A model of urban transport management

Paolo Ferrari*
University School of Engineering of Pisa, Via Diotisalvi 2, 56126 Pisa, Italy

Received 31 March 1998; received in revised form 9 June 1998

Abstract

This paper shows a scheme of urban transport management aimed at solving the externalities due to road traffic. In this scheme both private and public transport modes are managed by a planning authority having control of some decision variables: road pricing, transit ticket prices, and the service characteristics of transit. The multimodal transport system is subject to some constraints: physical and environmental capacity constraints, and budget constraints. In some cases an upper bound is imposed on the ticket price, in order to help people who are captive of transit. The planning authority fixes the level of all user charges and the transit service characteristics, in such a way that the average transport generalized cost is minimized, and at the same time the transport system is in equilibrium and all constraints are satisfied. A mathematical model of this management scheme is presented, and it is applied to the bimodal network of an Italian town, in order to show the effects that the management policy, and in particular the constraints imposed on the transport system, have on urban transport cost. Among other things the results obtained show, in accordance with the opinion of other authors, that the market of urban public transport is a niche market, whose properties, along with the characteristics of transit systems, are essentially determined by the physical and environmental capacity constraints of the road transport system. © 1998 Elsevier Science Ltd. All rights reserved.

Keywords: Urban transport management; Road pricing; Capacity constraints

1. Introduction

An individual or a system create externalities when the actions they carry out produce positive or negative effects on other individuals or systems, and the latter do not pay or receive any compensation for these effects. In these cases the individual or the system behave inefficiently, because they do not minimize the social cost of their actions.

* Tel: +39-50-550421; fax: +39-50-553573; e-mail: p.ferrari@ing.unipi.it

0191-2615/98/$—see front matter © 1998 Elsevier Science Ltd. All rights reserved.
PII: S0191-2615(98)00027-7
Negative externalities are usually associated with road traffic. A driver who travels a road causes damages for other users, and generally for all people living in the neighbourhood, without their receiving any compensation for these damages. Following the conventional theory, the damages an additional driver causes for the other road users are measured by the difference between the marginal social cost borne by road users as a whole and the private cost of the driver. As a consequence of this externality, both the number of cars which travel a road network, and the way they use the paths connecting the various O/D pairs, are different from those which would maximize the total utility.

Road pricing, i.e. the imposition of a tax on the use of roads, equal to the difference between social marginal and private cost, was proposed in the twenties in order to eliminate this externality [see e.g. Morrison (1986) for a survey of the subject].

Beckmann et al. (1956) have shown that, in case of road networks with separable cost functions, a means to levy this tax is the imposition on each link of the network of a toll equal to the product of the flow travelling the link by the derivative of the link cost function. In this case the user equilibrium flow pattern is an untolled system optimum, i.e. a flow pattern which maximizes the benefit of users as a whole if one computes the benefit without considering the toll paid by users. Recently Hearn and Motakury (1998) have shown that a set of congestion tolls exists which can attain the same goal, and various criteria can be used for choosing a particular toll in this set.

Some remarks can be made on this theory of road pricing. First of all, one can observe that, while the purpose of the theory is to maximize the utility of drivers, the function which is actually maximized neglects the toll paid by them; and if this toll were included in the driver costs, one could easily verify that the cost deriving from the imposition of road pricing is greater than the benefits drivers receive. Thus the behaviour of the transport system is inefficient.

Another remark concerns the fact that the damage deriving from congestion is slight while traffic remains below a critical threshold, which is usually called road capacity. When this threshold is exceeded there is high risk of flow instability, with long queues and great delays. These are the most important consequences of congestion, but the damages they produce cannot be measured by the difference between social marginal and private cost, because the cost functions cannot be defined beyond the capacity threshold, i.e. in the flow range where instability makes transport costs highly dispersed.

Taking this fact into account, a different theory of road pricing has recently been proposed (Ferrari, 1995, 1997), which calculates the additional costs to be imposed on links of a road network in such a way that an equilibrium flow pattern can be obtained which satisfies the capacity constraints. These constraints are not only physical, but even environmental ones.

The environmental damage produced by road traffic concerns the emission of pollutants, the noise, the intrusive occupancy of urban spaces by parking vehicles. It is usually admitted that, when car flow is below a threshold, which is called environmental capacity (Minister of Transport, 1963), the damage produced by traffic is slight, and in any case less than the benefits towns receive from it. When this threshold is exceeded, there is high risk of very great damages.

So this theory considers road pricing as a technical tool to obtain a good behaviour of road transport systems. In order to prevent capacity constraints from being violated, road users have to pay an amount of money which is used outside the transport system, but they do not receive any compensation for this payment. In this way road traffic creates a positive externality which penalizes the drivers. Thus road transport systems used by high demand present this peculiar
characteristic: they create negative externalities if the demand is not controlled, and positive externalities if the satisfaction of capacity constraints is imposed by the demand control. In both cases their behaviour is inefficient.

One of the ways the externalities produced by a system can be solved is the enlargement of the system, in such a way that all the consequences of the actions it carries out remain inside it (Siglitz, 1986).

Such a system could be created if, as an example, the urban road system included an Authority having the power to assign all money raised by road pricing to the improvement of road capacity. However this is often difficult, because the physical layout of the towns makes it impossible to modify the geometric characteristics of the streets, especially in the central areas of European cities. On the other hand it would be possible to internalize the externalities if this enlarged system included all the urban transport modes, both private and public, which should be managed by a planning Authority having control of some decision variables: road pricing, fares of transit and the service characteristics of the latter. The planning Authority chooses the values of the decision variables in such a way that the utility of the system users as a whole is maximized, and at the same time the system is in equilibrium, and the physical and environmental capacity constraints, along with the budget constraints, are satisfied.

