The Value of Flexibility in Robust Location-Transportation Problems

Amir Ardestani-Jaafari, Erick Delage
HEC Montréal, Montréal, H3T 2A7, Canada
{amir.ardestani-jaafari@hec.ca, erick.delage@hec.ca}

This article studies a multi-period capacitated fixed-charge location-transportation problem in which, while the location and capacity of each facility need to be determined immediately, the determination of final production and distribution of products can be delayed until actual orders are received in each period. In contexts where little is known about future demand, robust optimization, namely using a budgeted uncertainty set, becomes a natural method to employ in order to identify meaningful decisions. Unfortunately, it is well known that these types of multi-period robust decision problems are computationally intractable. To overcome this difficulty, we propose a set of tractable conservative approximations to the problem that each exploits to a different extent the idea of reducing the flexibility of the delayed decisions. While all of these approximation models outperform previous approximation models that have been proposed for this problem, each of them also has the potential to reach a different level of compromise between efficiency of resolution and quality of the solution. A row generation algorithm is also presented in order to address problem instances of realistic size. We also demonstrate that full flexibility is often unnecessary to reach nearly, or even exact, optimal robust locations and capacities for the facilities. Finally, we illustrate our findings with an extensive numerical study where we evaluate the effect of the amount of uncertainty on the performance and structure of each approximate solutions that can be obtained.

Key words: transportation, facility location, robust optimization, flexibility, conservative approximation, demand uncertainty

.

1. Introduction

Transportation planning can be decomposed in three different levels (Crainic and Laporte 1997): strategic transportation planning, tactical transportation planning, and operational transportation planning. At the highest level of management, an important decision consists in determining the geographical locations of factories, suppliers and warehouses. Determination of facility location, such as hub locations, supplier locations, air freight hub locations, railway station locations, etc., can significantly impact the design of the strategic networks. Recognizing this fact, researchers (e.g., Christensen et al. (2013) and Abouee-Mehrizi et al. (2014)) have been developing integrated models in order to have better control on the interactions between facility location decisions and transportation strategies.

The traditional way of describing the location-transportation problem (LTP) has been to assume a deterministic environment. In a deterministic setting, *i.e.,* when there is no uncertainty about problem data, a multi-period capacitated fixed-charge LTP, with L facility locations, N customer locations, and T periods, can take the form of the following mixed integer linear program (MILP):

$$\text{(Deterministic)} \quad \underset{I,Z,Y,P}{\text{maximize}} \quad \sum_{t=1}^{T}\sum_{i=1}^{L}\sum_{j=1}^{N}(\eta - d_{ij})Y_{ij}^t - c^T P^t - (C^T Z + K^T I) \tag{1a}$$

$$\text{subject to} \quad \sum_i Y_{ij}^t \le D_j^t \,,\, \forall j \in \{1,2,\ldots,N\}\,,\, \forall t \in \{1,2,\ldots,T\} \tag{1b}$$

$$\sum_j Y_{ij}^t \le P_i^t \,,\, \forall i \in \{1,2,\ldots,L\}\,,\, \forall t \in \{1,2,\ldots,T\} \tag{1c}$$

$$P^t \le Z\,,\, \forall t \in \{1,2,\ldots,T\} \tag{1d}$$

$$Y^t \ge 0\,,\, \forall t \in \{1,2,\ldots,T\} \tag{1e}$$

$$Z \le MI\,,\, I \in \{0,1\}^L\,, \tag{1f}$$

where $Z \in \mathbb{R}^L$, $Y \in \mathbb{R}^{L\times N\times T}$, $P \in \mathbb{R}^{L\times T}$, and with M as a constant chosen large enough. This MILP integrates the optimization of both "strategic" and "operational" decisions. At the strategic level, it includes for each candidate location $i = 1,2,\ldots,L$, the binary decision I_i denoting whether a facility should be opened or not, and the continuous decision Z_i denoting the production capacity of the facility. Once these are decided, operational decisions over a horizon of $t = 1,2,\ldots,T$ include for each period t, P_i^t denoting how many goods are produced at each i-th facility and Y_{ij}^t denoting how many goods are shipped from facility i to customers at location j. The demand during period t for location $j = 1,2,\ldots,N$ is characterized by D_j^t. The total profit generated by the company is computed based on sales revenue, with $\eta > 0$ the unit price of goods, on construction costs, composed for a given facility i with size Z_i of a fixed cost K_i and variable costs $C_i Z_i$, on production costs c_i for each facility i, and finally on transportation costs, with d_{ij} being the unit cost for any shipment from location i to j. Note that each parameter η, d_{ij}, and c_i could alternatively be considered time dependant.

In model (1), all parameters are considered to be known exactly at the time of making the strategic decision. In practice however, some parameters, in particular the exact size of each demand D_j^t,[1] is rarely known at the moment of building the facilities. In recent years, studies made in a number of field of applications (Bertsimas et al. (2011), Gabrel et al. (2014b)) have demonstrated the effectiveness of robust optimization (RO) for handling uncertainty especially in cases where there is no valid argument to justify the choice of a distribution model. A naïve application of robust optimization to LTP under demand uncertainty might lead to the following robust counterpart (RC):

$$
\begin{aligned}
\text{(RC)} \quad \underset{I,Z,Y,P}{\text{maximize}} \quad & \sum_t \sum_i \sum_j (\eta - d_{ij}) Y_{ij}^t - c^T P^t - (C^T Z + K^T I) && \text{(2a)} \\
\text{subject to} \quad & \sum_i Y_{ij}^t \leq D_j^t \,, \forall D \in \mathcal{D}, \forall j \,, \forall t && \text{(2b)} \\
& \text{(1c)} - \text{(1f)} \,,
\end{aligned}
$$

where $\mathcal{D}$ is the uncertainty set for the vector composed of all the demands $(D^1, D^2, \ldots, D^T)$.

Although it can be shown that the RC model can be reformulated as a MILP if $\mathcal{D}$ is polyhedral, the solution that it provides will often appear to be overly conservative, *i.e.,* it might suggest to open only a few facilities (if any at all) with very limited capacity. This is actually due to the fact that the RC model completely disregards how operational decisions, namely size of production and deliveries, are delayed and can exploit the information that becomes available about the demand. This motivates the use of the following multi-period robust location-transportation problem (MRLTP) model:

$$
\begin{aligned}
\text{(MRLTP)} \quad \underset{I,Z}{\text{maximize}} \quad & \min_{D \in \mathcal{D}} \sum_t h_t(I, Z, D^t) - (C^T Z + K^T I) && \text{(3a)} \\
\text{subject to} \quad & Z \leq MI \,, \, I \in \{0,1\}^L \,, && \text{(3b)}
\end{aligned}
$$

where $h_t(I, Z, D^t)$ is the profit generated during period t, once the demand is revealed for this period, and is defined as

$$
\begin{aligned}
h_t(I, Z, D^t) = \max_{Y^t, P^t} \quad & \sum_i \sum_j (\eta - d_{ij}) Y_{ij}^t - c^T P^t && \text{(4a)} \\
\text{subject to} \quad & \sum_i Y_{ij}^t \leq D_j^t \,, \forall j && \text{(4b)} \\
& \sum_j Y_{ij}^t \leq P_i^t \,, \forall i && \text{(4c)} \\
& P^t \leq Z && \text{(4d)} \\
& Y^t \geq 0 \,, && \text{(4e)}
\end{aligned}
$$

which in particular captures the fact that since it is assumed that goods cannot be stored (or demand backlogged) from one period to the other, neither at the facility nor at the demand locations, it is always possible to design an optimal transportation and production plan that depend only on the currently realized demand.

Finally, we make the common assumption that the demand vector D is known to lie in a budgeted uncertainty set (see Bertsimas and Sim (2004)), *i.e.,* that each D_i lies in an interval and that at most Γ of the terms across all locations and time periods can take extreme values.

While it appears that the MRLTP does implement as much flexibility as is needed in this problem, Atamtürk and Zhang (2007) established that evaluating the objective is already computationally intractable when $T = 1$. In this paper, we present a set of six conservative approximation models to the problem that each exploits to a different extent the idea of reducing the flexibility of the delayed decisions. These models will allow us to explore empirically the compromises that need to be made between flexibility/conservatism and "tractability"[2]. Overall, we consider this article to make the following contributions:

1. We present a set of tractable conservative approximations of the MRLTP that each employ different form of application of affine adjustments proposed in Ben-Tal et al. (2004) and Chen and Zhang (2009). While we do demonstrate empirically for the first time how significant the improvements can be in terms of the quality of the approximate solutions for the MRLTP, especially when comparing to Baron et al. (2011)'s robust model, we also establish conditions under which some of the simplest approximation schemes already provide optimal solutions. These theoretical results rely on carefully adapting the arguments presented in Ben-Tal et al. (2004) and Bertsimas and Goyal (2012) to our multi-period setting.

2. Two of our formulations, namely what will be referred as ELAARC and HD-ELAARC, also provide valuable insights about how better conservative approximation models can be obtained in robust multi-stage optimization problems. With ELAARC, this is done by creating affine adjustment only after replacing the recourse problem by an equivalent penalized formulation. With HD-ELAARC, this is done by letting the affine adjustments depend on the whole history even though an optimal recourse policy is known to be independent of the history. These two ideas might serve many other instances of robust multi-stage decision problems.

3. We propose a row generation algorithm that employs a parsimonious choice of valid inequalities in order to accelerate the resolution of one of our most complex approximation model, while being easily adaptable to any of our other formulations. Our implementation of this algorithm allows us to reduce by a factor between 16 and 260 the solution time of larger instances and to solve instances with 20 periods, 15 facility locations, and 30 demand locations in less than 3 hours while an exact method could not converge after running for more than 48 hours.

4. We perform an extensive numerical study in order to analyse the value of flexibility and the robustness-performance trade-off that can be achieved by each approximation model. Furthermore, we provide some insights about the general structure of the decisions that are proposed by each approximation model on a large set of problem instances.

The remainder of the paper is organized as follows. In Section 2, we review prior work about the robust location transportation problem under demand uncertainty. In Section 3, we present six new tractable approximation models for the MRLTP. In Section 4, we establish the relation between the bounds that are obtained using each approximation model and identify conditions under which some of the models return exact solutions. Next, we present in Section 5 the details of a decomposition scheme that can be used to accelerate the resolution of larger size models. In Section 6, we provide numerical results and finally, the conclusions and possible future research directions are presented in Section 7.

2. Prior work

To the best of our knowledge, Atamtürk and Zhang (2007) were the first to study a model related to the two-stage robust location-transportation problem (TRLTP), a special case of MRLTP with a single-period $T = 1$, for an application of network flow and design problem where their objective was to minimize worst-case cost over a budgeted uncertainty set. They compared a two-stage robust optimization model with a stochastic program where the objective of the stochastic program was to minimize the sum of the first-stage cost and expected value of the second-stage cost. When distribution was captured by 200 demand scenarios, they showed that while the solution of the two-stage robust optimization model increased the expected cost by 1.1% it actually decreased by 29.1% the cost incurred under the worst-case scenario. They identified the TRLTP as a special case of the modelling framework and after recognizing that their problem was NP-hard, proposed to use a cutting plane algorithm to reach a global optimum.

Recently Gabrel et al. (2014a) and Zeng and Zhao (2013) proposed two cutting plane methods to solve a TRLTP exactly under the budgeted uncertainty set with an integer budget. Gabrel et al. (2014a) showed that the adversarial problem in the TRLTP could be reformulated as a MILP. The master problem of TRLTP could then be tackled using Kelley's cutting plane algorithm given that optimality cuts are provided using a MILP solver. Zeng and Zhao (2013) seem to have improved on the solution time by employing a column-and-constraint generation (C&CG) algorithm instead of Kelley's cutting plane algorithm. Finally, in a similar transportation problem, Lei et al. (2015) proposed a two-level cutting plane method for a two-stage mobile facility fleet sizing and routing problem wherein fleet sizing and routing plan are determined in the first stage and allocation of demands to the mobile facilities are determined in second stage. Although there is empirical evidence that these exact resolution methods are efficient, the adversarial problem that is solved in each case takes the form of a MILP that is inherently NP-hard. There is therefore always a risk of having to endure unbearable computation times before obtaining solutions to any specific problem instance.

In Baron et al. (2011), the authors can be considered to have proposed the first tractable conservative approximation of the MRLTP model. In their paper, the authors proposed a robust optimization model in which static (*i.e.,* inflexible) production and fractional transportation policies are optimized. Indeed, they replaced the Y_{ij}^t variables with $X_{ij}^t D_j^t$ which reflects the notion that X_{ij}^t is the proportion of demand at location j and time t that is satisfied by the facility at location i. Specifically, their proposed fractional variable-based (FVB) model takes the form:

$$\text{(FVB)}\quad \underset{I,Z,X,P}{\text{maximize}}\ \min_{D\in\mathcal{D}} \sum_t\sum_i\sum_j(\eta-d_{ij})X_{ij}^tD_j^t - c^TP^t-(C^TZ+K^TI) \tag{5a}$$

$$\text{subject to}\ \sum_j X_{ij}^tD_j^t \le P_i^t\,,\ \forall D\in\mathcal{D}\,,\ \forall i\,,\ \forall t \tag{5b}$$

$$P^t\le Z\,,\ \forall t \tag{5c}$$

$$\sum_i X_{ij}^t \le 1\,,\ \forall j\,,\ \forall t \tag{5d}$$

$$X^t\ge 0\,,\ \forall t \tag{5e}$$

$$Z\le MI\,,\ I\in\{0,1\}^L. \tag{5f}$$

They next studied the impact of two types of uncertainty set, box and ellipsoidal sets, on the structure of the robust solution and compared it to the nominal one. In particular, they paid special attention to the number of opened facilities, the total capacity of facilities, the number of deliveries made from each facility to the customer locations under different scenarios. Surprisingly, the following example highlights the fact that the solution of the FVB model might abandon opportunities of making profits that are arbitrarily large even with respect to the worst-case scenario. In contrast, the simpler RC model actually does not suggest as much of a conservative solution for the same instances. On the other hand, some might argue that the FVB model provides a transportation policy that can easily be interpreted.

Example 1. Consider an example of MRLTP with $T=1$ and two customers such that $D\in[\bar{D}\pm\hat{D}]$ where $\bar{D}=10000$ and $\hat{D}=5000$. The location of customers is considered as the candidate location of facilities, $L=2$. The open facility will cover demand, if possible, with $\eta=1$, $c_i=0.1$, $C_i=0.1$ and $K_i=3000$ for all i and the transportation cost between locations is equal to 1. We assume that the budget is $\Gamma=2$ which leads to a box uncertainty set. As it is shown in Appendix A, the optimal value of RC model (2) is equal to 1000 but the optimal value of the FVB model is zero in this example. This indicates that while the RC model suggests opening the two facilities which leads to a worst-case profit of 1000, the FVB model closes everything down. When scaling every parameters in the objective function by some $\alpha>0$, FVB will let go of an arbitrarily large opportunity to make profit. Intuitively, the over-conservatism of the FVB model is due to the fact that any feasible candidate for production must satisfy the largest possible demand, because of (5b), while the worst-case profits that end up being measured in (5a) actually account for the lowest demand. This necessarily leads the FVB model to imply that a lot of the production will be wasted once one attempts to satisfy even a small amount of demand.

Recently, Bertsimas and de Ruiter (2015) proposed applying affine adjustments on a dual reformulation of the TRLTP and show improved computation time when compared to applying the same type of adjustment on the original TRLTP. Given that the two types of applications of these adjustments are shown to be equivalent, it is likely that their methods could be used to improve resolutions time of the models that are proposed in this work if one wishes to avoid using dedicated decomposition schemes. Yet, we are still convinced at the time of writing this article that it is necessary to employ row generation algorithms of the type presented here to obtain solutions to the larger size instances of MRLTP problem in a reasonable amount of time.

3. Six conservative tractable approximations

In what follows, we provide six progressive ways of improving the quality of the solution obtained from the RC and FVB model. Each of them will employ the idea of affine adjustments from Ben-Tal et al. (2004) and a version of the splitting based uncertainty set extensions from Chen and Zhang (2009) to exploit to a different extent the fact that the operational decisions P and Y can be adjusted to the realization of the demand. The type of flexibility added by our models can be divided in three classes. Similarly to what is done in the FVB model, the first class of approximation models, called "customer-driven", will adjust the size of a delivery to a customer simply based on information about that customer's demand, *i.e.,* that $Y_{ij}^t := \pi_{ij}^t(D_j)$ with $\pi_{ij}^t : \mathbb{R} \to \mathbb{R}$. In opposition, the second class of approximation models, called "market-driven" will be more flexible and attempt to optimize delivery policies that take into account the state of the market as a whole, *i.e.,* that $Y_{ij}^t := \pi_{ij}^t(D^t)$ with $\pi_{ij}^t : \mathbb{R}^N \to \mathbb{R}$. This second class will necessarily lead to models that are more computationally demanding yet have the potential of identifying better performing strategies. We will finally introduce a final class of approximation models, referred as "history-driven", that will attempt to exploit the full history of demand although we did not yet identify an improvement there that motivates the added computational burden. Note that in presenting each of the approximation models we omit to derive and spell out the finite dimensional MILP reformulation that would be obtained by applying duality theory to each robust constraint and objective function for the sake of keeping the presentation compact.

