



HAL
open science

Road network pricing and design for ordinary and hazmat vehicles: Integrated model and specialized local search

Francisco López-Ramos, Stefano Nasini, Armando Guarnaschelli

► **To cite this version:**

Francisco López-Ramos, Stefano Nasini, Armando Guarnaschelli. Road network pricing and design for ordinary and hazmat vehicles: Integrated model and specialized local search. *Computers and Operations Research*, 2019, 109, pp.170-187. 10.1016/j.cor.2019.05.006 . hal-02510066

HAL Id: hal-02510066

<https://hal.science/hal-02510066>

Submitted on 22 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Road Network Pricing and Design for Ordinary and Hazmat Vehicles: Integrated Model and Specialized Local Search

Francisco López-Ramos^a, Stefano Nasini^{b,*}, Armando Guarnaschelli^c

^aTechnological Institute of Celaya, Mexican National Council for Science and Technology, Guanajuato, México

^bDept. of Economics and Quantitative Methods, IÉSEG School of Management (LEM - CNRS UMR 9221), Lille, France

^cSchool of Transportation Engineering, Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile

Abstract

In the context of vehicle transportation in congested roads, we propose an optimization framework to integrate the operator decisions on network pricing, regulation, and expansion, while accounting for the shipments of hazardous materials. Current research trends only provide partial modeling integrations of the well-known toll optimization, hazmat transportation, and network design problems. However, the growing complexity of traffic management requires a stronger coordination in the operator decisions. In this paper, a mixed-integer non-linear bi-level problem is introduced to model this integration. The model considers a road network operator (acting as a *leader*), who maximizes its profit –the toll income minus the costs from roads construction and risk exposure to hazmat transportation–, and vehicles (acting as a *follower*), who minimize their travel costs –due to traffic congestion and toll charges. We introduce a reformulation approach that approximates this complex integrated problem with arbitrary precision and apply a specialized local search to exploit the structure of such reformulation. This combined resolution strategy relies upon a binary-search-based procedure, which sequentially updates the road prices intervals in such a way that the operator profit is monotonically improved. The effectiveness of the proposed approach is shown on a variety of structural configurations and economic settings, involving 1620 instances tested on the well-known Sioux Falls road network.

Keywords: Toll setting, Network design, Hazmat traffic regulation, Bi-level optimization, Specialized local search.

1. Introduction

The complexity of the current traffic management is progressively intensifying along with the rapid population growth on congested roads. As a consequence, the decision-making processes of road network operators are being encumbered by an increasing number of simultaneous difficulties, such as traffic congestion and the exposure of the population to risk of hazardous materials.

Contextually, investments on new roads and toll charging have been widely adopted strategies to alleviate these problems. In the operations research literature, these strategies have been modeled as the *Network Design Problem* (NDP, from now on) and the *Toll Optimization Problem* (TOP, from now on). While the NDP refers to ordinary vehicles, when hazmat vehicles are studied, instead, the name *Hazmat Transportation Problem* (HTP, from now on) is adopted. The HTP is often regarded as a separate class of problems, whose mathematical programming formulations generally mirror different forms of regulatory policies on shipments with hazardous materials, aiming at the minimization the total network risk.

The need for the integration of traffic management policies has been reflected in recent research trends that consider unified modeling approaches of either NDP and HTP [10] or TOP and HTP [33, 40, 18] or the NDP and TOP [45, 44]. However, these trends only provide partial integrations, while a unified formulation to serve as a modeling framework for a comprehensive range of operator decisions is still missing.

This work presents a Mixed-Integer Non-linear Bi-level Problem (MINLBP, from now on) to simultaneously determine the Road Network Pricing, Regulation, and Expansion (RNPRE, from now on) for ordinary and hazmat vehicles under traffic congestion. On the one hand, the network operator (acting as a *leader*) manages: 1) the investment on new road links, 2) the tolling of certain road links, and 3) the regulation of the access of hazmat vehicles to specific road links. These three interventions are done in such a way that the vehicle flows

*Corresponding author: s.nasini@ieseg.fr

are profitable and reduce the population risk to hazmat exposure. On the other hand, vehicles (acting as a *follower*) want to minimize their total travel costs from toll charges and travel time.

As noted by Dempe & Dutta [12], this class of models can be seen as non-convex programs with implicitly defined feasible regions. Candler & Townsley [9] has shown that the inducible region (i.e. the joint-solution space of the leader and follower agents) of a MINLBP is non-convex and Single-Level Reformulations (SLRs, from now on) are required to construct their algebraical characterization. In our context, these reformulations replace the vehicles best response with a set of constraints characterizing their optimality conditions [32, 46, 19, 29, 4].

Building on the SLR of this novel integration of NDP, TOP and HTP, we propose an approximation procedure that allows for the tractability of the two sources of non-linearity. The first source is the convex cost function at the follower level, driven by the usage of the well-known Bureau of Public Road function [34] (BPR, from now on) to model traffic congestion. The second source of non-linearity results from the bi-linear terms associated to the follower’s optimality conditions at the SLR.

Therefore, a Mixed-Integer Linear-Programming (MILP, from now on) is built in four steps:

- The follower’s problem is formulated using the Lin & Leong theory [31] to model the user equilibrium under congestion for multiple vehicle categories. This allows the leader to interdict the access of certain paths to certain categories of vehicles (see Appendix A).
- A piecewise linear approximation of the congestion function [15] is applied to the follower’s problem, where the approximation lines are obtained using the Imamoto & Tang method [26] (see Appendix B).
- A SLR is then constructed building on the strong duality theorem to characterize the optimality of the linearized follower’s problem. This contains a collection of bi-linear terms resulting from the products between toll and flow variables.
- Finally, toll variables are replaced with binary indicators denoting a discrete collection of toll levels [29, 33, 44], while the new bi-linear terms involving the product between discretized tolls and vehicle flows are linearized using the method suggested by Fourer [22].

The application of this four-step procedure results in a computationally challenging MILP that approximates the proposed RNPRES problem with arbitrary precision, by enlarging the number of terms in the linearization. A trade-off between problem size and solution quality is addressed by the usage of a specialized local search algorithm, which exploits the proposed linearization approach. This relies upon a binary-search-based procedure, which sequentially updates the road prices intervals in such a way that the operator profit is monotonically improved.

On the experimental side, a computational test involving 1620 instances on three extended variants of the Sioux Falls road network (each having different amount of new toll links) have been built for assessing the solution quality and computational performance under different scenarios. The results reveal the appropriateness of the proposed reformulation and linearization approach, under different structural configurations and economic settings (i.e the investment budget, the monetary value of time, the monetary value of risk, the number of lines approximating the congestion function, and the number of toll levels). We observe that the usage of the specialized local search algorithm improves the obtained MILP solution for any fixed number of linearized terms and attain better toll values in the vast majority of analyzed instances.

The rest of the paper is organized as follows. Section 2 surveys the relevant studies on TOP, NDP, and HTP problems, highlighting the contributions of the proposed approach. Section 3 presents the mathematical formulation of the bi-level optimization model. Section 4 develops a mixed-integer linear problem to approximate the bi-level model. Section 5 shows the solution approach to improve the initial tolls found by the MILP in a wise manner. Section 6 describes the networks used in the computational tests. Section 7 describes the computational experiments and shows their results. The paper finishes with some conclusions and directions for further research. In Appendix A and B, supplementary materials are provided to integrate the exposition and support the theoretical and empirical results. All mathematical proofs of propositions have been reported in Appendix C.

2. Related Works

This section revises the main streams of contributions corresponding to the TOP, NDP and HTP classes of models and the different partial integrations of these models that have been so far proposed in the operations research literature.

Toll Optimization Problem. In a seminal paper, Labbé et al.[29] describe a bi-level program where the leader’s objective maximizes the toll income in a road network, while the follower’s best response is a traffic flow minimizing the costs from toll charging and travel time.¹ The authors propose a SLR by replacing the lower level problem with its dual and then linearize the resulting primal-dual relationship constraint, giving rise to a MILP. With a view to extending the TOP to more realistic scenarios, Elström et al.[15] include operational costs on toll links as a function of vehicle flows, and transform the bi-level problem into a single-level problem by a set of VI-constraints.² Recently, Kamgaing-Kuiteing et al.[27] compare arc-flow and path-flow formulations of the TOP with fixed and elastic demand cases and perform a sensitivity analysis on their computational times with respect to demand elasticity and percentage of toll links.

Network Design Problem. The NDP is first modeled as a bi-level program by Gao et al.[23], who define the leader’s choices as the determination of new road links to minimize the system traveling costs; whereas the follower (representing the vehicles) is design to characterize the user equilibrium. Fontaine & Minner [19] reformulate the NDP as a mixed-integer linear problem using the same approach as Labbé et al. [29], and propose an improved Benders decomposition in which an acceleration technique is devised for solving Benders subproblems efficiently. In a recent paper, Fontaine & Minner [20] apply this decomposition method to an extended version of the NDP, where different vehicle flow patterns are considered in a time-varying fashion. Analogously, Bagloee et al.[2] study the interaction of two types of vehicles in the follower’s problem, and propose a hybrid algorithm that combines a generalized Benders decomposition with a Branch & Bound procedure. Finally, Wang & Zhang [39] extend the NDP framework by studying the interaction among multiple leaders operating in distinct regions of the road network. This interaction is analyzed using three bi-level programming models, each one representing different behaviors of the leaders.

Hazmat Transportation Problem. The HTP is first formulated as a bi-level problem by Kara & Verter [28], who define the leader’s objective as the minimization of the risk exposure to hazmat, subject to the follower’s best response, representing trucks attempting to minimize the transportation costs. Trucks are grouped into categories, and one route is determined for each category. Few years later, Erkut & Alp [16] introduce a path-addition heuristic to solve the HTP, while Erkut & Gzara [17] analyze the stability of the problem solution using this methodology. The authors show that more stable solutions are obtained when the leader’s goal considers both risk exposure and transportation costs. Bianco et al. [5] use Erkut & Gzara’s methodology to solve a bi-level problem in which the leader (representing a regional authority) minimizes the risk induced over the entire network; whereas, the followers (representing a local authority) minimize the risk over their local jurisdictions. The authors also prove the stability of the solution. Fontaine & Minner [21] apply a tailored Benders Decomposition to solve this problem to optimality for moderate-sized networks. Finally, Sun et al. [37] introduce uncertainty into the risk in order to design a robust HTP. Erkut & Gzara’s methodology has been used in combination with a Lagrangian relaxation heuristic in order to solve large-scale instances from real road networks.

Partial integrations. The literature about the modeling integration of TOP, NDP and HTP can be grouped into three categories: i) works dealing with the NDP and HTP, ii) works facing to the TOP and HTP, and iii) works studying the NDP and TOP.

In the first category, Chiou [10] presents a min-max bi-level model to effectively regulate risk associated with hazmat transportation while minimizing travel cost for all vehicles under demand uncertainty. In this model, a risk-averse signal setting policy is proposed against the worst-case realization of uncertainty in travel demand. A min-max single-level model and a bundle-like solution method are devised to solve the model.

Regarding the second category, Marcotte et al. [33] prove that toll setting is a good policy to regulate the use of roads for dangerous good shipments. The authors extend the model of Kara & Verter [28], incorporating tolls as decision variables of the leader. Wang et al. [40] incorporate the regulation of ordinary vehicles to evaluate the hazmat exposure through a duration-population-frequency risk function. Finally, Esfandeh et al.[18] extend the latter model by incorporating toll-free links into the leader problem, where a combination of risk and toll charging is minimized. An equilibrium-decomposed optimization heuristic is introduced to solve the resulting bi-level problem.

Finally, in the third category, Yang et al. [45] study the competition among private operators when tolling links in the same road network. In that problem, operators are designed as to maximize their own profits

¹In that original model there are already upper and/or lower bounds for tolls, and some links are toll-free.

²Elström et al. [15] also provide a lower bound of this reformulated problem by solving an approximated MILP.

	Studied Problems			Construction Costs	Traffic Congestion
	TOP	NDP	HTP		
Chiou [10]		✓	✓		
Marcotte et al. [33]	✓		✓		
Wang et al. [40]	✓		✓		
Esfandeh et al. [18]	✓		✓		✓
Yang et al. [45]	✓	✓		✓	✓
Xu et al. [44]	✓	✓			
This work	✓	✓	✓	✓	✓

Table 1: Summary of the main features incorporated in the works integrating some of the studied problems

evaluating their toll incomes and costs of constructing new road links for attracting vehicles. Two heuristics are proposed for solving the problem, and their convergence is demonstrated. A single operator version of that problem is studied by Xu et al. [44], by proposing a mixed-integer quadratic program in which complementary slackness conditions are relaxed.

Table 1 summarizes the main aspects related to the works integrating some of the studied problems. One can observe that none of these works integrate the TOP, NDP and HTP problems. Additionally, the features related to the evaluation of the construction costs of the leader and the traffic congestion have only been considered by Yang et al.[45], and Esfandeh et al.[18] and Yang et al.[45], respectively. Finally, the traffic regulation feature has been never studied before in the context of TOP, NDP or HTP problems. However, there exists an extended user equilibrium model in [31] demonstrating that users can reach an equilibrium even when an authority prevents them from using some links.

In this work, when dealing with bi-linear terms involving tolls and vehicle flows, a specialized local search algorithm is introduced to exploits the proposed linearization approach, based on a binary-search-based procedure. As shown in Section 7, this procedure allows improving the toll levels found by the approximated MILP with a limited computational effort.

3. Problem formulation

3.1. Assumptions

Consider a road network in which some vehicles carry hazardous materials, inducing risk exposure to people living near the road. To alleviate the entailed social cost, while maximizing profit, the network operator may take three types of decisions. One is to charge tolls on the different available paths, while not exceeding a maximum affordable charge. Another is to build new toll roads (subject to an investment budget). Finally, the network operator can restrain the access to certain types of vehicles (i.e., ordinary or hazmat vehicles) in certain roads.

As a response to the operator’s decisions, vehicles decide their routes by minimizing the total travel costs, including time cost and toll charges. Thus, two decision makers interact: the network operator, who aims at maximizing its profit; and the moving vehicles, who want to minimize their travel costs throughout the network.

A list of modeling assumptions is made explicitly hereafter.

- The risk exposure to hazmat for traveling vehicles is ignored as it is considered that the number of affected vehicles is much less compared to the number of people living near the hazmat roads [28, 16, 17, 45, 5].
- Traveling vehicles cooperate with the network operator when costly-equivalent alternative travel routes are available for different operator decisions, so that the operator can choose the one that provides the highest benefit. This is in line with optimistic formulations for the TOP problem [30, 44].
- The vehicles’ travel costs are separable and strictly non-decreasing function of link flows, based on the BPR function [34], as widely applied to the transportation literature [23, 15, 19].
- For each origin-destination pair, the number of traveling vehicles is known, and there exists a set of road links that are neither regulated by the network operator nor under construction through which all vehicles can reach their destinations.

- The network operator is a profit maximizer, who internalizes the externality costs induced by the hazmat risk. This is an important departure from the business setting proposed by [19], where the operators goal was to minimize the total travel time in the network (i.e. system-optimum), acting as a *benevolent social planner*.
- The traffic flows are specified in regular vehicle units. This does not prevent from taking into consideration different vehicle sizes as there are procedures to measure them in PCUs (Passenger Car Units) as described in [1].

3.2. Notation

Before the mathematical model is presented, we introduce some notation. Throughout this paper, unless otherwise explained, vectors will be denoted in boldface letters; whereas, sets will be denoted in calligraphic letters. Latin letters will be used for the primal variables and parameters (as listed below), whereas Greek letters will be used for the dual variables associated with constraints of the primal follower problem.

Sets:

\mathcal{N}	set of nodes (with $ \mathcal{N} = n$);
\mathcal{A}	set of links between nodes (with $\mathcal{A} \subseteq \mathcal{N} \times \mathcal{N}$);
\mathcal{A}^T	set of toll links (with $\mathcal{A}^T \subset \mathcal{A}$);
\mathcal{A}^F	set of toll-free links (with $\mathcal{A}^F = \mathcal{A} \setminus \mathcal{A}^T$);
\mathcal{A}^N	set of future possible links (with $\mathcal{A}^N \subset \mathcal{A}^T$);
\mathcal{V}	ordered set of vehicle types (with $ \mathcal{V} = V$), classifying the levels of risk;
\mathcal{K}^v	set of origin-destination pairs for vehicles of type $v \in \mathcal{V}$ (with $\mathcal{K} \subseteq \mathcal{N} \times \mathcal{N}$);

Exogenous parameters:

c_a	cost of constructing a new link $a \in \mathcal{A}^N$;
$g^{v,k}$	number of vehicles of origin-destination pair $k \in \mathcal{K}^v$ traveling from origin $o(k) \in \mathcal{N}$ to destination $d(k) \in \mathcal{N}$ associated with level of risk $v \in \mathcal{V}$;
n_a	number of people exposed to hazmat risk on link $a \in \mathcal{A}$;
m_v	monetary cost for each individual exposed to level of risk $v \in \mathcal{V}$.

