Netnomics (2009) 10:141–160
DOI 10.1007/s11066-008-9021-2

Joint pricing and design of urban highways with spatial and user group heterogeneity

Loukas Dimitriou · Theodore Tsekeris · Antony Stathopoulos

Accepted: 14 August 2008 / Published online: 12 September 2008

Abstract This paper addresses the joint optimization of capacity investments and toll charges imposed on multi-group users in monopolistic private highways within general road networks. A game-theoretic formulation is provided that leads to a nonconvex bilevel program. The proposed modeling framework handles several complex issues raised in realistic applications, such as regulations on the levels of tolls and service, and the discrete nature of highway capacity, using a genetic optimization technique. Real-application results show the importance of considering the spatial heterogeneity of prices, and the tradeoff between investments and pricing strategies in regulated private highways.

Keywords Road pricing · Network design · Regulated private highways · Urban networks · Genetic optimization

L. Dimitriou (✉) · A. Stathopoulos
Department of Transportation Planning and Engineering,
School of Civil Engineering, National Technical University of Athens,
Iroon Polytechniou 5, 15773 Athens, Greece
e-mail: lucdimit@central.ntua.gr

A. Stathopoulos
e-mail: a.stath@transport.ntua.gr

T. Tsekeris
Centre for Planning and Economic Research (KEPE),
Amerikis 11, 10672 Athens, Greece
e-mail: tsek@kepe.gr

1 Introduction

In the current context of deregulation and privatization of network industries, such as those of transport, telecommunication and electricity, the involvement of the private sector has led to significant changes in the infrastructure investment and pricing strategies. In relation to the road transport networks, the growing demand for mobility and the need for restraining public expenditures have led many governments worldwide to adopt privatization mechanisms for financing new or expanded road projects. In particular, the build–operate–transfer (BOT) concession schemes allow private investors to build new highways and operate them by collecting toll charges for a given number of years (concession period), sufficient to attain an agreed level of investment benefit, and then transfer them to the government. In this way, the funding sources are implicitly transferred from the tax-payers to the users of the specific roads.

For the successful implementation of such schemes, a revenue-maximizing firm should consider the tradeoff between the cost of road capacity provision and the revenues generated by the toll charges imposed on users. Namely, the private-sector operator should take into account both the level of capacity and prices in order to determine the expected volume of highway users and the toll revenues. Also, a number of constraints are usually required to be satisfied in practice, in order to meet firm objectives and increase the political and social acceptability of such schemes. Specifically, system designers/operators, who are subject to budgetary constraints, require imposing charges to recover construction costs and afford the operating and maintenance costs. Furthermore, the capacity choices affect the level of service, either directly by relieving congestion or by influencing pricing decisions and, in turn, traffic conditions in the concession period. Thus, the government/regulator may wish to exercise control on the minimum required level of service, in addition to the constraints imposed on toll rates. Especially, it may sometimes be an acceptable (and politically prudent) practice to set tolls at levels which would attract such an amount of traffic that meets certain mobility and environmental targets, rather than maximizes revenue.

Despite the evident need for jointly addressing the optimal road capacity and pricing decisions, these issues have mostly been treated separately in the current literature. Particularly, the capacity investment plans are commonly investigated through the framework of the Network Design Problem (NDP), which encompasses a class of bilevel programming or mathematical programming with equilibrium constraints (MPEC) models. These models seek to optimize the system designer objectives subject to a variety of physical, budgetary, reliability and other constraints, while taking into account user responses [1, 2]. Similar modeling approaches are employed to examine optimal toll rates under a range of pricing tactics and behavioral assumptions, within the framework of the toll design problem (TDP) [3–7].

This paper presents a general modeling framework for simultaneously determining the optimal level of capacity provision and user charges in toll roads, through suitably integrating the two types of problems mentioned previously (i.e., the NDP and the TDP). The proposed framework encompasses several alternative network design and pricing schemes, incorporating the spatial heterogeneity of prices as well as the heterogeneity among user groups with different characteristics. It also handles a number of realistic complications concerning the interactions among the system designer/operator, the government/regulator and the users, including the constraints imposed on the level of tolls and service.

Section 2 discusses crucial issues involved in the joint determination of network design and pricing strategies. Section 3 presents a game-theoretic formulation of the current modeling framework. Section 4 describes an evolutionary solution procedure for addressing the complexity of the joint design and pricing problem. Section 5 reports the computational experiments and results obtained from a real-world urban network application of the model, and Section 6 summarizes and concludes.

2 Issues in joint decisions of optimal road investment and pricing

The process of simultaneously determining the optimal capacity investments and prices in toll roads has been considered only in a few studies in the existing literature [8–11]. The studies are mostly restricted to simple networks, having either parallel or serial links. The given problem has been also examined for the case of general inter-urban road networks in [12–14] under alternative project objectives and market condition. These models have considered a standard, deterministic user equilibrium (DUE) traffic assignment procedure with elastic demand to represent the responses of users to optimal flat (uniform) tolls and capacity investments. Despite that the aforementioned studies have demonstrated the need for jointly considering toll and capacity decisions, several of their modeling assumptions can be regarded as oversimplified, departing from the realistic conditions underlying the design and pricing of private (BOT) road networks, and the behavior of their users.

For this reason, the present paper provides a number of extensions to the existing approaches dealing with the joint network design and pricing problem. Specifically, the (here, revenue-maximizing) strategies of the system designer/operator take into account the responses of users of multiple groups. In this way, they can consider a range of alternative responses due to the heterogeneity in the value of travel time, based on different incomes and trip purposes, in comparison to the other studies, which treat traffic as homogeneous. Such a treatment circumvents potential problems of underestimating the benefit of road pricing for the users. In addition, the present approach incorporates variations in the perceived generalized travel cost and route choice uncertainty of multi-class users, so that provide more realistic assumptions about the assignment of travel demand onto the road network. These features give rise to a multi-class stochastic user equilibrium (SUE) assignment procedure with elastic travel demand.

Furthermore, the current framework provides an extension of previous work [15] to allow handling various types of spatially differentiated (entry–exit–based as well as entry–exit-based) toll pricing strategies. Such types of strategies differ from the commonly adopted uniform toll strategy which ignores the spatial heterogeneity in the charges of users with different origin–destination (O–D) travel patterns. The inclusion of this element of spatial heterogeneity into the pricing strategies can arguably yield a more efficient and fair charging mechanism, leading to increased public acceptability of the BOT schemes and, hence, raising the demand for highway usage (see Section 5).

