Journal of Industrial and Management Optimization

https://www.aimspress.com/journal/jimo

JIMO, 22(12): 6638–6667.
DOI:10.3934/jimo.2026222
Received: 25 August 2026
Revised: 02 November 2026
Accepted: 11 November 2026
Published: 18 November 2026

Research article

Material-feasible and dependence-informed distributionally robust lot-sizing for multi-level BOM planning

Jiayu Shen[1,*] **and Kai Zhu**[2]

[1] Department of Public Basic Courses, Nanjing University of Industry Technology, Nanjing 210023, Jiangsu, China

[2] School of Automobile and Traffic Engineering, Jiangsu University of Technology, Changzhou 213001, Jiangsu, China

* **Correspondence:** Emails: 2018100895@niit.edu.cn, fjcyue007@126.com.

Abstract: Multi-level bill-of-materials (BOM) planning is exposed to both demand uncertainty and production or procurement yield uncertainty. Two modeling issues are especially important in this setting. First, a component shortage should restrict the production of its parent items rather than appear only as a penalized negative inventory. Second, paired demand–yield observations contain cross-source dependence information that is lost when ambiguity is imposed separately on the two marginals. This paper developed a material-feasible and dependence-informed distributionally robust lot-sizing model for multi-level BOM planning. Internal components and subassemblies were subject to hard material-availability constraints, whereas finished-product demand may be backlogged. Under box support, the robust material-feasibility conditions admitted an exact finite linear counterpart. To use the dependence information in paired observations without imposing an independence assumption, the Wasserstein ground metric was constructed from a regularized covariance geometry of standardized joint demand–yield samples. Standard Wasserstein duality then yielded an exact semi-infinite representation. The material-feasible BOM structure reduced the combinatorial dimension of the sample-wise separation problem because sign-selection binaries are required only for finished-product positions whose surplus can change sign over the support. Computational experiments calibrated ambiguity radii model by model and varied dependence by fixed-marginal re-pairing. The full-versus-block metric ablation showed that cross-source dependence changes the cost–risk trade-off: under favorable dependence, the full metric reduces the conservatism of the block metric, whereas under neutral and adverse dependence, the paired comparisons did not establish a statistically clear advantage for either joint distributionally robust optimization (DRO) model. Across tree, assembly, and shared-component BOM structures, the master–separation procedure returned explicit optimality certificates under a common time budget.

Keywords: material requirements planning; multi-level lot-sizing; material feasibility; demand–yield dependence; Wasserstein distributionally robust optimization

Mathematics Subject Classification: 90C17, 90C15, 90B30, 90B05

1. Introduction

Material requirements planning (MRP) translates finished-product requirements into time-phased replenishment decisions for subassemblies, components, and raw materials connected through a multi-level bill-of-materials (BOM) structure [1, 2]. The mathematical foundation of this planning problem is closely related to dynamic lot sizing and multi-echelon inventory control. Wagner and Whitin [3] established the classical dynamic lot-sizing trade-off between setup and inventory costs, Zangwill [4] developed network formulations with backlogging and multi-echelon production, and Clark and Scarf [5] provided a foundational echelon-inventory view for multi-echelon systems. Subsequent work extended lot-sizing formulations to capacitated, multi-item, multi-level, and scheduling settings [6–10]. Broad reviews by Karimi et al. [11], Jans and Degraeve [12], and Buschkuehl et al. [13] document the development of this literature and its increasingly rich industrial variants.

For BOM-based planning, these models are especially relevant because a production decision for a parent item generates dependent requirements for its components. A shortage of a low-level item can therefore restrict several downstream production decisions, whereas excessive replenishment can create inventory at multiple levels of the network. This technological coupling makes material feasibility different from an ordinary item-wise shortage accounting device. In particular, if a negative component balance is merely penalized while parent production remains unchanged, the mathematical model may describe a low-cost plan that cannot actually be executed. A physically meaningful BOM model should therefore distinguish internal material availability from finished-product backlog and should state the within-period timing convention under which component replenishment becomes available for parent consumption.

Uncertainty complicates this requirement. Stochastic lot-sizing has long addressed random demand and service requirements [14, 15], while stochastic programming provides the general modeling framework for optimization under a prescribed probability law or scenario distribution [16]. More recent work has incorporated multi-level BOM structures, service constraints, and production flexibility [17]. A separate literature has studied random production or procurement yield. Early contributions include periodic-review and random-yield inventory models [18, 19], joint demand–yield settings with substitution [20], and reviews of multiple lot sizing under random yield [21]. Recent production-inventory research has examined demand and yield randomness jointly [22], while robust and distributionally robust lot-sizing models have been developed specifically for yield uncertainty [23–25]. These studies establish the operational importance of both demand and yield uncertainty, but they do not by themselves resolve how dependence between the two sources should be represented in a multi-level BOM network.

Robust optimization and distributionally robust optimization provide alternatives when a fully specified probability law is unavailable. Classical robust optimization protects decisions against all realizations in a prescribed uncertainty set [26, 27]. Distributionally robust optimization (DRO) instead optimizes against the worst probability distribution in an ambiguity set. Moment-based ambiguity sets and general distributionally robust convex optimization were developed, among others, by Delage and Ye [28] and Wiesemann et al. [29]. Wasserstein DRO is particularly attractive in data-driven settings because the ambiguity set can be centered directly at an empirical distribution; standard duality and finite-sample theory provide a tractable analytical basis for many such models [30, 31]. For BOM planning, however, the choice of ambiguity geometry matters because demand and yield are heterogeneous uncertainty sources and because their empirical pairing may contain operationally relevant cross-source information.

Demand and yield may indeed be statistically dependent. Demand surges can coincide with supplier congestion, material-quality deterioration, or process stress, while favorable operating conditions may support both stronger demand fulfillment and higher yield. Dependence between demand and random supply has been studied in inventory settings [32], and Chen and Xie [33] considered a data-driven setting with both demand and yield randomness. For data-driven BOM planning, paired historical observations are therefore more informative than two independently sampled marginal data sets. At the same time, simply concatenating demand and yield into a joint vector does not force every distribution in a Wasserstein ball to preserve the empirical covariance. The modeling question is therefore not whether the ambiguity set can be claimed to maintain the observed dependence exactly, but how cross-source information in the paired data should enter the transportation geometry used to measure plausible distributional perturbations.

As an illustrative industrial setting, consider a multi-stage assembly operation in which finished products require several layers of subassemblies and purchased components. A surge in customer orders may coincide with supplier congestion, quality problems, or production overload, reducing the usable yield of upstream materials precisely when more components are required. Conversely, favorable operating conditions may allow high demand to coincide with relatively high usable yield. These mechanisms motivate examining different demand–yield dependence regimes, although their actual strength is industry- and data-dependent. This paper addresses the dependence issue together with the material-feasibility issue. We retain paired demand–yield observations, standardize their heterogeneous coordinates, and use a regularized covariance-informed transformation in the Wasserstein ground metric. The resulting ambiguity set does not impose the empirical covariance as a hard moment equality; rather, empirical cross-dependence enters the transportation geometry and therefore the cost of moving probability mass away from observed joint samples. In parallel, internal components and subassemblies satisfy robust material-availability constraints, while finished-product demand may be backlogged and penalized. The within-period sequence is stated explicitly: usable replenishment becomes available within the bucket before parent consumption and external demand are charged; deterministic lead times can be represented by shifting the corresponding indices.

Relation to recent work. Recent studies have investigated stochastic and robust planning in multi-level production systems. Thevenin et al. [34] developed stochastic MRP models under demand uncertainty. Sereshti et al. [17] examined production flexibility and service-level constraints in multi-level lot sizing. Schlenkrich and Parragh [35] studied capacitated multi-echelon lot sizing with setup carry-over under uncertain demand. Metzker et al. [25] investigated data-driven distributionally robust lot sizing under yield uncertainty, while Schlenkrich et al. [36] developed a progressive hedging approach for multi-stage stochastic lot sizing. These studies address closely related planning challenges but emphasize different uncertainty sources and decision structures. Our work focuses on the joint treatment of demand and yield uncertainty, support-robust internal BOM feasibility, and covariance-informed Wasserstein transportation geometry. Table 1 summarizes these methodological distinctions.

Our contributions are fourfold. First, we formulate a multi-level lot-sizing model that enforces hard internal material availability under joint demand and yield uncertainty, with an exact linear robust counterpart under box support. Second, we incorporate cross-source dependence into a covariance-informed Wasserstein ground metric without imposing fixed covariance constraints on candidate distributions. Third, material feasibility limits surplus-sign binaries in the separation mixed-integer linear

program (MILP) to sign-indefinite finished-product positions. Finally, we evaluate these mechanisms through calibrated full-versus-block comparisons, repeated experiments, computational certificates, and sensitivity analyses.

The remainder of the paper is organized as follows. Section 2 defines the BOM system, the within-period timing convention, joint uncertainty, and robust material feasibility. Section 3 presents the dependence-informed Wasserstein ambiguity set and the distributionally robust lot-sizing model. Section 4 derives the finite material-feasibility counterpart, states the standard Wasserstein dual characterization, and develops the exact master–separation algorithm. Section 5 describes the computational design and numerical evaluation. Section 6 concludes the paper.

Table 1. Comparison with recent MRP and lot-sizing research based on principal modeling features.

Study	Main uncertainty	Main methodological focus
Thevenin et al. [34]	Demand	Stochastic MRP; two-/multi-stage planning and rolling horizon
Sereshti et al. [17]	Demand	Multi-level BOM flexibility and service-level constraints
Schlenkrich & Parragh [35]	Demand	Capacitated multi-echelon lot sizing; stochastic and robust setup carry-over
Metzker et al. [25]	Yield	Data-driven distributionally robust multi-item lot sizing
Schlenkrich et al. [36]	Demand	Multi-stage stochastic lot sizing; progressive hedging
This study	Joint demand and yield	Hard internal BOM feasibility; covariance-informed Wasserstein metric; reduced separation binaries

2. Problem description

This section defines the multi-level BOM planning system, the joint demand–yield uncertainty, and the distinction between finished-product backlog and internal material feasibility. The main notation is summarized in Table 2.

Table 2. Main notation.

