How do different payment schemes affect public transport concessions and social welfare? A microeconomic model

Juan Pablo Sepúlveda, Pontificia Universidad Católica de Chile, jasepulveda@uc.cl
Patricia Galilea, Pontificia Universidad Católica de Chile, pgalilea@ing.puc.cl

Abstract

In public transport concessions, the contract that relates the operator of the services with the regulating entity defines the transport system. In particular, the mechanism of payment to the operator, established in the contract, encourages it to act in a certain way in the provision of the service. Thus, the design of contracts – how to pay the operator – largely determines the performance of the system in question.

Different payment mechanisms have been implemented in world’s transport systems. These are based on criteria that can be grouped into four types, among others: fixed-price, cost-plus, payment for operation variables and payment per transported passenger. Several authors have studied the advantages and disadvantages of distinct types of mechanism (Batarce & Galilea, 2013; Briones & Gómez-Lobo, 2014; Gagnepain & Ivaldi, 2002; Dalen & Gómez-Lobo, 2003). For the case of Transantiago (the public transport system of Santiago de Chile), the payment scheme considers operation variables and served demand; in addition, it establishes fines in case of failure to comply with performance indicators. The establishment of fines and/or bonuses in the contract is being used worldwide and its analysis could give insights on how to incentivize operators to achieve greater social welfare.

In this study, we are interested in comparing the effect of different payment mechanisms on the performance of the system and on social welfare. We are also interested in measuring the effect of incorporating into the contract bonuses and fines for compliance of performance indicators. Thus, we develop a microeconomic model that includes several elements, some of which are mentioned below.

We consider a generalized demand function for trips, which includes the fare, waiting time and travel time of users. We also consider a cost function of operators that depends on the frequency offered and the capacity of vehicles. In addition, we include a term associated with the social cost of the subsidy required to operate the system. In this way, we define a social welfare function that includes all those terms. In the same way, we define a utility function of the operator, given by its payment minus its costs.

In Transantiago, the payment to operators is composed, on average, of 70% for transported passengers and 30% for operational variables; in addition, this payment is higher than the system fare income, so a subsidy is required. In this context, and with a
rational behavior of the operator, our results indicate that the payment for operational variables is an incentive for the operator to produce more frequency, which would allow the regulator to replicate a socially optimum scenario. Our model also indicates that the payment per passenger transported is an incentive to the operator to pick up passengers, but from a certain level increasing it does not report greater social benefit. Hence, more complex (non-linear) payment schemes can be formulated that meet the conditions of incentives, but which mean less expenditure of public funds.

Our findings support the idea that mixed schemes are more convenient than those of pure operation, in bus systems similar to Transantiago; and we provide a microeconomic framework that can be used in future studies. We give insights on how to put the incentives so that the operators are aligned with the social objectives. The new Transantiago contracts, to be released in the next tendering process, are promising in this regard. Among the changes to the payment scheme is a bonus (up to 10% of the payment without bonus) for compliance of indicators, which goes in the right direction in terms of incentives.

Keywords: contract design, public transport concessions, bonuses/fines in contracts, incentives.
1. Introduction

In public transport concessions, the contract that relates the operator of the services with the regulating entity defines the transport system. In particular, the mechanism of payment to the operator, established in the contract, encourages it to act in a certain way in the provision of the service. Thus, the design of contracts – how to pay the operator – largely determines the performance of the system in question.

Different payment mechanisms have been implemented in world’s transport systems. These are based on criteria that can be grouped into four types, among others: fixed-price, cost-plus, payment for operation variables and payment per transported passenger. Several authors have studied the advantages and disadvantages of distinct types of mechanism (Batarce & Galilea, 2013; Briones & Gómez-Lobo, 2014; Gagnepain & Ivaldi, 2002; Dalen & Gómez-Lobo, 2003). For the case of Transantiago, the public transport system of Santiago de Chile, the payment scheme considers operation variables and served demand; in addition, it establishes fines in case of failure to comply with performance indicators.

The establishment of fines and/or bonuses in the contract is being used worldwide and its analysis could give insights on how to incentivize operators to achieve greater social welfare. In this study, we are interested in comparing the effect of different payment mechanisms on the performance of the system and on social welfare. Thus, we develop a microeconomic model and make simulations that allow us to contrast results.

The article is organized as follows. In section 2 we present a review of the relevant bibliography in the topics of transport service design and cost estimation, which are related to contracts. Together with this review, we give a brief account of the evolution of Transantiago contracts and their effects on the performance of the system, an evolution that inspires this work. In section 3 we present the model that we are going to evaluate. In section 4 we show the simulations performed and the results obtained from them. Finally, in section 5 we draw conclusions about these results and their implications.

