

# Facility Location: A Robust Optimization Approach

Opher Baron, Joseph Milner

Rotman School of Management, University of Toronto, Toronto, Ontario M5S 3E6, Canada  
 opher.baron@rotman.utoronto.ca, milner@rotman.utoronto.ca

Hussein Naseraldin

Department of Industrial Engineering and Management, ORT Braude College of Engineering, Karmiel 21982, Israel, nhussein@braude.ac.il

In this research, we apply robust optimization (RO) to the problem of locating facilities in a network facing uncertain demand over multiple periods. We consider a multi-period fixed-charge network location problem for which we find (1) the number of facilities, their location and capacities, (2) the production in each period, and (3) allocation of demand to facilities. Using the RO approach we formulate the problem to include alternate levels of uncertainty over the periods. We consider two models of demand uncertainty: demand within a bounded and symmetric multi-dimensional box, and demand within a multi-dimensional ellipsoid. We evaluate the potential benefits of applying the RO approach in our setting using an extensive numerical study. We show that the alternate models of uncertainty lead to very different solution network topologies, with the model with box uncertainty set opening fewer, larger facilities. Through sample path testing, we show that both the box and ellipsoidal uncertainty cases can provide small but significant improvements over the solution to the problem when demand is deterministic and set at its nominal value. For changes in several environmental parameters, we explore the effects on the solution performance.

*Key words:* facility location; robust optimization; uncertainty; robust counterpart

*History:* Received: November 2008; Accepted: June 2010 by Panos Kouvelis, after 1 revision.

## 1. Introduction

Facility location problems are often strategic in nature and entail long-term decisions, exposing firms to many uncertainties during the operational lifetime of a facility. When solving such problems, a firm may have to determine the number of facilities to open, their locations, and their capacities. Because of the high fixed costs incurred in changing a network of facilities, a firm may be limited in the frequency in which it reexamines these strategic decisions. After determining its facility network, the firm is often relegated to making operational decisions such as determining production quantities, service levels, and allocation of supply to demand. Thus, facility location problems often have a two-stage approach to solving them, location followed by evaluation. They are thus amenable to solution using various techniques of decision making under uncertainty.

There are alternate approaches to solve such problems. In general, as discussed below in our literature review, one may either assume some stochastic information about the possible outcomes or assume that uncertain information lies within some mathematical structure with no distributional information about it. In the former case, one would then optimize some expectation of the cost or profit of a system. In the latter case, one often solves for a solution that is

robust to the uncertainty in the problem statement. We take this approach.

This paper considers a facility location problem using a robust optimization (RO) modeling approach where demand is uncertain and no probability distribution is assumed. We consider the problem where a firm must establish fixed facilities at a charge in an initial period. Further, with a linear cost per unit, the firm must establish the maximum production capacity for each facility. In subsequent periods, the firm observes demand from several nodes and must determine the allocation of demand and production at open facilities. We allow for production costs at the facilities and delivery charges (assumed based on distance to the demand nodes). The production of each facility in each period is bounded by the maximum production capacity initially established. As a result, we allow for lost demand.

To tackle this problem, we apply the RO approach that differs from previous work on robust facility location as discussed in the literature review. This approach was independently developed by El-Ghaoui and Lebret (1997) and Ben-Tal and Nemirovsky (1998). The key issue in RO is how to model the uncertainty. Typically, the uncertainty is confined to a specific mathematical structure in order to ensure model tractability. The RO approach has been applied in recent years in operational, engineering, and financial contexts, e.g., pricing,

inventory, truss-design, portfolio selection, and scheduling. We give a brief tutorial on RO in section 2.

This paper contributes to the literature by formulating robust models for a facility location problem, and by providing insights into the solution structure. We consider a profit-maximizing objective where the firm must trade off establishing sufficient capacity with uncertain (and possibly low) revenue in the future. The mathematical program developed finds solutions that simultaneously address these alternatives. We present a tractable formulation that can be tested in a consistent manner against an approach where uncertainty is ignored. We study the problem through extensive numerical testing and provide a road map to understanding how RO improves system performance. In these tests we consider a semi-realistic problem with 15 nodes and up to 20 periods of demand. While realistic problems may be larger than this, alternate approaches, such as stochastic programming, cannot efficiently solve even this limited-size problem. We indicate environmental parameters that have greater influence on the performance. To the best of our knowledge, this work is the first to apply the RO approach in the context of facility location problems.

The paper is organized as follows: In section 2, we review some of the relevant literature and present the principles of RO. In section 3, we describe the setting, model the problem, and discuss our assumptions. In section 4, we compare the performance of alternate RO models and that of the nominal model in order to demonstrate the expected potential benefits of applying the RO approach to facility location. Finally, in section 5, we highlight the insights gained in this study and propose several future research directions.

## 2. Literature Review

In this section, we briefly review some of the relevant studies in two streams of research that are related to our problem: facility location and RO. The problem of facility location is an important strategic one and many models have been developed to approach it. For texts on the subject, we refer the reader to Mirchandani and Francis (1990), Daskin (1995), and Drezner and Hamacher (2002). In this paper, we consider a variant of the capacitated fixed-charge multi-period facility location problem where the production capacity of each facility must be determined before observing demand during the horizon. That is, we consider the problem under uncertainty.

There is a considerable body of literature on facility location under uncertainty. The recent review article by Snyder (2007) describes two types of problems: stochastic location problems and robust location problems. For the former, as in general stochastic optimization problems, the solution concept is to optimize the

expected value of the objective function. Objective functions studied include minimizing the expected travel cost such as in stochastic versions of the  $p$ -median problem (Mirchandani et al. 1985, Weaver and Church 1983) and maximizing the expected profit of a capacitated  $p$ -median problem or the profit of an inventory–location model (Daskin et al. 2002, Snyder et al. 2007). The closest model in structure to our problem is that of Louveaux (1986), who studies a capacitated fixed-charge location problem with stochastic information on demands, prices, and costs. By including a budget constraint, the author makes the problem equivalent to a capacitated  $p$ -median problem. Researchers have also studied versions of these problems that balance mean and variance of performance (see, e.g., Verter and Dincer 1992). Stochastic optimization approaches also include attempts to maximize the probability that an event such as the total distance from a location is within a limit (see, e.g., Berman and Wang 2006).

In contrast, robust location problems seek to determine locations that are in some sense robust to uncertainty in problem parameters. Often this takes the form of a worst-case performance over a set of possible scenarios or intervals of uncertainty. For these problems, minimax-regret and minimax-cost are applied. For example, Averbakh and Berman (1997) consider a minimax-regret formulation of the weighted  $p$ -center problem on a network with uncertain weights lying in known intervals, establishing that it may be solved by solving  $n+1$  deterministic  $p$ -center problems. Subsequently, Averbakh and Berman (2000a) consider the minimax-regret 1-center problem with uncertain node weights and edge lengths, both restricted to known intervals. They establish a polynomial-time algorithm for the case of a tree network and an  $\varepsilon$ -optimal solution for general networks. Similarly, Averbakh and Berman (2000b) consider the minimax-regret 1-median problem with interval uncertainty of the nodal demands and present a polynomial-time algorithm for the general problem.

Alternate definitions of robustness are also applied to location problems. For example, Carrizosa and Nickel (2003) propose a robust positioning of a facility on the plane, where robustness is defined as the minimal change in the uncertain parameters such that a solution location becomes inadmissible with respect to a total cost constraint. Kouvelis et al. (1992) consider finding solutions where the relative regret under any scenario is limited to be less than some percentage. This measure is also applied in Gutierrez and Kouvelis (1995) to the problem of selecting suppliers when considering exchange rate uncertainty. Snyder and Daskin (2006) apply this measure (they refer to it as  $p$ -robustness) to the  $p$ -median problem and the uncapacitated facility location problem. They solve for the minimum expected cost solution that is  $p$ -robust.

The RO approach was developed independently by El-Ghaoui and Lebter (1997), El-Ghaoui et al. (1998), and Ben-Tal and Nemirovsky (1998, 1999). A recent textbook compiling the research is Ben-Tal et al. (2009). The approach provides less conservative solutions than earlier worst-case solutions provided by robust mathematical programming approaches (e.g., Soyster 1973) by trading off some of the conservatism with improvement in the objective function by bounding the set of values uncertain parameters could achieve. A key feature of the RO approach is its tractability, which depends on the structure of the uncertainty set. Ben-Tal and Nemirovsky (1998, 1999) have shown that, if the uncertainty set is described as a box or an ellipsoid, then the robust formulation of the problem is tractable. (Here a box is defined by a set of linear inequalities while an ellipsoid is given by a quadratic relation among the uncertain parameters.) Bertsimas and Sim (2006) discuss the tractability of different types of robust problems. Bertsimas et al. (2004) explore the robust counterpart of a problem with an uncertainty set that is described by an arbitrary norm.

RO has been used in various areas such as portfolio management, e.g., Ben-Tal et al. (2000), dynamic pricing, e.g., Adida and Perakis (2006), contracts in the supply chain, e.g., Ben-Tal et al. (2005), project management, e.g., Cohen et al. (2007), and inventory management, e.g., Bertsimas and Thiele (2006).