In this way all the externalities would be eliminated. The satisfaction of capacity constraints would prevent the negative externalities, while there would not be positive externalities, because all the amount of money paid by transport users would remain inside the system. In particular the amount obtained from road pricing could be used for improving the road capacity, if this were possible, and thus for making the capacity constraints less tight; or for improving the quality of service of public transport, so that it would become advantageous for some drivers to shift from private cars to transit. In both cases the value of road pricing would be less than that drivers should pay if the amount of money collected were used outside the transport system.

A mathematical model of this urban transport management scheme is the subject of this paper. It shows that the optimization of the objective function of the system is a problem of mathematical programming with equilibrium constraints. This field of applied mathematics [see e.g. Anandalingam and Friesz (1992) for an introduction to the subject] includes many transportation problems, which take into account both supplier and user behaviour: the network design problem (Abdulaal and LeBlanc, 1979; Marcotte, 1983; LeBlanc and Boyce, 1986), the combined problem of signal control and traffic assignment (Marcotte, 1983; Fisk, 1984), the problem of traffic assignment and traffic control in general freeway-arterial corridor systems (Yang and Yagar, 1994), the optimal pricing and/or frequencies in transit systems (Fisk, 1986; Constantin and Florian, 1993; Miyagi and Suzuki, 1996). All these problems consider only one transport mode and are formulated as bilevel programming problems, for which many algorithms have been proposed. Some of them are based on descent approaches for the upper level problem with gradient information obtained from the lower level problem (Falk and Liu, 1995). Another approach (Abdulaal and LeBlanc, 1979) transforms the bilevel programming problem into a sequence of nonlinear programming problems, by using in each step of the sequence the solution of the lower level problem obtained in the previous step. Following Yang and Yagar (1994), one could transform the bilevel programming problem into a sequence of linear programming problems formulated by linearizing the objective function and the constraints of the upper level problem. The linearization procedure of this algorithm uses the derivatives of flows with respect
to decision variables obtained by a sensitivity analysis (Tobin and Friesz, 1988) of the lower level problem.

On the other hand the present paper deals with a bimodal transport system and proposes a new method for solving the programming problem with equilibrium constraints. The latter is transformed into a nonlinear programming problem, whose objective function and constraints are expressed as explicit functions of the decision variables by using a regression technique.

The paper is organized as follows. Section 2 presents the model. Sections 3 and 4 deal with the capacity constraints and the budget constraints. Section 5 is dedicated to the equilibrium constraint, which is represented by a variational inequality. Section 6 analyses the optimization problem, and Section 7 shows a solution procedure. Section 8 presents an application of the model to a real urban transport system; it shows the influence that the constraints imposed on the system, some of which are of political nature, have on the values of the decision variables and of the objective function. Section 9 summarizes the results of the paper and presents some concluding remarks.

2. The model of the transport system management

Consider an urban area, divided in $k$ zones, and let $G(N,L)$ be the graph of the bimodal transport network serving the area, where $N$ is the set of nodes and $L$ the set of links. $L^c$ is the set of links travelled by private cars, $L^b$ is the set of links of the alternative transport mode. When transit is the alternative transport mode, links in $L^b$ are pedestrian links used for going from the origin centroids to stop nodes, line haul links travelled by transit vehicles, boarding links and alighting links. In some cases the alternative transport mode is a mixed one, where some links in $L^b$ are travelled by cars: travellers directed to the central area of a town use their private cars for going from the origin centroids to parking places located at the border of the central area, and then continue on foot or by transit as far as their final destination. Taking this fact into account, in the following we will name public transport any transport mode used in place of private cars: it can be a transit mode, or a mixed one, or just a pedestrian one.

Each zone is identified by a centroid, and $w = (i,j)_{i \neq j}$, represents an ordered pair of centroids; $W$ is the set of pairs $w$ and $n$ is their number. We consider the steady-state of the system during a unit time period in which the total transport demand $d_w$ between any $w \in W$ is fixed.

We suppose that a car driver travelling between $i$ and $j$ chooses a path connecting $i$ and $j$, whereas a user of public transport chooses a hyperpath (Nguyen and Pallottino, 1988).

Let:

- $d^c_w = \text{the transport demand by car between } w$
- $d^b_w = \text{the transport demand by public transport between } w$
- $d_w = d^b_w + d^c_w$
- $P^c_w = \text{the set of paths connecting } w \text{ by car}$
- $P^b_w = \text{the set of hyperpaths connecting } w \text{ by public transport}$
- $P^c = \text{the set of all paths in } G \text{ travelled by car}$
- $P^b = \text{the set of all hyperpaths in } G \text{ travelled by public transport}$
- $P = P^c \cup P^b$
\[ M^c = \text{the number of paths } p \in P^c \]
\[ M^b = \text{the number of hyperpaths } p \in P^b \]
\[ M = M^c + M^b \]
\[ h_p = \text{the flow on path or hyperpath } p \in P \]
\[ h = \text{the vector of flows on all paths and hyperpaths } p \in P. \]

We have:
\[
d_w^c = \sum_{p \in P_w^c} h_p \quad d_w^b = \sum_{p \in P_w^b} h_p \quad \forall w \in W
\]

so that the demand set of the transport system is:
\[
\Omega = \left\{ h : h \in \mathcal{R}^M, h \geq 0, \sum_{p \in P_w^c} h_p + \sum_{p \in P_w^b} h_p = d_w, \forall w \in W \right\}
\]

Both the transport modes connecting \( w \in W \) are characterized by a vector of attributes, \( X_w^c \) and \( X_w^b \) respectively. The components of these vectors are the various components of the monetary cost of the journey (car operating cost, road pricing, transit fares, etc.) and the various components of the journey time (in-vehicle time, waiting time, on-foot time). The systematic utilities \( U_w^c \) and \( U_w^b \) relative to the two transport modes are linear functions of \( X_w^c \) and \( X_w^b \):
\[
U_w^c = \theta^c X_w^c \quad U_w^b = \theta^b X_w^b
\]

where \( \theta \) is a vector of coefficients and \( ' \) indicates the transpose of a vector.