2. Literature review

We divide this section into two parts. In the first one we show a review of the literature that reaches this work. In the second part we give an overview of the history of Transantiago contracts, which is relevant for the development of our model.
2.1. Contracts in transport concessions

We will provide a brief overview of the literature on contracts in transport markets. This summary is presented in three parts. The first part deals with part of the theoretical literature on economics and regulation. The second part gives an idea of the empirical literature on contracts of transport concessions. The third part presents some of the theoretical and numerical works specifically related to public transport markets.

There is a very relevant theoretical literature about contracts that regulate concessions. The provision of public services through concessions is characterized by information asymmetries, where the operator has more information about project costs than the regulator. A fundamental work is that of Laffont and Tirole (1986), who show that, under risk neutrality, the optimal contract is linear in the costs of the project. This optimal scheme has at its extremes fixed-price and cost-plus contracts.

The optimal contract of Laffont and Tirole is rather theoretical: it is difficult to calculate in real contexts and requires much information. However, Rogerson (2003) showed that, under certain mathematical conditions, the contracts of extremes, which are simpler than the optimal contract, allow to achieve much of the efficiency of it. Thus, fixed-price and cost-plus contracts may be good enough in regulatory contexts, and are widely used in practice. These works, however, do not consider the particularities of transport markets - for example, the costs of users due to the time they spend on transport. Our work addresses some of these characteristics.

Regarding the empirical literature about contracts in transport concessions, most of the work consists of estimating cost functions of transport systems and deriving measures of efficiency. Some of these articles are Gagnepain and Ivaldi (2002), Dalen and Gómez-Lobo (2003), and Batarce and Galilea (2013). As an example, the latter estimate cost functions for Transantiago operators, with contracts prior to the renegotiation of 2012. They also estimate an aggregate demand model and compare different price schemes.

A conceptual work that includes essential elements of public transport contracts is that of Briones and Gómez-Lobo (2014). They analyze incentives to the operator associated with different payment schemes - based on demand or supply. A payment associated with the number of passengers carried by the operator encourages him to stop at the stops and pick up the passengers; and to offer high frequency and regularity, if demand is high. But it also faces greater risks (demand, financial) and encourages it to an unsafe driving of vehicles when competing for passengers. On the other hand, a payment associated with the seats-kilometer traveled reduces the demand risk of the operator and encourages him to drive safely, but he no longer has incentives to pick up passengers.
Another branch of the transport literature is the one focused on models that are exclusive of transport, in particular of public transport. It emphasizes in the difference between the costs of the users of the transport system and the costs of the operators. This modeling has become more complex over the years; different authors have added new elements. Some of these papers correspond to Mohring (1972), Jansson (1980) and Jara-Díaz and Gschwender (2005, 2009), who show their results through parameter simulations. Although this literature deals with the characteristics of the transport markets, it does not consider the design of the contracts - the payment to the operators - that is the fundamental part of our work.

The model we formulate in this article is based on the works previously mentioned. We consider cost functions of users and operators similar to these authors. We add the part of the problem related to the demand and also consider the social cost of the subsidy granted to the system, but the main contribution of our work is the incorporation of contract design into microeconomic modeling. We study theoretically how the problem of the planner and the operator changes when we consider different payment schemes to the operator.

2.2. Transantiago contracts

Transantiago began its operations in February 2007, which ended 15 years of micros amarillas (references about this old system in Díaz et al., 2004). Transantiago meant a radical change in the life of the city, since it reformed almost all the aspects of the previous system: modification of the structure of services, new form of payment for the users, integration of fares of buses and subway, and change in the scheme of payment to the operators, among other aspects. The first months since the implementation of the system were disastrous (Muñoz et al., 2009), with gradual improvements over the years.

For this work, we are interested in the evolution of Transantiago contracts and payment schemes. Here we will give a brief summary to contextualize our methodology; for a more detailed account read Briones and Gómez-Lobo (2014). In 2007, when the system started, the contracts only considered the demand served as payment variable. In theory, an amount was paid per passenger transported, but since that amount was adjusted according to a referential demand, the real risk of demand was very low (less than 10%). In practice, it meant that the payment was almost fixed, so there was no incentive to cater demand or offer a good service. While these contracts also included fines for non-compliance of service variables, they were not enforced for two reasons: fines were too high to be executed and the regulator did not have the technology to measure the service variables.
Given the poor results of the system, the contracts were quickly renegotiated. In mid-2007, a discount was added to the payment scheme for non-compliance of the offer of seats stipulated in the operational plan, which this time could be measured. The risk of demand was also slightly increased. These changes improved the quality of the service, but not enough. A deeper reform came in 2008 with the incorporation of frequency and regularity indicators to contracts, and fines for non-compliance. In 2009, the indicator of the seat offer was modified by a seat-kilometer, which measures the real product of the service provided by the operator. By linking the payment to operational variables, Transantiago's service level improved considerably: higher frequency, shorter waiting times and travel times (see figure), but the biggest reform was yet to come. For more details on the indicators and their effects on the performance of Transantiago, see Beltrán et al. (2013).

