" of the Alcaldía Mayor de Bogotá.

[70] The zone and inter-zone services are local services, some of which are feeder services to the trunk routes, although they are not all feeder services.

and routes from the old system as well as the coexistence of two electronic payment cards, making the use of the new system cumbersome for users.[71]

Payment method

In the case of the SITP, the payment mechanism is different for trunk services operating in the exclusive corridors than for zone and inter-zone services.

For trunk services, the payment formula is:

$$RT = f(Q)_T \times \sum_k \left\{ \left[(TMVT_k \times NoVehT_k)/4,3 \right] + \left[\left[(OE_{KmT} \times TKMT_k) - CFR_k \right] \times KMStronc_k \right] \right\}$$

where,

k: indexes the type of vehicle (articulated, bi-articulated or normal 80 passenger bus).

RT: Payment to trunk operator

$f(Q)_T$:   Quality of service function for trunk services

$TMVT_k$:   Payment per vehicle type k in the fleet

$NoVehT_k$:   Number of buses type k in the operator's fleet

OE_{KmT}:   Number equal to or less than one that the concessionaire bid in the tendering process that multiplies the payment per kilometer parameter

$TKMT_k$:   Payment per kilometer parameter for vehicle type k (original value fixed in the bidding documents but is then indexed for cost inflation to calculate payments each period)

$KMStronc_k$:   Kilometers of service by bus type k effectively offered during the payment period

CFR_k: Is an adjustment coefficient for reserve fleet.

The function *f(Q)* is:

$$f(Q) = Max[0.4 \times FIR + 0.6 \times FIP, 0.970]$$

where *FIR* is an index that measures compliance with frequency requirements in the operational plan and *FIP* measures compliance with regularity (punctuality). As can be seen the penalty for non-compliance with frequency and punctuality can be at most 3% of revenues due in each period.

---

[71] See news at: http://www.eltiempo.com/colombia/bogota/ARTICULO-WEB-NEW_NOTA_INTERIOR-12367085.html

In sum, the payment mechanism for trunk services is based on operational variables (vehicles and kilometers offered) and there is no demand-risk. Service quality is enforced through the *f(Q)* performance function but cannot be lower than 0.97.

The payment mechanism for trunk services contrasts with the payment mechanism for zone and inter-zone services. In this case the payment formula is:

$$RZONA_i = \left\{ f(Q)_{Zonali} \times \left[\sum_k \left((TMVZ_{i,k} \times VEH_{i,k} / 4,3) + (TKMZ_{i,k} \times KM_{i,k}) \right) + \left(OE_{PasZonai} \times TPASZ_i \times PP_i \right) \right] \right\} - ARTZ_i$$

Where most of the variables are analogous to the case of trunk services except they refer to Zone *i*. One notable difference is that in this case there is an additional term that depends on the number of passengers transported (*PP*). Also, in this case the *OE* bidding variable multiplies the price paid per passenger transported (*TPASZ$_i$*) instead of the payment per kilometer.[72] The *f(Q)* function for the case of zone and inter-zone services is defined exactly the same as in the case of trunk services.

Therefore, in the case of zone and inter-zone services, operators face demand risk which is approximately 20% of projected income. This is consistent with the conceptual discussion presented in this report whereby it is more important to make operators face demand risk in services outside of exclusive infrastructure, like exclusive corridors, where monitoring performance and payment by users is more costly.

Penalties and fines

The contract establishes a long list of performance measures that if not met give rise to fines. However, it is interesting to note that the amount of fines that can be applied in any payment period is capped to be 10% of payment due.

Driver contractual restrictions

The contracts establish that drivers' compensation cannot be linked to the number of passengers transported. However, it is interesting to note that the contract also established the requirement of keeping a record of each driver detailing the number of kilometers driven, accidents, fines, qualifications, training and any customer complaints.

Summary of experience

The case of the SITP is interesting precisely because the incentives for performance of trunk services are slightly different from zone and inter-zone services. Only in this last case does an operator face some demand risk. As mentioned above, this is consistent with the optimal assignment of incentives according to the differing monitoring costs of each type of service. Since it is more expensive to monitor the zonal services it is rational to provide more incentives through some demand risk for these operators.

---

[72] The last term of the above formula (*ARTZ$_i$*) is an adjustment for non-trunk operations and only applies if the concessionaire is in charge of operating zonal terminals.