Symbol	Description
$\mathcal{I}$	Set of all items
$\mathcal{F}$	Set of finished products for which external-demand backlog is allowed
$\mathcal{I}^{\text{int}} = \mathcal{I} \setminus \mathcal{F}$	Set of internal items (subassemblies, components, and raw materials) subject to hard material feasibility
$\mathcal{T}$	Set of planning periods
$\mathcal{P}(j)$	Set of components required by parent item j
$C(i)$	Set of parent items that consume component i
a_{ij}	Units of component i required per unit of parent item j
x_{it}, z_{it}	Planned production/procurement quantity and setup indicator
I_{i0}	Initial inventory of item i
D_{it}, Y_{it}	Uncertain external demand and yield factor
$\xi = (D, Y)$	Joint demand–yield uncertainty vector

Continued on the next page

Symbol	Description
Ξ	Box support of ξ
$\widehat{\xi}^n$	The n-th paired historical demand–yield observation
$S_{it}(x,\xi)$	Cumulative end-of-period material surplus
$\Phi_{\mathrm{MF}}(x,\xi)$	Material-feasible inventory and finished-product backlog cost
R	Diagonal marginal standardization matrix
$\widehat{\Sigma}$	Empirical covariance matrix of standardized paired observations, partitioned into demand, yield, and cross blocks
W_{sh}	Shrinkage covariance-informed transport transformation
$W_{\mathrm{sh}}^{\mathrm{blk}}$	Cross-block-ablated shrinkage transformation retaining within-source covariance only
$\widehat{\mathbb{P}}_N$	Empirical distribution of paired joint observations
$\mathcal{P}_N(\varepsilon)$	Wasserstein ambiguity set centered at $\widehat{\mathbb{P}}_N$
K_i	Fixed setup cost of item i
c_i	Unit production or procurement cost of item i
h_i	Unit inventory-holding cost of item i
p_i	Unit backlog penalty for finished product i
M_{it}	Maximum lot size of item i in period t
r_i	Unit resource consumption of item i
q_i	Fixed resource consumption for setting up item i
B_t	Available resource capacity in period t
$\underline{d}_{it}, \overline{d}_{it}$	Lower and upper demand bounds
$\underline{y}_{it}, \overline{y}_{it}$	Lower and upper yield bounds
N	Number of paired training observations
ε	Wasserstein ambiguity radius
λ	Nonnegative Wasserstein dual variable
η_n	Sample-wise epigraph variable
$\mathcal{X}_{\mathrm{MF}}$	Material-feasible first-stage decision set
$\mathcal{V}_n$	Finite set of adverse realizations for sample n
θ_n	Optimal value of the sample-wise separation problem

2.1. BOM network, decisions, and within-period timing

Let $\mathcal{I} = \{1, \ldots, I\}$ be the item set and $\mathcal{T} = \{1, \ldots, T\}$ the planning horizon. The BOM graph is assumed to be acyclic. The set $\mathcal{F} \subseteq \mathcal{I}$ contains terminal finished products whose external demand can be backlogged, and $\mathcal{I}^{\mathrm{int}} = \mathcal{I} \setminus \mathcal{F}$ contains internal items whose availability constrains downstream production. If item i is a component of parent item j, then $a_{ij} > 0$ is the required BOM coefficient. We write

$$\mathcal{P}(j) = \{i \in \mathcal{I} : a_{ij} > 0\}, \qquad C(i) = \{j \in \mathcal{I} : a_{ij} > 0\}.$$

For each $i \in \mathcal{I}$ and $t \in \mathcal{T}$, the first-stage plan chooses $x_{it} \geq 0$ and $z_{it} \in \{0, 1\}$, with

$$0 \leq x_{it} \leq M_{it} z_{it}.$$

The usable output from the planned quantity is $Y_{it} x_{it}$. We use a zero-lead-time bucket convention. Within each period, items can be processed in a topological order of the acyclic BOM: replenishment of a lower-level item becomes available before that item is consumed by its parents, and finished-product external demand is charged after the required internal consumption. This convention makes the usual same-bucket material balance physically interpretable. The zero-lead-time convention is adopted to focus on joint demand–yield uncertainty and material feasibility. Fixed positive lead times could be represented by appropriately shifting material receipt periods and accounting for initial pipeline inventory. Such an extension would require corresponding changes to the inventory-balance equations and their robust counterparts. Stochastic lead times, particularly those caused by congestion or capacity disruptions, are not covered by the present formulation.

2.2. Joint demand–yield uncertainty

The uncertain vector is

$$\xi = (D, Y), \qquad D = \{D_{it} : i \in \mathcal{I}, t \in \mathcal{T}\}, \qquad Y = \{Y_{it} : i \in \mathcal{I}, t \in \mathcal{T}\}.$$

Its support is defined explicitly as the box

$$\Xi = \left\{(D, Y) : \underline{d}_{it} \leq D_{it} \leq \overline{d}_{it},\ \underline{y}_{it} \leq Y_{it} \leq \overline{y}_{it},\ i \in \mathcal{I},\ t \in \mathcal{T}\right\},$$

where $\underline{d}_{it} \geq 0$ and $0 < \underline{y}_{it} \leq \overline{y}_{it} \leq 1$. In the usual BOM setting, external demand is zero for most internal items; the notation nevertheless allows bounded external demand when required by an application.

Suppose N paired historical observations are available:

$$\widehat{\xi}^n = (\widehat{D}^n, \widehat{Y}^n), \qquad n = 1, \ldots, N.$$

The pairing contains empirical cross-source information. We do not assume that every probability distribution considered later has exactly the empirical covariance. Instead, Section 3 uses the paired observations to construct the transportation geometry of the Wasserstein ambiguity set.

2.3. Period balances, cumulative surplus, and material feasibility

To make within-period timing explicit, define the end-of-period net material position recursively by

$$S_{i0} = I_{i0}, \qquad S_{it}(x, \xi) = S_{i,t-1}(x, \xi) + Y_{it} x_{it} - D_{it} - \sum_{j \in C(i)} a_{ij} x_{jt}. \tag{2.1}$$

Expanding the recursion gives the cumulative representation

$$S_{it}(x, \xi) = I_{i0} + \sum_{\tau=1}^{t} Y_{i\tau} x_{i\tau} - \sum_{\tau=1}^{t} D_{i\tau} - \sum_{\tau=1}^{t} \sum_{j \in C(i)} a_{ij} x_{j\tau}. \tag{2.2}$$

Thus the cumulative expression is not used to replace the physical flow interpretation; it is the closed-form expansion of the period-by-period balance under the stated timing convention.

For a finished product $i \in \mathcal{F}$, negative surplus represents backlog, so

$$S_{it}^{+}(x,\xi) = \max\{S_{it}(x,\xi), 0\}, \qquad S_{it}^{-}(x,\xi) = \max\{-S_{it}(x,\xi), 0\}.$$

For an internal item $i \in \mathcal{I}^{\text{int}}$, negative surplus is not treated as a soft shortage. Parent production must be supported by material availability, and we impose

$$S_{it}(x,\xi) \geq 0, \qquad i \in \mathcal{I}^{\text{int}},\ t \in \mathcal{T},\ \xi \in \Xi. \tag{2.3}$$

Thus insufficient component or subassembly material cannot be compensated for merely by paying a shortage penalty. Under the static first-stage planning interpretation adopted here, upstream yield risk is absorbed ex ante through material buffers and production/procurement quantities that guarantee executable parent plans throughout the operational support Ξ. The guarantee is therefore conditional on the chosen support bounds; excessively wide physical bounds would create an intentionally conservative design and should not be interpreted as a free statistical confidence statement.

The random inventory and finished-product backlog cost for a material-feasible plan is

$$\Phi_{\text{MF}}(x,\xi) = \sum_{i\in\mathcal{I}^{\text{int}}}\sum_{t\in\mathcal{T}} h_i S_{it}(x,\xi) + \sum_{i\in\mathcal{F}}\sum_{t\in\mathcal{T}} \left[h_i S_{it}^{+}(x,\xi) + p_i S_{it}^{-}(x,\xi)\right]. \tag{2.4}$$

Since (2.3) guarantees $S_{it} \geq 0$ for internal items, their holding terms are linear. The piecewise-linear positive/negative representation is needed only for finished products.

2.4. Deterministic planning constraints and cost

The deterministic setup and production/procurement cost is

$$C(x,z) = \sum_{i\in\mathcal{I}}\sum_{t\in\mathcal{T}} (K_i z_{it} + c_i x_{it}).$$

Let $\mathcal{X}_0$ collect additional deterministic restrictions such as capacity, procurement budgets, storage limits, or lot-size bounds. For example,

$$\sum_{i\in\mathcal{I}} r_i x_{it} + \sum_{i\in\mathcal{I}} q_i z_{it} \leq B_t, \qquad t \in \mathcal{T}.$$

The capacity parameter B_t may vary across periods, allowing the model to represent known time-dependent resource availability and deterministic bottlenecks. However, the present formulation assumes that these capacities are specified before the first-stage decisions are made. Unexpected capacity losses, machine breakdowns, and congestion-induced delays would require additional uncertainty descriptions and potentially adjustable production decisions. Such effects are outside the scope of the current model.

We first define the conventional deterministic decision set

$$\mathcal{X} = \{(x,z) : 0 \leq x_{it} \leq M_{it} z_{it},\ z_{it} \in \{0,1\},\ (x,z) \in \mathcal{X}_0\},$$

and the material-feasible set

$$\mathcal{X}_{\text{MF}} = \left\{(x,z) \in \mathcal{X} : S_{it}(x,\xi) \geq 0\ \forall \xi \in \Xi,\ i \in \mathcal{I}^{\text{int}},\ t \in \mathcal{T}\right\}. \tag{2.5}$$

All models compared later use the same material-feasibility convention so that numerical differences can be attributed to uncertainty modeling rather than to allowing physically infeasible component shortages in some benchmarks.

3. Dependence-informed joint demand–yield DRO model

3.1. Empirical joint distribution and covariance-informed transport geometry

The empirical distribution of the paired observations is

$$\widehat{\mathbb{P}}_N = \frac{1}{N}\sum_{n=1}^{N} \delta_{\widehat{\xi}^n}.$$

A joint empirical distribution retains the observed pairing at its center, but a Wasserstein ball around that center can contain distributions with different dependence structures. We therefore do not describe a joint Wasserstein ball as automatically "preserving" dependence. Instead, empirical cross-source information is allowed to shape the transportation geometry.

Let $m = 2|\mathcal{I}||\mathcal{T}|$ be the displayed dimension of ξ; coordinates that are deterministic on Ξ can be removed in implementation. Let $R = \mathrm{diag}(r_1, \ldots, r_m)$ standardize the heterogeneous demand and yield coordinates, with $r_q = 1/\max\{\widehat{s}_q, \kappa_q\}$, where $\widehat{s}_q$ is the sample standard deviation and $\kappa_q > 0$ is a scale-compatible floor used only for nearly deterministic coordinates. Define

$$\widehat{u}^n = R(\widehat{\xi}^n - \bar{\xi}), \qquad \bar{\xi} = \frac{1}{N}\sum_{n=1}^{N} \widehat{\xi}^n,$$

and the empirical covariance matrix

$$\widehat{\Sigma} = \frac{1}{N-1}\sum_{n=1}^{N} \widehat{u}^n (\widehat{u}^n)^\top = \begin{bmatrix} \widehat{\Sigma}_{DD} & \widehat{\Sigma}_{DY} \\ \widehat{\Sigma}_{YD} & \widehat{\Sigma}_{YY} \end{bmatrix}.$$

Since the uncertainty dimension can be much larger than the number of historical samples, $\widehat{\Sigma}$ is typically singular. Rather than introduce an ad hoc ridge parameter, we use a data-driven linear shrinkage estimator. Let

$$\widehat{\mu} = \frac{\mathrm{tr}(\widehat{\Sigma})}{m}, \qquad \widetilde{\Sigma} = (1-\widehat{\gamma})\widehat{\Sigma} + \widehat{\gamma}\widehat{\mu} I_m, \tag{3.1}$$

where $\widehat{\gamma} \in [0,1]$ is the Ledoit–Wolf shrinkage intensity estimated from the standardized paired observations [37]. The resulting whitening-type transformation is

$$W_{\mathrm{sh}} = \widetilde{\Sigma}^{-1/2} R, \tag{3.2}$$

and the transformed ℓ_1 ground metric is

$$d_{\mathrm{DI}}(\xi, \xi') = \|W_{\mathrm{sh}}(\xi - \xi')\|_1 . \tag{3.3}$$

Conditional on the training data, W_{sh} is fixed and nonsingular whenever $\widehat{\gamma} > 0$ and $\widehat{\mu} > 0$; deterministic coordinates can be removed before constructing the metric. Hence d_{DI} is a norm-induced metric. The inverse square root makes perturbations along empirically high-variation joint directions less expensive than perturbations along weakly observed directions, while shrinkage stabilizes that geometry when $m \gg N$.

The off-diagonal blocks $\widehat{\Sigma}_{DY}$ and $\widehat{\Sigma}_{YD}$ are the only terms that encode demand–yield cross-covariance after marginal standardization. Shrinkage attenuates these empirical correlations by the common factor $1-\widehat{\gamma}$, but does not change their sign or introduce artificial cross-source dependence. Thus a historically frequent high-demand/low-yield direction remains less expensive under the full geometry than an otherwise comparable direction unsupported by the paired data. This is the precise modeling claim: dependence informs the transport cost; it is not imposed as a hard covariance equality on every distribution in the ambiguity set.