The closest work to ours in applying robustness concepts to facility location is Daskin et al. (1997) and Chen et al. (2006). Daskin et al. (1997) define a concept of  $\alpha$ -reliability. For a given  $0 < \alpha < 1$ , they find the solution to the  $p$ -median problem that minimizes the maximum regret, subject to a constraint that the scenarios used in defining the regret have a total probability of occurring of at least  $\alpha$ . This notion of  $\alpha$  is similar to the ellipsoid uncertainty set considered in RO in the sense that unlikely scenarios may be excluded while some total amount of uncertainty is included in the problem. Chen et al. (2006) introduce the concept of mean-excess regret, in comparison with the minimax-regret presented in Daskin et al. (1997). They show that the mean-excess regret model is more tractable, computationally speaking, thus much more appealing to real-world problems. While in the minimax-regret approach the  $\alpha$ -quantile of regret is minimized, with the  $\alpha$ -reliable mean-excess regret approach, the authors seek to minimize the expected weighted regret of all scenarios in the  $\alpha$ -set.

### 2.1. RO—A Short Tutorial

In this section, we provide a quick reference to the principles of RO. For further details, we refer the reader to Ben-Tal et al. (2009) and references therein.

Readers who are familiar with this approach can safely skip this section. Consider the following linear program (LP):

$$(LP) \quad \min_{x \in \mathcal{X}} c^T x, \quad (1)$$

$$\text{s.t.} \quad Ax \geq b,$$

where  $x \in \mathcal{X} \subseteq \mathbb{R}^n$  is the vector of decision variables,  $b \in \mathbb{R}^m$  is the right-hand side parameter vector,  $c \in \mathbb{R}^n$  is the vector of objective function coefficients, and  $A \in \mathbb{R}^{m \times n}$ , with elements  $a_{ij}$ , is the constraint coefficient matrix.

In a typical problem like (LP), we assume that  $c$ ,  $A$ , and  $b$  are deterministic and we solve the problem and obtain an optimal solution. The RO approach considers some of the data parameters as uncertain, yet lying within a set that expresses limits on the uncertainty. That uncertainty set then defines the limits on uncertainty that a solution will be immunized against. That is, the solution  $x$  will address any possible uncertainty lying within the set. In the RO approach the (LP) is transformed into a robust counterpart by replacing each constraint that has uncertain coefficients with a constraint that reflects the incorporation of the uncertainty set. In the following description, we focus on uncertainty in row  $i$  in the constraint matrix; uncertainty in the objective coefficients can be accommodated as well. Let  $\tilde{a}_{ij}$  denote an uncertain entry in the matrix. We consider two types of uncertainty sets, a box uncertainty set and an ellipsoidal uncertainty set.

**2.1.1. Box Uncertainty.** Under box uncertainty,  $\tilde{a} = \{\tilde{a}_{ij}\}_{i=1, \dots, m, j=1, \dots, n}$  is unknown but bounded in a box of the form  $U^B = \{\tilde{a}_{ij} \in \mathbb{R} : |\tilde{a}_{ij} - \bar{a}_{ij}| \leq \varepsilon G_{ij} : i = 1, 2, \dots, m, j = 1, 2, \dots, n\}$ , where  $\bar{a}_{ij}$  is the mean or nominal value of  $\tilde{a}_{ij}$ ,  $G_{ij}$  represents the uncertainty scale particular to the given entry, and  $\varepsilon$  is the uncertainty level common across all the entries (Ben-Tal et al. 2005). The specific case where  $G_{ij} = \bar{a}_{ij}$  corresponds to the case where the relative deviation from the nominal is at most  $\varepsilon$ . That is, for each  $i$  and  $j$ ,  $\tilde{a}_{ij}$  lies in the interval  $[\bar{a}_{ij} - \varepsilon|\bar{a}_{ij}|, \bar{a}_{ij} + \varepsilon|\bar{a}_{ij}|]$ . Below, we adopt this case as we focus on the relative uncertainty around positive demand values in our network.

For a constraint  $i$ ,  $\sum_j \tilde{a}_{ij} x_j \geq b_i$ , we only need to augment the left-hand side (LHS) of the equation to reflect the uncertainty set in the formulation. Formally, in the augmented constraint we require, for a given solution  $x$ , that

$$\min_{\tilde{a} \in U^B} \left\{ \sum_{j=1}^n \tilde{a}_{ij} x_j \right\} \geq b_i \quad \text{or}$$

$$\min_{\tilde{a}_{ij} : |\tilde{a}_{ij} - \bar{a}_{ij}| \leq \varepsilon |\bar{a}_{ij}|} \left\{ \sum_{j=1}^n \tilde{a}_{ij} x_j \right\} \geq b_i.$$

Given the structure of  $U^B$ , the optimal solution of the optimization on the LHS is

$$\sum_{j=1}^n \bar{a}_{ij} x_j - \varepsilon \sum_{j=1}^n |\bar{a}_{ij}| |x_j| \geq b_i, \quad (2)$$

which can be reformulated as

$$\sum_{j=1}^n \bar{a}_{ij} x_j - \varepsilon \sum_{j=1}^n |\bar{a}_{ij}| y_j \geq b_i, \quad (3)$$

$$-y_j \leq x_j \leq y_j \quad \text{for all } j. \quad (4)$$

The constraints (1) are now replaced by (3) and (4), the robust counterpart constraints, and the robust LP is solved.

The box uncertainty set provides a tighter constraint and is thus more conservative than the original (LP). To see this, assume that in problem (LP)  $a_{ij} = \bar{a}_{ij}$  for all  $i, j$  (the coefficients are at their nominal values). Then the LHS of each constraint (3) is smaller than its counterpart in (LP).

**2.1.2. Ellipsoidal Uncertainty.** Under ellipsoidal uncertainty,  $\tilde{a}_{ij}$  is chosen from a set that expresses a measure of the uncertainty spread. As argued in Bertal and Nemirovsky (1998, 1999), the box uncertainty set may lead to conservative solutions, as it allows all the uncertain parameters to take on their worst values at the same time. Consider constraint  $i$ ,  $\tilde{a}_i^T x \geq b_i$ , where  $\tilde{a}_i = \{\tilde{a}_{i1}, \dots, \tilde{a}_{im}\}$ ,  $i = 1, 2, \dots, m$ . As an alternative to the box uncertainty set, the uncertain parameters can be characterized by both a nominal vector,  $\bar{a}_i$ , and a positive definite matrix,  $C_i \in R^{n \times n}$ , that expresses the relationship between alternate dimensions of the uncertainty set. Consider the set

$$U_i^E = \{a_i \in R^n : (a_i - \bar{a}_i)^T C_i^{-1} (a_i - \bar{a}_i) \leq \Omega_i^2\},$$

where  $\Omega_i$  is a safety parameter indicating the amount of uncertainty, with respect to constraint  $i$ , to be covered by the robust solution. Because  $C$  is positive definite, the set  $U_i^E$  is an ellipsoid.

For a candidate solution vector  $x \in R^n$ , if we restrict uncertain parameters,  $\tilde{a}_i$ , to be in the ellipsoid  $U_i^E$ , we have the augmented constraint

$$\min_{\tilde{a}_i \in U_i^E} \tilde{a}_i^T x \geq b_i. \quad (5)$$

The LHS of (5) can be solved using the Karush-Kuhn-Tucker conditions to find

$$\tilde{a}_i = \bar{a}_i - \frac{\Omega_i}{\sqrt{x^T C_i x}} C_i x.$$

Then analogous to the transformation of (2) into (3) and (4), the robust counterpart to the constraint

$\tilde{a}_i^T x \geq b_i$  implied by (5) is

$$\bar{a}_i^T x - \Omega_i \sqrt{x^T C_i x} \geq b_i. \quad (6)$$

If one were to assign stochastic uncertainty to the coefficients of matrix  $A$ , the solution in (6) suggests that  $\bar{a}_i$  and  $C_i$  would represent the expected values,  $E[\tilde{a}_i]$ , and the covariance matrix,  $E[(\tilde{a}_i - \bar{a}_i)(\tilde{a}_i - \bar{a}_i)^T]$ , respectively, and thus, for a given solution vector,  $x$ , one can consider  $\tilde{a}_i^T x$  to be a random variable with expected value  $\mu_x = \bar{a}_i^T x$  and standard deviation

$$\begin{aligned} \sigma_x &= \sqrt{E[\tilde{a}_i^T x - \mu_x]^2} \\ &= \sqrt{x^T E[(\tilde{a}_i - \bar{a}_i)(\tilde{a}_i - \bar{a}_i)^T] x} \\ &= \sqrt{x^T C_i x}. \end{aligned}$$

In this case, (6) implies that the constraint holds with  $\Omega_i$  standard deviations of slack.

In summary, by using an ellipsoid set, we restrict all of the parameters from obtaining their worst-case values simultaneously. Moreover, if we are to interpret the ellipsoid as representing the mean and covariance structure of the uncertain parameters with a Gaussian distribution, then in the ellipsoidal uncertainty case, the constraints are stochastically guaranteed to a given level through an appropriate choice of  $\Omega$ .

### 3. A Multi-Period Facility Location Problem

In this paper, we consider the application of RO to a capacitated multi-period fixed-charge facility location problem. A firm seeks to locate facilities on a connected graph and establish the maximum capacity of these facilities. These decisions are made once at the start of a time horizon and constitute the strategic level decisions of the firm. Then, in each period, the firm observes demand from nodes in the network and must determine how much demand to satisfy and from which facilities. These operational-level decisions are made subject to the strategic decisions; demand can only be served from open facilities and the total demand served from a facility must be less than its maximum capacity. The firm incurs costs for opening facilities, establishing capacity, producing in a period at a facility, and shipping from a facility to demand nodes. It receives a revenue for each satisfied unit of demand. The objective is to maximize the total system profit.

