Two-Stage Stochastic Optimization for Capacitated Facility Location Under Demand Uncertainty

Sofiat O. Ajide
ORCID ID - 0009-0005-5804-3753
September 2026

Abstract

Facility location decisions are typically made before demand is fully known, yet most applied studies solve a single deterministic model using expected or nominal demand. This paper formulates and solves a capacitated facility location problem using a real academic benchmark instance, then extends it to a two-stage stochastic program in which facility-opening decisions are made under uncertainty and allocation decisions are made as recourse once demand is realized. Using the standard evaluation framework from stochastic programming, we compute the Value of the Stochastic Solution (VSS) and the Expected Value of Perfect Information (EVPI) for this instance, and test the robustness of both quantities across alternative demand-scenario structures and shortage-penalty assumptions. We find that the deterministic solution, while 3.2% cheaper up front, leaves an expected 1,230 units of demand unmet once evaluated honestly against demand uncertainty, and that accounting for uncertainty at the facility-opening stage is worth $82,933 in expected cost (VSS), consistently an order of magnitude larger than the value of having perfect foresight about which scenario will occur (EVPI), across every scenario structure and penalty assumption tested. This gap between VSS and EVPI indicates that the primary benefit of stochastic optimization here comes from committing to a more robust first-stage decision, not from reacting optimally after the fact, a finding consistent with related results reported in the humanitarian logistics literature.

Keywords: Facility Location; Stochastic Programming; Optimization Under Uncertainty; Value of the Stochastic Solution; Two-Stage Recourse; Sensitivity Analysis.

1. Introduction

Facility location decisions, such as where to open warehouses, depots, or service centers, are among the most consequential and least reversible decisions an organization makes. Once a facility is built or leased, its location and capacity are fixed for a planning horizon typically measured in years, while the demand it must serve fluctuates continuously in response to market conditions, seasonality, competitive activity, and, in some sectors, disruptions that are difficult to forecast even a single quarter in advance. Despite this, a large share of applied facility location studies solve a deterministic model using a single estimate of demand, either the historical average or a forecast, and treat the resulting solution as though it were reliable.

This simplification is not merely a modeling convenience; it can materially misstate the risk associated with a location decision. A deterministic model, by construction, cannot express the possibility that realized demand will exceed the capacity of the facilities selected, nor can it express the cost of the operational failures, stockouts, emergency shipments, or lost service, that follow when it does. A facility network that appears optimal, or even comfortably efficient, under a single demand forecast may in fact be fragile, performing acceptably only within a narrow band of outcomes around that forecast, and expensively or infeasibly outside it.

This paper examines the consequences of that simplification directly, using a real, established academic benchmark instance for the capacitated facility location problem. We first solve the deterministic version of the problem using nominal demand, then extend it to a two-stage stochastic program in which the decision of which facilities to open is made before demand is known, while the decision of how to allocate customers to open facilities is made afterward, once a demand scenario is realized. We then apply two standard metrics from the stochastic programming literature, the Value of the Stochastic Solution (VSS) and the Expected Value of Perfect Information (EVPI), to quantify precisely what is gained, and what is not gained, by explicitly modeling uncertainty rather than optimizing around a single demand estimate. Because both metrics depend on modeling choices, specifically how demand scenarios are structured and how the cost of unmet demand is penalized, we additionally test the sensitivity of our findings to both choices, to establish whether the qualitative conclusion holds beyond a single arbitrary parameterization.

The contribution of this paper is not a new algorithm or a new class of model; both the deterministic capacitated facility location problem and its two-stage stochastic extension are well established in the operations research literature, and are reviewed in Section 2. The contribution is a complete, reproducible computational case study on a real benchmark instance that (i) demonstrates concretely how much a deterministic facility location decision can understate real operating risk, (ii) separates that risk into the portion attributable to poor upfront planning (VSS) versus the portion attributable to irreducible uncertainty about the future (EVPI), and (iii) tests whether that separation is a robust structural feature of the instance or an artifact of a particular set of modeling assumptions.

The remainder of this paper is organized as follows. Section 2 reviews the relevant literature on deterministic and stochastic facility location, robust alternatives to stochastic programming, solution methods for large-scale stochastic programs, and applications of the VSS/EVPI framework in related domains. Section 3 presents the mathematical formulation of the deterministic and stochastic models and defines the evaluation metrics. Section 4 describes the benchmark instance and the scenario generation procedure. Section 5 reports the main computational results. Section 6 reports sensitivity analyses on scenario structure and shortage-penalty magnitude. Section 7 discusses the findings and their practical implications. Section 8 discusses limitations, and Section 9 concludes.

2. Literature Review

2.1 Deterministic Facility Location

The facility location problem, in its many variants, is one of the most extensively studied problem classes in operations research, encompassing the uncapacitated and capacitated facility location problems, the p-median problem, and the p-center problem, among others. The capacitated facility location problem addressed in this paper, in which each candidate facility has a fixed capacity and a fixed opening cost, and each customer's demand must be served, in whole or in part, by one or more open facilities, is NP-hard in general, and has motivated a substantial body of exact and heuristic solution methods. The benchmark instances used in this paper originate from Beasley's foundational work on large-scale capacitated warehouse location [1], which introduced a widely used family of test instances, generated to reflect realistic combinations of facility capacity, fixed cost, and customer allocation cost, and which remain in active use as a benchmark for new solution methods more than three decades later.

2.2 Facility Location Under Uncertainty

Facility location under uncertainty has a substantial dedicated literature, comprehensively surveyed by Snyder [2], who catalogs the range of modeling paradigms used to handle uncertain parameters in location models, including two-stage and multi-stage stochastic programming, robust optimization, and chance-constrained formulations, and who highlights that the choice among these paradigms typically depends on whether reliable probability distributions for the uncertain parameters are available, and on the decision-maker's attitude toward risk. The two-stage stochastic programming framework applied in this paper, including the Value of the Stochastic Solution and Expected Value of Perfect Information metrics used to evaluate it, follows the standard treatment in Birge and Louveaux's foundational text on stochastic programming [3], which formalizes these metrics as a general framework for quantifying the benefit of incorporating uncertainty into a decision model, independent of the specific application domain.