For later identification of cross-source dependence, define

$$\widehat{\Sigma}^{\mathrm{blk}} = \begin{bmatrix} \widehat{\Sigma}_{DD} & 0 \\ 0 & \widehat{\Sigma}_{YY} \end{bmatrix}$$

and apply the *same* shrinkage intensity and target scale,

$$\widetilde{\Sigma}^{\mathrm{blk}} = (1-\widehat{\gamma})\widehat{\Sigma}^{\mathrm{blk}} + \widehat{\gamma}\widehat{\mu}I_m, \qquad W_{\mathrm{sh}}^{\mathrm{blk}} = (\widetilde{\Sigma}^{\mathrm{blk}})^{-1/2}R. \tag{3.4}$$

The full and block-ablated metrics therefore differ only through the demand–yield cross blocks; marginal scaling, within-demand covariance, within-yield covariance, shrinkage intensity, and the shrinkage target are held fixed.

3.2. Wasserstein ambiguity set

For probability distributions $\mathbb{P}$ and $\mathbb{Q}$ supported on Ξ, let

$$W_{d_{\mathrm{DI}}}(\mathbb{P},\mathbb{Q}) = \inf_{\Pi\in\mathcal{M}(\mathbb{P},\mathbb{Q})} \int_{\Xi\times\Xi} d_{\mathrm{DI}}(\xi,\xi')\,\Pi(d\xi,d\xi'),$$

where $\mathcal{M}(\mathbb{P},\mathbb{Q})$ is the set of couplings with marginals $\mathbb{P}$ and $\mathbb{Q}$, and $\mathcal{P}(\Xi)$ denotes the probability measures supported on Ξ. For $\varepsilon \geq 0$, define

$$\mathcal{P}_N(\varepsilon) = \left\{\mathbb{P}\in\mathcal{P}(\Xi) : W_{d_{\mathrm{DI}}}(\mathbb{P},\widehat{\mathbb{P}}_N) \leq \varepsilon\right\}. \tag{3.5}$$

The radius controls the amount of distributional perturbation measured in the dependence-informed geometry. Because different benchmark models operate in different uncertainty dimensions, Section 5 calibrates their radii separately under a common statistical criterion rather than assigning the same numerical radius to every model.

3.3. Worst-case cost and material-feasible DRO model

For $(x,z)\in\mathcal{X}_{\mathrm{MF}}$, define

$$Q_\varepsilon^{\mathrm{MF}}(x) = \sup_{\mathbb{P}\in\mathcal{P}_N(\varepsilon)} \mathbb{E}_{\mathbb{P}}[\Phi_{\mathrm{MF}}(x,\xi)]. \tag{3.6}$$

The proposed model is

$$\min_{(x,z)\in\mathcal{X}_{\mathrm{MF}}} \left\{C(x,z) + Q_\varepsilon^{\mathrm{MF}}(x)\right\}. \tag{MF-JY-DRO}$$

The first-stage plan must therefore be materially feasible for every supported realization, while the expected holding and finished-product backlog cost is protected against distributional misspecification inside the Wasserstein ambiguity set.

The two robustness layers play deliberately different roles. Material feasibility is support-robust: dependence information is not used to justify physically infeasible component plans. Statistical dependence instead shapes the distributional worst-case cost among plans that are executable throughout Ξ. Thus the Wasserstein radius controls economic distributional robustness, whereas the support bounds control the strength of the physical material-feasibility guarantee.

3.4. Relation to benchmark models

All benchmark models use the same BOM constraints and the same material-feasible set $\mathcal{X}_{\mathrm{MF}}$. They differ only in how demand and yield uncertainties are represented. This common feasible set is important: otherwise, a model that permits negative internal material could appear cheaper merely because it is allowed to violate the physical BOM.

If demand and yield are replaced by their sample means

$$\bar{D} = \frac{1}{N}\sum_{n=1}^{N} \widehat{D}^n, \qquad \bar{Y} = \frac{1}{N}\sum_{n=1}^{N} \widehat{Y}^n,$$

we obtain the deterministic benchmark

$$\min_{(x,z)\in\mathcal{X}_{\mathrm{MF}}} \left\{ C(x,z) + \Phi_{\mathrm{MF}}(x,\bar{\xi}) \right\}, \qquad \bar{\xi} = (\bar{D}, \bar{Y}).$$

Setting $\varepsilon = 0$ yields the sample average approximation (SAA) benchmark

$$\min_{(x,z)\in\mathcal{X}_{\mathrm{MF}}} \left\{ C(x,z) + \frac{1}{N}\sum_{n=1}^{N} \Phi_{\mathrm{MF}}(x,\widehat{\xi}^n) \right\}.$$

D-DRO treats only D as distributionally ambiguous and fixes yield at $\bar{Y}$, whereas Y-DRO treats only Y as ambiguous and fixes demand at $\bar{D}$. The proposed MF-JY-DRO instead operates on the paired vector $\xi = (D, Y)$.

To isolate the value of demand–yield cross-dependence in the transport geometry, we additionally define a cross-block-ablated joint DRO benchmark using $W_{\mathrm{sh}}^{\mathrm{blk}}$ from (3.4). It uses the same paired empirical distribution, the same joint uncertainty dimension, and the same within-demand and within-yield covariance blocks as MF-JY-DRO, but sets $\widehat{\Sigma}_{DY} = \widehat{\Sigma}_{YD} = 0$. Hence the difference between the full and block-ablated joint models is attributable to cross-source covariance information rather than to marginal scaling, within-source serial correlation, or the use of paired samples itself.

4. Structural reformulation and certified solution framework

This section separates two ingredients that play different roles. The robust BOM material-feasibility constraints admit a problem-specific finite linear counterpart. The worst-case expectation is then represented by standard Wasserstein duality and solved by a master–separation procedure. The structural benefit of material feasibility is that internal-item surplus never changes sign, which reduces the binary dimension of the separation problem.

4.1. Finite counterpart of robust material feasibility

For $i \in \mathcal{I}^{\text{int}}$ and $t \in \mathcal{T}$, define the lower envelope

$$\underline{S}_{it}(x) = I_{i0} + \sum_{\tau=1}^{t} \underline{y}_{i\tau} x_{i\tau} - \sum_{\tau=1}^{t} \overline{d}_{i\tau} - \sum_{\tau=1}^{t} \sum_{j \in C(i)} a_{ij} x_{j\tau}. \tag{4.1}$$

Proposition 1. *Under the box support Ξ, constraint* (2.3) *is equivalent to the finite family of linear inequalities*

$$\underline{S}_{it}(x) \geq 0, \qquad i \in \mathcal{I}^{\text{int}},\ t \in \mathcal{T}. \tag{4.2}$$

Proof. For fixed x, expression (2.2) is componentwise nondecreasing in each yield factor $Y_{i\tau}$ that contributes to item i's replenishment and nonincreasing in each external-demand component $D_{i\tau}$. The parent-consumption terms are first-stage quantities and are therefore fixed with respect to ξ. Hence the minimum of $S_{it}(x, \xi)$ over the box Ξ is attained at $Y_{i\tau} = \underline{y}_{i\tau}$ and $D_{i\tau} = \overline{d}_{i\tau}$ for $\tau \leq t$, yielding $\underline{S}_{it}(x)$. Therefore $S_{it}(x, \xi) \geq 0$ for every $\xi \in \Xi$ if and only if $\underline{S}_{it}(x) \geq 0$. □

Proposition 1 turns the physical material-feasibility requirement into deterministic linear inequalities that can be placed directly in the master problem. The result is exact for the stated operational support box, not a probabilistic approximation. Its conservatism is therefore controlled by the support bounds rather than by the Wasserstein radius: widening Ξ strengthens the physical execution guarantee but can force larger ex-ante buffers. We henceforth use

$$\mathcal{X}_{\text{MF}} = \left\{ (x, z) \in \mathcal{X} : \underline{S}_{it}(x) \geq 0,\ i \in \mathcal{I}^{\text{int}},\ t \in \mathcal{T} \right\}.$$

When internal items have zero external demand, the corresponding $\overline{d}_{i\tau}$ terms vanish.

4.2. Wasserstein dual characterization

The following representation is a direct application of standard Wasserstein/Kantorovich duality rather than a new duality result. For fixed x, define

$$\theta_n(x, \lambda) = \sup_{\xi \in \Xi} \left\{ \Phi_{\text{MF}}(x, \xi) - \lambda d_{\text{DI}}(\xi, \widehat{\xi}^n) \right\}.$$

Under compact Ξ and the continuity of $\Phi_{\text{MF}}(x, \cdot)$, standard Wasserstein duality [30, 31] gives

$$Q_{\varepsilon}^{\text{MF}}(x) = \inf_{\lambda \geq 0} \left\{ \lambda \varepsilon + \frac{1}{N} \sum_{n=1}^{N} \theta_n(x, \lambda) \right\}. \tag{4.3}$$

Consequently, MF-JY-DRO is equivalent to

$$\begin{aligned} \min_{x, z, \lambda, \eta} \quad & C(x, z) + \lambda \varepsilon + \frac{1}{N} \sum_{n=1}^{N} \eta_n \\ \text{s.t.} \quad & \eta_n \geq \Phi_{\text{MF}}(x, \xi) - \lambda d_{\text{DI}}(\xi, \widehat{\xi}^n), \quad \xi \in \Xi,\ n = 1, \ldots, N, \\ & (x, z) \in \mathcal{X}_{\text{MF}}, \qquad \lambda \geq 0. \end{aligned} \tag{4.4}$$

Formulation (4.4) is exact but semi-infinite; no claim is made that this representation is by itself computationally tractable.

4.3. Finite master problem

Let $\mathcal{V}_n$ be the current finite set of adverse realizations associated with sample n, and let $\widetilde{\xi}^{n\ell}$ denote realization $\ell \in \mathcal{V}_n$. For finished products, we introduce epigraph variables $r_{it}^{n\ell}$ satisfying

$$r_{it}^{n\ell} \geq h_i S_{it}(x, \widetilde{\xi}^{n\ell}), \qquad r_{it}^{n\ell} \geq -p_i S_{it}(x, \widetilde{\xi}^{n\ell}), \qquad i \in \mathcal{F}.$$

Internal-item holding cost remains linear because Proposition 1 guarantees nonnegative surplus for every supported realization. The finite master problem is

$$\begin{aligned}
\min_{x,z,\lambda,\eta,r} \quad & C(x,z) + \lambda\varepsilon + \frac{1}{N}\sum_{n=1}^{N} \eta_n \\
\text{s.t.} \quad & \eta_n \geq \sum_{i\in\mathcal{I}^{\text{int}}}\sum_{t\in\mathcal{T}} h_i S_{it}(x, \widetilde{\xi}^{n\ell}) + \sum_{i\in\mathcal{F}}\sum_{t\in\mathcal{T}} r_{it}^{n\ell} - \lambda d_{\text{DI}}(\widetilde{\xi}^{n\ell}, \widehat{\xi}^{n}), && n = 1,\dots,N,\ \ell \in \mathcal{V}_n, \\
& r_{it}^{n\ell} \geq h_i S_{it}(x, \widetilde{\xi}^{n\ell}), && i \in \mathcal{F},\ t \in \mathcal{T},\ n, \ell, \\
& r_{it}^{n\ell} \geq -p_i S_{it}(x, \widetilde{\xi}^{n\ell}), && i \in \mathcal{F},\ t \in \mathcal{T},\ n, \ell, \\
& \underline{S}_{it}(x) \geq 0, && i \in \mathcal{I}^{\text{int}},\ t \in \mathcal{T}, \\
& (x,z) \in \mathcal{X}, \qquad \lambda \geq 0.
\end{aligned} \tag{4.5}$$

This is a mixed-integer linear program when $\mathcal{X}$ is mixed-integer linear and the adverse realizations are fixed. It is a relaxation of (4.4) because only finitely many epigraph constraints are enforced. In formulation (4.5), $C(x,z)$ represents the deterministic setup and production costs, while $\lambda\varepsilon$ accounts for the Wasserstein ambiguity radius. The variables η_n represent sample-wise worst-case cost contributions. For each retained adverse realization, $r_{it}^{n\ell}$ captures the larger of the holding and backlog costs for finished product i in period t. The internal-item holding costs remain linear under hard material feasibility. The constraints $\underline{S}_{it}(x) \geq 0$ enforce robust component availability, whereas $(x,z) \in \mathcal{X}$ imposes the deterministic planning restrictions. The finite master provides a lower-bound relaxation of the semi-infinite problem.