We assume that the capacity is infinitely divisible and cannot be adjusted during the horizon, though production can be adjusted without cost from period to period. Demand that is fulfilled is done so by direct shipment from a production facility or several facilities to the demand node (i.e., we do not solve for delivery tours). We

assume that inventory cannot be carried from period to period. This may be plausible in a service or make-to-order firm, as well as in a high volume JIT production environment. However, it is a limitation of the model in an environment where inventory may be built up, e.g., to address seasonal demand.

### 3.1. Nominal Problem Formulation

We formulate the nominal problem assuming that all data in the problem are deterministic and known, including future demand. Let  $G(N, A)$  be a connected graph with  $|N|$  nodes and arc set  $A$ . Let  $d_{ij}$  be the distance, and, without loss of generality, the cost of shipping units, between nodes  $i$  and  $j$  for each  $i, j \in N$ . We assume  $d_{ij}$  is symmetric and  $d_{ii} = 0$  for all  $i \in N$ . Let  $T$  be the length of the horizon and  $D_{it}$ ,  $i \in N$ ,  $t = 1, \dots, T$  be the demand in period  $t$  at node  $i$ .

Let  $K_i$  be the cost of opening a facility at node  $i$ . Let  $C_{i0}$  be the cost per unit of capacity established at the beginning of the horizon and let  $c_{it}$  be the cost per unit of production at node  $i$  in period  $t$ . Fulfilled demand generates a revenue of  $\eta > 0$  per unit. Unsatisfied demand is lost. We do not include a discount factor for value over time; one may be readily included.

We define the following decision variables: Let  $I_i = 1$  if a facility is opened at node  $i$ ;  $I_i = 0$  otherwise. Let  $Z_{i0}$  be the (maximum) capacity of an open facility at node  $i$ . Let  $X_{ijt}$  be the proportion of demand at node  $j$  in period  $t$  that is satisfied by an open facility at node  $i$  ( $0 \leq X_{ijt} \leq 1$ ). For expositional ease, we also introduce the decision variable  $Z_{it}$ , the production at node  $i$  in period  $t$ . Finally, let  $\tau$  be the profit over the horizon. The problem is formulated as follows:

$$(P') \max_{x, Z, I, Z_0, \tau} \tau,$$

$$\text{s.t.}$$

$$\sum_{i=1}^N \sum_{j=1}^N \sum_{t=1}^T (\eta - d_{ij}) D_{jt} X_{ijt} - \sum_{i=1}^N \sum_{t=1}^T c_{it} Z_{it} \quad (7a)$$

$$- \sum_{i=1}^N (C_{i0} Z_{i0} + K_i I_i) \geq \tau,$$

$$\sum_{j=1}^N D_{jt} X_{ijt} \leq Z_{it} \quad \text{for all } i, t, \quad (7b)$$

$$\sum_{i=1}^N X_{ijt} \leq 1 \quad \text{for all } j, t, \quad (7c)$$

$$Z_{i0} \leq M I_i \quad \text{for all } i, \quad (7d)$$

$$Z_{it} \leq Z_{i0} \quad \text{for all } i, t, \quad (7e)$$

$$X_{ijt} \geq 0; I_i \in \{0, 1\} \quad \text{for all } i, j, t. \quad (7f)$$

Constraint (7a) represents the objective, which is placed in the constraint set to aid the RO formulation below. The first term expresses the revenue less delivery costs; the second term, the production costs; the third, the capacity and facility location cost. Constraint (7b) guarantees the demand satisfied by a facility is less than its production. Constraint (7c) implies at most 100% of the demand at a node is fulfilled. Constraint (7d) ensures that capacity is only available at open facilities. Constraint (7e) implies the production at a facility in each period is less than the established maximum capacity.

If the values of all the parameters were known, problem (P') could be solved as a mixed integer linear program (MILP) for the optimal strategic decisions and operational decisions. However, if the problem data were uncertain, that solution would not necessarily be optimal. Let  $\tilde{D}_{jt}$ ,  $j \in N$ ,  $t = 1, \dots, T$  be the uncertain demand at node  $j$  in period  $t$ . (For simplicity, we consider only demand uncertainty. Other uncertainties such as price, shipping cost, or production cost uncertainty can readily be included.) Using the RO approach, we now reformulate P' to reflect the allowed uncertainty and solve for both the strategic decisions (location and capacity) and the operational decisions (production and allocation).

### 3.2. The Robust Problem—Box Uncertainty

For the case of box uncertainty, we assume that demand in each period is unknown and bounded on a symmetric interval around a known nominal value. That is, let  $\tilde{D}_{jt}$  be the uncertain demand from node  $j$  in period  $t$  and let  $\bar{D}_{jt}$  be its nominal value, i.e., the center point of the interval. Then we assume  $\tilde{D}_{jt}(1 - \varepsilon_t) \leq \bar{D}_{jt} \leq \tilde{D}_{jt}(1 + \varepsilon_t)$  for  $0 \leq \varepsilon_t \leq 1$  for all  $t$ , where  $\varepsilon_t$  measures the uncertainty size in period  $t$ . We let  $U_{jt}^B = [\tilde{D}_{jt}(1 - \varepsilon_t), \tilde{D}_{jt}(1 + \varepsilon_t)]$ ,  $\bar{U}_t^B = U_{1t}^B \times \dots \times U_{Nt}^B$ , and  $U^B = \bar{U}_1^B \times \dots \times \bar{U}_T^B$ . We now transform problem (P') into a new problem expressing uncertainty by substituting  $\tilde{D}_{jt}$  for  $D_{jt}$  in constraints (7a) and (7b) and then augmenting it with the constraint  $\tilde{D}_{jt} \in U^B$  for all  $j \in N$  and  $t = 1, 2, \dots, T$ .

The augmented constraint for (7a) is

$$\min_{\tilde{D}_{jt} \in U^B} \left\{ \sum_{i=1}^N \sum_{j=1}^N \sum_{t=1}^T \tilde{D}_{jt} (\eta - d_{ij}) X_{ijt} \right\} - \sum_{i=1}^N \sum_{t=1}^T c_{it} Z_{it}$$

$$- \sum_{i=1}^N (C_{i0} Z_{i0} + K_i I_i) \geq \tau. \quad (8)$$

For a given candidate solution  $X \geq 0$ , the optimal solution to the optimization problem on the LHS of (8) is  $\tilde{D}_{jt}(1 - \varepsilon_t)$  if  $(\eta - d_{ij}) \geq 0$ , and  $\tilde{D}_{jt}(1 + \varepsilon_t)$  if  $(\eta - d_{ij}) < 0$ . However, if  $(\eta - d_{ij}) < 0$ , then  $X_{ijt} = 0$  in the optimal solution to (P') with (7a) replaced by (8). Thus, we can ignore the second case and find the

robust counterpart of constraint (7a) to be

$$\sum_{i=1}^N \sum_{j=1}^N \sum_{t=1}^T (\eta - d_{ij}) \bar{D}_{jt} (1 - \varepsilon_t) X_{ijt} - \sum_{i=1}^N \sum_{t=1}^T c_{it} Z_{it} - \sum_{i=1}^N (C_{i0} Z_{i0} + K_i I_i) \geq \tau. \quad (9)$$

Next, consider constraint (7b) for a given  $i$  and  $t$ . Noting that the inequality is now “ $\leq$ ,” we augment the constraint by solving for the maximum of the LHS over the uncertainty set, leading to

$$\max_{\tilde{D}_{jt} \in U^B} \left\{ \sum_{j=1}^N \tilde{D}_{jt} X_{ijt} \right\} \leq Z_{it}.$$

Noting  $\tilde{D}_{jt} \geq 0$ , the constraint implies

$$\sum_{j=1}^N \bar{D}_{jt} (1 + \varepsilon_t) X_{ijt} \leq Z_{it} \quad \text{for all } i, t. \quad (10)$$

Therefore, the robust counterpart of problem (P') with uncertainty set  $U^B$  is given by

$$(RC^{Box}) \max_{\mathbf{x}, \mathbf{Z}, \mathbf{I}, \mathbf{Z}_0, \tau} \tau$$

subject to constraints (7c)–(7f), (9), and (10).

We let  $M = \max_t \left\{ \sum_{j=1}^N \bar{D}_{jt} (1 + \varepsilon_t) \right\}$  be an appropriately large number in (7d). Problem (RC<sup>Box</sup>) is an MILP that can be solved effectively using off-the-shelf software.

### 3.3. The Robust Problem—Ellipsoidal Uncertainty

Consider the case where the uncertain data are contained in an ellipsoid. As discussed in section 2.1 such an ellipsoidal uncertainty set can also be characterized by a mean/variance approach. That is, one can replace the uncertain parameters by their means plus or minus a given number of standard deviations expressing the decision maker's preferences. We define the uncertainty set and then reformulate the robust counterpart of each constraint with uncertain elements in (P').

As before, let  $\tilde{D}_{jt}$  be the uncertain demand at node  $j$  in period  $t$ . For notational ease, let  $\tilde{D}_t = \{\tilde{D}_{jt}\}_{j=1 \dots N} \in \mathbb{R}^N$  and  $\tilde{D} = \{\tilde{D}_t\}_{t=1 \dots T} \in \mathbb{R}^{NT}$  be vectors of the demand. Let  $\bar{D}_{jt}$  be the nominal value and let  $\Sigma$  of size  $(NT \times NT)$  denote the covariance matrix of  $\tilde{D}$ . (As we assume independence of demands over time and location for our test below,  $\Sigma$  is diagonal with entries  $\sigma_{jt}^2$ .)