![Figure 1: Evolution of the average waiting time (min) in peak periods of weekdays. Source: Extracted from Beltrán et al. (2013).](image)

In 2012 the payment scheme was modified: To the payment per transported passengers was added a payment for kilometers traveled, which is deflated by the indicator of compliance of seats-kilometer. In this way, the payment for operational variables went from being a fine or discount to directly counting the operator’s income. In fact, the payment for kilometers traveled is, on average, 30% of the operator’s income.
3. Model

In this section we present the model we use for our analysis of operator incentives. We organize the section in three parts. First, we outline all the elements that define the model. In the second part, we model the first-best scenario, where we ignore the utilitarian behavior of the operator. In the third part, we model the scenario in which the operator makes his own decisions and analyze the incentives behind those.

3.1. Definitions

Consider a market of public transport trips and a generalized price $P(q)$ for making a trip, which includes the fare, the cost of the waiting time and the cost of the travel time of users, and that depends on the demand $q$ in passengers/hour. This demand is distributed uniformly over time.

The provision of the transport service is carried out by a single company, whose cost of operating a fleet of $B$ vehicles of capacity $C$ seats (or passengers) is $B(C_1 + C_2 C)$, where $C_1$ is the fixed cost of operating the capacity and $C_2$ is the marginal cost, both in $$/hour. The utility of the operator is given by the following expression:

$$U = t - B(C_1 + C_2 C)$$

where $t$ is the payment to the operator (defined in its contract), whose functional form we will explain later.

Users of the system experience a cost associated with the use of their time, in addition to the payment of the fee. We will consider a cost associated with the waiting time and a cost associated with travel time. The average waiting time of a passenger is $k/f$, where $f$ is the operating frequency, in vehicles/hour, and $k$ is a parameter associated with the regularity of the vehicles. Thus, the total waiting time of the system (of all passengers) will be $qk/f$, and the cost of that waiting time will be $C_E qk/f$, where $C_E$ is the social value of the waiting time. Similarly, the cost of travel time is given by $C_V t_V q$, where $t_V$ is the travel time and $C_V$ is the social value of travel time.

Users pay a fare for the use of the transport service, which is collected by the regulator (not delivered to the operator). We derive an expression for the system fare, based on the generalized price of a trip and the costs of the users:

$$P = \text{fare} + \frac{k}{f} + C_V t_V$$

$$\text{fare} = P - \frac{k}{f} - C_V t_V$$
and the fare income of the system will be \((P - C_E k/f - C_V t_V)q\).

If the payment to the operator is greater than the system fare income, a state subsidy is required to complete the payment. A money unit of subsidy has a cost for the state of \(1 + \lambda\) money units, where \(\lambda\) is the cost of public funds.

In this way, social welfare \(W\) is defined by the following expression:

\[
W = \int_0^q P(x) \, dx - \left( P(q) - C_E \frac{k}{f} - C_V t_V \right) q - C_E \frac{k}{f} q - C_V t_V q + t - B(C_1 + C_2 C) \\
- (1 + \lambda) \left( t - \left( P(q) - C_E \frac{k}{f} - C_V t_V \right) q \right)
\]

where the first four terms correspond to the consumer surplus, the two following terms to the producer surplus, and the last term to the social cost of the subsidy. We can write this expression more abbreviated, which is as shown below:

\[
W = \int_0^q P(x) \, dx - C_E \frac{k}{f} q - C_V t_V q - B(C_1 + C_2 C) - \lambda \left( t - \left( P(q) - C_E \frac{k}{f} - C_V t_V \right) q \right)
\]

An assumption underlying this expression is that the payment to the operator is greater than the system fare income, that is, the term associated with the subsidy is negative. Otherwise, we would be saying that the system generates income to the planner, which is not the case of most transport systems. In the extreme case the payment to the operator is equal to the system revenue, which corresponds to a second-best scenario.