Operator's decision variables:

t_a	unitary toll at each link $a \in \mathcal{A}^T$;
w_a^v	if a link $a \in \mathcal{A}^T$ has been opened for vehicles of level of risk $v \in \mathcal{V}$;
y_a	if a new road has been constructed in the potential link location $a \in \mathcal{A}^N$.

Vehicles's decision variables:

$x_a^{v,k}$	flow of vehicles of level of risk $v \in \mathcal{V}$ related to origin-destination $k \in \mathcal{K}^v$ going through $a \in \mathcal{A}$.
-------------	---

To further clarify the definition of sets and parameters, it is worth mentioning that the set of future possible links \mathcal{A}^N constitutes both a subset of the total set of links \mathcal{A} (i.e., $\mathcal{A}^N \subset \mathcal{A}$) and a subset of the set of toll links (i.e., $\mathcal{A}^N \subset \mathcal{A}^T$). Therefore, \mathcal{A}^T represents the set of both existing and future possible links to which we can charge a positive price. Beside, it should be noted that the ordered set of vehicle types \mathcal{V} encodes the information about the different risk levels, being the risk-free type (ordinary vehicles) associated with the first element in \mathcal{V} . As a consequent, the monetary costs verify $m_0 \leq m_1 \leq \dots \leq m_V$.

Finally, to conclude the list of used notation we define $\mathcal{K} = \cup_{v \in \mathcal{V}} \mathcal{K}^v$ and introduce the vector forms $\mathbf{t} = [t_a : a \in \mathcal{A}^T] \in \mathbb{R}^{|\mathcal{A}^T|}$, $\mathbf{y} = [y_a : a \in \mathcal{A}^N] \in \{0, 1\}^{|\mathcal{A}^N|}$, $\mathbf{w} = [w_a^v : a \in \mathcal{A}^T, v \in \mathcal{V}] \in \{0, 1\}^{|\mathcal{A}^T| |\mathcal{V}|}$ and $\mathbf{x} = [x_a^{v,k} : a \in \mathcal{A}, v \in \mathcal{V}, k \in \mathcal{K}^v] \in \mathbb{R}^{|\mathcal{A}| |\mathcal{V}| |\mathcal{K}|}$; as well as the expressions $x_a^v = \sum_{k \in \mathcal{K}^v} x_a^{v,k}$, for all $v \in \mathcal{V}$, and $x_a = \sum_{v \in \mathcal{V}} x_a^v$, for all $a \in \mathcal{A}$.

3.3. Modelling Approach

In this section, we introduce the general modeling framework for the integrated RNPRES problem. We first define the operator's and vehicles' payoffs to capture their respective revenue and cost structures.

$$\begin{aligned} \text{Network operator's profit:} \quad P^O(\mathbf{t}, \mathbf{y}, \mathbf{w}, \mathbf{x}) &= \sum_{a \in \mathcal{A}^T} t_a x_a - \sum_{a \in \mathcal{A}^N} c_a y_a - \sum_{a \in \mathcal{A}} n_a \sum_{v \in \mathcal{V}} m_v x_a^v \\ \text{Vehicles transportation cost at link } a: \quad P_a^V(t_a, x_a) &= \chi_{\{a \in \mathcal{A}^T\}} t_a + u_a^0 \left(1 + r_a (x_a/q_a)^4\right) \end{aligned}$$

where $\chi_{\{a \in \mathcal{A}^T\}}$ is an indicator function, and u_a^0 , r_a and q_a are exogenous tuning parameters for the traveling time to be specified for each application.

The network operator's payoff is the total profit consisting of the revenue from toll value (t_a , for $a \in \mathcal{A}^T$) multiplied by vehicles flows (x_a , for $a \in \mathcal{A}^T$), minus the two sources of cost: (i) the cost associated with the construction of new links ($c_a y_a$, for $a \in \mathcal{A}^N$) and (ii) the cost related to risk exposure to hazmat transportation. This second cost is a social cost that a benevolent operator must internalize. This is computed as n_a (the size of the population in the neighborhood of link a), while m_v renders the market value of such risk.

The vehicles transportation cost is divided into toll charges established by the operator (first term) and the time cost defined as a convex function of the amount of flow (second term). The latter is approached by the well-known BPR function [34], where parameter u_a^0 represents the market value of the travel time under free-flow conditions (i.e. in absence of other vehicles), r_a attaches the relative importance of congestion to the travel time, and q_a defines the link capacity in terms of the number of vehicles going through.

The RNPRES problem can be defined as a leader-follower game on a multi-vehicle network $\mathcal{G} = (\mathcal{V}, \mathcal{K}, \mathcal{N}, \mathcal{A})$. Mathematically, this problem can be formulated as the following bi-level model:

$$\max_{\mathbf{t}, \mathbf{y}, \mathbf{w}, \mathbf{x}} P^O(\mathbf{t}, \mathbf{y}, \mathbf{w}, \mathbf{x}), \quad \text{subject to: } (\mathbf{t}, \mathbf{y}, \mathbf{w}) \in \Phi \quad \text{and} \quad \mathbf{x} \in \Psi(\mathbf{t}, \mathbf{w}) \quad (1)$$

where Φ is the set of feasible network operator choices for pricing, regulation and expansion, which are characterized by the following system of linear constraints:

$$\sum_{a \in \mathcal{A}^N} c_a y_a \leq B \quad (2a)$$

$$t_a \leq T_a, \quad \forall a \in \mathcal{A}^T \quad (2b)$$

$$w_a^{v-1} \geq w_a^v \quad \forall a \in \mathcal{A}^T \quad v = 2 \dots |\mathcal{V}| \quad (2c)$$

$$y_a \geq w_a^v, \quad \forall a \in \mathcal{A}^N, \quad v \in \mathcal{V} \quad (2d)$$

$$y_a \in \{0, 1\}, \quad \forall a \in \mathcal{A}^N \quad (2e)$$

$$w_a^v \in \{0, 1\}, \quad \forall a \in \mathcal{A}^T, \quad v \in \mathcal{V} \quad (2f)$$

$$t_a \geq 0, \quad \forall a \in \mathcal{A}^T \quad (2g)$$

Here, the constraint (2a) fixes a budgetary limit (B) to the operator investment in new road construction;³ (2b) set a maximum toll level in each road link (T_a) reflecting public regulation in urban tarries; (2c) implement a regulatory policy imposing only access restriction to vehicles with higher level of hazmat;⁴ (2d) prevent vehicles from accessing the roads that have not been built; (2e)–(2g) set the variable domains.

The vehicle best response $\Psi(\mathbf{t}, \mathbf{w})$ results from the route choice model [41], which assumes that users are in a simultaneous equilibrium (i.e. no driver can unilaterally reduce the travel cost by shifting to another route). As described by Sheffi [35], the equilibrium solution relies on an homogeneous population of users. However, a recent study by Lin & Leong [31] derives the equilibrium solution for the case when specified paths are not available for a certain type of road user. Building on [31], the mathematical programming formulation for the

³The budget constraint is frequently adopted in network design problems [2, 19, 20, 23, 39] to contain the range of investment possibilities within the liquidity accessibility (i.e. the network operator ability to access funding), as well as to mirror public regulations on investment plans.

⁴This is in line with the empirical setting of congested urban areas, where the network operators act conditionally on local and governmental policies.

user equilibrium in the presence of differentiated route channels for vehicles of different types (as induced by the operators's choices on $\mathbf{w} \in \{0, 1\}^{|\mathcal{A}^T| \times |\mathcal{V}|}$) can be defined as follows:

$$\Psi(\mathbf{t}, \mathbf{w}) = \underset{\mathbf{x}}{\operatorname{argmin}} \sum_{a \in \mathcal{A}} \int_0^{\sum_{v \in \mathcal{V}} \sum_{k \in \mathcal{K}^v} x_a^{v,k}} P_a^V(t_a, \omega) d\omega \quad (3a)$$

$$\text{subject to: } \sum_{a \in \mathcal{A}} N_{i,a} x_a^{v,k} = s_i^{v,k}, \quad \forall i \in \mathcal{N}, v \in \mathcal{V}, k \in \mathcal{K}^v \quad (3b)$$

$$x_a^v \leq M_{a,v}^{(w)} w_a^v, \quad \forall a \in \mathcal{A}^T, v \in \mathcal{V} \quad (3c)$$

$$x_a^{v,k} \geq 0, \quad \forall a \in \mathcal{A}, v \in \mathcal{V}, k \in \mathcal{K}^v \quad (3d)$$

The objective function (3a) is a purely mathematical construct ensuring that road users choose their routes selfishly. Constraints (3b) define the balance conditions using the link-node incidence matrix $N \in \{-1, 0, 1\}^{|\mathcal{N}| \times |\mathcal{A}|}$, where a -1/1 is set in the entries where a link goes out/comes in a node, and 0 if a link is not connected with a node. The right-hand-side of (3b) is defined as

$$s_i^{v,k} = \begin{cases} g^{v,k} & \text{if } i \text{ is the destination of } k \\ -g^{v,k} & \text{if } i \text{ is the origin of } k \\ 0 & \text{otherwise} \end{cases},$$

where $g^{v,k}$ denotes the number of vehicles associated with an origin-destination pair $k \in \mathcal{K}^v$. Next constraints (3c) are a linear equivalent version of the interdiction type of constraints introduced by [31]. These constraints prevent the circulation of vehicle flow for a given type, so that if for a certain road the operator does not give access to a certain type of user ($w_a^v = 0$), the flow resulting from that type of user is zero ($x_a^v = 0$), whereas if the regulatory policy of the network operator keeps that arc within choice set of that type of user ($w_a^v = 1$), this constraint is not bounding as long as $M_{a,v}^{(w)}$ is sufficiently large.⁵ Finally, constraints (3d) specify non-negativity on flow variables. A detailed derivation of the user equilibrium for problem (3a) is provided in Appendix A.

Let $x_a = \sum_{v \in \mathcal{V}} \sum_{k \in \mathcal{K}^v} x_a^{v,k}$. So, when integrating the vehicle's payoff from a link $a \in \mathcal{A}$ over each arc flow, the follower's objective function becomes:

$$\begin{aligned} \sum_{a \in \mathcal{A}} \int_0^{x_a} P_a^V(t_a, \omega) d\omega &= \sum_{a \in \mathcal{A}} \int_0^{x_a} t_a d\omega + u_a^0 \int_0^{x_a} \left(1 + r_a \left(\frac{\omega}{q_a}\right)^4\right) d\omega \\ &= \sum_{a \in \mathcal{A}^T} t_a x_a + \sum_{a \in \mathcal{A}} u_a^0 \left(x_a + \frac{r_a}{5(q_a)^4} (x_a)^5\right) \end{aligned} \quad (4)$$

Therefore, the resulting model for road network pricing, regulation and expansion embeds in an unified framework the complexity of current traffic management in congested urban area, allowing the leader to select over \mathbf{t} , \mathbf{y} and \mathbf{w} variables; and the follower, to respond to the leader choices based on its equilibrium conditions. While a comprehensive computational analysis of this problem is carried out in Section 7, to gain a conceptual understanding of the network operator profit, the following small example provides an illustrative case of a follower's problem that can be analytically solved.

Example 1. Consider a two-node network consisting of a single origin-destination pair and two travel links, a toll-free link (already constructed) and a toll link to be built with toll levels t . Two types of vehicles are circulating across the network: ordinary (risk-free) vehicles and hazmat vehicles. Again the binary indicator w^v is required to be equal to one if the link to be built is opened to vehicles of type (risk level) v and zero otherwise. Recall that x_a^v is used to refer to the flow of vehicles of type v circulating through link a . Assuming a unitary demand, the flow balance for vehicles of type v is $x_1^v + x_2^v = 1$, so that we can eliminate two variables by defining a new flow variable x^v , in such a way that $x_1^v = x^v$ and $x_2^v = 1 - x^v$. Using (4), we construct the follower problem for this toy example as:

⁵An adjusted value of this big-M is $\sum_{k \in \mathcal{K}_a^v} g^{v,k}$, where $\mathcal{K}_a^v \subset \mathcal{K}^v$ is the subset of origin-destination pairs related to vehicle type v , whose origin and destination are not the respective head and tail nodes of the link a .

$$\Psi(t, \mathbf{w}) = \begin{cases} \underset{x^1, x^2}{\operatorname{argmin}} t(2 - x^1 - x^2) + \frac{r_1}{5}(x^1 + x^2)^5 + \frac{r_2}{5}(2 - x^1 - x^2)^5 & \text{if } w^2 = 1 \\ \underset{x^1, x^2}{\operatorname{argmin}} t(1 - x^1) + \frac{r_1}{5}(x^1 + 1)^5 + \frac{r_2}{5}(1 - x^1)^5 & \text{if } w^1 = 1 \text{ and } w^2 = 0 \\ \frac{r_1}{5(q_1)^4} 2^5 & \text{if } y = 0 \end{cases}$$

To keep notation short, we have assumed that $u_a^0 = q_a = 1$, for each route. By letting $[\hat{x}^1(t, \mathbf{w}), \hat{x}^2(t, \mathbf{w})] \in \Psi(t, \mathbf{w})$ and assuming $m_1 = 0$ (as ordinary vehicles bring no risk), the operator's payoff can be expressed as a function of the toll level t under the three alternative accessibility decisions $(w^1, w^2) \in \{(1, 1), (1, 0), (0, 0)\}$:

$$P^O(\mathbf{t}, \mathbf{y}, \mathbf{w}, \mathbf{x}) = \begin{cases} t(2 - \hat{x}^1(t, [1, 1]) - \hat{x}^2(t, [1, 1])) - c - n_1 \hat{x}^1(t, [1, 1]) - n_2(1 - \hat{x}^1(t, [1, 1])) & \text{if } w^1 = w^2 = 1 \\ t(1 - \hat{x}^1(t, [1, 0])) - c - n_1 \hat{x}^1(t, [1, 0]) - n_2(1 - \hat{x}^1(t, [1, 0])) & \text{if } w^1 = 1 \text{ and } w^2 = 0 \\ -n_{a1} & \text{if } y = 0 \end{cases}$$

Fig. 1 illustrates operator's payoff for every value of t and for each construction decision, fixing $r_1 = 1$ and $r_2 = 100$ (i.e. the impact of congestion to the travel time is 100 times higher in the toll-free path compared to the paid path).

(a) Population size in the neighborhood of the toll-free road $n_1 = 100$.

(b) Population size in the neighborhood of the toll-free road $n_1 = 10$.

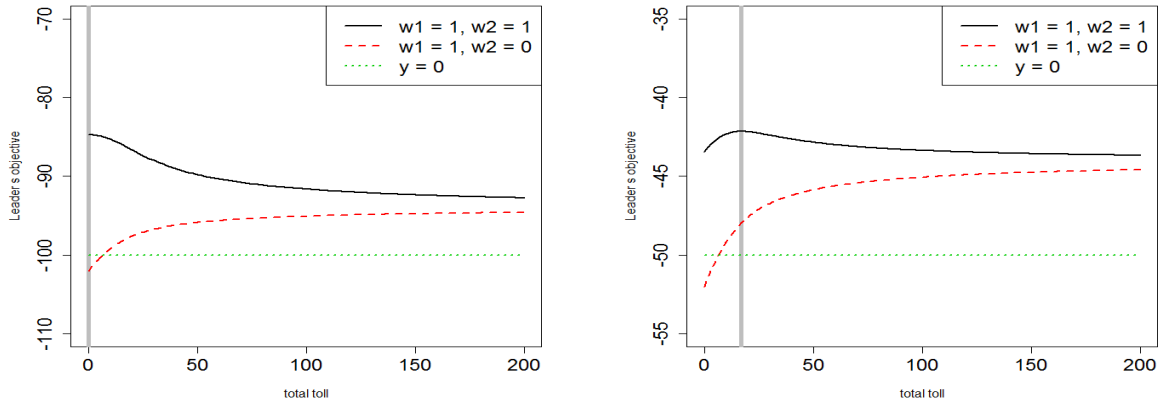


Figure 1: Network operator's payoff as a function of the toll level t , for different combinatorial decisions about the construction and accessibility of the new road. Two different parameterizations are considered, concerning the Population size in the neighborhood of the toll-free road: $n_1 = 10$ (left panel), $n_1 = 100$ (right panel).

In this stylized example, one can see that when the population size in the neighborhood of the toll-free road is equal to 100, the best operator choice is to build the new road, so that both categories of vehicles have access with 0 charge. By contrast, when the population size in the neighborhood of the toll-free road is equal to 10, the operator charges 16.8 monetary units for traveling through that new road.

