A MICROECONOMIC INTERPRETATION FOR THE SYSTEM OPTIMAL TRAFFIC ASSIGNMENT PROBLEM WITH NONADDITIVE PATH COST

Louis de Grange
Department of Industrial Engineering, Diego Portales University, Santiago, Chile.
Telephone: (56-2) 676-2400; e-mail: **louis.degrange@udp.cl**

Juan Carlos Muñoz
Department of Transport Engineering and Logistics, Pontificia Universidad Católica de Chile, Telephone: (56-2) 354 4270; e-mail: **jcm@ing.puc.cl**

Rodrigo Troncoso
Facultad de Gobierno, Universidad del Desarrollo, Santiago de Chile
Telephone: (56-2) 377-4817; e-mail: **rtroncoso@lyd.org**

ABSTRACT

Using a Bergson-Samuelson welfare function we outline a microeconomic interpretation of the effects of the nonlinearity in the time/cost relationship for travelers in a congested transport network. It is demonstrated that a marginal cost traffic flow assignment following Wardrop's second principle, although it minimizes the total cost of a transport network, may reduce social welfare compared to the market equilibrium assignment based on Wardrop's first principle. A welfare-maximizing assignment model is presented and used to show that if the travellers' utility functions are linear, the assignment that maximizes social welfare will be the same as the assignment that minimizes total network cost, but if users' utility functions are non-linear (reflecting the traditional non-satiation and diminishing marginal utility axioms), the two assignments will be different. It is further shown that the effects of this non-linearity are such that a welfare-maximizing assignment will meet with less user resistance than a minimum total network cost assignment.

Keywords: traffic assignment, Wardrop, minimum cost, social welfare, utility function.

1. INTRODUCTION

For decades, Wardrop's second principle (Wardrop, 1952) has been considered the paradigm that should govern vehicle flow patterns in a congested road network whenever the goal is an efficient solution in terms of network users' total aggregate cost or trip time. Since this approach assigns vehicles in a such a way as to minimize total network cost, it is to be preferred to the market or user equilibrium approach, expressed as Wardrop's first principle (Wardrop, 1952), in which each user chooses a route independently of the others in order to minimize his or her individual cost. In traffic assignments following the second principle, regarded as socially optimal, drivers cooperate with one other and in certain cases sacrifice individual benefit in order to reduce total network cost or trip time. Note that under both approaches, assignment is based on the costs network users face rather than the welfare they obtain.

In Bell and Iida (1997) it was demonstrated that if each traveller unilaterally chooses the route that maximizes his or her individual utility function independently of all other users, the result will be a market equilibrium based on Wardrop's first principle. In other words, maximizing individual utility produces the same market equilibrium as minimizing individual cost. However, this equivalence is not necessarily satisfied in a centrally planned assignment that minimizes total network cost as opposed to one that maximizes social welfare.

In this paper, we use a microeconomic framework to describe situations where a marginal cost assignment may reduce social welfare compared to a market equilibrium assignment. It will be demonstrated that if travellers have standard utility functions (increasing and concave in consumption), an assignment based on Wardrop's second principle may generate a reduction in social welfare even though total network costs are lower. The basis for the proof is that moving from a market equilibrium to a minimum total cost assignment implies a reassignment of resources that will make some travellers better off but others worse off. If we define social welfare as the sum of the individual utilities (Bergson, 1938; Samuelson, 1956; Boadway and Bruce, 1984), the non-linearity of the latter due to the non-satiation and diminishing marginal utility axioms may result in a welfare loss to those made worse off that is greater than the welfare gain to those made better off. In other words, although total travel time is shorter under Wardrop's second principle, total welfare for the travellers as a whole might be reduced.

Non-additivity between costs and routes in arcs has been studied previously. Gabriel and Bernstein (1997) formulated a traffic equilibrium model relaxing the assumption that the cost of the routes is the linear sum of the costs of the arcs that compose it, i.e. arches and costs are non additive within a route, generating different traffic equilibrium. This approach is extended to the optimal system problem, but it omits a microeconomic interpretation based on welfare economics. A similar approach was developed by Larsson et al. (2002). However, unlike Gabriel and Bernstein (1997), where the route cost function is based on "money", in the formulation of Larsson et al. (2002), the route cost is expressed in terms of time. Lo and Chen (2000), Han and Lo (2004) and Agdeppa et al. (2007) have developed algorithms to solve this problem.

We use a similar framework as in Gabriel and Bernstein (1997), but we develop and interpret the model from a microeconomic perspective, using tools from the consumer and welfare theories.

As defined by Easa (1991), traffic assignment is "the process of allocating a set of present or future trip interchanges, known as origin-destination (OD) demands, to a specified transportation network." The results of traffic assignments are an essential consideration in many planning and transport system design processes such as infrastructure project scenario evaluation, environmental impact analysis and highway design and pricing. The criteria for making these assignments will influence the decisions of the transport authorities and planners and in turn impact the system users' welfare, which should be the ultimate objective of the central planning exercise.

Due to its simplicity, traffic assignment based on Wardrop's second principle has been the principal aim of many methodologies for the analysis of issues such as road pricing (Arnott and Small, 1994; Button and Verhoef, 1998; Yang and Meng, 2002; Verhoef, 2002; Shepherd and Sumalee, 2004; Santos, 2004; Yanga *et al.*, 2010), transport network design and bi-level optimization problems (Marcotte, 1983; Newbery, 1989; Yang and Bell, 1998; Brotcorne *et al.*, 2001; Mun *et al.*, 2003; Koh *et al.*, 2009), transport market regulation (Smith, 1979; Small, 1992; Small and Gómez-Ibáñez, 1998; Parry, 2002).