Another strong assumption typically employed in the existing literature (e.g., see [11–14]) treats road capacity as being adjustable in continuous increments. Provided that, in reality, the number of lanes or links is discrete, the expression of capacity as a discrete variable provides a solution that is infrastructure related, which may involve greater physical intuition and bearing in the road network design for investment planning purposes, in comparison to the solution of the corresponding continuous version of the problem. The resulting problem is referred to here as the joint discrete network design and pricing problem (JDNDPP). However, the adoption of discrete units of capacity leads to a mixed-integer problem, which further complicates the solution procedure. For this reason, an evolutionary optimization approach is adopted here to provide an efficient solution for the complex JDNDPP.

Moreover, as it will be shown in the following section, the game-theoretic formulation of the proposed framework enables to identify interactions among multiple players, i.e. the system designer/operator, the government/regulator and the users. The representation of these interactions allows the evaluation of different highway design and pricing options, through altering the components (objectives and constraints) of the problem setup. In turn, more realistic BOT investment schemes can be produced that enhance their attractiveness to the urban system stakeholders (users, public agencies and private investors).

3 Game-theoretic formulation of the problem

The processes of road network design and pricing can be considered as a class of two-stage Stackelberg games with perfect information among alternative players. These games recognize the principal players, which are the system designer/operator and the system users, and they can be generally expressed as bilevel programming problems. Such a type of formulation is commonly met in the literature when considering separately the NDP [16] and the TDP [17]. In the present problem, at the upper level, the system operator (the ‘leader’), taking into account a number of constraints, integrates within the design and pricing tactics the non-cooperative responses of users of different classes (the ‘followers’). These responses can involve changes in the trip-making and route choice behavior, as captured here through a multi-class SUE assignment with elastic demand. The specific traffic assignment model is formulated as an unconstrained minimization problem [18], which is performed at the lower level.

In this paper, the private highway market is regarded as monopolistic and subject to governmental control on the level of tolls and service. In this way, a third player is added, i.e. the government (or regulator), whose decisions affect the performance of the other two players. Also, it is assumed that the set (number and location) of the highway links has already been identified before the problem is formulated and solved. This assumption typically holds for a realistic network design process that takes place in an urban environment, as in the present case. Therefore, the problem seeks to estimate the (uniform or spatially differentiated) toll charges and/or the number of link lanes from the private-sector designer/operator point of view, so that the tolled highway, which competes with free alternative roads, will attract such a portion of multi-class users that optimizes the profit or net revenue.

Consider a network $G(N, A)$ composed of a set of N nodes and A links, which connect the origin zone r with destination zone s, and q_m^{rs} be the demand of the users of class m for moving between the O–D pair r–s. In this context, each class m corresponds to a particular user group having an assumed common value of travel time (VOTT), which may reflect similar socio-economic and travel characteristics, which reflect the value of travel time (VOTT), which may reflect similar socio-economic and travel characteristics. As it is typically adopted in the literature, this study assumes that, for each user group, travelers share the same discrete VOTT probability distributions [19]. It is noted that other user group definitions may be additionally adopted, according to the nature of each application, such as the type of vehicle (private car and truck) in relation to the damage causing to the pavement condition.

Also, consider the link travel time function $t_a(x_a)$ as being positive and monotonically increasing with traffic flow x_a at each link $a \in A$. Then, the complete form of the bilevel optimization framework, which refers to the JDNDPP, for the design and pricing of a private highway that constitutes part of the network can be expressed in the upper-level problem as follows:

$$\max_{p,\, y} R(p, y) = \sum_{a \in \widehat{A}} E\{p_a f_a - \lambda V_a(y_\alpha(w_a))\} \tag{1}$$

$$\text{subject to} \quad w_a \in \{0, 1, \ldots, \ell\}, \quad \forall a \in \widehat{A} \tag{2}$$

$$p_{\min} \le p_a \le p_{\max}, \quad \forall a \in \widehat{A} \tag{3}$$

$$\sum_{a \in \widehat{A}} (V_a(y_\alpha(w_a))) \le B, \quad \forall a \in \widehat{A} \tag{4}$$

$$x_a(p, y)/y_a \le L \quad \forall a \in \widehat{A} \tag{5}$$

while the SUE link flow conditions are estimated at the lower-level problem:

$$\min_{x,q} Z(x,\ q) = \sum_{rsakm} \text{VOTT}_m\, \delta^{rs}_{a,km}\, t_a(x_a)\ x_a - \sum_{rsakm} \text{VOTT}_m\, \delta^{rs}_{a,km} \int_0^{x_a} t_a(w)\, dw$$
$$+ \sum_{rsm} \text{VOTT}_m D^{-1}_{rsm}\ \left(q^{rs}_m\right) D_{rsm}\left(S^{rs}_m(x)\right) - \sum_{rsm} S^{rs}_m(x)\ D_{rsm}\left(S^{rs}_m(x)\right)$$
$$+ \sum_{rsm} \int_0^{q^{rs}_m} \text{VOTT}_m\ D^{-1}_{rsm}(q)\, dq - \sum_{rsm} q^{rs}_m\, \text{VOTT}_m\ D^{-1}_{rsm}\left(q^{rs}_m\right) \tag{6}$$

In the upper-level problem, R denotes the expected net revenue function of the private-sector highway designer/operator, which is the objective function of the complete form of the JDNDPP. In this function, the first component refers to the revenues raised from the toll charges, p_a, imposed on the entry volume f_a of users. In the case of uniform (spatially non-differentiated) toll pricing, the revenues are independent of the selection of entry points or the entry–exit travel pattern followed by users in the highway.

In the case of entry-based toll pricing, the revenues depend on the selection of the entry point, and index a refers to the highway access (entry) link $a \in \widehat{A} \subset A$ with capacity y_a, whose entry node is used by travelers to access the highway, where $\widehat{A}$ is the set of highway links. In the case of entry–exit-based toll pricing, the revenues depend on the entry–exit travel pattern and, hence, index a refers to every (entry, intermediate or exit) link of the highway. The latter case can be regarded as the most efficient and fair pricing tactic, since the highway users are charged on the basis of the pay-as-you-drive principle. Specifically, given the linear nature of the highway network, the link-specific tolls in the entry–exit-based pricing strategy give rise to path costs which are link-additive, assuming that users value time linearly. Hence, each user is charged for the specific path following (or the amount of capacity consuming) along the highway. It is noted that in the entry-based pricing strategy, f_a (that differs from the link flow x_a) stands for the amount of users entering the highway at the entry point of link α. In the entry–exit-based pricing strategy, which charges the use of every link of the highway, f_a coincides with x_a.

The second component of the objective function (1) corresponds to the monetary expenditures V_a for capacity provision at highway link $a \in \widehat{A}$, where λ is a parameter that transfers the capital cost of the project into unit period cost, depending on the number of years of operating the project by the private sector. The scalar w_a is an integer decision variable which determines the number of lane additions in link $a \in \widehat{A}$, up to a physical threshold ℓ, as shown in relationship (2). The scalars $p_{\min}$ and $p_{\max}$ in relationship (3) denotes the minimum and maximum allowable toll charges, which are controlled by the government. The budgetary restrictions are represented in inequality (4), where B is the total available highway construction budget.