If we make some assumptions about the operation of the service, it is possible to obtain expressions for fleet size and travel time. In particular, we will use the terms proposed by Jara-Díaz and Gschwender (2009). Suppose that the service consists of a circular line having a total length of \(L\) kilometers, and that passengers travel a distance of \(d\) kilometers of it (with \(d < L\)). Suppose also that the vehicles take a time of \(T\) hours to cross the whole line (excluding the stops), that each passenger takes a time of \(t_s\) hours to get on the vehicle and that the time of boarding dominates to the time of descent\(^1\). The demand \(q\) is distributed homogeneously along the line.

With these assumptions the cycle time of the system is defined by \(T + t_s q / f\), where \(q / f\) is the number of passengers that board each vehicle in the line. If we use that the frequency is the quotient of the fleet size with the cycle time, we can derive the following expression for the fleet size:

\[
B = f t + t_s q
\]

---

\(^1\) This means that the board and descent occur in parallel, and that the boarding time is greater than the descent time. Thus, the only relevant time is the boarding one, because the descent time is included in it.
In addition, the travel time experienced by each passenger is the part of the cycle time that corresponds to his trip, i.e. the part $d/L$ of the cycle time. Thus, the travel time of each passenger is given by:

$$ t_V = \frac{d}{L} \left( T + t_s \frac{q}{f} \right) $$

With these expressions for $B$ and $t_V$, the social welfare function is as follows:

$$ W = \int_0^q P(x) dx - C_E \frac{k}{f} q - C_V \frac{d}{L} \left( T + t_s \frac{q}{f} \right) q - (Tf + t_s q)(C_1 + C_2 C) 
- \lambda \left( t - \left( P(q) - C_E \frac{k}{f} - C_V \frac{d}{L} \left( T + t_s \frac{q}{f} \right) \right) q \right) $$

### 3.2. First-Best Scenario

Let's consider the problem of maximizing social welfare, without operator financing constraints. The solution of this problem has associated the maximum possible social welfare, given the values of the parameters, but also has associated a negative utility of the operator. This implies that such a solution is not feasible in reality, but serves to compare with other scenarios. The mathematical formulation of the problem and its solution are shown below:

$$ \max_{q,f,C} W $$

$$ s.t. \quad \frac{q d}{f L} \leq C $$

The imposed restriction is of operation and indicates that the capacity of the vehicles must be greater than or equal to the necessary capacity for all the passengers who are waiting in the stops to be able to board the next vehicle to arrive. The optimum capacity will be equal to that necessary capacity, that is, the minimum where that number of passengers can fit. Increasing capacity only increases operator costs; there is no benefit for greater capacity, according to our definition of social welfare. So,

$$ C^{FB} = \frac{q d}{f L} $$

We replace this expression in the social welfare function:
\[ W = \int_0^q P(x) \, dx - C_E \frac{k}{f} q - C_V \frac{d}{L} \left( T + t_s \frac{q}{f} \right) q - (T f + t_s q) \left( C_1 + C_2 \frac{q}{f} \frac{d}{L} \right) \]
\[ - \lambda \left( t - \left( P(q) - C_E \frac{k}{f} - C_V \frac{d}{L} \left( T + t_s \frac{q}{f} \right) \right) \right) \]

We reorder the terms in a convenient way for the algebra that follows:

\[ W = \int_0^q P(x) \, dx - \left( (1 + \lambda) C_E k + ((1 + \lambda) C_V + C_2) \frac{d}{L} t_s q \right) \frac{q}{f} \]
\[ - \left( (C_V + C_2) \frac{d}{L} T + t_s C_1 \right) q - T C_1 f - \lambda (t - P(q) q) \]

The first-order condition and the optimum frequency are as follows:

\[ \frac{\partial W}{\partial f} = \left( 1 + \lambda \right) C_E k + (1 + \lambda) C_V + C_2 \frac{d}{L} t_s q \right) \frac{q}{f^2} - T C_1 - \lambda \frac{\partial t}{\partial f} = 0 \]

\[ f_{FB} = \frac{q}{\sqrt{T C_1 + \lambda \frac{\partial t}{\partial f} \left( 1 + \lambda \right) C_E k + (1 + \lambda) C_V + C_2 \frac{d}{L} t_s q}} \]

In this part we must make an assumption about the functional form of payment \( t \). In section 2 we said that the payment in Transantiago is linear in the passengers transported and in the kilometers traveled. Since we are considering the frequency and capacity as the operating variables, we will use those to model the payment, rather than the kilometers traveled. In an ideal corridor the latter are directly proportional to the frequency and the capacity, so it is a good approach the one we are using. In particular, we will say that the payment has the following expression:

\[ t = t_0 + \alpha_q q + \alpha_f f \]

where \( \alpha_f \) is the payment for vehicles/hour produced, \( \alpha_q \) is the payment per transported passenger and \( t_0 \) is a fixed payment (which can be zero).