Single level reformulations (SLRs) are one of the most used techniques for solving these class of problems (see, for instance, [42, 11, 12, 13, 24, 47] for a theoretical and practical background of this type of reformulations). The SLRs require to determine the optimality conditions of the lower level problem, and to append them to the upper level problem. As claimed by Dempe & Dutta [12] (see Theorem 2.1), an equivalent relationship exists between the (global) solution of the original MINLBP and the one of its single level reformulation, as long as (3) is convex and satisfies the Slater's constraint qualification.⁶

Building on this result, the next section studies the characterization of the optimal network operator choices with the usage of a single level reformulation for problem (1)-(3). Additionally, some linearization techniques are applied to bi-linear terms, appearing as a consequence of this reformulation, to obtain a mixed-integer linear programming problem.

⁶This is a sufficient condition for strong duality in convex problems (see Proposition 8.7 in [8]).

4. The RNPRES problem approximated by a mixed-integer linear problem

The RNPRES problem (1)-(3) is a MINLBP containing two sources of non-linearities: (i) the time cost defined as a convex function of the amount of flow, (ii) the bi-linear terms involving tolls and vehicle flows representing the leader's income. In the following subsections, a three-step procedure is proposed to build a MILP approximating the original MINLBP. This MILP is built replacing both non-linear terms, and the follower's objective function with systems of linear constraints in higher dimensions.

- In Subsection 4.1, the follower's problem is linearized by approaching the congestion term using a piecewise linear approximation.
- In Subsection 4.2, a SLR is applied to this linearized follower's problem by replacing the follower's objective function with a new set of constraints and variables representing the strong duality relationship between primal and dual forms in the optimum.
- In Subsection 4.3, toll variables are replaced with binary indicators denoting a discrete collection of possible toll levels to be charged to vehicles.

The goodness of the approximated MILP hinges on its size, so a trade-off between accuracy and efficiency will be computationally studied in Section 7.

4.1. Linear approximation of the follower's problem

The follower's objective (4) is an increasing, convex and non-linear function with respect to the total arc flow. Hereafter, we consider a collection $\mathcal{L} = \{1 \dots L\}$ of piecewise linear approximations to give an estimate from below of the time cost. Let $u_a(x_a) = u_a^0 \frac{r_a}{5(q_a)^4} (x_a)^5$ and define a collection of L pivots $\delta_1 \dots \delta_L$. A first-order Taylor expansion of $u_a(x_a)$ at point d_ℓ is uniquely characterized by the corresponding slope $d_a^\ell = u_a^0 r_a (d_\ell)^4$ and intercept $e_a^\ell = -\frac{4}{5(q_a)^4} u_a^0 r_a (d_\ell)^5$, so that, for any link $a \in \mathcal{A}$, we have:

$$u_a(x_a) = u_a^0 \frac{r_a}{5(q_a)^4} (x_a)^5 \geq \tilde{u}_a(x_a) = \max_{\ell \in \mathcal{L}} \{d_a^\ell x_a + e_a^\ell\} \quad (5)$$

For a given road link $a \in \mathcal{A}$, an illustrative example is given in Fig. 2 for $L = 3$. The three approximation lines are specified by the pivots δ_ℓ and knots $(\alpha_{\ell-1}, \tilde{u}_a(\alpha_{\ell-1}))$ and $(\alpha_\ell, \tilde{u}_a(\alpha_\ell))$.

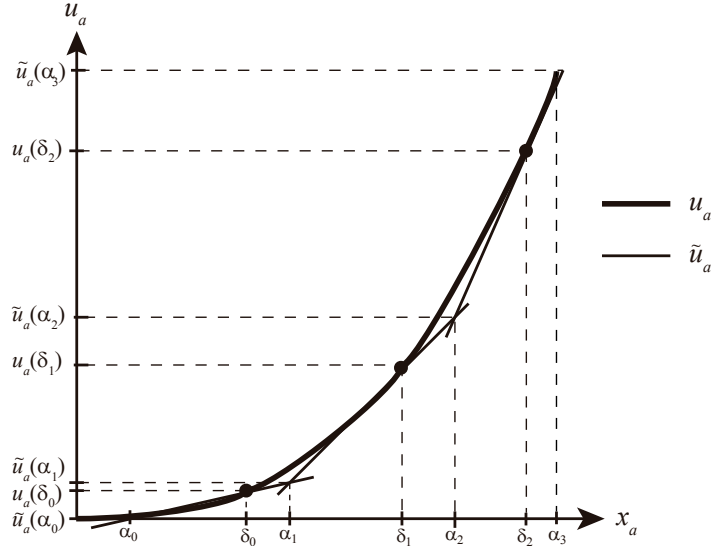


Figure 2: Example of a linear piecewise approximation with 3 lines

Note that, the piecewise linear function coincides with the time cost function in the pivots $\delta_1 \dots \delta_L$, whereas the maximum deviation values are in the knots. These are determined using the method of Imamoto & Tang [26], described in detail in Appendix B. Since the time cost function is convex, the piecewise linear function (5) can be transformed into the following set of linear inequalities:

$$\tilde{u}_a \geq d_a^\ell x_a + e_a^\ell, \quad \forall a \in \mathcal{A}, \ell \in \mathcal{L}. \quad (6)$$

These constraints are appended to the follower's problem (3), and the non-linear term of the objective function is replaced with the new variables \tilde{u}_a , resulting in the following linearized follower's problem:

$$\tilde{\Psi}_P(\mathbf{t}, \mathbf{w}) = \underset{\mathbf{x}, \tilde{\mathbf{u}}}{\operatorname{argmin}} \quad \sum_{a \in \mathcal{A}^T} t_a x_a + \sum_{a \in \mathcal{A}} (u_a^0 x_a + \tilde{u}_a) \quad (7a)$$

$$\text{subject to: } \sum_{a \in \mathcal{A}} N_{i,a} x_a^{v,k} = s_i^{v,k}, \quad [\alpha_i^{v,k}] \quad \forall i \in \mathcal{N}, v \in \mathcal{V}, k \in \mathcal{K}^v \quad (7b)$$

$$x_a^v \leq M_{a,v}^{(w)} w_a^v, \quad [\beta_a^v] \quad \forall a \in \mathcal{A}^T, v \in \mathcal{V} \quad (7c)$$

$$\tilde{u}_a \geq d_a^\ell x_a + e_a^\ell, \quad [\gamma_a^\ell] \quad \forall a \in \mathcal{A}, \ell \in \mathcal{L} \quad (7d)$$

$$x_a^{v,k} \geq 0, \quad \forall a \in \mathcal{A}, v \in \mathcal{V}, k \in \mathcal{K}^v \quad (7e)$$

$$\tilde{u}_a \geq 0, \quad \forall a \in \mathcal{A} \quad (7f)$$

Problem (7) consists of a multicommodity network flow model with the inclusion of $|\mathcal{L}|$ min-max constraints. The values of $\alpha_i^{v,k}$, β_a^v and γ_a^ℓ are the dual variables associated with constraints (7b)-(7d), respectively.

In its general form, the problem of the network operator (1) is a non-convex optimization problem with an implicitly defined feasible region. When it comes to its linearized version, we know that $\tilde{\Psi}_P(\mathbf{t}, \mathbf{y})$ approaches $\Psi(\mathbf{t}, \mathbf{w})$, as $|\mathcal{L}|$ grows large. This results in a trade-off between problem size and accuracy of the approximation.

The existence of a solution to this bi-level problem with linearized follower can be deduced based on the work in [3, 25, 11], as established in the next proposition.

Proposition 1. *Consider the linearized network operator problem, where $\mathbf{x} \in \Psi(\mathbf{t}, \mathbf{w})$ is replaced with $\mathbf{x} \in \tilde{\Psi}_P(\mathbf{t}, \mathbf{w})$. Then, for any L , as long as the network of toll-free links \mathcal{A}^F is connected, the network operator can always find at least one optimal combination of toll levels \mathbf{t} , accessibility pattern \mathbf{w} and constructed links \mathbf{y} .*

A proof of this proposition can be found on Appendix C. In the rest of the paper, the focus will be on the approximated bi-level problem, where $\mathbf{x} \in \Psi(\mathbf{t}, \mathbf{w})$ is replaced with $\mathbf{x} \in \tilde{\Psi}_P(\mathbf{t}, \mathbf{w})$ in (1).

4.2. Single Level Reformulation

As already discussed in Subsection 3.3, generally adopted techniques for solving MINLBPs rely on a single level reformulation [32, 46, 19, 29]. This reformulation is based on the optimality conditions of the lower level problem (i.e., the vehicle's best response from the linearized version, $\tilde{\Psi}_P(\mathbf{t}, \mathbf{w})$, in our case), which are represented by a set of constraints and variables establishing the relationship between primal and dual forms in the follower's optimum.

First, we consider the dual form of the linearized follower's problem (7):

$$\tilde{\Psi}_D(\mathbf{t}, \mathbf{w}) = \underset{\alpha, \beta, \gamma}{\operatorname{argmax}} \quad \sum_{v \in \mathcal{V}} \sum_{k \in \mathcal{K}^v} \sum_{i \in \mathcal{N}} s_i^{v,k} \alpha_i^{v,k} - \sum_{a \in \mathcal{A}^T} \sum_{v \in \mathcal{V}} M_{a,v}^{(w)} w_a^v \beta_a^v + \sum_{a \in \mathcal{A}} \sum_{\ell \in \mathcal{L}} e_a^\ell \gamma_a^\ell \quad (8a)$$

$$\text{subj. to } \alpha_{j(a)}^{v,k} - \alpha_{i(a)}^{v,k} + \beta_a^v + \sum_{l \in \mathcal{L}} d_a^l \gamma_a^l \geq -t_a - u_a^0, \quad \forall a \in \mathcal{A}^T, v \in \mathcal{V}, k \in \mathcal{K}^v \quad (8b)$$

$$\alpha_{j(a)}^{v,k} - \alpha_{i(a)}^{v,k} + \sum_{l \in \mathcal{L}} d_a^l \gamma_a^l \geq -u_a^0, \quad \forall a \in \mathcal{A}^F, v \in \mathcal{V}, k \in \mathcal{K}^v \quad (8c)$$

$$-\sum_{l \in \mathcal{L}} \gamma_a^l \geq -1, \quad \forall a \in \mathcal{A} \quad (8d)$$

$$\beta_a^v \geq 0, \quad \forall a \in \mathcal{A}^T, v \in \mathcal{V} \quad (8e)$$

$$\gamma_a^\ell \geq 0, \quad \forall a \in \mathcal{A}, \ell \in \mathcal{L} \quad (8f)$$

where the notation $i(a)$ and $j(a)$ has been introduced to denote the first and second endpoint of either $a \in \mathcal{A}^T$ or $a \in \mathcal{A}^F$.

Since (7) is a linear program; building on strong duality, the vehicle's best response $\mathbf{x} \in \tilde{\Psi}(\mathbf{t}, \mathbf{w})$ can be established by requiring the equality between primal payoff and its dual. The resulting single level problem is then as follows:

$$\max_{\mathbf{t}, \mathbf{y}, \mathbf{w}, \mathbf{x}, \tilde{\mathbf{u}}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}} \quad \sum_{a \in \mathcal{A}^T} t_a x_a - \sum_{a \in \mathcal{A}^N} c_a y_a - \sum_{a \in \mathcal{A}} n_a \sum_{v \in \mathcal{V}} m_v x_a^v \quad (9a)$$

subj. to

$$(C1) \left\{ \begin{array}{ll} \sum_{a \in \mathcal{A}^N} c_a y_a \leq B & (9b) \\ t_a \leq T_a, & \forall a \in \mathcal{A}^T \quad (9c) \\ y_a \geq w_a^v, & \forall a \in \mathcal{A}^N, v \in \mathcal{V} \quad (9d) \end{array} \right.$$

$$(C2) \left\{ \begin{array}{ll} \alpha_{j(a)}^{v,k} - \alpha_{i(a)}^{v,k} + \beta_a^v + \sum_{l \in \mathcal{L}} d_a^l \gamma_a^l \geq -t_a - u_a^0, & \forall a \in \mathcal{A}^T, v \in \mathcal{V}, k \in \mathcal{K}^v \quad (9e) \\ \alpha_{j(a)}^{v,k} - \alpha_{i(a)}^{v,k} + \sum_{l \in \mathcal{L}} d_a^l \gamma_a^l \geq -u_a^0, & \forall a \in \mathcal{A}^F, v \in \mathcal{V}, k \in \mathcal{K}^v \quad (9f) \\ \sum_{l \in \mathcal{L}} \gamma_a^l \leq 1, & \forall a \in \mathcal{A} \quad (9g) \end{array} \right.$$

$$(C3) \left\{ \begin{array}{ll} \sum_{a \in \mathcal{A}} N_{i,a} x_a^{v,k} = s_i^{v,k}, & \forall i \in \mathcal{N}, v \in \mathcal{V}, k \in \mathcal{K}^v \quad (9h) \\ x_a^v \leq M_{a,v}^{(w)} w_a^v, & \forall a \in \mathcal{A}^T, v \in \mathcal{V} \quad (9i) \\ \tilde{u}_a \geq d_a^\ell \sum_{v \in \mathcal{V}} \sum_{k \in \mathcal{K}^v} x_a^{v,k} + e_a^\ell, & \forall a \in \mathcal{A}, \ell \in \mathcal{L} \quad (9j) \end{array} \right.$$

$$(C4) \left\{ \sum_{v \in \mathcal{V}} \sum_{k \in \mathcal{K}^v} \sum_{i \in \mathcal{N}} s_i^{v,k} \alpha_i^{v,k} - \sum_{a \in \mathcal{A}^T} \sum_{v \in \mathcal{V}} M_{a,v}^{(w)} w_a^v \beta_a^v + \sum_{a \in \mathcal{A}} \sum_{\ell \in \mathcal{L}} e_a^\ell \gamma_a^\ell = \sum_{a \in \mathcal{A}^T} t_a x_a + \sum_{a \in \mathcal{A}} (u_a^0 x_a + \tilde{u}_a) \quad (9k) \right.$$

$$(C5) \left\{ \begin{array}{ll} t_a \geq 0, & \forall a \in \mathcal{A}^T \quad (9l) \\ y_a \in \{0, 1\}, & \forall a \in \mathcal{A}^N \quad (9m) \\ w_a^v \in \{0, 1\}, & \forall a \in \mathcal{A}^T, v \in \mathcal{V} \quad (9n) \\ x_a^{v,k} \geq 0, & \forall a \in \mathcal{A}, v \in \mathcal{V}, k \in \mathcal{K}^v \quad (9o) \\ \beta_a^v \geq 0, & \forall a \in \mathcal{A}^T, v \in \mathcal{V} \quad (9p) \\ \gamma_a^\ell \geq 0, & \forall a \in \mathcal{A}, \ell \in \mathcal{L} \quad (9q) \\ \tilde{u}_a \geq 0, & \forall a \in \mathcal{A} \quad (9r) \end{array} \right.$$

where (C1) characterize the feasible operator choices, (C2) contain the follower's primal feasibility, (C3) ensure the follower's dual feasibility, (C4) is the strong duality condition to characterize the follower's optimality, (C5) define variable domains.

4.3. Building a feasible solution by linearizing the bi-linear terms

The single level reformulation problem (9a)–(9r) is non-linear and non-convex because of the presence of two types of bi-linear terms: the ones involving $w_a^v \beta_a^v$ in the strong duality condition (12a), and the ones involving $t_a x_a$ in the objective function (9a) and strong duality condition (12a).⁷

⁷Some literature works propose convex envelopes to linearize bi-linear terms (see, for instance, Esfandeh et al.[15]). However, these approaches require the inclusion of a large number of variables and constraints, preventing the convexified model to be solved for real-sized networks.

In this subsection, we present a linearization procedure which allows building a feasible solution of problem (9a)–(9r). This solution can be made arbitrarily close to the optimal solution at the expense of the problem size and, consequently, its computational efficiency.

The linearization of $w_a^v \beta_a^v$ is done by noticing that the big-M type constraint (9i) is equivalent to $w_a^v \beta_a^v = 0$, as $\beta_a^v = 0$, when $w_a^v = 1$ for sufficiently large values of M (inactive bound). Thus, for a large positive scalar $M_{a,s}^{(\beta)}$, we can introduce the constraint:

$$\beta_a^v \leq M_{a,s}^{(\beta)}(1 - w_a^v) \quad \forall a \in \mathcal{A}^T, v \in \mathcal{V}. \quad (10)$$

As for the second bi-linear terms, we propose a discrete approximation approach including new variables and constraints (this is exploited by the specialized local search procedure presented in the next section). Let us define a set of toll levels $\mathcal{S} = \{1 \dots S\}$, and approximate the toll variables t_a with the aid of binary variables z_a^s (denoting if toll level $s \in \mathcal{S}$ is charged at road link $a \in \mathcal{A}^T$), as follows:

$$t_a = \sum_{s \in \mathcal{S}} p_a^s z_a^s, \quad \forall a \in \mathcal{A}^T \quad \text{and} \quad \sum_{s \in \mathcal{S}} z_a^s = 1, \quad \forall a \in \mathcal{A}^T \quad (11)$$

where the value associated with each toll level is set by parameter p_a^s , so toll variables t_a are replaced with toll levels z_a^s , for a finite number of levels $s \in \mathcal{S} = \{1 \dots S\}$.

