Advances in Industrial Engineering, December 2026, 60(2): 517-539
DOI: 10.22059/aie.2026.419080.1995

RESEARCH PAPER

Endogenous Bottleneck Capacity with Inventory Risk under Demand Uncertainty

Jalal Delaram[1 *]

[1]*Assistant Professor, School of Industrial Engineering, College of Engineering, University of Tehran, Tehran, Iran.*

Received: 21 July 2026, Revised: 16 August 2026, Accepted: 19 September 2026
© The Author(s) retain the copyright 2026

Abstract

Production bottlenecks are central to the performance of manufacturing and service systems, yet their economic analysis typically assumes exogenous capacity. This paper studies a single-period, multi-product production system with stochastic demand. In this system, products compete for a shared bottleneck whose capacity is determined endogenously through costly investment. Extending the classical bottleneck framework, we incorporate inventory holding costs and salvage values, allowing inventory risk to directly influence both production and capacity decisions. We characterize the optimal solution and show that bottleneck capacity behaves as an endogenous economic resource whose shadow price equals the marginal cost of capacity investment. The resulting optimality conditions yield a capacity-adjusted newsvendor rule that links inventory overage risk and bottleneck usage to production quantities. Comparative statics demonstrate that higher holding costs discourage capacity expansion, while greater salvage opportunities justify larger investments. A numerical example illustrates the analytical results and highlights the importance of coordinating production planning, inventory management, and capacity investment under demand uncertainty.

Keywords: Newsvendor Models; Shared Resources; Capital Investment; Inventory Overage; Stochastic Demand.

1 Introduction

Production systems frequently operate under capacity limitations that restrict the flow of work and determine overall system performance. Such constraints, often referred to as bottlenecks, play a critical role in shaping production decisions, product mix, and profitability. When demand is uncertain, the economic implications of bottlenecks become even more pronounced, as production decisions must balance capacity scarcity against the risk of excess inventory (Francas et al., 2024; Yang, 2021).

The economic analysis of bottlenecks has a long history in operations research (Arnott et al., 1990; Irma-Delia & Santiago-Omar, 2025; Wang et al., 2005). A key contribution in this literature is Lawrence and Buss (1995), who demonstrate that bottleneck capacity affects optimal production quantities through a shadow price that acts as an implicit marginal cost under stochastic demand. This insight established an important link between bottleneck management and classical inventory models. Subsequent studies have further examined production planning under shared capacity constraints, yet most of this work treats capacity as

* Corresponding author: (Jalal Delaram)
Email: delaram@ut.ac.ir

fixed and focuses on short-term allocation decisions rather than strategic capacity determination.

In practice, bottleneck capacity is rarely exogenous. Expanding or relieving a bottleneck typically requires costly investment, and the profitability of such investments depends on demand uncertainty and inventory risk. In particular, unsold production may incur holding costs or be salvaged at a reduced value, which directly affects the economic attractiveness of both production and capacity expansion. However, existing models largely abstract from these inventory considerations when analyzing bottleneck-constrained systems.

This paper addresses the following research question: How should a firm jointly determine production quantities and capacity investment when products compete for a shared constrained resource and demand is uncertain? To answer this question, we develop a single-period, multi-product model in which production decisions are made before demand realization, excess inventory incurs holding costs, and unsold units may generate salvage revenue. The level of the constrained resource is endogenized through a convex investment cost. We characterize the optimal solution and show that production quantities satisfy a generalized critical-fractile condition in which inventory risk and resource usage jointly determine marginal profitability. We further show that the shadow value of the constrained resource equals the marginal investment cost, providing a clear economic interpretation of capacity scarcity.

Our work is closely related to but distinct from the capacity-constrained newsvendor literature. In particular, Van Mieghem (2007) studies risk mitigation through resource diversification and flexibility in newsvendor networks, but treats capacity as exogenous or as a discrete investment choice. By contrast, our model endogenizes continuous capacity investment and explicitly links it to inventory overage risk through holding costs and salvage values. Unlike prior models where the shadow price of capacity is merely a scarcity indicator, we show that the shadow price equals the marginal cost of capacity investment, i.e., $\lambda^* = G'(B^*)$. This result provides a direct economic link between capacity investment decisions and production planning under demand uncertainty.

Our analysis yields several insights. Higher inventory holding costs discourage investment in the constrained resource, while greater salvage opportunities mitigate downside risk and justify larger capacity levels. A numerical illustration complements the analytical results and highlights how these effects shape product mix and investment decisions. While our analysis is restricted to a single-period setting for analytical tractability, the model provides a foundation for dynamic extensions with inventory carryover and capacity adjustment costs.