Being the payment to the operator linear in the frequency, we can express the optimum frequency and the social welfare that includes it explicitly. If the payment is non-linear, such expressions will generally be implicit.

The expression of the optimal frequency is now:

\[ f_{FB} = \frac{q}{\sqrt{T C_1 + \lambda \alpha_f \left( (1 + \lambda) C_E k + (1 + \lambda) C_V + C_2 \right) \frac{d}{L} t_s q}} \]
We can see that the optimum frequency decreases with the payment associated to the frequency: by increasing the society’s expense in paying the operator, it is optimal to provide a lower frequency. Regarding the variation of $\lambda$, its effect is not clear: the parameter is in the numerator and denominator of the expression. The simulation of the next section will give us some insight into this effect.

We replace the last expression in the social welfare function:

$$W = \int_0^q P(x)dx - 2\sqrt{(TC_1 + \lambda \alpha_f)\left((1 + \lambda)C_Ek + ((1 + \lambda)C_V + C_2)\frac{d}{L}t_s q\right)}q$$

$$- \left((C_V + C_2)\frac{d}{L}T + t_s C_1\right)q - \lambda(t_0 + \alpha_q q - P(q)q)$$

The first-order condition of the demand is as follows:

$$\frac{\partial W}{\partial q} = P(q) - \frac{\partial}{\partial q}\left(2\sqrt{(TC_1 + \lambda \alpha_f)\left((1 + \lambda)C_Ek + ((1 + \lambda)C_V + C_2)\frac{d}{L}t_s q\right)}q\right)$$

$$- (C_V + C_2)\frac{d}{L}T - t_s C_1 - \lambda\left(\alpha_q - P'(q)q - P(q)\right) = 0$$

From this condition we can deduce that the optimal demand decreases with the increase of $\lambda$, $\alpha_f$ or $\alpha_q$. All these parameters are associated with costs (the negative terms) and their increases imply a greater burden for the regulator mainly. It is also worth noting that the term associated with subsidy cost (in absolute value) is increasing in $q$, when $\lambda > 0$.

Second order conditions of frequency and demand are met.

In general, the utility of the operator is negative in the first-best scenario (unless the payment parameters are arbitrarily high), i.e.:

$$U_{FB}^F = t - (Tf_{FB}F + t_s q_{FB}^F)(C_1 + C_2C_{FB}^F) < 0$$

3.3. Maximizing profits

Let us now consider the problem of maximizing the utilities of the operator, as opposed to the problem of maximizing social welfare. The solution of this problem has associated a social welfare smaller than the previous problem, because to increase the profits of the operator it is incurred in a social loss. The formulation of the problem and its solution are as follows:
\[
\max_{q, f, c} U = t - (T f + t_s q)(C_1 + C_2 C)
\]

s.t. \[\frac{q \, d}{f \, L} \leq C\]

The optimal capacity is, as in the first-best problem, the minimum capacity where all the passengers that are waiting at the stops when the vehicle arrives fit:

\[C^o = \frac{q \, d}{f \, L}\]

We replace this expression in the utility function:

\[
U = t - (T f + t_s q) \left(C_1 + C_2 \frac{q \, d}{f \, L}\right)
\]

The first-order condition and the optimum frequency are as follows:

\[
\frac{\partial U}{\partial f} = \alpha_f - T C_1 + t_s C_2 \frac{q^2 \, d}{f^2 \, L} = 0
\]

\[f^o = q \sqrt{\frac{t_s C_2 d / L}{T C_1 - \alpha_f}}\]

The optimum frequency increases with the frequency payment. We assume that \(T C_1 > \alpha_f\). If this were not so, the optimum frequency would tend to infinity. The second order condition for this expression is met. We conveniently replace the expression in the utility function:

\[
U = t_0 + \left(\alpha_q + \alpha_f \sqrt{\frac{t_s C_2 d / L}{T C_1 - \alpha_f}}\right) q - \left(\frac{T C_1 \sqrt{t_s C_2 d / L}}{T C_1 - \alpha_f} + \sqrt{(T C_1 - \alpha_f) C_2 \frac{d}{L} t_s + C_2 \frac{d}{L} T + t_s C_1}\right) q
\]

where the negative term corresponds to the cost function of the operator. This cost function is linear in demand. This is a good result if we consider that cost estimates of different transport systems, including Transantiago, indicate functions that have constant economies of scale.