Replacing t_a with $\sum_{s \in \mathcal{S}} p_a^s z_a^s$ in all occurrences of (9a)–(9r), results in new bi-linear terms involving products between binary and continuous variables. As suggested by Fourer [22], these terms can be straightforwardly linearized for each $a \in \mathcal{A}^T$ by the inclusion of a continuous non-negative variable φ_a bounded by $\sum_{s \in \mathcal{S}} p_a^s z_a^s x_a$. It turns out that the strong duality condition is expressed as follows:

$$(C4') \left\{ \begin{array}{ll} \sum_{v \in \mathcal{V}} \sum_{k \in \mathcal{K}^v} \sum_{i \in \mathcal{N}} s_i^{v,k} \alpha_i^{v,k} + \sum_{a \in \mathcal{A}} \sum_{\ell \in \mathcal{L}} e_a^\ell \gamma_a^\ell = \sum_{a \in \mathcal{A}^T} \varphi_a + \sum_{a \in \mathcal{A}} (u_a^0 x_a + \tilde{u}_a) & (12a) \\ \varphi_a \geq 0, & \forall a \in \mathcal{A}^T \quad (12b) \\ \varphi_a \leq p_a^s x_a + M_{a,s}^{(\varphi,c)}(1 - z_a^s), & \forall a \in \mathcal{A}^T, s \in \mathcal{S} \quad (12c) \\ \varphi_a \geq p_a^s x_a - M_{a,s}^{(\varphi,d)}(1 - z_a^s), & \forall a \in \mathcal{A}^T, s \in \mathcal{S} \quad (12d) \\ \beta_a^v \leq M^{(\beta)}(1 - w_a^v) & \forall a \in \mathcal{A}^T, v \in \mathcal{V} \quad (12e) \end{array} \right.$$

where $M_{a,s}^{(\varphi,c)}$ and $M_{a,s}^{(\varphi,d)}$ are large positive scalars disabling constraints (12c) and (12d) respectively, when not active. In the next section, a dynamic update of these parameters is defined to reduce the number of explored branch-and-bound nodes.

Based on the described linearization, problem (9a)–(9r) is approximated by the following MILP:

$$\max_{z, y, \varphi, x, \tilde{u}, \alpha, \beta, \gamma} \Upsilon = \sum_{a \in \mathcal{A}^T} \varphi_a - \sum_{a \in \mathcal{A}^N} c_a y_a - \sum_{a \in \mathcal{A}} n_a \sum_{v \in \mathcal{V}} m_v x_a^v, \quad \text{s.t.} \quad (C1)-(C3), (C4'), (C5) \quad (13)$$

Observe that, even for small-sized networks, linearized single-level reformulations of the form (13) result in large MILP instances. Therefore, an efficient solution approach is studied in the next section, and numerically tested in Section 7.

5. Local search procedure based on binary-searching on toll levels

The MILP (13) provides a feasible solution of (9a)–(9r), which we can try to improve by increasing the number of toll levels $|\mathcal{S}|$. However, this increase should be carefully done as the problem size grows substantially with $|\mathcal{S}|$. In fact, passing from $|\mathcal{S}|$ to $|\mathcal{S}| + 1$ entails an increase of $|\mathcal{A}^T|$ in the number of variables, thus making computationally challenging the definition of a large number of discrete tolls. Conversely, using small $|\mathcal{S}|$ may lead to poor solutions, if compared to their continuous counterpart.

Therefore, the interest is in providing reasonably good feasible solutions of (9a)–(9r), while solving (13) with a small number of toll levels. To do so, a specialized local search exploiting the linearization strategy in (11) is proposed. This strategy builds on the workings of the well known binary search [43], a search that iteratively shrinks by half the toll interval $([0, \mathbf{T}])$, in which the toll charges are determined.

Our application of the binary search works as follows. For a given link $a \in \mathcal{A}^T$, and a number of toll levels $|\mathcal{S}|$, an initial partition of $[0, T_a] = [p_a^1, p_a^2) \cup [p_a^2, p_a^3) \cup \dots \cup [p_a^{|\mathcal{S}|-1}, p_a^{|\mathcal{S}|})$ is set. In each iteration, an instance of (13) is solved with the given toll levels $p_a^1, \dots, p_a^{|\mathcal{S}|}$. The opposite half of the current interval in which the toll level is selected is discarded for the next iteration, and the partition is updated, distributing the toll levels uniformly on the shrunk toll interval. This procedure is repeated until the relative gap of the network profit between successive iterations is as small as desired.

Algorithm 1 formalizes the pseudo-code of the proposed procedure for a given value of $|\mathcal{S}|$ and \mathbf{T} . Parameter ϵ is used to set the tolerable gap between consecutive values of the network profit (Υ).

Algorithm 1 Binary-search-based approach to solve large instances of (13).

```

1: Input:  $j = 0, \mathcal{S}, \mathcal{G}, \epsilon$ ; Output:  $\mathbf{p}, \mathbf{z}, \mathbf{y}, \mathbf{x}, \tilde{\mathbf{u}}$ ;
2: Initialize incumbent objective:  $\Upsilon^0 = -\infty$ ;
3: Initialize toll levels:  $p_a^s = \frac{(s-1)T_a}{|\mathcal{S}|-1}$ , for  $a \in \mathcal{A}^T, s \in \mathcal{S}$ ;
4: repeat
5:    $j++$ ;
6:   Solve (13) and let  $\Upsilon^j$  be the optimal objective;
7:   for all  $a \in \mathcal{A}^T$  do
8:     if  $\sum_{s=1}^{\lceil |\mathcal{S}|/2 \rceil} z_a^s = 0$  then
9:        $p_a^1 = \frac{1}{2} (p_a^1 + p_a^{|\mathcal{S}|})$ ;
10:    else
11:       $p_a^{|\mathcal{S}|} = \frac{1}{2} (p_a^1 + p_a^{|\mathcal{S}|})$ ;
12:    end if
13:    for all  $s \in 2 \dots |\mathcal{S}| - 1$  do
14:       $p_a^s = p_a^1 + \frac{s-1}{|\mathcal{S}|-1} (p_a^{|\mathcal{S}|} - p_a^1)$ ;
15:    end for
16:  end for
17: until  $\frac{\Upsilon^j - \Upsilon^{(j-1)}}{\Upsilon^j} \leq \epsilon$ 

```

The way in which toll levels are updated allows the objective function in (13) to increase monotonically.

Lemma 2. *At the j^{th} iteration of Algorithm 1, let $P^{(j)} = \{\mathbf{p}^{1,(j)}, \mathbf{p}^{2,(j)}, \dots, \mathbf{p}^{|\mathcal{S}|,(j)}\}$ be the ordered sets of available toll levels, and let $\hat{\mathbf{p}}^{(j)} \in P^{(j)}$ be the optimal toll level from the solution of (13) at the j^{th} iteration of Algorithm 1. We claim that $\hat{\mathbf{p}}^{(j)} \in P^{(j+1)}$, i.e. the optimal toll level at the j^{th} iteration of Algorithm 1 is contained in the sets of available toll levels of iteration $(j+1)^{\text{th}}$.*

Theorem 3. *For each iteration j of Algorithm 1, we have $\Upsilon^{(j+1)} \geq \Upsilon^j$.*

A proof of lemma 2 and Theorem 3 can be found on Appendix C.

Example 2. *We consider four hypothetical iterations of the proposed procedure for a given link $a \in \mathcal{A}^T$ in the case $|\mathcal{S}| = 5$ (i.e., 5 toll levels are established). The changes in the toll interval can be described as follows:*

- Initially, the toll interval is set to $[0, T_a]$, so the toll levels are set to $p_a^1 = 0, p_a^2 = T_a/4, p_a^3 = 2T_a/4, p_a^4 = 3T_a/4, p_a^5 = T_a$. The solution of (13) at the first iteration selects the toll level p_a^1 , located in the left half-interval $[0, T_a/2]$, as denoted by the black circle.
- As a result, the toll interval is shrunk to $[0, T_a/2]$, and the toll levels are now updated to $p_a^1 = 0, p_a^2 = T_a/8, p_a^3 = 2T_a/8, p_a^4 = 3T_a/8, p_a^5 = T_a/2$. The solution of (13) at the second iteration selects the toll level p_a^4 , located in the right half-interval of $[T_a/4, T_a/2]$.
- Thus, the current interval $[0, T_a/2]$ is shrunk to $[T_a/4, 2T_a/4]$, and the toll levels are now updated to $p_a^1 = T_a/4, p_a^2 = 5T_a/16, p_a^3 = 3T_a/8, p_a^4 = 7T_a/16, p_a^5 = T_a/2$. The solution of (13) at the third iteration selects the toll level $p_a^{v,1}$, located in the left half-interval of $[T_a/4, 3T_a/8]$.
- Now, the interval is shrunk to $[T_a/4, 3T_a/8]$, and the toll levels updated to $p_a^1 = T_a/4, p_a^2 = 9T_a/32, p_a^3 = 10T_a/32, p_a^4 = 11T_a/32, p_a^5 = 3T_a/8$. The solution of (13) at the fourth iteration, selects p_a^4 .

Fig. 3 illustrates provides a graphical illustration for these four iterations.

To sum up, in this illustrative example, bounding the toll value from $[0, T_a]$ to $[T_a/4, 3T_a/8]$ requires solving four times problem (13) with $|\mathcal{S}| = 5$ toll levels.

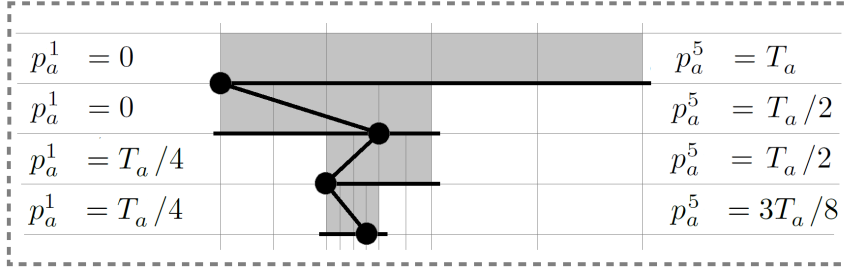


Figure 3: Illustration of four iterations of the binary-search-based procedure for a given link $a \in \mathcal{A}^T$, and the case $|\mathcal{S}| = 5$. The selected toll level is denoted by a black circle.

The main computational effort in Algorithm 1 relies on the solution of problem (13) at line 7. For this purpose, the optimality bound provided in Theorem 3 establishes a valid inequality to be appended to problem (13) from the second to the last iteration of Algorithm 1. As shown in Section 7, the inclusion of this valid inequality produces a substantial reduction of the explored branch-and-bound nodes when large-scale instances of problem (13) are solved. As a consequence, the first iteration of Algorithm 1 (the constructive heuristics) results computationally more challenging than the subsequent ones.⁸

Although the specialized local search presented in this section constitutes a heuristic method, whose solution has no guarantee of global optimality, it results to be extremely effective in practice, as shown in Section 7. In fact, the linearization of the bi-linear term in constraint (C4) requires a large collection of binary variables that make problem (13) computationally challenging. To counterbalance the trade-off between accuracy and efficiency in the construction of the set \mathcal{S} of toll levels, the proposed binary-search based procedure sequentially solves small instances of problem (13), based on a parsimonious inclusion of toll levels.

6. Data selection and processing

To show the effectiveness of the proposed methodology, three extensions of the Sioux Falls road network (each having different amounts of available candidate toll links to be built) have been generated. These extensions are devised for assessing both the solution properties and the computational performance under a variety of structural configurations. The original Sioux Falls road network has 24 nodes and 76 arcs, together with 552 origin-destination pairs that are included in our analysis.⁹

The three extensions of the Sioux Falls road network are depicted in figures 4-6 for the respective cases of 8, 16 and 24 candidate toll links to be built.¹⁰

⁸Note that adjusted values of $M_{a,s}^{(\beta)}$, $M_{a,s}^{(\varphi,c)}$ and $M_{a,s}^{(\varphi,d)}$ in the big-M type of constraints might also serve to reduce the number of explored branch-and-bound nodes, while preventing the potential infeasibility induced by the rounding procedure of many numerical solvers. Building on constraints (12c) and (12d), the values of $M_{a,s}^{(\beta)}$, $M_{a,s}^{(\varphi,c)}$, and $M_{a,s}^{(\varphi,d)}$ can be dynamically updated at each iteration of Algorithm 1. Bounding the left-hand side of each big-M constraint to their maximum values, $M_{a,s}^{(\beta)}$, $M_{a,s}^{(\varphi,c)}$ and $M_{a,s}^{(\varphi,d)}$ are set at the j^{th} iteration of Algorithm 1 as follows:

$$M_{a,s}^{(\beta)} = \sum_{a' \in \mathcal{A}^T} \left(u_{a'}^0 + \sum_{l \in \mathcal{L}} d_{a'}^l + P_{a'}^{|\mathcal{S}|} \right), \quad M_{a,s}^{(\varphi,c)} = (\max\{p_a^q(j) : q \in \mathcal{S}\} - p_a^s(j)) G_a \quad \text{and} \quad M_{a,s}^{(\varphi,d)} = p_a^s(j) G_a, \quad (14)$$

where $G_a = \sum_{v \in \mathcal{V}, k \in K_a^v} g^{v,k}$ is the maximum flow through link a and $p_a^s(j)$ the s^{th} toll level for arc a at the j^{th} iteration of Algorithm 1.

⁹The basic data of this road network has been taken from the repository of the *Transportation Networks for Research Core Team* [36], as this network has been already used in previous operations research studies, since the work of Suwansirikul et al. [38].

¹⁰The selection of the candidate toll links from the pairs of locations in the Sioux Falls road network (corresponding to the green dotted lines in figures 4-6) has not been done randomly. To build a realistic scenario of urban transportation, each candidate location for a new road link has been set evaluating the total free flow time (i.e. the total travel times for all origin-destination pairs without congestion on a saturated network) using 552 possible links (76 existing and 476 potential new arcs). This information is used in a simplified network design problem (consisting on a standard multi-commodity formulation with binary link construction) in which we pick the K link locations among the $\mathcal{N} \times \mathcal{N}$ which minimize the total construction cost, while meeting minimum achievable travel time for the network. These K locations correspond to the green dotted lines in figures 4-6 for the respective cases of 8, 16 and 24.

To fit the modeling approach for RNPRES problems, the Sioux Falls road network data set has been enriched by the additional information concerning the construction costs, the hazmat transportation risk (for the population living near arcs) and the toll bounds for each route link (resulting from legal requirements that the network operator must fulfill).

The construction costs are estimated assuming an average toll per km of link length and using a 5-year amortization horizon. The travel demand assigned to one new link in the simplified network design problem should be able to cope with the total construction costs of the link within the 5 years of continuous use. As the RNPRES problem is solved assuming a typical working day, these costs are normalized accordingly.

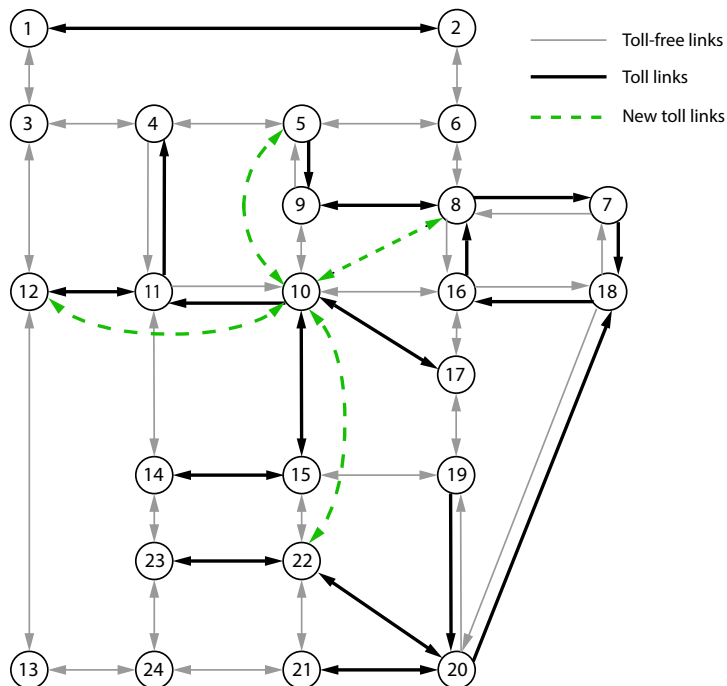


Figure 4: Sioux Falls road network with 8 potential new links, 14 existing toll-free links and 13 existing payment links.