The remainder of the paper is organized as follows. Section 2 reviews recent and related papers, then provides the research gap and our contribution. Section 3 introduces the model, formulate it formally, and derives the optimal production and capacity decisions. Section 4 discusses the results and presents the managerial insights. Section 5 concludes the findings, presents limitations of this work with directions for future research.

2 Literature Review

This study relates to two major streams of the operations research and operations management literature: research on bottlenecks in production systems and the extensive body of work on newsvendor-type models under uncertainty. By combining these streams, the paper addresses a gap concerning the endogenous economic determination of bottleneck capacity under inventory risk.

2.1 Bottlenecks in Production Systems

Bottlenecks have long been recognized as critical determinants of system performance in manufacturing and service operations (Irma-Delia & Santiago-Omar, 2025; Thorstenson &

Ramani, 2020; Wang et al., 2005). A substantial literature examines bottlenecks from an operational perspective, focusing on scheduling, throughput analysis, and performance improvement. More recent work emphasizes the dynamic and non-static nature of bottlenecks in complex production systems. For example, Lai et al. (2021) developed data-driven methods for identifying shifting bottlenecks in non-serial manufacturing lines, demonstrating significant performance gains in industrial applications. Subramaniyan et al. (2018) addressed throughput bottlenecks using the real-time machine data collection. The paper proposes a data-driven algorithm to better identify bottleneck groups and provide diagnostic insights. The algorithm is based on the active period theory of throughput bottleneck analysis. Such studies highlight the practical importance of bottlenecks but primarily treat capacity as technologically fixed and focus on detection rather than economic decision-making.

From a planning and improvement perspective, the Theory of Constraints (TOC) provides influential managerial guidance on bottleneck management. Wu et al. (2020) extended TOC to stochastic environments using queueing and simulation models and show that improving the nominal bottleneck may not always be optimal once variability and improvement costs are considered. While this work introduces economic trade-offs into bottleneck improvement, capacity levels themselves remain largely exogenous, and inventory risk is not explicitly modeled. Buddas (2014) studied bottlenecks in the humanitarian supply chain and operations. The focus in this paper is set on the activities of aid supply procurement and aid consolidation into standardized deliveries of humanitarian aid.

Tang (2019) proposed a new method based on overall equipment effectiveness to identify the bottlenecks of complex manufacturing systems. The new method considers not only the output of a manufacturing system but also the influences from its upstream and downstream systems. Lawrence and Buss (1995) provided one of the earliest economic analyses of production bottlenecks under stochastic demand. Lawrence showed that bottleneck capacity enters optimal production decisions through a shadow price that acts as an implicit marginal cost. This result establishes a conceptual bridge between bottleneck analysis and inventory theory. However, bottleneck capacity is assumed fixed, and inventory overage risk plays no role in shaping capacity decisions.

2.2 Newsvendor and Risk-Based Production Models

The newsvendor model is a foundational framework for analyzing production and ordering decisions under demand uncertainty (Hamzaoui et al., 2023; Mandal & Pal, 2021). Classical formulations capture the trade-off between underage and overage costs through holding and salvage values. Numerous extensions consider multiple products, pricing decisions, substitution effects, and robustness to uncertainty. For example, Murray et al. (2012) studied a multi-product newsvendor model with endogenous pricing and cross-elastic demand, developing computationally tractable solution methods for large-scale problems. Similarly, Zhang et al. (2021) analyzed a robust multi-product newsvendor with stochastic substitution rates, deriving tractable solutions under uncertainty sets.

Another strand considers capacity-constrained or shared-resource newsvendor models, where multiple products compete for limited production capacity (Yu et al., 2026). These models typically derive generalized critical-fractile conditions that allocate a fixed capacity across products under uncertainty (Van Mieghem, 2007). Van Mieghem (2007) provides a comprehensive analysis of risk mitigation in newsvendor networks through resource diversification, flexibility, and hedging, but treats capacity as exogenous or as a binary investment choice. While this literature captures important interactions between demand risk and capacity scarcity, capacity itself is usually treated as exogenously given, and investment decisions are not modeled in a continuous endogenous framework (Modak et al., 2025; Ostovari et al., 2025). Our model extends this line of research by endogenizing capacity investment and

linking it explicitly to inventory overage risk, yielding a capacity-adjusted newsvendor rule where the shadow price equals the marginal investment cost.

Recent work on production–inventory planning under uncertainty (Yin et al., 2016), such as Fleuren et al. (2025), incorporates complex network structures, multiple capacity constraints, and non-stationary demand using stochastic programming approaches. These models are well suited for large-scale planning but often sacrifice analytical transparency and do not isolate the economic role of bottlenecks or their shadow prices (Gharaei et al., 2020).