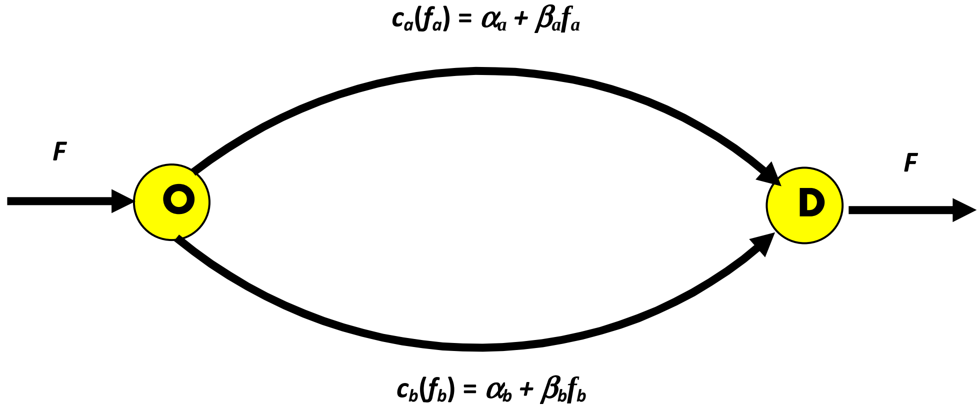
The remainder of this paper is organized as follows. Section 2 introduces a simple model of a small road network that provides a graphical explanation of the social welfare loss incurred by changing from a market equilibrium to a marginal cost or minimum total cost assignment. Section 3 formulates an equivalent optimization problem for traffic assignment based on maximizing social welfare à la Bergson-Samuelson instead of minimizing total cost, analyses the problem's optimality conditions and compares them to those for Wardrop's second principle. Section 4 we present some properties for the linear cost functions. Section 5 summarizes the study's findings and presents the conclusions.

2. A SIMPLE ROAD NETWORK MODEL

2.1 Description of network and its equilibrium

Consider the simple example in Figure 1, in which homogeneous (i.e., identical) travellers with linear cost functions are assigned to a road network with a single origin-destination (OD) pair and just two alternative routes or arcs.

Figure 1
Simple road network



As can be seen in the figure, the total flow of travellers between origin O and destination D is F, where $f_a + f_b = F$. The average cost of each arc is linear in own flow. The parameters α_a and α_b are the free-flow costs of their respective arcs while β_a and β_b are the marginal effects of the arcs' respective flows on individual arc cost.

Network equilibrium according to Wardrop's first principle (or the traffic equilibrium) is obtained by equating the average costs of the two arcs and adding the flow conservation constraint as follows (Beckmann *et al.*, 1956):

$$c_a(f_a) = c_b(f_b) \rightarrow \alpha_a + \beta_a f_a = \alpha_b + \beta_b f_b \tag{1}$$

$$f_a + f_b = F \tag{2}$$

Assuming that in equilibrium, both routes have traffic flow (*i.e.*, excluding border solutions), the solution to problem (1)-(2) is

$$f_a^* = \frac{(\alpha_b - \alpha_a) + \beta_b F}{\beta_a + \beta_b} \quad , \qquad f_b^* = F - f_a^* \tag{3}$$

By contrast, a traffic assignment following Wardrop's second principle, based on minimization of cost over the entire network, is derived by equating marginal costs of the arcs subject to the flow conservation constraint:

$$\frac{\partial\left(c_a(f_a)\cdot f_a\right)}{\partial f_a} = \frac{\partial\left(c_b(f_b)\cdot f_b\right)}{\partial f_b} \rightarrow \alpha_a + 2\beta_a f_a = \alpha_b + 2\beta_b f_b \tag{4}$$

$$f_a + f_b = F \tag{5}$$

Assuming positive flows, the solution of problem (4)-(5) is

$$f_a^{**} = \frac{(\alpha_b - \alpha_a) + 2\beta_b F}{2(\beta_a + \beta_b)} \quad , \qquad f_b^{**} = F - f_a^{**} \tag{6}$$

Solution (6) generates total network costs below those of solution (3), and is therefore the optimal network assignment.

2.2 Travellers' utility functions and the social welfare function

Assume that each traveller i on the network has a utility function $U(g_i)$, where g_i is the amount of a good consumed. The functions are increasing (non-satiation axiom) at a decreasing rate in consumption of the good (diminishing marginal utility axiom). Thus, $\frac{\partial U}{\partial g_i} > 0$ and $\frac{\partial^2 U}{\partial g_i^2} < 0$.

For the study of flow assignment we also assume that time is the only good. We therefore define $g_i = t_i - c_i$, where t_i is the total time available to individual i (*e.g.*, 24 hours per day) and c_i is the individual's travel time. This implies that g_i is the individual's free time.

The social welfare for all travellers collectively as a function of the free time $g_a = t - c_a$ and $g_b = t - c_b$ of the individuals using arcs a and b respectively in Figure 1 can be represented graphically in terms of iso-welfare curves (Samuelson, 1956). Since the network model assumes homogeneous individuals, the utility of travellers on arc a will be $U_a = U\left(g_a\right)$ and that of travellers on arc b will be $U_b = U\left(g_b\right)$.

As explained in the Introduction, if each traveller unilaterally chooses a network route so as to maximize individual utility, then together they will generate the same market equilibrium as would the minimization of individual cost, but this equivalence will not necessarily be maintained by a centrally planned traffic assignment. In other words, minimization of total network costs may lead to different results than those obtained by maximizing social welfare based on individual utilities.

To analyze this point in greater depth, consider the following social welfare function of the type described by Bergson (1938), Samuelson (1956) and Boadway and Bruce (1984):

$$W\left(U_1, U_2, ..., U_n\right) = \sum_{i=1}^{n} \lambda_i U_i \tag{7}$$

where W is social welfare, U_i is the utility of individual i and λ_i is the relative importance of each individual.

If we assume that all individuals are equally important $\left(\lambda_i = 1, \forall i\right)$, then, still using Figure 1 as our example, we define the following social welfare function:

$$W = \sum_{m} f_m U_m = f_a \cdot U\left(g_a\right) + f_b \cdot U\left(g_b\right) = f_a \cdot U_a + f_b \cdot U_b \tag{8}$$

This function can be plotted as iso-welfare curves in a coordinate space whose axes are the variables g_a and g_b. Since traffic flows f_a and f_b depend on costs c_a and c_b, respectively, and $g_i = t - c_i$, flows f_a and f_b can be expressed as functions of g_a and g_b, that is, $f_a = \dfrac{t - g_a - \alpha_a}{\beta_a}$ and $f_b = \dfrac{t - g_b - \alpha_b}{\beta_b}$.