Recall in our discussion of the box uncertainty case, the multi-dimensional unit box was given by the absolute, normalized deviations, i.e.,

$$U^B = \left\{ \tilde{D} \in \mathbb{R}^{NT} \mid -1 \leq \frac{\tilde{D}_{jt} - \bar{D}_{jt}}{\varepsilon_t \bar{D}_{jt}} \leq 1 \quad \text{for all } j, t \right\}.$$

To construct the ellipsoid set, we consider the total-normalized-squared deviations. Therefore, for the constraint (7a), we define the following uncertainty set:

$$U_1^E = \left\{ \tilde{D} \in \mathbb{R}^{NT} \mid \sum_{j=1}^N \sum_{t=1}^T \left[ \frac{\tilde{D}_{jt} - \bar{D}_{jt}}{\varepsilon_t \bar{D}_{jt}} \right]^2 \leq \Omega_1^2 \right\}, \quad (11)$$

where  $\Omega_1$  is the safety parameter applied to constraint (7a) in (P'). Observe that if we set  $\Omega_1 = 1$  in (11),  $U_1^E$  is the largest ellipsoid contained in  $U^B$ , and if we set  $\Omega_1 = \sqrt{NT}$ ,  $U_1^E$  is the smallest ellipsoid containing  $U^B$ .

Letting  $\sigma_{jt} = \varepsilon_t \bar{D}_{jt}$  and defining  $\Sigma$  to be the  $NT \times NT$  diagonal matrix with non-zero entries  $\sigma_{jt}^2$ , then

$$U_1^E = \left\{ \tilde{D} \in \mathbb{R}^{NT} \mid (\tilde{D} - \bar{D})^T \Sigma^{-1} (\tilde{D} - \bar{D}) \leq \Omega_1^2 \right\}.$$

We now consider the augmented constraint

$$\min_{\tilde{D} \in U_1^E} \left\{ \sum_{i=1}^N \sum_{j=1}^N \sum_{t=1}^T (\eta - d_{ij}) \tilde{D}_{jt} X_{ijt} - \sum_{i=1}^N \sum_{t=1}^T c_{it} Z_{it} - \sum_{i=1}^N (C_{i0} Z_{i0} + K_i I_i) \geq \tau. \right.$$

Let  $V_{jt} = \sum_{i=1}^N (\eta - d_{ij}) X_{ijt}$ . Similar to (6) we arrive at the robust counterpart constraint

$$\sum_{j=1}^N \sum_{t=1}^T \bar{D}_{jt} V_{jt} - \Omega_1 \sqrt{\sum_{j=1}^N \sum_{t=1}^T \sigma_{jt}^2 V_{jt}^2} - \sum_{i=1}^N \sum_{t=1}^T c_{it} Z_{it} - \sum_{i=1}^N (C_{i0} Z_{i0} + K_i I_i) \geq \tau. \quad (12)$$

Next, we linearize (12) by considering a decision variable,  $W$ , with  $W = \sqrt{\sum_{j=1}^N \sum_{t=1}^T \sigma_{jt}^2 V_{jt}^2}$ . This facilitates solution by standard conic optimization (subject to a transformation detailed below). Then the set of constraints

$$\sum_{j=1}^N \sum_{t=1}^T \bar{D}_{jt} V_{jt} - \Omega_1 W - \sum_{i=1}^N \sum_{t=1}^T c_{it} Z_{it} - \sum_{i=1}^N (C_{i0} Z_{i0} + K_i I_i) \geq \tau, \quad (13a)$$

$$V_{jt} = \sum_{i=1}^N (\eta - d_{ij}) X_{ijt}, \quad (13b)$$

$$W = \sqrt{\sum_{j=1}^N \sum_{t=1}^T \sigma_{jt}^2 V_{jt}^2} \quad (13c)$$

form the robust counterpart constraint for constraint (7a) in case of uncertainty set  $U_1^E$ .

Under the assumption that  $U_1^E \subseteq \mathbb{R}^{NT+}$  (the positive orthant), and given the objective of maximizing  $\tau$ , we

can relax (13b) to

$$V_{jt} \leq \sum_{i=1}^N (\eta - d_{ij}) X_{ijt}. \quad (13b')$$

Similarly we can relax (13c) to

$$W \geq \sqrt{\sum_{j=1}^N \sum_{t=1}^T \sigma_{jt}^2 V_{jt}^2}. \quad (13c')$$

For each of the constraints (7b), we can also formulate robust counterpart constraints under ellipsoidal uncertainty. Let

$$U_{2t}^E = \left\{ \tilde{D} \in \mathbb{R}^{NT} \mid \sum_{j=1}^N \left[ \frac{\tilde{D}_{jt} - \bar{D}_{jt}}{\varepsilon_t \bar{D}_{jt}} \right]^2 \leq \Omega_{2t}^2 \right\}.$$

Here  $\Omega_{2t}$  is the safety parameter associated with the uncertainty in period  $t$ . Again we let  $\sigma_{jt} = \varepsilon_t \bar{D}_{jt}$  for each  $t$ , and let  $C_t$  be the  $N \times N$  diagonal matrix with non-zero entries  $\sigma_{jt}^2$ . The augmented constraint for (7b) for a given  $i$  and  $t$  is

$$\max_{\tilde{D}_{jt} \in U_{2t}^E} \left[ \sum_{j=1}^N \tilde{D}_{jt} X_{ijt} \right] \leq Z_{it} \quad \text{for all } i, t.$$

We use a similar approach as in (6) to find the robust counterpart constraint:

$$\sum_{j=1}^N \bar{D}_{jt} X_{ijt} + \Omega_{2t} \sqrt{\sum_{j=1}^N \sigma_{jt}^2 X_{ijt}^2} \leq Z_{it} \quad \text{for all } i, t.$$

Note that we now add in  $\Omega_{2t}$  multiples of the uncertainty to reflect the changed direction of the inequality ( $\leq$ ) in (7b). Similar to above, we let  $Q_{it} = \sqrt{\sum_{j=1}^N \sigma_{jt}^2 X_{ijt}^2}$ . Noting that  $c_{it} > 0$  in the objective function, so that  $Z_{it}$  is to be minimized, we can express  $Q_{it}$  as an inequality, so that the robust counterpart constraints for constraints (7b) are

$$\sum_{j=1}^N \bar{D}_{jt} X_{ijt} + \Omega_{2t} Q_{it} \leq Z_{it} \quad \text{for all } i, t, \quad (14)$$

$$Q_{it} \geq \sqrt{\sum_{j=1}^N \sigma_{jt}^2 X_{ijt}^2} \quad \text{for all } i, t. \quad (15)$$

Observe that (13c') and (15) express conic quadratic constraints. This follows because the variables in their respective vectors  $[\{V_{jt}\}_j, W]$  and  $[\{X_{ijt}\}_j, Q_{it}]$  are elements of Lorentz cones, i.e.,  $[\{V_{jt}\}_j, W] \in L^{NT+1}$  and  $[\{X_{ijt}\}_j, Q_{it}] \in L_{it}^{N+1}$ , where  $L^y$  is the Lorentz cone of dimension  $y$ . Note, however, that variable  $X_{ijt}$  appears in  $L_{it}^{N+1}$  and, because of (13b), also in  $L^{NT+1}$ . This proves problematic in solving the resulting conic

optimization problem. Therefore, we define  $v_{jt} \geq \sigma_{jt} V_{jt}$  for all  $j$  and  $t$ , and  $\chi_{ijt} \geq \sigma_{jt} X_{ijt}$  for all  $i, j$ , and  $t$ . We then replace (13c') by  $W \geq \sqrt{\sum_{j=1}^N \sum_{t=1}^T v_{jt}^2}$  and (15) by  $Q_{it} \geq \sqrt{\sum_{j=1}^N \chi_{ijt}^2}$ . The complete robust counterpart of (P') is

$$(RC^{Ell}) \quad \max \tau$$

$$\sum_{j=1}^N \sum_{t=1}^T \bar{D}_{jt} V_{jt} - \Omega_1 W - \sum_{i=1}^N \sum_{t=1}^T c_{it} Z_{it} - \sum_{i=1}^N (C_{i0} Z_{i0} + K_i I_i) \geq \tau, \quad (16a)$$

$$\sum_{j=1}^N \bar{D}_{jt} X_{ijt} + \Omega_{2t} Q_{it} \leq Z_{it} \quad \text{for all } i, t, \quad (16b)$$

$$V_{jt} \leq \sum_{i=1}^N (\eta - d_{ij}) X_{ijt} \quad \text{for all } j, t, \quad (16c)$$

$$v_{jt} \geq \sigma_{jt} V_{jt} \quad \text{for all } j, t, \quad (16d)$$

$$\chi_{ijt} \geq \sigma_{jt} X_{ijt} \quad \text{for all } i, j, t, \quad (16e)$$

$$W \geq \sqrt{\sum_{j=1}^N \sum_{t=1}^T v_{jt}^2} \quad (16f)$$

$$Q_{it} \geq \sqrt{\sum_{j=1}^N \chi_{ijt}^2} \quad \text{for all } i, t, \quad (16g)$$

and constraints (7c)–(7f).