Santoso et al. [4] apply a closely related two-stage stochastic programming approach to supply chain network design, jointly determining facility locations and capacities under uncertain demand, supply, and cost parameters, and address the resulting large-scale stochastic program using a sampling-based solution method, since realistic supply chain instances may involve far more scenarios than can be enumerated directly. The present paper applies the same conceptual two-stage framework at a smaller, fully reproducible scale that permits direct enumeration of scenarios, with an explicit focus on decomposing the value of the stochastic solution into its VSS and EVPI components, and testing the robustness of that decomposition, on a single, real benchmark instance.

2.3 Robust Optimization as an Alternative Paradigm

Where reliable probability distributions for demand are not available, or where a decision-maker prefers to guard against worst-case outcomes rather than optimize an expectation, robust optimization offers an alternative to stochastic programming. Baron, Milner, and Naseraldin [5] apply robust optimization to a multi-period facility location problem under demand uncertainty modeled as bounded, rather than probabilistic, and show that robust solutions can provide small but consistent improvements over a deterministic nominal solution, while differing structurally, in the number and size of facilities opened, from both the nominal and stochastic solutions. Gülpınar, Pachamanova, and Çanakoğlu [6] similarly apply robust approximations to a capacitated facility location problem with a chance-constrained service-level requirement, comparing performance under both normally distributed and ambiguously distributed demand, and find that robust formulations offer favorable worst-case performance at a generally modest cost in expected performance relative to a stochastic or deterministic benchmark. These robust approaches represent a natural extension of the present analysis: since this paper finds that VSS substantially exceeds EVPI, meaning the value of a robust first-stage decision dominates the value of resolving uncertainty, a robust optimization formulation, which is explicitly designed to produce first-stage decisions that perform well across a range of outcomes without requiring a full probability distribution, may be a particularly well-suited alternative for this type of instance, and is noted as a direction for future work in Section 9.

2.4 Solution Methods for Large-Scale Stochastic Programs

The two-stage stochastic program in this paper is solved directly, as a single deterministic equivalent mixed-integer program over all scenarios jointly, an approach that is tractable because the number of scenarios considered is small. For applications with a very large or continuous scenario space, direct enumeration becomes computationally intractable, and approximate methods are required. The most widely used such method is the Sample Average Approximation (SAA) method, formalized by Kleywegt, Shapiro, and Homem-de-Mello [7], in which the true expected-cost objective is approximated using a finite Monte Carlo sample of scenarios, the resulting sampled problem is solved exactly, and the process is repeated across multiple samples to obtain both a candidate solution and a statistical estimate of its optimality gap. This method underlies the sampling-based solution approach used by Santoso et al. [4] for large-scale supply chain network design, and would be the natural extension of the present paper's methodology to instances with a richer, more realistic demand distribution than the small discrete scenario sets used here.

2.5 Applications of the VSS/EVPI Framework

The VSS/EVPI framework applied in this paper has been used to evaluate the value of stochastic optimization across a range of application domains beyond commercial facility location. Rawls and Turnquist [8] apply a two-stage stochastic programming model to pre-position emergency relief supplies before a natural disaster, jointly determining facility locations, inventory levels, and post-disaster distribution decisions under uncertainty in demand, transportation network damage, and pre-stocked supply damage, and explicitly compute the value of the stochastic solution to demonstrate that pre-positioning decisions informed by a range of disaster scenarios meaningfully outperform decisions based on a single expected scenario. Balcik and Beamon [9] more broadly review facility location models for humanitarian relief, noting that the central challenge in this domain, committing capacity in advance of an uncertain and potentially catastrophic event, is structurally identical to the challenge examined in this paper, differing primarily in the severity of the uncertainty and the human, rather than purely financial, cost of a shortage. The consistency between the qualitative finding reported by Rawls and Turnquist, that upfront stochastic planning materially outperforms deterministic planning under uncertainty, and the finding reported in this paper is notable given the substantial difference in application context, commercial warehouse location versus emergency relief pre-positioning, and suggests that the VSS-dominates-EVPI pattern observed here may reflect a more general property of capacitated location problems under demand uncertainty, rather than an artifact specific to this instance.

3. Problem Formulation

3.1 Deterministic Capacitated Facility Location Problem

Let I be the set of candidate facility locations and J be the set of customers. For each facility $i \in I$, let f_i denote its fixed opening cost and Q_i its capacity. For each customer $j \in J$, let d_j denote its demand, and let c_{ij} denote the cost of serving all of customer j's demand from facility i. Decision variable $y_i \in \{0,1\}$ indicates whether facility i is opened, and decision variable $x_{ij} \in [0,1]$ indicates the fraction of customer j's demand served by facility i. The deterministic model is:

$$\text{minimize} \quad \Sigma_{i \in I} f_i y_i \;+\; \Sigma_{i \in I}\Sigma_{j \in J} c_{ij} x_{ij} \tag{1}$$

subject to:

$$\Sigma_{i \in I} x_{ij} = 1 \quad \forall j \in J \tag{2}$$

$$\Sigma_{j \in J} d_j x_{ij} \leq Q_i y_i \quad \forall i \in I \tag{3}$$

$$x_{ij} \geq 0 \quad \forall i \in I, j \in J \quad ; \quad y_i \in \{0,1\} \quad \forall i \in I \tag{4}$$

Constraint (2) ensures every customer's demand is fully served. Constraint (3) ensures no facility serves more demand than its capacity, and links allocation to the facility being open. The objective (1) minimizes total fixed and allocation cost.