4.4. Sample-wise separation with dependence-informed distance

Let $(x^k, z^k, \lambda^k, \eta^k)$ be the incumbent master solution. For sample n, separation solves

$$\theta_n(x^k, \lambda^k) = \sup_{\xi\in\Xi} \left\{ \Phi_{\text{MF}}(x^k, \xi) - \lambda^k d_{\text{DI}}(\xi, \widehat{\xi}^n) \right\}. \tag{4.6}$$

For a fixed master solution, formulation (4.6) identifies the most adverse demand–yield realization for historical sample n. The first term measures the resulting inventory-holding and finished-product backlog costs, while the second term penalizes deviations from the observed sample according to the dependence-informed transportation metric. The separation problem maximizes their difference over the uncertainty support. If its optimal value exceeds the corresponding master variable η_n by more than the prescribed tolerance, the resulting adverse realization generates a violated epigraph constraint that is added to the master problem.

Write the rows of W_{sh} as $w_q^\top$, $q = 1,\dots,m$, and introduce the nonnegative absolute-value epigraph variables s_q, whose constraints are specified in (4.7). For $\lambda^k > 0$, the negative objective coefficient

forces these variables to attain their absolute-value bounds at optimality. When $\lambda^k = 0$, the distance penalty vanishes and the formulation remains exact.

For fixed x^k, define support-based surplus bounds for each finished product:

$$\overline{S}_{it}^{k} = I_{i0} + \sum_{\tau=1}^{t} \overline{y}_{i\tau} x_{i\tau}^{k} - \sum_{\tau=1}^{t} \underline{d}_{i\tau} - \sum_{\tau=1}^{t} \sum_{j \in C(i)} a_{ij} x_{j\tau}^{k},$$

$$\underline{S}_{it}^{k} = I_{i0} + \sum_{\tau=1}^{t} \underline{y}_{i\tau} x_{i\tau}^{k} - \sum_{\tau=1}^{t} \overline{d}_{i\tau} - \sum_{\tau=1}^{t} \sum_{j \in C(i)} a_{ij} x_{j\tau}^{k}.$$

Partition the finished-product positions into

$$\mathcal{J}_{+}^{k} = \{(i,t) \in \mathcal{F} \times \mathcal{T} : \underline{S}_{it}^{k} \geq 0\},$$

$$\mathcal{J}_{-}^{k} = \{(i,t) \in \mathcal{F} \times \mathcal{T} : \overline{S}_{it}^{k} \leq 0\},$$

and

$$\mathcal{J}_{\pm}^{k} = (\mathcal{F} \times \mathcal{T}) \setminus (\mathcal{J}_{+}^{k} \cup \mathcal{J}_{-}^{k}).$$

The holding and backlog branches are known in advance on $\mathcal{J}_{+}^{k}$ and $\mathcal{J}_{-}^{k}$; a sign binary is needed only on $\mathcal{J}_{\pm}^{k}$. For each $(i,t) \in \mathcal{J}_{\pm}^{k}$, introduce $u_{it}^{+}, u_{it}^{-} \geq 0$ and $b_{it} \in \{0,1\}$. Their exact linear constraints are included in the following separation MILP.

The exact separation MILP is

$$\begin{aligned}
\theta_n(x^k, \lambda^k) = \max \quad & \sum_{i \in \mathcal{I}^{\mathrm{int}}} \sum_{t \in \mathcal{T}} h_i S_{it}^{k}(D,Y) \\
& + \sum_{(i,t) \in \mathcal{J}_{+}^{k}} h_i S_{it}^{k}(D,Y) + \sum_{(i,t) \in \mathcal{J}_{-}^{k}} (-p_i) S_{it}^{k}(D,Y) \\
& + \sum_{(i,t) \in \mathcal{J}_{\pm}^{k}} (h_i u_{it}^{+} + p_i u_{it}^{-}) - \lambda^k \sum_{q=1}^{m} s_q \\
\text{s.t.} \quad & u_{it}^{+} - u_{it}^{-} = S_{it}^{k}(D,Y), && (i,t) \in \mathcal{J}_{\pm}^{k}, \\
& 0 \leq u_{it}^{+} \leq \overline{S}_{it}^{k} b_{it}, \quad 0 \leq u_{it}^{-} \leq -\underline{S}_{it}^{k}(1 - b_{it}), && (i,t) \in \mathcal{J}_{\pm}^{k}, \\
& s_q \geq w_q^{\top}(\xi - \widehat{\xi}^{n}), \quad s_q \geq -w_q^{\top}(\xi - \widehat{\xi}^{n}), && q = 1, \ldots, m, \\
& \xi \in \Xi, \quad s_q \geq 0, \quad b_{it} \in \{0,1\}, && (i,t) \in \mathcal{J}_{\pm}^{k}.
\end{aligned} \tag{4.7}$$

Here

$$S_{it}^{k}(D,Y) = I_{i0} + \sum_{\tau=1}^{t} Y_{i\tau} x_{i\tau}^{k} - \sum_{\tau=1}^{t} D_{i\tau} - \sum_{\tau=1}^{t} \sum_{j \in C(i)} a_{ij} x_{j\tau}^{k}.$$

The transport term remains linearizable without integer variables because each transformed deviation

$$w_q^{\top}(\xi - \widehat{\xi}^{n}) = \sum_{r=1}^{m} (W_{\mathrm{sh}})_{qr} (\xi_r - \widehat{\xi}_r^{n})$$

is affine in ξ. The full matrix can make these constraints denser than under a diagonal or block-sparse transformation, but it does not itself add integer variables.

Proposition 2. *At iteration k, the exact separation problem requires exactly one surplus-sign binary for each sign-indefinite finished-product position in $\mathcal{J}_{\pm}^{k}$, and none for internal items or sign-definite finished-product positions. Hence*

$$|\mathcal{J}_{\pm}^{k}| \le |\mathcal{F}||\mathcal{T}| < |\mathcal{I}||\mathcal{T}|$$

whenever the BOM contains at least one internal item. Relative to the original all-item soft-shortage formulation, material feasibility therefore yields the structural upper bound $|\mathcal{F}||\mathcal{T}|$, with additional exact presolve reduction to $|\mathcal{J}_{\pm}^{k}|$.

Proof. Internal-item surplus is nonnegative over Ξ by Proposition 1. For finished products, the support bounds determine the sign whenever $\underline{S}_{it}^{k} \ge 0$ or $\overline{S}_{it}^{k} \le 0$. Only when $\underline{S}_{it}^{k} < 0 < \overline{S}_{it}^{k}$ can both branches of the holding/backlog maximum be active, which requires one binary selector. □

4.5. Theoretical properties and the cutting-plane algorithm

We next establish the theoretical properties underlying the master–separation framework. These results distinguish standard Wasserstein duality from the BOM-specific material-feasibility result in Proposition 1 and the binary-reduction result in Proposition 2.

Assumption 1. *The support Ξ is nonempty and compact, and the ground metric d_{DI} is continuous on $\Xi \times \Xi$.*

Assumption 1 holds for the box support defined in Section 2. For every fixed material-feasible plan x, each internal surplus term is affine in ξ, while finished-product positive and negative parts are continuous piecewise-linear functions. Thus $\Phi_{\mathrm{MF}}(x, \xi)$ is continuous and bounded on Ξ.

Proposition 3. *Under Assumption 1, for every fixed x satisfying material feasibility,*

$$Q_{\varepsilon}^{\mathrm{MF}}(x) = \inf_{\lambda \ge 0} \left\{ \lambda\varepsilon + \frac{1}{N} \sum_{n=1}^{N} \sup_{\xi \in \Xi} \left[\Phi_{\mathrm{MF}}(x, \xi) - \lambda d_{\mathrm{DI}}(\xi, \widehat{\xi}^{n}) \right] \right\}.$$

Proof. For fixed material-feasible x, the function $\Phi_{\mathrm{MF}}(x, \cdot)$ is continuous and bounded on compact Ξ. The ground metric d_{DI} is continuous because it is the ℓ_1-norm of a fixed linear transformation. The standard strong-duality result for Wasserstein distributionally robust expectations therefore applies [30, 31], yielding exactly the stated representation. The proposition is used as a standard analytical tool; the paper does not claim the duality identity itself as a new result. □

Proposition 4. *Let $Z^{\star}$ be the optimal value of the exact semi-infinite formulation* (4.4)*, and let Z^{k} be the optimal value of the finite master problem* (4.5) *at iteration k. Then*

$$Z^{k} \le Z^{\star}.$$

If one or more violated adverse realizations are added to the sets $\mathcal{V}_n$, then

$$Z^{k+1} \ge Z^{k}.$$

The proof is provided in the separate supplementary Appendix A.

For an incumbent solution $(x^k, z^k, \lambda^k, \eta^k)$, define the sample-wise violations

$$v_n^k = \theta_n(x^k, \lambda^k) - \eta_n^k, \qquad n = 1, \ldots, N,$$

and the maximum violation

$$\Delta^k = \max_{n=1,\ldots,N} v_n^k.$$

If $\Delta^k \leq 0$, all semi-infinite epigraph constraints are satisfied. If $\Delta^k > 0$, a maximizer of at least one separation problem provides a valid violated constraint for the master.

Theorem 1. *Suppose Assumption 1 holds, and the finite master problem and all sample-wise separation problems are solved to global optimality at iteration k. If the cutting-plane method terminates with*

$$\theta_n(x^k, \lambda^k) \leq \eta_n^k + \delta, \qquad n = 1, \ldots, N,$$

then the returned first-stage solution (x^k, z^k) has an objective value within δ of the global optimum of the exact formulation (4.4)*.*

The proof is provided in the separate supplementary Appendix A.

Remark 1. *The sample-wise stopping rule is sufficient but not necessary. The weaker aggregate condition*

$$\sum_{n=1}^{N} \max\{v_n^k, 0\} \leq N\delta$$

leads to the same δ-optimality bound after averaging. We retain the sample-wise rule because it gives a transparent cut-generation test for each empirical observation.

Remark 2. *Material feasibility is not delegated to the separation oracle. Constraints* (4.2) *are enforced in every master iteration and are valid for the entire support Ξ. Separation is used only for the worst-case expectation. This division avoids generating plans that temporarily violate component availability and also explains why internal-item sign binaries disappear from the oracle.*

Algorithm 1. Master–separation framework for MF-JY-DRO.

1. **Initialization.** Set $\mathcal{V}_n = \{0\}$ and $\widetilde{\xi}^{n0} = \widehat{\xi}^n$ for every n. Choose $\delta > 0$ and set $k = 0$.
2. **Master.** Solve (4.5) to obtain $(x^k, z^k, \lambda^k, \eta^k)$. The master already enforces robust material feasibility through (4.2).
3. **Separation.** For each $n = 1, \ldots, N$, solve (4.7); let $\xi^{n,k}$ and $\theta_n(x^k, \lambda^k)$ be the maximizer and objective value, respectively. The N oracle problems can be solved independently.
4. **Termination.** If $\theta_n(x^k, \lambda^k) \leq \eta_n^k + \delta$ for all n, return (x^k, z^k).
5. **Cut generation.** For each violated sample n, create a new index ℓ, set $\widetilde{\xi}^{n\ell} = \xi^{n,k}$, add ℓ to $\mathcal{V}_n$, increment k, and return to Step 2.