3.1. Customer-driven affine adjustments

Our first approximation model will stem from the realization that in the recourse problem of the MRLTP, namely Problem (4), the inequality constraint (4c) will be active at optimum and can therefore be replaced with an equality constraint. This argument motivates replacing P_i^t with $\sum_j X_{ij}^t D_j^t$ for all i and t in the FVB model (5). Using this replacement, our model effectively fully adapts variable P_i^t to the revealed demand, which was an important issue with the FVB model. In order to ensure that we obtain a tighter approximation than with the RC model, we also propose replacing the fractional adjustment, $Y^t_{ij} := X^t_{ij} D^t_j$ with $Y^t_{ij} := X^t_{ij} D^t_j + W^t_{ij}$ for all i, j, and t. The motivation for using W^t_{ij} is that there are some cases as shown in Example 1, wherein RC provides a tighter solution than FVB. Introducing the variable W^t_{ij}, namely the "static" component of the transportation policy, enables us to guarantee that this revised model always provides a solution that is at most as conservative as the solution of the RC model (see Proposition 2 for more details). Overall, these modifications lead to our revised fractional-variable based (RFVB1) model:

$$
\begin{aligned}
\text{(RFVB1)} \quad \underset{I,Z,X,W}{\text{maximize}} \quad & \min_{D \in \mathcal{D}} \sum_t \sum_i \sum_j (\eta - d_{ij} - c_i)(X^t_{ij} D^t_j + W^t_{ij}) - (C^T Z + K^T I) && \text{(6a)} \\
\text{subject to} \quad & \sum_i X^t_{ij} D^t_j + W^t_{ij} \leq D_j \,, \forall \, D \in \mathcal{D} \,, \forall \, j \,, \forall \, t && \text{(6b)} \\
& \sum_j X^t_{ij} D^t_j + W^t_{ij} \leq Z_i \,, \forall \, D \in \mathcal{D} \,, \forall \, i \,, \forall \, t && \text{(6c)} \\
& X^t_{ij} D^t_j + W^t_{ij} \geq 0 \,, \forall \, D \in \mathcal{D} \,, \forall \, i \,, \forall \, j \,, \forall \, t && \text{(6d)} \\
& Z \leq M I \,, \, I \in \{0,1\}^L \,. && \text{(6e)}
\end{aligned}
$$

We next exploit an extended description of the budgeted uncertainty set proposed in Chen and Zhang (2009) in order to optimize customer-driven transportation policies that have piecewise-linear structure (we also refer the reader to Georghiou et al. (2015) for details about techniques involving non-linear decision structures). Specifically, we employ a lifting of the demand uncertainty space

$$
\mathcal{D} = \left\{ D \in \mathbb{R}^{N \times T} \,\middle|\, \exists (D^+, D^-) \in \mathcal{D}_2, \, D = \bar{D} + D^+ - D^- \right\}
$$

where

$$
\mathcal{D}_2 = \left\{ (D^+, D^-) \in \mathbb{R}^{N \times T} \times \mathbb{R}^{N \times T} \,\middle|\, \begin{array}{c} \exists (\delta^+, \delta^-) \in \mathbb{R}^{N \times T} \times \mathbb{R}^{N \times T}, \, \delta^+ \geq 0, \, \delta^- \geq 0, \, \|\delta^+ + \delta^-\|_\infty \leq 1, \\ \|\delta^+ + \delta^-\|_1 \leq \Gamma, \, D^{t+}_j = \hat{D}^{t+}_j \delta^{t+}_j, \, D^{t-}_j = \hat{D}^{t-}_j \delta^{t-}_j \, \forall \, j \, \forall \, t \end{array} \right\}.
$$

As illustrated in Figure 1, this lifting allows one to define different affine policies for positive perturbations than those defined for negative perturbations thus giving rise to the possibility of a better performing non-linear adjustment. For example, by letting $W_{ij} = \alpha \bar{D}^t_j$, $X^{t+}_{ij} = 0$ and $X^{t+}_{ij} = -\alpha$ for some $0 \leq \alpha \leq 1$, the lifting implements the policy $Y^t_{ij} := \alpha \min(D^t_j ; \bar{D}^t_j)$ (see Figure 1(c)) which can make better use of the capacity Z_i that is made available.

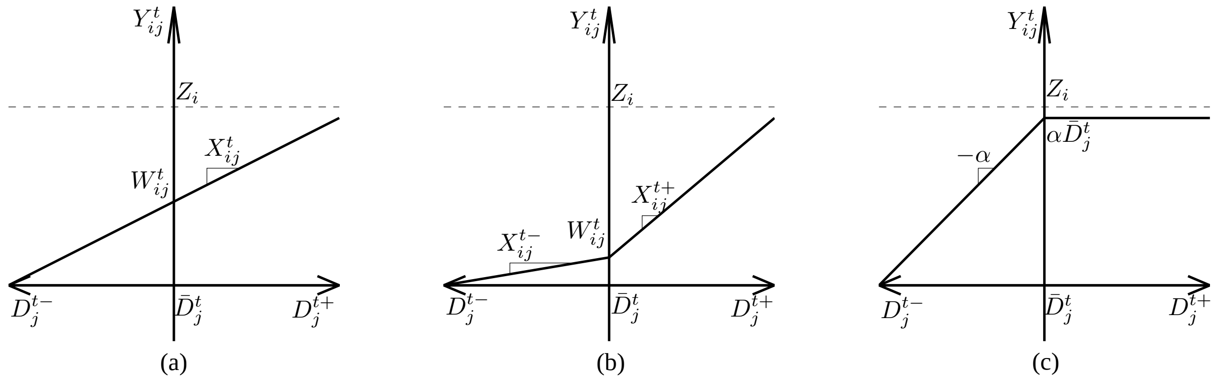


Figure 1 **Illustrative comparison of an affine adjustment in (a) and an affine adjustment on the lifted space (D_j^{t+}, D_j^{t-}) in (b). Finally, (c) presents an example of lifted adjustment that implements $Y_{ij}^t := \alpha \min(D_j^t; \bar{D}_j^t)$ in order to make better use of available capacity.**

This manipulation of the model leads to our second revision of the fractional-variable based (RFVB2) model:

$$(\text{RFVB2}) \quad \underset{I,Z,X^+,X^-,W}{\text{maximize}} \quad \min_{(D^+,D^-)\in\mathcal{D}_2} \sum_t \sum_i \sum_j (\eta - d_{ij} - c_i)(X_{ij}^{t+} D_j^{t+} + X_{ij}^{t-} D_j^{t-} + W_{ij}^t) - (C^T Z + K^T I) \tag{7a}$$

$$\text{subject to} \quad \sum_i X_{ij}^{t+} D_j^{t+} + X_{ij}^{t-} D_j^{t-} + W_{ij}^t \leq D_j^t \,, \forall (D^+, D^-) \in \mathcal{D}_2 \,, \forall j \,, \forall t \tag{7b}$$

$$\sum_j X_{ij}^{t+} D_j^{t+} + X_{ij}^{t-} D_j^{t-} + W_{ij}^t \leq Z_i \,, \forall (D^+, D^-) \in \mathcal{D}_2 \,, \forall i \,, \forall t \tag{7c}$$

$$X_{ij}^{t+} D_j^{t+} + X_{ij}^{t-} D_j^{t-} + W_{ij}^t \geq 0 \,, \forall (D^+, D^-) \in \mathcal{D}_2 \forall i \,, \forall j \,, \forall t \tag{7d}$$

$$Z \leq MI \,, I \in \{0,1\}^L \,. \tag{7e}$$

3.2. Market-driven affine adjustments

We now provide three approximation models that will attempt to exploit full market information in making deliveries. The first of these attempts can be considered as a direct application of the AARC framework for the MRLTP as it was initially introduced by Ben-Tal et al. (2004). In such a framework, the adaptive policies for later stage decisions are considered to be restricted to the set of affine functions of the uncertain parameters. In the context of this problem, this means that each adaptive policy of the MRLTP model (3) should take the form $Y_{ij}^t := (X_{ij}^t)^T D^t + W_{ij}^t$ with $X_{ij}^t \in \mathbb{R}^N$ and $W_{ij}^t \in \mathbb{R}$. In other words, this means that the delivery for a customer j can depend on all the orders that are made in this market. Intuitively, this added flexibility might be beneficial considering that the amount of production of each facility is constrained by its capacity Z_i, therefore an increase in demand from a nearby customer might justify reducing the amount of goods to transport to a further customer in order to improve profitability. We note that similarly as before the variable P_i^t of the MRLTP model (3) will be replaced by $\sum_j Y_{ij}^t$ in all of our proposed approximations. When restricting our search to affine policies of the D^t vector, the approximation model takes the following form:

$$
\begin{aligned}
\text{(AARC)} \quad & \underset{I,Z,X,W}{\text{maximize}} \ \min_{D\in\mathcal{D}} \sum_t\sum_i\sum_j(\eta-d_{ij}-c_i)\left((X_{ij}^t)^T D^t + W_{ij}^t\right) - (C^T Z + K^T I) && \text{(8a)}\\
& \text{subject to} \ \sum_i (X_{ij}^t)^T D^t + W_{ij}^t \le D_j^t\,,\ \forall D\in\mathcal{D}\,,\ \forall j\,,\ \forall t && \text{(8b)}\\
& \sum_j (X_{ij}^t)^T D^t + W_{ij}^t \le Z_i\,,\ \forall D\in\mathcal{D}\,,\ \forall i\,,\ \forall t && \text{(8c)}\\
& (X_{ij}^t)^T D^t + W_{ij}^t \ge 0\,,\ \forall D\in\mathcal{D}\,,\ \forall i\,,\forall j\,,\ \forall t && \text{(8d)}\\
& Z \le MI, I\in\{0,1\}^L\,. && \text{(8e)}
\end{aligned}
$$

Similar to what was done to obtain the RFVB2 model, AARC can be improved by lifting the uncertainty set. LAARC of MRLTP (3) can be obtained by considering policies that are affine in the pair of perturbations $(D^+, D^-) \in \mathcal{D}_2$, namely $Y_{ij}^t := (X_{ij}^{t+})^T D^{t+} + (X_{ij}^{t-})^T D^{t-} + W_{ij}^t$ with $X_{ij}^{t+} \in \mathbb{R}^N$, $X_{ij}^{t-} \in \mathbb{R}^N$, and $W_{ij}^t \in \mathbb{R}$. This new approximation model takes the more sophisticated form:

(LAARC)

$$
\begin{aligned}
& \underset{I,Z,X^+,X^-,W}{\text{maximize}} \ \min_{(D^+,D^-)\in\mathcal{D}_2} \sum_t\sum_i\sum_j(\eta-d_{ij}-c_i)\left((X_{ij}^{t+})^T D^{t+} + (X_{ij}^{t-})^T D^{t-} + W_{ij}^t\right)\\
& \qquad\qquad\qquad\qquad - (C^T Z + K^T I) && \text{(9a)}\\
& \text{subject to} \ \sum_i (X_{ij}^{t+})^T D^{t+} + (X_{ij}^{t-})^T D^{t-} + W_{ij}^t \le D_j\,,\ \forall\,(D^+,D^-)\in\mathcal{D}_2,\ \forall j,\ \forall t && \text{(9b)}\\
& \sum_j (X_{ij}^{t+})^T D^{t+} + (X_{ij}^{t-})^T D^{t-} + W_{ij}^t \le Z_i\,,\ \forall\,(D^+,D^-)\in\mathcal{D}_2,\ \forall i,\ \forall t && \text{(9c)}\\
& (X_{ij}^{t+})^T D^{t+} + (X_{ij}^{t-})^T D^{t-} + W_{ij}^t \ge 0\,,\forall\,(D^+,D^-)\in\mathcal{D}_2,\ \forall i,\forall j,\ \forall t && \text{(9d)}\\
& Z \le MI,\ I\in\{0,1\}^L\,. && \text{(9e)}
\end{aligned}
$$

Now, we propose an extension to the LAARC, referred as the ELAARC model, which will benefit from a manipulation of a multi-period robust optimization model which to the best of our knowledge is presented for the first time. The key idea is to reformulate the recourse problem (4) in a way that relaxes the constraint that is plagued by uncertainty without compromising the authenticity of the model. Namely, let us consider the following equivalent reformulation

$$
h_t(I,Z,D^t) = \max_{Y^t,P^t} \quad \sum_i\sum_j(\eta-d_{ij}-c_i)Y_{ij}^t - \sum_j B_j\theta_j^t \tag{10a}
$$

$$\text{subject to} \quad \sum_i Y_{ij}^t \le D_j^t + \theta_j^t\,,\ \forall j \tag{10b}$$

$$\sum_j Y_{ij}^t \le Z_i\,,\ \forall i \tag{10c}$$

$$Y^t \ge 0\,,\ \theta^t \ge 0\,, \tag{10d}$$

where $Y^t \in \mathbb{R}^{L\times N}$, $\theta^t \in \mathbb{R}^N$ and where each B_j is a marginal penalty for violating constraint (4b) that is chosen large enough for the optimal value of the optimization problem to remain the same. We refer the reader to Appendix B for a proof that the assignment $B_j = \max_i(\eta - c_i - d_{ij})\ \forall j$ meets this criterion.

Similarly as for the LAARC model, we adjust the deliveries based on the lifted uncertainty space, $Y_{ij}^t := (X_{ij}^{t+})^T D^{t+} + (X_{ij}^{t-})^T D^{t-} + W_{ij}^t$, furthermore we adjust each new auxiliary variable θ_j according to $\theta_j^t := S_j^{t+} D_j^{t+} + S_j^{t-} D_j^{t-}$ in order to obtain the ELAARC approximation model

(ELAARC)

$$\underset{\substack{I,Z,X^+,X^-,\\ W,S^+,S^-}}{\text{maximize}} \quad \min_{(D^+,D^-)\in\mathcal{D}_2} \sum_t \sum_i \sum_j \left((X_{ij}^{t+})^T D^{t+} + (X_{ij}^{t-})^T D^{t-} + W_{ij}^t\right)$$

$$-(C^T Z + K^T I) - \sum_t \sum_j B_j (S_j^+ D_j^{t+} + S_j^- D_j^{t-}) \tag{11a}$$

$$\text{subject to} \quad \sum_i (X_{ij}^{t+})^T D^{t+} + (X_{ij}^{t-})^T D^{t-} + W_{ij}^t \le D_j^t$$

$$+ S_j^{t+} D_j^{t+} + S_j^{t-} D_j^{t-}\,,\ \forall (D^+, D^-) \in \mathcal{D}_2\,,\ \forall j\,,\ \forall t \tag{11b}$$

$$S_j^{t+} D_j^{t+} + S_j^{t-} D_j^{t-} \ge 0\,,\ \forall (D^+, D^-) \in \mathcal{D}_2\,, \forall j\,,\ \forall t \tag{11c}$$

$$(9c) - (9e)\,, \tag{11d}$$

where $S^+ \in \mathbb{R}^{N\times T}$ and $S^- \in \mathbb{R}^{N\times T}$. Finally, one might realize that when using this lifted uncertainty space, the worst-case analysis of this optimization model really only depends on negative adversarial perturbations. This will be an interesting feature to exploit when the time comes to implement and solve the model.

Proposition 1. *The LAARC and ELAARC approximation models can respectively be reduced to the following two optimization problems:*

(LAARC2)

$$\underset{I,Z,X^-,W}{\text{maximize}} \quad \min_{D^-\in\mathcal{D}_3} \sum_t \sum_i \sum_j (\eta - d_{ij} - c_i)\left((X_{ij}^{t-})^T D^{t-} + W_{ij}^t\right) - (C^T Z + K^T I)$$

$$\text{subject to} \quad \sum_i (X_{ij}^{t-})^T D^{t-} + W_{ij}^t \le \bar{D}_j^t - D_j^{t-}\,,\ \forall D^- \in \mathcal{D}_3\,, \forall j\,,\ \forall t$$

$$\sum_j (X_{ij}^{t-})^T D^{t-} + W_{ij}^t \le Z_i\,,\ \forall D^- \in \mathcal{D}_3\,, \forall i\,,\ \forall t$$

$$(X_{ij}^{t-})^T D^{t-} + W_{ij}^t \ge 0\,,\ \forall D^- \in \mathcal{D}_3\,, \forall i\,, \forall j\,,\ \forall t$$

$$Z \le MI,\ I \in \{0,1\}^L\,,$$

and

(ELAARC2)

$$\underset{I,Z,X^-,W,S^-}{\text{maximize}} \quad \min_{D^-\in\mathcal{D}_3} \sum_t\sum_i\sum_j(\eta-d_{ij}-c_i)\left((X_{ij}^{t-})^T D^- + W_{ij}\right)$$
$$-(C^TZ+K^TI)-\sum_t\sum_j B_j(S_j^{t-}D_j^{t-}) \qquad (12a)$$
$$\text{subject to} \quad \sum_i (X_{ij}^{t-})^T D^- + W_{ij} \le \bar{D}_j^t - D_j^{t-} + S_j^{t-}D_j^{t-}\,,\,\forall\, D^-\in\mathcal{D}_3\,,\,\forall\, j\,,\,\forall\, t \qquad (12b)$$
$$\sum_j (X_{ij}^{t-})^T D^- + W_{ij} \le Z_i\,,\,\forall\, D^-\in\mathcal{D}_3\,,\forall\, i\,,\,\forall\, t \qquad (12c)$$
$$(X_{ij}^{t-})^T D^- + W_{ij} \ge 0\,,\,\forall\, D^-\in\mathcal{D}_3\,,\,\forall\, i\,,\forall\, j\,,\,\forall\, t \qquad (12d)$$
$$S_j^{t-}D_j^{t-} \ge 0\,,\,\forall\, D^-\in\mathcal{D}_3\,,\,\forall j\,,\,\forall\, t \qquad (12e)$$
$$Z \le MI\,,\, I\in\{0,1\}^L\,, \qquad (12f)$$

where

$$\mathcal{D}_3 = \left\{ D^-\in\mathbb{R}^{N\times T} \middle| \exists\delta^-\in\mathbb{R}^{N\times T},\, 0\le\delta^-\le 1,\, \sum_{t=1}^T\sum_{j=1}^N \delta_j^{t-}\le\Gamma,\, D_j^{t-}=\hat{D}_j^t\delta_j^{t-}\,\forall\, j\,\forall\, t \right\}.$$

Proof: First, one can easily confirm that both LAARC and ELAARC reduce respectively to LAARC2 and ELAARC2 when the uncertainty set $\mathcal{D}$ is replaced with the following uncertainty set

$$\mathcal{D}_2' := \mathcal{D}_2 \cap \{(D^+,D^-)\in\mathbb{R}^{N\times T}\times\mathbb{R}^{N\times T}\,|\, D^+=0\}\,.$$

Since $\mathcal{D}_2'\subset\mathcal{D}_2$, it is clear that the optimal values of LAARC2 and ELAARC2 are respectively at least as large as the optimal value of LAARC and ELAARC. Looking more specifically at the LAARC2 model, given any optimal solution (I^*, Z^*, X^{-*}, W^*), it is possible to reconstruct a feasible solution for LAARC, simply considering $X^{+*}=X^{-*}$, that achieves the same objective value as the optimal value identified by LAARC2. Hence, this reconstructed solution is optimal for LAARC. Note that in confirming feasibility of this reconstructed solution the difficulty resides in establishing whether the robust demand constraint is satisfied. Namely, that for all $j=1,2,\ldots,N$ and for all t on can confirm that

$$\max_{(D^+,D^-)\in\mathcal{D}_2}\sum_i((X_{ij}^{t+*})^T D^{t+} + (X_{ij}^{-*t})^T D^{t-} + W_{ij}^{t*}) - \bar{D}_j^t - D_j^{t+} + D_j^{t-}$$
$$= \max_{(D^+,D^-)\in\mathcal{D}_2}\sum_i((X_{ij}^{t-*})^T(D^{t+}+D^{t-}) + W_{ij}^{t*}) - \bar{D}_j^t - D_j^{t+} + D_j^{t-}$$
$$\le \max_{(D^+,D^-)\in\mathcal{D}_2}\sum_i((X_{ij}^{t-*})^T(D^{t+}+D^{t-}) + W_{ij}^{t*}) - \bar{D}_j^t + (D_j^{t+} + D_j^{t-})$$
$$= \max_{(0,D^-)\in\mathcal{D}_2}\sum_i((X_{ij}^{t-*})^T D^{t-} + W_{ij}^{t*}) - \bar{D}_j^t + D_j^{t-}$$
$$= \max_{D^-\in\mathcal{D}_3}\sum_i((X_{ij}^{t-*})^T D^{t-} + W_{ij}^{t*}) - \bar{D}_j^t + D_j^{t-} \le 0\,,$$

where we exploited the fact that, for all $(D^+, D^-) \in \mathcal{D}_2$, D^+ is non-negative. Finally an exactly similar argument can be made to confirm that the optimal solution of ELAARC2 can be used to obtain an optimal solution to ELAARC, simply by letting $X^{+*} = X^{-*}$ and $S^{+*} = S^{-*}$. $\square$

3.3. History driven affine adjustments

For completeness, we finally highlight the fact that, in a multi-period setting, one can suppose that an even more flexible transportation strategy can be obtained by employing affine adjustments that depend jointly on all previous realization of the demand until the implementation of the transportation decision. Mathematically speaking, the injection of such additional flexibility leads to the following structures. For all t and j, in the case of the direct AARC approach one gets $Y_{ij}^t := \sum_{t'=1}^t (X_{ij}^{tt'})^T D^{t'} + W_{ij}^t$, while the history driven version of LAARC would employ $Y_{ij}^t := \sum_{t'=1}^t (X_{ij}^{tt'+})^T D^{t'+} + (X_{ij}^{tt'-})^T D^{t'-} + W_{ij}^t$. Finally, the ELAARC model could additionally employ $\theta_j^t := \sum_{t'=1}^t S_j^{tt'} D_j^{t+} + S_j^{tt'} D_j^{t+}$. We present below the history-driven version of ELAARC in its reduced form.