If \( -\alpha \) is the coefficient of the car operating cost, Eq. (3) can be written:
\[
U_w^c = -\alpha t_w^c \quad U_w^b = -\alpha t_w^b
\]

where
\[
t_w^c = -\frac{1}{\alpha} \theta^c X_w^c \quad t_w^b = -\frac{1}{\alpha} \theta^b X_w^b
\]

are the generalized trip costs between \( w \) on the two transport modes expressed in monetary units.

If the random residuals of the utility are independent Weibull variables, demand \( d_w \) divides between the two transport modes following a logit model. So we have:
\[
d_w^c = d_w \frac{\exp(-\alpha t_w^c)}{\exp(-\alpha t_w^c) + \exp(-\alpha t_w^b)} = \frac{d_w}{1 + \exp[\alpha(t_w^c - t_w^b)]}
\]

from which:
where:

\[ \lambda_w(d_w^c) = \frac{1}{\alpha} \ln \frac{d_w - d_w^c}{d_w^c} \]

\[ \Lambda_w(h) = \frac{1}{\alpha} \ln \frac{h_p}{\sum_{p \in P_w} h_p} \]

We suppose that the urban transport system as a whole is managed by an Authority which, given the layout of transit lines and the plan of car circulation, controls the system by means of a set of decision variables composed of transit fares, the transit service frequencies and road pricing imposed on some car links.

The system has to satisfy some constraints: constraints of physical and environmental capacity of the car system, constraints of physical capacity of the transit system, budget constraints of the urban transport system as a whole, and equilibrium constraint of flows travelling the bimodal network.

Let \( r \) be the vector of the road pricing values, \( \pi \) the vector of transit fares, \( v \) the vector of the service frequencies of the transit lines. The vector \( x \) of the decision variables is:

\[ x = (r', \pi', v')' \]
\[ R^v = \alpha_1 \sum_{i \in I} f_i r_i + \alpha_2 \sum_{j \in J} f_j \pi_j \]  

(10)

where \( \alpha_1 \) and \( \alpha_2 \) are two coefficients which transform the unit time income into annual income.

### 3.2. Annual operating cost

The annual operating cost \( C^v \) is the sum of the transit management cost, which is a function \( \varphi(v) \) of the service frequency vector \( v \), and of the cost of collecting charges from road users, which is supposed to be proportional to the sum of flows \( f_i \) on links \( i \in I \). So we have:

\[ C^v = \varphi(v) + \beta \sum_{i \in I} f_i \]  

(11)

where \( \beta \) is a coefficient.

### 3.3. Management budget

The annual deficit \( D \) of the transport system is the difference between \( C^v \) and \( R^v \):

\[ D = C^v - R^v \]  

(12)

If there is any public funding which contributes to financing the transport system, \( D \) may be positive. On the other hand, if \( D \) were negative, the difference \( R^v - C^v \) could be used for improving the capacity of road transport system: e.g. for increasing the parking capacity of the central area of the town.

Let \( A \) be the annual public funding, \( B \) the upper bound imposed for political reasons on the amount that the planning Authority may draw every year from the revenues of the transport system in order to finance the improvement of the road system. We have the following constraints:

\[ D \leq A \quad -D \leq B \]  

(13)

### 4. Capacity constraints

#### 4.1. Capacity constraints of the private car system

The network has \( \mu \) capacity constraints which concern the links on which road pricing is applied, and which are linear functions of flows on these links (Ruberti, 1995):

\[ \sum_{i \in I} \beta_{ri} f_i - H_r \leq 0 \quad r \in (1, 2 \to \mu). \]  

(14)
In general $H_r$ is the sum of two terms: a value $H_0^r$, which depends on the present capacity of the transport system, and a value $\Delta H_r$ which can be obtained by improving the road system. $\Delta H_r$ is supposed to be proportional to the available financing, i.e. to the positive difference between the public funding $A$ and the budget deficit $D$:

$$\Delta H_r = \eta \mu_r (A - D) = \eta \mu_r \left[A - \varphi(v) - \beta \sum_{i \in I} f_i + \alpha_1 \sum_{i \in I} f_i r_i + \alpha_2 \sum_{j \in J} f_j \pi_j \right]$$

(15)

where $\eta = 1$ if $A - D \geq 0$, 0 otherwise, and $\mu_r$ is a constant, which is a characteristic of each constraint.

4.2. Capacity constraints of public transport

Let:

- $T$ = the set of transit lines
- $B_t$ = the capacity of a transit vehicle travelling on line $t \in T$
- $J_t$ = the set of links travelled by line $t \in T$
- $v_t$ = the frequency of line $t \in T$
- $f_j$ = the flow on link $j \in J$.

We have the following constraints:

$$\frac{f_j}{v_j} - B_t \leq 0 \forall j \in J_t, \forall t \in T$$

(16)

5. Equilibrium constraint

We suppose that the cost on any link of the bimodal network is a function of flows travelling the links of both the transport modes. So, given a vector $h \in \Omega$, the cost $C_p$ of a path $p$ connecting by car a $w$ pair is a function of $h$ and of vector $r$ of road pricing; and the cost $C_p$ of a hyperpath connecting $w$ by public transport is a function of $h$ and of vectors $\pi$ and $v$ of tickets prices and transit frequencies (all costs are expressed in monetary units):

$$C_p = C_p(h, r) \quad \forall p \in P^w_c \quad \forall w \in W$$

$$C_p = C_p(h, \pi, v) \quad \forall p \in P^h_w \quad \forall w \in W$$

(17)

Given a vector $x = (r', \pi', v')'$ of decision variables, $x \in \mathcal{R}^m$, a vector $\tilde{h} \in \Omega$ is an equilibrium flow vector if the following conditions are satisfied:

$$\tilde{h} > 0 \Rightarrow C_p(\tilde{h}, r) - t^c_w = 0 \quad \forall p \in P^c_w \quad \forall w \in W$$

$$\tilde{h}_p = 0 \Rightarrow C_p(\tilde{h}, r) - t^c_w \geq 0$$

(18)
\[ \tilde{h}_p > 0 \Rightarrow C_p(\tilde{h}, \pi, v) - t^b_w = 0 \quad \forall p \in P^b_w \quad \forall w \in W \tag{19} \]
\[ \tilde{h}_p = 0 \Rightarrow C_p(\tilde{h}, \pi, v) - t^b_w \geq 0 \]

where \( t^c_w \) and \( t^b_w \) are the journey costs, which verify Eq. (7).