Here we let  $M$  in (7d) be sufficiently large, e.g.,  $M = N \max_{j,t} [D_{jt} \in U^E]$ .

In this formulation when we consider the demand revenues in (16a), the uncertainty expresses the possible downside of lower than expected revenues. In doing so, we limit the overall uncertainty in all periods under consideration through the summation in (16f). In comparison, when considering the production required in (16b), the uncertainty expresses the alternate downside of high requirements for capacity. In this case we limit the demand required to be served in each period separately through (16g). The effect of these is similar to the box uncertainty case—we require sufficient capacity in each period to serve the demand expected, but assume lower revenues in finding our profits than would be implied by these high demands.

Problem  $(RC^{Ell})$  is a conic quadratic program. There are several software programs that efficiently solve conic quadratic programs, e.g., the MOSEK

optimization tools—<http://www.mosek.com>. We note that problem  $(RC^{Box})$  has  $N^2T+N(T+2)+1$  variables, of which  $N$  are binary, and  $N(3T+1)+1$  constraints. Problem  $(RC^{Ell})$  has  $2N^2T+N(4T+2)+2$  variables, of which  $N$  are binary, and  $N^2T+N(6T+1)+2$  constraints, of which  $NT+1$  are conic quadratic constraints. This increase in the number of variables and constraints heavily impacts the size of the mixed-integer problems that can be solved. We turn next to solving examples of these programs to evaluate how the alternate robust models affect the objective value and the solutions.

### 4. Numerical Study

In this section, we conduct numerical experiments to understand how the solutions provided by the nominal problem and the RO models, using either box or ellipsoidal uncertainty sets, differ. We consider how they differ in the number of facilities opened, their capacities, and the network topology determined by the solution. In addition, we study the expected profit under each model as given by a simulation of demand sample paths. Before presentation of the results, we describe the test environment in detail.

#### 4.1. Test Environment

We randomly generate  $N$  nodes on a unit square, representing the demand points and potential facility locations. The nominal demand at each demand point in each period,  $\bar{D}_{jt}$ , is independent of all others and is assumed constant over the  $T$  periods. We let  $N = 15$  and  $T = 20$ . The nominal demand at each location is drawn from the uniform distribution, Uniform [17,500, 22,500].

We use the revenue and cost parameters summarized in Table 1 for the base case. The revenue is set at  $\eta = 1$ . The delivery cost,  $d_{ij}$ , equals the Euclidean distance between nodes  $i$  and  $j$ . The expected distance between any two nodes is approximately 0.5 and the expected maximum distance is approximately 1 (see, e.g., Larson and Odoni 2007). We let  $C_{i0} = 0.1$ ,  $c_{it} = 0.1$  ( $\forall i, t$ ) and  $K = 50,000$ . Using these values, the expected margin per unit per period is approximately 0.3 if one location is opened and 0.9 if all  $N$  locations are opened. The cost of opening a facility to serve the

average demand is approximately between 3 and 10 periods of demand.

The initial uncertainty sets are generated as follows. In the base case we let  $\gamma = 0.15$  be the initial (first period) uncertainty in the demand, and let  $\varepsilon_t = \gamma + (1 - \gamma)\varepsilon_{t-1}$  with  $\varepsilon_0 = 0$ . For the box uncertainty set we assume  $U_{jt}^B = \bar{D}_{jt}[1 \pm \varepsilon_t]$ . Because  $\varepsilon_t$  is a concave, increasing function in  $t$ ,  $U_{jt}^B$  is increasing in  $t$ . In Figure 1, the solid lines depict the boundaries of  $U_{jt}^B$ . The marked points depict a sample path that might occur (described below). For the ellipsoidal uncertainty set we let the safety parameters  $\Omega_1$  and  $\Omega_{2t}$  for  $t = 1, \dots, T$  equal 1. This creates the largest ellipsoid contained in the box  $U^B$  (the inscribed ellipsoid).

We compare the solutions of the nominal problem to those of the box and ellipsoidal uncertainty problems. We do so by first considering the topology of the solutions (number of facilities established, number of edges utilized, etc.). Second, we compare the profit of the solutions found.

#### 4.2. Comparing the Topology of the Solutions

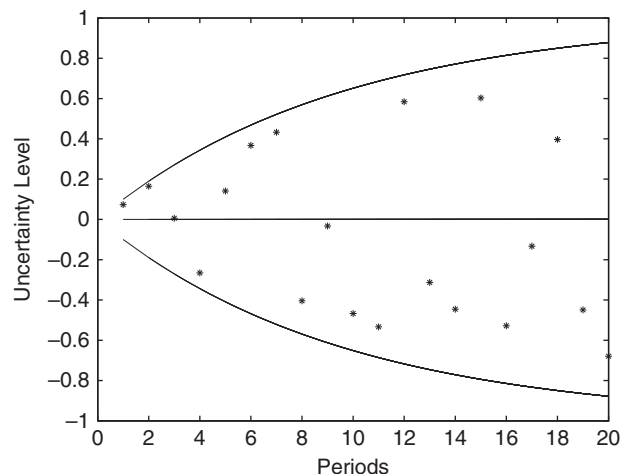
We compare the number of facilities opened, the capacity of the facilities, and the connectivity of the graph for the three models. To do so, we generate 250 random graphs by choosing node locations uniformly over the unit square. For each graph we solve the nominal, box uncertainty, and ellipsoidal uncertainty models. For each graph we generate 10 sample paths for the demand for a total of 2500 sample paths. The sample paths are generated by sampling over the box uncertainty set for each time period  $t$  (the area between the curves in Figure 1). Sampling is conducted using either a bell-shaped beta(2,2) distribution, uniform distribution, or U-shaped beta(1/2, 1/2) distribution.

In Figure 2, we present typical solutions of the models. In the figures, a star denotes an open facility

Table 1 Parameter Values in Base Case

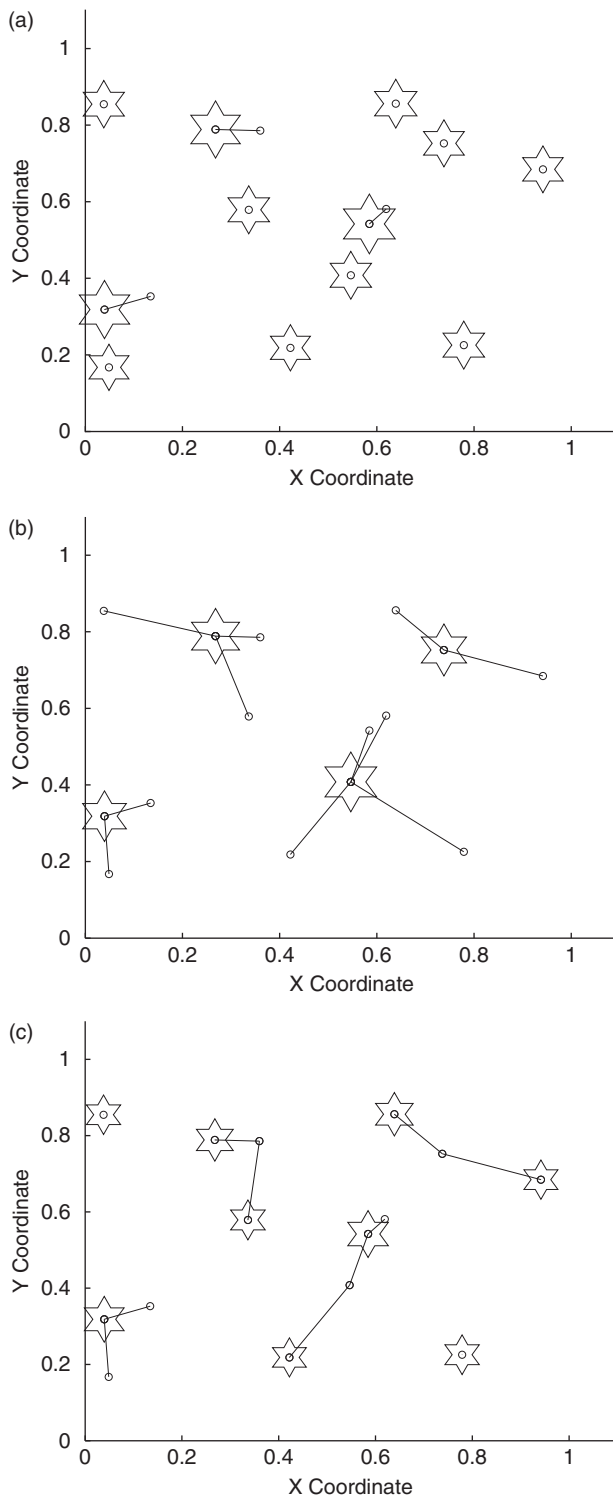
$\eta$	1
$N$	15
$T$	20
$c_{it}$	0.1
$C_{i0}$	0.1
$K$	50,000
$D_j$	$\sim U[17500, 22500]$

Figure 1 Box Uncertainty Set given by  $\gamma = 15\%$  and an Instance of a Sample Path in the Box





**Figure 2** Network Representations of the Solution to an Instance of (a) the Nominal Model, (b) Box Uncertainty Case, (c) Ellipsoidal Uncertainty Case



(co-located with a demand node), a circle represents a demand node, and a line indicates that in at least one period the facility is used to satisfy demand at the node. The size of the stars is proportional to the ca-

capacity of the facility. In Table 2, we present the mean number of facilities, mean capacity, and mean number of connections per open facility. The last is defined as

$$\frac{\sum_j \sum_i I\{\sum_t X_{ijt} > 0\}}{\sum_j I_j},$$

where  $I\{\}$  is the indicator function. We also present the percentage of the demand covered by each of the three solutions. Results are given for the case of  $\eta = 1, 3,$  and  $6$ .