(HD-ELAARC)

$$\underset{I,Z,X^-,W,S^-}{\text{maximize}} \quad \min_{D^- \in \mathcal{D}_3} \sum_t \sum_i \sum_j (\eta - d_{ij} - c_i)(\sum_{t'=1}^t (X_{ij}^{tt'-})^T D^{t'-} + W_{ij}^t) - (C^T Z + K^T I) - \sum_t \sum_j B_j (\sum_{t'=1}^t S_j^{tt'-} D_j^{t'-}) \tag{13a}$$

$$\text{subject to} \quad \sum_i \sum_{t'=1}^t (X_{ij}^{tt'-})^T D^{t'-} + W_{ij}^t \le \bar{D}_j^t - D_j^{t-} + \sum_{t'=1}^t S_j^{tt'-} D_j^{t'-}, \forall D^- \in \mathcal{D}_3, \forall j, \forall t \tag{13b}$$

$$\sum_j \sum_{t'=1}^t (X_{ij}^{tt'-})^T D^{t'-} + W_{ij}^t \le Z_i, \forall D^- \in \mathcal{D}_3, \forall i, \forall t \tag{13c}$$

$$\sum_{t'=1}^t (X_{ij}^{tt'-})^T D^{t'-} + W_{ij}^t \ge 0, \forall D^- \in \mathcal{D}_3, \forall i, \forall j, \forall t \tag{13d}$$

$$\sum_{t'=1}^t S_j^{tt'-} D_j^{t'-} \ge 0, \forall D^- \in \mathcal{D}_3, \forall j, \forall t \tag{13e}$$

$$Z \le MI, I \in \{0,1\}^L, \tag{13f}$$

where for each i, j, t, and $t' \le t$, we have that $X_{ij}^{tt'-} \in \mathbb{R}^N$ and $S_j^{tt'-} \in \mathbb{R}$.

While we will show in our numerical experiments that such history driven models can be used to obtain even tighter bounds than their non-history driven version, we note two important drawbacks. First, from a computational perspective the number of parameters that need to be optimized using this type of adjustment scales in the order of $O(LN^2T^2)$. Perhaps as importantly, the decision rules that are obtained with this model will suggest strategies which structure is incoherent with the most natural structure that would be used by optimal fully flexible strategies, namely the fact that the transportation policy for time t only depend on the realized demand for time t. For these two reasons, we will later omit to present a complete numerical analysis of this model.

4. Theoretical analysis of robust approximation models

In this section, we are interested in demonstrating theoretically how solutions of better quality can be obtained by using an approximation model that offers more flexibility for the delayed decisions. In particular, we start by establishing what are the respective qualities of the bounds that are obtained from each model regarding the worst-case profit of a candidate solution for facility locations and capacities.

PROPOSITION 2. *Given some fixed values for the strategic decision vectors* $I \in \{0,1\}^L$ *and* $Z \in \mathbb{R}^L$*, let* $f_{RC}(I,Z)$*,* $f_{MRLTP}(I,Z)$*,* $f_{FVB}(I,Z)$*,* $f_{RFVB1}(I,Z)$*,* $f_{RFVB2}(I,Z)$*,* $f_{AARC}(I,Z)$*,* $f_{LAARC}(I,Z)$*,* $f_{ELAARC}(I,Z)$ *and* $f_{HD\text{-}ELAARC}(I,Z)$ *respectively be the value of the objective functions of approximation models* (2)*,* (3)*,* (5)*,* (6)*,* (7)*,* (8)*,* (9)*,* (11)*, and* (13) *when the rest their respective decision variables are optimized. The following partial ordering is satisfied for any values of* I *and* Z*:*

$$f_{RC}(I,Z) \le f_{RFVB1}(I,Z) \le f_{RFVB2}(I,Z) \le f_{LAARC}(I,Z) \le f_{ELAARC}(I,Z) \le f_{MRLTP}(I,Z),$$
$$f_{FVB}(I,Z) \le f_{RFVB1}(I,Z) \le f_{AARC}(I,Z) \le f_{LAARC}(I,Z)$$
$$f_{ELAARC}(I,Z) \le f_{HD\text{-}ELAARC}(I,Z) \le f_{MRLTP}(I,Z),.$$

Proof: The function $f_{ELAARC}(I,Z)$ provides a lower bound on true worst-case profit $f_{\mathrm{MRLTP}}(I,Z)$ since the adjustable variables that appear in problem (10) are limited to affine function of uncertain parameter. The ELAARC model reduces to the LAARC model when the value of variables S_j^{t+} and S_j^{t-} are forced to take the value zero for all j and t. One can also show that the LAARC model reduces to the AARC model when the constraint $X_{ij}^{t+} = -X_{ij}^{t-}\,\forall i,j,t$, is added, thus leading to a lower evaluation of the worst-case multi-period profit. The LAARC model also reduces to the RFVB2 model when adding the constraints that each terms of $X_{ij}^t \in \mathbb{R}^N$ equals zero except for the j-th term. A similar set of constraints make the AARC model reduce to RFVB1 model. The RFVB2 model reduces to the RFVB1 model under similar conditions than those that make LAARC reduce to AARC. Lastly, one can show that RFVB1 upper bounds RC since the optimization model becomes equivalent to RC when we force $X = 0$.

Next, assuming that W is fixed to zero, one can show that the evaluation of worst-case profit obtained from the RFVB1 model is larger than the evaluation from the FVB model since one can replace constraint (6c) with

$$\sum_j D_j^t X_{ij}^t \le P_i^t\,,\,\forall D \in \mathcal{D}\,,\,\forall i\,,\,\forall t \;\&\; P_i^t \le Z_i\,,\,\forall i\,,\,\forall t\,,$$

after letting $P \in \mathbb{R}^{L\times T}$ be a set of additional decision variables of the model and since the objective function of the RFVB1 model has the following property:

$$\begin{aligned}
&\min_{D\in\mathcal{D}} \sum_t\sum_i\sum_j(\eta - d_{ij} - c_i)D_j^t X_{ij}^t - (C^T Z + K^T I) \\
&\quad = \min_{D\in\mathcal{D}} \sum_t\sum_i\sum_j(\eta - d_{ij} - c_i)D_j^t X_{ij}^t - (C^T Z + K^T I) + \sum_t c^T P^t - c^T P^t \\
&\quad = \min_{D\in\mathcal{D}} \sum_t\sum_i\sum_j(\eta - d_{ij})D_j^t X_{ij}^t - (C^T Z + K^T I) + \sum_t\sum_i c_i(P_i^t - \sum_j D_j^t X_{ij}^t) - c^T P^t \\
&\quad \geq \min_{D\in\mathcal{D}} \sum_t\sum_i\sum_j(\eta - d_{ij})D_j^t X_{ij}^t - (C^T Z + K^T I) - \sum_t c^T P^t \,.
\end{aligned}$$

In this derivation, the last inequality comes from the robust constraint $\sum_j D_j^t X_{ij}^t \leq P_i^t\ \forall D \in \mathcal{D}$ for all i and t. Since this last expression is the objective function of the FVB model, it is clear that the optimal value of this problem will be lower than the value of the RFVB1 model. Now, given that in fact the RFVB1 optimizes the objective function over all W instead of forcing this decision variable to zero as assumed earlier, it necessarily will increase even further the difference between the two bounds.

Finally, while it is clear that $f_{\text{ELAARC}}(I,Z) \leq f_{\text{HD-ELAARC}}(I,Z)$ since the ELAARC model is equivalent to the HD-ELAARC after we introduce the constraint that $X_{ij}^{tt'} = 0$ for all $t \neq t'$, the case for $f_{\text{HD-ELAARC}}(I,Z) \leq f_{\text{MRLTP}}(I,Z)$ needs a little more explanations. To clarify this relation, one needs to remember that for all $D \in \mathbb{R}^{N\times T}$,

$$\begin{aligned}
\sum_t h_t(I,Z,D^t) = \max_{\{Y^t,P^t\}_{t=1}^T} \quad & \sum_t\sum_i\sum_j(\eta - d_{ij})Y_{ij}^t - c^T P^t && \text{(14a)}\\
\text{subject to} \quad & \sum_i Y_{ij}^t \leq D_j^t \,,\ \forall j\,,\ \forall t && \text{(14b)}\\
& \sum_j Y_{ij}^t \leq P_i^t \,,\ \forall i\,,\ \forall t && \text{(14c)}\\
& P^t \leq Z\,,\ \forall t && \text{(14d)}\\
& Y^t \geq 0\,,\ \forall t\,, && \text{(14e)}
\end{aligned}$$

where all temporal decision variables are optimized jointly in a way that can exploit the full information about D, although it is unnecessary to do so because the problem decomposes. Yet, from this perspective, if we replace each Y_t with a history-driven affine function $Y_{ij}^t := \sum_{t'=1}^t (X_{ij}^{tt'})^T D^{t'} + W_{ij}^t$, we necessarily obtain an under evaluation of $\sum_t h_t(I,Z,D^t)$. Note that this argument further indicates that the affine adjustment for each Y^t does not need to be non-anticipative in order to generate a valid lower bound on worst-case profits. □

The result presented in proposition (2) can easily be used to establish guarantees with respect to the optimized bound on worst-case profit that are evaluated by each model.

COROLLARY 1. *Let* f^*_{RC}*,* f^*_{MRLTP}*,* f^*_{FVB}*,* f^*_{RFVB1}*,* f^*_{RFVB1}*,* f^*_{AARC}*,* f^*_{LAARC}*,* f^*_{ELAARC} *and* $f^*_{HD\text{-}ELAARC}$ *respectively be the optimal value of* (2), (3), (5), (6), (7), (8), (9), (11), *and* (13). *The following partial ordering is always satisfied:*

$$f^*_{FVB} \le f^*_{RFVB1} \le f^*_{RFVB2} \le f^*_{LAARC} \le f^*_{ELAARC} \le f^*_{HD\text{-}ELAARC} \le f^*_{MRLTP}$$
$$f^*_{RC} \le f^*_{RFVB1} \le f^*_{AARC} \le f^*_{LAARC}\,.$$

Together, these results show that more sophisticated models of this list always provide better conservative approximation of the optimal value MRLTP model (See Figure 2). In fact, anytime one approximation model in this list returns exactly the optimal value of the MRLTP, all models that are higher or equal to it in this ordering are guaranteed to return an exact optimal solution and exact optimal worst-case bound.

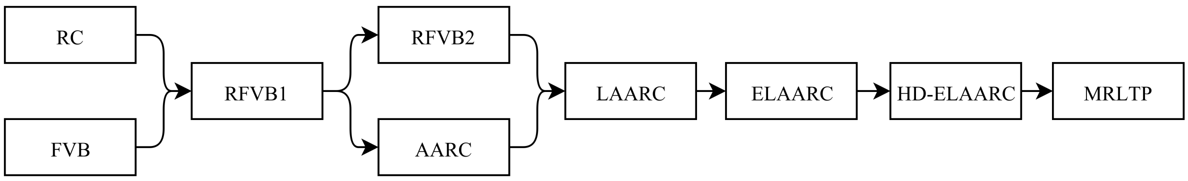


Figure 2 **Partial ordering of the quality of bounds obtained from the different approximation models. Each arrow connects an approximation model to an approximation model that returns a tighter optimized bound for the optimal worst-case profit of the MRLTP model.**

In the following theorem, we present conditions under which some of the proposed approximation models are exact and refer the reader to Appendix C for a detailed proof.

THEOREM 1. *The MRLTP model* (3) *is equivalent to:*

- *RFVB1, RFVB2, AARC, LAARC, ELAARC, and HD-ELAARC when* $C = 0$*,*
- *RC, RFVB1, RFVB2, AARC, LAARC, ELAARC, and HD-ELAARC when* $\Gamma = NT$*,*
- *LAARC, ELAARC, and HD-ELAARC when* $\Gamma = 1$*.*

Intuitively, for the cases of $C = 0$ and $\Gamma = NT$, the proof relies on exploiting the fact that the optimization model used to evaluate $f_{\mathrm{MRLTP}}(Z, I)$ can be shown to reduce to a problem in which the uncertainty decomposes over a number of constraints so that an equivalence between static and adjustable decisions identified in Ben-Tal et al. (2004) can be exploited. Otherwise, in the case of $\Gamma = 1$, our proof follows in spirit the arguments used to support Theorem 1 of Bertsimas and Goyal (2012) yet must address differently the fact that none of the delayed decision variables is a mapping of the whole multi-temporal demand vector. We believe this proof contains elements that might pave the way to a possible extension of the result in Bertsimas and Goyal (2012).

Overall, Corollary 1 and Theorem 1 imply that LAARC, ELAARC, and HD-ELAARC not only provide tighter bounds than all other proposed approximation models but also are optimal for MRLTP for a number of interesting situations.

5. Improving numerical efficiency using row generation algorithm

In this section, we propose a row generation algorithm as a solution method for ELAARC2 which we expect will be more computationally efficient than feeding the MILP reformulation of the model directly to an off the shelf MILP solver. Therefore, we reformulated ELAARC based on the following theorem which proof can be found in Appendix D.

THEOREM 2. *The reduced ELAARC model is equivalent to*

$$\begin{aligned} \underset{I,Z,\rho}{\text{maximize}} \quad & \rho - (C^T Z + K^T I) && \text{(15a)}\\ \text{subject to} \quad & \rho \leq g(Z) && \text{(15b)}\\ & Z \leq MI\,,\, I \in \{0,1\}^L\,, && \text{(15c)} \end{aligned}$$

where $g(Z)$ is defined as

$$\begin{aligned} \min_{\substack{\delta^-,\theta,\lambda,\psi\\ \Theta,\Lambda,\Psi}} \quad & -(C^T Z + K^T I) + \sum_t \sum_i Z_i \theta_i^t + \sum_t \sum_j \lambda_j^t \bar{D}_j^t - \sum_t \sum_j \Lambda_{jj}^t \hat{D}_j^t && \text{(16a)}\\ \text{subject to} \quad & \theta_i^t + \lambda_j^t \geq \eta - c_i - d_{ij}\,, \forall i\,, \forall j\,, \forall t && \text{(16b)}\\ & \Theta_{ik}^t + \Lambda_{jk}^t \geq (\eta - c_i - d_{ij})\delta_k^{t-}\,, \forall i\,, \forall j\,, \forall k\,, \forall t && \text{(16c)}\\ & \sum_k \Theta_{ik}^t \leq \Gamma\theta_i^t, \Theta_{ik}^t \leq \theta_i^t,\ \forall i, \forall k, \forall t && \text{(16d)}\\ & \sum_k \Lambda_{jk}^t \leq \Gamma\lambda_j^t, \Lambda_{jk}^t \leq \lambda_j^t,\ \Lambda_{jk}^t \leq B_j \delta_j^{t-}\,, \forall j\,, \forall k\,, \forall t && \text{(16e)}\\ & \sum_k \Theta_{ik}^t + \lambda_{jk}^t - (\eta - c_i - d_{ij})\delta_k^{t-} \leq \Gamma\left(\theta_i^t + \lambda_j^t - \psi_{ij}^t - (\eta - c_i - d_{ij})\right)\,, \forall i\,, \forall j \forall t && \text{(16f)}\\ & \Theta_{ik}^t + \lambda_{jk}^t - (\eta - c_i - d_{ij})\delta_k^{t-} \leq \theta_i^t + \lambda_j^t - \psi_{ij}^t - (\eta - c_i - d_{ij})\,, \forall i\,, \forall j\,, \forall k\,, \forall t && \text{(16g)}\\ & 0 \leq \delta^- \leq 1\,, \sum_t \sum_j \delta_j^{t-} \leq \Gamma && \text{(16h)}\\ & \lambda \geq 0\,, \Lambda \geq 0\,, \theta \geq 0\,, \Theta \geq 0\,, \psi \geq 0\,, \Psi \geq 0\,, && \text{(16i)} \end{aligned}$$

with $\delta^- \in \mathbb{R}^{N\times T}$, $\theta \in \mathbb{R}^{L\times T}$, $\lambda \in \mathbb{R}^{N\times T}$, $\psi \in \mathbb{R}^{L\times N\times T}$, $\Theta \in \mathbb{R}^{L\times N\times T}$, $\Lambda \in \mathbb{R}^{N\times N\times T}$, and $\Psi \in \mathbb{R}^{L\times N\times N\times T}$.