When the finite master problems are solved globally, their optimal objective values form a nondecreasing sequence of lower bounds on the exact DRO optimum. Globally solved separation problems provide exact tests of the omitted epigraph constraints, and Theorem 1 establishes the resulting optimality guarantee under these conditions. In numerical implementation, however, the master and separation MILPs may terminate before global optimality is established. We therefore distinguish the exact analytical reformulation from its time-limited numerical solution. Valid solver lower bounds and feasible first-stage solutions evaluated using separation upper bounds provide an a posteriori optimality-gap certificate, even when individual separation problems remain unresolved.

4.6. Computational implications

The master contains the robust material-feasibility inequalities (4.2) explicitly, so it never proposes a plan that relies on negative internal material. The separation problems remain independent across empirical samples and are therefore parallelizable. Relative to a formulation that treats shortage at every item as a positive/negative surplus choice, Proposition 2 removes sign binaries for all internal items and also presolves sign-definite finished-product positions. The full covariance-informed matrix W_{sh} changes the continuous absolute-value constraints in the distance term but does not introduce additional integer variables. The binary reduction is characterized structurally by Proposition 2; the computational study reports solve time, iterations, cuts, and certified optimality gaps to assess the behavior of the resulting master–separation procedure.

5. Computational experiments

The computational study is designed to answer two questions raised by the model development. First, does demand–yield cross-dependence change the planning decisions once the marginal demand and yield samples are held fixed? Second, can the material-feasible master–separation algorithm provide useful optimality certificates for different BOM structures under a common computational budget? The study therefore emphasizes controlled dependence identification and certified computation rather than an environment constructed to favor a particular joint-risk model.

5.1. Experimental design, calibration, and benchmarks

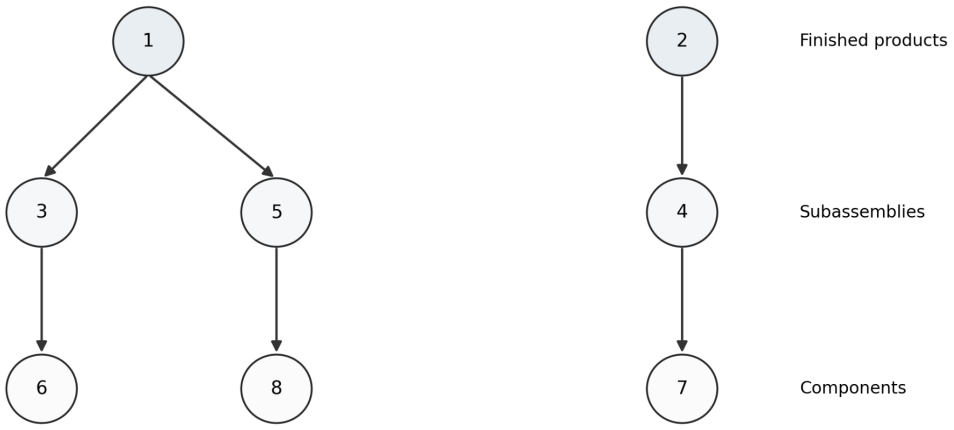


Figure 1. Topology of the eight-item BOM used in the main computational experiments. Nodes 1 and 2 are finished products, nodes 3–5 are subassemblies, and nodes 6–8 are lower-level components. The numerical BOM coefficients for the ten replications are reported in Table 3.

The main experiment uses an eight-item, six-period tree BOM with two finished products. Each replication contains $N = 40$ paired training observations and an independent test sample of $M = 600$ observations. Ten independent replications are generated, and all competing models within a replication

use the same marginal demand and yield samples and the same test pool. This common-random-number design makes within-replication model differences directly comparable.

BOM structure and experimental parameters. The eight-item tree BOM contains two finished products $\mathcal{F} = \{1, 2\}$, three subassemblies $\{3, 4, 5\}$, and three lower-level components $\{6, 7, 8\}$. Its parent–component relationships are $(1, 3)$, $(2, 4)$, $(1, 5)$, $(3, 6)$, $(4, 7)$, and $(5, 8)$, where (j, i) indicates that component i is required by parent item j. The corresponding BOM topology is illustrated in Figure 1. The BOM topology remains fixed across all ten independent replications, while the six nonzero coefficients are generated by successive integer draws from $\{1, 2\}$, using MATLAB's `rng(100+r,'twister')` for replication $r = 0, \ldots, 9$. Table 3 reports the complete set of nonzero BOM coefficients for all ten replications. All remaining BOM coefficients are zero.

Table 3. Complete nonzero BOM coefficients for the ten independent replications.

Replication	a_{31}	a_{42}	a_{51}	a_{63}	a_{74}	a_{85}
0	2	1	1	2	1	1
1	2	2	1	1	2	2
2	2	2	1	2	2	2
3	1	1	1	2	2	1
4	1	1	2	1	1	2
5	1	1	2	1	2	1
6	1	2	1	2	2	2
7	1	2	2	1	1	1
8	1	1	2	1	2	2
9	2	1	2	1	2	2

The item-level cost parameters, initial inventories, and production limits are summarized in Table 4. The setup, production, and holding cost parameters are fixed across periods and replications. For every item and period, the production or procurement upper bound is $M_{it} = 600$. The resource-constraint parameters are $r_i = 1$ and $q_i = 0$ for all items, with $B_t = 1500$ in every period. Consequently, the aggregate resource constraint reduces to $\sum_{i \in \mathcal{I}} x_{it} \leq 1500$.

Table 4. Cost, inventory, and production parameters for the eight-item BOM.

Items	K_i	c_i	h_i	p_i	I_{i0}	M_{it}
1–2	110	8.5	1.5	45	0	600
3–5	49	3.0	0.4	0	11	600
6–8	53	3.5	0.5	0	14	600

Demand and yield sample generation. The nominal demand of finished product i in period t is defined as

$$\mu_{it} = (16 + 4(i-1))\left[1 + 0.12 \sin\left(\frac{\pi(t-1)}{5}\right) + 0.03(t-1)\right], \quad i \in \mathcal{F}.$$

External demand for internal items is zero. The demand support is $[0.65\mu_{it}, 1.45\mu_{it}]$ for finished products and $\{0\}$ for internal items. The usable-yield support is $[0.82, 1]$ for every item and period.

For each sample trajectory n, a common Gaussian factor is generated according to

$$G_{n1} \sim \mathcal{N}(0,1), \qquad G_{nt} = 0.45G_{n,t-1} + \sqrt{1-0.45^2}\, Z_{nt}^{G}, \quad t = 2, \ldots, 6,$$

where the innovations Z_{nt}^{G} are independent standard normal random variables.

The demand and yield observations are generated by

$$\begin{aligned} D_{it}^{n} &= \mathrm{clip}_{[0.65\mu_{it},\, 1.45\mu_{it}]}\left(\mu_{it}[1 + 0.18G_{nt} + 0.08Z_{it}^{D,n}]\right), && i \in \mathcal{F}, \\ Y_{it}^{n} &= \mathrm{clip}_{[0.82,\,1]}\left(0.94 - 0.055\left[0.65G_{nt} + \sqrt{1-0.65^2}Z_{it}^{Y,n}\right]\right), && i \in \mathcal{I}, \end{aligned}$$

where $Z_{it}^{D,n}$ and $Z_{it}^{Y,n}$ are mutually independent standard normal innovations, independent of the common factor, and $\mathrm{clip}_{[a,b]}(v) = \min\{b, \max\{a, v\}\}$.

The common Gaussian factor introduces temporal and cross-item dependence, together with an initial demand–yield dependence structure. The resulting observations are subsequently re-paired to construct the prescribed stress dependence levels while preserving the empirical marginal sample sets.

For replication $r = 0, \ldots, 9$, the random seeds are $10{,}000 + r$ for the training sample pool, $20{,}000 + r$ for the independent test pool, $1000 + r$ for training-sample re-pairing, $2000 + r$ for test-sample re-pairing, and $5000 + r$ for bootstrap calibration. Each replication contains $N = 40$ training observations and $M = 600$ test observations. The Wasserstein radii are calibrated using 100 bootstrap resamples and the 90% empirical quantile.

Dependence is varied without changing either empirical marginal multiset. For each training pool, demand observations and complete yield trajectories are re-paired using the stress scores

$$g_D^n = \frac{1}{|\mathcal{F}||\mathcal{T}|}\sum_{i\in\mathcal{F}}\sum_{t\in\mathcal{T}}\frac{\widehat{D}_{it}^{n} - \bar{D}_{it}}{s_{it}^{D}}, \qquad g_Y^n = -\frac{1}{|\mathcal{I}||\mathcal{T}|}\sum_{i\in\mathcal{I}}\sum_{t\in\mathcal{T}}\frac{\widehat{Y}_{it}^{n} - \bar{Y}_{it}}{s_{it}^{Y}}.$$

The realized stress dependence is

$$\rho_{\mathrm{stress}} = \mathrm{Corr}\left(g_D^n, g_Y^{\pi(n)}\right), \tag{5.1}$$

where positive values mean that high-demand observations tend to be paired with low-yield observations. The target values are -0.8, 0, and 0.8. Across the ten independent replications, the mean realized test correlations are -0.796, -0.001, and 0.797, respectively. The realized correlations are calculated separately for the training and test samples. Owing to the smaller training sample size, the training correlations may deviate from their target values, whereas the mean test correlations closely match the prescribed dependence levels. Thus, the experiment spans favorable, approximately neutral, and adverse demand–yield dependence while preserving the marginal samples. The target stress correlations $\rho_0 \in \{-0.8, 0, 0.8\}$ are deliberately selected for controlled sensitivity analysis; they are not estimates obtained from industrial observations. Because the yield stress score is defined using negative standardized yield, a positive stress correlation corresponds to high demand being paired with low usable yield, whereas a negative value represents a more favorable demand–yield pairing. The re-pairing procedure preserves the original marginal sample pools while changing their cross-source association. These scenarios are intended to reveal the operational effects of dependence rather than to represent empirically calibrated industry-wide correlation levels.

The permutation π is constructed by rank-based matching. For a sample pool of size L, let R_n be the ascending rank of the demand stress score g_D^n, and define $q_n = \Phi^{-1}((R_n - 0.5)/L)$, where Φ denotes the standard normal cumulative distribution function. An independent Gaussian vector u is orthogonalized against q by replacing it with $u - (u^\top q)/(q^\top q)q$. After standardizing both vectors to the unit sample standard deviation, define $z_n = \rho_0 q_n + \sqrt{1-\rho_0^2}\, u_n$. The ascending ranks of z_n are matched to those of the yield stress scores, assigning entire yield trajectories to demand observations. This preserves the marginal samples and the within-trajectory dependence structure.

Every DRO benchmark receives a model-specific Wasserstein radius. For model m, $B = 100$ bootstrap resamples are drawn from its empirical sample, and

$$\varepsilon_m = \mathrm{Quantile}_{0.90}\left\{ W_{d_m}\left(\widehat{\mathbb{P}}_N^{(m,b)}, \widehat{\mathbb{P}}_N^{(m)}\right) : b = 1, \ldots, B \right\}. \tag{5.2}$$

The same 90% bootstrap quantile rule is therefore used across models, but the numerical radii are allowed to differ because the uncertainty spaces and transport metrics differ. The average calibrated D-DRO radius is 4.722, while the Y-DRO radius ranges from 17.044 to 17.067 across the three dependence cases. The corresponding Block-JDRO and Full-JDRO radii are reported in Table 6. For each replication and dependence level, Block-JDRO and Full-JDRO use the same Ledoit–Wolf shrinkage intensity; they differ only in whether the demand–yield cross-covariance blocks are removed or retained in the transportation geometry.