3.2 Two-Stage Stochastic Extension with Recourse

Demand is now treated as uncertain, realized as one of a finite set of scenarios S, with scenario s occurring with probability p_s and customer j's demand under scenario s denoted $d_j(s)$. Facility-opening decisions y_i are first-stage decisions, made once, before the scenario is known. Allocation decisions $x_{ij}(s)$ are second-stage recourse decisions, made separately for each scenario after demand is realized. To guarantee feasibility under every scenario, a shortage variable $u_j(s)$ is introduced, representing the fraction of customer j's demand left unmet under scenario s, penalized at rate π per unit of unmet demand. The two-stage stochastic program is:

$$\textit{minimize} \quad \Sigma_{i \in I} f_i y_i \;+\; \Sigma_{s \in S} p_s \, \Sigma_{i \in I}\Sigma_{j \in J} c_{ij}(s) x_{ij}(s) \tag{5}$$

subject to:

$$\Sigma_{i \in I} x_{ij}(s) + u_j(s) = 1 \quad \forall j \in J, s \in S \tag{6}$$

$$\Sigma_{j \in J} d_j(s) x_{ij}(s) \leq Q_i y_i \quad \forall i \in I, s \in S \tag{7}$$

plus the shortage penalty term $\pi \cdot \Sigma_s \, p_s \, \Sigma_j \, d_j(s) u_j(s)$ added to the objective (5), and non-negativity/binary restrictions on x, u, and y as in the deterministic model. This is the deterministic equivalent formulation of the two-stage stochastic program: since S is a small, finite, enumerable set of scenarios in this study, the entire stochastic program can be solved directly as a single mixed-integer program spanning all scenarios jointly, without resorting to decomposition or sampling methods.

3.3 Illustrative Small Example

Before presenting the evaluation metrics formally, it is useful to walk through a small hand-worked example illustrating why a deterministic solution can fail under uncertainty. Consider a toy instance with two candidate facilities and three customers. Facility A has capacity 100 and fixed cost 50; Facility B has capacity 60 and fixed cost 40. Customers 1, 2, and 3 have nominal demands of 40, 30, and 30 units respectively, for a total nominal demand of 100 units, exactly matching Facility A's capacity alone. Suppose allocation costs are such that serving all three customers from Facility A alone is cheapest under nominal demand: a deterministic model, minimizing cost under this single demand estimate, would open only Facility A, at a total cost of 50 plus allocation cost, since opening Facility B as well would add an unnecessary 40 in fixed cost with no nominal-case benefit.

Now suppose demand is uncertain: with probability 0.5, demand is exactly nominal (100 total); with probability 0.5, each customer's demand is 20% higher (120 total). Under the high-demand scenario, Facility A's capacity of 100 is insufficient for total demand of 120, so 20 units must either go unserved or be served by a second facility that is not open. The deterministic (EV) solution, having opened only Facility A, cannot avoid this shortfall in the high-demand scenario. A two-stage stochastic model, anticipating both scenarios when the facility-opening decision is made, would instead compare the cost of opening both facilities from the start, guaranteeing sufficient capacity in both scenarios, against the expected cost of opening only Facility A and incurring a shortage penalty in the high-demand scenario. If the shortage penalty is large enough, as it typically is when unmet demand represents lost sales, contractual penalties, or emergency fulfillment, the stochastic model will choose to open both facilities despite the additional fixed cost, precisely the qualitative pattern observed at full scale in Sections 5 and 6, where the RP solution opens one more facility than the EV solution specifically to avoid capacity shortfall in higher-demand scenarios.

This toy example also illustrates why VSS is necessarily non-negative: the EV solution's facility plan is, by definition, feasible and optimal for the nominal scenario, so its expected cost across all scenarios (EEV) can never be lower than the cost of a plan (RP) that was optimized with full knowledge of the scenario distribution from the start. The full-scale results in Section 5 exhibit exactly this relationship, with EEV (1,156,822.95) exceeding RP (1,073,889.86).

3.4 Computational Approach

All models in this paper, the deterministic problem, the base three-scenario stochastic problem, and the sensitivity variants with five scenarios and alternative shortage penalties, are solved as exact mixed-integer programs using the SCIP solver via Google OR-Tools' linear solver interface [10], with no decomposition, relaxation, or heuristic approximation. This is computationally feasible here because the number of scenarios considered is small (three or five) and the underlying deterministic problem itself, with 16 binary facility variables, is modest in size; the stochastic extension multiplies the number of continuous allocation and shortage variables by the number of scenarios but does not increase the number of binary variables, since facility-opening decisions remain first-stage and scenario-independent. The deterministic model solves in approximately 0.09 seconds; the three-scenario stochastic model, with 2,550 continuous variables against the deterministic model's 800, solves in approximately 0.19 seconds; and the five-scenario model, with 4,250 continuous variables, solves in a comparable 0.19 seconds. This confirms that, for the scenario counts considered in this study, solving the deterministic equivalent formulation directly imposes negligible computational overhead relative to the deterministic baseline. For applications requiring dozens or hundreds of scenarios to adequately represent a continuous or high-dimensional demand distribution, this direct approach would eventually become intractable, motivating the sampling-based methods discussed in Section 2.4.

3.5 Evaluation Metrics

Three related solutions are computed and compared, following the standard framework in Birge and Louveaux [3]:

The Recourse Problem (RP) solution solves the full two-stage stochastic program (5)-(7) directly, choosing y to minimize expected cost across all scenarios.

The Expected Value (EV) solution solves the deterministic problem (1)-(4) using nominal (mean) demand only, then that solution's facility-opening decisions y_{EV} are evaluated against every scenario to obtain the Expected result of using the EV solution (EEV), i.e. the true expected cost of committing to a facility plan chosen without regard to uncertainty.

The Wait-and-See (WS) solution solves the deterministic problem separately for each scenario, as though the scenario were known in advance, and averages the resulting costs; it represents a hypothetical upper bound on solution quality that is not achievable in practice, since facility decisions must be made before demand is known.