We can therefore rewrite (8) as

$$W = \left(\frac{t - g_a - \alpha_a}{\beta_a}\right) U_a + \left(\frac{t - g_b - \alpha_b}{\beta_b}\right) U_b \tag{9}$$

The social iso-welfare curve for a given value of W satisfies the condition

$$dW = \left(\frac{\partial f_a}{\partial g_a} U_a + \frac{\partial U_a}{\partial g_a} f_a \right) dg_a + \left(\frac{\partial f_b}{\partial g_b} U_b + \frac{\partial U_b}{\partial g_b} f_b \right) dg_b = 0 \qquad (10)$$

2.3 Comparative statics of equilibria

In traffic equilibrium (Wardrop's first principle), $c_a = c_b$ and therefore

$$\frac{g_a}{g_b} = 1 \qquad (11)$$

The set of feasible solutions for the two route flows is determined by $f_a + f_b = F$. Since $f_a = \frac{t - g_a - \alpha_a}{\beta_a}$ and $f_b = \frac{t - g_b - \alpha_b}{\beta_b}$, the set of feasible solutions can be expressed in terms of g_a and g_b as

$$g_a = \underbrace{t\left(\frac{\beta_a + \beta_b}{\beta_b} \right) - F\beta_a - \alpha_a - \alpha_b \left(\frac{\beta_a}{\beta_b} \right)}_{K} - g_b \left(\frac{\beta_a}{\beta_b} \right) \qquad (12)$$

$$g_a = K - g_b \left(\frac{\beta_a}{\beta_b} \right), \text{ and therefore } \frac{dg_a}{dg_b} = -\frac{\beta_a}{\beta_b} \qquad (13)$$

The traffic equilibrium is thus given by the intersection of the straight lines defined by (11) and **¡Error! No se encuentra el origen de la referencia.**. In Figure 2 this equilibrium is shown as point A on iso-welfare curve W^A.

The optimal assignment that minimizes total network costs (Wardrop's second principle) is obtained when $\frac{\partial\left(c_a\left(f_a\right)\cdot f_a\right)}{\partial f_a} = \frac{\partial\left(c_b\left(f_b\right)\cdot f_b\right)}{\partial f_b}$, that is, when the marginal costs of the two routes as defined in (4) are equal. We then get the following relationship between the individual users' free time:

$$g_a = \left(\frac{\alpha_b - \alpha_a}{2} \right) + g_b \qquad (14)$$

It follows, therefore, that the intersection of lines (14) and **¡Error! No se encuentra el origen de la referencia.** gives the socially optimal assignment for the network (Wardrop's second principle).

Note that line (14) is just a parallel shift of line (11). If $\alpha_a > \alpha_b$, the shift is rightwards from (11); if, on the other hand, $\alpha_a < \alpha_b$, the shift would be leftwards. Furthermore, if $\alpha_a = \alpha_b$, the market equilibrium would coincide with the socially optimal assignment.

Assuming that $\alpha_a > \alpha_b$, consider a small shift of straight line (14). In other words, α_a is just slightly greater than α_b. This generates the assignment shown in Figure 2 as point *B* and travellers' social welfare has increased ($W^B > W^A$).

If, however, $\alpha_a < \alpha_b$, the shift moves the minimum-cost assignment to point *C* and social welfare decreases ($W^C < W^A$). And if we again assume $\alpha_a > \alpha_b$ but that α_a is much greater than α_b, the minimum-cost assignment at point *D* as shown in Figure 3 is once more at a lower social welfare level ($W^D < W^A$).

Figure 2
Social welfare in a marginal cost assignment (Wardrop's second principle)

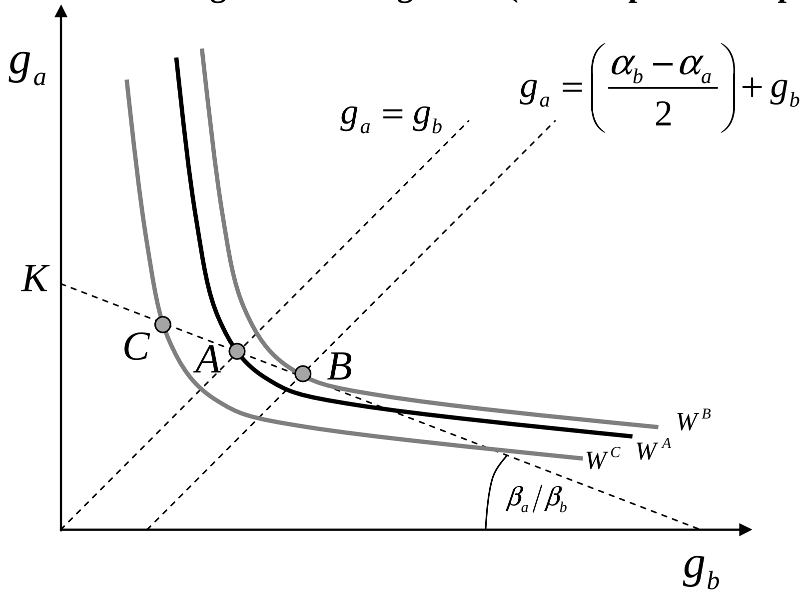
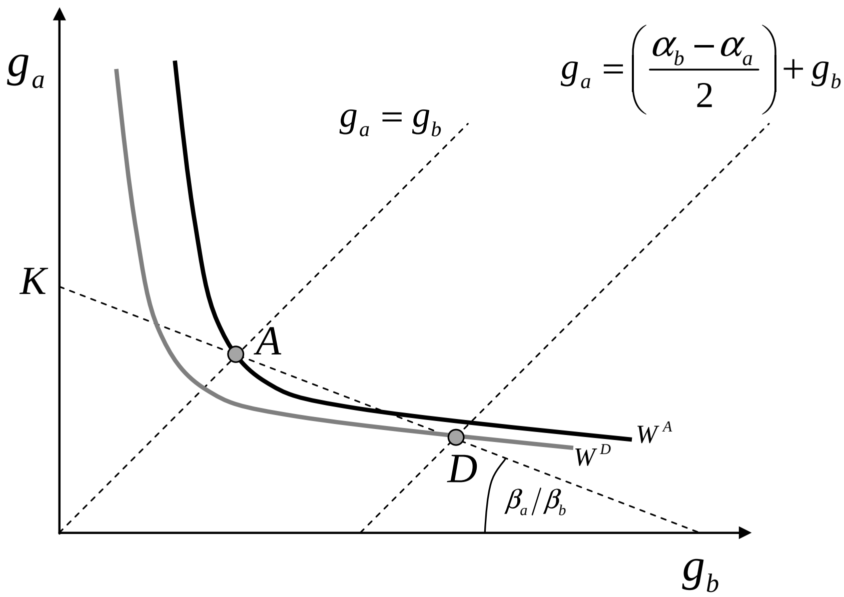