Most multi-product newsvendor models assume independent demands across products for analytical tractability (Murray et al., 2012; Yu et al., 2026). However, some studies have considered correlated demands. For example, Van Mieghem (2007) analyzes risk mitigation through resource diversification when demands are correlated, showing that negatively correlated demands can provide hedging benefits. Zhang et al. (2021) consider substitution effects in a robust multi-product newsvendor model, which implicitly accounts for demand correlations through substitution rates. Our model assumes independent demands to maintain analytical tractability and isolate the effects of inventory risk on capacity investment.

2.3 Real Options and Capacity Investment under Uncertainty

A related body of literature examines capacity investment decisions through the lens of real options theory. Dixit and Pindyck (1994) provide a comprehensive framework for investment under uncertainty, emphasizing the irreversibility of capital investments and the value of waiting. In their framework, the option to delay investment creates an opportunity cost that must be accounted for when evaluating capacity expansion projects. Trigeorgis (1996) extends this analysis to incorporate multiple sources of uncertainty and strategic interactions, identifying various types of real options including the option to defer, expand, contract, or abandon operations.

The real options perspective is particularly relevant for capacity investment decisions, as capacity expansions are typically irreversible and involve significant sunk costs. The option to delay investment allows firms to wait for better information about demand conditions, reducing the risk of overinvestment. This option value is higher when demand uncertainty is greater or when the investment cost is highly irreversible.

Our model complements the real options literature by focusing on the joint determination of production quantities and capacity levels in a multi-product newsvendor setting with inventory risk. While the real options approach emphasizes the timing of investment (when to invest), our model focuses on the optimal level of capacity (how much to invest) given the current demand uncertainty. A natural extension would be to combine these perspectives by allowing firms to delay capacity investment until demand uncertainty resolves.

A related body of literature examines capacity investment decisions through the lens of real options theory (Dixit & Pindyck, 1994; Trigeorgis, 1996). This perspective emphasizes the irreversibility of capital investments and the value of waiting for better information before committing to capacity expansion. The option to delay investment creates an opportunity cost that must be accounted for when evaluating capacity projects. Our model focuses on the optimal level of capacity (how much to invest) given current demand uncertainty, complementing the real options focus on the timing of investment (when to invest).

2.4 Comparison with Closest Models

To clarify the positioning of our contribution relative to the most closely related studies, Table 1 summarizes the key differences in assumptions, decision variables, and main results across selected models from the literature.

Table 1. Comparison of Model Features and Contributions.

Study	Capacity Treatment	Inventory Risk	Demand Structure	Decision Variables	Key Results
Lawrence and Buss (1995)	Fixed	Not modeled	Stochastic	Production quantities	Shadow price as implicit marginal cost
Van Mieghem (2007)	Fixed (flexibility choices)	Not modeled	Stochastic	Resource allocation, flexibility investment	Risk mitigation through diversification
Murray et al. (2012)	Fixed	Standard newsvendor	Stochastic	Production quantities, prices	Endogenous pricing with capacity constraints
Zhang et al. (2021)	Fixed	Standard newsvendor	Stochastic (robust)	Production quantities	Robust multi-product newsvendor with substitution
Fleuren et al. (2025)	Multiple fixed constraints	Complex inventory costs	Stochastic (non-stationary)	Production-inventory planning	Stochastic programming approach
This Paper	**Endogenous (costly investment)**	**Explicitly modeled (holding costs and salvage values)**	**Stochastic (independent)**	**Production quantities and capacity investment**	*Generalized critical fractile with capacity shadow price; $\lambda = G'(B)$*; capacity-adjusted newsvendor rule**

As shown in Table 1, the key distinguishing features of our model are:

1. Endogenous capacity investment: Unlike prior studies that treat capacity as fixed or exogenously given, we allow the firm to choose the optimal capacity level through a convex investment cost function. This endogenization enables us to derive the relationship $\lambda^* = G'(B^*)$, linking the shadow price directly to the marginal cost of capacity.
2. Explicit modeling of inventory risk: We incorporate holding costs and salvage values directly into the objective function, allowing inventory overage risk to influence both production and capacity decisions. This integration yields a generalized newsvendor rule in Equation (11) that accounts for both capacity scarcity and inventory risk.
3. Unified analytical framework: By combining multi-product production decisions, inventory risk, and endogenous capacity investment within a single analytical model, we provide a tractable framework that captures the interdependencies between these decisions. Our results show that inventory parameters affect not only production quantities but also the optimal capacity level and its shadow price.
4. Corrected comparative statics: We establish the directional effects of holding costs and salvage values on production quantities, capacity, and the shadow price, providing clear managerial guidance for capacity investment under demand uncertainty.