Based on Theorem 2, we propose employing a row generation algorithm to solve ELAARC2 wherein one goes through the following steps:

Row generation algorithm

Step #1: Set $UB = \infty$ and $LB = -\infty$. Solve the deterministic model with $D = \bar{D}$ to obtain an initial set of facility location $\dot{I}^{(1)}$ and capacities $\dot{Z}^{(1)}$. Let $\kappa = 1$.

Step #2: Solve the following sub-problem

$$\text{(SP)} \quad \underset{\delta^-,\theta,\lambda,\psi,\Theta,\Lambda,\Psi}{\text{minimize}} \quad \sum_t \sum_i \dot{Z}^{(\kappa)} \theta_i^t + \sum_t \sum_j \lambda_j^t \bar{D}_j^t - \sum_t \sum_j \Lambda_{jj}^t \hat{D}_j^t$$
$$\text{subject to} \quad (16b) - (16i).$$

Set $\dot{\theta}^{(\kappa)}$, $\dot{\lambda}^{(\kappa)}$, $\dot{\Lambda}^{(\kappa)}$, and $(\dot{\delta}^-)^{(\kappa)}$ to their respective value based on optimal solution of the above SP model. Let ρ^* be the optimal value of the above SP model. Set $LB = \max(LB, \rho^* - (C^T \dot{Z}^{(\kappa)} + K^T \dot{I}^{(\kappa)}))$.

Step #3: Let $\kappa := \kappa + 1$ and solve the following master problem:

$$\text{(MP)} \underset{I,Z,\rho}{\text{maximize}} \quad \rho - (C^T Z + K^T I) \tag{17a}$$

$$\text{subject to} \quad \rho \le \sum_t \sum_i (\dot{\theta}_i^t)^{(l)} Z_i + \sum_t \sum_j (\dot{\lambda}_j^t)^{(l)} \bar{D}_{jt} - \sum_t \sum_j (\dot{\Lambda}_{jj}^t)^{(l)} \hat{D}_{jt} \quad \forall l \in \{1, 2, \ldots, \kappa - 1\} \tag{17b}$$

$$Z \le MI, \; I \in \{0,1\}^L. \tag{17c}$$

Let $\dot{I}^{(\kappa)}$, $\dot{Z}^{(\kappa)}$, and $\rho^{(\kappa)}$ take on the values of any optimal solution of the master problem (MP). Let $UB = \rho^{(\kappa)} - (C^T \dot{Z}^{(\kappa)} + K^T \dot{I}^{(\kappa)})$.

Step #4: If $UB - LB \le \varepsilon$ then terminate and return $\dot{Z}^{(\kappa)}, \dot{I}^{(\kappa)}$ and $\rho^{(\kappa)}$ as the optimal solution, otherwise repeat from Step #2. (Note that the termination condition can also be verified at the end of Step #2.)

One can actually improve the convergence speed of the algorithm by exploiting a specific type of valid inequalities for the ELAARC problem. Consider that in order for a triplet (I, Z, ρ) to be feasible in problem (15), for any $\{(D^-)^{(l)}\}_{l \in \Omega} \subset \mathcal{D}_3$, there must exist an assignment for X^-, W, and S^- such that the following constraint is satisfied:

$$\rho \le \sum_t \sum_i \sum_j (\eta - d_{ij} - c_i) \left((X_{ij}^{t-})^T (D^{t-})^{(l)} + W_{ij}^t \right) - \sum_t \sum_j B_j (S_j^{t-} (D_j^{t-})^{(l)}) \,,\, \forall l \in \Omega$$
$$\sum_i (X_{ij}^{t-})^T (D^{t-})^{(l)} + W_{ij}^t \le (D_j^{t-})^{(l)} \,,\, \forall l \in \Omega \,,\, \forall j \,, \forall t$$
$$\sum_j (X_{ij}^{t-})^T (D^{t-})^{(l)} + W_{ij}^t \le Z_i \,,\, \forall l \in \Omega \,,\, \forall i \,, \forall t$$
$$(X_{ij}^{t-})^T (D^{t-})^{(l)} + W_{ij}^t \ge 0 \,,\, \forall l \in \Omega \,,\, \forall i \,,\, \forall j \,, \forall t \,.$$

This gives rise to the idea of replacing the master problem with

$$\begin{aligned}
\text{(MP')} \quad \underset{I,Z,\rho,X^-,W,S^-}{\text{maximize}} \quad & \rho - (C^T Z + K^T I) \\
\text{subject to} \quad & \rho \leq \sum_t \sum_i \sum_j (\eta - d_{ij} - c_i)\left((X_{ij}^{t-})^T D^{t-} + W_{ij}^t\right) \\
& \qquad - \sum_t \sum_j B_j S_j^{t-} D^{t-}, \; \forall D^- \in \mathcal{D}_4^\kappa \\
& \sum_i (X_{ij}^{t-})^T D^{t-} + W_{ij}^t \leq D_j^{t-}, \; \forall D^- \in \mathcal{D}_4^\kappa, \; \forall j, \forall t \\
& \sum_j (X_{ij}^{t-})^T D^{t-} + W_{ij}^t \leq Z_i, \; \forall D^- \in \mathcal{D}_4^\kappa, \; \forall i, \forall t \\
& (X_{ij}^{t-})^T D^{t-} + W_{ij}^t \geq 0, \; \forall D^- \in \mathcal{D}_4^\kappa, \; \forall i, \; \forall j, \forall t \\
& \text{(17b)} - \text{(17c)},
\end{aligned}$$

for some well chosen finite set of feasible demand realization $\mathcal{D}_4^\kappa$. In particular, our implementation uses $\mathcal{D}_4^\kappa$ as the set that simply contains the most recently identified worst-case demand $D_j^- := \bar{D}_j - \hat{D}_j(\delta_j^-)^{(\kappa)}$.

One can observe in Table 1, the effect of including such valid inequalities in the decomposition scheme on a set of four problem instances of different sizes. In particular, it might come as a surprise to realize how much the number of iterations is reduced with this simple improvement.

Remark 1. One might alternatively consider the following classical decomposition scheme for robust optimization problem. Start by obtaining the solution of ELAARC for the nominal demand. Then, identify the worst-case realization for the objective and each constraint. Finally, iterate until convergence including in each new iteration the worst-case demand that was generated for each constraint in the previous rounds. Unfortunately, this procedure is somewhat inefficient because of the large difference between the large size of the scenario-based version of ELAARC, which also holds binary variables, and the small size of the linear programming problems that provide the next worst-case demand.

6. Numerical results

In this section, we evaluate the proposed approximation models on a set of randomly generated problem instances. The questions we seek to address are:

- What are the computational requirements of each approximation models and of the proposed row generation algorithm? (Section 6.1),
- What is the impact of varying the amount of uncertainty on the quality of the robust strategy and of the optimized bound proposed by each approximation model? (Section 6.2),
- What is the potential of each model with respect to trading-off average performance and robustness? (Section 6.3),

Table 1 Impact of valid inequalities on row generation algorithm

T	L	N	Γ%	# of iteration			Time (sec)		
				Without VI	With VI	Imp. %	Without VI	With VI	Imp. %
			10	34	16	53	6	3	50
			30	46	37	20	11	8	27
			50	30	27	10	7	6	14
1	10	20	70	27	18	33	5	2	60
			90	23	4	83	5	<1	>80
			100	24	2	92	4	<1	>75
			Avg.	31	17	48	6	<3.5	>42
			10	257	163	37	88	46	48
			30	193	177	8	65	52	20
			50	164	135	18	70	49	30
1	20	40	70	141	93	34	72	57	21
			90	105	20	81	60	22	63
			100	95	2	98	26	<1	>96
			Avg.	159	93	46	64	<37.8	>41
			10	162	58	64	25	6	76
			30	174	89	49	28	15	46
			50	181	91	50	29	18	38
10	10	10	70	159	34	79	25	5	80
			90	159	3	98	26	<1	>96
			100	147	2	99	23	<1	>96
			Avg.	164	46	73	26	<7.6	>71
			100	368	63	83	534	88	84
			30	392	109	72	647	143	78
			50	476	121	75	707	173	76
10	15	15	70	521	99	81	783	134	83
			90	542	15	97	800	20	98
			100	514	2	100	760	2	100
			Avg.	469	68	85	705	93	86

- Are there interesting insights about the structure of the robust decisions that are proposed by each approximation model, namely in terms of number open facilities and total capacity of open facilities, and of statistics about the amount of demand that is covered and the amount of unused capacity under different scenarios? (Section 6.4).

Each of these experiments will employ different sets of problem instances generated randomly according to the following procedure. We randomly generate N nodes on a unit square, representing the demand points and choose randomly L nodes of this N nodes as candidate facility locations. The respective unit transportation cost between a facility and a customer location, d_{ij}, is simply considered equal to the Euclidean distance between the two. For each facility i we draw a value for each parameters η, C_i, and K_i at random uniformly and independently from the intervals $[1.5, 2]$, $[0.5, 0.1]$, $[0, 50000]$ respectively, while the production cost parameter is simply set as $c_i = 0.5$. The specific characterization of demand uncertainty is also randomly generated as follows: for each demand location j and period of time t, the nominal demand $\bar{D}_j^t$ is generated uniformly from the interval $[0, 20000]$ and the maximum demand perturbation is set to $\hat{D}_j^t = \varepsilon_j^t \bar{D}_j^t$ where ε_j^t is drawn randomly between 0.15 and 1.

6.1. Computational Analysis

In this subsection, we compare the computational time associated to the resolution of each approximation model when implemented directly using Optimization Programming Language (OPL) within IBM ILOG CPLEX Optimization Studio 12.6.1, while for ELAARC we also evaluated the performance of our novel row generation algorithm. We are especially interested in comparing these computational times to the computational requirements associated with the exact column and constraint (C&CG) algorithm[3] presented in Zeng and Zhao (2013) for varying size of problem instances and budget of uncertainty Γ.

Table 2 focuses on single-period problems and presents the computation times of three problem instances of different size: the "small" size instance had 10 facility and 10 demand locations, the "medium" size instance had 10 facility and 20 demand locations, and finally the large size instance had 50 facility and 100 demand locations. For each instance, we measured the impact of varying the budget of uncertainty among different proportions of the total number of locations. A second set of computational experiments involved three multi-period instances of different size: the "small" instance had 10 periods, 10 facility and 10 demand locations, while the largest instance had 20 periods, 15 facility and 30 demand locations. Again, we attempted to measure the impact of varying the budget of uncertainty but this time among different proportions of $T \times N$, which is namely the size of the uncertain vector D in each problem instance.

Our first observation is that the customer-driven models (*i.e.,* FVB, RFVB1 and RFVB2) benefit from strong computationally efficiency and can actually be solved even in the case of large problem instances in at most a few seconds. While the market-driven models are more computationally demanding, we observe a significant reduction of computational efforts for the LAARC2 and ELAARC2 models when compared to AARC due to the use of the reduced form identified in Proposition 1. It also appears that for medium size instances the ELAARC2 model becomes slightly easier to solve than LAARC2 even though it involves a larger set of decision variables and constraints. Otherwise, although these two market-driven models can be solved in less than an hour for the medium single-period and multi-period instances, it becomes impossible to obtain a solution during our 48 hours time frame for the largest single-period and multi-period instances. One can obviously explain the difficulty of resolving market-driven models by the fact that the number of degrees of freedom for the affine adjustment grow at the rate of $O(LN^2T)$ instead of $O(LNT)$ for customer-driven models. Comparatively, we observe perhaps with surprise that the C&CG algorithm requires much less efforts than any of these direct implementations. This seems to indicate that the efficiency of the decomposition scheme that is used by C&CG compensates for the fact that C&CG requires the solution of a number of outer and inner mixed integer linear programs. This leaves us with the question of whether our conservative approximation models could also benefit from some well-designed decomposition scheme.

Indeed, looking at the “Row gen.” column of both table, we remark that the time needed to solve the ELAARC2 model can be significantly improved using our proposed row generation algorithm. To be precise, we estimate that this algorithm is responsible for reducing the computation requirements by a factor at least between 16 and 260 (see multi-period instance with $\Gamma = 90\%$ where we have $48 \times 3600/663 = 260$) depending on the size of Γ. Practically speaking, we see that this algorithm allows us to identify robust approximate solutions for the largest single-period and multi-period instances in less than three hours (with an average of less than an hour and a half). In comparison, there is also evidence that the C&CG algorithm is unable to converge in less than 48 hours for the single-period instance when Γ equals 30% and 50% of the number of locations, while it is unable to do so for the large multi-period instance when Γ is greater than 30% of the total number of uncertain parameters (except for the trivial case of box uncertainty). One might finally observe that except for AARC the computational time of all models initially increased as the budget was increased but later decreases back to a lower delay. The reason of this trend might be related to number of extreme points of uncertainty set $\mathcal{D}_3$ which is known to contain the worst-case realizations for at least most of these models.

Regarding the resolution of HD-ELAARC, our experiments indicated that solving this model directly with a MILP solver typically takes about 30 minutes ($80\times$ more difficult than solving ELAARC2) for small size multi-period problem (*i.e.,* $T = 10$, $L = 15$, and $N = 15$)). Because of time limitation, we were unable to experiment with larger problem instances.

Conclusions: While both RFVB1 and RFVB2 models can be solved almost as efficiently as the FVB model, market-driven models should only be solved using standard optimization software when the problem instance is of medium size. For larger sized problem, the use of a row generation algorithm is needed and highly effective for these models. This allows us to provide nearly-exact robust solutions (as shown in the next subsection) for problems where exact solutions are unobtainable. It appears however that much more algorithmic efforts are needed to provide solutions to HD-ELAARC for problems of such large size.

6.2. Optimality gap analysis

In this subsection, we attempt to compare empirically the increasing quality of the approximate robust solutions that are obtained from the different conservative approximation models. Our hope

Table 2 Computational time (in seconds) needed for indentifying approximate and exact robust solutions for three single-period instances of increasing sizes and varying level of budgets (in % of total number of uncertain parameters). The dash "-" denotes situations where the method did not converge in less than 48 hours.

L	N	Γ%	FVB	RFVB1	RFVB2	AARC	LAARC2	ELAARC2	Row gen.	C&CG
10	20	10	<1	<1	<1	4	3	9	3	<1
		30	<1	<1	<1	2	6	10	8	1
		50	<1	<1	<1	11	7	7	6	1
		70	<1	<1	<1	6	13	13	2	1
		90	<1	<1	<1	24	18	28	<1	<1
		100	<1	<1	<1	219	2.6	9	<1	<1
		Avg.	<1	<1	<1	44	8	13	<3.7	<1
20	40	10	<1	<1	<1	521	415	303	46	8
		30	<1	<1	<1	272	264	166	52	11
		50	<1	<1	<1	283	275	191	49	50
		70	<1	<1	<1	581	523	398	57	19
		90	<1	<1	<1	1,747	1,308	1,287	22	3
		100	<1	<1	<1	69,394	2,326	1,011	<1	<1
		Avg.	<1	<1	<1	12,050	852	559	<44	<15
50	100	10	<1	2	6	-	-	-	3,241	8,465
		30	<1	4	11	-	-	-	4,563	-
		50	<1	4	9	-	-	-	8,460	-
		70	<1	5	4	-	-	-	3,781	7,682
		90	<1	4	6	-	-	-	1,382	7
		100	<1	2	2	-	-	-	<1	2
		Avg.	<1	3.5	6.3	-	-	-	<3,572	-

is to quantify, from the perspective of worst-case analysis, what is the actual value in employing a more flexible model. The subsection's development is threefold. We first investigate in single-period problem instances the impact of changing the size of the potential demand perturbations ε and of the uncertainty budget Γ on the quality of these solutions. We then perform a similar analysis for the multi-period setting. Finally, we confirm that there exists multi-period problem instances for which the history-driven model HD-ELAARC can indeed be used to obtain better approximate robust solution than the non-history-driven alternatives.

It is worth clarifying that in what follows, every problem instance was generated using the procedure presented earlier in the introduction of this section with a single exception concerning the size of the potential demand perturbations ε which was fixed to specific values in order to monitor the effect of this parameter. Furthermore, in discussing our finding we will refer to the following values which are worth defining precisely.

- The "optimized worst-case bound" of a conservative approximation model refers to the best lower bound on worst-case profit that can be achieved according to this model. Mathematically, for some model $\mathcal{M} \neq \text{MRLTP}$, this is measured using $f^*_{\mathcal{M}}$

Table 3 Computational time (in seconds) needed for identifying approximate and exact robust solutions for three multi-period instances of increasing sizes and varying level of budgets (in % of total number of uncertain parameters). The dash "-" denotes situations where the method did not converge in less than 48 hours.

T	L	N	Γ%	FVB	RFVB1	RFVB2	AARC	LAARC2	ELAARC2	Row gen.	C&CG
10	10	10	10	<1	<1	<1	25	11	10	6	3
			30	<1	<1	<1	32	25	19	15	1
			50	<1	<1	<1	41	38	21	18	1
			70	<1	<1	<1	115	19	29	5	1
			90	<1	<1	<1	103	23	31	<1	<1
			100	<1	<1	<1	61	32	27	<1	<1
			Avg.	<1	<1	<1	63	25	23	<8.8	<1.5
10	15	15	10	<1	<1	<1	500	342	428	88	1
			30	<1	<1	<1	3,497	1,813	1,916	143	12
			50	<1	<1	<1	4,749	2,770	2,662	173	9
			70	<1	<1	<1	4,815	3,360	3,048	134	36
			90	<1	<1	<1	5,140	3,933	3,681	20	8
			100	<1	<1	<1	6,316	4,431	4,120	2	2
			Avg.	<1	<1	<1	4,170	2,775	2,643	63	11
20	15	30	10	<1	<1	<1	-	-	-	3,781	184
			30	<1	<1	<1	-	-	-	5,646	-
			50	<1	<1	<1	-	-	-	10,567	-
			70	<1	<1	<1	-	-	-	4,445	-
			90	<1	<1	<1	-	-	-	663	-
			100	<1	<1	<1	-	-	-	1	<1
			Avg.	<1	<1	<1	-	-	-	4184	-

- The "achieved worst-case profit" of a strategic decision refers to the actual worst-case profit achieved if this strategic decision is applied. Mathematically for a strategic decision $(I^*_{\mathcal{M}}, Z^*_{\mathcal{M}})$ obtained using model $\mathcal{M}$, this is measured using $f_{\text{MRLTP}}(I^*_{\mathcal{M}}, Z^*_{\mathcal{M}})$.
- The "optimal worst-case profit" of a problem instance refers to the best worst-case profit that can be achieved for this instance. Mathematically, it is measured using f^*_{MRLTP} and obtained in our experiments by solving the C&CG algorithm (see Endnote 3).
- The "relative optimized bound gap" of a conservative approximation model refers to the relative difference between the optimal worst-case profit for this problem instance and the optimized worst-case bound of this model. Mathematically, for some model $\mathcal{M} \neq$ MRLTP, it is measured using $(f^*_{\text{MRLTP}} - f^*_{\mathcal{M}})/f^*_{\text{MRLTP}}$.
- The "relative sub-optimality" of a strategic decision refers to the relative difference between the optimal worst-case profit for this problem instance and the achieved worst-case profit of this decision. Mathematically, for a strategic decision $(I^*_{\mathcal{M}}, Z^*_{\mathcal{M}})$ obtained using model $\mathcal{M}$, it is measured using $(f^*_{\text{MRLTP}} - f_{\text{MRLTP}}(I^*_{\mathcal{M}}, Z^*_{\mathcal{M}}))/f^*_{\text{MRLTP}}$.