We simplify the expression of utility and derive the first-order condition of demand:
\[ \frac{\partial U}{\partial q} = \alpha_q - 2 \sqrt{(TC_1 - \alpha_f)C_2 \frac{d}{L} t_s} - \left( C_2 \frac{d}{L} T + t_s C_1 \right) = 0 \]

For this condition to have a "closed" solution, it would be required that the payment for the marginal passenger were decreasing, which is not the case of Transantiago or other transportation systems with similar payment mechanisms. With a linear payment on demand, the optimal number of passengers to be transported by the operator is infinite or zero, depending on the value of that payment. This implies that the payment per passenger must be greater than the marginal cost of attending it, so that the operator has the incentive to pick up the passenger. This does not mean that the operator will serve infinite passengers. The operator’s solution represents only a potential demand, which is modified by the real demand curve of the market and by the fare set by the regulator.

Let us briefly discuss the functional form of the previous problem. If the payment per passenger were decreasing, the potential demand that the operator would serve would be a finite number, and the solution of the complete problem (shown below) could be maintained the same. A payment scheme of this type would imply greater social welfare, because the total payment to the operator would be lower. A different specification of this problem is that the marginal cost is increasing, and not constant as it results from our definitions. In such a case, and even though the payment per passenger is constant, the operator’s solution will be finite.

Once the incentive problem is resolved, the regulator will set the fare (or, equivalently in our model, the number of passengers) that maximizes social welfare under the operator’s frequency choice. We replace the expression of the latter in the social welfare function:

\[
W = \int_0^q P(x) dx - \left( (1 + \lambda)C_E k + (1 + \lambda)C_V + C_2 \right) \frac{d}{L} t_s q \left( \frac{q}{f} \right)^{\frac{1}{\sigma}}
- \left( C_V + C_2 \right) \frac{d}{L} T + t_s C_1 \right) q - \left( T C_1 + \lambda \alpha_f \right) f^{\circ} - \lambda \left( t_0 + \alpha_q q - P(q)q \right)
\]

\[
W = \int_0^q P(x) dx - \left( (1 + \lambda)C_E k + (1 + \lambda)C_V + C_2 \right) \frac{d}{L} t_s q \left( \frac{TC_1 - \alpha_f}{t_s C_2 d / L} \right)
- \left( C_V + C_2 \right) \frac{d}{L} T + t_s C_1 \right) + \left( T C_1 + \lambda \alpha_f \right) \left( \frac{t_s C_2 d / L}{TC_1 - \alpha_f} \right) q
- \lambda \left( t_0 + \alpha_q q - P(q)q \right)
\]

The first-order condition for \( q \) in this scenario is as follows:
\[
\frac{\partial W}{\partial q} = P(q) - \left((1 + \lambda)C_V + C_2\right) \frac{d}{L} T - t_s C_1
\]

\[
- \left(\frac{TC_1 - \alpha_f}{t_s C_2 d/L}\right) - (C_V + C_2) \frac{d}{L} T - t_s C_1
\]

The terms associated with the costs of the operator and the users are constant in \( q \), and with a generalized price decreasing in \( q \) the solution of this condition is finite. With respect to the variation of the parameters, in the cases of \( \lambda \) and \( \alpha_q \) the behavior of the solution is similar to the first-best scenario one: the optimal demand decreases with the increase of these parameters. In the case of \( \alpha_f \), the expression above shows that the terms would move in different directions, but from the simulation that we will show in the next section we will see that the effect of the increase of the optimal demand dominates.

4. Simulation and results

In this section we will simulate data for the parameters of our model and we will obtain results for the variables of the problems raised in the previous section. We will compare different payment schemes with these results. The parameters of the operator cost function and simulated ideal corridor are based on Jara-Díaz and Gschwender (2009), and are shown in the following table:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( C_1 ) ($/h)</td>
<td>5000</td>
</tr>
<tr>
<td>( C_2 ) ($/h)</td>
<td>100</td>
</tr>
<tr>
<td>( d ) (km)</td>
<td>10</td>
</tr>
<tr>
<td>( L ) (km)</td>
<td>60</td>
</tr>
<tr>
<td>( t_s ) (h)</td>
<td>0.00139</td>
</tr>
<tr>
<td>( T ) (h)</td>
<td>3</td>
</tr>
<tr>
<td>( C_E ) ($/h)</td>
<td>2200</td>
</tr>
<tr>
<td>( C_V ) ($/h)</td>
<td>740</td>
</tr>
<tr>
<td>( k )</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Table 1: Parameters of the operator cost function and the ideal corridor.

The generalized price function we will consider is \( P(q) = 1200 - 0.1q \). This is a linear function that is easy to modify for sensitivity analysis.