Figure 3
Social welfare in a marginal cost assignment (Wardrop's second principle)



2.4 Assignment maximizing social welfare

From (10) we have

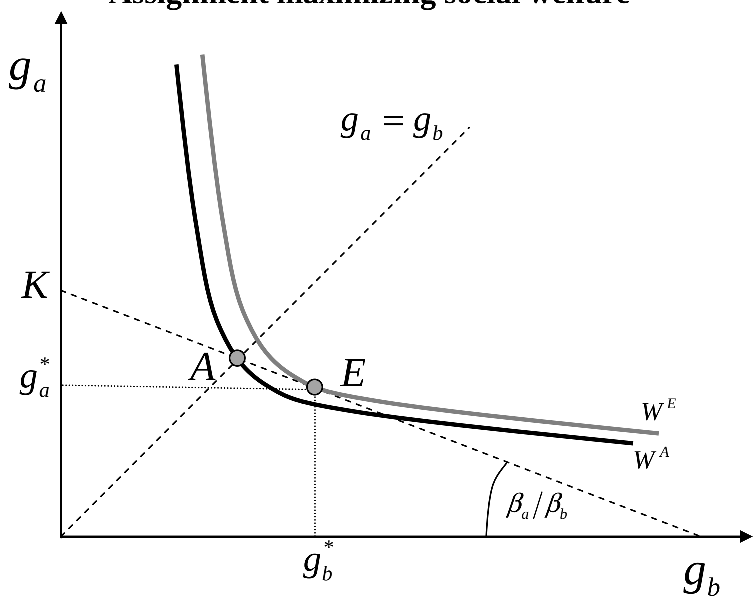
$$\frac{dg_a}{dg_b} = -\frac{\left(\dfrac{\partial f_b}{\partial g_b} U_b + \dfrac{\partial U_b}{\partial g_b} f_b\right)}{\left(\dfrac{\partial f_a}{\partial g_a} U_a + \dfrac{\partial U_a}{\partial g_a} f_a\right)} \tag{15}$$

Since the condition expressed by **¡Error! No se encuentra el origen de la referencia.** must be satisfied, the optimality condition for obtaining the flows that maximize social welfare is

$$\frac{\beta_a}{\beta_b} = \frac{\left(\dfrac{\partial f_b}{\partial g_b} U_b + \dfrac{\partial U_b}{\partial g_b} f_b\right)}{\left(\dfrac{\partial f_a}{\partial g_a} U_a + \dfrac{\partial U_a}{\partial g_a} f_a\right)} \tag{16}$$

From (16) and **¡Error! No se encuentra el origen de la referencia.** we obtain the solutions g_a^* and g_b^* that maximize social welfare and from which we can directly derive the corresponding welfare-maximizing traffic flows f_a^* and f_b^*. This solution is shown as point *E* in Figure 4 (not to be confused with point *B* in Figure 2).

Figure 4
Assignment maximizing social welfare



In order that a traffic equilibrium in this example be the same as a social-welfare maximizing assignment (that is, point *E* in Figure 4 coinciding with point *A*), the optimality condition (16) and $g_a = g_b$ must hold simultaneously. A numerical example for the network in in Figure 1 comparing the assignment that minimizes total time with the one maximizing total welfare is set out in Appendix *A*.

If the individual travellers have linear utility functions, such as one taking the form $U_i = \eta + \rho g_i$ ($\rho > 0$), the optimality condition (16) becomes

$$\frac{\beta_a}{\beta_b} = \frac{\left(\frac{-1}{\beta_b}(\eta + \rho g_b) + \rho\left(\frac{t - g_b - \alpha_b}{\beta_b}\right)\right)}{\left(\frac{-1}{\beta_a}(\eta + \rho g_a) + \rho\left(\frac{t - g_a - \alpha_a}{\beta_a}\right)\right)} = \frac{\beta_a}{\beta_b}\frac{\left(\rho(t - g_b - \alpha_b) - (\eta + \rho g_b)\right)}{\left(\rho(t - g_a - \alpha_a) - (\eta + \rho g_a)\right)} \tag{17}$$

$$\rho(t - g_b - \alpha_b) - (\eta + \rho g_b) = \rho(t - g_a - \alpha_a) - (\eta + \rho g_a) \tag{18}$$

$$g_a = \left(\frac{\alpha_b - \alpha_a}{2}\right) + g_b \tag{19}$$

As can be observed, (19) is the same as (14), the marginal cost assignment that minimizes the total network cost. Therefore, if travellers have linear utility functions, the minimum cost network assignment is equivalent to the assignment that maximizes social welfare.

3. COMPARISON OF COST MINIMIZING AND WELFARE MAXIMIZING ASSIGNMENT

In this section, we formulate traffic assignment problems based on the minimization of total cost and on maximization of welfare and compare their respective optimality conditions.

3.1 Total cost minimization problem

The formulation of the minimum total cost problem uses the nomenclature set out in Nagurney (2000). Assuming the cost of an arc c_a is positive, increases with flow and is a function only of its own flow (f_a), the problem is specified as follows:

$$\min_{\{h_p^w\}} \quad C = \sum_{a \in A} c_a(f_a)\, f_a \tag{20}$$

s.t.:

$$\sum_{p \in P_w} h_p^w = T_w \qquad (\mu_w) \qquad \forall w \in W \tag{21}$$

$$f_a = \sum_{w \in W} \sum_{p \in P_w} \delta_{ap} h_p^w \qquad \forall a \in A \tag{22}$$

$$h_p^w \geq 0 \qquad \forall p \in P_w, \forall w \in W \tag{23}$$

c_a: cost of arc *a*.
f_a: flow in arc *a*.
δ_{ap}: parameter equal to 1 if arc *a* belongs to route *p*, otherwise 0.
h_p^w: flow on route *p* between origin-destination pair *w*.
T_w: number of trips (fixed demand) between origin-destination pair *w*.