6.2.1. Impact of size of potential perturbation on optimality gap

We consider 100 randomly generated problem instances with $L = 10, N = 10$, and $T = 1$. Table 4 presents the average (taken over the set of 100 instances) relative optimized bound gap and the average relative sub-optimality gap for the solutions (*i.e.,* identified strategies for I and Z) of both customer-driven and market-driven type models under different budgets of uncertainty Γ when the demand intervals are forced to a relatively small size, *i.e.,* $\varepsilon = 0.15$. Similarly, Tables 5 and 6 present the same statistics, on the same set of instances, but with medium size $\varepsilon = 0.30$, and large size $\varepsilon = 0.45$ demand intervals.

Regarding the quality of the optimized worst-case bound, one might first observe in these tables that as indicated by Corollary 1, the optimized bounds always improve when one uses a more flexible approximation model. One might further remark that the most significant improvements appear to occur exactly when passing to models that implement the most significant changes in terms of added flexibility and resulting computational needs, namely from the FVB model to the RFVB1 model and later by passing to a market-driven model. When we look at the results for the FVB model and other customer-driven models, we observe that RFVB1 and RFVB2 models reduce by factors of 8 and 13 respectively the quality of the optimized worst-case bound offered by the FVB model. In particular, one might notice that when $\varepsilon = 0.30$, the company always identifies profitability in servicing its customers under the RFVB1 and RFVB2 models while the FVB model suggests shutting down all facilities at $\Gamma = 4$. This is serious evidence that the FVB model is overly conservative. Furthermore, it appears that a significant gain is achieved with the introduction of market-driven policies such that the proposed optimized worst-case bounds are on average always less than 0.59% from being exact. Although the added value of using the LAARC and ELAARC models is not very pronounced (refer to underlined and **bold** entries respectively), the difference becomes more remarkable as the size of demand intervals is increased. Regarding sensitivity to the size of Γ and ε, one might notice that the quality of the optimized worst-case bounds for FVB decreases when the budget of uncertainty increases unlike the other models. We also estimate the quality of the other model's optimized bound to be less affected by the growth of the size of demand intervals.

Regarding the quality of the approximate robust solution itself, we can confirm that employing more flexible adjustments clearly improves the chances of identifying good strategic decisions. For instance, in Table 6 where there is large demand intervals, for $\Gamma = 5$, the FVB model always suggests to construct no facilities foregoing all chances of making any profit (*i.e.,* a 100% worst-case profit loss) while ELAARC provides strategic decisions that on average achieve a worst-case profit that is only 0.28% from being the optimal worst-case profit achievable. ELAARC also provides a guaranteed lower bound on worst-case profits that is on average only 0.50% lower than the optimal worst-case profit. It can also be observed that all of our proposed methods provide optimal robust solutions for the case with $\Gamma = N$ as predicted by Theorem 1, moreover the LAARC and ELAARC models' solutions are also optimal when $\Gamma = 1$.

Table 7 provides additional statistics about the relative sub-optimality of the different solutions proposed by each approximation model in the 3000 problem instances that are surveyed in Tables 4, 5, and 6. Specifically, the table indicates for a number of different percentage gaps, the proportion of instances for which each model was able to identify an approximate robust solution which relative sub-optimality was within that given gap. Each proportion can be interpreted as the likelihood that the solution obtained from a model achieves a worst-case profit that is within some percentage away from being optimal. The table also presents what was the average and maximum relative sub-optimality gap for each model. In particular, one can observe that the flexibility of ELAARC gives it the best chances of providing a solution that achieves a certain level of relative sub-optimality. Yet, one can also remark that in terms of maximum relative sub-optimality gap, LAARC was able to perform slightly better. This serves us as a reminder that optimizing a tighter lower bound on an objective value does not guarantee that a solution of better quality will be obtained, however in most cases one can certainly say that it serves as a great proxy. It is worth also noting that the limited additional flexibility of RFVB1 and RFVB2, compared to FVB, has a significant pay off in terms of relative sub-optimality. For instance, the proportion of problem instances where a guaranteed profit is wasted decreases from 75.0% to almost 1% with the RFVB1 and RFVB2 models. Finally, LAARC and ELAARC never forego on the potential of making positive profit in any of these instances.

6.2.2. Optimality gap analysis in multi-period problems

We consider 100 randomly generated problem instances with $L = 10, N = 10$, and either three or five periods. Table 8 presents the same statistics as in Table 4 but for a set of 100 problem instances with three periods $T = 3$ while the demand perturbation size is forced to $\varepsilon = 0.3$. Alternatively, Table 9 presents the same statistics for $T = 5$. In these tables we observe a similar trend as before except for the perhaps unexpected fact that the RFVB2 model seems to provide on average better quality solutions and bounds than AARC. In particular, as underlined in both tables the average relative optimized bound gap is always very close to 1 or 2% (see the underlined entries) while this same statistic rises to values that are close to 4 or 7% with AARC. This seems to indicate that the flexibility provided by AARC, namely adapting to market information, is less useful than the flexibility provided by RFVB2, namely reacting differently for positive than for negative perturbations. The same observation can be made when comparing the two models' average relative sub-optimality gap. One can additionally confirm that LAARC and ELAARC both provide the best quality solutions and optimized worst-case bound. Furthermore, ELAARC is able to slightly

Table 4 Average relative optimized bound gap (Bound gap) and average relative sub-optimality gap (Opt. gap) for the solutions obtained from each approximation model under different values of budget when ε=0.15

Γ	FVB		RFVB1		RFVB2		AARC		LAARC		ELAARC	
	Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap
1	31.0	11.9	11.3	4.88	5.47	2.78	0.00	0.00	0	0	0	0
2	52.7	25.5	11.4	4.3	7.32	3.49	0.04	0.03	0.02	0.01	0.02	0.01
3	66.5	36.2	8.55	4.44	7.02	2.39	0.08	0.07	0.04	0.04	**0.03**	0.03
4	76.4	42.7	5.84	3.38	5.61	2.48	0.11	0.09	0.05	0.04	**0.04**	0.04
5	83.2	47.2	3.75	2.34	3.75	2.35	0.11	0.08	0.04	0.03	0.04	0.03
6	87.5	49.9	2.15	1.49	2.15	1.49	0.08	0.06	0.01	0.01	0.01	0.01
7	90.0	54.1	1.08	0.83	1.08	0.83	0.06	0.05	0.01	0.01	0.01	0.01
8	91.2	53.9	0.41	0.35	0.41	0.35	0.03	0.03	0.01	0.01	0.01	0.01
9	91.5	55.5	0.09	0.08	0.09	0.08	0.01	0.01	0.00	0.00	0.00	0.00
10	91.5	55.2	0	0	0	0	0	0	0	0	0	0

Table 5 Average relative optimized bound gap (Bound gap) and average relative sub-optimality gap (Opt. gap) for the solutions obtained from each approximation model under different values of budget when ε=0.30

Γ	FVB		RFVB1		RFVB2		AARC		LAARC		ELAARC	
	Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap
1	55.0	26.2	22.6	10.2	11.3	6.02	0.00	0.00	0	0	0	0
2	82.4	47.7	24.4	10.5	15.5	7.53	0.12	0.11	0.06	0.05	**0.05**	0.04
3	95.4	74.5	19.2	11.7	15.7	6.14	0.36	0.25	0.12	0.09	**0.08**	0.08
4	99.7	95.4	13.6	8.96	13.1	6.69	0.59	0.38	0.15	0.11	**0.12**	0.11
5	100	100	9.07	6.45	9.07	6.45	0.79	0.46	0.15	0.1	**0.13**	0.09
6	100	100	5.31	4.01	5.31	4.01	0.75	0.45	0.11	0.08	**0.09**	0.06
7	100	100	2.68	2.14	2.68	2.14	0.50	0.33	0.05	0.05	0.05	0.05
8	100	100	1.02	0.9	1.02	0.9	0.2	0.17	0.03	0.03	0.03	0.03
9	100	100	0.22	0.22	0.22	0.22	0.04	0.04	0.00	0.00	0.00	0.00
10	100	100	0	0	0	0	0	0	0	0	0	0

Table 6 Average relative optimized bound gap (Bound gap) and average relative sub-optimality gap (Opt. gap) for the solutions obtained from each approximation model under different values of budget when ε=0.45

Γ	FVB		RFVB1		RFVB2		AARC		LAARC		ELAARC	
	Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap
1	70.7	34.7	34.8	16.5	17.6	9.34	0.00	0.00	0	0	0	0
2	95.5	75.0	39.7	20.4	25.5	14.3	0.20	0.14	0.20	0.14	**0.11**	0.09
3	100.0	98.9	32.6	22.22	26.87	10.74	1.67	0.90	0.40	0.30	**0.26**	0.21
4	100	100	23.9	17.6	23.2	14.0	3.08	1.54	0.57	0.41	**0.44**	0.37
5	100	100	16.2	12.3	16.2	12.3	3.57	1.82	0.56	0.34	**0.50**	0.28
6	100	100	9.67	7.65	9.67	7.65	2.89	1.61	0.46	0.25o	**0.40**	0.22
7	100	100	4.84	4.03	4.84	4.03	1.70	1.15	0.31	0.27	**0.29**	0.25
8	100	100	1.80	1.57	1.80	1.57	0.61	0.52	0.11	0.06	0.11	0.06
9	100	100	0.36	0.36	0.36	0.36	0.11	0.07	0.01	0.01	0.01	0.01
10	100	100	0	0	0	0	0	0	0	0	0	0

Table 7 Proportion of the 3000 problem instances analysed in Tables 4, 5, and 6 where the relative sub-optimality gap of each approximation model was within a certain range. Average gap and maximum gap are also reported.

Gap range	FVB	RFVB1	RFVB2	AARC	LAARC	ELAARC
$=0\%$	0.0	32.7	32.7	68.2	83.7	85.7
$\leq 0.1\%$	0.0	35.7	36.4	77.0	90.7	92.6
$\leq 1\%$	1.3	47.5	50.6	92.8	98.8	99.2
$\leq 10\%$	7.4	83.4	88.5	99.9	100.0	100.0
$=100\%$	75.0	1.1	1.0	0.0	0.0	0.0
Avg. gap	79.81	5.91	4.7	0.26	0.05	0.04
Max gap	100	100	100	34.67	3.32	4.82

tighten its optimized bound (as shown in **bold**) and obtain solutions that are slightly less sub-optimal when Γ equals 30% of the total number of uncertain parameters. It finally appears based on this experiment that when one uses other models than the FVB model, the quality of the approximate robust solutions improves, for any fixed percentage of uncertainty budget, as the number of time periods increases. This appears a little counter-intuitive yet one might conjecture from this empirical evidence that as the horizon becomes longer, it becomes easier to hedge (or perhaps hide from) the risks related to demand perturbation so that approximation models become more effective at identifying good strategies.

Table 10 repeats the analysis of Table 7 in presenting further statistics regarding the relative sub-optimality of the solutions that are obtained from the different conservative approximation models. All statistics that are presented were assessed on the 1000 problem instances covered in Tables 8 and 9. Again, we see significant improvement for passing from the FVB model to RFVB1 (with the maximum gap being reduced from 100% to 8.37%), and very good odds (*i.e.,* 99.6%) of achieving less than a 1% relative sub-optimality gap with LAARC or ELAARC. Yet, one should realize that the odds of achieving an exact solution with both of these models reduces significantly in this set of multi-period problem instances, namely a reduction from above 84% when $T=1$ to less than 13.2% in this set of multi-period problems.

Table 8 Average relative optimized bound gap (Bound gap) and average relative sub-optimality gap (Opt. gap) for the solutions obtained from each approximation model under different values of budget when $T=3$ and ε=0.3

T	L	N	Γ%	FVB		RFVB1		RFVB2		AARC		LAARC		ELAARC	
				Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap
3	10	10	0.1	51.99	6.84	7.67	3.79	0.45	0.20	4.29	1.52	0.06	0.05	0.06	0.04
			0.3	86.47	23.37	9.53	2.89	1.47	0.60	7.24	2.02	0.37	0.21	**0.35**	0.20
			0.5	85.60	23.35	6.94	1.60	2.42	0.86	4.44	1.56	0.71	0.36	**0.70**	0.36
			0.7	84.70	23.75	3.01	1.32	2.15	0.56	1.73	0.97	0.60	0.35	0.60	0.35
			0.9	84.26	24.46	0.29	0.18	0.28	0.18	0.20	0.14	0.10	0.08	0.10	0.08

Table 9 Average relative optimized bound gap (Bound gap) and average relative sub-optimality gap (Opt. gap) for the solutions obtained from each approximation model under different values of budget when $T = 5$ and ε=0.3

T	L	N	Γ%	FVB		RFVB1		RFVB2		AARC		LAARC		ELAARC	
				Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap
5	10	10	0.1	60.43	3.76	5.30	3.00	<u>0.26</u>	0.10	4.22	2.18	0.03	0.02	0.03	0.02
			0.3	73.11	5.11	5.51	2.82	<u>0.68</u>	0.32	4.65	2.12	0.17	0.10	**0.16**	0.10
			0.5	70.42	6.43	4.56	1.65	<u>1.09</u>	0.55	3.33	1.03	0.31	0.15	**0.30**	0.14
			0.7	68.49	8.25	2.61	0.60	<u>1.19</u>	0.50	1.53	0.52	0.33	0.16	0.33	0.16
			0.9	67.50	9.59	0.36	0.15	0.31	0.11	0.18	0.09	0.06	0.04	0.06	0.04

Table 10 Proportion of the 1000 problem instances analysed in Tables 8 and 9 where the relative sub-optimality gap of each approximation model was within a certain range. Average gap and maximum gap are also reported.

Gap range	FVB	RFVB1	RFVB2	AARC	LAARC	ELAARC
$= 0\%$	0.0	1.1	2.7	1.0	12.6	13.2
$\leq 0.1\%$	0.2	10.1	23.6	14.6	56.0	56.9
$\leq 1\%$	2.0	43.8	92.2	56.2	99.6	99.6
$\leq 10\%$	67.6	100.0	100.0	100.0	100.0	100.0
$= 100\%$	5.7	0.0	0.0	0.0	0.0	0.0
Avg.	15.97	1.77	0.4	1.17	0.15	0.14
Max gap	100	8.37	3.54	6.18	1.07	1.05

6.2.3. Optimized bound gap reduction using HD-ELAARC

Table 11 presents the relative optimized bound gap and the relative sub-optimality gap for the solutions of the ELAARC and HD-ELAARC models in a specific multi-period instance where $T = 3$, $L = 10$, and $N = 10$ drawn according to the procedure used previously for different values of the uncertainty budget Γ. The relative gaps that are reported confirm that HD-ELAARC has the potential to identify a tighter optimized worst-case bound for the MRLTP problem and consequently provides approximate robust solution that slightly improves the relative sub-optimality gap. Yet, we consider this improvement to be somewhat small for passing from a model which size grows with $O(LN^2T)$ to $O(LN^2T^2)$.

Conclusions: It appears based on this analysis that RFVB2 and ELAARC2 are the two models that have the best to offer, compared to other models in their respective class, in terms of trading-off speed of resolution and robustness of the facility location strategy that they are able to identified. Additionally, we observed that customer-driven based approximation models, in particular the FVB model, are sensitive to the size of potential perturbations while the performance of market-driven based models appear to be a little more stable. It also appears that the performance of solutions of conservative approximation schemes somehow benefit from longer horizon problems in which there might be more opportunities to edge or hide from the risk. On the other hand, it appears much more difficult to close the sub-optimality gap in larger problems with the type flexibility that is

Table 11 Relative optimized bound gap (Bound gap) and relative sub-optimality gap (Opt. gap) for the solutions obtained from ELAARC and HD-ELAARC under different values of budget with $\varepsilon = 0.3$

	ELAARC		HD-ELAARC	
Γ %	Bound gap	Opt. gap	Bound gap	Opt. gap
10	0.32	0.32	0.11	0.11
20	0.60	0.45	0.40	0.29
30	0.77	0.48	0.38	0.21
40	0.74	0.34	0.25	0.09
50	0.93	0.14	0.40	0.07
60	0.83	0.40	0.38	0.17
70	0.44	0.21	0.22	0.13
80	0.42	0.19	0.17	0.07
90	0.17	0.17	0.04	0.04
100	0.00	0.00	0.00	0.00

found in customer and market-driven adjustments. There might still be however some hope to close this gap with a history-driven model like HD-ELAARC, yet one would be left with the challenge of designing efficient decomposition scheme for this model.

6.3. Robustness-performance trade-off

In this subsection, we study the robustness and performance of the approximate robust solutions obtained using our different approximation models in a pair of experiment. While the first experiment involves a set of 100 medium size single-period problem instances where $L = 10$ and $N = 20$, the second one involves a set of 100 large size single-period problem instances where $L = 50$ and $N = 100$. Each problem instance is generated according to the procedure described in the introduction of this section. Unlike what was done in the numerical studies of previous sections, we do not wish to evaluate the worst-case performance of the solutions that are obtained but rather estimate what type of balance these solutions can achieve in terms of compromise that needs to be made between potential protection against risk (captured by a percentile) and potential expected profit. In particular, for each problem instance, we evaluate the statistical performance of each approximate robust solution on a set of 100 demand scenarios. To obtain each of these scenario, each customer's demand is independently generated from its respective demand interval using a uniform distribution. In the instances of larger size, due to the duration of the resolution process, we limit our study to the FVB, RFVB1 and RFVB2 models.