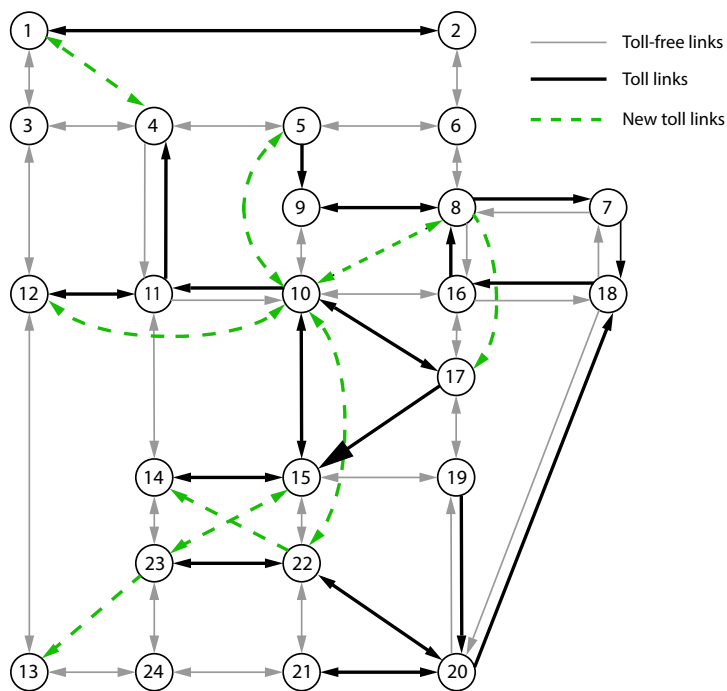


Figure 5: Sioux Falls road network with 16 potential new links, 14 existing toll-free links and 13 existing payment links.

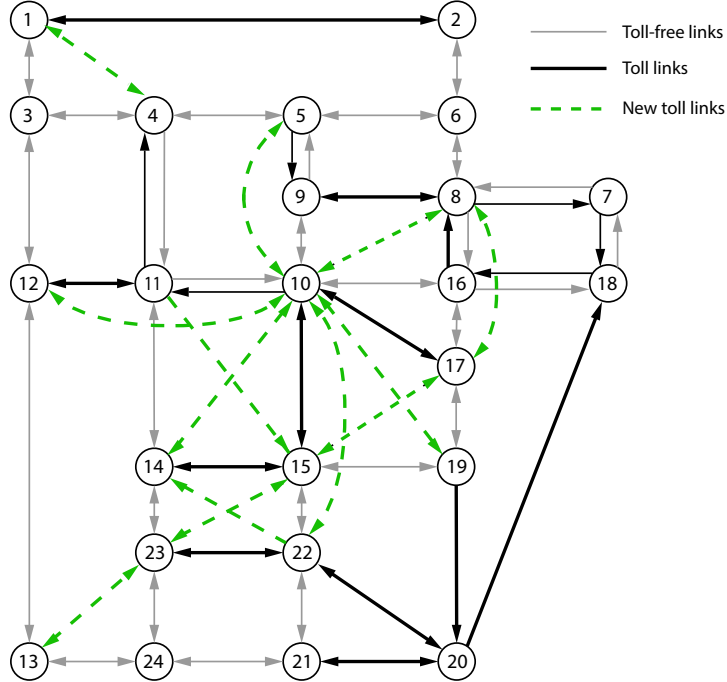


Figure 6: Sioux Falls road network with 32 potential new links, 14 existing toll-free links and 13 existing payment links.

As for the number of people exposed to hazmat (n_a , for $a \in A$), considering that Sioux falls is a metropolitan network (i.e., the longer the length the higher the population living nearby), this amount is randomly generated using a uniform distribution between 0.03 and 0.08 and multiplied by the link length. When using this estimation for every link in the Sioux falls network, the sum $\sum_{a \in A} n_a$ is almost linearly proportional to the population estimated for this network. Using the total amounts per origin-destination pairs reported in the same repository [36], the 70% was assigned to ordinary vehicles and the 30% to hazmat vehicles.

These extended data are summarized in Table 2.

Parameter	Average	Max	Min	SD
T_a	1	2	0.4	0.31
r_a	0.15	0.15	0.15	0
q_a	12343	25900	4948	5962
$g^{0,k}$	478	3080	70	487
$g^{1,k}$	205	1320	30	209
n_a	85	200	40	31
c_a	462	850	180	199

Table 2: Summary statistics of the additional parameters for the Sioux Falls road network, including the average, maximum and minimum values and the standard deviation across the set of arcs.

7. Numerical Experiments

Building on the solution mechanism described in Section 5, this section analyzes traffic management policies for the Sioux Falls networks depicted in Fig. 4-6, with a view to supporting the effectiveness of the proposed methodology to solve large-scale instances of road network pricing, regulation, and expansion problems. To do so, two computational experiments have been conducted, covering a variety of structural configurations and economic settings.

The first experiment (analyzed in Subsection 7.1) focuses on the impact of the reformulation and linearization approach on the specialized local search algorithm. It involves a battery of 1620 instances of RNPPE problems, generated by the following parameters configurations:

- $|\mathcal{L}|$: number of approximation lines for the BPR function, with values $\{2, 4, 6, 8\}$;
- $|\mathcal{S}|$: number of toll levels in the discretization (11), with values $\{2, 4, 8, 16, 32\}$;
- $|\mathcal{A}^T|$: number of candidate toll links in the Sioux Falls road networks, with values $\{8, 16, 24\}$;

For each parameter configuration $3^3 = 27$ replicates have been generated by setting m , u^0 , and B at three levels (within a predefined range).

As the core reformulation and linearization strategy for the RNPRES problem relies on the selection of $|\mathcal{S}|$ and $|\mathcal{L}|$, this experiment pivots on testing the benefit of applying Algorithm 1 (where each iteration is performed based on the ILOG CPLEX 12.8 implementation of the Branch-and-cut method).

The second experiment (analyzed in Subsection 7.2) has been design to assess the effect of exogenous business settings on the road links creation and travel time. It involves a battery of 81 instances of RNPRES problems with $|\mathcal{L}| = 8$ and $|\mathcal{S}| = 32$, generated by the following parameters configurations:

- m : monetary cost of the exposure to risk, with three values $\{0.001, 0.005, 0.01\}$;
- u^0 : monetary cost of congestion for each individual, with three values $\{0.001, 0.005, 0.01\}$;
- B : construction budget, with three values $\{0.1, 0.5, 0.9\}$;
- ρ : homogeneous discount for all the construction costs, with three values $\{0.25, 0.5, 0.75\}$;

All optimization procedures are coded in AMPL platform and solved using IBM CPLEX 12.8.0 on a HP ProDesk 600 G3 SFF with Intel(R) Core(TM) i5-7500 CPU and 32 GB of RAM. [The integrality tolerance of the CPLEX rounding procedure has been set at 1.0e-6.](#)

7.1. Network operator profit versus computational performance

The motivating idea behind the numerical assessment presented in this subsection is to show that, although the linearization approach proposed in Subsection (4.3) gives rise to a $|\mathcal{S}|$ -levels discrete approximation of the original RNPRES problem, for each value of $|\mathcal{S}|$ the initial solution can be improved with limited computational effort.

Tables 3-5 show the solutions and performances of the 1620 problem instances corresponding to extensions of the Sioux Falls road network with 8 (Fig. 4), 16 (Fig. 5) and 24 (Fig. 6) available candidate toll links, respectively. The values of each row are averaged over the $3^3 = 27$ instances corresponding to the combinations of m , u^0 and B . As a stopping condition for Algorithm 1, we use a $1.0e - 6$ gap between consecutive iterations (i.e. $(\Upsilon^j - \Upsilon^{(j-1)})/\Upsilon^j \leq 1.0e - 6$). Beside, a $1.0e - 6$ MIP-GAP has been set for each internal iteration, with 3648720 (twelve hours) of time limit. CPU times are reported in seconds.

Scenario		Initial solution		Local search			Local search performance			
$ \mathcal{L} $	$ \mathcal{S} $	OF_{NO}	T_{CPU}	OF_{NO}	Iter	T_{CPU}	Δ_{OF}^{AV}	Δ_{OF}^{Max}	$\#\Delta_{OF} > 0$	ΔT_{CPU}
2	2	32717823	154	33201871	3.93	241	1.82 %	3.2 %	92.59 %	87
	4	32718294	70	32743218	3.30	138	0.14 %	0.58 %	48.15 %	68
	8	32717778	283	33016762	4.22	355	0.98 %	1.26 %	44.44 %	82
	16	32717984	320	34531234	3.56	612	4.59 %	36 %	62.96 %	293
	32	35321842	1128	37729245	3.00	1512	12.35 %	36.05 %	48.15 %	484
4	2	43510171	393	43873118	3.78	444	0.89 %	3 %	85.19 %	51
	4	43546865	386	43653973	2.93	470	0.23 %	1.11 %	40.74 %	84
	8	43546864	256	43721514	3.30	324	0.40 %	0.46 %	55.56 %	67
	16	42337094	542	42424696	3.48	760	0.22 %	0.56 %	48.15 %	219
	32	42307599	392	42383794	3.11	562	0.19 %	1.21 %	48.15 %	170
6	2	47649462	700	47657193	3.33	945	0.08 %	0.86 %	85.19 %	245
	4	47654792	131	47702408	2.52	492	0.09 %	0.13 %	25.92 %	361
	8	47655281	223	47702887	2.70	549	0.10 %	0.25 %	37.04 %	326
	16	46332173	704	46424158	2.93	930	0.19 %	0.2 %	48.15 %	226
	32	47642970	422	47893731	3.00	543	0.50 %	0.7 %	59.26 %	121
8	2	48871285	1849	49114435	3.33	1983	0.48 %	0.77 %	81.48 %	135
	4	50261528	696	50511711	2.70	831	0.49 %	0.66 %	55.55 %	135
	8	50263561	1075	50352970	3.26	1100	0.43 %	0.86 %	48.15 %	25
	16	46332173	704	46424158	2.67	1279	0.18 %	1.73 %	29.63 %	576
	32	47118003	2133	48363471	2.67	2355	2.65 %	58.34 %	48.15 %	222

Table 3: Performance results of the local-search procedure defined in Algorithm 1 on the Sioux Falls road network with 8 available candidate toll links, as depicted on Fig. 4.

Scenario		Initial solution		Local search			Local search performance				
$ \mathcal{L} $	$ \mathcal{S} $	OF_{NO}	T_{CPU}	OF_{NO}	Iter	T_{CPU}	Δ_{OF}^{AV}	Δ_{OF}^{Max}	$\#\Delta_{OF} > 0$	ΔT_{CPU}	
2	2	36960097	199	36997059	2.52	278	0.1 %	0.2 %	33.33 %	79	
	4	36959469	617	36997068	3.00	755	0.11 %	0.25 %	48.15 %	138	
	8	36960027	351	37144689	2.52	426	0.50 %	0.52 %	37.04 %	75	
	16	35089191	1111	36995992	3.00	1289	10.04 %	268 %	40.74 %	177	
	32	35321842	1128	36996537	2.74	1362	17.25 %	191 %	44.44 %	234	
4	2	47321735	151	47474340	2.52	214	0.32 %	0.43 %	40.74 %	63	
	4	49204202	1036	49351523	2.70	1237	0.30 %	0.31 %	29.63 %	200	
	8	49204869	1446	49249900	2.26	1637	0.10 %	0.11 %	25.93 %	191	
	16	47455880	1351	49233683	2.67	1621	3.85 %	100 %	44.44 %	270	
	32	45787315	1444	49253199	2.59	1714	7.51 %	100 %	44.44 %	270	
6	2	53849248	97	53904637	2.56	182	0.12 %	0.51 %	33.33 %	85	
	4	53856049	1079	53908808	2.33	1256	0.11 %	0.12 %	29.63 %	178	
	8	53854030	1179	53906169	2.37	1311	0.10 %	0.12 %	29.63 %	132	
	16	53782893	5318	53903467	2.37	6030	0.17 %	1 %	37.04 %	712	
	32	48387209	3122	52278917	2.59	5688	18.50 %	417 %	51.85 %	2566	
8	2	56799077	95	57065263	2.44	176	0.77 %	1.51 %	37.04 %	80	
	4	56800029	2420	56912343	2.52	2855	0.22 %	0.42 %	40.74 %	434	
	8	54709852	5896	54934609	2.63	6211	0.50 %	0.61 %	48.15 %	225	
	16	53782893	5318	53903467	2.33	7444	0.42 %	0.51 %	25.93 %	2126	
	32	54572380	7391	55710934	2.22	7641	1.55 %	28 %	22.22 %	250	

Table 4: Performance results of the local-search procedure defined in Algorithm 1 on the Sioux Falls road network with 16 available candidate toll links, as depicted on Fig. 5.

Scenario		Initial solution		Local search			Local search performance				
$ \mathcal{L} $	$ \mathcal{S} $	OF_{NO}	T_{CPU}	OF_{NO}	Iter	T_{CPU}	Δ_{OF}^{AV}	Δ_{OF}^{Max}	$\#\Delta_{OF} > 0$	ΔT_{CPU}	
2	2	36618508	207	36628151	2.81	298	0.03 %	31 %	55.56 %	91	
	4	37851698	807	37889824	2.56	934	0.10 %	0.24 %	37.04 %	127	
	8	37851684	989	37889889	2.56	1058	0.10 %	0.19 %	48.15 %	69	
	16	35236347	1123	37852013	2.67	1495	7.34 %	100 %	51.85 %	372	
	32	34317704	2878	37889337	2.67	3195	11.06 %	191 %	44.44 %	317	
4	2	50393582	140	50443988	2.56	224	0.1 %	0.2 %	40.74 %	83	
	4	50395635	1118	50444449	2.44	1208	0.11 %	0.27 %	33.33 %	90	
	8	50394961	713	50444376	2.41	834	0.10 %	0.4 %	22.22 %	121	
	16	41539305	1385	47821928	2.63	1895	23.53 %	368 %	48.15 %	511	
	32	50390947	1555	50444854	3.33	2359	0.11 %	0.22 %	66.67 %	804	
6	2	55022575	103	55208552	2.52	180	0.37 %	7.44 %	48.15 %	77	
	4	55156669	1864	55209391	2.52	1976	0.10 %	0.12 %	44.44 %	112	
	8	55155279	1476	55218894	2.85	2014	0.13 %	0.49 %	48.15 %	539	
	16	53237921	3374	55209938	2.74	4263	3.82 %	100 %	59.26 %	889	
	32	53134432	2831	54822191	2.52	5085	4.98 %	112 %	51.85 %	2254	
8	2	58169769	125	58231711	2.56	171	0.10 %	0.15 %	44.44 %	46	
	4	58173787	1597	58239800	2.44	2094	0.11 %	0.17 %	44.44 %	497	
	8	52303049	3217	55753579	2.41	4148	0.10 %	0.12 %	40.74 %	935	
	16	53237921	3374	55209938	2.33	4247	20.52 %	542 %	29.63 %	873	
	32	45404869	4157	58173208	2.26	4537	69.16 %	911 %	25.93 %	380	

Table 5: Performance results of the local-search procedure defined in Algorithm 1 on the Sioux Falls road network with 24 available candidate toll links, as depicted on Fig. 6.

From left to right of each table, the first two columns report the number of approximation lines ($|\mathcal{L}|$) and intervals of discretization for tolls ($|\mathcal{S}|$). The third and fourth columns contain the information concerning the average network operator profit (OF_{NO}) and average CPU time (T_{CPU}) required to solve problem (13) at the first iteration (initial solution) of the local-search procedure defined in Algorithm 1. The fifth to seventh columns show the results having performed all the iterations of the search (including the computation of the initial solution) in terms of the improved network operator profit (OF_{NO}), the number of performed iterations (Iter) and the total CPU time (T_{CPU}). The final group of columns (the eighth, ninth and tenth columns) provide a summary assessment of the impact of using the binary-search based approach as a local search to improve the obtained solution, by considering the average (Δ_{OF}^{AV}) and the maximum (Δ_{OF}^{Max}) percentage of increase in the network operator profit and the percentage of instances (over the all 27 generated configurations at each combination of $|\mathcal{L}|$ and $|\mathcal{S}|$) for which the local search has been able to improve the initial solution ($\#\Delta_{OF} > 0$).

A first remark on tables 3-5 is the particularly small increase in the CPU time consumption required to

perform the local search on the initial solution. This is due to the availability of valid inequalities based on the optimality bounds provided in Theorem 3.