P_w: set of routes between origin-destination pair w.
W: set of origin-destination pairs in network.
A: set of arcs in network.

The total cost of route p is defined as $C_p = \sum_{a \in A} c_a\left(f_a\right)\delta_{ap}$. The objective function of problem (20) can then be written as a function of route costs (Nagurney, 2000):

$$\min_{\{h_p^w\}} \quad C = \sum_{w \in W} \sum_{p \in P_w} C_p h_p^w = \sum_{w \in W} \sum_{p \in P_w} \hat{C}_p^w \tag{24}$$

Expression (24) explicitly assumes that the value of time is constant (and obviously positive), allowing equivalence between the objective of minimizing the total network time and the objective of minimizing the total monetary cost of the network. This assumption is not present when using the criterion of maximizing social welfare as outlined in section 3.2

The first-order optimality condition of (24) is

$$\frac{\partial \hat{C}_p^w}{\partial h_p^w} \begin{cases} = \mu_w, \ \forall h_p^w > 0 \\ \geq \mu_w, \ \forall h_p^w = 0 \end{cases} \tag{25}$$

where $\frac{\partial \hat{C}_p^w}{\partial h_p^w} = \frac{\partial C_p}{\partial h_p^w} h_p^w + C_p$ denotes the marginal cost of an additional unit of flow on route p between OD pair w. The parameters μ_w are the Lagrange multipliers for constraints (21), and represent the increase in total system cost C if demand between OD pair w increases by a unit. Thus, $\mu_w = \frac{\partial C}{\partial T_w} > 0$.

The left-hand side of optimality condition (25) can be rewritten as

$$\frac{\partial \hat{C}_p^w}{\partial h_p^w} = \frac{\partial C_p}{\partial h_p^w} h_p^w + C_p = \frac{\partial}{\partial h_p^w}\left(\sum_{a \in A} c_a \delta_{ap}\right) h_p^w + C_p = \left(\sum_{a \in A} \frac{\partial c_a}{\partial f_a}\frac{\partial f_a}{\partial h_p^w}\delta_{ap}\right) h_p^w + C_p \tag{26}$$

$$\frac{\partial \hat{C}_p^w}{\partial h_p^w} = \left(\sum_{a \in A} \frac{\partial c_a}{\partial f_a}\delta_{ap}\right) h_p^w + C_p \tag{27}$$

The term $\left(\sum_{a \in A} \frac{\partial c_a}{\partial f_a}\delta_{ap}\right) h_p^w$ represents the cost externality generated on route p when route flow increases. The assignment defined by (25) and (27) thus equalizes the marginal costs (including the associated externalities) of all routes used for each OD pair.

3.2 Social welfare maximization problem

The formulation of the problem based on generating maximum social welfare follows the criteria developed by Bergson (1938), Samuelson (1956) and Boadway and Bruce (1984). The problem is specified as follows:

$$\max_{\{h_p^w\}} \quad W = \sum_{w\in W}\sum_{p\in P_w} U_p h_p^w = \sum_{w\in W}\sum_{p\in P_w} \hat{U}_p^w \tag{28}$$

where U_p is the utility perceived by users of route *p* between OD pair *w*, and $\hat{U}_p^w = U_p h_p^w$ is the total utility or welfare of travellers on route *p* between OD pair *w*. Since users are assumed to be homogeneous, the individual utility functions can be written as $U_p = U\left(g_p\right)$.

The optimality condition for problem (28) is

$$\frac{\partial \hat{U}_p^w}{\partial h_p^w}\begin{cases} = \gamma_w, \forall h_p^w > 0 \\ \le \gamma_w, \forall h_p^w = 0 \end{cases} \tag{29}$$

where $\frac{\partial \hat{U}_p^w}{\partial h_p^w} = \frac{\partial U_p}{\partial h_p^w} h_p^w + U_p$ denotes the marginal utility of the social welfare of travellers on route *p* between OD pair *w*. The parameter γ_w measures the reduction in total social welfare of system *W* if demand on pair *w* increases by one unit. The sign of γ_w depends on the level of route congestion.

The optimality conditions for (29) and (25) will in general produce different flow assignments along the routes (see example in Section 2).

Since $\hat{U}_p^w = U_p h_p^w$, from (29) we have

$$\frac{\partial U_p}{\partial h_p^w} h_p^w + U_p \begin{cases} = \gamma_w, \forall h_p^w > 0 \\ \le \gamma_w, \forall h_p^w = 0 \end{cases} \tag{30}$$

Applying the chain rule to (30), we obtain

$$\frac{\partial U_p}{\partial C_p}\frac{\partial C_p}{\partial h_p^w} h_p^w + U_p \begin{cases} = \gamma_w, \forall h_p^w > 0 \\ \le \gamma_w, \forall h_p^w = 0 \end{cases} \tag{31}$$

$$\frac{\partial U_p}{\partial C_p}\frac{\partial \sum_a \left(\delta_{ap} c_a\right)}{\partial h_p^w} h_p^w = \frac{\partial U_p}{\partial C_p}\left(\sum_{a\in A}\delta_{ap}\frac{\partial c_a}{\partial f_a}\frac{\partial f_a}{\partial h_p^w}\right) h_p^w = \frac{\partial U_p}{\partial C_p}\left[\sum_{a\in A}\left(\delta_{ap}\frac{\partial c_a}{\partial f_a}\right)\right] h_p^w \tag{32}$$

The expression $\frac{\partial C_p}{\partial h_p^w} h_p^w = \left[\sum_{a\in A} \left(\delta_{ap} \frac{\partial c_a}{\partial f_a} \right) \right] h_p^w$ is the same as (27). The expression $\frac{\partial U_p}{\partial C_p} = -\frac{\partial U_p}{\partial g_p}$ can be interpreted as the negative of the marginal utility of income. The presence of the term $\frac{\partial U_p}{\partial C_p}$ in (31) is the equivalent of incorporating a weight factor for the cost externality $\frac{\partial C_p}{\partial h_p^w} h_p^w$.