This comparison highlights that our model fills a gap in the literature by unifying bottleneck economics and inventory theory within a tractable analytical framework, extending the classical analysis of Lawrence and Buss (1995) and offering new insights into coordinated production, inventory, and capacity decisions under uncertainty.

We assume that each demand distribution $F_i(\cdot)$ is continuous and differentiable with density $f_i(\cdot) = F_i'(\cdot) > 0$ on the support of D_i. This differentiability assumption justifies the derivative-based optimization approach and the use of first-order KKT conditions. For non-differentiable demand distributions (e.g., discrete empirical distributions), the KKT conditions would need to be expressed using subgradient methods; we leave this extension for future research.

2.5 Research Gap and Contribution

Despite these advances, the literature lacks an analytical framework that jointly models inventory risk, multi-product production decisions, and endogenous bottleneck capacity investment. Existing bottleneck studies emphasize operational identification or improvement, while newsvendor and capacity-constrained models focus on allocating fixed capacity rather

than determining its optimal level. Conversely, capacity investment models typically abstract from detailed inventory overage risk and salvage considerations.

This paper fills this gap by developing a single-period, multi-product model in which production quantities and bottleneck capacity are chosen jointly under stochastic demand. By incorporating holding costs and salvage values, the model allows inventory risk to directly influence both production decisions and capacity investment. The analysis shows that bottleneck capacity behaves as an endogenous economic resource whose shadow price equals the marginal cost of investment. The resulting optimality conditions yield a generalized, capacity-adjusted newsvendor rule that links inventory economics to bottleneck scarcity.

By unifying bottleneck economics and inventory theory within a tractable analytical framework, this study extends the classical analysis of Lawrence and Buss (1995) and provided new insights into how firms should coordinate production planning, inventory management, and capacity investment under demand uncertainty.

3 Problem Formulation

We consider a single-period decision problem faced by a firm that produces multiple products using a shared constrained resource. At the beginning of the period, the firm simultaneously chooses production quantities for each product and the capacity level of the bottleneck resource, before demand is realized. Product demands are uncertain and independently distributed, and unsold production may incur holding costs or generate salvage value. The bottleneck resource limits total feasible production and can be expanded at a convex investment cost. This section formally defines the model primitives, decision variables, and assumptions. Section 3.1 introduces the demand and cost structure, Section 3.2 formulates the expected profit function, and Section 3.3 presents the resulting optimization problem.

3.1 Problem Description and Assumptions

We consider a single-period production system in which a firm produces multiple products that compete for a shared constrained resource (a bottleneck). Production quantities and bottleneck capacity are chosen before demand is realized. Product-specific demands are stochastic, and unsold inventory at the end of the period incurs holding costs and may be salvaged at a reduced value. Bottleneck capacity is expandable through costly investment and determines the total amount of the constrained resource available for production. Table 2 summarizes the notation used throughout the paper, including indices, parameters, and decision variables.

Table 2. Nomenclature

Indices		Unit	Range
i	product index, $i = 1, \dots, n$		
Parameters			
p_i	unit selling price of product i	\$/unit	10 – 1,000
c_i	unit production cost of product i	\$/unit	5 – 500
h_i	unit holding cost for unsold inventory of product i	\$/unit	1 – 50
v_i	unit salvage value for unsold inventory of product i, with $0 \le v_i < p_i$	\$/unit	0 – 900
a_i	amount of bottleneck capacity required to produce one unit of product i	capacity units	0.5 – 5
D_i	random demand for product i with cumulative distribution function $F_i(.)$	units	100 – 10,000
$G(B)$	strictly convex cost function for installing bottleneck capacity B, with $G'(B) > 0$, $G''(B) > 0$, and $G(B) \to \infty$ as $B \to \infty$	\$	Varies by industry
Decision Variables			
q_i	Production quantity of product i, $q_i \ge 0$	units	0 – 10,000
B	installed bottleneck capacity, $B \ge 0$	capacity units	0 – 50,000

We assume that each demand distribution ($F_i(\cdot)$) is continuous and differentiable with density ($f_i(\cdot) = F_i'(\cdot) > 0$) on the support of ($D_i$). This differentiability assumption justifies the derivative-based optimization approach. For non-differentiable distributions (e.g., discrete empirical distributions), the KKT conditions would need to be expressed using subgradients; we leave this extension for future research.