From these three quantities, two standard measures follow:

$$VSS = EEV - RP \tag{8}$$

$$EVPI = RP - WS \tag{9}$$

VSS quantifies the cost of ignoring uncertainty when making the first-stage decision; EVPI quantifies the cost of not knowing the future, even when uncertainty is properly modeled. Both are non-negative by construction, and together they decompose the total gap between naive planning and perfect foresight into an actionable component (VSS, addressed by better modeling at the planning stage) and an irreducible component (EVPI, addressed only by better forecasting or by building flexibility to react after the fact).

4. Data and Scenario Generation

The base network is instance cap41 from Beasley's OR-Library capacitated warehouse location data set [1], a real, established academic benchmark comprising 16 candidate facilities and 50 customers, each candidate facility with capacity 5,000 units and fixed opening cost 7,500 (one facility carries zero fixed cost, representing, for instance, an already-owned site with no additional opening investment required), and a published per-customer allocation cost for serving that customer's full nominal demand from each facility. Total nominal demand across all 50 customers is 58,268 units against 80,000 units of total candidate capacity, so the instance is feasible but capacity-constrained enough that facility selection meaningfully affects both cost and risk; a network with substantially more spare capacity than demand would be expected to show a smaller VSS, since even a suboptimal facility plan would have enough slack capacity to absorb higher-than-expected demand without shortage.

Demand uncertainty is modeled using three discrete scenarios in the base case, generated for this study rather than drawn from a published stochastic data set, since no standard stochastic extension of this benchmark exists: a Low scenario at 80% of nominal demand (probability 0.25), a Nominal scenario at 100% of nominal demand (probability 0.50), and a High scenario at 120% of nominal demand (probability 0.25), applied uniformly across all customers. This base scenario structure is deliberately simple, applying the same multiplier to every customer rather than allowing demand to vary independently or in spatially correlated patterns across the customer base; Section 6 tests whether the paper's qualitative conclusions are sensitive to this simplification by considering a wider three-scenario spread and an alternative five-scenario structure. The shortage penalty π is set, in the base case, to three times the average per-unit allocation cost across the instance, representing the cost of emergency or backup fulfillment for unmet demand; Section 6 also tests sensitivity to this parameter directly.

5. Computational Results

All models were solved to proven optimality using the SCIP mixed-integer solver via Google OR-Tools. Table 1 reports the wait-and-see cost for each individual scenario in the base case.

Scenario	Demand Multiplier	Probability	Wait-and-See Cost ($)
Low	0.80	0.25	794,295.84
Nominal	1.00	0.50	1,040,454.73
High	1.20	0.25	1,399,757.19

Table 1. Wait-and-See Optimal Cost by Demand Scenario

Table 2 reports the full set of solutions and derived metrics for the base case.

Quantity	Value ($)	Facilities Open
EV solution (deterministic, nominal demand)	1,040,454.73	13
EEV (EV solution evaluated under uncertainty)	1,156,822.95	13 (fixed)
RP (stochastic program optimum)	1,073,889.86	14
WS (wait-and-see expected cost)	1,068,740.62	varies by scenario
VSS = EEV − RP	82,933.09	—
EVPI = RP − WS	5,149.23	—

Table 2. Summary of Deterministic and Stochastic Solutions (Base Case)

The EV solution opens 13 of 16 facilities and achieves the lowest cost of any single solution (1,040,454.73) when evaluated only against nominal demand. However, when the EV solution's fixed facility plan is evaluated against all three demand scenarios, including the High scenario in which 13 facilities' combined capacity (65,000 units) is insufficient for realized demand (69,922 units), the EEV rises to 1,156,822.95, and an expected 1,230.4 units of demand go unmet even after accounting for the most cost-effective feasible reallocation. The RP solution opens 14 facilities, a plan that is more expensive under nominal demand alone but leaves zero expected shortage across all three scenarios, for an expected cost of 1,073,889.86.

The resulting VSS of $82,933.09 indicates that committing to a facility plan that explicitly accounts for demand uncertainty, rather than optimizing for the nominal case alone, is worth over $82,000 in expected cost for this instance, an 8.0% reduction relative to the EEV. By contrast, the EVPI of $5,149.23 is an order of magnitude smaller, only 0.48% of RP, indicating that once a reasonably robust facility plan is in place, there is comparatively little additional value in knowing exactly which demand scenario will occur in advance.

5.1 Facility Plan Comparison

Figure 1 shows which facilities are opened under each solution. The EV and RP plans agree on 13 facilities; the RP solution additionally opens Facility 15, the single change responsible for eliminating expected shortage under the High scenario.