The economic interpretation of the optimality condition (30)-(31) is different from that of (25)-(27). Whereas the latter simply includes the externalities of the arcs constituting the routes taken (independently of the free-flow cost), the former weights the externalities by the marginal utility of income. Since we have assumed traditional utility functions subject to the non-satiation and diminishing marginal utility axioms, that is, $\frac{\partial U_p}{\partial g_p} > 0$ and $\frac{\partial^2 U_p}{\partial g_p^2} < 0$, the routes with higher individual cost in the optimal system assignment will be multiplied by a higher marginal utility of free time than routes demanding lower individual times.

We therefore expect that the maximum welfare assignment will reduce flow on the higher individual cost routes and increase it on the lower cost ones. In this sense, we may argue that assignments based on social welfare maximization à la Bergson-Samuelson will meet less resistance on the part of users than those based on Wardrop’s second principle.

As explained in section 3.1, this difference occurs because the problem (24) of minimum total cost (criterion that is based on Wardrop's second principle) assumes that the value of time is constant, which does not occur in problem (28), since the marginal utility of travel time may vary between different origin-destination pairs in the network.

3.3 Weighted total cost minimization problem

It is possible to modify problem (24) such that it considers a varying value of time that depends on the cost of route *p* for different pairs *w*. Considering a value of time ϕ_p^w, we can pose the following optimization problem:

$$\min_{\{h_p^w\}} \quad \tilde{C} = \sum_{w\in W} \sum_{p\in P_w} \phi_p^w C_p h_p^w = \sum_{w\in W} \sum_{p\in P_w} \phi_p^w \hat{C}_p^w \tag{33}$$

where ϕ_p^w is a weight for the value of time and satisfies $\frac{\partial \phi_p^w}{\partial C_p} > 0, \forall w, p \in P^w$. This last condition is a consequence of the axiom of decreasing marginal utility of consumption of goods (leisure in this case). Then, it follows that $\frac{\partial \phi_p^w}{\partial h_p^w} = \frac{\partial \phi_p^w}{\partial C_p} \frac{\partial C_p}{\partial h_p^w} > 0$.

The first order condition of problema (33) is:

$$\frac{\partial\left(\phi_p^w \hat{C}_p^w\right)}{\partial h_p^w} = \hat{C}_p^w \frac{\partial \phi_p^w}{\partial h_p^w} + \phi_p^w \left(\frac{\partial \hat{C}_p^w}{\partial h_p^w}\right) \begin{cases} = \tilde{\mu}_w, \forall h_p^w > 0 \\ \leq \tilde{\mu}_w, \forall h_p^w = 0 \end{cases} \tag{34}$$

where:

$$\frac{\partial\left(\phi_p^w \hat{C}_p^w\right)}{\partial h_p^w} = \left(C_p^w h_p^w\right) \frac{\partial \phi_p^w}{\partial h_p^w} + \phi^p \left(\frac{\partial C_p}{\partial h_p^w} h_p^w + C_p\right) = \left(C_p \frac{\partial \phi_p^w}{\partial C_p} + \phi^p\right) \frac{\partial C_p}{\partial h_p^w} h_p^w + \phi_p^w C_p \tag{35}$$

Replacing $U_p = \phi_p^w C_p$ we obtain directly:

$$\frac{\partial U_p}{\partial h_p^w} = \frac{\partial\left(\phi_p^w C_p\right)}{\partial h_p^w} \rightarrow \frac{\partial U_p}{\partial C_p} \frac{\partial C_p}{\partial h_p^w} = C_p \frac{\partial \phi_p^w}{\partial C_p} \frac{\partial C_p}{\partial h_p^w} + \phi_p^w \frac{\partial C_p}{\partial h_p^w} \tag{36}$$

From expressions (31), (34) and (36) it is possible to conclude that it exists a vector of weights ϕ_p^w that allows us to transform the problem of minimizing total costs in and equivalent problem of maximizing total welfare.

4. SOME PROPERTIES OF THE LINEAR COST FUNCTIONS

4.1 Simple road network with linear costs

For the simple network of Section 2, optimality condition (25) would be expressed as follows:

$$\frac{\partial c_a}{\partial f_a} f_a + c_a = \frac{\partial c_b}{\partial f_b} f_b + c_b \tag{37}$$

Since $\frac{\partial c_a}{\partial f_a} = \beta_a$ and $\frac{\partial c_b}{\partial f_b} = \beta_b$ and also $f_a = \frac{c_a - \alpha_a}{\beta_a}$ and $f_b = \frac{c_b - \alpha_b}{\beta_b}$, if we then substitute these into (37) we obtain

$$\beta_a \left(\frac{2c_a - \alpha_a}{\beta_a}\right) = \beta_b \left(\frac{2c_b - \alpha_b}{\beta_b}\right) \rightarrow 2\left(c_a - c_b\right) = \left(\alpha_a - \alpha_b\right) \tag{38}$$

Given that in market equilibrium $c_a = c_b$, it will also be true that $\alpha_a = \alpha_b$. Therefore, if the fixed costs of the alternative routes are equal, the market equilibrium assignment will minimize total network time.