Six models are compared. DET replaces uncertain quantities by their training-sample means. SAA minimizes empirical average cost. D-DRO treats only demand as distributionally ambiguous and fixes the yield at its sample mean; Y-DRO does the converse. Block-JDRO uses the joint demand–yield vector but removes the demand–yield cross-covariance blocks from the shrinkage covariance before constructing the ground metric. Full-JDRO retains those cross blocks and is the proposed dependence-informed specification. All six models use the same hard material-feasible set $\mathcal{X}_{\mathrm{MF}}$, so differences in finished-product service cannot be generated by allowing infeasible internal-component shortages.

For a test sample $\{\xi^q\}_{q=1}^M$, the reported total cost is

$$\mathrm{TC}(x,z) = C(x,z) + \frac{1}{M}\sum_{q=1}^{M} \Phi_{\mathrm{MF}}(x, \xi^q).$$

Finished-product fill rate is computed from newly incurred backlog increments rather than by summing cumulative backlog repeatedly. If $B_{it}^q = S_{it}^-(x,\xi^q)$ and $B_{i0}^q = 0$, define

$$\Delta B_{it}^q = \max\{B_{it}^q - B_{i,t-1}^q, 0\}, \qquad \mathrm{FR} = 1 - \frac{\sum_{q,i\in\mathcal{F},t} \Delta B_{it}^q}{\sum_{q,i\in\mathcal{F},t} D_{it}^q}.$$

The 95% upper-tail shortage measure is the mean of the largest 5% of scenario shortage costs. Total cost and upper-tail shortage are treated as the primary outcomes; average shortage cost and fill rate are reported as supporting service measures.

All optimization models are implemented in MATLAB. The DRO models are solved with `intlinprog` using the master–separation procedure of Section 4 and a common 60-second total time budget in the main experiment. The target certified relative gap is 8%, but the time limit takes precedence; hence, the achieved a posteriori certificate is reported rather than treating the target as an

unconditional guarantee. The experiments were run on an AMD Ryzen 5 5600 processor with 32 GB RAM. For each DRO run, we report the best available global lower bound L and a valid upper bound U obtained from a feasible first-stage solution. The reported relative optimality gap is

$$\text{Gap} = \frac{\max\{0, U - L\}}{\max\{1, |U|\}}.$$

The upper-bound calculation uses certified separation bounds when available and conservative support-based bounds when a separation MILP remains unresolved within the computational budget. Consequently, the reported gap quantifies the remaining numerical optimality uncertainty and does not imply that every separation problem has been solved to global optimality. In particular, a positive finite-time gap must not be interpreted as an exact numerical optimum.

5.2. Out-of-sample performance across dependence regimes

Table 5 reports the out-of-sample performance over ten independent replications, with sample standard deviations shown in parentheses. The results reveal a clear trade-off between total planning cost and shortage protection. DET and SAA achieve relatively low total costs but incur substantially higher shortage losses than the joint DRO models. For example, under adverse dependence ($\rho_{\text{stress}} = 0.8$), the mean shortage costs of DET and SAA are 1708.9 and 1480.9, respectively, compared with 8.5 for Block-JDRO and 15.2 for Full-JDRO. The corresponding mean fill rates increase from 0.8710 and 0.8882 to 0.9992 and 0.9987, respectively. These results demonstrate that joint demand–yield robustness substantially improves service performance at the expense of higher total costs.

Table 5. Out-of-sample performance under controlled demand–yield dependence. Entries are mean (standard deviation) over ten independent replications.

ρ_{stress}	Model	Total cost	Shortage cost	Fill rate	95% tail shortage
−0.8	DET	10487.2 (1297.4)	1536.2 (380.1)	0.8804 (0.0263)	4813.8 (611.0)
−0.8	SAA	10323.3 (1292.3)	1272.5 (281.9)	0.8986 (0.0197)	4241.4 (551.4)
−0.8	D-DRO	10703.4 (1290.9)	499.5 (183.9)	0.9575 (0.0156)	2275.9 (585.8)
−0.8	Y-DRO	10527.2 (1329.1)	757.2 (206.0)	0.9359 (0.0168)	3319.2 (589.4)
−0.8	Block-JDRO	13909.9 (1633.3)	0.3 (0.3)	1.0000 (0.0000)	5.5 (6.0)
−0.8	Full-JDRO	12340.3 (1549.2)	76.0 (81.3)	0.9932 (0.0073)	567.1 (498.2)
0	DET	10574.5 (1291.0)	1620.7 (370.6)	0.8757 (0.0248)	5521.8 (589.6)
0	SAA	10404.6 (1281.4)	1382.2 (219.1)	0.8928 (0.0149)	4961.2 (493.3)
0	D-DRO	10770.7 (1298.4)	564.6 (191.1)	0.9531 (0.0157)	2748.0 (661.4)
0	Y-DRO	10553.3 (1287.7)	956.0 (231.2)	0.9218 (0.0170)	4308.0 (693.4)
0	Block-JDRO	13728.6 (1488.8)	12.3 (32.1)	0.9989 (0.0028)	134.1 (292.0)
0	Full-JDRO	13508.8 (1308.6)	15.8 (23.5)	0.9986 (0.0021)	207.4 (277.1)
0.8	DET	10665.7 (1283.5)	1708.9 (359.7)	0.8710 (0.0234)	6297.6 (578.3)
0.8	SAA	10496.6 (1277.6)	1480.9 (202.9)	0.8882 (0.0135)	5711.1 (486.2)
0.8	D-DRO	10845.7 (1300.5)	637.2 (193.5)	0.9483 (0.0155)	3386.7 (721.6)
0.8	Y-DRO	10646.5 (1297.1)	1075.6 (206.4)	0.9136 (0.0143)	5154.6 (647.2)
0.8	Block-JDRO	13949.4 (1583.0)	8.5 (10.7)	0.9992 (0.0010)	155.5 (169.6)
0.8	Full-JDRO	13876.6 (1661.8)	15.2 (35.2)	0.9987 (0.0030)	185.5 (334.2)

The second pattern concerns the effect of cross-source dependence on the performance of the two joint DRO models. Block-JDRO retains within-demand and within-yield covariance information but removes the demand–yield cross-covariance blocks, whereas Full-JDRO incorporates the complete covariance structure into the transportation metric. As shown in Figure 2, the mean total cost of Block-JDRO remains relatively stable across the three dependence regimes, ranging from 13,728.6

to 13,949.4. In contrast, the mean total cost of Full-JDRO increases from 12,340.3 under favorable dependence to 13,876.6 under adverse dependence. The largest difference between the two models occurs under favorable dependence, where Full-JDRO achieves a mean cost reduction of 1569.6. Under approximately neutral and adverse dependence, the mean cost differences are substantially smaller. These findings indicate that the economic impact of cross-source covariance information depends on the underlying demand–yield relationship.

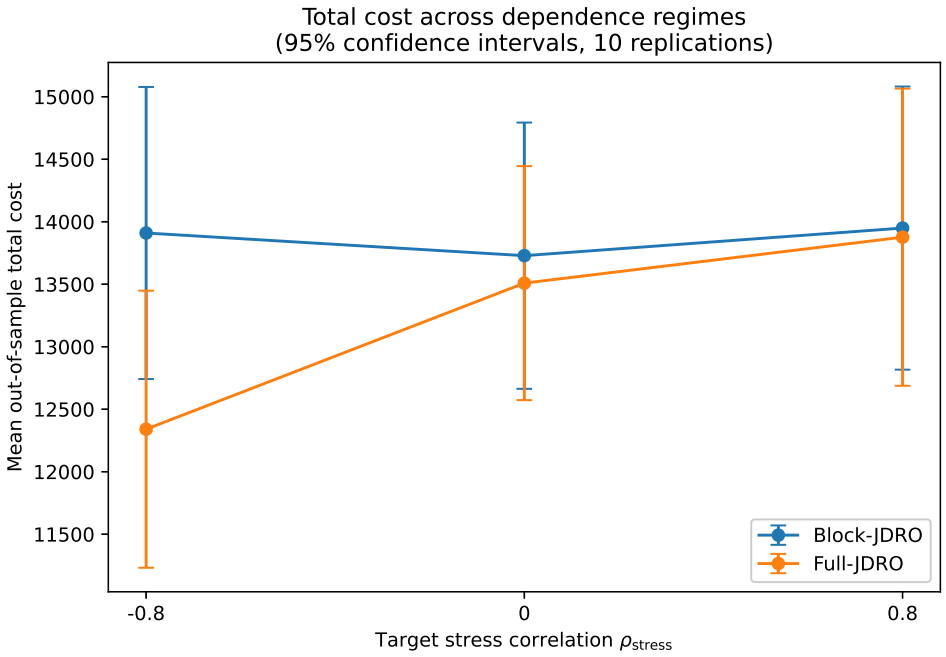


Figure 2. Mean out-of-sample total cost for Block-JDRO and Full-JDRO across demand–yield dependence regimes. Error bars indicate 95% confidence intervals based on ten independent replications.

The upper-tail shortage results in Table 5 further confirm the cost–service trade-off between Full-JDRO and Block-JDRO, particularly under favorable demand–yield dependence.

5.3. Paired dependence ablation and solver certificates

Because Block-JDRO and Full-JDRO share the same training and test samples within each replication, their performance differences are evaluated using paired comparisons. Table 6 reports the mean differences and corresponding 95% confidence intervals over ten independent replications. Under favorable dependence ($\rho_{\text{stress}} = -0.8$), Full-JDRO reduces the mean total cost by 1569.6, with a 95% confidence interval of $[-2041.1, -1098.1]$. This cost reduction is accompanied by an increase of 561.6 in upper-tail shortage cost, with a confidence interval of $[207.2, 916.1]$. The corresponding mean fill-rate difference is -0.0068, indicating a measurable cost–service trade-off. Under approximately neutral and adverse dependence, the confidence intervals for both total cost and shortage-risk differences contain zero. These results indicate that the economic value of cross-source dependence information is regime-dependent, with the most pronounced cost–risk trade-off observed under favorable dependence.

Table 6. Paired comparison between Full-JDRO and Block-JDRO over ten independent replications. All performance differences are calculated as Full-JDRO minus Block-JDRO. The 95% confidence intervals are based on paired Student's t statistics with nine degrees of freedom.

Panel A: Dependence levels and calibrated Wasserstein radii

Target ρ	Mean test ρ	$\bar{\varepsilon}_{\text{Block}}$	$\bar{\varepsilon}_{\text{Full}}$
−0.8	−0.796	23.57	22.33
0	−0.001	23.85	22.70
0.8	0.797	23.62	22.37

Panel B: Paired out-of-sample performance differences

Target ρ	Performance measure	Mean difference	95% confidence interval
−0.8	Total cost	−1569.6	[−2041.1, −1098.1]
	Shortage cost	75.7	[17.7, 133.8]
	Tail shortage (95%)	561.6	[207.2, 916.1]
	Fill rate	−0.0068	[−0.0120, −0.0016]
0	Total cost	−219.8	[−467.1, 27.5]
	Shortage cost	3.5	[−12.3, 19.4]
	Tail shortage (95%)	73.3	[−82.8, 229.4]
	Fill rate	−0.0003	[−0.0017, 0.0010]
0.8	Total cost	−72.8	[−449.5, 303.8]
	Shortage cost	6.6	[−11.4, 24.6]
	Tail shortage (95%)	30.0	[−102.0, 162.1]
	Fill rate	−0.0006	[−0.0021, 0.0010]

The calibrated Wasserstein radii in Table 6 differ systematically between the two joint DRO models. The average radii of Full-JDRO are smaller than those of Block-JDRO across all three dependence regimes, reflecting the influence of cross-source covariance information on the transportation geometry. Across the 120 DRO solves in the expanded main experiment, the mean relative optimality gaps are 4.64%, 8.99%, 7.83%, and 7.67% for D-DRO, Y-DRO, Block-JDRO, and Full-JDRO, respectively. The corresponding maximum gaps are 7.58%, 12.03%, 10.54%, and 11.91%, while the mean runtimes are 6.2, 53.3, 49.3, and 48.0 seconds. Several runs terminate at the computational time limit or with an uncertain separation-oracle status. Therefore, the reported numerical comparisons reflect the solutions obtained under the prescribed computational budget, with the associated optimality-gap and solver-status information retained for assessing solution quality.