The importance of applying the proposed local-search procedure to the solution obtained by solving (13) emerges from the analysis of the evolution of the network operator profit, when passing from small to large values of $|\mathcal{S}|$. In fact, while the optimal objective function of (13) does not necessarily increase when incrementing the number of toll levels¹¹, the CPU times increase substantially. Therefore, since the analyst is unaware of the approximate location of optimal toll levels, a wrong selection of S (even if big) might have substantially negative consequence for the quality of the resulting solution, as well as for the computational performance. This is due to the fact that the collection of S toll levels $\mathbf{p}^{1,(1)}, \mathbf{p}^{2,(1)}, \dots, \mathbf{p}^{S,(1)}$ is not entirely contained in the collection of S' toll levels $\mathbf{p}^{1,(1)}, \mathbf{p}^{2,(1)}, \dots, \mathbf{p}^{S',(1)}$, even when $S < S'$. As noticeable from Tables 3-5, the proposed local search method improves significantly the feasible solution of (9a)–(9r) for the vast majority of analyzed instances using $|\mathcal{S}|$ from 16 to 32, and as the number of candidate new toll links increases. The search also allows restoring (in the vast majority of cases) higher OF values when an *unlucky* $|\mathcal{S}|$ is selected by the analyst. This improvement of the initial solution is obtained with a computational cost moving from 25 to 2566 seconds. Frequently, this improvement is obtained between two and three local search iterations.

Beside, a prominent role on the problem size and on the accuracy of the obtained solution is played by the value of $|\mathcal{L}|$. Fig. 7 provides a graphical illustration to support the goodness of the piece-wise linear approximation when $|\mathcal{L}|$ grows large, using the same collection of 1620 problem instances introduced in the beginning of this section.

We observe that the optimal network operator profit gets progressively stable when $|\mathcal{L}|$ increases, for the three configurations of the Sioux Falls road network with 8, 16 and 24 available candidate toll links. This suggests that the piecewise linear approximation estimated using the method of Imamoto & Tang properly mimics the actual travel time function when $|\mathcal{L}|$ is at least equal to 8.

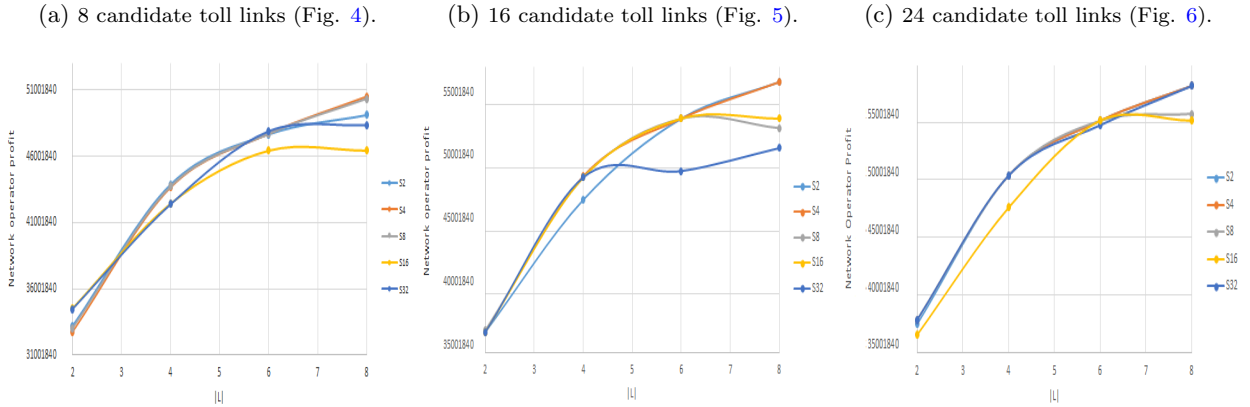


Figure 7: The effect of $|\mathcal{L}|$ on the optimal network operator profit, for each fixed value of $|\mathcal{S}|$.

As discussed in Section 5, the main computational effort in Algorithm 1 relies on the solution of problem (13) at line 7. For this purpose, the optimality bound provided in Theorem 3 establishes a valid inequality to be appended to problem (13) at each iteration of Algorithm 1.

Table 6 shows the effect of the inclusion of this valid inequality on the three study networks. From left to right, it is indicated the pair of number of approximation lines ($|\mathcal{L}|$) and intervals of discretization for tolls ($|\mathcal{S}|$) used, the total CPU times spent by the binary search with and without the inclusion of the valid inequality (T_{CPU}^{VI} and T_{CPU} , respectively), and the corresponding increase of the CPU time spent by the binary search (ΔT_{CPU}) in percentage.

On average, the inclusion of the valid inequality within the local search gives rise to a gap of 24%, 59% and 56%, for the road networks with 8 (Fig. 4), 16 (Fig. 5) and 24 (Fig. 6) available candidate toll links, respectively. In only two cases, the increase in CPU time has been negative, meaning that the total resolution without the usage of the valid inequality has spent less time. It is also worth-noticing that the goodness of the valid inequalities does not depend on the particular selection of $|\mathcal{L}|$ and $|\mathcal{S}|$. However, it seems more

¹¹A monotonic growth would be guaranteed only under the condition that all toll levels are incremented keeping fixed the already defined ones (which was the underlying intuition of Algorithm 1).

profitable as the number of candidate new toll links increases.

Scenario		8 candidate toll links			16 candidate toll links			24 candidate toll links		
$ \mathcal{L} $	$ \mathcal{S} $	T_{CPU}	T_{CPU}^{VI}	ΔT_{CPU}	T_{CPU}	T_{CPU}^{VI}	ΔT_{CPU}	T_{CPU}	T_{CPU}^{VI}	ΔT_{CPU}
2	2	175	241	-38 %	1068	278	74 %	2398	298	88 %
	4	151	138	9 %	1997	755	62 %	2463	934	62 %
	8	723	355	51 %	2035	426	79 %	2197	1058	52 %
	16	750	612	18 %	2627	1289	51 %	2037	1495	27 %
	32	2151	1512	30 %	3264	1362	58 %	3621	3195	12 %
4	2	561	444	21 %	1465	214	85 %	2116	224	89 %
	4	508	470	7 %	2966	1237	58 %	3949	1208	69 %
	8	773	324	58 %	3252	1637	50 %	2937	834	72 %
	16	951	760	20 %	3432	1621	53 %	5231	1895	64 %
	32	1513	562	63 %	4058	1714	58 %	3110	2359	24 %
6	2	963	945	2 %	3884	182	95 %	2254	180	92 %
	4	514	492	4 %	8303	1256	85 %	5119	1976	61 %
	8	1441	549	62 %	3720	1311	65 %	4527	2014	56 %
	16	1192	930	22 %	9231	6030	35 %	5487	4263	22 %
	32	963	543	44 %	10420	5688	45 %	3772	5085	-35 %
8	2	2017	1983	2 %	2525	176	93 %	8948	171	98 %
	4	1558	831	47 %	5348	2855	47 %	4646	2094	55 %
	8	1603	1100	31 %	7321	6211	15 %	8284	4148	100 %
	16	2849	1279	55 %	7600	7444	2 %	8824	4247	52 %
	32	2355	2355	0 %	10758	7641	29 %	13875	4537	67 %

Table 6: The effect of the valid inequality on the CPU time for each pair of values of $|\mathcal{L}|$ and $|\mathcal{S}|$.

Finally, to better explore the effect of the number of discretized toll levels and number of pivots of the PWL approximation on the performance of the proposed solution mechanism, a statistical analysis on the computational experiment is carried out. Tables 7 and 8 report the results of two regression analysis for the increase in the network operator profit and the number of improved solutions having applied the local search procedure.

Covariate	Estimate	Std. Error	Pr(> t)
Intercept	-5.772	10.147	0.569
$ \mathcal{L} $	1.373	1.852	0.458
$ \mathcal{S} $	1.200	0.614	0.049 *
$ \mathcal{A}^T $	0.420	0.587	0.474
$ \mathcal{L} : \mathcal{S} $	-0.288	0.112	0.010 **
$ \mathcal{L} : \mathcal{A}^T $	-0.115	0.107	0.281
$ \mathcal{S} : \mathcal{A}^T $	-0.067	0.035	0.056 .
$ \mathcal{S} : \mathcal{L} : \mathcal{A}^T $	0.021	0.006	0.001 ***

Table 7: Regression model for the increase in the network operator profit, having applied the local search procedure. The notation ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

From Table 7, it is important to notice that individually, only $|\mathcal{S}|$ has an impact on the increase of the network operator profit when applying the local search on the initial solution of (13). Conceptually, a better representation of toll levels allows a better profit for the network operator. The interactions also play a role on the improvement of this profit, particularly, $|\mathcal{L}| : |\mathcal{S}|$ (the interaction between $|\mathcal{L}|$ and $|\mathcal{S}|$) and $|\mathcal{S}| : |\mathcal{L}| : |\mathcal{A}^T|$ (the interaction among $|\mathcal{L}|$, $|\mathcal{S}|$ and $|\mathcal{A}^T|$) are both significant factors.

Covariate	Estimate	Std. Error	Pr(> t)
Intercept	1.002	0.493	0.042 *
$ \mathcal{L} $	-0.085	0.089	0.341
$ \mathcal{S} $	-0.043	0.029	0.149
$ \mathcal{A}^T $	-0.071	0.028	0.013 *
$ \mathcal{L} : \mathcal{S} $	0.003	0.005	0.609
$ \mathcal{L} : \mathcal{A}^T $	0.005	0.005	0.338
$ \mathcal{S} : \mathcal{A}^T $	0.003	0.001	0.041 *
$ \mathcal{S} : \mathcal{L} : \mathcal{A}^T $	-0.000	0.000	0.235

Table 8: Regression model for the number of improved solutions having applied the local search procedure. The notation ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

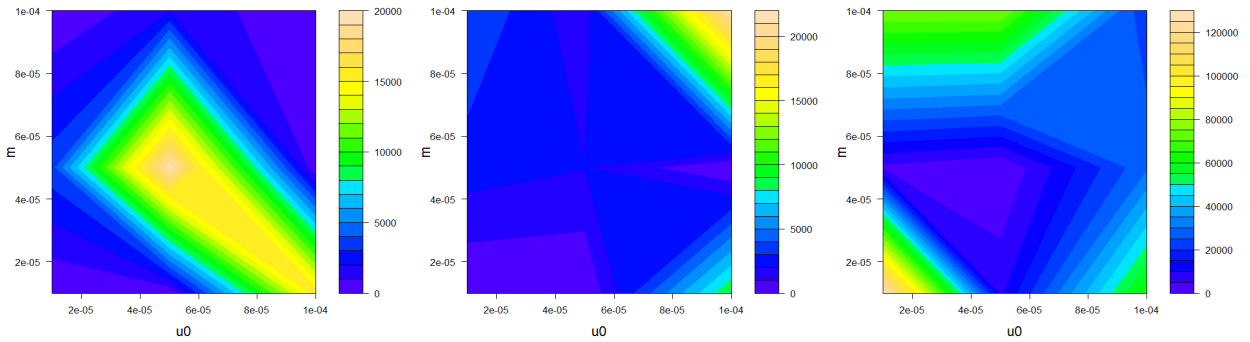
Concerning the number of improved solutions, the effect of $|\mathcal{L}|$, $|\mathcal{S}|$ and $|\mathcal{A}^T|$ is not statistically evident. In fact, as reported in Table 8, $|\mathcal{L}|$ and $|\mathcal{S}|$ seem to play a negative role, though the p-values suggest no significant impact. However, the number of toll-links ($|\mathcal{A}^T|$) has a significant negative effect on the probability that the local search improves the initial solution when $|\mathcal{S}|$ is small, whereas the effect of $|\mathcal{A}^T|$ becomes positive when $|\mathcal{S}|$ is large.

7.2. Traffic management sensitivity

To gain a conceptual understanding of the combinatorial nature of the RNPRE problem, Example 1 has considered a stylized case, showing how the optimal operator choices vary under different exogenous business settings. To better assess the effect of these conditions on the road links creation and travel time, a parameter sensitivity analysis is carried out in this subsection.

Building on the solution procedure applied in the previous subsection, the contour plots in Fig. 8 show the total construction expenditure (top panels) and the travel time (bottom panels) for nine combinations of m (monetary cost of the exposure to risk) and u^0 (monetary cost of congestion for each individual), averaged over three levels of B (investment budget). Each panel corresponds to a different level of ρ (homogeneous discount for the construction costs with respect to the initial costs presented in Section 6). Panel (a) and (d) are associated with $\rho = 0.25$, panel (b) and (e) are associated with $\rho = 0.5$, and panel (c) and (f) are associated with $\rho = 0.75$.

(a) 25% reduction (construct. cost). (b) 50% reduction (construct. cost). (c) 75% reduction (construct. cost).



(d) 25% reduction (travel time). (e) 50% reduction (travel time). (f) 75% reduction (travel time).

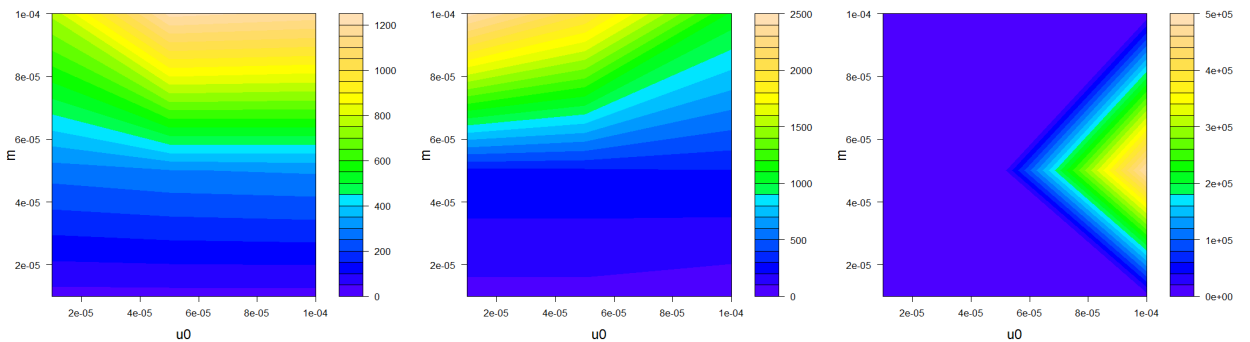


Figure 8: The effect of m , u^0 and ρ on the optimal construction expenditure and travel time for ordinary vehicles, for each fixed value of $|\mathcal{S}|$.

Fig. 8 reveals a complex interaction between m , u^0 , B and ρ . When the construction prices are sufficiently high (i.e. $\rho = 0.25$ and $\rho = 0.5$), the travel time increases with m . This is possibly due to the fact that the network operator pay a bigger effort in designing roads to minimize the exposure to risk when its monetary value is high, with a negative impact on the travel time for ordinary vehicles. Regarding the investment in new roads, the optimal expenditure has a complex dependency patterns with respect to m , u^0 and ρ . This argument is supported by the statistical analysis reported in tables 9 and 10.

Covariate	Estimate	Std. Error	Pr(> t)	
Intercept	-3.490e+04	2.508e+04	0.168	
m	3.075e+08	2.815e+08	0.278	
u^0	-1.970e+07	2.815e+08	0.944	
B	-7.734e-02	2.418e-01	0.750	
ρ	1.169e+03	2.418e-01	0.006	**
$m : u^0$	-3.378e+11	2.260e+12	0.882	
$m : B$	3.072e+02	1.963e+03	0.876	
$m : \rho$	-7.533e+06	3.177e+06	0.049	*
$u^0 : B$	5.839e+02	1.963e+03	0.767	
$u^0 : \rho$	-4.055e+05	4.077e+06	0.921	
$B : \rho$	9.643e-04	3.540e-03	0.786	

Table 9: Regression model for the construction expenditure. The notation ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Covariate	Estimate	Std. Error	Pr(> t)	
Intercept	1.887e+04	7.888e+04	0.811	
m	-1.157e+09	8.853e+08	0.195	
u^0	2.127e+08	8.853e+08	0.810	
B	5.387e-03	7.605e-01	0.994	
$m : u^0$	-1.656e+12	7.108e+12	0.816	
ρ	-7.198e+02	1.296e+03	0.580	
$m : B$	-2.382e+01	6.173e+03	0.996	
$m : \rho$	3.740e+07	1.282e+07	0.004	**
$u^0 : B$	-4.453e+01	6.173e+03	0.994	
$u^0 : \rho$	-3.210e+06	1.282e+07	0.803	
$B : \rho$	-5.989e-05	1.113e-02	0.995	

Table 10: Regression model for the travel time. The notation ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

The fixed construction costs seem to be the most evident driving factor to bust the construction expenditure (estimate: $1.169e + 03$, p-value: 0.006), in particular when the monetary value of the risk exposure is low. However, when the network operator strongly internalizes the externality costs from risk exposure (i.e. large values of m), the total expenditure in construction might decrease with construction costs (estimate: $-7.533e + 06$, p-value: 0.049). Again, the positive interaction $m : \rho$ suggests that the travel time increases with m , when the construction costs are sufficiently high (estimate: $3.740e + 07$, p-value: 0.004).