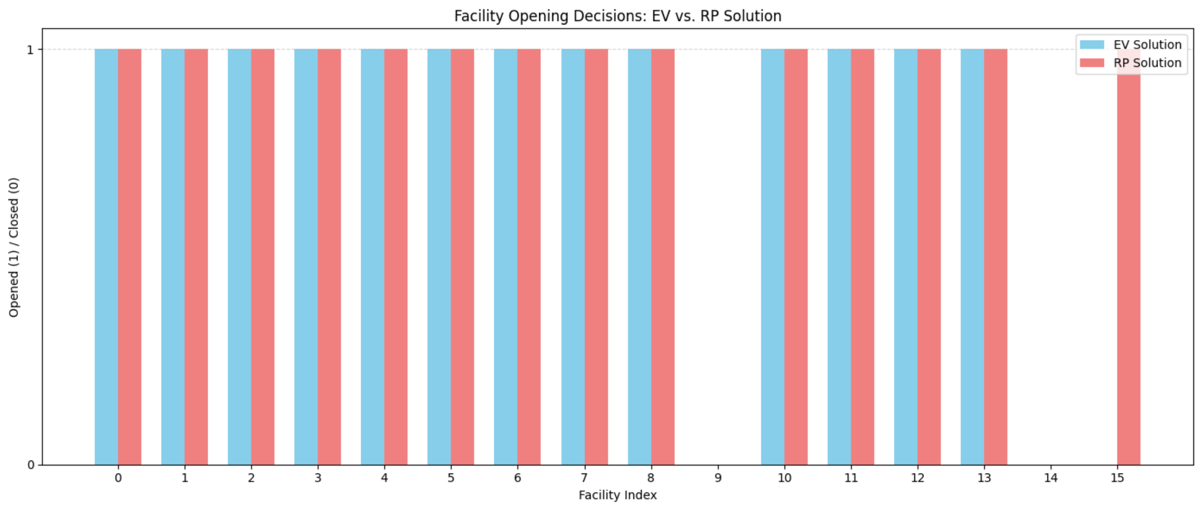

Figure 1. Facility Opening Decisions: EV vs. RP Solution

Figure 2 shows, for each plan's fixed facility set, the resulting capacity utilization under each demand scenario, computed by re-solving only the allocation (recourse) problem with facilities held fixed. Under the High scenario, every open facility in the EV plan reaches 100% utilization, meaning no facility retains any spare capacity, which is precisely why unmet demand becomes unavoidable once realized demand exceeds nominal by any margin. The RP plan's High-scenario row is nearly identical, except Facility 6 remains at 98% utilization rather than 100%, a small but non-zero buffer made possible by Facility 15 absorbing part of the load that would otherwise fall on Facilities 6 and 13. Comparing the Nominal and Low rows across the two plans shows the same reallocation effect at smaller scale: Facility 13's utilization is consistently lower under the RP plan than under the EV plan, since Facility 15 shares part of its demand once it is available.

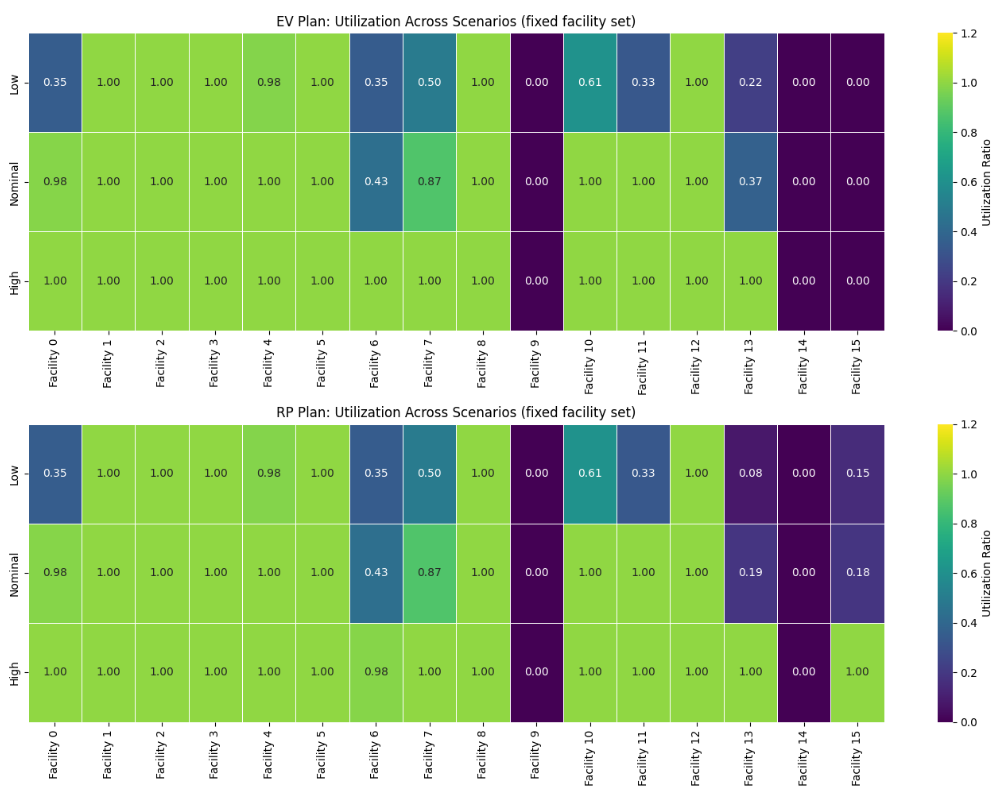

Figure 2. Facility Utilization Under the EV and RP Plans, Across Demand Scenarios

These figures make the mechanism behind the VSS finding concrete: the EV plan is not merely more expensive to run under high demand, it is structurally saturated, with zero slack anywhere in the network once demand rises. The RP plan trades a small amount of additional fixed cost for a small amount of spare capacity in exactly the facility where it is needed, which is sufficient to avoid shortage entirely at the scenario probabilities and magnitudes considered in this study.

5.2 Computational Performance

Table 2a reports solve time and model size for the deterministic model and the base and sensitivity stochastic models, confirming that exact solution via the deterministic equivalent formulation remains computationally trivial at the scenario counts considered in this study.

Model	Binary Vars	Continuous Vars	Constraints	Solve Time (s)
Deterministic	16	800	66	0.09
Stochastic, 3 scenarios	16	2,550	198	0.19
Stochastic, 5 scenarios	16	4,250	330	0.19

Table 2a. Model Size and Solve Time by Formulation

All solve times were recorded on a standard cloud compute instance with no parallelization beyond SCIP's default configuration; none of the models required more than a fraction of a second to solve to proven optimality, and solve time does not scale linearly with the roughly threefold increase in continuous variables between the deterministic and five-scenario models, likely reflecting that the additional variables introduced by each scenario are structurally similar and do not meaningfully increase the difficulty of the branch-and-bound search for a problem of this size.

6. Sensitivity Analysis

6.1 Sensitivity to Scenario Structure

Since the base-case scenario structure was generated for this study rather than estimated from historical demand data, it is important to test whether the paper's central finding, that VSS substantially exceeds EVPI, is a robust feature of this instance or an artifact of that particular scenario choice. Table 3 reports RP, VSS, and EVPI under the base three-scenario structure, an alternative five-scenario structure with a wider ±30% spread and finer granularity, and a three-scenario structure with the base probabilities but a wider ±30% demand spread.