Relationship (5) introduces the regulatory control of the government on the minimum required level of service (mobility target) in the set of constraints, where L is the maximum allowable flow-to-capacity (x_a/y_a) ratio for each highway link $a \in \widehat{A}$. This operational target is used to ensure a desired balance between infrastructure supply and highway utilization rate that enhances mobility, in terms of reducing travel times between the various O–D pairs, after the new network configuration. It should be noted here that the simultaneous consideration of budget and level-of-service requirements in the set of constraints may result in an unfeasible solution. This possibility stresses the need for careful selection of the bound of each problem constraint.

At the lower-level problem, Z expresses the objective function of network users of different classes m, in terms of their value of travel time VOTT_m, who seek to minimize their perceived generalized travel cost. The traffic assignment procedure assumes that users have elastic demand and their route choice behavior is consistent with the SUE link flow conditions, in the sense that no traveler can improve his/her perceived travel time by unilaterally changing routes [20]. The binary parameter $\delta^{rs}_{a,km}$ takes the value 1, if link a is part of the path k of the feasible path set K_{rs} followed by users of group m between r–s, or 0 otherwise. Assuming that the demand function D_{rsm} is nonnegative and strictly decreasing with respect to the cost of paths between r–s, then $q^{rs}_m = D_{rsm}\left(S^{rs}_m\right)$ and $S^{rs}_m = D^{-1}_{rsm}\left(q^{rs}_m\right)$, where D^{-1}_{rsm} is the inverse demand function and S^{rs}_m is the perceived travel cost function. The latter function is expressed in relation to the expected value E of the total path travel cost C^{rs}_{km}, as follows:

$$S^{rs}_m(x) = E\left\lfloor \min_{k \in K^{rs}} \left\{C^{rs}_{km}\right\} \mid C^{rs}(x) \right\rfloor, \tag{7}$$

with

$$\frac{\partial S^{rs}_m\left(C^{rs}\right)}{\partial C^{rs}_{km}} = P^{rs}_{km}, \tag{8}$$

where P^{rs}_{km} denotes the probability that users of class m select path k between r–s pair. Then, the measure of probability P^{rs}_{km} depends on the following utility function:

$$U^{rs}_{km} = -\theta\, C^{rs}_{km} + \varepsilon^{rs}_{km} \tag{9}$$

where U^{rs}_{km} expresses the utility of users of class m selecting path k between r–s pair, θ is the path cost perception parameter and ε^{rs}_{km} is a random error term, independent and identically distributed (iid) for all routes, which is here assumed to follow a Gumbel distribution, hence, yielding a logit model. The path travel cost C^{rs}_{km} is expressed in monetary terms, as a composite function of the value of travel time (VOTT) and toll charge:

$$C^{rs}_{km} = \sum_{a \in A} \text{VOTT}_m\, \delta^{rs}_{a,km}\, t\left(x_a\right) + \sum_{a \in \widehat{A}} \delta^{rs}_{a,km}\, p_a \tag{10}$$

The adoption of the SUE assumption denotes that the resulting equilibrium flows correspond to the most probable (expected) flow pattern. The effect of this stochasticity on the performance of the upper-level problem is represented through the expected value operator in Eq. 1.

The estimation of the demand responses of users to changes in their path travel costs is based on the following relationship [21]:

$$D_{rsm}^{(n)} = D_{rsm}^{0} \exp\left(uC_{rs}\right), \ \forall\, r, s, m \tag{11}$$

where D_{rsm}^{0} refers to the potential demand (or the demand at zero cost) expressing the maximum desire for travel of users of class m for the $r-s$ pair and u is a scaling parameter (here, a relatively large elasticity value $u = -0.3$ is used, since the path cost is expressed in monetary terms).

4 The evolutionary solution approach

The current formulation of the joint network design and pricing problem results in a non-deterministic polynomial-time (NP-hard) mixed-integer programming problem of increased computational complexity. This is because of the nonlinearity of the functions involved in the upper and lower-level problems, and the combinatorial selection of the highway link lanes. The nonlinearity and nonconvexity portend the existence of local solutions, which imply that it might be difficult to solve for a global optimum (or adequately near-optimum) solution and its uniqueness conditions cannot be met. In view of the difficulty in applying the standard algorithmic approaches for search of the global optimum, this study adopts an evolutionary computation approach, based on a genetic algorithm (GA). GAs have been widely used in various bilevel programming problems [22] and, specifically, in addressing the discrete NDP [1, 23] and the TDP [6, 17].

These algorithms employ population-based stochastic, global search mechanisms suitable for the solution of bilevel mixed-integer programs with multiple constraints, requiring information only about the performance of a 'fitness' function for various candidate states [24]. The steps of the current iterative solution procedure are given in the pseudo-code shown in Table 1. The GA population is composed here of 50 individuals, each of them corresponding to alternative binary codings of number of lanes and toll charge combinations.

After the random selection of an initial population, in accordance with the problem constraints, the expected revenue function (1) is mapped into a fitness function. Subsequently, the solution of the multi-class SUE assignment model with elastic demand provides the equilibrium state of the system for each candidate solution. Then, the fitness function is evaluated for each candidate solution, and the genetic operations are employed to produce an improved population. The reproduction operator employs a tournament selection between three candidate individuals. The crossover operation uses a relatively high crossover rate equal to 80%, augmented with an elitism strategy (10% of the best individuals from the ten preceding generations feeds the current generation), which enhances the probability of the individuals with good performance to exchange genetic information. A mutation rate equal to 5% is used to diminish the probability of finding a false peak. The convergence of the GA is considered to be achieved when the mean of

Table 1 The steps of the evolutionary optimization procedure used for solving the network design and pricing problems

Steps	Description
Step 1 (initialization)	Produce an initial random population of candidate feasible solutions (toll charges and link lanes) and select the properties of the genetic operators
DO UNTIL convergence	
Step 2 (lower-level problem)	Perform path enumeration and produce the SUE link flow pattern for each candidate solution
Step 3 (performance evaluation)	Check for the consistency of constraints and estimate the fitness function for each candidate solution
Step 4 (upper-level problem)	Produce a genetically improved population of candidate solutions through the stochastic selection of the 'fittest' solution set, crossover operation among the selected individuals, elitist movement of individuals from previous generations and mutation of individuals

the current population performance coincides with the best solution so far obtained over ten successive generations, or no significant improvement of the best solution is obtained over 20 successive generations, or a maximum number of 50 generations has been performed. Table 2 provides a summary of the settings of the current GA operators whose specification is based on [24].