Let’s start by looking at the effect of the variation of \( \lambda \) on the frequency of the first-best scenario. In the following table we show the different values of that frequency for two values of the demand: \( q = 2000 \) and \( q = 10000 \). In the first case we see that the
frequency increases with the increase of $\lambda$, and in the second case it decreases. In any case, the frequency variation is very little, so $\lambda$ is not a critical parameter in this sense.

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>$f^{FB}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$q = 2000$</td>
</tr>
<tr>
<td>0</td>
<td>14,090</td>
</tr>
<tr>
<td>0.1</td>
<td>14,113</td>
</tr>
<tr>
<td>0.2</td>
<td>14,132</td>
</tr>
<tr>
<td>0.3</td>
<td>14,148</td>
</tr>
<tr>
<td>0.4</td>
<td>14,162</td>
</tr>
<tr>
<td>0.5</td>
<td>14,175</td>
</tr>
<tr>
<td>0.6</td>
<td>14,186</td>
</tr>
<tr>
<td>0.7</td>
<td>14,195</td>
</tr>
<tr>
<td>0.8</td>
<td>14,204</td>
</tr>
<tr>
<td>0.9</td>
<td>14,212</td>
</tr>
<tr>
<td>1</td>
<td>14,219</td>
</tr>
</tbody>
</table>

Table 2: First-Best frequency for different values of the cost of public funds.

Let’s now look at the demand charts for the first-best scenario. First consider the case where $\lambda = 0$, that is, the subsidy has no cost to the state, or public funds are infinite. The following chart shows this case, typical of the literature on transport economics (Jara-Díaz and Gschwender, 2009, among others). The average cost of the user (UAC) and total cost (TAC) are decreasing, and the total marginal cost (TMC) is also. Since the optimal demand is at the intersection of the price with the total marginal cost, and that the intersection is below the total average cost, a subsidy is required for the system to operate at that point.

![Figure 2: Cost curves and optimum demand in the first-best scenario when $\lambda = 0$.]
Second, consider the case where $\lambda > 0$, in particular $\lambda = 0.1$. The following graphic shows this case for some parameters of the payment to the operator, in particular $\alpha_f = 8000$ and $\alpha_q = 90$. We can see that when public funds have a cost, the average and marginal costs cease to be decreasing in the interest range and begin to increase at some point. The optimal demand is less in this case than when $\lambda = 0$, because the payment to the operator has a social cost.

![Figure 3: Cost curves and optimum demand in the first-best scenario when $\lambda = 0.1$.](image)

Let's see now what happens to the operator's decisions. First, we show a graph where we compare the first-best frequency with the operator's frequency for three values of $\alpha_f$ (the frequency payment). We use $\lambda = 0$ to focus only on the incentive effect. It is seen that, with the increase of this incentive, the frequency of the operator increases, to the point of reaching the first-best frequency for $\alpha_f$ large.
Second, let's look at the graph of the operator problem. For him to have incentives to pick up passengers, the marginal benefit (MB) of taking them should be greater than the associated marginal cost (OMC). This marginal cost is constant because, as we derive in the previous section, the operator's private cost is linear. And with a linear payment scheme, as shown in the graph, the operator will choose to take of all the demand that is available (once the fare has been set). The parameters considered are the same as above: $\alpha_f = 8000$ and $\alpha_q = 90$. 

![Figure 4: Comparison between the first-best frequency and operator's frequency.](image)

![Figure 5: Marginal cost and marginal benefit curves of the operator in the maximizing profit scenario.](image)
Third, let's look at cost curves when the operator chooses its frequency, compared to the costs of the first-best scenario. Since the frequency chosen by the operator is lower than the first-best (with $\alpha_f = 8000$), the average cost of the user is higher than its first-best pair. As for the cost of the operator, this is lower than its first-best pair, but the average total cost increases because the increase of the cost of the users is the dominant effect. We consider $\lambda = 0.1$, $\alpha_q = 90$ and $\alpha_f = 8000$.

Let’s compare the optimal demands of the first-best scenario and the operator’s. The following graph shows the marginal costs of both scenarios, and we can see that the operator optimum demand is less than the socially optimal one.

Figure 6: Comparison for average cost curves when using the first-best frequency and operator's frequency.

Figure 7: Optimum demand in the first-best and maximizing profit scenarios.
Now we are going to compare different values of the payment parameters with respect to the performance of the system and the associated served demand. In the following table we show three payment schemes, the two scenarios we have discussed so far and the results for each combination: frequency, demand, operator profit and social welfare. We consider $\lambda = 0,1$.