Scenario Structure	RP ($)	Facilities	VSS ($)	EVPI ($)	VSS/EVPI
3-scenario, ±20% (base case)	1,073,889.86	14	82,933.09	5,149.23	16.1
5-scenario, ±30%	1,086,761.62	16	94,020.34	9,431.28	10.0
3-scenario, wide ±30%	1,112,430.10	16	173,113.57	10,034.36	17.3

Table 3. Sensitivity of Results to Demand Scenario Structure

Both alternative scenario structures increase the absolute magnitude of VSS and EVPI relative to the base case, which is expected, since a wider or more granular demand spread increases both the potential cost of a poorly chosen facility plan and the potential value of perfect information about which scenario will occur. Critically, however, the ratio of VSS to EVPI remains in a similar range across all three structures, between 10.0 and 17.3, indicating that the qualitative finding of this paper, that the value of a robust first-stage decision substantially exceeds the value of perfect foresight, is not an artifact of the specific base-case scenario probabilities and magnitudes chosen, but a structural feature of this facility network under demand uncertainty more broadly. The wider three-scenario structure also causes the stochastic-optimal solution to open all 16 candidate facilities, indicating that as the range of plausible demand outcomes widens, the model responds by maximizing available capacity headroom rather than selecting an intermediate number of facilities.

6.2 Sensitivity to Shortage Penalty Magnitude

The shortage penalty π is a modeling assumption representing the cost of failing to serve demand, whether through lost sales, emergency fulfillment, or service-level violations, and is not directly observed in the underlying benchmark instance. Table 4 reports results across a range of penalty multipliers, from 1.5 to 8 times the average per-unit allocation cost in the instance.

Shortage Penalty	RP ($)	Facilities	Expected Shortage (EEV, units)	VSS ($)
1.5× avg. unit cost	1,072,431.44	14	1,230.4	13,797.94
2× avg. unit cost	1,073,889.85	14	1,230.4	35,870.72
3× avg. unit cost (base case)	1,073,889.86	14	1,230.4	82,933.09
5× avg. unit cost	1,073,889.86	14	1,230.4	177,057.85
8× avg. unit cost	1,073,889.85	14	1,230.4	318,244.98

Table 4. Sensitivity of Results to Shortage Penalty Magnitude

Two findings stand out. First, as expected, VSS increases substantially with the shortage penalty, from $13,798 at the lowest penalty tested to $318,245 at the highest, since VSS directly reflects the cost of the EV solution's expected shortage, and a higher penalty makes that shortage more costly. Second, and less expected, the stochastic-optimal facility plan itself, and the resulting RP cost, are almost entirely insensitive to the penalty magnitude: 14 facilities are opened at every penalty level tested, and RP cost varies by less than $1,500 across a more than five-fold range of penalty values. This indicates that the RP solution's choice to open 14 facilities is not narrowly dependent on correctly estimating the cost of a shortage; rather, once the penalty is large enough to make any shortage clearly undesirable, which even the lowest multiplier tested achieves, the model converges on the same robust facility plan. This is a practically useful property, since a real decision-maker is unlikely to know the exact cost of a stockout, and would otherwise be concerned that the model's recommendation depends sensitively on a hard-to-estimate parameter.

7. Discussion

The gap between VSS and EVPI in this instance, and its robustness across the sensitivity analyses in Section 6, carries a clear practical implication: the primary risk in this facility location problem comes from making a fragile upfront commitment, not from an inability to react optimally once demand is observed. This is a meaningful distinction for a decision-maker choosing where to invest limited planning effort. If EVPI had been the larger quantity, the practical lesson would favor investing in better demand forecasting or building flexible, reconfigurable capacity. Since VSS dominates here, robustly across every scenario structure and penalty assumption tested, the stronger lesson is that the facility-opening decision itself should be stress-tested against a range of demand scenarios before being finalized, even if that means committing to a plan that looks suboptimal under any single point forecast.

It is also worth noting what the EV solution's failure actually represents. A 13-facility network is 3.2% cheaper than the 14-facility RP solution under nominal demand, a gap that could easily look like a clear, attractive cost saving to a decision-maker evaluating only a single forecast. The stochastic analysis reframes that apparent saving as a bet against demand variability, one that, for this instance, is not worth taking once the expected cost of shortage and reallocation under higher demand is properly accounted for.

This paper's finding that VSS substantially exceeds EVPI in a commercial facility location context is consistent with Rawls and Turnquist's [8] finding, in the very different context of disaster relief pre-positioning, that stochastic planning informed by a range of scenarios meaningfully outperforms planning based on a single expected scenario. The consistency of this qualitative pattern across two structurally similar but substantively very different applications, routine commercial distribution planning versus emergency humanitarian logistics, suggests that the VSS-dominates-EVPI relationship may be a more general feature of capacitated location problems with a hard capacity constraint and a meaningful cost of shortage, rather than a coincidence specific to either study's instance.

This observation also connects to the robust optimization literature reviewed in Section 2.3: since the practical value in this problem lies substantially in the first-stage decision's robustness, rather than in reacting optimally to resolved uncertainty, a robust optimization formulation, which directly optimizes first-stage decisions for worst-case or bounded-uncertainty performance without requiring a full probability distribution over scenarios, may achieve similar practical benefit to the stochastic programming approach used here, at a potentially lower modeling burden, since it would not require the decision-maker to specify scenario probabilities that, as this instance's scenarios illustrate, are themselves difficult to know precisely in practice.