Provided that GAs belong to the class of stochastic approximation optimization methods, multiple exhaustive repetitions of the algorithm have been performed here to increase the probability that local-optimal traps have been avoided and the resulting solution is adequately near-optimal. It is also noted that the given GA structure could be well applied to address similar road pricing and design problems in both urban and inter-city highway networks of much larger dimensions than the present one (see Section 5).

The objective function (1) is composed of two conflicting components, giving rise to a multi-objective problem setup, which implies that the solutions form a so-called Pareto Set (or Pareto Frontier). Namely, the maximum benefit (maximum profit) can be derived by multiple, properly selected combinations of road capacity additions and toll rates. Such combinations can range from the provision of small capacity additions which are expensive to use, to the provision of large capacity additions servicing a high level of demand with lower user charges. In this sense, the proposed framework can identify at least

Table 2 Selection of the evolutionary characteristics of the genetic algorithm

Genetic operation	Characteristic
Initialization	Random
Selection	Tournament selection among three candidates
Crossover	80% (n-point)
Mutation	5%
Elitism	10%
Stopping criteria	Max. 50 generations, or no improvement after 20 consecutive generations

one optimal solution to the joint design and pricing problem. Alterations to the GA could be employed so that the optimal solution set (i.e. those solutions composing the final generation) to yield a Pareto Frontier [25].

5 Computational experiments and results

5.1 Experimental setup of the study

The proposed modeling framework is implemented into a suitably selected part of the urban road network of Athens, Greece, that is composed of primary and secondary roads, which are linked with a closed orbital urban highway, called Attiki Odos. The network (see Fig. 1) covers the most densely populated region along the highway, where the heaviest daily traffic volumes are observed. It is composed of 54 (internal and connecting) links servicing the demand represented by a 10×10 O–D matrix. Table 3 presents the physical and operating characteristics of the internal links of the network under study, with each link lane having a capacity equal to 1,500 veh/h. Attiki Odos operates under a BOT concession scheme by a private firm, which currently imposes uniform toll charges on users (2.70 €/private car) at all highway access points.

Based on the socio-economic and travel characteristics of the case study area, two VOTT user classes are identified. The first class I has an hourly $\text{VOTT}_{\text{I}} = 4.0$ €, representing predominantly commuters and business travelers, and refers to the 80% of the traveler population, while the second class II has

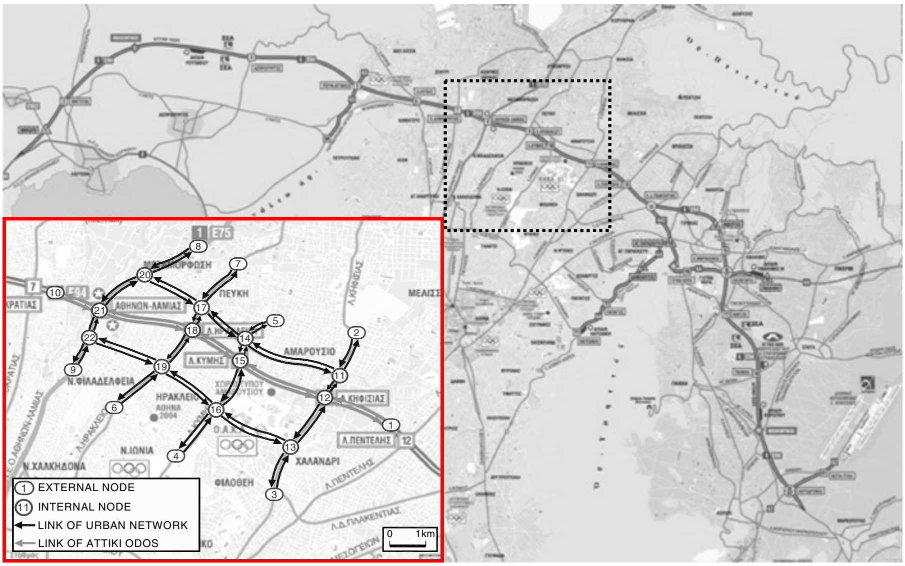


Fig. 1 Map of the study area, and configuration and coding of the urban network and tolled highway

Table 3 Physical and operating characteristics of the network under study

Link	Link length (m)	Link free flow speed (km/h)	Link lanes[a]
12-11	600	80	3
11-12	600	80	3
13-12	1,300	80	3
12-13	1,300	80	3
11-14	2,080	50	1
14-11	2,080	50	1
12-15	1,960	100	3
15-12	1,960	100	3
13-16	2,050	80	2
16-13	2,050	80	2
14-15	400	50	2
15-14	400	50	2
15-16	1,200	50	2
16-15	1,200	50	2
14-17	1,170	30	1
17-14	1,170	30	1
15-18	1,270	100	3
18-15	1,270	100	3
16-19	1,780	50	1
19-16	1,780	50	1
17-18	530	50	1
18-17	530	50	1
18-19	1,310	50	1
19-18	1,310	50	1
17-20	1,810	30	1
20-17	1,810	30	1
18-21	1,990	100	3
21-18	1,990	100	3
19-22	1,820	30	1
22-19	1,820	30	1
21-22	760	100	3
22-21	760	100	3
20–21	890	100	3
21–20	890	100	3

[a]The number of lanes in the links of the highway refers to the maximum number of lanes physically allowable and currently operating in reality.

an hourly $\text{VOTT}_{\text{II}} = 1.5$ €, representing discretionary trips, and refers to the 20% of the traveler population. The minimum and maximum toll levels are set equal to $p_{\min} = 0$ € and $p_{\max} = 7$ €. The optimal toll and capacity choices are examined for the private highway links ($21 \leftrightarrow 18 \leftrightarrow 15 \leftrightarrow 12$) based on the demand pattern in a representative (design hour) travel period. In order to calculate the travel time t_a at link a, the well-known Bureau of Public Roads (BPR) function is used, as follows:

$$t_a\left(x_a\right) = t_a^0\left(1 + \mu\left(\frac{x_a}{G_a}\right)^{\beta}\right), \ \forall a \in A \tag{12}$$

where t_a^0 is the link travel time at free-flow conditions, μ and β are parameters referring to local operating conditions (in this study, $\mu = 0.15$ and $\beta = 4$) and G_a is the maximum traffic capacity at link a.

The feasible path choice set K^{rs} is based on a k-shortest/efficient paths selection (here $k = 3$, except for the highway connections between nodes 21 and 12 where $k = 1$). The selection of $k = 3$, given the specific network structure, helps avoid loops or other problems in estimating a feasible and efficient path choice set. For the general case of selecting a representative path choice set, various algorithms have been proposed in the literature [26].