Let \( C(h, x) - \Lambda(h) : \mathcal{R}^M \to \mathcal{R}^M \) be the vector valued function whose components are \( C_p(h, r) - \Lambda_w(h) \forall p \in P^c_w \) and \( C_p(h, \pi, v) \forall p \in P^b_w, \forall w \in W \). We have the following theorem:

**Theorem.** A vector \( \tilde{h} \in \Omega \) is an equilibrium flow vector if and only if it satisfies the following variational inequality (VI):

\[ [C(\tilde{h}, x) - \Lambda(\tilde{h})](h - \tilde{h}) \geq 0 \quad \forall h \in \Omega \tag{20} \]

**Proof.** It easy to verify that \( \tilde{h} \) is a solution of VI(20) if and only if it is a solution of the following minimum problem:

\[ \min [\varphi(h) : h \in \Omega] \text{ where } \varphi(h) = [C(\tilde{h}, x) - \Lambda(\tilde{h})](h - \tilde{h}) \tag{21} \]

We write the nonnegativity constraints and the demand constraints which define the set \( \Omega \) as follows:

\[ s_p(h) = -h_p \leq 0 \quad \forall p \in P \tag{23} \]
\[ G_w(h) = \sum_{p \in P^c_w} h_p + \sum_{p \in P^b_w} h_p - d_w = 0 \forall w \in W \tag{24} \]

Since \( \varphi(h) \) and constraints (23) and (24) are linear, \( \tilde{h} \) is a solution of problem (21) if and only if it verifies the following KKT conditions:

\[ C(\tilde{h}, x) - \Lambda(\tilde{h}) + \sum_{p \in P} u_p \nabla s_p(\tilde{h}) + \sum_{w \in W} \rho_w \nabla G_w(\tilde{h}) = 0 \tag{25} \]

where multipliers \( u_p \forall p \in P \) are nonnegative and verify the complementary slackness (CS) condition \( \sum_{p \in P} u_p s_p(\tilde{h}) = 0 \), whereas multipliers \( \rho_w \forall w \in W \) may have any sign.

The \( p \) component of Eq. (25), \( p \in P^b_w \), is:

\[ C_p(\tilde{h}, \pi, v) - u_p + \rho_w = 0 \tag{26} \]

and it coincides with the equilibrium condition (19). As a matter of fact, if \( \tilde{h}_p > 0p \in P^b_w \), we have from the CS condition that \( u_p = 0 \) and thus \( C_p(\tilde{h}, \pi, v) + \rho_w = 0 \); on the other hand, if \( \tilde{h}_p = 0p \in P^b_w \), we have \( u_p \geq 0 \) and thus \( C_p(\tilde{h}, \pi, v) + \rho_w \geq 0 \), so that \( -\rho_w \) is the cost by public transport between \( w \) at equilibrium:
\[-\rho_w = t^b_w \quad \forall w \in W \]  

The \( p \) component of Eq. (25), \( p \in P^c_w \), is:

\[
C_p(h, r) - \Lambda_w(h) - u_p + \rho_w = 0
\]

and it coincides with the equilibrium condition (18). As a matter of fact, if \( h_p > 0p \in P^c_w \), we have \( u_p = 0 \) and thus \( C_p(h, r) - \Lambda_w(h) + \rho_w = 0 \); on the other hand, if \( h_p = 0p \in P^c_w \), we have \( u_p \geq 0 \) and thus \( C_p(h, r) - \Lambda_w(h) + \rho_w \geq 0 \), where, taking into account Eq. (7) and Eq. (27), \( \Lambda_w(h) - \rho_w \) is the trip cost \( t^c_w \) by car between \( w \) at equilibrium. □

6. Optimum values of decision variables

We write flows on links \( i \in L^c \) and \( j \in L^b \) as functions of vector \( h \):

\[
f_i = \sum_{p \in P^c} h_p \delta_{ip} \tag{29}
\]

\[
f_j = \sum_{p \in P^b} h_p \alpha_{jp} \delta_{jp} \tag{30}
\]

where \( \delta_{ip} \) and \( \delta_{jp} \) are 1 if links \( i \) and \( j \) belong to path or hyperpath \( p \) respectively, 0 otherwise; and \( \alpha_{jp} \), which a function of \( v \), is the probability that link \( j \) is used in the hyperpath \( p \) (Nguyen and Pallottino, 1988).

By substituting Eq. (29) and Eq. (30) into Eq. (10)Eq. (11)Eq. (14) and Eq. (16), we can write the budget and capacity constraints as functions of \( h \) and \( x \). If we put:

\[
g_b(h, x) = C^y - R^y = \varphi(v) + \sum_{i \in I} \sum_{p \in P^c} (\beta h_p \delta_{ip} - \alpha_1 h_p r_i \delta_{ip}) - \alpha_2 \sum_{j \in J} \sum_{p \in P^b} h_p \delta_{jp} \alpha_{jp} \pi_j
\]

the budget constraints can be written as follows:

\[
g_b(h, x) - A \leq 0
\]

\[
-g_b(h, x) - B \leq 0
\]

If we put:

\[
g_s(h, x) = \sum_{i \in I} \sum_{p \in P^c} \beta_{ij} h_p \delta_{ip} - H_r^0 - \eta \mu_c [A - g_b(h, x)] \quad r \in (1, 2 \ldots \mu) \tag{33}
\]

where, as usual, \( \eta = 1 \) if \( A - g_b(h, x) \geq 0 \), 0 otherwise,
the capacity constraints can be written as follows:

\[ g_r(h, x) \leq 0 \quad r \in (1, 2 \ldots \mu) \]  
\[ g_j(h, x) - B^t \leq 0 \quad \forall j \in J \forall t \in T \]  

We suppose that the Authority which manages the urban transport system chooses the values of decision variables in such a way that the average user satisfaction \( S \) is maximized, and at the same time the system is in equilibrium and the budget and capacity constraints are satisfied. On the hypothesis that utility is a Weibull random variable, \( S \) is given by (see e.g. Daganzo, 1979, pp. 12–13):

\[
S = \frac{1}{\sum_{w \in W} d_w} \left[ \sum_{w \in W} d_w \ln \left( \exp(-\alpha t_b^w) + \exp(-\alpha t_c^w) \right) \right]
\]  

where, as usual, \( t_b^w \) and \( t_c^w \) are transport costs at equilibrium between \( w \) by private car and by public transport, respectively.