The overall picture from the analysis in Fig. 8 and tables 9 and 10 is the one of a complex interaction patterns between the exogenous economic conditions and the global outcomes, emerging from the multiple operator choices over road network pricing, regulation and expansion.

8. Conclusions and Further Research

This work presents a mixed-integer non-linear bi-level problem to simultaneously determine the road network pricing, regulation and expansion for ordinary and hazmat vehicles under traffic congestion. This enables the network operator (corresponding to the leader) to consider the most significant trade-offs in the investment on new roads, such as the flow of ordinary and hazmat vehicles (which result from the follower best response to the operator choices), charging different tolls for each road according to the overall impact on the entailed risk level, while restricting the use of network links to the dangerous goods shipments. Therefore, building on this comprehensive traffic management policy, the proposed modeling framework allows reaching a balance between the profit of the leader, the travel costs of the follower, and the externality costs induced by the exposure to hazmat risk.

As described in Section 4, we follow a reformulation and linearization strategy in which the designed MINBLP is approximated by a MILP with arbitrary precision. This is done in three steps. First, the travel time function of the vehicles is approached using a piecewise linear function based on the method of Imamoto & Tang [26]. Second, the linearized followers' problem is replaced with a set of constraints and additional variables based on the strong duality theorem. Finally, tolls are discretized using binary variables denoting the toll levels.

Exploiting the proposed linearization strategy, the approximated MILP is solved within a specialized local search, which sequentially improves the toll levels found by the initial resolution of the MILP. This is built on the idea of a binary-search-based procedure, that sequentially shrinks the toll interval in which the toll levels are contained.

The reported results on three network variants of the Sioux Falls road network show that the local search can improve the leader's profit up to a 502 % while not spending more than 30 minutes. On the contrary, when those improvements are not significant, the CPU times are negligible. The level of improvement seems to depend on the number of toll levels. In fact, the network profits grows significantly for values of $|\mathcal{S}|$ between 16 and 32, and as the number of candidate new toll links increases. The inclusion of the valid inequality reduces significantly the CPU times of the binary search, being able to obtain the same toll values up to 98% less of CPU time.

To the best of our knowledge, the integration of road network pricing, regulation and expansion for multiple vehicles categories has not been addressed in previous research. In proposing this novel modeling framework, this work has pay attention in reformulation and linearization strategy, with a view to providing an efficient resolution mechanism to help road network operators to take informs decision.

Regarding further research, the following lines are proposed: 1) study of the demand by time periods [20], 2) demand uncertainty [10], 3) robust network design [7], 4) multiple leaders [45], incorporating congestion and vehicles into hazmat risk [40, 18], and 5) [development of tailored bounding methods for big-M's \[33\]](#). In all lines, further steps into the modeling integrations can still be made, placing this work at the beginning of a research line aiming at providing road network operators with the ability to take simultaneously decision over the complex patterns of the traffic management.

Acknowledgements

The authors would like to express their gratitude to Marijn Verschelde, who shared with us his view and constructive comments. We are also thankful to the suggestions and recommendations from the participants at the 2018 International Workshop on Bi-level Programming, who provided rich insights that helped improve the paper.

References

- [1] A.S. Ahuja. *Development of Passenger Car Equivalents for Freeway Merging Sections*. University of Nevada, Las Vegas, 2007.
- [2] Saeed Asadi Bagloee, Majid Sarvi, and Michael Patriksson. A hybrid branch-and-bound and benders decomposition algorithm for the network design problem. *Computer-Aided Civil and Infrastructure Engineering*, 00:1–25, 2016.
- [3] Jonathan F Bard. Optimality conditions for the bilevel programming problem. *Naval Research Logistics (NRL)*, 31(1):13–26, 1984.
- [4] Francisco Benita, Francisco López-Ramos, and Stefano Nasini. A bi-level programming approach for global investment strategies with financial intermediation. *European Journal of Operational Research*, 274(1):375–390, 2019.
- [5] L Bianco, M Caramia, and S Giordani. A bilevel flow model for hazmat transportation network design. *Transportation Research Part C*, 17:175–196, 2009.
- [6] Mustapha Bouhtou, Stan van Hoesel, Anton F van der Kraaij, and Jean-Luc Lutton. Tariff optimization in networks. *INFORMS journal on computing*, 19(3):458–469, 2007.
- [7] Luis Cadarso, Esteve Codina, Escudero Laureano F., and Ángel Marín. Rapid transit network design: considering recovery robustness and risk aversion measure. *Transportation Research Procedia*, 22:255–264, 2017.
- [8] Giuseppe C Calafiore and Laurent El Ghaoui. *Optimization models*. Cambridge university press, 2014.
- [9] Wilfred Candler and Robert Townsley. A linear two-level programming problem. *Computers & Operations Research*, 9:59–76, 1982.

- [10] Suh-Wen Chiou. A risk-averse signal setting policy for regulating hazardous material transportation under uncertain travel demand. *Transportation Research Part D*, 50:446–472, 2017.
- [11] Stephan Dempe. *Foundations of bilevel programming*. Springer Science & Business Media, 2002.
- [12] Stephan Dempe and Joydeep Dutta. Is bilevel programming a special case of a mathematical program with complementarity constraints? *Mathematical programming*, 131(1-2):37–48, 2012.
- [13] Stephan Dempe, Vyacheslav Kalashnikov, Gerardo A Pérez-Valdés, and Nataliya Kalashnykova. *Bilevel programming problems: Theory, algorithms and applications to energy networks*. Springer, 2015.
- [14] Sophie Dewez, Martine Labbé, Patrice Marcotte, and Gilles Savard. New formulations and valid inequalities for a bilevel pricing problem. *Operations research letters*, 36(2):141–149, 2008.
- [15] J Elström, Sumalee Agachai, and HK Lo. Optimizing toll locations and levels using a mixed integer linear approximation approach. *Transportation Research Part B*, 46:834–854, 2012.
- [16] E Erkut and O Alp. Designing a road network for hazardous materials shipments. *Computers & Operations Research*, 34:1389–1405, 2007.
- [17] E Erkut and F Gzara. Solving the hazmat transport network design problem. *Computers & Operations Research*, 35:2234–2247, 2008.
- [18] T Esfandeh, C Kwon, and R Batta. Regulating hazardous materials transportation by dual toll pricing. *Transportation Research Part B*, 83:20–35, 2016.
- [19] Pirmin Fontaine and Stefan Minner. Benders decomposition for discrete–continuous linear bilevel problems with application to traffic network design. *Transportation Research Part B: Methodological*, 70:163–172, 2014.
- [20] Pirmin Fontaine and Stefan Minner. A dynamic discrete network design problem for maintenance planning in traffic networks. *Annals of Operations Research*, 253:757–772, 2017.
- [21] Pirmin Fontaine and Stefan Minner. Benders decomposition for the hazmat transport network design problem. *European Journal of Operational Research*, 267:996–1002, 2018.
- [22] Robert Fourer. Strategies for not linear optimization. *5th INFORMS Optimization Society Conference*, Session TA-01:6–8 March, 2014.
- [23] Z Gao, J Wu, and Sun H. Solution algorithm for the bi-level discrete network design problem. *Transportation Research Part B*, 39:479–495, 2005.
- [24] Pablo Garcia-Herreros, Lei Zhang, Pratik Misra, Erdem Arslan, Sanjay Mehta, and Ignacio E Grossmann. Mixed-integer bilevel optimization for capacity planning with rational markets. *Computers & Chemical Engineering*, 86:33–47, 2016.
- [25] Patrick T Harker and Jong-Shi Pang. Existence of optimal solutions to mathematical programs with equilibrium constraints. *Operations research letters*, 7(2):61–64, 1988.
- [26] A Imamoto and B Tang. Optimal piecewise linear approximation of convex functions. *In: Proceedings of the World Congress on Engineering and Computer Science*, Citeseer:1191–1194, 2008.
- [27] Aimé Kamgaing-Kuiteing, Patrice Marcotte, and Gilles Savard. Network pricing of congestion-free networks: The elastic and linear demand case. *Transportation Science*, 51:791–1029, 2017.
- [28] Bahar Y Kara and Vedat Verter. Designing a road network for hazardous materials transportation. *Transportation Science*, 38:188–196, 2004.
- [29] Martine Labbé, Patrice Marcotte, and Gilles Savard. A bilevel model of taxation and its application to optimal highway pricing. *Management Science*, 44:1608–1622, 1998.
- [30] Martine Labbé and Alessia Violin. Bilevel programming and price setting problems. *4OR*, 11(1):1–30, 2013.

- [31] Dung-Ying Lin and Pak-Weng Leong. An n-path user equilibrium for transportation networks. *Applied Mathematical Modelling*, 38(2):667–682, 2014.
- [32] Zhi-Quan Luo, Jong-Shi Pang, and Daniel Ralph. *Mathematical programs with equilibrium constraints*. Cambridge University Press, 1996.
- [33] P Marcotte, A Mercier, G Savard, and Verter V. Toll policies for mitigating hazardous materials transport risk. *Transportation Science*, 43:228–243, 2009.
- [34] Bureau of Public Roads. Traffic assignment manual. *U.S. Department of Commerce, Urban Planning Division. Washington, DC*, 1964.
- [35] Yosef Sheffi. *Urban transportation networks*, volume 6. Prentice-Hall, Englewood Cliffs, NJ, 1985.
- [36] Ben Stabler, Hillel Bar-Gera, and Elizabeth Sall. Transportation networks for research core team. <https://github.com/bstabler/TransportationNetworks>, 2018. Online; accessed 09 February 2018.
- [37] Longsheng Sun, Mark H Karwan, and Changhyun Kwon. Robust hazmat network design problems considering risk uncertainty. *Transportation Science*, 50:1188–1203, 2016.
- [38] Chaisak Suwansirikul, Terry Friesz L, and Roger Tobin L. Equilibrium decomposed optimization: A heuristic for the continuous equilibrium network design problem. *Transportation Science*, 21:254–263, 1987.
- [39] H Wang and X Zhang. Game theoretical transportation network design among multiple regions. *Annals of Operational Research*, 249:97–117, 2017.
- [40] J Wang, Y Kang, C Kwon, and R Batta. Dual toll pricing for hazardous materials transport with linear delay. *Networks and Spatial Economics*, 12:147–165, 2012.
- [41] John Glen Wardrop. Some theoretical aspects of road traffic research. In *Inst Civil Engineers Proc London/UK/*, 1952.
- [42] Ue-Pyng Wen and YH Yang. Algorithms for solving the mixed integer two-level linear programming problem. *Computers & Operations Research*, 17(2):133–142, 1990.
- [43] Jr Williams and F Louis. A modification to the half-interval search (binary search) method. In *Proceedings of the 14th Annual Southeast Regional Conference*, ACM-SE 14, pages 95–101, New York, NY, USA, 1976. ACM.
- [44] Meng Xu, Guangmin Wang, Susan Grant-Muller, and Ziyou Gao. Joint road toll pricing and capacity development in discrete transport network design problem. *Transportation*, 44:731–752, 2017.
- [45] H Yang, F Xiao, and Huang H. Private road competition and equilibrium with traffic equilibrium constraints. *Journal of Advanced Transportation*, 43:21–45, 2008.
- [46] JJ Ye, DL Zhu, and QJ Zhu. Exact penalization and necessary optimality conditions for generalized bilevel programming problems. *SIAM Journal on Optimization*, 7(2):481–507, 1997.
- [47] M Hosein Zare, Juan S Borrero, Bo Zeng, and Oleg A Prokopyev. A note on linearized reformulations for a class of bilevel linear integer problems. *Annals of Operations Research*, pages 1–19, 2017.

Appendix A: On the application of the N -path user equilibrium

This appendix reviews the N -path user equilibrium model developed in [31] and applies it to the follower problem (3). Using a similar notation as in [35], we can explicitly define from the incidence structure of matrix N a collection of paths \mathcal{P}^k connecting each origin-destination pair $k \in \mathcal{K}$. Beside, we define two indicators $\Delta_{a,p}^k$ (for $a \in \mathcal{A}$, $k \in \mathcal{K}$, $p \in \mathcal{P}^k$) and $\widehat{\Delta}_{i,p}^k$ (for $i \in \mathcal{N}$, $k \in \mathcal{K}$, $p \in \mathcal{P}^k$). The first indicator verifies $\Delta_{a,p}^k = 1$ if the arc $a \in \mathcal{A}$ belongs to the path $p \in \mathcal{P}^k$ connecting the origin-destination pair $k \in \mathcal{K}$, and $\Delta_{a,p}^k = 0$ otherwise. The second indicator verifies $\widehat{\Delta}_{i,p}^k = 1$ if the node $i \in \mathcal{N}$ belongs to the path $p \in \mathcal{P}^k$ connecting the origin-destination pair $k \in \mathcal{K}$, and $\widehat{\Delta}_{i,p}^k = 0$ otherwise.

Let $f_{v,p}^k$ be the flow of path $p \in \mathcal{P}^k$ for user type $v \in \mathcal{V}$, between the origin-destination pair $k \in \mathcal{K}^v$ and note that $x_a^{v,k} = \sum_{p \in \mathcal{P}^k} f_{v,p}^k \Delta_{a,p}^k$. Then, the path-based balance conditions and the arc-based balance conditions can be equivalently formulated as¹²:

$$\sum_{p \in \mathcal{P}^k} f_{v,p}^k = \sum_{i \in \mathcal{N}} \sum_{p \in \mathcal{P}^k} s_i^{v,k} \widehat{\Delta}_{i,p}^k, \quad (\forall v \in \mathcal{V}, k \in \mathcal{K}^v) \quad \text{and} \quad \sum_{a \in \mathcal{A}} N_{i,a} x_a^{v,k} = s_i^{v,k} \quad (\forall i \in \mathcal{N}, v \in \mathcal{V}, k \in \mathcal{K}^v).$$

Similarly, the arc interdiction constraints can be equivalently formulated as

$$(1 - w_a^v) \sum_{p \in \mathcal{P}^k} f_{v,p}^k \Delta_{a,p}^k = 0, \quad (\forall v \in \mathcal{V}, k \in \mathcal{K}^v, a \in \mathcal{A}^T) \quad \text{and} \quad x_a^v \leq M w_a^v \quad (\forall a \in \mathcal{A}^T, v \in \mathcal{V}).$$

Then, the follower problem (3) can be equivalently written as

$$\min_{\mathbf{x}} \quad \sum_{a \in \mathcal{A}} \int_0^{\sum_{v \in \mathcal{V}} \sum_{k \in \mathcal{K}^v} \sum_{p \in \mathcal{P}^k} f_{v,p}^k \Delta_{a,p}^k} P_a^V(t_a, \omega) d\omega \quad (.1a)$$

$$\text{subject to:} \quad \sum_{p \in \mathcal{P}^k} f_{v,p}^k = \sum_{i \in \mathcal{N}} \sum_{p \in \mathcal{P}^k} s_i^{v,k} \widehat{\Delta}_{i,p}^k, \quad [\widehat{\alpha}_v^k] \quad \forall v \in \mathcal{V}, k \in \mathcal{K}^v \quad (.1b)$$

$$(1 - w_a^v) \sum_{p \in \mathcal{P}^k} f_{v,p}^k \Delta_{a,p}^k = 0 \quad [\widehat{\beta}_{av}^k] \quad \forall v \in \mathcal{V}, k \in \mathcal{K}^v, a \in \mathcal{A}^T \quad (.1c)$$

$$f_{v,p}^k \geq 0 \quad \forall v \in \mathcal{V}, k \in \mathcal{K}^v, p \in \mathcal{P}^k \quad (.1d)$$

Constraints (.1c) extend the traditional user equilibrium model by regulating the access of each class of users to certain arcs. Note that they differ from the path-availability constraints introduced in [31], since problem (3) requires arc interdiction instead of path interdiction. However, as stated by the next proposition, (.1a)–(.1d) still verifies the conventional perspective of user equilibrium in which no road user can unilaterally make route changes to improve her/his travel time.