Managerial implications. The dependence-informed metric does not provide a universally superior policy. Under favorable demand–yield dependence, Full-JDRO achieves a lower average total cost than Block-JDRO, but these savings are accompanied by higher upper-tail shortage losses. Managers with greater tolerance for occasional service disruptions may therefore favor the lower-cost policy, whereas applications facing substantial shortage consequences may place greater weight on tail-risk

protection. Under neutral and adverse dependence, the paired confidence intervals do not establish a consistent performance advantage for either joint model. Consequently, the choice of transportation geometry should reflect both observed dependence and the relative importance of cost efficiency and service reliability.

Table 7. Computational comparison of the unreduced and reduced separation formulations over 90 matched instances.

Performance measure	Unreduced	Reduced
Mean number of explicit sign binaries	48.00	4.07
Fixed internal-item sign binaries	36	0
Mean number of nonfixed sign binaries	12.00	4.07
Mean solution time (s)	0.2215	0.2080
Mean relative oracle gap	2.91×10^{-13}	1.70×10^{-13}
Mean branch-and-bound nodes	6.16	6.14

Computational effect of the separation reduction. To evaluate the structural reduction established in Proposition 2, we compare the reduced and unreduced separation formulations using 90 matched instances covering ten independent replications, three dependence regimes, and three training observations per regime. Both formulations use identical incumbent decisions, uncertainty samples, transportation metrics, and solver settings. As reported in Table 7, the mean number of explicit sign-selection binaries decreases from 48 to 4.07, representing a reduction of 91.5%. In the unreduced formulation, 36 internal-item binaries are fixed to their known signs to ensure numerical stability, leaving 12 nonfixed binaries. Relative to these 12 binaries, the reduced formulation achieves a further reduction of 66.1%. The two formulations produce consistent optimal objective values across all 90 instances, with a maximum absolute difference of 3.60×10^{-8}. The mean solution time decreases from 0.2215 to 0.2080 seconds, while the average branch-and-bound node counts remain nearly identical. These findings demonstrate the substantial reduction in mixed-integer formulation size achieved by exploiting BOM material feasibility, with a modest runtime improvement on the small-scale test instances.

5.4. BOM scalability and material-feasibility verification

A short scalability study evaluates the two joint models on tree, assembly, and shared-component BOMs under adverse dependence. The training sample remains $N = 40$ and the target stress correlation is 0.8. Because this experiment is intended to assess the certified algorithm rather than estimate fine out-of-sample performance differences, it uses one replication, $M = 300$ test scenarios, $B = 60$ bootstrap resamples, and a common 60-second solve limit. Table 8 reports the resulting certificates and cutting-plane statistics.

Table 8. Certified computational performance across BOM structures under adverse dependence.

Instance	Items	Periods	Model	ε	Gap (%)	CPU (s)	Iter.	Cuts
Tree-8	8	6	Block-JDRO	23.44	7.97	33.0	4	96
Tree-8	8	6	Full-JDRO	22.30	9.31	17.4	3	72
Assembly-12	12	8	Block-JDRO	38.67	3.70	29.5	4	96
Assembly-12	12	8	Full-JDRO	35.51	4.66	34.2	4	96
Shared-14	14	8	Block-JDRO	43.37	0.19	17.0	4	96
Shared-14	14	8	Full-JDRO	40.13	0.50	19.0	4	96

The algorithm returns an explicit certificate within the one-minute budget for every tested BOM topology. The Tree-8 cases terminate with gaps of 7.97% and 9.31% for Block-JDRO and Full-JDRO, respectively. The Assembly-12 gaps are 3.70% and 4.66%, while the Shared-14 cases attain gaps below 0.6%. All cases require only three or four master iterations and at most 96 accumulated cuts. These results do not establish monotone computational scaling—instance difficulty depends on both topology and the incumbent sign pattern—but they show that the certified master–separation procedure remains operational on the three representative BOM structures considered here.

Table 9. Extended scalability results for Full-JDRO under different BOM configurations and computational budgets.

Instance	I	T	Budget (s)	LB	UB	Gap (%)	CPU (s)	Iter.	Cuts
Shared-20	20	8	120	89482	93970	4.78	37.24	4	64
Shared-30	30	8	120	90019	94359	4.60	110.94	4	64
Shared-40	40	8	120	77944	94902	17.87	120.51	3	32
Shared-40	40	8	240	90517	94902	4.62	206.23	4	64
Shared-20	20	12	120	201480	211930	4.93	120.24	4	48
Shared-30	30	12	180	178550	212560	16.00	180.09	3	32
Tree-30	30	8	120	66720	85423	21.89	120.07	3	32
Assembly-30	30	8	120	88980	93317	4.65	106.41	4	64

Table 9 extends the Full-JDRO evaluation to 20–40 items and planning horizons of 8–12 periods. The computational difficulty increases with instance size, horizon length, and BOM characteristics. For Shared-40 with eight periods, extending the computational budget from 120 to 240 seconds reduces the certified optimality gap from 17.87% to 4.62%. The 30-item, 12-period instance retains a 16.00% gap after approximately 180 seconds, illustrating the difficulty of longer planning horizons. When separation problems remain unresolved, conservative support-based bounds preserve valid optimality certificates. From a managerial perspective, computation time is therefore a planning resource: larger instances may justify longer solution budgets, while hard material feasibility remains enforced even when global optimality has not been established.

Finally, hard material feasibility is numerically satisfied throughout the study. Across the entire main experiment, the maximum recorded internal-component deficit is 2.43×10^{-14}; in the scalability experiment, it is below 9.86×10^{-15}. These values are at numerical precision and confirm that the reported plans do not obtain their finished-product service performance by allowing negative internal component balances. This implementation check is important because the role of material feasibility is structural: internal shortages are ruled out by the feasible set, while the reported shortage and fill-rate measures refer only to finished products.

Overall, the experiments support a narrower conclusion than uniform dominance. Model-specific bootstrap calibration removes the arbitrary common-radius comparison; fixed-marginal re-pairing

isolates cross-source dependence; and the block-versus-full ablation shows that retaining demand–yield cross-covariance changes the cost–risk trade-off systematically as the dependence moves from favorable to adverse. The computational certificates and machine-precision internal deficits provide complementary evidence that these comparisons are obtained under hard internal material feasibility rather than through a soft component-shortage mechanism.

5.5. Sensitivity to the Wasserstein radius and uncertainty-support width

We investigate the sensitivity of Full-JDRO and Block-JDRO to two distinct sources of conservatism: the Wasserstein ambiguity radius and the physical uncertainty-support width. The experiments use the Tree-8 instance with six planning periods, 40 training observations, and 300 out-of-sample scenarios under a target stress correlation of 0.8. Model-specific Wasserstein radii are calibrated using 60 bootstrap resamples at a 90% confidence level. All sensitivity experiments employ a common 120-second computational budget, and the resulting finite-time optimality gaps are reported.

Table 10. Sensitivity to the Wasserstein radius on the Tree-8 instance.

Model	Multiplier	ε	Test cost	Fill (%)	Gap (%)	CPU (s)
Full	0.50	11.15	11040.58	98.03	6.03	33.49
Full	0.75	16.73	12924.38	99.94	5.50	62.43
Full	1.00	22.30	12868.35	99.92	7.55	31.24
Full	1.25	27.88	13197.28	99.97	7.05	45.17
Full	1.50	33.45	13322.49	99.96	6.31	43.54
Block	0.50	11.72	11000.67	97.82	7.78	28.51
Block	0.75	17.58	12707.81	99.91	6.22	54.30
Block	1.00	23.44	12812.88	99.93	7.18	52.44
Block	1.25	29.31	13135.04	99.95	7.26	42.35
Block	1.50	35.17	13177.72	99.95	6.85	41.91

We first vary the Wasserstein radius by multipliers of 0.50, 0.75, 1.00, 1.25, and 1.50 relative to each model's independently calibrated baseline. As shown in Table 10, increasing the radius generally raises out-of-sample operating costs while improving service protection. For Full-JDRO, the average out-of-sample cost increases from 11040.58 to 13322.49 as the radius multiplier increases from 0.50 to 1.50, whereas the fill rate increases from 98.03% to 99.96%. Most of the service improvement occurs between multipliers 0.50 and 0.75, with diminishing additional benefits at larger radii. The out-of-sample performance is not strictly monotonic, which does not contradict the monotonicity of the exact robust optimal value with respect to the ambiguity radius. These findings illustrate the economic trade-off between distributional robustness and operating cost.

We next examine sensitivity to the width of the operational uncertainty support by applying multipliers of 1.0, 1.2, and 1.4. Demand intervals are enlarged around their original midpoints, while yield intervals are enlarged subject to the physical bounds of zero and one. The training and test samples, transportation metrics, and calibrated Wasserstein radii are held fixed. Widening the support affects both the hard material-feasibility constraints and the domain of the worst-case expectation. Accordingly, this experiment measures their combined effect rather than isolating the cost of hard material feasibility alone.

Table 11. Sensitivity to uncertainty-support width on the Tree-8 instance.

Model	Width	LB	UB	Test cost	Holding	Fill (%)	Gap (%)
Full	1.0	13637.80	14751.15	12868.35	1699.44	99.92	7.55
Full	1.2	15033.34	16147.86	13046.91	1715.71	99.65	6.90
Full	1.4	16464.65	17611.77	15330.71	2144.04	99.99	6.51
Block	1.0	13719.05	14780.15	12812.88	1654.30	99.93	7.18
Block	1.2	14962.68	16162.70	13882.82	1944.78	99.97	7.42
Block	1.4	16332.97	17423.68	14735.42	2115.14	99.995	6.26

Table 11 shows that expanding the physical uncertainty support substantially increases the robust objective. For both models, the reported lower–upper bound intervals at successive width levels do not overlap, providing evidence that the optimal robust cost increases despite the remaining finite-time optimality gaps. At the 1.4 width multiplier, average out-of-sample cost increases by approximately 19.1% for Full-JDRO and 15.0% for Block-JDRO compared with their respective baseline values. The reported inventory-holding costs also increase. These results quantify the economic burden associated with protecting internal material availability against a wider set of demand and yield realizations. The fill rate does not increase monotonically: Full-JDRO exhibits a temporary service reduction at the 1.2 multiplier. Nevertheless, all tested solutions maintain internal material deficits within numerical precision. Thus, wider physical support imposes a greater economic burden without necessarily producing proportional improvements in realized finished-product service. From a managerial perspective, these results highlight the importance of selecting physically credible uncertainty bounds. Overly wide box supports may require substantial additional material protection while yielding limited improvements in realized service. Practitioners should therefore calibrate support bounds using available process knowledge and operational limits rather than enlarging them indiscriminately. More refined joint or budgeted uncertainty supports may reduce conservatism, but their tractability and feasibility counterparts would require separate analysis.

Taken together, the two sensitivity experiments distinguish distributional conservatism controlled by the Wasserstein radius from structural conservatism imposed by hard material-feasibility constraints. The numerical comparisons are based on one replication and are interpreted as illustrative sensitivity evidence rather than statistical proof of universal model superiority.