5.2 Results of the pricing strategies with fixed capacity

In the first set of the experiments, the highway configuration is regarded as predetermined, with three lanes per direction (see Table 3), and only pricing decisions are made, ignoring service requirements. Namely, only the first component of objective function (1) concerning the toll revenues is considered, while the network design-related constraints in Eqs. 2 and 4 and level-of-service constraint in Eq. 5 are omitted. As mentioned in Section 3, three alternative pricing strategies are considered, i.e. those of a uniform toll charge and of spatially differentiated toll charges based on entry selection and entry–exit selection.

Table 4 indicates the increased revenues obtained from the entry-based toll pricing (by 13.2%) and entry–exit-based toll pricing (by 20.5%), in comparison to the revenues obtained from the uniform toll pricing. This is because flat tolls are independent of the spatial structure of congestion conditions in the highway and charge the same amount for both short and long-distance trips. This tactic discourages a significant portion of users (especially those with an assumed lower VOTT) to use the highway, resulting in diversion towards the free arterial routes. The diversion behavior entails reduction of the total demand for highway travel and toll revenues. Therefore, the spatially differentiated toll pricing strategies and, especially, the entry–exit-based user charging mechanism, are more profitable than the flat toll pricing.

In addition, the spatially differentiated pricing schemes are found to lead to a considerable reduction of the total network travel cost in monetary units, that is, 4.6% for the entry-based toll pricing and 5.5% for the entry–exit-based toll pricing, compared to the uniform toll pricing (see Table 4). This outcome can be attributed to the fact that the spatially differentiated toll strategies, particularly the entry–exit-based charging mechanism, allow to the ‘leader’ of the game (the operator) a greater flexibility to identify and suitably

Table 4 Results of the alternative highway pricing strategies with fixed capacity

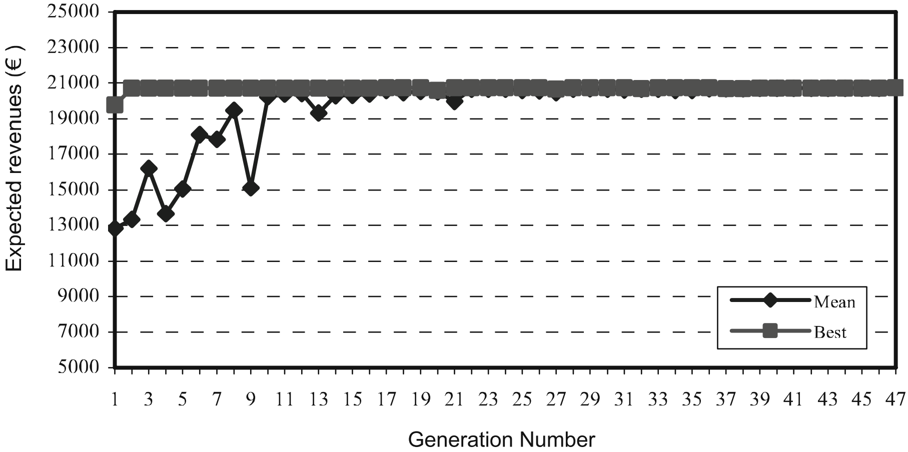
Problem	Expected revenues (€)	Total travel cost (€)	Social cost (€)
Flat toll pricing	20,719	37,508	16,789
Entry toll pricing	23,447	35,788	12,341
Entry–exit toll pricing	24,964	35,455	10,491

adjust prices for those critical entry points (or entry–exit pairs respectively) for which the users have increased willingness-to-pay. In this way, the operator obtains an increased ability to control the highway access and prevailing traffic conditions, thus ensuring a more efficient utilization of the highway capacity, while delivering a larger benefit to the users of the free links of the urban network.

An appropriate measure of social cost is also used, which equals the total travel cost minus the expected net revenues (which cancel out the toll revenues with the capacity investments). Although here the operator is a private firm, such a measure allows determining the welfare implications of different road pricing (and design) strategies, in terms of the total cost to the society in satisfying the given overall travel demand. Table 4 signifies the important reduction of the social cost, which reflects the gain in the net social benefit, resulting from the imposition of entry-based toll pricing (by 26.5%) and entry–exit-based toll pricing (by 37.5%), compared to the social cost resulting from the imposition of uniform toll pricing.

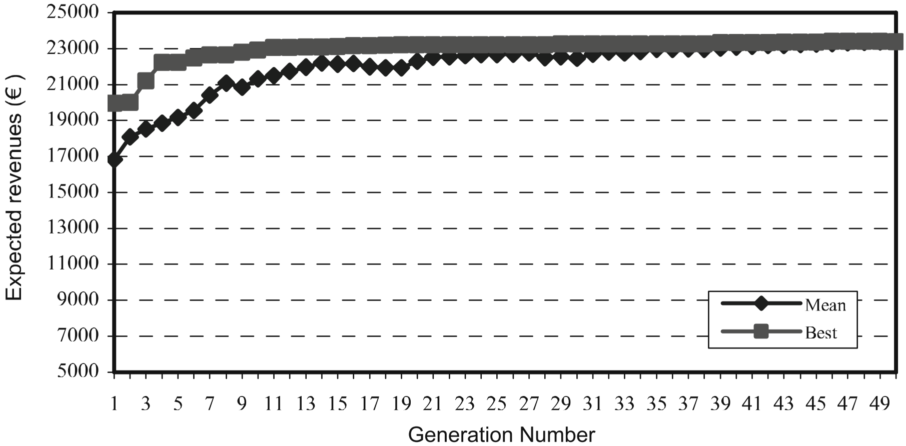
It should be noted that the current setup addresses the problem of the optimal design and pricing of a new road infrastructure, which unavoidably adds to social welfare, compared with the case of no additional infrastructure (even if its use is expensive), since it provides new travel options. In addition to the traffic conditions and toll price, multiple other (welfare-improving) criteria could be used in the set of constraints, in the form of additional regulations, to control the system operation from the state's point of view. On the contrary, when the problem setup concerns the management of existing road infrastructure, the profit maximization objective does not necessarily add to social welfare. For instance, private road operators tend to impose high toll rates which, in turn, reduce travel demand, yielding lower level of social welfare.

From the computational point of view, the convergence of the flat toll pricing is evidently much faster than that of the entry-based and entry–exit-based toll pricing. This is because flat toll pricing constitutes a univariate optimization problem, constrained to a narrow 0–7 € decision set, whose near-optimal solution can be approximated with an increased probability by a population of 50 individuals within the three to four first generations (see Fig. 2a), while the rest generations are used to fine tune it. In contrast, the entry-based and entry–exit-based toll pricing strategies constitute multivariate optimization problems requiring an increased computational effort with an average of 20 generations to converge (see Fig. 2b, c). The CPU time required in a standard single-core PC for each GA individual is 25 s, which yields about 20 min for each generation. As it is shown in the figures, the paths of the optimal (best) solution and mean GA performance basically coincide after about 30 generations, which implies that further genetic evolution is unlikely to produce an improved solution and, hence, the selection of 50 generations can be regarded as sufficient for the algorithm convergence.