If we denote \( p_b^w \) and \( p_c^w \) any path and hyperpath which connect \( w \) and are travelled by users at equilibrium, \( t_b^w \) and \( t_c^w \) can be expressed as a function of \( x \) and of a corresponding equilibrium vector \( h(x) \) in the following way:

\[
t_b^w = C_{p_b^w}(h(x), \pi, v) \quad t_c^w = C_{p_c^w}(h(x), r) \quad w \in W
\]  

where, as usual, \( C_{p_b^w} \) and \( C_{p_c^w} \) are the costs on hyperpath \( p_b^w \) and path \( p_c^w \) respectively.

By substituting Eq. (38) into Eq. (37), satisfaction \( S \) can be written as a function of \( x \) and \( h(x) \) as follows:

\[
S(h(x), x) = \frac{1}{\sum_{w \in W} d_w} \left[ \sum_{w \in W} d_w \ln \left( \exp\left[ -\alpha C_{p_b^w}(h(x), \pi, v) \right] + \exp\left[ -\alpha C_{p_c^w}(h(x), r) \right] \right) \right]
\]  

Let \( \Psi(x) \) be the set of solutions of VI (20) for each \( x \in \mathcal{R}_+^m \), i.e. the set of \( h(x) \) which verify:

\[
[C(h(x), x) - \Lambda(h(x), x)][h - h(x)] \geq 0 \quad \forall h \in \Omega
\]  

and let \( Z \subset \mathcal{R}^{M+m} \) be the set of vectors \((h'(x), x')' : x \in \mathcal{R}_+^m, h(x) \in \Psi(x)\), which satisfy the budget and capacity constraints (31), (32), (35), (36).

Consider the function \( \Phi(h(x), x) \) defined in the set \( Z \), where \( \Phi(h(x), x) = -S(h(x), x) \), and \( S(h(x), x) \) is given by Eq. (39).

The optimum vector \( \bar{x} \) of decision variables is solution of the following minimum problem:

\[
\min_{x} \Phi(h(x), x) \text{s.t.} (h(x)', x')' \in Z
\]
7. The solution of the minimum problem

It is easier to obtain a solution of problem (41) if one expresses both the objective function and the constraints in terms of link flows instead of flows on paths and hyperpaths, even if this procedure needs the introduction of other variables, i.e. the vector of car demand and the total waiting cost of public transport (Nguyen and Pallottino, 1988).

Let:
- \( f^c \) = the vector of flows on links \( i \in L^c \): \( f^c_i \) is given by Eq. (29)
- \( d^c \) = the vector of car demand \( d^c_w \forall w \in W \)
- \( c^c_i(f, r) \) = the cost of link \( i \in L^c \).

The problem can be simplified further if one supposes that the car link cost functions \( c^c_i(f, r) \) are separable, i.e. \( c^c_i(f, r) = c^c_i(f_i^c, r) \), and the costs of non-boarding links of public transport are independent of flows, so that the cost of a hyperpath is constant, and all users of public transport between each \( w \in W \) travel through the unique hyperpath of minimum cost \( t^h_w(\pi, v) \) connecting \( w \). In this case, taking into account that:

\[
\begin{align*}
\Lambda_w(h) &= \lambda_w(d^c_w) \\
C_p(h, x) &= \sum_{i \in L^c} c^c_i(f^c_i, r) \delta_{ip} \forall p \in P^c \\
f^c_i &= \sum_{p \in P^c} h_p d_{ip} \\
d^c_w &= \sum_{p \in P^c_w} h_p 
\end{align*}
\]

Section 6 (40) can be written:

\[
\sum_{i \in L^c} c^c_i(f^c_i(x), r)(f^c_i - f^c_i(x)) - \sum_{w \in W} [\lambda_w(d^c_w(x)) + t^h_w(\pi, v)](d^c_w - d^c_w(x)) \geq 0
\]

where \( f^c_i(x) = \sum_{p \in P^c} h_p(x) d_{ip} \) and \( d^c_w(x) = \sum_{p \in P^c_w} h_p(x) \).

It is easy to verify that the solution \([f^c(x), d^c(x)]\) of VI (42) is the unique solution of the following minimum problem:

\[
\min_{f^c, d^c} \tilde{F}(f^c, d^c, x) \text{s.t.} f^c_i = \sum_{p \in P^c} h_p d_{ip} \forall i \in L^c, d^c_w = \sum_{p \in P^c_w} h_p, d^c_w \leq d^c_w \forall w \in W
\]

where \( \tilde{F}(f^c, d^c, x) = \sum_{i \in L^c} \int c_i(u, r) du - \sum_{w \in W} \int [t^h_w(\pi, v) + \lambda_w(z)] dz \) is a strictly convex function.
Let $d^b_w(x) = d_w - d^c_w(x) \forall w \in W$, and let $f^b(x)$ be the vector of flows on links $j \in L^b$ obtained by assigning $d^b_w(x) \forall w \in W$ to the hyperpaths of minimum cost $t^b_w(\pi, v) \forall w \in W$. Let $\tilde{Z}$ be the set of vectors $\tilde{z} = [f^c(x), f^b(x), x]$ which satisfy the budget and capacity constraints (13), (14), (16), and let $Y$ be the set of vectors $y = [f^c(x), x]$ where $f^c(x)$, $x$ are components of any $\tilde{z} \in \tilde{Z}$.