In Figure 3, we report the average expected profit and the average 10th percentile profit of each approximation model's solution as total budget for the uncertainty set is varied. The same results are also presented in Figure 4 to highlight what type of compromise can be achieved by adjusting the budget of uncertainty. Considering that a common criticism of robust optimization approaches has been that it provides overly conservative solutions, it might come as a surprise that our results show that a flexible robust optimization approach with an appropriately calibrated uncertainty set (e.g. the LAARC model with $\Gamma = 1$) will provide solutions that outperform the solutions of the deterministic model (1), obtained by setting $\Gamma = 0$, in terms of both expected profit and risk exposure as measured through the 10th percentile. Another interesting observation is that overly conservative solutions might often actually be the result of not injecting enough flexibility in the robust optimization model, as is the case for the FVB and RFVB1 models. The figures clearly show that whether the instance is small or large, it is always worth employing the slightly more sophisticated RFVB2 model to achieve significantly better risk and return trade-off. Figure 3(a) also demonstrate how performance is improved by employing market-driven models.

Conclusions: Our experiments clearly show that whether the instance is small or large, it is always worth employing the slightly more sophisticated RFVB2 model to achieve significantly better risk and return trade-off. Figure 4 also demonstrates how performance is improved by employing market-driven models yet this was not confirmed on large problem instances due to the heavier computational requirements of the resolution methods for these models.

6.4. Decision structure

In this subsection, we study the strategies that are obtained from our approximation models. In particular, we look at characteristics such as the number of facilities that are opened and the total production capacities that is installed. To perform this analysis, we replicate the experiments that were done in section 6.3 with $L = 10$ and $N = 20$. Statistics of these experiments are reported in Table 12. In particular, its first set of rows indicates the proportion of problem instances where at least one facility location was proposed for different levels of uncertainty budgets. Once again, the over conservatism of the FVB model can be observed as the model refuses to open any facility in 43% of instances for a relatively small value of $\Gamma = 2$. In contrast, the proportion of problem instances where no facilities are selected is below 15% for all other approximation models. In the other two sets of rows of Table 12, we reports the number of open facilities and total capacity of the proposed solution averaged over the instances where at least one facility location was selected. Regarding the strategies proposed by each models, one might notice that more flexible models always propose opening a larger number of facilities. However, the same cannot be said of total capacity. In particular, it appears that when $\Gamma = 1$ market-driven models are a bit more cautious with respect to the capacity of its facilities. Increasing the amount of uncertainty has the natural effect of encouraging a smaller number of smaller facilities. It might also be worth underlining the fact that although the FVB model tends to propose the smallest number of facilities, it is misled to promote much larger ones. We believe all these results reaffirm the added value that is obtained by including more flexible policies in the robust optimization model.

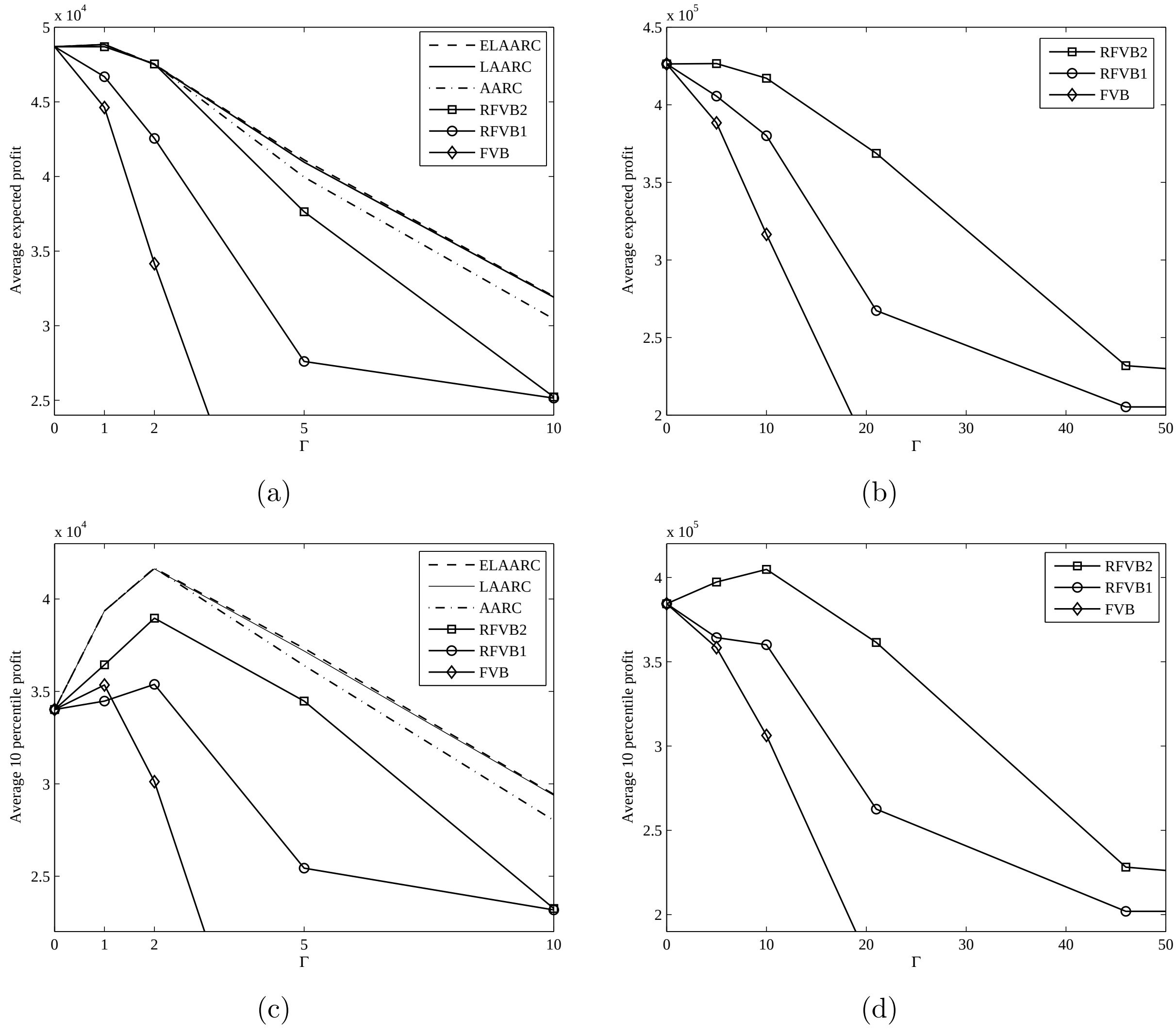
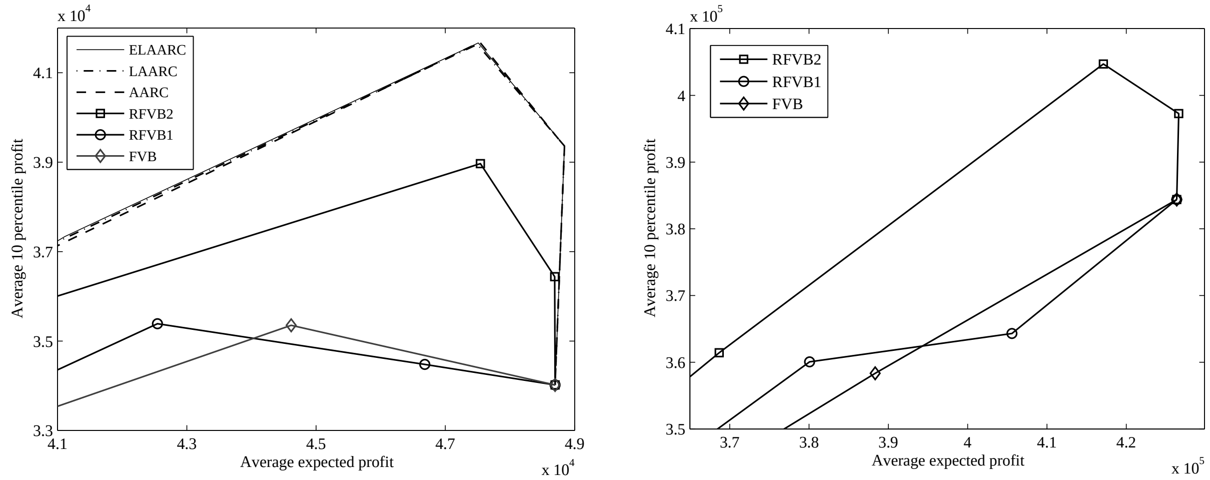


Figure 3 **Average expected and 10th percentile profit achieved by the different robust methods on 100 problem instances when adjusting the level of conservativeness Γ. Figure (a) and (c) present the average expected and average 10th percentile profit respectively for medium sized instances with $L = 10$ and $N = 20$ while (b) and (d) present the same statistics for for large sized instances with $L = 50$ and $N = 100$. Note that in (a) and (b) the curves for LAARC and ELAARC were combined since the performances where indistinguishable.**

We conclude this numerical study with Table 13 which describes how much each approximation model is able to cover the realized demand and make efficient use of its capacity as the uncertainty budget Γ is increased. The first observation one can make is that the percentage of covered demand and the percentage of unused capacity displays increased caution, *i.e.,* a decrease of both percentages, as the models account for increased uncertainty through Γ. We also observe that market-driven models have less unused capacity and cover a larger percentage of demand than



(a) (b)

Figure 4 Average expected profit versus average 10th percentile profit achieved by the different robust methods on 100 problem instances when adjusting the level of conservativeness Γ. Figure (a) presents the achieved risk-return trade-off for instances of medium size while (b) presents it for instances of large size. Note that in (a) the curves for LAARC and ELAARC were combined since the performances where indistinguishable.

other models. Among the customer-driven model, the RFVB2 model appears to use a strategy that resembles much more to the strategies of the market-driven models.

Conclusions: In sum, market driven model propose strategies that open more facilities of smaller sizes. This strategy seems to allow the decision maker to have more flexibility to choose where goods will be shipped from to meet a certain customer's demand. Meanwhile smaller capacities also protects the company from suffering a high rate of unsold products. We also observe that the strategies obtained from market-driven models make better use of the available capacity and cover a larger percentage of the demand than other models.

Table 12 Statistics describing the structure of approximate robust strategies is a set of 100 single-period problem instances with $L = 10$ and $N = 20$

	Γ	FVB	RFVB1	RFVB2	AARC	LAARC	ELAARC
# of instances with open facilities	1	75%	88%	93%	95%	95%	95%
	2	57%	86%	90%	94%	94%	94%
	5	10%	85%	85%	91%	91%	91%
Average # of open facility	1	1.56	1.66	1.86	1.86	1.86	1.86
	2	1.21	1.56	1.82	1.81	1.83	1.83
	5	1.20	1.49	1.61	1.62	1.70	1.71
Average total capacity	1	170227	167479	171089	164867	164797	164806
	2	146404	135752	156456	153134	153582	153905
	5	112252	78879	109826	119999	124363	125493

Table 13 Proportion of demand that is covered and total capacity that is unused averaged over the 100 demand scenarios from each problem instance and over a set of 100 problem instances

Γ	Title	FVB	RFVB1	RFVB2	AARC	LAARC	ELAARC
1	Unused capacity (%)	1.18	2.06	1.52	0.74	0.74	0.74
	Covered demand (%)	63.11	72.46	79.25	79.05	79.02	79.03
2	Unused capacity (%)	0.17	0.79	0.74	0.20	0.21	0.22
	Covered demand (%)	41.93	58.15	70.85	73.02	73.21	73.36
5	Unused capacity (%)	0.00	0.00	0.02	0.00	0.00	0.00
	Covered demand (%)	5.80	34.23	47.43	55.55	57.60	58.12

7. Conclusion

In this paper, we have studied a multi-period robust location-transportation problem with demand uncertainty which was characterized using the budgeted uncertainty set. In order to overcome the known computational difficulty of resolution of this model, we presented six new conservative approximation models that each implement to a different extent the flexibility in the delayed production and transportation decisions. We believe these models, and in particular the RFVB2, ELAARC and HD-ELAARC models, are especially relevant to the transportation literature as the only conservative approximation model that had been presented prior to this work was the FVB model which as Example 1 and our empirical results demonstrated is overly conservative. While this conservativity can easily be corrected for by adding a small amount flexibility to the delayed decisions as is done in the customer-driven RFVB2 model, the solution quality is drastically improved using market-driven models such as the ELAARC. The quality is even further improved using history driven models, *i.e.,* HD-ELAARC, although the number of decision variables in this model quickly makes it prohibitive. As portrayed by Table 14, improving solution quality comes at a price in terms of computational requirements. Therefore, we developed a row generation algorithm that enables us to solve market-driven based approximations for large instances.

Table 14 Summary of the trade-off between flexibility of the adjustments, complexity of the model, and quality of the solution in a multi-period setting. Note that we lack significant evidence about the magnitude of the improvement in quality for HD-ELAARC.

Model	Variables			Total number of variables	Average opt. gap
	P_i^t	Y_{ij}^t	θ_j^t		
HD-ELAARC	$\sum_i Y_{ij}^t$	$\sum_{t'=1}^t \sum_k X_{ijk}^{tt'-} D_k^{t'-} + W_{ij}^t$	$\sum_{t'=1}^t S_j^{tt'-} D_j^{t'-}$	$O(LN^2T^2)$	N/A
ELAARC2	$\sum_i Y_{ij}^t$	$\sum_k X_{ijk}^{t-} D_k^{t-} + W_{ij}^t$	$S_j^{t-} D_j^{t-}$	$O(LN^2T)$	$\sim 0.14\%$
LAARC2	$\sum_i Y_{ij}^t$	$\sum_k X_{ijk}^{t-} D_k^{t-} + W_{ij}^t$	0	$O(LN^2T)$	$\sim 0.15\%$
AARC	$\sum_i Y_{ij}^t$	$\sum_k X_{ijk}^{t} D_k^{t} + W_{ij}^t$	0	$O(LN^2T)$	$\sim 1.17\%$
RFVB2	$\sum_i Y_{ij}^t$	$X_{ij}^{t+} D_j^{t+} + X_{ij}^{t-} D_j^{t-} + W_{ij}^t$	0	$O(LNT)$	$\sim 0.40\%$
RFVB1	$\sum_i Y_{ij}^t$	$X_{ij}^t D_j^t + W_{ij}^t$	0	$O(LNT)$	$\sim 1.77\%$
FVB	P_i^t	$X_{ij}^t D_j^t$	0	$O(LNT)$	$\sim 15.97\%$

A side product of our analysis is to have identified conditions under which full flexibility is not necessary in order to obtain a solution of the best quality possible. This is summarized in Table 15.

Table 15 Conditions for approximation models to identify optimal robust strategic decisions

Condition	RC	FVB	RFVB1	RFVB2	AARC	LAARC	ELAARC	HD-ELAARC
$C=0$	×	×	✓	✓	✓	✓	✓	✓
$\Gamma=1$	×	×	×	×	×	✓	✓	✓
$\Gamma=N$	✓	×	✓	✓	✓	✓	✓	✓

Finally, our numerical study compares the performances of the proposed approximation models in terms of sub-optimality of the approximate robust solution, resolution time, achievable risk-return trade-off, and structure of optimal robust decisions.

Although our work focuses on a location-transportation problem, we expect our methods to be applicable to many other multi-stage robust optimization problem with right-hand side uncertainty that appear in the field of transportation, such as network transportation problem, (e.g. Atamtürk and Zhang (2007)), supply chain network design problem, (e.g. Tsiakis et al. (2001)) and hub location-transportation problem (e.g. Oktal and Ozger (2013)).

As a closing remark, one extension of our models that is worth mentioning arises in situations where some facilities may be shut-down due to a disruption such as a natural disasters. While we refer the reader to An et al. (2014) and references therein for more details on location reliability problems, a simple approach consists of considering a set of binary parameters γ_j^t which indicates whether facility j is shut-down at time t. One can then replace the maximum production constraint in problem (4) with $P_i^t \leq (1-\gamma_i^t)Z_i\,, \forall\, i$, and consider the profit for each period to be a function of I, Z, D^t, and γ^t. If one assumes that the vector of disruption γ lies in a budgeted uncertainty set that is independent from the budgeted uncertainty set used for D, then since $h_t(I, Z, D^t, \gamma^t)$ is concave in γ^t for any fixed values of I, Z, and D^t, one can actually relax γ to be a vector of fractional value without affecting the model and then employ any version of our different form of adjustments. For instance, an AARC model would employ the transportation policy $Y_{ij}^t := (X_{ij}^t)^T D^t + (O_{ij}^t)^T \gamma^t + W_{ij}^t$. Alternatively, an ELAARC approach might also employ affine adjustments for penalized excess variables that are used to relax the production constraints. It remains unclear however what might be sufficient conditions for any of these conservative approximations to return exact solutions in this context.

Endnotes

1. While sources of uncertainty other than demand might affect the performance of facility location decisions and it might be interesting to account for them, in this paper we focus on demand uncertainty as we expect it to have the most impact on the quality of the decision that needs to be made. See for instance Delage et al. (2014) where the authors argue that simply using the expected values of parameters that appear in the objective function already generates solutions that can be considered robust for such multi-period problems.

2. Note that in this paper we will consider a model to be tractable if it can be reformulated as a mixed integer linear program of finite dimension.

3. The column and constraint generation algorithm proposed in Zeng and Zhao (2013) was implemented using the two-stage representation of our multi-period problem where the recourse problem takes the form presented in (14) and exploited a reduction that relies on the one sided uncertainty set presented in $\mathcal{D}_3$.

References

Abouee-Mehrizi, Hossein, Oded Berman, Mohammad Reza Baharnemati. 2014. Designing production-inventory-transportation systems with capacitated cross-docks. *Transportation Science* **48**(1) 121–135.

An, Yu, Bo Zeng, Yu Zhang, Long Zhao. 2014. Reliable p-median facility location problem: two-stage robust models and algorithms. *Transportation Research Part B: Methodological* **64** 54–72.

Atamtürk, Alper, Muhong Zhang. 2007. Two-stage robust network flow and design under demand uncertainty. *Operations Research* **55**(4) 662–673.

Baron, Opher, Joseph Milner, Hussein Naseraldin. 2011. Facility location: A robust optimization approach. *Production and Operations Management* **20**(5) 772–785.

Ben-Tal, Aharon, Alexander Goryashko, Elana Guslitzer, Arkadi Nemirovski. 2004. Adjustable robust solutions of uncertain linear programs. *Mathematical Programming* **99**(2) 351–376.

Bertsimas, Dimitris, David Brown, Constantine Caramanis. 2011. Theory and applications of robust optimization. *SIAM Review* **53**(3) 464–501.

Bertsimas, Dimitris, Frans de Ruiter. 2015. Duality in two-stage adaptive linear optimization: Faster computation and stronger bounds. *working paper* .