We make several observations. First, compared with the nominal model, both robust models with the box and ellipsoidal uncertainty sets open fewer facilities with larger capacities, with the box model opening far fewer. Because the box uncertainty set contains the ellipsoid set, the degree of uncertainty is greater in the former. Therefore, it must respond simultaneously to both higher potential demand (resulting in larger capacity) and lower potential revenue (resulting in fewer facilities, lowering the fixed cost). As such, we find the strategic cost (the initial cost of establishing the facilities and their capacities) for the box uncertainty model to be 40% of the cost in the nominal model. In comparison, the ellipsoidal model strategic cost is 90% of that of the nominal model. Second, the ellipsoidal uncertainty model establishes more edges per open facility than the nominal model. Importantly, it allows more than one facility to serve the same demand node, expressing flexibility in the service to increase robustness to demand uncertainty. This stands in contrast to both the nominal and box uncertainty models that almost always serve each demand node from a unique facility. Finally, we observe that on the sample paths, the two robust models cover 100% of the demand compared with 95% for the nominal model. This indicates the robust models are performing as expected in terms of their ability to address uncertainty.

### 4.3. Comparing the Profit of the Solutions

In this subsection, we compare the profit of the robust models to that of the nominal model as we change various parameters of the model. For each of the models, we solve the associated mathematical programs for  $I_i, Z_{i0}$  (the strategic decisions of location and capacity), and for  $Z_{it},$  and  $X_{ijt}$  (the operational decisions of production and allocation). In practice, however, while the firm cannot change the strategic decisions under our assumptions, it will, upon observation of demand, determine the active operational decisions, say  $\hat{Z}_{it}$  and  $\hat{X}_{ijt}$ . That is, given a solution,  $I_i^*$  and  $Z_{i0}^*$ , for each period  $t$ , we assume the firm observes realized demand  $\hat{D}_{it}$  and solves the following

**Table 2 Comparing the Solution Topology of the Three Models**

	Model	Mean # of open facilities	Mean capacity	Mean # edges/facility	% of demand covered		
					Uniform	Bell-shape	U-shape
$\eta = 1$	Nominal	10.74	28,221	0.42	96.99	94.99	94.58
	Box	3.48	140,350	3.91	100	100	100
	Inscribed ellipsoid	8.68	59,737	14.36	100	100	100
$\eta = 3$	Nominal	10.74	28,221	0.42	96.74	95.08	94.62
	Box	4.24	136,425	3.25	100	100	100
	Inscribed ellipsoid	8.66	59,822	14.34	100	100	100
$\eta = 6$	Nominal	10.74	28,221	0.42	96.74	95.08	94.62
	Box	4.44	134,485	2.94	100	100	100
	Inscribed ellipsoid	8.66	59,759	14	100	100	100

operational problem for the operational profit,  $\Pi_t$ :

$$(P^{opr}) \quad \Pi_t = \max_{\mathbf{x}, \mathbf{z}} \sum_{i=1}^N \sum_{j=1}^N (\eta - d_{ij}) \hat{D}_{jt} \hat{X}_{ijt} - \sum_{i=1}^N c_{it} \hat{Z}_{it},$$

s.t.

$$\sum_{j=1}^N \hat{D}_{jt} \hat{X}_{ijt} \leq \hat{Z}_{it} \quad \text{for all } i,$$

$$\sum_{i=1}^N \hat{X}_{ijt} \leq 1 \quad \text{for all } j, t,$$

$$\hat{Z}_{it} \leq Z_{i0}^* \quad \text{for all } i, t,$$

$$\hat{X}_{ijt} \geq 0 \quad \text{for all } i, j, t.$$

We compare the performance of the strategic solutions with respect to sample paths  $\{\hat{D}_{it}\}_i$  for the uncertain demand. To do so, we generate sample paths for the demand at the demand nodes and then optimize  $(P^{opr})$  for the operational decisions  $\hat{Z}_{it}$  and  $\hat{X}_{ijt}$  for each of the models using the bell-shaped, uniform, and U-shaped distributions. This allows us to estimate the operating costs associated with the strategic solution. Throughout these tests, we consider a single-network topology (the location of facilities). The nominal demand for the base case in the tests is 20,000 units per period for each node (as opposed to sampled from the range [17,500, 22,500]). All other parameters are as given in Table 1. For each of the distributions, we generate 1000 sample paths.

First, we consider how the base case box and ellipsoidal uncertainty robust models perform as the nominal demand grows for the various demand distributions. Second, we consider the trade-off between robustness of the solution and performance vs. the nominal solution. Third, we vary the horizon length. Finally, we vary the fixed charge.

**4.3.1. Profit vs. Demand Growth and Uncertainty.** In Table 3, we present the percentage increase (or decrease) in the profit for the base case box and ellipsoidal uncertainty models relative to the nominal model. We do so for the case with constant nominal demand and two cases where the nominal demand increases: demand growing by 500 units per period (a 2.5% growth rate) and demand growing at 1000 units per period (a 5% growth rate). We also present the standard deviation of the increase in profits. In addition, for each sample path, we solve for the optimal solution of the nominal problem,  $(P)$ , with the demand determined by the realized sample path. Doing so provides an upper bound on the possible profit achieved. We present the percentage of the difference between the upper bound and the nominal solution given by the solution value (UB). That is,

**Table 3 Percentage Increase in Profit over the Nominal Model, Associated Standard Deviations and Implied Percentage of Upper Bound Achieved for the Base Case Box and Inscribed Ellipsoidal Uncertainty Models**

Case	Box uncertainty			Inscribed ellipsoidal uncertainty		
	Bell	Uniform	U-shape	Bell	Uniform	U-shape
Base case						
Mean %	-0.48	1.65	3.77	6.12	8.28	10.34
Standard deviation	0.98	1.20	1.61	1.16	1.43	1.89
% of upper bound	<0	17.29	34.69	93.03	95.59	96.47
Growth +500×t						
Mean %	-5.18	-3.71	-2.16	2.15	3.72	5.49
Standard deviation	0.67	0.78	0.95	0.80	1.03	1.35
% of upper bound	<0	<0	<0	66.86	78.12	85.15
Growth +1000×t						
Mean %	-7.41	-6.44	-5.39	1.43	2.85	4.42
Standard deviation	0.41	0.51	0.62	0.74	1.00	1.26
% of upper bound	<0	<0	<0	54.19	72.03	81.33

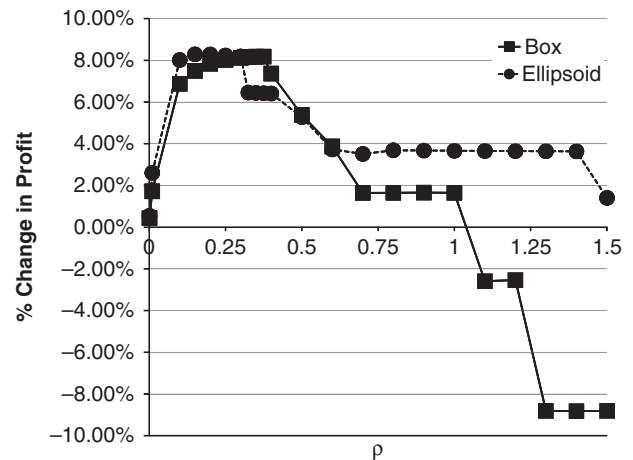
(solution value – nominal)/(upper bound – nominal). For example, in the box uncertainty case when demand is uniformly distributed over the uncertainty set, we observe a 1.65% increase in average profit, which corresponds to achieving 17.29% of the possible increase in the profit. Alternately, if the demand was drawn from a bell-shaped distribution, there is a decrease of 0.48% in the average profit and thus does not correspond to any percentage increase.

We observe that the (inscribed) ellipsoidal uncertainty case provides significantly better solutions to the operational problem faced than that of the nominal or box uncertainty cases. This result is not surprising as it emphasizes the difference in the nature of the solutions between the two robust models. The box is defined to immunize the solution against any possible outcome within the assumed uncertainty set, whereas the inscribed ellipsoid does not immunize the solution against realizations in the corners of the box. For the demand distributions, we test under the assumption of independent sampling; the inscribed ellipsoid immunizes the solution against realizations more likely to happen. Thus, the box model adopts a very conservative approach, establishing centralized facilities that can be deployed to meet any demand. However, the ellipsoidal model establishes multiple, flexible facilities, close to those given by the nominal solution, but sized to address potential uncertainty. From a sample path standpoint, such a model appears to have better performance. On average, the ellipsoidal model captures 80.3% of the benefit possible as given by the upper bound.

We observe that the best performance for each case (both in mean profit and percentage of upper bound) is when expected demand is constant over time. When there is a trend in the growth over the period (either at 500 or 1000 per period), the benefit of both RO models diminishes. In the box uncertainty model, in each of the demand growth cases, the performance is worse than the nominal model, though that of the ellipsoidal model is still positive. Further, as would be expected, both RO models perform better as the likelihood of more extreme demand increases, i.e., as the sampling distribution changes from a bell-shaped to uniform to U-shaped distribution.