<table>
<thead>
<tr>
<th>Scheme</th>
<th>First-Best</th>
<th>Max U</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency (veh/h)</td>
<td>32</td>
<td>11.5</td>
</tr>
<tr>
<td>Demand (pax/h)</td>
<td>6.350</td>
<td>6.150</td>
</tr>
<tr>
<td>Profit ($)</td>
<td>-43.480</td>
<td>46.720</td>
</tr>
<tr>
<td>Welfare ($)</td>
<td>1.947.000</td>
<td>1.382.418</td>
</tr>
<tr>
<td>$t = 95q + 8000f$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frequency (veh/h)</td>
<td>31.4</td>
<td>32.1</td>
</tr>
<tr>
<td>Demand (pax/h)</td>
<td>6.330</td>
<td>6.330</td>
</tr>
<tr>
<td>Profit ($)</td>
<td>151.450</td>
<td>151.450</td>
</tr>
<tr>
<td>Welfare ($)</td>
<td>1.928.000</td>
<td>1.928.000</td>
</tr>
<tr>
<td>$t = 95q + 14100f$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frequency (veh/h)</td>
<td>33</td>
<td>7.2</td>
</tr>
<tr>
<td>Demand (pax/h)</td>
<td>6.380</td>
<td>5.790</td>
</tr>
<tr>
<td>Profit ($)</td>
<td>-312.600</td>
<td>-24.400</td>
</tr>
<tr>
<td>Welfare ($)</td>
<td>1.972.000</td>
<td>825.000</td>
</tr>
</tbody>
</table>

Table 3: Results for three different payment schemes in the two scenarios.

The first payment scheme is the one we used to display the graphs. For that scheme, the first-best frequency is quite different from that of the operator. While demand does not seem to differ so much, social welfare does decline considerably in the operator's scenario compared to the first-best. The utility of the operator is negative in the first scenario and smaller than in the second one, but it is enough to grant a fixed payment (which does not depend on the demand or the frequency) that covers the loss to enable the operation in these conditions. Such a fixed payment would diminish social welfare a little (because the state pay it), but does not change the solution of the problem; so, it is easy to implement.

The second scheme in comparison increases the payment associated with the frequency with respect to the first scheme. With this new value the operator will choose a frequency similar to that of the first-best, and the optimum demand of the two scenarios will be the same. Thus, the operator's profits and social welfare also coincide. In this way, the regulator has a tool to modify the incentives of the operator and forces him to produce efficiently in social terms. It is worth noting that the solution is not the same as the previous scheme, the welfare is a little lower with this scheme that pays more since the state would have a higher expense by paying more for the frequency.

The third scheme in comparison does not pay per frequency. The payment per passenger transported was increased to 140 to ensure compliance with the incentive to take passengers, commented earlier. The solution to the first-best scenario is similar to that of
the other schemes, but the operator solution is significantly worse than the first-best scenario. Without incentives to operate, the company decides to produce a low frequency to lessen its losses. We insist that the negative profits can be alleviated with a fixed payment, but we show the results without that additional payment to compare.

5. Conclusions

In this article we show a microeconomic model of public transport that we developed to compare different payment schemes to the operator of a concession. The payment schemes modeled are of a linear type in the passengers transported and in the frequency of operation, schemes inspired by the current mechanism of payment of the public transport of Santiago de Chile. The results of the simulations indicate that the frequency payment is an incentive for the operator to produce more frequency, which would allow the regulator to replicate a socially optimum scenario. Our model also indicates that the payment per passenger transported is an incentive to the operator to pick up passengers, but from a certain level increasing it does not report greater social benefit. Hence, more complex (non-linear) payment schemes can be formulated that meet the conditions of incentives, but which mean less expenditure of public funds.

We can conclude working directions in the development of public policies regarding public transport. As indicated in the previous paragraph, the payment for operational variables has a significant impact on the performance of the system and on social welfare. Hence, this element should be given more weight in the contracts that regulate service concessions. In section 2 we commented that the payment for kilometers traveled represents 30% of the total payment to the operator, so there is room to work in this regard. The payment per transported passenger plays an important role in giving incentives relative to the passenger himself, but it may be overestimated in Santiago's current transport system.

This study can serve as a basis for starting to model the economy of contracts in real public transport systems, which are more complex. There are several elements that can be included in an upcoming study: the network of services, the network structure (here we only consider an ideal corridor), fare evasion (which in Santiago is a critical problem), among others. Another extension for this work is to consider real data that may allow concluding more precise recommendations. In particular, when several operators interact in the same system (which is the case in Santiago), it is possible to compare their cost structures and operating decisions, in order to encourage everyone to provide a service in a social efficient way.
References