4.2 General network with linear costs

The result in 4.1 can be easily generalized to networks with linear costs and the same free-flow route times from optimality condition (25), which states that

$$\frac{\partial \hat{C}_p^w}{\partial h_p^w} = \frac{\partial \hat{C}_q^w}{\partial h_q^w}, \quad \forall (p,q) \in P_w, h_p^w > 0, h_q^w > 0 \tag{39}$$

Since $\frac{\partial \hat{C}_p^w}{\partial h_p^w} = \sum_{a \in A} \frac{\partial (c_a \cdot f_a)}{\partial f_a} \delta_{ap}$ we get

$$\frac{\partial \hat{C}_p^w}{\partial h_p^w} = \sum_{a \in A} \left(c_a + \frac{\partial c_a}{\partial f_a} f_a \right) \delta_{ap} = \sum_{a \in A} (c_a \delta_{ap}) + \sum_{a \in A} \left(\frac{\partial c_a}{\partial f_a} f_a \delta_{ap} \right) \tag{40}$$

Given that $\frac{\partial c_a}{\partial f_a} = \beta_a$ and $f_a = \frac{c_a - \alpha_a}{\beta_a}$, (40) reduces to

$$\frac{\partial \hat{C}_p^w}{\partial h_p^w} = \sum_{a \in A} (c_a \delta_{ap}) + \sum_{a \in A} \left(\beta_a \left(\frac{c_a - \alpha_a}{\beta_a} \right) \delta_{ap} \right) \tag{41}$$

$$\frac{\partial \hat{C}_p^w}{\partial h_p^w} = 2C_p - \sum_{a \in A} (\alpha_a \delta_{ap}) \tag{42}$$

Now imposing optimality condition (39), we obtain

$$\frac{\partial \hat{C}_p^w}{\partial h_p^w} = \frac{\partial \hat{C}_q^w}{\partial h_q^w} \rightarrow 2C_p - \sum_a (\alpha_a \delta_{ap}) = 2C_q - \sum_a (\alpha_a \delta_{aq}) \tag{43}$$

Thus, if the free-flow costs of the routes used between OD pair *w* are the same $\left(\sum_{a \in A} (\alpha_a \delta_{ap}) = \sum_{a \in A} (\alpha_a \delta_{aq}) \right)$, then from (43) we arrive at

$$C_p = C_q, \quad \forall (p,q) \in P_w, h_p^w > 0, h_q^w > 0 \tag{44}$$

We have thus demonstrated that if the arc cost functions are linear, the market equilibrium assignment (Wardrop's first principle) is consistent with a marginal cost assignment that minimizes total system cost (Wardrop's second principle) as long as the free-flow costs of the routes used are all the same. The optimality conditions (29) and (25) will be equivalent regardless of the type of arc cost functions as long as the users have linear utility functions. Indeed, if $U_p = \eta + \rho g_p$ ($\rho > 0$) and all users are equal with the same amount of free time, it is easily demonstrated that

$$\sum_w \sum_p U_p h_p^w = \sum_w \sum_p (\eta + \rho g_p) h_p^w = \eta \sum_w \sum_p h_p^w + \rho \sum_w \sum_p g_p h_p^w \tag{45}$$

$$\sum_w \sum_p U_p h_p^w = \eta \sum_w T_w + \rho \sum_w \sum_p (T - C_p) h_p^w \tag{46}$$

$$\sum_w \sum_p U_p h_p^w = (\eta + \rho T) \sum_w T_w - \rho \sum_w \sum_p C_p h_p^w \quad (47)$$

Since $(\eta + \rho T) \sum_w T_w$ is constant and $\rho > 0$, $\max \sum_w \sum_p U_p h_p^w = \min \sum_w \sum_p C_p h_p^w$.

5. CONCLUSIONS

Many urban transport planning processes have developed methodologies for optimizing resource use based on criteria such as Wardrop's second principle, according to which transport networks are designed or traffic flows assigned so as to minimize travellers' total cost. However, the results thus obtained may differ if instead of minimizing cost, the objective is to maximize social welfare.

Following a framework similar to that of Gabriel and Bernstein (1997), this paper demonstrated that if we define a social welfare function based on the sum of individual traveller utilities à la Bergson-Samuelson, market equilibrium in a transport network as defined by Wardrop's first principle may generate a social welfare level higher than that achieved by a traffic assignment based on Wardrop's second principle. In other words, minimizing total network costs may result in lower social welfare.

The minimization of total costs for a given trip can lead to users being assigned to routes whose cost to them individually is higher than alternative routes due to the former routes' relatively lower marginal cost for the network as a whole. Because of declining marginal utility, the impact on these users' utility per unit of time is greater than that experienced by those assigned to routes with shorter trip times. Since marginal utility is key to the maximum welfare assignment, this assignment reduces the cost (and welfare) difference between users made better off and those made worse off. Thus, a Bergson-Samuelson welfare-maximizing assignment should meet with less resistance from users than one following Wardrop's second principle that minimizes total system cost. If, however, the individual travellers' utility functions are linear, the second Wardrop principle assignment will coincide with the social-welfare maximizing assignment for any route arc cost function.

Finally, it was demonstrated that if the route arcs have linear cost functions that depend on own flow, and if the alternative routes all have the same free-flow costs, the market equilibrium assignment will be the same as the one that minimizes total system cost. In these cases, the market will assign resources optimally even under congestion.

REFERENCES

Agdeppa, R. P.; Yamashita, N. and Fukushima, M. (2007). The traffic equilibrium problem with nonadditive costs and its monotone mixed complementarity problem formulation. Transportation Research, 41B, 862–874.

Arnott, R. and Small, K. (1994). The economics of traffic congestion. American Scientist, 82, 446 - 455.

Beckmann, M., McGuire, C. B. and C. B. Winsten. (1956). Studies in the Economics of Transportation. Yale University Press, New Haven, CT.

Bell, M.G.H. and Iida, Y. (1997). Transportation Network Analysis. John Wiley, Chichester, UK.

Bergson, A. (1938). A Reformulation of Certain Aspects of Welfare Economics', Quarterly Journal of Economics, 52, (2), 310-334.

Boadway, R. and N.Bruce. (1984). Welfare Economics Basil Blackwell, Cambridge.

Brotcorne, L.; Labbé, M.; Marcotte, P. and Savard, G. (2001). A Bilevel model for toll optimization on a multicommodity transportation network. Transportation Science, 35, 345-358.

Button K.J. and Verhoef E.T. (eds.) (1998) Road Pricing, Traffic Congestion and the Environment: Issues of Efficiency and Social Feasibility Edward Elgar, Cheltenham.

Easa, S. M. (1991). Traffic assignment in practice: overview and guidelines for users. Journal of Transportation Engineering, 117, 602-622.

Gabriel, S. and Bernstein, D. (1997). The traffic equilibrium problem with nonadditive path costs, Transportation Science, 31, 337–348.