7.1 Managerial Implications

For a practitioner responsible for a real facility location decision, three concrete implications follow from these results. First, a deterministic cost comparison between two candidate facility plans, of the kind that appears in this instance as a 3.2% nominal-case saving from opening one fewer facility, should not be treated as decisive without also evaluating each plan's performance under plausible higher-demand scenarios; this paper's results show that such a comparison can reverse entirely once uncertainty is accounted for. Second, the sensitivity analysis in Section 6.2 suggests that a practitioner need not know the exact cost of a stockout to benefit from this type of analysis: the facility plan recommended by the stochastic model was unchanged across a more than fivefold range of shortage-penalty assumptions, so a reasonable, even approximate, estimate of shortage cost is sufficient to identify a materially more robust facility plan. Third, the sensitivity analysis in Section 6.1 suggests that a practitioner without confident, validated demand scenario probabilities should not be deterred from applying this framework, since the qualitative recommendation, and the relative dominance of VSS over EVPI, remained stable across every scenario structure tested; the specific dollar magnitudes of VSS and EVPI are sensitive to scenario assumptions, but the higher-level conclusion, that robust upfront planning matters more than reactive flexibility for this type of network, is not.

8. Limitations

This study has several limitations that bound the scope of its conclusions. The demand scenarios used here, in both the base case and the sensitivity analysis, are generated for illustration rather than drawn from historical demand data or a validated forecasting model; a real application would calibrate scenario probabilities and magnitudes from actual demand history, and Section 6.1 shows that while the qualitative VSS-EVPI relationship appears robust to reasonable variation in scenario structure, the exact magnitudes reported here should not be treated as precise estimates for any real network.

The scenario structure applies a uniform multiplier across all customers in every configuration tested, which does not capture independent or spatially correlated demand variation across a real customer base; a network in which some customers' demand increases while others' decreases might show a different, and plausibly smaller, VSS, since aggregate demand variability would be partially diversified away rather than moving in lockstep. The shortage penalty, while shown in Section 6.2 to have limited influence on the facility plan actually selected, does directly determine the magnitude, though not the qualitative dominance, of VSS, and a real application should estimate this parameter from actual shortage costs rather than an illustrative multiple of allocation cost.

Finally, this analysis, including its sensitivity tests, is conducted entirely on a single benchmark instance and its variants; the relative magnitude of VSS versus EVPI is known in the stochastic programming literature to vary considerably across problem structures, particularly with the ratio of total capacity to total demand and the structure of allocation costs, and the consistency with Rawls and Turnquist's [8] finding noted in Section 7, while suggestive, is not a substitute for testing across a larger and more diverse set of benchmark instances.

9. Conclusion

This paper formulated and solved a two-stage stochastic extension of the capacitated facility location problem on a real academic benchmark instance, and applied the standard VSS and EVPI framework to decompose the value of explicitly modeling demand uncertainty. For this instance, accounting for uncertainty at the facility-opening stage was worth substantially more than knowing the future in advance, and this relationship was found to be robust across multiple alternative demand-scenario structures and shortage-penalty assumptions, indicating that upfront planning robustness, not reactive flexibility, is the primary lever available to a decision-maker facing this type of demand risk. This finding is consistent with related results in the humanitarian logistics literature, suggesting it may reflect a more general property of capacitated location problems under demand uncertainty.

Future work could extend this analysis in several directions: testing independently varying, spatially correlated, or continuously distributed demand scenarios, ideally calibrated to real historical demand data, rather than the discrete generated scenarios used here; applying the Sample Average Approximation method to scale this analysis to instances with a richer scenario space than can be directly enumerated; formulating and solving a robust optimization counterpart to this problem, following the approaches of Baron et al. [5] and Gülpınar et al. [6], to test whether it achieves comparable practical benefit to the stochastic programming approach with less reliance on precisely specified scenario probabilities; and testing across a larger and more diverse set of OR-Library benchmark instances to determine whether the VSS-dominates-EVPI relationship observed here generalizes across a broader range of network structures and capacity-to-demand ratios.

References

[1] J.E. Beasley, "An algorithm for solving large capacitated warehouse location problems," European Journal of Operational Research, vol. 33, pp. 314-325, 1988.

[2] L.V. Snyder, "Facility location under uncertainty: A review," IIE Transactions, vol. 38, no. 7, pp. 547-564, 2006.

[3] J.R. Birge and F. Louveaux, Introduction to Stochastic Programming, 2nd ed., Springer, New York, 2011.

[4] T. Santoso, S. Ahmed, M. Goetschalckx, and A. Shapiro, "A stochastic programming approach for supply chain network design under uncertainty," European Journal of Operational Research, vol. 167, no. 1, pp. 96-115, 2005.

[5] O. Baron, J. Milner, and H. Naseraldin, "Facility location: A robust optimization approach," Production and Operations Management, vol. 20, no. 5, pp. 772-785, 2011.

[6] N. Gülpınar, D. Pachamanova, and E. Çanakoğlu, "Robust strategies for facility location under uncertainty," European Journal of Operational Research, vol. 225, no. 1, pp. 21-35, 2013.

[7] A.J. Kleywegt, A. Shapiro, and T. Homem-de-Mello, "The sample average approximation method for stochastic discrete optimization," SIAM Journal on Optimization, vol. 12, no. 2, pp. 479-502, 2002.

[8] C.G. Rawls and M.A. Turnquist, "Pre-positioning of emergency supplies for disaster response," Transportation Research Part B: Methodological, vol. 44, no. 4, pp. 521-534, 2010.

[9] B. Balcik and B.M. Beamon, "Facility location in humanitarian relief," International Journal of Logistics Research and Applications, vol. 11, no. 2, pp. 101-121, 2008.

[10] L. Perron and V. Furnon, "OR-Tools," Google Optimization Tools, 2024. [Online]. Available: https://developers.google.com/optimization.

Code and Data Availability

Code, data, and full computational results supporting this paper, including the sensitivity analyses reported in Section 6, are available at github.com/sofiatajide/stochastic-facility-location.