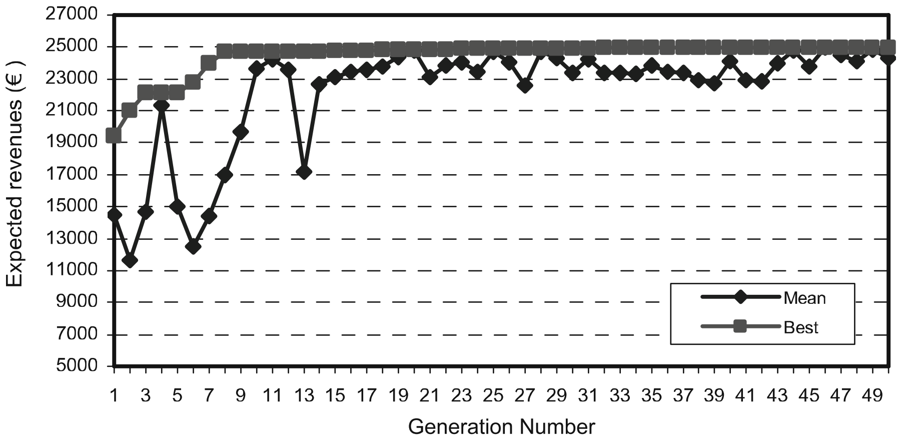
Expected revenues (€)
25000
23000
21000
19000
17000
15000
13000
11000
9000
7000
5000
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31 33 35 37 39 41 43 45 47
Generation Number
Mean
Best

(a)



Expected revenues (€)
25000
23000
21000
19000
17000
15000
13000
11000
9000
7000
5000
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31 33 35 37 39 41 43 45 47 49
Generation Number
Mean
Best

(b)



Expected revenues (€)
27000
25000
23000
21000
19000
17000
15000
13000
11000
9000
7000
5000
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31 33 35 37 39 41 43 45 47 49
Generation Number
Mean
Best

(c)

◀ **Fig. 2** Convergence diagram of the highway pricing problem with fixed capacity when charging: **a** flat tolls, **b** entry–exit-based tolls and **c** entry-exit-based tolls

5.3 Results of the pricing and design strategies

In the following set of experiments, the problem extends to the joint optimization of the (spatially differentiated) toll pricing and network design, in terms of the number of highway lanes. Such an optimization approach allows examining the strategic pricing and investment behavior of the system designer/operator in the long run. This extended problem yields a mixed-integer bilevel program, as that formulated in Section 3. The estimation of the construction and maintenance costs normalized for the peak hour and per lane and kilometer is made with the following equation:

$$E^k = \frac{K^k}{L_m}\left(\frac{r_0\,(1+r_0)^n}{(1+r_0)^n - 1}\right), \tag{13}$$

where E^k is the normalized peak hour construction and maintenance cost per kilometer, K^k is the total construction and maintenance cost for the whole period, $L_m = 365 \times n \times 24 \times p^h$ are the daily hours of operation normalized in peak hours by the factor p^h and r_0 is the interest rate for the payback period n (here equal to 10 years). The construction (plus land acquisition) and maintenance costs amount to 50 M€/Km. Hence, the monetary expenditures for each highway link are calculated as $V_a = h_a \times \ell \times E^k$, with h_a be the length of link $a \in \widehat{A}$.

The first JDNDPP setup, with entry-based tolls and entry–exit-based tolls, ignores constraint (5) on service requirements. Then, the JDNDPP considers a government regulation on the level of service, by requiring a flow-to-capacity ratio lower than 90% ($L = 0.90$) for each highway link. Table 5 shows that the expected net revenue obtained from the joint network design and differentiated toll pricing is dropped, in comparison to that obtained from applying differentiated tolls with fixed capacity in the actual network situation. The reduced net revenue is due to the inclusion of investment costs in the objective function (1). Also, the total travel cost is increased since the optimal number of highway lanes found in the experiments is less than that *a priori*

Table 5 Results of the alternative highway pricing and design strategies without and with service requirements

Problem	Expected net revenues (€)	Total travel cost (€)	Social cost (€)
Entry toll and network design	22,781	36,001	13,220
Entry toll and network design with service requirements	22,504	35,905	13,401
Entry–exit toll and network design	23,061	36,361	13,300
Entry–exit toll and network design with service requirements	22,988	35,867	12,879

considered in solving the standalone problem of spatially differentiated toll pricing (see Section 5.2), which implies a decrease in the network performance.

Similar to the case of differentiated toll pricing with fixed capacity, the results indicate that the joint network design and pricing with entry–exit-based tolls leads to increased revenues, in comparison to the joint network design and pricing with entry-based tolls (see Table 5). The increased net revenues are obtained for both the cases of omitting and incorporating service requirements. In contrast with the case of including service regulation, the imposition of entry–exit-based tolls in the joint pricing and design problem attracts such an