**4.3.2. Robustness–Performance Trade-Off.** As we observe above, the model with the ellipsoidal uncertainty set finds a solution that performs better than the model with the box uncertainty set that circumscribes it, even when the actual demand is drawn from a distribution with support over the entire box. Again, this is not surprising as the solution for the ellipsoid case does not immunize the solution against all of the uncertainty contained in the box. Next, we investigate how by assuming a smaller uncertainty set

**Figure 3** Percentage Increase (Relative to the Nominal Model) in Profit for Models with Box and Ellipsoid Uncertainty Sets, as the Value of  $\rho$  Changes



(either for the box or the ellipsoid models) one can find a better performing solution. That is, we will immunize the solution against a smaller uncertainty set; however, in testing we continue to sample from the base case box uncertainty set.

Consider first the box model. In the base case, we assumed the box was defined by an initial uncertainty  $\gamma = 0.15$ , and then defined  $\varepsilon_t$  for  $t = 1, \dots, i$ . Suppose that we vary the value of  $\gamma$ , and therefore  $\varepsilon_t$ , to define a new box uncertainty set. Let  $\rho$  be the ratio of the volume of the box uncertainty set to the volume of the base case box, i.e.,  $\rho = 100\%$  in the base case. For each value of  $\rho$ , we can also generate an ellipsoid that is equal in volume to that specific box. To do so, as above, we let  $\sigma_{jt} = \varepsilon_t \bar{D}_{jt}$  as  $\varepsilon_t$  varies. Then, rather than using the inscribed ellipsoid, by properly scaling the safety parameters  $\Omega_1$  and  $\Omega_{2t}$  we equate the multi-dimensional volumes of the box and ellipsoid. For the base case with  $N = 15$  nodes and  $T = 20$  periods, we find  $\Omega_1 = 8.4784$  and  $\Omega_{2t} = 2.1327$  (see chapter 2 of Ben-Tal et al. 2009 for details on such scaling). Observe this ellipsoid is larger than the inscribed ellipsoid defined by  $\Omega_1 = \Omega_{2t} = 1$ .

In Figure 3, we plot the percentage increase (relative to the nominal model) in the robust solution profit as we vary  $\rho$  from 0.1% to 150% assuming the demand is sampled uniformly over the base case box uncertainty set. We observe that the box model's maximum improvement is approximately 8.18% at approximately  $\rho = 0.35$ , and the ellipsoid model's is approximately 8.28% at  $\rho = 0.15$ . The "best" ellipsoid captures infrequent, but potentially important, demand observations toward the edges of its volume. Intuitively, for the box to do the same it would need to be larger; thus, the larger volume of the "best" box. For smaller  $\rho$ , the robust models do not immunize the solution against much uncertainty and therefore perform close to the nominal model. For

**Table 4** Percentage Increase in Profits over the Nominal Model and Implied Percentage of Upper Bound Achieved for the  $\rho = 0.15$  and  $\rho = 0.35$  Cases

$\rho$	Case	Equal volume					
		Box uncertainty			Ellipsoidal uncertainty		
		Bell	Uniform	U-shape	Bell	Uniform	U-shape
0.15	Base case						
	Mean %	5.88	7.50	8.81	6.19	8.29	10.22
	% of upper bound	89.53	86.23	82.12	94.17	95.78	95.55
	Growth $+500 \times t$						
	Mean %	2.59	3.81	5.07	2.67	4.21	5.92
	% of upper bound	82.22	82.03	78.37	84.16	88.94	91.86
0.35	Base case						
	Mean %	6.12	8.18	10.05	4.28	6.44	8.54
	% of upper bound	93.00	94.48	93.26	65.31	73.47	79.32
	Growth $+500 \times t$						
	Mean %	2.19	3.73	5.40	2.05	3.63	5.41
	% of upper bound	67.58	78.17	83.51	62.78	75.67	83.92
0.35	Base case						
	Mean %	1.50	2.88	4.35	0.80	2.22	3.79
	% of upper bound	57.34	72.81	80.26	33.26	55.89	68.36

Cases with shading indicate comparable results.

larger values of  $\rho$ , the performance decreases as the models increasingly immunize the solution against uncertainties that are unlikely to occur. In particular, we observe an expected profit loss for the box case for  $\rho > 100\%$ , i.e., for boxes larger than the base case from which demand is sampled, there is no longer a trade-off; greater robustness only results in lowered profits. The results indicate that for small  $\rho$  both models can balance robustness with profits.

In Table 4, we present the results for the box and ellipsoid models for  $\rho = 0.15$  and  $\rho = 0.35$ , corresponding roughly to the best performance for each model. We note the similarity of the ellipsoid model at  $\rho = 0.15$  to the box model at  $\rho = 0.35$  (shown in the shaded regions in the table). Also note the similarity to the ellipsoid inscribed in base case box, given in Table 3. That is, the inscribed ellipsoid performs very nearly as well as the best possible ellipsoid, at least for the test case. As above, the performance for each model increases as the demand spreads away from the mean (U-shaped vs. bell-shaped), and decreases as the rate of demand growth over time increases. We note that in all these cases the ellipsoid performs slightly better than the box model.

**4.3.3. Varying Horizon Length,  $T$ .** We investigate how the robust solutions compare with the nominal

**Table 5** Percentage Increase in Profit for Box and Ellipsoidal Uncertainty Sets when Varying  $T$

Number of periods	$\rho$							Base case
	0.1	0.15	0.2	0.25	0.3	0.35	0.4	
<b>Box</b>								
1	-0.09	-0.24	-0.46	-0.72	-1.00	-1.30	-1.60	-5.25
2	0.38	0.35	0.26	0.14	-0.01	-0.16	-0.31	-2.11
3	1.09	1.04	0.97	0.87	0.75	0.63	0.50	-0.29
4	1.35	1.53	1.57	1.55	1.49	1.40	1.31	0.20
5	1.94	2.19	2.27	2.26	2.21	2.14	2.06	0.01
10	4.20	4.69	4.98	4.66	4.66	4.65	4.64	3.21
15	5.46	6.18	6.55	6.33	5.85	5.90	5.20	2.02
20	6.88	7.49	7.83	8.01	8.11	8.17	7.35	1.65
Inscribed								
<b>Ellipsoid</b>								
1	0.00	-0.05	-0.11	-0.20	-0.30	-0.42	-0.56	-0.76
2	0.36	0.38	0.35	0.28	0.19	0.08	-0.03	-0.15
3	1.01	1.06	1.04	0.98	0.91	0.09	0.01	-0.04
4	0.72	1.53	1.59	1.58	1.53	1.45	1.37	1.35
5	1.49	2.29	2.32	2.28	2.20	2.11	2.01	2.03
10	4.77	5.27	4.85	4.80	4.73	4.68	4.61	4.72
15	6.68	7.05	6.64	6.03	5.98	5.20	5.15	6.59
20	8.02	8.28	8.27	8.22	8.17	6.42	6.40	8.26

Maximum profit increase for each  $T$  is shaded.

solution as the length of the horizon changes. Results are reported in Table 5. We present the box and ellipsoid models for  $\rho \in [0.1, 0.4]$ , the base case ( $\rho = 1.0$ ), and the inscribed ellipsoid for the base case. Increase in profit is reported vs. the nominal with uniform sampling over the base case box. We observe that for a given  $\rho$ , as the number of periods increases, the profit increases vs. the nominal for both the box and ellipsoidal cases (with the exception of  $\rho = 1$ ). This result is similar to that of Bertsimas and Thiele (2006). In the table, we emphasize the best performance for each horizon length for the two models. We observe that as the horizon increases, a larger box uncertainty set is required to achieve better performance, while an ellipsoid with  $\rho = 0.15$  is almost always the best value. Further, we observe that for the higher number of periods, the ellipsoidal model performs better than the box at their respective maxima. The flexibility provided by the ellipsoidal uncertainty model at  $\rho = 0.15$  allows the firm to find solutions that can better accommodate greater variation in the demand found over a longer horizon, whereas the box size must increase.