Proposition 4. Consider the travel time along each path, which is computed as

$$c_p^k = \sum_{a \in \mathcal{A}} P_a^V(t_a, x_a) \Delta_{a,p}^k, \quad \forall p \in \mathcal{P}^k, k \in \mathcal{K}$$

We claim that the optimality conditions of problem (.1a)–(.1d) verify the N -path user equilibrium conditions:

¹²See [6, 14, 27] for a more detailed comparison between the alternative use of path-based formulations and arc-based formulation.

$$\begin{aligned}
(a) \quad & f_{v,p}^k \left[c_p^k - \left(\widehat{\alpha}_v^k + (1 - w_a^v) \widehat{\beta}_{bv}^k \right) \right] = 0 & \forall v \in \mathcal{V}, k \in \mathcal{K}^v, p \in \mathcal{P}^k \\
(b) \quad & c_p^k - \left(\widehat{\alpha}_v^k + (1 - w_a^v) \widehat{\beta}_{bv}^k \right) \geq 0 & \forall v \in \mathcal{V}, k \in \mathcal{K}^v, p \in \mathcal{P}^k \\
(c) \quad & \sum_{p \in \mathcal{P}^k} f_{v,p}^k = \sum_{i \in \mathcal{N}} \sum_{p \in \mathcal{P}^k} s_i^{v,k} \widehat{\Delta}_{i,p}^k & \forall v \in \mathcal{V}, k \in \mathcal{K}^v \\
(d) \quad & (1 - w_a^v) \sum_{p \in \mathcal{P}^k} f_{v,p}^k \Delta_{a,p}^k = 0 & \forall v \in \mathcal{V}, k \in \mathcal{K}^v, a \in \mathcal{A}^T \\
(e) \quad & f_{v,p}^k \geq 0 & \forall v \in \mathcal{V}, k \in \mathcal{K}^v, p \in \mathcal{P}^k
\end{aligned}$$

A proof of this proposition can be found on [Appendix C](#). Note that for each vehicle of type $v \in \mathcal{V}$, conditions (a) and (b) are valid for every path $p \in \mathcal{P}^k$ connecting the origin-destination pair $k \in \mathcal{K}^v$. They imply that either the flow on that specific path is zero (in which case the travel time on that path c_p^k shouldn't be smaller than the total shadow price $\widehat{\alpha}_v^k + (1 - w_a^v) \widehat{\beta}_{bv}^k$), or the flow on that path is positive (in which case the travel time on that path c_p^k is equal to the total shadow price $\widehat{\alpha}_v^k + (1 - w_a^v) \widehat{\beta}_{bv}^k$).

In both cases, the total shadow price corresponds to the minimum path travel time between the analyzed origin-destination pairs. This is consistent with the user equilibrium principle; in fact, there are two types of paths connecting origin-destination pairs: those on which the travel time is equal to the total shadow price, and those on which the travel time is not smaller to the total shadow price. Thus, a vehicle cannot improve its travel time by unilaterally changing routes.

Appendix B: Imamoto & Tang's method

The approximation lines are determined using the method of Imamoto & Tang's [26] because it provides the less approximation average error. Next Algorithm 2 shows the pseudo-code of this method, with input parameters α_0 and $\alpha_{|\mathcal{L}|}$ specifying the minimum and maximum values for the total link flow x_a associated with a road link $a \in A$, and ϵ denoting its maximum approximation average error.

Algorithm 2 Pseudo-code of Imamoto & Tang's [26] method for computing the approximation lines.

1: Input: $u_a, \mathcal{L}, \alpha_0, \alpha_{|\mathcal{L}|}, \epsilon$; Output: d_a, e_a ;
2: Initialize counter: $j = 1$;
3: Initialize correction step: $\Delta^{(j)} = 1$;
4: Initialize pivots: $\delta_l^{(j)} = \alpha_0 + \frac{l-1.5}{|\mathcal{L}|}(\alpha_{|\mathcal{L}|} - \alpha_0), \forall l \in \mathcal{L}$;
5: **repeat**
6: $\alpha_0^{(j)} = \alpha_0$;
7: $\alpha_{|\mathcal{L}|}^{(j)} = \alpha_{|\mathcal{L}|}$;
8: **for all** $l \in \mathcal{L}$ **do**
9: $\alpha_l^{(j)} = \left(u_a(\delta_{l-1}^{(j)}) - u_a(\delta_l^{(j)}) + \delta_l^{(j)} u'_a(\delta_l^{(j)}) - \delta_{l-1}^{(j)} u'_a(\delta_{l-1}^{(j)}) \right) \left(u'_a(\delta_l^{(j)}) - u'_a(\delta_{l-1}^{(j)}) \right)^{-1}$;
10: $\tilde{u}_l^{(j)} = u'_a(\delta_{l-1}^{(j)})(\alpha_l^{(j)} - \delta_{l-1}^{(j)}) + u_a(\delta_{l-1}^{(j)})$;
11: $\phi_l^{(j)} = \tilde{u}_l^{(j)} - u_a(\alpha_l^{(j)})$;
12: **end for**
13: $\phi_{max}^{(j)} = \max_{l \in \mathcal{L}} |\phi_l^{(j)}|$;
14: $\phi_{min}^{(j)} = \min_{l \in \mathcal{L}} |\phi_l^{(j)}|$;
15:
16: **if** $\phi_{max}^{(j)} > \phi_{max}^{(j-1)}$ **then**
17: $j --$;
18: $\Delta^{(j)} = \frac{\Delta^{(j+1)}}{2}$;
19:
20: **end if**
21: **for all** $l \in \mathcal{L}$ **do**
22: $\delta_l^{(j+1)} = \delta_l^{(j)} + \Delta^{(j)} \left(\phi_{l+1}^{(j)} - \phi_l^{(j)} \right) \left(\frac{\phi_{l+1}^{(j)}}{\alpha_{l+1}^{(j)} - \delta_l^{(j)}} + \frac{\phi_l^{(j)}}{\delta_l^{(j)} - \alpha_l^{(j)}} \right)^{-1}$;
23: **end for**
24: $j ++$;
25: **until** $\left| \frac{\phi_{max}^{(j)}}{\phi_{min}^{(j)}} - 1 \right| < \epsilon$
26: **for all** $l \in \mathcal{L}$ **do**
27: $d_a^l = u'_a(\delta_l^{(j)})$;
28: $e_a^l = u_a(\delta_l^{(j)}) - \delta_l^{(j)} u'_a(\delta_l^{(j)})$;
29: **end for**

Proposition 1.

Proof. The proof consists in showing first the existence of a feasible solution and then the existence of an optimal one.

First, for any mapping $\Phi : \mathbb{R} \rightarrow \mathbb{R}$, we define its graph as $\text{gph}\Phi := \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m \mid y \in \Phi(x)\}$. Then, as long as the network of toll-free links \mathcal{A}^F is connected, we note that for any $(\mathbf{t}, \mathbf{y}, \mathbf{w}) \in \Phi$, the follower's feasible region (defined by (3b)-(3d)) is a non-empty compact polytope. Thus, as the $\Psi(\mathbf{t}, \mathbf{w})$ is also non-empty, the MINLBP (1) possesses at least one feasible solution (i.e., the inducible region is non-empty).

Second, for a fixed \mathbf{w} we introduce the set-valued mapping $\tilde{\Psi}_P^{\mathbf{w}}(\mathbf{t}) := \tilde{\Psi}_P(\mathbf{t}, \mathbf{w})$ for any $\mathbf{t} \in [\mathbf{0}, \mathbf{T}]$. It is well-known from the theory of linear parametric optimization that $\text{gph}\Psi_{\Theta}^{\mathbf{z}} := \{(\mathbf{t}, \mathbf{x}) \mid \mathbf{t} \in [\mathbf{0}, \mathbf{T}], \mathbf{x} \in \tilde{\Psi}_P^{\mathbf{w}}(\mathbf{t})\}$ is the union of finitely many polyhedral sets and, thus, closed. Consequently, based on Theorem 3.1 in [11], the set

$$\text{gph}\Psi = \{(\mathbf{x}, \mathbf{t}, \mathbf{w}) \in \Phi \mid \mathbf{x} \in \tilde{\Psi}_P(\mathbf{t}, \mathbf{w})\} = \bigcup_{\mathbf{w} \in \{0,1\}^q} \{(\mathbf{x}, \mathbf{t}, \mathbf{w}) \in \Phi \mid (\mathbf{x}, \mathbf{w}) \in \text{gph}\tilde{\Psi}_P^{\mathbf{w}}\}$$

is closed (as it is the union of finitely many polyhedral sets). Then, the inducible region $IR = \{(\mathbf{t}, \mathbf{y}, \mathbf{w}) \in \Phi \mid (\mathbf{x}, \mathbf{w}) \in \text{gph}\tilde{\Psi}_P^{\mathbf{w}}\}$ is closed and bounded. Therefore, if we combine the above arguments, we have that IR is non-empty and compact. Next to it, by means of the Weierstrass theorem, as the leader's objective is continuous with respect to the continuous variables \mathbf{t} and \mathbf{x} , the linearized network operator problem (1) (in which $\mathbf{x} \in \Psi(\mathbf{t}, \mathbf{w})$ is replaced with $\mathbf{x} \in \tilde{\Psi}_P(\mathbf{t}, \mathbf{w})$) admits a global optimal solution for any L . \square

Lemma 2.

Proof. At the j^{th} iteration of Algorithm 1, we define the ordered sets of points $P_0^{(j)} = \{\mathbf{p}^{1,(j)}, \mathbf{p}^{2,(j)}, \dots, \mathbf{p}^{[\lceil |\mathcal{S}|/2 \rceil, (j)}\}$ and $P_1^{(j)} = \{\mathbf{p}^{[\lceil |\mathcal{S}|/2 \rceil + 1, (j)}, \mathbf{p}^{[\lceil |\mathcal{S}|/2 \rceil + 2, (j)}, \dots, \mathbf{p}^{|\mathcal{S}|, (j)}\}$ as the collection of toll levels contained in the first and second half of the toll interval, respectively, so that $P^{(j)} = P_0^{(j)} \cup P_1^{(j)}$. Once $\hat{\mathbf{p}}^{(j)}$ is known, points in $P^{(j+1)}$ are generated as follows:

$$\mathbf{p}^{s, (j+1)} = L(j) + \mathbf{T} \frac{s-1}{2j(|\mathcal{S}|-1)}, \quad \text{where } L(j) = \begin{cases} \mathbf{p}^{1, (j)} + \mathbf{T} & \text{if } \hat{\mathbf{p}}^{(j)} \in P_0^{(j)} \\ \frac{1}{2} (\mathbf{p}^{1, (j)} + \mathbf{p}^{|\mathcal{S}|, (j)}) & \text{if } \hat{\mathbf{p}}^{(j)} \in P_1^{(j)} \end{cases}$$

Note that the toll levels $P^{(j)}$ are homogeneously distributed, by construction, within the interval $[\mathbf{p}^{1, (j)}, \mathbf{p}^{|\mathcal{S}|, (j)}]$, and independently of whether $\hat{\mathbf{p}}^{(j)} \in P_0^{(j)}$ or $\hat{\mathbf{p}}^{(j)} \in P_1^{(j)}$. So, the pointwise distances between successive toll levels are $D(j) = \mathbf{T} \frac{1}{2j(|\mathcal{S}|-1)}$. Moreover, since either $\hat{\mathbf{p}}^{(j)} \in P_0^{(j)}$ or $\hat{\mathbf{p}}^{(j)} \in P_1^{(j)}$, we have that $\hat{\mathbf{p}}^{(j)}$ can be expressed as $L(j+1) + 2kD(j)$ for some $k = 1 \dots (|\mathcal{S}|-1)$. Consequently, $\hat{\mathbf{p}}^{(j)} \in P^{(j+1)}$. \square

Theorem 3.

Proof. Let $\chi^{(j)}$ be the feasible region of problem (13) at the j^{th} iteration of Algorithm 1, defined in terms of variables $\mathbf{t}, \mathbf{w}, \mathbf{y}, \boldsymbol{\varphi}, \mathbf{x}, \hat{\mathbf{u}}, \boldsymbol{\alpha}, \boldsymbol{\beta}$, and $\boldsymbol{\gamma}$ (by projecting \mathbf{z} into \mathbf{t}), and define $\hat{\mathbf{v}}^{(j)}, \hat{\mathbf{p}}^{(j)}$ and $\Upsilon^{(j)}$ as its corresponding vector of optimal variables, vector of optimal toll levels and optimal objective value, respectively. We also consider the collection of points $\{\mathbf{p}^{1, (j)}, \mathbf{p}^{2, (j)}, \dots, \mathbf{p}^{|\mathcal{S}|, (j)}\}$ associated with the search interval $[\mathbf{p}^{1, (j)}, \mathbf{p}^{|\mathcal{S}|, (j)}]$ of Algorithm 1 at the j^{th} iteration. Based on Lemma 2, the half-interval contraction for the toll interval $[\mathbf{p}^{1, (j+1)}, \mathbf{p}^{|\mathcal{S}|, (j+1)}]$ retains the optimal toll levels from the previous iteration, i.e $\hat{\mathbf{p}}^{(j)} \in \{\mathbf{p}^{1, (j+1)}, \mathbf{p}^{2, (j+1)}, \dots, \mathbf{p}^{|\mathcal{S}|, (j+1)}\} \subset [\mathbf{p}^{1, (j+1)}, \mathbf{p}^{|\mathcal{S}|, (j+1)}]$. This implies that $\hat{\mathbf{v}}^{(j)} \in \chi^{(j+1)}$, so that $\Upsilon^{(j+1)} \geq \Upsilon^{(j)}$. \square

Proposition 4.

Proof. Consider the Lagrangian function associated with (.1a)-(1d):

$$L(\mathbf{f}, \hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\beta}}) = z[\mathbf{x}(\mathbf{f})] + \sum_{v \in \mathcal{V}} \sum_{k \in \mathcal{K}^v} \hat{\alpha}_v^k \left(\sum_{i \in \mathcal{N}} \sum_{p \in \mathcal{P}^k} s_i^{v,k} \hat{\Delta}_{i,p}^k - \sum_{p \in \mathcal{P}^k} f_{v,p}^k \right) - \sum_{v \in \mathcal{V}} \sum_{k \in \mathcal{K}^v} \sum_{a \in \mathcal{A}^T} \hat{\beta}_{av}^k \left((1 - w_a^v) \sum_{p \in \mathcal{P}^k} f_{v,p}^k \Delta_{a,p}^k \right)$$

where

$$z[\mathbf{x}(\mathbf{f})] = \sum_{a \in \mathcal{A}} \int_0^{\sum_{v \in \mathcal{V}} \sum_{k \in \mathcal{K}^v} \sum_{p \in \mathcal{P}^k} f_{v,p}^k \Delta_{a,p}^k} P_a^V(t_a, \omega) d\omega$$

For the problem of minimizing $L(\mathbf{f}, \widehat{\boldsymbol{\alpha}}, \widehat{\boldsymbol{\beta}})$ subject to non-negativity constraints (.1d), the optimality conditions are

$$f_{v,p}^k \frac{\partial L(\mathbf{f}, \widehat{\boldsymbol{\alpha}}, \widehat{\boldsymbol{\beta}})}{\partial f_{v,p}^k} = 0 \quad \text{and} \quad \frac{\partial L(\mathbf{f}, \widehat{\boldsymbol{\alpha}}, \widehat{\boldsymbol{\beta}})}{\partial f_{v,p}^k} \geq 0 \quad \forall p \in \mathcal{P}, v \in \mathcal{V}, k \in \mathcal{K}^v.$$

with few algebraical manipulation we get

$$\frac{\partial L(\mathbf{f}, \widehat{\boldsymbol{\alpha}}, \widehat{\boldsymbol{\beta}})}{\partial f_{v,p}^k} = c_p^k - (\widehat{\alpha}_v^k + (1 - w_a^v) \widehat{\beta}_{av}^k) \quad \forall p \in \mathcal{P}, v \in \mathcal{V}, k \in \mathcal{K}^v.$$

Therefore, we conclude that the optimal solutions of (.1a)–(.1d) must verify

$$\begin{aligned} (a) \quad & f_{v,p}^k \left[c_p^k - \left(\widehat{\alpha}_v^k + (1 - w_a^v) \widehat{\beta}_{bv}^k \right) \right] = 0 & \forall v \in \mathcal{V}, k \in \mathcal{K}^v, p \in \mathcal{P}^k \\ (b) \quad & c_p^k - \left(\widehat{\alpha}_v^k + (1 - w_a^v) \widehat{\beta}_{bv}^k \right) \geq 0 & \forall v \in \mathcal{V}, k \in \mathcal{K}^v, p \in \mathcal{P}^k \\ (c) \quad & \sum_{p \in \mathcal{P}^k} f_{v,p}^k = \sum_{i \in \mathcal{N}} \sum_{p \in \mathcal{P}^k} s_i^{v,k} \widehat{\Delta}_{i,p}^k & \forall v \in \mathcal{V}, k \in \mathcal{K}^v \\ (d) \quad & (1 - w_a^v) \sum_{p \in \mathcal{P}^k} f_{v,p}^k \Delta_{a,p}^k = 0 & \forall v \in \mathcal{V}, k \in \mathcal{K}^v, a \in \mathcal{A}^T \\ (e) \quad & f_{v,p}^k \geq 0 & \forall v \in \mathcal{V}, k \in \mathcal{K}^v, p \in \mathcal{P}^k \end{aligned}$$

□