If we remember that $t^c_w$ in Eq. (37) is the cost of any path $p^c_w$ connecting $w$ and travelled by users at equilibrium, we have that the average user satisfaction $S$ can be written as follows:

$$S[f^c(x), x] = \frac{1}{\sum_{w \in W} d_w \sum_{v \in W} d_v \ln \left( \exp[-\alpha t^b_w(\pi, v)] + \exp \left[ -\alpha \sum_{i \in L^c} c^c_i(f^c_i(x), r) \delta_{ipw} \right] \right)}$$

so that the optimum vector $\tilde{x}$ of decision variables is the solution of the following minimum problem:

$$\min_{x} \tilde{\Phi}[f^c(x), x] \text{s.t.} [f^c(x), x] \in Y$$

where $\tilde{\Phi}[f^c(x), x] = -S[f^c(x), x]$ is a nonconvex function, so that problem (45) is a bilevel non-linear nonconvex problem.

In case the components of $x$ are few, one can compute the expressions of the objective function and of constraints as explicit functions of $x$ by solving the lower level problem (43) for a wide range of $x$ values, and then computing the regressions of the values of the objective function and of constraints on the corresponding $x$ values. In this way, by substituting the regression equations for the actual functions, the bilevel programming problem (45) can be reduced to a nonlinear programming problem with very good approximation, as we will see in the next section.

Problem (43) is a problem of assignment of elastic demand to the network of private cars. The solution $f^c(\hat{x}), d^c(\hat{x})$, corresponding to a particular value $\hat{x}$ of $x$, can be obtained by using one of the methods which have been proposed for the assignment of elastic demand: e.g. the algorithm of excess-demand (Sheffy, 1985, p. 148), which transforms the problem of assignment of elastic demand into a problem of assignment of fixed demand to an augmented network; the details of the procedure can be found in Ferrari (1996). Journey costs $t^b_w(\hat{x}) \forall w \in W$ are obtained by computing, for $x = \hat{x}$, the hyperpaths of minimum cost connecting each $w \in W$, while flows $f^b(\hat{x})$ are calculated by assigning demand $d^b(\hat{x}) = d - d^c(\hat{x})$ to the hyperpaths of minimum cost. So it is possible to compute the values of the objective function and of the budget and capacity constraint functions for $x = \hat{x}$.

In the problem which this paper deals with, the components of $x$ are actually few. As a matter of fact a town, even if large, is divided in two or three areas, on each of which a different value or road pricing is imposed; for instance, the congestion charging scheme proposed for London (Richards et al., 1996) divides the city area in three parts: Central London and two annular areas around it. The transit lines are usually assembled in few groups, each of which has a different service frequency, and ticket price is the same for all the lines.

Thus the method of solving the bilevel problem (45) by computing the expressions of the objective function and of constraints as explicit functions of $x$ has been used in the application of the model to a real network which is illustrated in the next section.
8. An application to a real case

The model of transport system management presented in this paper has been applied to the transport network of Pistoia, a town of Tuscany having around 90,000 inhabitants. During the peak hour in the morning about 11,000 passengers travel towards the central area of the town, using private cars and the bus lines of the transit system. The average number of passengers per car is 1.30. Twenty-five percent of the total transport demand is captive of the transit system, while 75% can choose between car and transit.

The parking supply in the central area satisfies only 35% of the demand, so that many cars park illegally. On the other hand the excess of car transport demand cannot be transferred to transit, because the dispersion of origin points in a large area around the town makes the use of transit very difficult for people living in this area.

Therefore a new public transport system has been planned, which is a mixed one that uses four interchange parking areas located at the border of the central area of the town. Drivers directed towards the central area have two alternatives: they can use their private cars as far as their final destinations; or they can go by car from the origin to one of the four interchange parking areas, and then reach on foot or by transit the final destination. Six bus lines serve the central area and connect the four parking areas between them. Taking into account the characteristics of the zones served by bus lines, the latter have been divided in two groups, having service frequency $v$ and $v/3$, respectively.

The network of public transport has 875 links and 353 nodes; 50 of them are O/D centroids.

A parking tax has been imposed on drivers who park their cars in the central area, which has been divided in two parts: the tax is $r$ in the inner part and $r/1.5$ in the other. It has been supposed that drivers pay the tax when they enter the central area and again when they leave, so that drivers are charged like transit users, who pay the ticket two times in their round trips.

It was supposed that the cost functions on car links are separable, and that the cost on links of public transport are independent of flows. On these hypotheses, given a vector $x = (r, \pi, v)^T$ of decision variables, the solution of VI (42) is the unique solution of minimum problem (43), and can be obtained by assigning the elastic car demand to an augmented network, as it was shown in the previous section. The augmented network has 190 nodes and 2316 links: 1862 of them are virtual links travelled by the demand $d^b_w$ on public transport.

The coefficients of the attributes in the expression (3) of utility function have the following values (times are measured in min, monetary costs in thousands of lire), estimated in an empirical survey carried out in an italian city (Comune di Napoli, 1992):

- journey time by car or by bus $\theta_1 = 0.035$/min
- journey time on foot $\theta_2 = 0.05$/min
- waiting time $\theta_3 = 0.06$/min
- park searching time $\theta_4 = 0.05$/min
- trip monetary cost $\theta_5 = 0.4$/1000 lire
- parking tax $\theta_6 = 0.5$/1000 lire

By solving problem (43) for a given vector $\hat{x}$ of decision variables, the corresponding values of the following variables at equilibrium have been computed:
(a) the annual budget $D$, i.e. the difference between annual operating cost and income;
(b) the ratio $T$ between the demand entering the central area by car and the total demand directed to the central area which is not captive of transit;
(c) the average satisfaction $S$;
(d) the maximum passenger flow on a bus for each line.