Han, D. and H. Lo. (2004). Solving Non-additive Traffic Assignment Problems: A Decent Methods for Co-coercive Variational Inequalities. European Journal of Operational Research, 159(3), 529-544.

Koh, A.; Shepherd, S. and Sumalee, A. (2009). Second best toll and capacity optimisation in networks: solution algorithm and policy implications. Transportation, 36, 147-165.

Larsson, T.; Lindberg, P.O.; Lundgren, J. and Patriksson, M.; and Rydergren, C. (2002). On traffic equilibrium models with a nonlinear time/money relation, in: M. Patriksson, M. Labbe (Eds.), Transportation Planning: State of the Art, Kluwer Academic Publishers.

Lo, H.K. and Chen, A. (2000). Traffic equilibrium problem with route-specific costs: Formulation and algorithms. Transportation Research Part, 34B, 493-513.

Marcotte, P. (1983). Network Optimization with Continuous Control Parameters. Transportation Science, 17, 181-197.

Mun, S., Konishi, K. and Yoshikawa. K. (2003). Optimal Cordon Pricing. Journal of Urban Economics, 54, 21–38.

Nagurney, A. (2000). Sustainable Transportation Networks, Edward Elgar Publishing, Cheltenham, England.

Newbery, D. M., (1989). Cost recovery from optimally designed roads. Economica, 56, 165-85.

Parry, I. W. H. (2002). Comparing the Efficiency of Alternative Policies for Reducing Traffic Congestion. Journal of Public Economics, 85,333–362.

Samuelson, P. (1956). Social Indifference Curves. The Quarterly Journal of Economics, 70, 1-22.

Santos, G. (2004). Urban Congestion Charging: A Second-Best Alternative. Journal of Transport Economics and Policy, 38,345–369.

Shepherd, S. P., and Sumalee. A. (2004). A Genetic Algorithm Based Approach to Optimal Toll Level and Location Problems. Networks and Spatial Economics, 4, 161–179.

Smith, M.J. (1979). The marginal cost pricing of a transportation network. Transportation Research, 13B, 237-242.

Small, K.A. (1992) Urban Transport Economics Harwood, Chur.

Small, K. A. and Gómez-Ibáñez, J. A. (1998). Road Pricing for Congestion Management: The Transition from Theory to Policy." In Road Pricing, Traffic Congestion and the Environment: Issue of Efficiency and Social Feasibility, edited by K. J. Button and E. T. Verhoef. Cheltenham: Edward Elgar, 213–246.

Verhoef, E.T. (2002). Second-best congestion pricing in general networks: heuristic algorithms for finding second-best optimal toll levels and toll points. Transportation Research, 36B, 707-729.

Wardrop, J. (1952). Some theoretical aspects of road traffic research. Proceeding of the Institute of Civil Engineers, 2, 325-378.

Yang, H. and Bell, M.G. (1998). Models and algorithms for road network design: a review and some new developments. Transport Reviews, 18, 257-278.

Yang, H. and Meng, Q. (2002). A note on highway pricing and capacity choice under a build-operate-transfer scheme. Transportation Research , 36A, 659-663.

Yanga, H.; Xua, W. and Heydecker,B. (2010). Bounding the efficiency of road pricing. Transportation Research, 46E, 90-108.

APPENDIX

An analytical example for a simple network

Assume that in the simple network in Figure 1, total fixed demand $f_a + f_b = 100$ and the cost functions for each arc are

$$c_a(f_a) = 50 + f_a \quad , \quad c_b(f_b) = 120 + 5f_b \tag{48}$$

The market equilibrium point (Wardrop's first principle) is then

$$f_a^* = 95 \ , \quad f_b^* = 5 \ , \quad c_a(f_a^*) = c_b(f_b^*) = 145 \tag{49}$$

The total cost of the network is $c_a(f_a^*) \cdot f_a^* + c_b(f_b^*) \cdot f_b^* = 14.500$. If the utility function is given by $U_i = \frac{(g_i)^{1-\theta}}{1-\theta} = \frac{(t-c_i)^{1-\theta}}{1-\theta}$ where t = 200 and θ = 0.7, the value of the social welfare function is $U(c_a^*) \cdot f_a^* + U(c_b^*) \cdot f_b^* = 1{,}109.1$.

The minimum total cost assignment (Wardrop's second principle) is

$$f_a^* = 89.17 \ , \quad f_b^* = 10.83 \ , \quad c_a(f_a^*) = 139.17 \ , \quad c_b(f_b^*) = 174.17 \tag{50}$$

The total network cost is $c_a(f_a^*) \cdot f_a^* + c_b(f_b^*) \cdot f_b^* = 14{,}295.8$ and the value of the welfare function is $U(c_a^*) \cdot f_a^* + U(c_b^*) \cdot f_b^* = 1{,}115.1$. Comparing (49) with (50), we find that in the latter case total cost falls by 204.2 units while social welfare rises by 6 units. Thus, in this example, moving to a minimum cost assignment also increases social welfare.

If, however, the cost for arc *a* is $c_a(f_a) = 100 + f_a$, the market equilibrium point is then

$$f_a^* = 86.7 \ , \quad f_b^* = 13.3 \ , \quad c_a(f_a^*) = c_b(f_b^*) = 186.7 \tag{51}$$

Total network cost is $c_a(f_a^*) \cdot f_a^* + c_b(f_b^*) \cdot f_b^* = 18{,}666.7$ and the value of the social welfare function is $U(c_a^*) \cdot f_a^* + U(c_b^*) \cdot f_b^* = 725$.

The minimum cost assignment is

$$f_a^* = 85 \ , \quad f_b^* = 15 \ , \quad c_a(f_a^*) = 185 \ , \quad c_b(f_b^*) = 195 \tag{52}$$

The total network cost is $c_a(f_a^*) \cdot f_a^* + c_b(f_b^*) \cdot f_b^* = 18{,}650$ and the value of the social welfare function is $U(c_a^*) \cdot f_a^* + U(c_b^*) \cdot f_b^* = 719.5$.

Comparing (51) and (52) we see that in the latter case, total cost fell by 16.7 units but this time social welfare also fell, by 5.5 units. Thus, in this example the minimum total cost assignment reduces social welfare.