Bertsimas, Dimitris, Vineet Goyal. 2012. On the power and limitations of affine policies in two-stage adaptive optimization. *Mathematical Programming* **134**(2) 491–531.

Bertsimas, Dimitris, Melvyn Sim. 2004. The price of robustness. *Operations Research* **52**(1) 35–53.

Chen, Xin, Yuhan Zhang. 2009. Uncertain linear programs: Extended affinely adjustable robust counterparts. *Operations Research* **57**(6) 1469–1482.

Christensen, Tue R. L., Kim Allan Andersen, Andreas Klose. 2013. Solving the single-sink, fixed-charge, multiple-choice transportation problem by dynamic programming. *Transportation Science* **47**(3) 428–438.

Crainic, Teodor Gabriel, Gilbert Laporte. 1997. Planning models for freight transportation. *European Journal of Operational Research* **97**(3) 409 – 438.

Delage, Erick, Sharon Arroyo, Yinyu Ye. 2014. The value of stochastic modeling in two-stage stochastic programs with cost uncertainty. *Operations Research* **62**(6) 1377–1393.

Gabrel, Virginie, Mathieu Lacroix, Cécile Murat, Nabila Remli. 2014a. Robust location transportation problems under uncertain demands. *Discrete Applied Mathematics* **164, 1**(1) 100 – 111.

Gabrel, Virginie, Cécile Murat, Aurélie Thiele. 2014b. Recent advances in robust optimization: An overview. *European Journal of Operational Research* **235**(3) 471 – 483.

Georghiou, Angelos, Wolfram Wiesemann, Daniel Kuhn. 2015. Generalized decision rule approximations for stochastic programming via liftings. *Mathematical Programming* **152**(1-2) 301–338.

Lei, Chao, Wei-Hua Lin, Lixin Miao. 2015. A two-stage robust optimization approach for the mobile facility fleet sizing and routing problem under uncertainty. *Computers & Operations Research* **In press**.

Oktal, Hakan, Asuman Ozger. 2013. Hub location in air cargo transportation: A case study. *Journal of Air Transport Management* **27**(1) 1 – 4.

Tsiakis, Panagiotis, Nilay Shah, Constantinos C. Pantelides. 2001. Design of multi-echelon supply chain networks under demand uncertainty. *Industrial & Engineering Chemistry Research* **40**(16) 3585–3604.

Zeng, Bo, Long Zhao. 2013. Solving two-stage robust optimization problems using a column-and-constraint generation method. *Operations Research Letters* **41**(5) 457 – 461.

Appendix A: Analytical solutions to RC and FVB models in Example 1

For the box uncertainty set, the RC model (2) takes the following form

$$\begin{aligned}
\underset{I,Z,Y,P}{\text{maximize}} \quad & \sum_i \sum_j (\eta - d_{ij}) Y_{ij} - cP - (C^T Z + K^T I) && \text{(19a)}\\
\text{subject to} \quad & \sum_i Y_{ij} \leq \bar{D}_j - \hat{D}_j \,, \forall j && \text{(19b)}\\
& \sum_j Y_{ij} \leq P_i \,, \forall i && \text{(19c)}\\
& P \leq Z \,,\; Z \leq MI && \text{(19d)}\\
& Y \geq 0 \,, \,, I \in \{0,1\}^L \,. && \text{(19e)}
\end{aligned}$$

In the optimal solution of RC model (19), the value of Y_{ij} is equal to zero, since $\eta - c_i - C_i - d_{ij} < 0$, for all i and j where $i \neq j$ and is equal to $\bar{D}_j - \hat{D}_j = 10000 - 5000 = 5000$ for all i and j when $i = j$. In sequence, the optimal value of variables P_i, Z_i and I_i are equal to 5000, 5000 and 1 for all i respectively. Therefore the optimal value of problem (19) is equal to 1000. On the other hand, the FVB model (5) with box uncertainty set takes the form as

$$\begin{aligned}
\underset{I,Z,X,P}{\text{maximize}} \quad & \sum_i \sum_j (\eta - d_{ij})(\bar{D}_j - \hat{D}_j) X_{ij} - \sum_i c_i P_i - (C^T Z + K^T I) && \text{(20a)}\\
\text{subject to} \quad & \sum_i X_{ij} \leq 1 \,, \forall j && \text{(20b)}\\
& \sum_j (\bar{D}_j + \hat{D}_j) X_{ij} \leq P_i \,, \forall i && \text{(20c)}\\
& P_i \leq Z_i \,, \forall i && \text{(20d)}\\
& X_{ij} \geq 0 \,, \forall i, \forall j && \text{(20e)}\\
& Z_i \leq M I_i ,\; I \in \{0,1\} \,, \forall i \,. && \text{(20f)}
\end{aligned}$$

Similar to what we conclude above, the optimal solution has $X_{ij}=0$ for all i and j when $i \neq j$. The optimal value of variable P_i is equal to $(\bar{D}_i + \hat{D}_i)X_{ii}$ for all i and the optimal value of variable Z_i is equal to that of variable P_i for all i. Therefore, the objective function (20a) can be reformulated as

$$\sum_i \eta(\bar{D}_i - \hat{D}_i)X_{ii} - \sum_i (c_i + C_i)(\bar{D}_i + \hat{D}_i)X_{ii} - \sum_i K_i 1_{\{X_{ii}>0\}} = \sum_i (2000X_{ii} - 3000 \times 1_{\{X_{ii}>0\}}) \leq 0\,,$$

where the last inequality comes from $\sum_i X_{ij} \leq 1$. Therefore the optimal value of problem (20) is equal to zero in this example.

Appendix B: Selecting large enough B for problem (10)

Lemma 1. *For any I, $Z \geq 0$, $D^t \geq 0$, the optimal value of problem* (10) *is equal to the optimal value of problem* (4) *when* $B_j = \max_i(\eta - c_i - d_{ij})\,,\forall j$.

Proof: First, as was argued earlier, in problem (4) there is always an optimal solution for which constraint (4c) is tight. This implies that the optimal value of problem (4) is the same as in

$$h_t(I,Z,D^t) = \max_{Y^t} \quad \sum_i \sum_j (\eta - d_{ij} - c_i)Y_{ij}^t \tag{21a}$$

$$\text{subject to} \quad \sum_i Y_{ij}^t \leq D_j^t\,,\forall j \tag{21b}$$

$$\sum_j Y_{ij}^t \leq Z\,,\forall i \tag{21c}$$

$$Y^t \geq 0\,. \tag{21d}$$

Now, given that this problem is feasible, strict duality applies so that its optimal value is equal to the optimal value of the following problem:

$$h_t(I,Z,D^t) = \min_{\lambda^t,\theta^t} \quad \sum_i Z_i\theta_i^t + \sum_j D_j^t \lambda_j^t \tag{22a}$$

$$\text{subject to} \quad \theta_i^t + \lambda_j^t \geq \eta - d_{ij} - c_i\,,\forall i\,,\forall j \tag{22b}$$

$$\lambda^t \geq 0,\ \theta^t \geq 0. \tag{22c}$$

where $\lambda^t \in \mathbb{R}^N$ and $\theta^t \in \mathbb{R}^L$ are dual variables for (21b) and (21c) respectively. It is implied from problem (22) that there is an optimal solution for which λ_j^t is smaller or equal to $\max_i(\eta - c_i - d_{ij})$ for all j and t, therefore one can add to problem (22) the constraint that $\lambda_j^t \leq \max_i(\eta - c_i - d_{ij})$ without affecting its optimal value. By applying duality theory a second time, one can easily confirm that he obtains exactly problem (10) with $\max_i(\eta - c_i - d_{ij})$ in place of every B_j. Hence, this completes the proof. □

Appendix C: Proof of Theorem 1

C.1. Proof of case $C=0$

First, if all $C_i = 0$, then it is necessarily the case that all Z_i's can be as large as MI_i. Next, we replace variables Y_{ij}^t with $X_{ij}^t D_j^t$ and P_i^t with $\sum_j D_j^t X_{ij}^t$ in the recourse problem (4) which makes the recourse problem equivalent to

$$h_t(I,MI,D^t) := \max_{X^t} \quad \sum_i \sum_j (\eta - d_{ij} - c_i)X_{ij}^t D_j^t \tag{23a}$$

$$\text{subject to} \quad \sum_i X_{ij}^t \leq 1\,,\, \forall j \tag{23b}$$

$$\sum_j X_{ij}^t D_j^t \leq M I_i\,,\, \forall i \tag{23c}$$

$$X_{ij}^t \geq 0\,, \tag{23d}$$

where $X^t \in \mathbb{R}^{L \times N}$ are the new decision variables for the t-th period. Constraint (23c) can be replaced by

$$X_{ij}^t \leq I_i\,,\, \forall i\,,\, \forall j\,, \tag{24}$$

since (23c) implies that there can be no shipment when binary variable I_i is equal to 0 and otherwise the shipment can be as large as M. Therefore, the objective function of the MRLTP can be reformulated as

$$\min_{D \in \mathcal{D}} \quad \max_X \quad \sum_t \sum_i \sum_j (\eta - d_{ij} - c_i) X_{ij}^t D_j^t \tag{25a}$$

$$\text{subject to} \quad \sum_i X_{ij}^t \leq 1\,,\, \forall j\,,\, \forall t \tag{25b}$$

$$X^t \leq I_i\,,\, \forall i\,,\, \forall j,\, \forall t \tag{25c}$$

$$X^t \geq 0\,,\, \forall t\,. \tag{25d}$$

Since both feasible sets for D and X are compact, based on Sion's minimax theorem, we can reverse the order of minimization over $\mathcal{D}$ and maximization over X, therefore problem (3) with $C = 0$ can be reduced to

$$\underset{I,X}{\text{maximize}} \quad -KI + \min_{D \in \mathcal{D}} \sum_t \sum_i \sum_j (\eta - d_{ij} - c_i) X_{ij}^t D_j^t \tag{26a}$$

$$\text{subject to} \quad \sum_i X_{ij}^t \leq 1\,,\, \forall j\,,\, \forall t \tag{26b}$$

$$X_{ij}^t \leq I_i\,,\, \forall i\,,\, \forall j\,,\, \forall t \tag{26c}$$

$$X^t \geq 0\,,\, \forall t \tag{26d}$$

$$I \in \{0,1\}^L\,. \tag{26e}$$

where $X^t \in \mathbb{R}^{L \times N}$. Note that problem (26) is equivalent to the RFVB1 model when W is fixed to zero in the later one, hence RFVB1 necessarily achieves an optimal value that is larger since it optimizes over W. Given that Proposition 2 states that RFVB1 optimizes a lower bound on worst-case profit, it is clear that the two models are therefore equivalent. Finally, following Corollary 1, all tighter approximation models are also equivalent to MRLTP. □

C.2. Proof of case $\Gamma = NT$

We recall the following theorem from (Ben-Tal et al. 2004).

THEOREM 3. *(Ben-Tal et al. 2004) The adjustable robust counterpart of two-stage robust optimization problem is equivalent to its RC approximation when the uncertainty affecting every one of the constraints is independent of the uncertainty affecting all other constraints (constraint-wise uncertainty).*

For any fixed I and Z the optimal value of the RC model (2) can be obtained by solving the following problem:

$$\begin{aligned} f_{\mathrm{RC}}(I,Z) := \max_{Y,P} \quad & \sum_t \sum_i \sum_j (\eta - d_{ij}) Y_{ij}^t - c^T P_t - (C^T Z + K^T I) \\ \text{subject to} \quad & \sum_i Y_{ij}^t \le D_j^t \,,\ \forall D \in \mathcal{D}, \forall j \,, \forall t \\ & \sum_j Y_{ij}^t \le P_i^t \,, \forall i \,, \forall t \\ & P^t \le Z \,, \forall t \\ & Y^t \ge 0 \,, \forall t \,. \end{aligned}$$

Noting that in this problem, when $\mathcal{D}$ is a box uncertainty set, the uncertainty does decompose constraint-wise. Hence, according to Theorem 3, the optimal value of this problem is equal to the optimal value of the following "wait-and-see" problem:

$$\max_{D \in \mathcal{D}} \ g(I,Z,D)$$

where

$$\begin{aligned} g(I,Z,D) := \qquad \max_{Y,P} \quad & \sum_t \sum_i \sum_j (\eta - d_{ij}) Y_{ij}^t - c^T P^t - (C^T Z + K^T I) \\ \text{subject to} \quad & \sum_i Y_{ij}^t \le D_j^t, \forall j \,, \forall t \\ & \sum_j Y_{ij}^t \le P_i^t \,, \forall i \,, \forall t \\ & P^t \le Z \,, \forall t \\ & Y^t \ge 0 \,, \forall t \,. \end{aligned}$$

In this problem, all decisions are made once all the information about D is obtained. This necessarily leads to an optimal value that is larger than if each (Y^t, P^t) was adjusted only based on the realized D^t. We thus conclude that

$$f_{\mathrm{RC}}(I,Z) \le f_{\mathrm{MRLTP}}(I,Z) \le \max_{D \in \mathcal{D}} g(I,Z,D) = f_{\mathrm{RC}}(I,Z) \,.$$

Furthermore, based on Corollary 1, RFVB1, RFVB2, AARC, LAARC, ELAARC are optimal in this case and equivalent to the following formulation:

$$\begin{aligned} \underset{I,Z,Y,P}{\text{maximize}} \quad & \sum_t \sum_i \sum_j (\eta - d_{ij} - c_i) Y_{ij}^t - (C^T Z + K^T I) \\ \text{subject to} \quad & \sum_i Y_{ij}^t \le \bar{D}_j^t - \hat{D}_j^t, \forall j, \forall t \\ & \sum_j Y_{ij}^t \le Z_i \,, \forall i, \forall t \\ & Y \ge 0 \,, I \in \{0,1\}^L \,. \quad \square \end{aligned}$$

C.3. Proof of case $\Gamma = 1$

We start by demonstrating that each $h_t(I, Z, D^t)$ is a concave function of D^t.

Lemma 2. *Let $g : \mathbb{R}^m \to \mathbb{R}$ be a function defined as*

$$g(x) := \begin{array}{rl} \max\limits_{y\in\mathbb{R}^n} & c^T y \\ \text{subject to} & Ay \le x \\ & y \in \mathcal{Y}, \end{array}$$

where $y \in \mathbb{R}^n$, for some $c \in \mathbb{R}^n$, some $A \in \mathbb{R}^{m\times n}$, and some compact convex set $\mathcal{Y} \subset \mathbb{R}^n$, and where infeasibility of the optimization problem is interpreted as returning the value $-\infty$. Then, $g(\cdot)$ is a concave function.

Proof: Consider two assignments x_1 and x_2 for which $g(x_1)$ and $g(x_2)$ are finite valued, we should show that $g(\theta x_1 + (1-\theta)x_2) \ge \theta g(x_1) + (1-\theta)g(x_2)$. To do so, first consider that since $g(\cdot)$ is finite valued at x_1 and x_2 and since $\mathcal{Y}$ is compact, there must exist some assignments y_1 and y_2 that respectively achieve the optimum of the optimization problems associated to $g(x_1)$ and $g(x_2)$. Now consider the following:

$$\begin{aligned} g(\theta x_1 + (1-\theta)x_2) &= \sup\{c^T y : y \in \mathcal{Y},\, Ay \le \theta x_1 + (1-\theta)x_2\} \\ &\ge c^T(\theta y_1 + (1-\theta)y_2) = \theta c^T y_1 + (1-\theta)c^T y_2 \\ &= \theta \sup\{c^T y : y \in \mathcal{Y},\, Ay \le x_1\} + (1-\theta)\sup\{c^T y : y \in \mathcal{Y},\, Ay \le x_2\} \\ &= \theta g(x_1) + (1-\theta)g(x_2)\,, \end{aligned}$$

where we used the fact that $y := \theta y_1 + (1-\theta)y_2$ is a valid assignment in the first supremum operation since $\mathcal{Y}$ is convex and $A(\theta y_1 + (1-\theta)y_2) = \theta(Ay_1) + (1-\theta)Ay_2 \le \theta x_1 + (1-\theta)x_2$. □

Since the function $\sum_t h_t(I, Z, D^t)$ is jointly concave in D and the budgeted uncertainty set is polyhedral, a worst-case demand necessarily occurs at one of the extreme points of $\mathcal{D}$. There are $2NT + 1$ extreme points in $\mathcal{D}$ when $\Gamma = 1$: *i.e.,* the nominal demand as the first extreme point, in other extreme points all customers demand get their nominal value for all periods except a single customer location at a single time period where the demand can be either equal to its largest amount or lowest amount. Let us identify each of these extreme point as $\{(D^t)^{(l,\tau)}\}_{(l,\tau)\in\Omega}$ with $\Omega := \{0, 1, \ldots, 2N\} \times \{1, \ldots, T\}$ and where:

$$(D^t)^{(l,\tau)} := \begin{cases} \bar{D}^t & l = 0 \text{ or } \tau \ne t \\ \bar{D}^t + e_l \hat{D}_l & l = 1, \ldots, N \\ \bar{D}^t - e_{l-N}\hat{D}_{l-N} & l = N+1, \ldots, 2N \end{cases},$$

with e_l as the vector of size N with all elements equal to 0 except for the l-th element which is equal to 1.