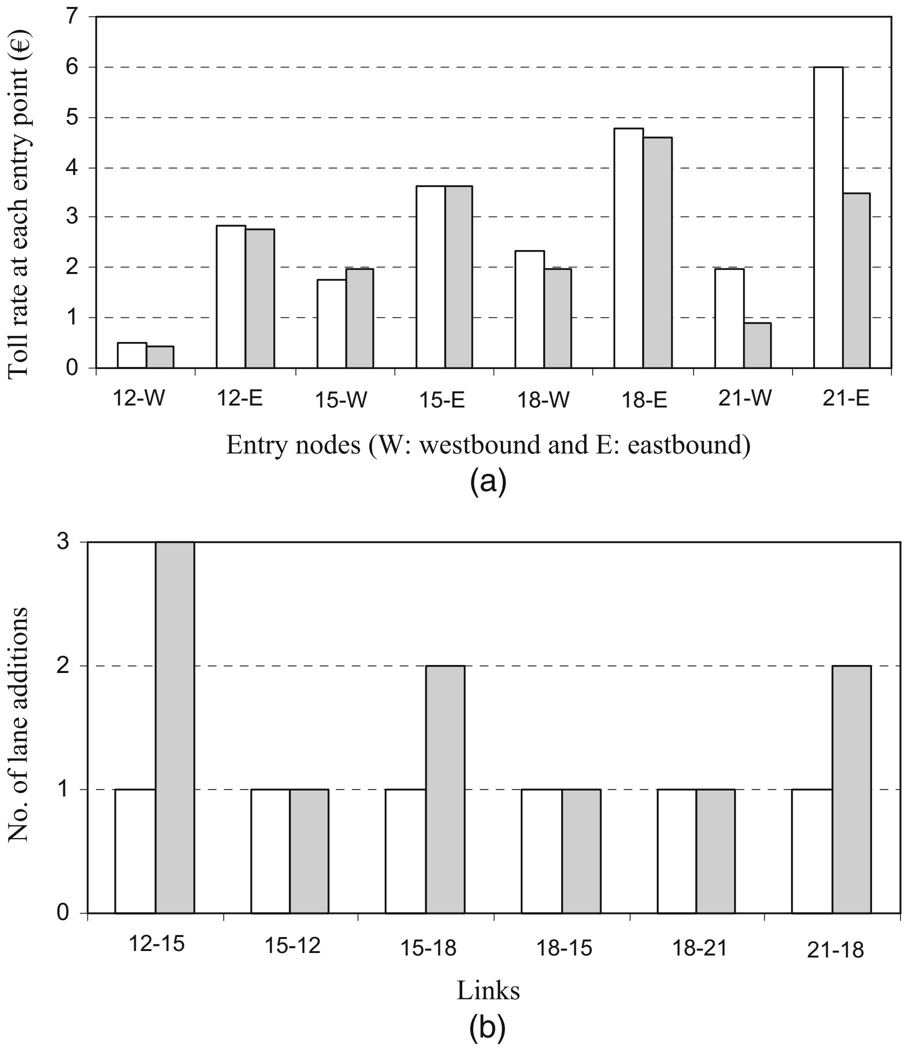


Fig. 3 Results of the entry-based toll pricing and design strategy: **a** toll rates and **b** lane additions, *without* service requirements (*white columns*) and *with* service requirements (*gray columns*)

added amount of traffic that induces higher total travel cost, compared to the imposition of entry-based tolls, when omitting service regulation. The service regulation entails additional capacity provision and/or higher toll rates than those required for revenue maximization. Therefore, the regulation is found here to decrease the expected net revenues, especially for the case of entry-based tolls. On the other hand, the service requirements 'guide' the solution procedure towards a desired level of system performance so that drop the total network travel cost, especially when charging entry–exit-based tolls (see Table 5).

Moreover, the results demonstrate the ability of the joint network design and entry–exit-based toll pricing with service requirements to achieve the lowest social cost, in comparison to all the other alternative BOT schemes (see

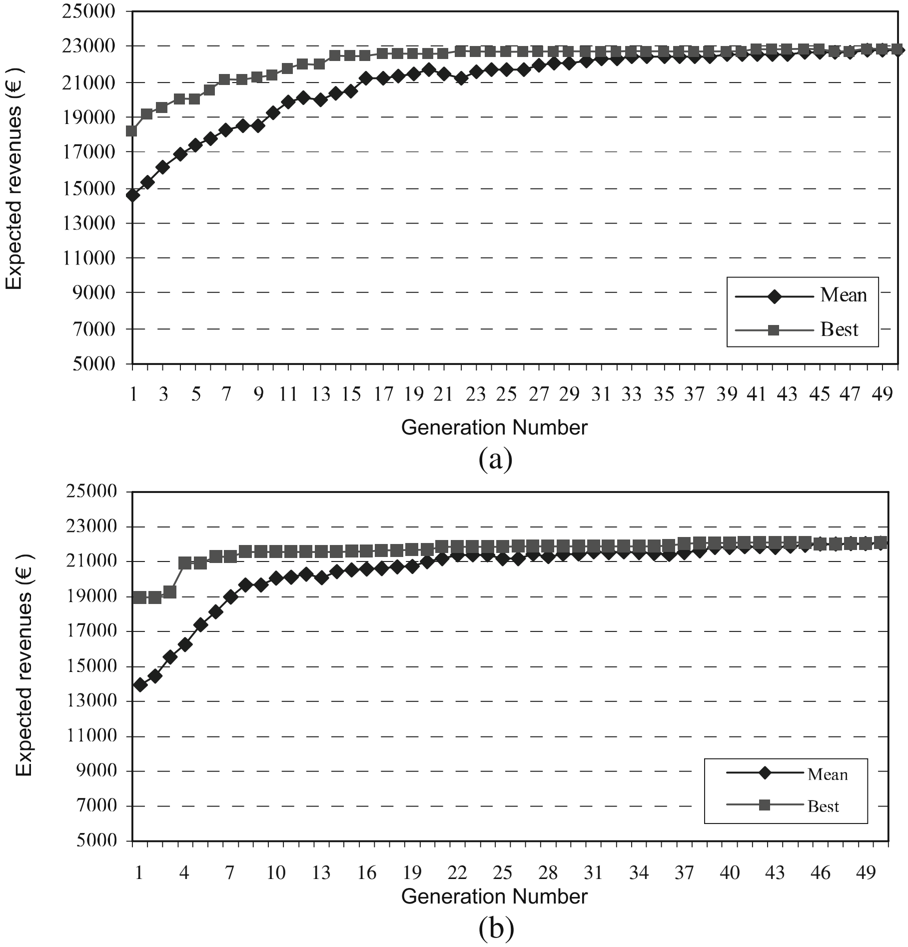


Fig. 4 Convergence diagram of the highway entry-based toll pricing and design problem: **a** without service requirements and **b** with service requirements

Table 5). This outcome signifies the importance of considering a suitable combination of design and pricing strategies with regulatory control to maximize the net social benefit, through simultaneously handling possible increase of total travel cost (due to added traffic induced by entry–exit-based tolls) and reduction of revenues (particularly due to the need for increasing capacity investments).

It should also be taken into account that regulations in service requirements may significantly change the strategy of the private firm on how to invest and charge users in the highway. Figure 3 indicates that, when no service requirements are imposed by the government/regulator, the number of lanes to be added are restricted to those being absolutely necessary (i.e. one lane per link) and the toll prices are set relatively high. On the other hand, when service requirements are introduced, then more lane additions are made in conjunction with reduced toll rates, so that allow the firm to recover investment cost through attracting more customers (users). It should be stressed here that the results depicted in Fig. 3 represent the output of two runs of the algorithm. Similar solutions, in terms of the designer/operator's profits, could be obtained from other algorithm runs producing different network configurations, i.e. combinations of lane additions and toll rates.

In terms of the computational cost, the joint network design and differentiated toll pricing problem requires about 20 generations for convergence, basically similar to the problem of entry-based and entry–exit-based toll pricing with fixed capacity. For demonstration purposes, Fig. 4 shows the convergence diagram of the highway entry-based toll pricing and design problem, while similar is the convergence pattern for the entry–exit-based toll pricing and design problem (not shown here), since both algorithms correspond to the same (bilevel mixed-integer programming) formulation, function objectives and constraints.

6 Summary and concluding remarks

The paper described the formulation and solution of alternative revenue-maximizing pricing and design strategies in a private highway which competes with free alternative arterials in a mixed-ownership, general road network. The study provides several important extensions to the current treatment of the joint road investment-pricing problem, which relate to the role of the heterogeneity of prices across access and entry–exit points, and of user groups, dependent upon their socio-economic status and trip purpose. Also, the problem setup takes into account the discrete nature of the capacity variable as well as variations in the perceived generalized travel cost and route choice uncertainty.