The calculation has been repeated for 20 sets of decision variables, and then the regressions of the previously quoted variables on the decision variables have been estimated. The regression relationships which have been obtained reproduce the actual functions with very high approximation. This is shown, for instance, by the regressions of $D$, $T$ and $S$ which are reported below along with the $R^2$ statistics and the $t$-Student values of coefficients ($D$ is measured in $10^9$ lire, $S$, $r$, $\pi$ in thousands of lire, $v$ in min$^{-1}$):

$$D = 7.585 + 121.64v - 7.344r + 0.776r^2 - 7.291\pi$$
$$R^2 = 0.9963 \quad t_v = 54.48 \quad t_r = -18.13 \quad t_{r^2} = 13.78 \quad t_{\pi} = -11.44$$

(46)

$$T = 0.693 - 0.211v - 0.103r + 0.003r^2 + 0.044\pi$$
$$R^2 = 0.9975 \quad t_v = -7.97 \quad t_r = -21.44 \quad t_{r^2} = 4.64 \quad t_{\pi} = 5.75$$

(47)

$$S = -0.399 - 0.02/v - 0.213r + 0.14r^2 - 0.163\pi$$
$$R^2 = 0.9956 \quad t_v = -16.42 \quad t_r = -20.74 \quad t_{r^2} = 9.58 \quad t_{\pi} = -10.08$$

(48)

As we said before, the maximum $T$ value consistent with the present parking supply in the central area of the town is 0.35. If $A$ is the value of public funding, the maximum amount that can be spent in increasing the parking supply is $A - D$, if $A - D > 0$. We have computed that an increase of 0.020 can be added to the ratio $T$ for each billion of lire invested each year for 12 years in increasing the parking supply. Let $\bar{T}$ be a bound imposed on $T$ for physical and environmental reasons. The ratio $T$ has to satisfy the following constraints:

$$T(r, \pi, v) - \bar{T} \leq 0$$

(49)

$$T(r, \pi, v) - 0.35 - 0.020\eta[A - D(r, \pi, v)] \leq 0$$

(50)

where $\eta = 1$ if $A - D \geq 0$, 0 otherwise, and $T(r, \pi, v)$ is given by Eq. (47).

Let $B$ be the upper bound imposed for political reasons on the amount that the planning Authority may draw every year from the revenues of the transport system in order to finance the increase of parking supply. We have the following constraints:

$$D(r, \pi, v) \leq A$$

(51)

$$-D(r, \pi, v) \leq B$$

(52)

where $D(r, \pi, v)$ is given by Eq. (46).
The maximum number of passengers per bus is 90. If \( f_k(r, \pi, v) \) is the maximum flow on a bus of line \( k \), and \( K \) is the set of bus lines, we have the following constraints:

\[
f_k(r, \pi, v) - 90 \leq 0 \quad \forall k \in K
\]  

(53)

where \( f_k(r, \pi, v) \) are given by the regression equations.

If social reasons require the imposition of an upper bound \( \pi_{\text{max}} \) on the price of the bus ticket, we have another constraint:

\[
\pi - \pi_{\text{max}} \leq 0
\]  

(54)

The objective function is \( \Phi(r, \pi, v) = -S(r, \pi, v) \) where \( S(r, \pi, v) \) is given by Eq. (48).

The optimum values of decision variables are the nonnegative values which minimize \( \Phi(r, \pi, v) \) in the set defined by constraints (49) \ldots (54). This constrained non linear programming problem has been transformed into a sequence of nonconstrained ones by using an augmented lagrangian penalty function (Fletcher, 1996, pp. 287–295) and each nonconstrained problem has been solved by the Powell method (Fletcher, 1996, pp. 89–92).

The optimum values of decision variables have been computed, on the hypothesis that no upper bound \( \bar{T} \) exists on \( T \) value, in the seven cases reported below, which are characterized by different values of \( A \) and \( B \). When \( \pi_{\text{max}} \) is not indicated, the constraint on bus ticket price has not been imposed. \( A \) and \( B \) are measured in billions of lire, \( \pi_{\text{max}} \) in thousands of lire.

Case 1: \( A = 0, B = 0 \)
Case 2: \( A = 0, B = -5 \)
Case 3: \( A = 0, B = -15 \)
Case 4: \( A = 5, B = 0 \)
Case 5: \( A = 5, B = 0, \pi_{\text{max}} = 2.0 \)
Case 6: \( A = 5, B = 0, \pi_{\text{max}} = 1.8 \)
Case 7: \( A = 5, B = -15 \)

The results obtained are reported in Table 1. It shows, for each case, the optimum values of decision variables, the proportion \( T \) of the demand entering the central area by car, the budget value \( D \) and the value of objective function \( \Phi = -S \). The latter has been expressed in monetary units dividing its value by the coefficient that affects the car operating cost in the expression (3) of utility function, so that it is a measure of the average transport generalized cost.

First of all we observe that the optimum values of the objective function in the various cases are not very different between them. This fact essentially depends on the small dimensions of the central area of the town: its centroids are connected between them and with the interchange parking areas by pedestrian paths whose costs are low and independent of the decision variables. In case the values of the decision variables should make the transport cost by car and by bus too high, users would travel on foot from the interchange parking areas to the centroids of the central area, and between these centroids; so that the transport cost would remain rather low in any case. In urban areas having the same characteristics of that studied in this paper, but with larger dimensions, the pedestrian paths would be longer, and thus more expensive; therefore we should
expect that the differences between the objective function values would have the same signs of those obtained in Pistoia, but would be more substantial.

It can be noted that, as it was to be expected, the optimum value of the objective function $\Phi$ decreases, other things being the same, if the transport system is partially financed by public funding. The optimum $\Phi$ is lower the greater is the portion of the system revenues dedicated to the improvement of parking supply, which is not bounded by physical or environmental constraints imposed on the proportion of users entering the central area by car. When the portion of the system revenues that the planning Authority may dedicate to the increase of parking supply is high, as in cases 3 and 7, the proportion of users that do not enter the central area by car is low. Since a large portion of people who do not travel by car are pedestrians, we can argue that, if there were no constraint on the increase of parking supply, the number of bus users would be very close to the number of people who have no availability of a private car.