Therefore, for some fixed I and Z and when the budget is equal to one, the optimal value of the MRLTP model is equivalent to

$$f_{\text{MRLTP}}(I, Z) = \max_{D\in\{D^{(l,\tau)}\}_{(l,\tau)\in\Omega}} \sum_t h_t(I, Z, D^t)\,.$$

Following this argument we have that

$$f_{\mathrm{MRLTP}}(I,Z) = \max_{Y,\rho} \quad \rho - \sum_i (C_i Z_i + K_i I_i) \tag{27a}$$

$$\text{subject to} \quad \rho \leq \sum_t \sum_i \sum_j (\eta - d_{ij} - c_i)(Y_{ij}^t)^{(l,\tau)} \ \forall (l,\tau) \in \Omega \tag{27b}$$

$$\sum_j (Y_{ij}^t)^{(l,\tau)} \leq Z_i \,,\, \forall i,\, \forall (l,\tau) \in \Omega \tag{27c}$$

$$\sum_i (Y_{ij}^t)^{(l,\tau)} \leq (D_j^t)^{(l,\tau)} \,,\, \forall j,\, \forall (l,\tau) \in \Omega \tag{27d}$$

$$(Y_{ij}^t)^{(l,\tau)} \geq 0 \,,\, \forall i, \forall j,\, \forall (l,\tau) \in \Omega \tag{27e}$$

$$(Y_{ij}^t)^{(l,\tau)} = (Y_{ij}^t)^{(0,t)} \,,\, \forall i,\, \forall j,\, \forall (l,\tau) \in \Omega \,,\, \forall t \neq \tau \,, \tag{27f}$$

where $(Y_{ij}^t)^{(l,\tau)} \in \mathbb{R}$ is the recourse decision when scenario (l,τ) occurs, and where the last constraint captures the fact that in the MRLTP model, the decisions for each Y^t only depend on D^t so that the transportation decision should be the same for all vertices where $D^t = 0$. After replacing the variables $(Y_{ij}^t)^{(l)} := (Y_{ij}^t)^{(l,\tau)} = (Y_{ij}^t)^{(0,t)}$, $\forall t \neq \tau$, we alternatively obtain:

$$f_{\mathrm{MRLTP}}(I,Z) = \max_{Y,\rho} \quad \rho - \sum_i (C_i Z_i + K_i I_i) \tag{28a}$$

subject to

$$\rho \leq \sum_{t \neq \tau} \sum_i \sum_j (\eta - d_{ij} - c_i)(Y_{ij}^t)^{(0)} + \sum_i \sum_j (\eta - d_{ij} - c_i)(Y_{ij}^\tau)^{(l)} \,,\, \forall (l,\tau) \in \Omega \tag{28b}$$

$$\sum_j (Y_{ij}^t)^{(l)} \leq Z_i \,,\, \forall i,\, \forall t,\, \forall l = 0, \ldots, 2N \tag{28c}$$

$$\sum_i (Y_{ij}^t)^{(l)} \leq (D_j^t)^{(l,t)} \,,\, \forall j,\, \forall t,\, \forall l = 0, \ldots, 2N \tag{28d}$$

$$(Y_{ij}^t)^{(l)} \geq 0 \,,\, \forall i, \forall j,\, \forall t,\, \forall l = 0, \ldots, 2N \,. \tag{28e}$$

Given that in the LAARC model the objective function and each robust constraint involves expressions that are linear in D, a similar argument as above can be used to also reformulate this model in terms of vertices of the budgeted uncertainty set. This leads to the following problem

$$\begin{aligned}
f_{\mathrm{LAARC}}(I,Z) = \max_{X^+,X^-,W,\rho} \quad & \rho - \sum_i (C_i Z_i + K_i I_i) \\
\text{subject to} \quad & \\
& \rho \leq \sum_t \sum_i \sum_j (\eta - d_{ij} - c_i)((X_{ij}^{t+})^T (D^{t+})^{(l,\tau)} + (X_{ij}^{t-})^T (D^{t-})^{(l,\tau)} + W_{ij}^t) \,,\, \forall (l,\tau) \in \Omega \\
& \sum_j ((X_{ij}^{t+})^T (D^{t+})^{(l,\tau)} + (X_{ij}^{t-})^T (D^{t-})^{(l,\tau)} + W_{ij}^t) \leq Z_i \,,\, \forall i \,,\, \forall t,\, \forall (l,\tau) \in \Omega \\
& \sum_i ((X_{ij}^{t+})^T (D^{t+})^{(l,\tau)} + (X_{ij}^{t-})^T (D^{t-})^{(l,\tau)} + W_{ij}^t) \leq (D_j^t)^{(l,\tau)} \,,\, \forall j \,,\, \forall t,\, \forall (l,\tau) \in \Omega \\
& (X_{ij}^{t+})^T (D^{t+})^{(l,\tau)} + (X_{ij}^{t-})^T (D^{t-})^{(l,\tau)} + W_{ij}^t \geq 0 \,,\, \forall i, \forall j,\, \forall (l,\tau) \in \Omega \,,
\end{aligned}$$

where we characterized the extreme points of $\mathcal{D}_2$ as

$$((D^{t+})^{(l,\tau)}, (D^{t-})^{(l,\tau)}) = \begin{cases} (0,0) & \text{if } t \neq \tau \text{ or } l = 0 \\ (e_l \hat{D}_l, 0) & \text{if } t = \tau \text{ and } l = 1, \ldots, N \\ (0, e_{l-N} \hat{D}_l) & \text{if } t = \tau \text{ and } l = N+1, \ldots, 2N \end{cases} \,,$$

and with $(D_j^t)^{(l,\tau)} := \bar{D}_j + (D_j^{t+})^{(l,\tau)} - (D_j^{t-})^{(l,\tau)}$.

By exploiting the definition of each $(D^t)^{(l,\tau)}$, one can show that the equation above reduces to

$$\begin{aligned} f_{\text{LAARC}}(I,Z) = \max_{X^+,X^-,W,\rho} \quad & \rho - \sum_i (C_i Z_i + K_i I_i) \\ \text{subject to} \quad & \\ & \rho \leq \sum_t \sum_i \sum_j (\eta - d_{ij} - c_i)((X_{ij}^{\tau+})^T (D^{\tau+})^{(l,\tau)} + (X_{ij}^{t-})^T (D^{\tau-})^{(l,\tau)}) \\ & \qquad + (\eta - d_{ij} - c_i) W_{ij}^t \,,\, \forall (l,\tau) \in \Omega \\ & \sum_j (X_{ij}^{t+})^T (D^{t+})^{(l,t)} + (X_{ij}^{t-})^T (D^{t-})^{(l,t)} + W_{ij}^t \leq Z_i \,,\, \forall i \,,\, \forall t \,,\, \forall l = 0,\ldots,2N \\ & \sum_i (X_{ij}^{t+})^T (D^{t+})^{(l,t)} + (X_{ij}^{t-})^T (D^{t-})^{(l,t)} + W_{ij}^t \leq (D_j^t)^{(l,t)} \,,\, \forall j \,,\, \forall t \,,\, \forall l = 0,\ldots,2N \\ & (X_{ij}^{t+})^T (D^{t+})^{(l,t)} + (X_{ij}^{t-})^T (D^{t-})^{(l,t)} + W_{ij}^t \geq 0 \,,\, \forall i, \forall j, \forall l = 0,\ldots,2N \,. \end{aligned}$$

and further manipulations leads to

$$f_{\text{LAARC}}(I,Z) = \max_{X^+,X^-,W,\rho} \quad \rho - \sum_i (C_i Z_i + K_i I_i) \tag{30a}$$

$$\text{subject to} \quad \rho \leq \sum_t \sum_i \sum_j (\eta - d_{ij} - c_i) W_{ij}^t \tag{30b}$$

$$\rho \leq \sum_t \sum_i \sum_j (\eta - d_{ij} - c_i) W_{ij}^t + (\eta - d_{ij} - c_i) X_{ijk}^{\tau+} \hat{D}_k^\tau, \begin{cases} \forall k = 1,\ldots,N \\ \forall \tau = 1,\ldots,T \end{cases} \tag{30c}$$

$$\rho \leq \sum_t \sum_i \sum_j (\eta - d_{ij} - c_i) W_{ij}^t + (\eta - d_{ij} - c_i) X_{ijk}^{\tau-} \hat{D}_k^\tau, \begin{cases} \forall k = 1,\ldots,N \\ \forall \tau = 1,\ldots,T \end{cases} \tag{30d}$$

$$\sum_j W_{ij}^t \leq Z_i \,,\, \forall i \,,\, \forall t \tag{30e}$$

$$\sum_j X_{ijk}^{t+} \hat{D}_k + W_{ij}^t \leq Z_i \,,\, \forall i \,,\, \forall t, \forall k = 1,\ldots,N \tag{30f}$$

$$\sum_j X_{ijk}^{t-} \hat{D}_k + W_{ij}^t \leq Z_i \,,\, \forall i \,,\, \forall t, \forall k = 1,\ldots,N \tag{30g}$$

$$\sum_i W_{ij}^t \leq \bar{D}_j \,,\, \forall j \,,\, \forall t \tag{30h}$$

$$\sum_i X_{ijk}^{t+} \hat{D}_k + W_{ij}^t \leq \bar{D}_j + \hat{D}_k 1_{\{j=k\}} \,,\, \forall j \,,\, \forall t, \forall k = 1,\ldots,N \tag{30i}$$

$$\sum_i X_{ijk}^{t-} \hat{D}_k + W_{ij}^t \leq \bar{D}_j - \hat{D}_k 1_{\{j=k\}} \,,\, \forall j \,,\, \forall t, \forall k = 1,\ldots,N \tag{30j}$$

$$W_{ij}^t \geq 0 \,,\, \forall i \,,\, \forall j \,,\, \forall t \tag{30k}$$

$$X_{ijk}^{t+} \hat{D}_k + W_{ij}^t \geq 0 \,,\, \forall i \,,\, \forall j \,,\, \forall t, \forall k = 1,\ldots,N \tag{30l}$$

$$X_{ijk}^{t-} \hat{D}_k + W_{ij}^t \geq 0 \,,\, \forall i \,,\, \forall j \,,\, \forall t, \forall k = 1,\ldots,N \,, \tag{30m}$$

where we made use of the fact that

$$(X_{ij}^{t+})^T (D^+)^{(l,\tau)} + (X_{ij}^{t-})^T (D^-)^{(l,\tau)} + W_{ij}^t = \begin{cases} W_{ij} & \text{if } t \neq \tau \text{ or } l = 0 \\ X_{ijl}^{t+} \hat{D}_l + W_{ij} & \text{if } t = \tau \text{ and } l = 1,\ldots,N \\ X_{ij(l-N)}^{t-} \hat{D}_{l-N} + W_{ij} & \text{if } t = \tau \text{ and } l = N+1,\ldots,2N \end{cases} .$$

In problem (30), we next reformulate the decision variables W_{ij}, X^{+}_{ijl} and $X^{-}_{i,j,l-N}$ as follows

$$W^t_{ij} \to \dot{Y}^t_{ij0},\ \forall i,\ \forall j,\ \forall t\,,$$
$$X^{t+}_{ijk} \to \frac{\dot{Y}^t_{ijk} - \dot{Y}^t_{ij0}}{\hat{D}_k},\ \forall i,\ \forall j,\ \forall t,\ \forall k = 1,\ldots,N\,,$$
$$X^{t-}_{ijk} \to \frac{\dot{Y}^t_{ij(N+k)} - \dot{Y}^t_{ij0}}{\hat{D}_k},\ \forall i,\ \forall j,\ \forall k = 1,\ldots,N\,.$$

Therefore, problem (30) can be reformulated as

$$f_{\mathrm{LAARC}}(I,Z) = \max_{\dot{Y},\rho}\ \ \rho - \sum_i (C_i Z_i + K_i I_i) \tag{31a}$$
$$\text{subject to}\ \ \rho \le \sum_{t\ne\tau}\sum_i\sum_j (\eta - d_{ij} - c_i)\dot{Y}^t_{ij0} + \sum_i\sum_j (\eta - d_{ij} - c_i)\dot{Y}^{\tau}_{ijl}\,,\ \forall\,(l,\tau)\in\Omega \tag{31b}$$
$$\sum_j \dot{Y}^t_{ijl} \le Z_i\,,\ \forall i\,,\ \forall t\,,\ \forall l = 0,\ldots,2N \tag{31c}$$
$$\sum_i \dot{Y}^t_{ijl} \le \bar{D}_j\,,\ \forall j\,,\ \forall t \tag{31d}$$
$$\sum_i \dot{Y}^t_{ijl} \le \bar{D}_j + \hat{D}_j 1_{\{j=l\}}\,,\ \forall j\,,\ \forall t\,,\ \forall l = 1,\ldots,2N \tag{31e}$$
$$\sum_i \dot{Y}^t_{ijl} \le \bar{D}_j - \hat{D}_j 1_{\{j=l-N\}}\,,\ \forall j\,,\ \forall t\,,\ \forall l = N+1,\ldots,2N \tag{31f}$$
$$\dot{Y}^t_{ijl} \ge 0\,,\ \forall i\,,\ \forall j\,,\ \forall t\,,\ \forall l = 0,\ldots,2N\,. \tag{31g}$$

A careful comparison of problems (28) and (31) can confirm that these are the same so that they will return the same optimal value and identify the same set of optimal solutions for Z and I.

Appendix D: Proof of Theorem 2

We first derive the robust-counterpart of constraint (12b) as

$$\exists s \in \mathbb{R}^{N\times T}, m \in \mathbb{R}^{N\times N\times T},$$
$$\sum_i W^t_{ij} + \Gamma s^t_j + \sum_k m^t_{jk} \le \bar{D}_{jt}\,,\ \forall j\,,\ \forall t \tag{32a}$$
$$s^t_j + m^t_{jj} \ge \hat{D}_{jt}(1 - S^{t-}_j + \sum_i X^t_{ijj})\,,\ \forall j\,,\ \forall t \tag{32b}$$
$$s^t_j + m^t_{jk} \ge \hat{D}_{kt}\sum_i X^t_{ijk},\ \forall j, \forall k \ne j\,,\ \forall t \tag{32c}$$
$$s \ge 0\,,\ m \ge 0\,, \tag{32d}$$

where $\forall\,k$ refers to $\forall\,k = 1,\ldots,N$ as will continue to be the case below. The condition described in (32a)-(32d) can be considered equivalent to the original constraint given that strict duality applies since $\mathcal{D}_3$ is non-empty when $\Gamma \ge 0$.

Similarly, we can derive the robust-counterpart of constraint (12c) as

$$\exists u \in \mathbb{R}^{L\times T}, v \in \mathbb{R}^{L\times N\times T},$$
$$\sum_j W^t_{ij} + \Gamma u^t_i + \sum_k v^t_{ik} \le Z_i\,,\ \forall i\,, \forall t \tag{33a}$$
$$u^t_i + v^t_{ik} \ge \hat{D}^t_k \sum_j X^{t-}_{ijk}\,,\ \forall i\,,\ \forall k\,,\ \forall t \tag{33b}$$
$$u \ge 0\,,\ v \ge 0\,, \tag{33c}$$

and the robust-counterpart of constraint (12d) as

$$\exists x \in \mathbb{R}^{L\times N\times T}, y \in \mathbb{R}^{L\times N\times N\times T},$$

$$-W_{ij}^t + \Gamma x_{ij}^t + \sum_k y_{ijk}^t \leq 0\,, \forall i\,, \forall j\,, \forall t \tag{34a}$$

$$x_{ij}^t + y_{ijk}^t \geq -\hat{D}_k^t X_{ijk}^{t-}\,, \forall i\,, \forall j\,, \forall k, \forall t \tag{34b}$$

$$x \geq 0\,, y \geq 0\,, \tag{34c}$$

and finally, the robust-counterpart of constraint (12e) as

$$S_j^{-t} \geq 0, \forall j\,, \forall t\,. \tag{35a}$$

Therefore the reduced ELAARC can be reformulated as

$$\underset{\substack{I,Z,X^-,W,S^-\\ s,m,u,v,x,y}}{\text{maximize}} \quad \min_{D^- \in \mathcal{D}_3} \sum_t \sum_i \sum_j (\eta - d_{ij} - c_i)(\sum_k X_{ijk}^{t-} D_k^{t-} + W_{ij}^t)$$

$$-(C^T Z + K^T I) - \sum_t \sum_j B_j S_j^{t-} D_j^{t-} \tag{36a}$$

$$(32\text{a}) - (32\text{d})\,, (33\text{a}) - (33\text{c})\,, (34\text{a}) - (34\text{c})\,, (35\text{a}) \tag{36b}$$

$$Z \leq MI,\; I \in \{0,1\}^L. \tag{36c}$$

Since $\mathcal{D}_3$ is compact and convex, one can apply Sion's minimax theorem to reverse the order of maximization over $\{X^-, W, S^-, s, m, u, v, x, y\}$ with the minimization over D^- and then replace the inner maximization by its dual minimization problem. The dual minimization problem joined with the minimization with respect to D^- leads to the following optimization model

$$\begin{aligned}
\min_{\substack{\delta^-,\theta,\lambda,\psi\\ \Theta,\Lambda,\Psi}} \quad & -(C^T Z + K^T I) + \sum_t \sum_i Z_i \theta_{it} + \sum_t \sum_j \lambda_{jt} \bar{D}_{jt} - \sum_t \sum_j \Lambda_{jjt} \hat{D}_{jt} \\
\text{subject to} \quad & \theta_i^t + \lambda_j^t - \psi_{ij}^t = \eta - c_i - d_{ij}\,, \forall i\,, \forall j\,, \forall t \\
& \Theta_{ik}^t + \Lambda_{jk}^t - \Psi_{ijk}^t = (\eta - c_i - d_{ij})\delta_k^{t-}\,, \forall i\,, \forall j\,, \forall k\,, \forall t \\
& \sum_k \Theta_{ik}^t \leq \Gamma \theta_i^t, \Theta_{ik}^t \leq \theta_i^t\,, \forall i\,, \forall k\,, \forall t \\
& \sum_k \Lambda_{jk}^t \leq \Gamma \lambda_j^t, \Lambda_{jk}^t \leq \lambda_j^t,\; \Lambda_{jk}^t \leq B_j \delta_j^{t-}\,, \forall j\,, \forall k\,, \forall t \\
& \sum_k \Psi_{ijk}^t \leq \Gamma \psi_{ij}^t, \Psi_{ijk}^t \leq \psi_{ij}^t\,, \forall i\,, \forall j\,, \forall k\,, \forall t \\
& 0 \leq \delta^- \leq 1\,, \sum_t \sum_j \delta_j^{t-} = \Gamma \\
& \lambda \geq 0\,, \Lambda \geq 0\,, \theta \geq 0\,, \Theta \geq 0\,, \psi \geq 0\,, \Psi \geq 0\,,
\end{aligned}$$

where $\lambda \in \mathbb{R}^{N\times T}$, $\Lambda \in \mathbb{R}^{N\times N\times T}$, $\theta \in \mathbb{R}^{L\times T}$, $\Theta \in \mathbb{R}^{L\times N\times T}$, $\psi \in \mathbb{R}^{L\times N\times T}$, and $\Psi \in \mathbb{R}^{L\times N\times N\times T}$ are the dual variables associated with constraints (32a), (32c)-(32d), (33a), (33b), (34a), and (34b) respectively.

Next, one can further reduce this optimization problem by replacing $\psi_{ij}^t = \theta_i^t + \lambda_j^t - (\eta - c_i - d_{ij})$ and $\Psi_{ijk}^t = \Theta_{ik}^t + \lambda_{jk}^t - (\eta - c_i - d_{ij})\delta_k^{t-}$ everywhere and obtain the model presented in the theorem. It is worth emphasizing that this replacement of variables reduces the rate of growth of the total number of decision variables of the model to $O(LNT)$ instead of $O(LN^2T)$. □