6. Conclusions

This paper develops a material-feasible and dependence-informed distributionally robust lot-sizing framework for multi-level BOM planning under joint demand and yield uncertainty. The principal modeling change is to distinguish physical internal material availability from finished-product service shortfall. Internal components and subassemblies must have nonnegative cumulative material surplus for every realization in the support, whereas finished-product demand may be backlogged. Under box support, these robust material-feasibility requirements have an exact finite linear counterpart obtained from the lower yield and upper demand bounds. The second modeling element is a covariance-informed Wasserstein ground metric constructed from standardized paired demand–yield observations. This construction lets cross-source dependence information affect transportation costs without claiming that

every candidate distribution preserves the empirical covariance. Standard Wasserstein duality provides the exact semi-infinite expectation representation; the paper's structural computational contribution comes from combining that representation with material feasibility. In particular, the sample-wise separation problem requires surplus-sign binaries only for finished-product positions whose support interval crosses zero, not for every item in the BOM. The computational results isolate these mechanisms under model-specific bootstrap calibration and fixed-marginal dependence perturbations. The full-versus-block metric ablation shows that cross-source dependence information changes the cost–risk trade-off rather than uniformly improving every performance measure: favorable dependence allows the full metric to reduce the block model's conservatism, whereas the results under neutral and adverse dependence show no statistically conclusive performance advantage for either joint model. Across the tested tree, assembly, and shared-component BOM structures, the master–separation procedure returns explicit optimality certificates under a common time budget, and internal material deficits remain at machine precision.

The present computational evaluation is based on reproducible synthetic instances rather than paired demand–yield records from an operating industrial facility. Consequently, the numerical results demonstrate the behavior of the proposed framework under controlled dependence scenarios but do not constitute empirical validation of the assumed correlation levels or of the achievable improvements in a specific industry. Validation using time-aligned industrial demand, production, and usable-yield records remains an important direction for future research. Several extensions remain natural. Incorporating positive deterministic lead times would require revised time-indexed material balances and explicit treatment of pipeline inventory. More challenging extensions include stochastic lead times and uncertain capacity bottlenecks, which may induce additional dependence between material availability and production decisions, and require new robust counterparts. Adjustable or rolling-horizon decisions could further reduce the conservatism of static material-feasibility protection. Finally, richer dependence descriptions could combine transportation geometry with explicit moment information when sufficiently large datasets become available.

Author contributions

Jiayu Shen: Conceptualization, methodology, formal analysis, investigation, and writing–original draft; Kai Zhu: Software (program debugging), visualization, and writing–review and editing. All authors have read and approved the final version of the manuscript.

Use of Generative-AI tools declaration

During the preparation and revision of this manuscript, the authors used ChatGPT (OpenAI) to assist with English language editing, manuscript polishing, LaTeX formatting, and checking the clarity and consistency of mathematical and computational descriptions. All AI-assisted suggestions were reviewed and revised as appropriate by the authors. The authors take full responsibility for the mathematical arguments, computational results, references, and overall content of the manuscript.

Acknowledgments

No specific funding was received for this study.

Data availability

The data used in the computational experiments are generated from the simulation settings described in this article. The generated datasets and computational codes are available from the corresponding author upon reasonable request.

Conflict of interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

References

1. J. Orlicky, *Material Requirements Planning: The New Way of Life in Production and Inventory Management*, McGraw-Hill, New York, 1975.
2. P. J. Billington, J. O. McClain, L. J. Thomas, Mathematical programming approaches to capacity-constrained MRP systems: Review, formulation and problem reduction, *Manag. Sci.*, **29** (1983), 1126–1141. https://doi.org/10.1287/mnsc.29.10.1126
3. H. M. Wagner, T. M. Whitin, Dynamic version of the economic lot size model, *Manag. Sci.*, **5** (1958), 89–96. https://doi.org/10.1287/mnsc.5.1.89
4. W. I. Zangwill, A backlogging model and a multi-echelon model of a dynamic economic lot size production system: A network approach, *Manag. Sci.*, **15** (1969), 506–527. https://doi.org/10.1287/mnsc.15.9.506
5. A. J. Clark, H. Scarf, Optimal policies for a multi-echelon inventory problem, *Manag. Sci.*, **6** (1960), 475–490. https://doi.org/10.1287/mnsc.6.4.475
6. H. Tempelmeier, S. Helber, A heuristic for dynamic multi-item multi-level capacitated lotsizing for general product structures, *Eur. J. Oper. Res.*, **75** (1994), 296–311. https://doi.org/10.1016/0377-2217(94)90076-0
7. A. Drexl, A. Kimms, Lot sizing and scheduling: Survey and extensions, *Eur. J. Oper. Res.*, **99** (1997), 221–235. https://doi.org/10.1016/S0377-2217(97)00030-1
8. H. Stadtler, Multilevel lot sizing with setup times and multiple constrained resources: Internally rolling schedules with lot-sizing windows, *Oper. Res.*, **51** (2003), 487–502. https://doi.org/10.1287/opre.51.3.487.14949
9. Y. Pochet, L. A. Wolsey, *Production Planning by Mixed Integer Programming*, Springer, New York, 2006. https://doi.org/10.1007/0-387-33477-7
10. S. Helber, F. Sahling, A fix-and-optimize approach for the multi-level capacitated lot sizing problem, *Int. J. Prod. Econ.*, **123** (2010), 247–256. https://doi.org/10.1016/j.ijpe.2009.08.022

11. B. Karimi, S. M. T. Fatemi Ghomi, J. M. Wilson, The capacitated lot sizing problem: A review of models and algorithms, *Omega*, **31** (2003), 365–378. https://doi.org/10.1016/S0305-0483(03)00059-8
12. R. Jans, Z. Degraeve, Modeling industrial lot sizing problems: A review, *Int. J. Prod. Res.*, **46** (2008), 1619–1643. https://doi.org/10.1080/00207540600902262
13. L. Buschkuehl, F. Sahling, S. Helber, H. Tempelmeier, Dynamic capacitated lot-sizing problems: A classification and review of solution approaches, *OR Spectr.*, **32** (2010), 231–261. https://doi.org/10.1007/s00291-008-0150-7
14. J. H. Bookbinder, J.-Y. Tan, Strategies for the probabilistic lot-sizing problem with service-level constraints, *Manag. Sci.*, **34** (1988), 1096–1108. https://doi.org/10.1287/mnsc.34.9.1096
15. S. Helber, F. Sahling, K. Schimmelpfeng, Dynamic capacitated lot sizing with random demand and dynamic safety stocks, *OR Spectr.*, **35** (2013), 75–105. https://doi.org/10.1007/s00291-012-0283-6
16. A. Shapiro, D. Dentcheva, A. Ruszczynski, *Lectures on Stochastic Programming: Modeling and Theory*, SIAM, Philadelphia, 2009. https://doi.org/10.1137/1.9780898718751
17. N. Sereshti, Y. Adulyasak, R. Jans, Managing flexibility in stochastic multi-level lot sizing problem with service level constraints, *Omega*, **122** (2024), 102957. https://doi.org/10.1016/j.omega.2023.102957
18. M. Henig, Y. Gerchak, The structure of periodic review policies in the presence of random yield, *Oper. Res.*, **38** (1990), 634–643. https://doi.org/10.1287/opre.38.4.634
19. C. A. Yano, H. L. Lee, Lot sizing with random yields: A review, *Oper. Res.*, **43** (1995), 311–334. https://doi.org/10.1287/opre.43.2.311
20. A. Hsu, Y. Bassok, Random yield and random demand in a production system with downward substitution, *Oper. Res.*, **47** (1999), 277–290. https://doi.org/10.1287/opre.47.2.277
21. A. Grosfeld-Nir, Y. Gerchak, Multiple lot sizing in production to order with random yields: Review of recent advances, *Ann. Oper. Res.*, **126** (2004), 43–69. https://doi.org/10.1023/B:ANOR.0000012275.01260.f5
22. P. Berling, D. R. Sonntag, Inventory control in production–inventory systems with random yield and rework: The unit-tracking approach, *Prod. Oper. Manag.*, **31** (2022), 2628–2645. https://doi.org/10.1111/poms.13706
23. P. Metzker, S. Thevenin, Y. Adulyasak, A. Dolgui, Robust optimization for lot-sizing problems under yield uncertainty, *Comput. Oper. Res.*, **149** (2023), 106025. https://doi.org/10.1016/j.cor.2022.106025
24. P. Metzker Soares, S. Thevenin, Y. Adulyasak, A. Dolgui, Adaptive robust optimization for lot-sizing under yield uncertainty, *Eur. J. Oper. Res.*, **313** (2024), 513–526. https://doi.org/10.1016/j.ejor.2023.08.036
25. P. Metzker, S. Thevenin, Y. Adulyasak, A. Dolgui, Distributionally robust optimization for the multi-period multi-item lot-sizing problems under yield uncertainty, *IEEE Trans. Autom. Sci. Eng.*, **22** (2025), 11731–11752. https://doi.org/10.1109/TASE.2025.3535945
26. A. Ben-Tal, A. Nemirovski, Robust convex optimization, *Math. Oper. Res.*, **23** (1998), 769–805. https://doi.org/10.1287/moor.23.4.769

27. D. Bertsimas, M. Sim, The price of robustness, *Oper. Res.*, **52** (2004), 35–53. https://doi.org/10.1287/opre.1030.0065

28. E. Delage, Y. Ye, Distributionally robust optimization under moment uncertainty with application to data-driven problems, *Oper. Res.*, **58** (2010), 595–612. https://doi.org/10.1287/opre.1090.0741

29. W. Wiesemann, D. Kuhn, M. Sim, Distributionally robust convex optimization, *Oper. Res.*, **62** (2014), 1358–1376. https://doi.org/10.1287/opre.2014.1314

30. P. Mohajerin Esfahani, D. Kuhn, Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations, *Math. Program.*, **171** (2018), 115–166. https://doi.org/10.1007/s10107-017-1172-1

31. R. Gao, A. J. Kleywegt, Distributionally robust stochastic optimization with Wasserstein distance, *Math. Oper. Res.*, **48** (2023), 603–655. https://doi.org/10.1287/moor.2022.1275

32. H. K. Okyay, F. Karaesmen, S. Özekici, Newsvendor models with dependent random supply and demand, *Optim. Lett.*, **8** (2014), 983–999. https://doi.org/10.1007/s11590-013-0616-7

33. Z. Chen, W. Xie, Regret in the newsvendor model with demand and yield randomness, *Prod. Oper. Manag.*, **30** (2021), 4176–4197. https://doi.org/10.1111/poms.13515

34. S. Thevenin, Y. Adulyasak, J.-F. Cordeau, Material requirements planning under demand uncertainty using stochastic optimization, *Prod. Oper. Manag.*, **30** (2021), 475–493. https://doi.org/10.1111/poms.13277

35. M. Schlenkrich, S. N. Parragh, Capacitated multi-item multi-echelon lot sizing with setup carry-over under uncertain demand, *Int. J. Prod. Econ.*, **277** (2024), 109379. https://doi.org/10.1016/j.ijpe.2024.109379

36. M. Schlenkrich, J.-F. Cordeau, S. N. Parragh, Progressive hedging for multi-stage stochastic lot sizing problems with setup carry-over under uncertain demand, *Comput. Oper. Res.*, **193** (2026), 107523. https://doi.org/10.1016/j.cor.2026.107523

37. O. Ledoit, M. Wolf, A well-conditioned estimator for large-dimensional covariance matrices, *J. Multivar. Anal.*, **88** (2004), 365–411. https://doi.org/10.1016/S0047-259X(03)00096-4



AIMS Press

© 2026 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0)