Finally one can observe that the objective function $\Phi$ increases if an upper bound is imposed on bus ticket price, and it is greater the lower is this bound.

### 9. Conclusions

This paper has shown a scheme of urban transport management aimed at solving the externalities due to road traffic. In this scheme both private and public transport modes are managed by a planning Authority having control of all user charges, i.e. road pricing and transit ticket price, and of the service characteristics of transit. The system is subject to some constraints: physical and environmental capacity constraints, and a budget constraint. The planning Authority fixes the level of the user charges and of the transit characteristics in such a way that the average transport generalized cost is minimized, and at the same time the system is in equilibrium and all constraints are satisfied. In some cases social reasons require that an upper bound is imposed on the ticket price, in order to help people who are captive of transit.

A mathematical model of this management scheme has been presented, and it has been applied to a real urban network, in order to show the effects that the management policy, and in particular the constraints imposed on the system, have on the optimum urban transport cost. The results obtained, even if they refer to a particular case, point out some characteristics of urban transport systems which probably have a more general significance. The optimum value of the average urban transport generalized cost decreases if a portion of the system revenues is dedicated to

<table>
<thead>
<tr>
<th>Case</th>
<th>$r$ (lire)</th>
<th>$\pi$ (lire)</th>
<th>$v$ (min$^{-1}$)</th>
<th>$T$</th>
<th>$D$ ($10^9$ lire)</th>
<th>$\Phi$ (lire)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4365</td>
<td>2113</td>
<td>0.206</td>
<td>0.350</td>
<td>0.00</td>
<td>3760</td>
</tr>
<tr>
<td>2</td>
<td>3391</td>
<td>2475</td>
<td>0.176</td>
<td>0.450</td>
<td>$-5.00$</td>
<td>3692</td>
</tr>
<tr>
<td>3</td>
<td>2444</td>
<td>2975</td>
<td>0.139</td>
<td>0.561</td>
<td>$-10.54$</td>
<td>3662</td>
</tr>
<tr>
<td>4</td>
<td>3153</td>
<td>2092</td>
<td>0.190</td>
<td>0.450</td>
<td>0.00</td>
<td>3445</td>
</tr>
<tr>
<td>5</td>
<td>3407</td>
<td>2000</td>
<td>0.200</td>
<td>0.422</td>
<td>1.41</td>
<td>3462</td>
</tr>
<tr>
<td>6</td>
<td>3930</td>
<td>1800</td>
<td>0.216</td>
<td>0.367</td>
<td>4.08</td>
<td>3502</td>
</tr>
<tr>
<td>7</td>
<td>1573</td>
<td>3043</td>
<td>0.118</td>
<td>0.647</td>
<td>$-9.87$</td>
<td>3410</td>
</tr>
</tbody>
</table>
the improvement of the car system capacity, e.g. by increasing the parking supply so that a
greater number of private cars can enter the central area of the town; if there were neither poli-
tical nor physical reasons which impose a threshold on the parking supply in the central area of
the town, the optimum of transport cost would be obtained with a number of bus users very close
to the number of people who have no availability of a private car. If a portion of revenues is used
to maintain low the level of the transit ticket price, in order to help people who are captive of
public transport, the average urban transport cost increases. One could argue from these results
that, if all people had availability of a car, and if there were no reason which imposes an upper
bound on the increase of the road system capacity, since the objective of the planning Authority
is to minimize the generalized transport cost, there would be little room for transit system.

It is people who are captive of transit, along with the physical and environmental constraints
that make it difficult to increase the road system capacity, who justify transit system from both
social and economical point of view. Two categories of people, those who are captive of transit,
and those who are directed to the central area of the town and do not want to pay the road pri-
cing imposed to satisfy the capacity constraints, constitute the market niche of transit system.

This conclusion is relative to the particular case examined in this paper. But also other authors
are of the opinion that the transit market is a niche market: see e.g. Hensher, 1996, who asserts
that ‘urban public transport is becoming a supplier of niche services, despite the continued
emphasis on providing across-the-board services’. Probably in the future the number of people
who are captive of transit will decline, and those remaining will be served by other forms of
public transport, e.g. a dial-a-ride system. Thus the characteristics of public transport, along with
those of its market, will be essentially determined by capacity constraints imposed on road
transport system.

Acknowledgements

The author is indebted to two anonymous referees and to the Associate Editor M.G.H. Bell for
their useful comments and suggestions on an earlier edition of this paper, which was financed in
part by Consiglio Nazionale delle Ricerche, Progetto Finalizzato Trasporti 2.

References

Press, New Haven, CT.
Constantin, I., Florian, M., 1993. Optimizing frequencies in a transit network: a nonlinear bilevel programming
approach. Centre de Recherche sur les Transports et Département d’Informatique et de Recherche Opérationnelle,
Université de Montréal, Montréal, Canada.
72.
Transportation and Traffic Theory, Volmuller and Hammerslag eds. VNY Science Press, Utrecht, The Netherlands, 
pp. 197–216.
Fisk, C.S., 1986. A conceptual framework for optimal transportation systems planning with integrated supply and 
demand models. Transportation Science 20, 37–47.
librium and Advanced Transportation Modelling, in press.
with user-optimal flows. Transportation Research 20B, 259–265.
approach with transportation network equilibrium constraints. Proceedings of the 7th World Conference on Trans-
of Operational Research 37, 176–186.
Richards, M., Gilliam, C., Larkinson, J., 1996. The London congestion charging research program: the program in 
overview. Traffic Engineering and Control 37, 66–71.
Yang, H., Yagar, S., 1994. Traffic assignment and traffic control in general freeway-arterial corridor systems. Trans-
portation Research 28B, 463–486.