A game-theoretic formulation is provided which yields a nonconvex bilevel nonlinear program, incorporating a number of (budget, physical, price and service) constraints. The complexity of the solution procedure is addressed with the use of an evolutionary optimization technique. The implementation corresponds to a suitably selected, realistic road network including a privately operated highway, in order to address practical issues related to infrastructure investment planning and pricing policies in urban areas.

In the case of highway pricing with fixed capacity, the results show the benefits of imposing spatially differentiated toll pricing strategies, especially entry–exit-based tolls, in terms of enhancing the net social benefit, increasing revenues and reducing total network travel cost, compared to the flat tolls. In the case of joint highway pricing and capacity provision, the entry–exit-based tolls are found to yield the largest operator's revenues, while service regulation can eliminate the increase of total travel cost, in comparison to that induced by the entry-based tolls. An appropriate combination of highway design and (entry–exit-based) tolling scheme with service requirements can induce the lowest social cost. Moreover, the adverse effect of service regulation on revenues could be possibly addressed through the joint consideration of advanced highway and network-wide traffic management measures, such as information provision, ramp metering and integrated traffic signal control, in order to further enhance the public acceptability of BOT schemes.

Finally, the current problem setup could be extended to allow for two or more competing firms operate multiple toll roads. The study of the (self-) organization of this competitive highway market should take into account profit and investment interrelations of road providers as well as demand interdependences among all network participants. The conditions of profitability and welfare gains would then be affected by strategic interactions between firms involving conflict, co-operation and co-ordination, the role of regulation and other technical and organizational aspects related to such circumstances. Another possible extension might consider the partitioning of the travel period of the study into a number of intervals to allow representing the dynamic network loading conditions and, hence, determining optimal time-varying toll rates for the highway entries or entry–exit pairs.

Acknowledgement The authors express their gratitude to anonymous reviewers for their fruitful comments and suggestions.

References

1. Kim, B. J., & Kim, W. (2006). An equilibrium network design model with a social cost function for multimodal networks. *Annals of Regional Science, 40*(3), 473–491.
2. Dimitriou, L., Tsekeris, T., & Stathopoulos, A. (2007). Evolutionary combinatorial programming for discrete road network design with reliability requirements. *Lecture Notes in Computer Science, 4448*, 678–687.
3. Yang, H., & Zhang, X. (2002). Multi-class network toll design problem with social and spatial equity constraints. *Journal of Transportation Engineering, 128*(5), 420–428.
4. Yang, H., Tang, W. H., Cheung, W. M., & Meng, Q. (2002). Profitability and welfare gain of private toll roads in a network with heterogeneous users. *Transportation Research Part A, 36*(6), 537–554.
5. Chen, M., Bernstein, D. H., & Spasovic, L. N. (2004). Toll-design problem with stochastic route choice. *Environment and Planning Design, 31*(5), 731–742.

6. Shepherd, S. P., & Sumalee, A. (2004). A genetic algorithm based approach to optimal toll level and location problem. *Networks and Spatial Economics, 4*(2), 161–179.
7. Yang, H., Zhang, X., & Meng, Q. (2004). Modeling private highways in networks with entry–exit based toll charges. *Transportation Research Part B, 38*(3), 191–213.
8. Keeler, T. E., & Small, K. A. (1977). Optimal peak-load pricing, investment and service levels on urban expressways. *Journal of Political Economy, 85*(1), 1–25.
9. Mills, G. (1995). Welfare and revenue divergence for a tolled link in a road network. *Journal of Transport Economics and Policy, 29*(2), 137–146.
10. Viton, P. A. (1995). Private roads. *Journal of Urban Economics, 37*(3), 260–289.
11. Verhoef, E. T., & Rouwendal, J. (2004). Pricing, capacity choice, and financing in transportation networks. *Journal of Regional Science, 44*(3), 405–435.
12. Yang, H., & Meng, Q. (2000). Highway pricing and capacity choice in a road network under a build–operate–transfer scheme. *Transportation Research Part A, 34*(3), 207–222.
13. Chen, A., Subprasom, K., & Ji, Z. (2006). A simulation-based multi-objective genetic algorithm (SMOGA) for BOT network design problem. *Optimization and Engineering, 7*(3), 225–247.
14. Chen, A., & Subprasom, K. (2007). Analysis of regulation and policy of private toll roads in a build-operate-transfer scheme under demand uncertainty. *Transportation Research Part A, 41*(6), 537–558.
15. Dimitriou, L., Tsekeris, T., & Stathopoulos, A. (2008). Genetic computation of road network design and pricing Stackelberg games with multi-class users. *Lecture Notes in Computer Science, 4974*, 670–679.
16. Ben-Ayed, O., Boyce, D. E., & Blair, C. E. (1998). A general bi-level programming formulation of the network design problem. *Transportation Research Part B, 22*(4), 311–318.
17. Yin, Y. (2000). Genetic-algorithms-based approach for bilevel programming models. *Journal of Transportation Engineering, 126*(2), 115–120.
18. Zhao, Y., & Kockelman, K. M. (2006). On-line marginal-cost pricing across networks: Incorporating heterogeneous users and stochastic equilibria. *Transportation Research Part B, 40*(5), 424–435.
19. Cantarella, G. E., & Binetti, M. (1998). Stochastic equilibrium traffic assignment with value of time distributed among users. *International Transactions in Operational Research, 5*(6), 541–553.
20. Daganzo, C. F., & Sheffi, Y. (1977). On stochastic models of traffic assignment. *Transportation Science, 11*(3), 253–274.
21. Yang, H., & Bell, M. G. H. (1997). Traffic restraint, road pricing and network equilibrium. *Transportation Research Part B, 31*(4), 303–314.
22. Colson, B., Marcotte, P., & Savard, G. (2005). Bilevel programming: a survey. 4OR. *Quarterly Journal of Operations Research, 3*(2), 87–107.
23. Poorzahedy, H., & Rouhani, O. M. (2007). Hybrid meta-heuristic algorithms for solving network design problem. *European Journal of Operational Research, 182*(2), 578–596.
24. Goldberg, D. E. (1989). *Genetic algorithms in search, optimization and machine learning*. Reading, MA: Addison-Wesley.
25. Osyczka, A., & Kundu, S. (1995). A new method to solve generalized multicriteria optimization problems using the simple genetic algorithm. *Structural and Multidisciplinary Optimization, 10*(2), 94–99.
26. Cascetta, E., Russo, F., Viola, F., & Vitetta, A. (2002). A model of route perception in urban road networks. *Transportation Research Part B, 36*(7), 577–592.