Appendix A: Core Solver Implementation

The following Python implementation, using Google OR-Tools' linear solver interface, implements the deterministic model (Section 3.1) and the two-stage stochastic model with shortage penalty (Section 3.2) used to produce every result reported in this paper. The complete file, including the parser for the OR-Library instance format and the script used to generate Tables 1 through 4, is available at the repository linked above.

```
from ortools.linear_solver import pywraplp


def parse_cap(filepath):
    with open(filepath) as f:
        tokens = f.read().split()
    tokens = iter(tokens)
    n = int(next(tokens))
    m = int(next(tokens))
    capacity, fixed_cost = [], []
    for _ in range(n):
        capacity.append(float(next(tokens)))
        fixed_cost.append(float(next(tokens)))
    demand, alloc_cost = [], []
    for _ in range(m):
        demand.append(float(next(tokens)))
        alloc_cost.append([float(next(tokens)) for _ in range(n)])
    return {"n": n, "m": m, "capacity": capacity, "fixed_cost": fixed_cost,
            "demand": demand, "alloc_cost": alloc_cost}



def solve_deterministic(data, demand):
    n, m = data["n"], data["m"]
    capacity, fixed_cost, alloc_cost = data["capacity"], data["fixed_cost"], data["alloc_cost"]
    nominal_demand = data["demand"]

    solver = pywraplp.Solver.CreateSolver('SCIP')
    y = [solver.BoolVar(f'y_{i}') for i in range(n)]
    frac = [[solver.NumVar(0, 1, f'frac_{i}_{j}') for i in range(n)] for j in range(m)]

    for j in range(m):
        solver.Add(sum(frac[j][i] for i in range(n)) == 1)
    for i in range(n):
        solver.Add(sum(frac[j][i] * demand[j] for j in range(m)) <= capacity[i] * y[i])

    obj = solver.Sum(fixed_cost[i] * y[i] for i in range(n)) + \
        solver.Sum(alloc_cost[j][i] / nominal_demand[j] * demand[j] * frac[j][i]
                   for j in range(m) for i in range(n))
    solver.Minimize(obj)
    solver.Solve()
    return {"cost": solver.Objective().Value(), "y": [v.solution_value() for v in y]}


def solve_stochastic(data, scenarios, shortage_penalty, fix_y=None):
    n, m = data["n"], data["m"]
    capacity, fixed_cost, alloc_cost = data["capacity"], data["fixed_cost"], data["alloc_cost"]
    nominal_demand = data["demand"]
    unit_cost = [[alloc_cost[j][i] / nominal_demand[j] for i in range(n)] for j in range(m)]
    scenario_demand = [[nominal_demand[j] * s["mult"] for j in range(m)] for s in scenarios]
    S = len(scenarios)

    solver = pywraplp.Solver.CreateSolver('SCIP')
    y = [solver.BoolVar(f'y_{i}') for i in range(n)] if fix_y is None else fix_y

    frac = [[[solver.NumVar(0, 1, f'frac_{s}_{j}_{i}') for i in range(n)] for j in range(m)] for s in
range(S)]
    shortage = [[solver.NumVar(0, 1, f'short_{s}_{j}') for j in range(m)] for s in range(S)]

    for s in range(S):
        for j in range(m):
            solver.Add(sum(frac[s][j][i] for i in range(n)) + shortage[s][j] == 1)
        for i in range(n):
            solver.Add(sum(frac[s][j][i] * scenario_demand[s][j] for j in range(m)) <= capacity[i] *
y[i])
```

```python
    fixed_term = (solver.Sum(fixed_cost[i] * y[i] for i in range(n)) if fix_y is None
                  else sum(fixed_cost[i] * y[i] for i in range(n)))
    recourse_term = solver.Sum(
        scenarios[s]["prob"] * unit_cost[j][i] * scenario_demand[s][j] * frac[s][j][i]
        for s in range(S) for j in range(m) for i in range(n)
    )
    shortage_term = solver.Sum(
        scenarios[s]["prob"] * shortage_penalty * scenario_demand[s][j] * shortage[s][j]
        for s in range(S) for j in range(m)
    )
    solver.Minimize(fixed_term + recourse_term + shortage_term)
    solver.Solve()

    total_shortage = sum(shortage[s][j].solution_value() * scenario_demand[s][j] * scenarios[s]["prob"]
                         for s in range(S) for j in range(m))
    y_val = [(v.solution_value() if fix_y is None else v) for v in y]
    return {"cost": solver.Objective().Value(), "y": y_val, "expected_shortage": total_shortage}


if __name__ == "__main__":
    data = parse_cap("cap41.txt")
    nominal_demand = data["demand"]

    scenarios = [
        {"name": "Low", "mult": 0.80, "prob": 0.25},
        {"name": "Nominal", "mult": 1.00, "prob": 0.50},
        {"name": "High", "mult": 1.20, "prob": 0.25},
    ]

    avg_unit_cost = sum(data["alloc_cost"][j][i] / nominal_demand[j]
                        for j in range(data["m"]) for i in range(data["n"])) / (data["m"] * data["n"])
    shortage_penalty = 3 * avg_unit_cost

    ev_sol = solve_deterministic(data, nominal_demand)
    y_ev = [round(v) for v in ev_sol["y"]]
    print("EV solution cost:", round(ev_sol["cost"], 2), "| facilities open:", sum(y_ev))

    eev = solve_stochastic(data, scenarios, shortage_penalty, fix_y=y_ev)
    print("EEV:", round(eev["cost"], 2), "| expected shortage:", round(eev["expected_shortage"], 2))

    rp = solve_stochastic(data, scenarios, shortage_penalty)
    print("RP:", round(rp["cost"], 2), "| facilities open:", sum(round(v) for v in rp["y"]),
          "| expected shortage:", round(rp["expected_shortage"], 2))

    ws_total = 0
    for s in scenarios:
        demand = [nominal_demand[j] * s["mult"] for j in range(data["m"])]
        sol = solve_deterministic(data, demand)
        ws_total += s["prob"] * sol["cost"]
        print(f"Scenario {s['name']}: wait-and-see cost = {sol['cost']:.2f}")
    print("WS:", round(ws_total, 2))

    vss = eev["cost"] - rp["cost"]
    evpi = rp["cost"] - ws_total
    print("\nVSS = EEV - RP =", round(vss, 2))
    print("EVPI = RP - WS =", round(evpi, 2))
```