**4.3.4. Varying Fixed Cost,  $K$ .** We explore the effect of changing the fixed cost  $K$  in Table 6. We highlight the best performance for the various values of  $\rho$  for

**Table 6** Number of Open Facilities and Percentage Increase in Profit for Box and Ellipsoidal Uncertainty Sets Varying  $K$ 

Case	$K$	Number of facilities		% increase in profit over nominal							
		Nominal	Base case	$\rho$							
				0.1	0.15	0.2	0.25	0.3	0.35	0.4	1.0 (base case)
Box	0	15	15	6.38	7.19	7.62	7.86	8.00	8.08	8.11	8.10
	10,000	15	11	6.57	7.40	7.84	8.09	8.23	8.32	8.35	6.23
	20,000	15	8	6.76	7.62	8.07	8.33	8.27	8.35	8.38	4.49
	40,000	15	8	7.00	7.65	8.02	8.22	6.28	7.99	6.31	4.49
	80,000	9	2	6.62	7.12	5.79	5.86	5.10	6.30	5.04	-5.16
	150,000	9	2	6.08	6.27	5.08	5.06	5.02	3.73	2.80	-5.90
	200,000	4	1	5.18	4.12	2.94	2.94	2.90	4.04	-0.33	-18.52
	250,000	3	1	4.59	4.77	3.35	1.33	1.31	2.07	1.30	-16.32
(Inscribed)											
Ellipsoid	0	15	15	7.89	8.13	8.15	8.12	8.09	8.06	8.04	8.13
	10,000	15	15	8.11	8.37	8.39	7.35	8.32	7.89	7.35	8.37
	20,000	15	11	8.35	8.42	8.09	6.56	7.18	6.58	6.56	7.25
	40,000	15	11	8.26	8.51	8.09	6.21	6.27	7.96	6.21	7.25
	80,000	9	6	7.53	7.34	6.46	4.96	5.35	6.27	4.96	7.32
	150,000	9	7	6.21	6.33	5.09	4.91	4.99	6.17	4.91	7.49
	200,000	4	4	5.32	4.14	2.86	2.84	2.91	4.05	2.84	5.42
	250,000	3	3	4.59	4.79	3.36	1.28	3.32	2.08	1.28	4.08

Maximum profit increase for each  $T$  is shaded.

each model. For smaller fixed costs, the box model's results are better for larger  $\rho$ . As in Table 5, the ellipsoid model best performance is generally close to  $\rho = 0.15$ , which is also similar to the performance of the inscribed ellipsoid for  $\rho = 1$ . Observe that as the value of  $K$  increases, the number of facilities opened decreases. For comparison purposes, we present the number of facilities opened in the nominal, base case box ( $\rho = 1$ ), and inscribed ellipsoids in the table. As in section 4.2, the base case box opens fewer (and larger) facilities than the robust model with the inscribed ellipsoid. We find that only a limited amount of uncertainty should be addressed, roughly comparable to that of the inscribed ellipsoid. As before, we find the value of  $\rho = 0.15$  robust in doing so for the ellipsoid. For the box, by reducing the uncertainty set, more facilities are opened and the benefits of flexibility are gained.

## 5. Discussion

The objective of this paper is to investigate potential means by which RO techniques could be used to solve facility location problems. Previous work on RO has not considered such problems, particularly our approach to profit maximization. The results of the work reemphasize the need to evaluate the nature of the uncertainty allowed. We show that the alternate models of uncertainty lead to very different solution

network topologies with the box uncertainty case opening fewer, larger facilities. We then generate sample paths in order to test how the alternate solutions to the strategic problem would perform in practice. For each sample path, we solve the resulting operational problem of determining production and allocation. We find that the robust model with the box uncertainty performs poorly with respect to balancing robustness with profit. In contrast, the model with the inscribed ellipsoid provides small but significant improvements in the average profit. We show, however, similarly good results are obtained by varying the size of the box or ellipsoid used to immunize the solution against uncertainty. Importantly, we find that a single ellipsoid is able to provide good solutions over a wide range of parameter choices.

The problem we solve is difficult as it is a mixed-integer conic program. Within our problem formulation, demand appears in two places: first, as a multiplier for the revenues determining the objective function value, and second as a multiplier for the production capacity for each of the nodes. As such, these lead to alternate requirements of the solution; namely, the ability to address large demands, while at the same time observing low revenues. The resulting conic quadratic program for the ellipsoidal case is large by the standards of conic programming (up to 20 periods and 15 nodes). While such a problem may be

considered to be of only semi-realistic size, we emphasize that other methods, particularly stochastic programming, would have great difficulty in solving problems of such size.

We find the solutions given by the robust model with ellipsoidal uncertainty set address the need for robustness directly by considering alternate links between facilities and nodes to accommodate the uncertain demand. Future work may consider additional uncertainty sets such as a polyhedra, e.g., an intersection of box with an  $L_1$  norm. These may prove to provide solutions to larger problems while restricting the uncertainty away from worst-case scenarios.

### Acknowledgments

Support for this research was provided by grants from the Natural Science and Engineering Research Council of Canada. The authors would like to thank the journal's editorial team for their many helpful suggestions.

### References

- Adida, E., G. Perakis. 2006. A robust optimization approach to dynamic pricing and inventory control with no backorders. *Math. Program., Ser. B* 107(1–2): 97–129.
- Averbakh, I., O. Berman. 1997. Minimax regret  $p$ -center location on a network with demand uncertainty. *Locat. Sci.* 5(4): 247–254.
- Averbakh, I., O. Berman. 2000a. Algorithms for the robust 1-center problem on a tree. *Eur. J. Oper. Res.* 123(2): 292–302.
- Averbakh, I., O. Berman. 2000b. Minimax regret median location on a network under uncertainty. *INFORMS J. Comput.* 12(2): 104–110.
- Ben-Tal, A., L. El-Ghaoui, A. Nemirovsky. 2009. *Robust Optimization*. Princeton University Press, Princeton, NJ.
- Ben-Tal, A., B. Golany, A. Nemirovsky, J-P. Vial. 2005. Retailer-supplier flexible commitments contracts—a robust optimization approach. *Manuf. Serv. Oper. Manage.* 7: 248–271.
- Ben-Tal, A., T. Margalit, A. Nemirovsky. 2000. Robust modeling of multi-stage portfolio problems. Frenk, H., H. Roos, T. Terlaky, S. Zhang, eds, *High-Performance Optimization*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 303–328.
- Ben-Tal, A., A. Nemirovsky. 1998. Robust convex optimization. *Math. Oper. Res.* 23(4): 769–805.
- Ben-Tal, A., A. Nemirovsky. 1999. Robust solutions of uncertain linear programs. *Oper. Res. Lett.* 25(1): 1–13.
- Berman, O., J. Wang. 2006. The 1-median and 1-antimedial problems with continuous probabilistic demand weights. *INFOR* 44: 267–283.
- Bertsimas, D., D. Pachamanova, M. Sim. 2004. Robust linear optimization under general norms. *Oper. Res. Lett.* 32(6): 510–516.
- Bertsimas, D., M. Sim. 2006. Tractable approximations to robust conic optimization problems. *Math. Program., Ser. B* 107(1–2): 5–36.
- Bertsimas, D., A. Thiele. 2006. A robust optimization approach to inventory theory. *Oper. Res.* 54(1): 150–168.
- Carrizosa, E., S. Nickel. 2003. Robust facility location. *Math. Methods Oper. Res.* 58(2): 331–349.
- Chen, G., M. S. Daskin, Z.-J. M. Shen, S. Uryasev. 2006. The  $\alpha$ -reliable mean-excess regret model for stochastic facility location modeling. *Nav. Res. Logist.* 53(7): 617–626.
- Cohen, I., B. Golany, A. Shtub. 2007. The stochastic time-cost tradeoff problem: A robust optimization approach. *Networks.* 49(2): 175–188.
- Daskin, M. 1995. *Network and Discrete Location: Models Algorithms and Applications*. John Wiley and Sons, New York.
- Daskin, M. S., C. R. Coullard, Z. M. Shen. 2002. An inventory-location model: Formulation, solution algorithm and computational results. *Ann. Oper. Res.* 110(1–4): 83–106.
- Daskin, M. S., S. M. Hesse, C. S. ReVelle. 1997.  $\alpha$ -reliable  $p$ -minimax regret: A new model for strategic facility location modeling. *Locat. Sci.* 5(4): 227–246.
- Drezner, Z., H. Hamacher. 2002. *Facility Location: Applications and Theory*. Springer, Berlin.
- El-Ghaoui, L., H. Lebert. 1997. Robust solutions to least-square problems to uncertain data matrices. *SIAM J. Matrix Anal. Appl.* 18(4): 1035–1064.
- El-Ghaoui, L., F. Oustry, H. Lebert. 1998. Robust solutions to uncertain semidefinite programs. *SIAM J. Optim.* 9(1): 33–52.
- Gutierrez, G. J., P. Kouvelis. 1995. A robustness approach to international sourcing. *Ann. Oper. Res.* 59(1): 165–193.
- Kouvelis, P., A. A. Kurawarwala, G. J. Gutierrez. 1992. Algorithms for robust single and multiple period layout planning for manufacturing systems. *Eur. J. Oper. Res.* 63(2): 287–303.
- Larson, R. C., A. R. Odoni. 2007. *Urban Operations Research*. Dynamic Ideas, Belmont, MA.
- Louveaux, F. V. 1986. Discrete stochastic location models. *Ann. Oper. Res.* 6(2): 23–34.
- Mirchandani, P. B., R. L. Francis. 1990. *Discrete Location Theory*. John Wiley and Sons, New York.
- Mirchandani, P. B., A. Oudjit, R. T. Wong. 1985. Multidimensional extensions and a nested dual approach for the  $m$ -median problem. *Eur. J. Oper. Res.* 21(1): 121–137.
- Snyder, L., M. S. Daskin. 2006. Stochastic  $p$ -robust location problems. *IIE Trans.* 38(11): 971–985.
- Snyder, L. V. 2007. Facility location under uncertainty: A review. *IIE Trans.* 38(7): 537–554.
- Snyder, L. V., M. S. Daskin, C.-P. Teo. 2007. The stochastic location model with risk pooling. *Eur. J. Oper. Res.* 179(3): 1221–1238.
- Soyster, A. L. 1973. Convex programming with set-inclusive constraints and applications to inexact linear programming. *Oper. Res.* 21(5): 1154–1157.
- Verter, V., M. C. Dincer. 1992. An integrated evaluation of facility location capacity acquisition and technology selection for designing global manufacturing strategies. *Eur. J. Oper. Res.* 60: 1–18.
- Weaver, J. R., R. L. Church. 1983. Computational procedures for location problems on stochastic networks. *Trans. Sci.* 17(2): 168–180.