We also assume $G(B)$ is strictly convex ($G''(B) > 0$) to ensure a unique optimal capacity level. The condition $G(B) \to \infty$ as $B \to \infty$ ensures that the objective function is bounded above and a finite optimal solution exists. Several specific functional forms are commonly used in practice:

(a) Quadratic form: $G(B) = kB^2$, where $k > 0$ is a cost parameter. This form is widely used due to its analytical tractability and yields a linear marginal cost $G'(B) = 2kB$. It is appropriate when capacity expansion costs increase quadratically with capacity.

(b) Power form: $G(B) = kB^{\alpha}$, with $k > 0$ and $\alpha > 1$. This form allows flexibility in the degree of convexity: larger α implies more rapidly increasing marginal costs. The marginal cost is $G'(B) = k\alpha B^{\alpha-1}$.

(c) S-shaped form: $G(B) = k_1B^2 + k_2B^3$ or a logistic-type function, which captures economies of scale at low capacity levels (initially concave) followed by diminishing returns at high capacity levels (convex).

The qualitative results of our model—binding capacity $B^* = \sum a_i q_i^*$, positive shadow price $\lambda^* = G'(B^*) > 0$, and the generalized critical fractile in Equation (11)—hold for any strictly convex and increasing $G(B)$. However, the magnitude of the effects (e.g., how much capacity changes in response to a change in holding costs) depends on the specific functional form and the curvature of $G(B)$. The quadratic form $G(B) = kB^2$ is used in the numerical example for its simplicity and tractability.

For analytical tractability, we assume that product demands D_i are independently distributed. This assumption is standard in the multi-product newsvendor literature and allows us to derive closed-form optimality conditions. However, in practice, product demands may be correlated due to common market factors, substitution effects, or complementarity.

We assume a risk-neutral decision maker who maximizes expected profit. This assumption is standard in the analytical newsvendor literature and allows us to derive closed-form optimality conditions. However, in practice, managers may exhibit risk aversion, particularly for capacity investment decisions that involve large, irreversible commitments.

The salvage value constraint $0 \le v_i < p_i$ ensures that the salvage value does not exceed the selling price, which is economically reasonable: a firm cannot sell a product in a secondary market for more than its original selling price. However, the salvage value v_i may exceed the production cost c_i. This can occur in several realistic scenarios: (i) the product has significant value in secondary markets (e.g., refurbished electronics, spare parts for machinery); (ii) the product can be remarketed to different customer segments (e.g., overstocked fashion items sold to discount retailers); or (iii) the product has alternative uses (e.g., raw materials that can be repurposed). In such cases, $v_i > c_i$ is possible and economically meaningful.

3.2 Mathematical Model

We now formulate the firm's expected profit function. For each product, realized sales depend on the minimum of production quantity and stochastic demand, generating underage and overage outcomes. Excess production incurs holding costs and may yield salvage value, while all produced units consume the shared bottleneck resource. Bottleneck capacity constrains total feasible production and is determined endogenously through a convex investment cost. The firm's problem is to choose production quantities and bottleneck capacity to maximize expected profit subject to the capacity constraint. Consider sales of product i are given by $\min\{q_i, D_i\}$, and the end-of-period leftover inventory is $(q_i - D_i)^+ = \max\{q_i -$

$D_i, 0\}$. The expected profit from product i is

$$\Pi_i(q_i) = \mathbb{E}[p_i \min(q_i, D_i)] - c_i q_i - h_i \mathbb{E}[(q_i - D_i)^+] + v_i \mathbb{E}[(q_i - D_i)^+] \quad (1)$$

To simplify Equation (1), we use two standard newsvendor identities. First, the expected sales (the minimum of production quantity and demand) is:

$$E[min(q_i, D_i)] = q_i - \int_0^{q_i} F_i(x)dx \quad (1a)$$

Second, the expected leftover inventory is:

$$E[(q_i - D_i)^+] = \int_0^{q_i} F_i(x)dx \quad (1b)$$

Substituting (1a) and (1b) into Equation (1) gives:

$$\begin{aligned} \Pi_i(q_i) &= p_i \left[q_i - \int_0^{q_i} F_i(x)dx \right] - c_i q_i - h_i \left[\int_0^{q_i} F_i(x)dx \right] + v_i \left[\int_0^{q_i} F_i(x)dx \right] \\ &= (p_i - c_i) q_i - (p_i - v_i + h_i) \int_0^{q_i} F_i(x)dx \end{aligned}$$

where the last step combines the integral terms: $p_i \int F_i(x)dx + h_i \int F_i(x)dx - v_i \int F_i(x)dx = (p_i - v_i + h_i) \int F_i(x)dx$.

Using standard newsvendor identities, Equation (1) can be rewritten as:

$$\Pi_i(q_i) = (p_i - c_i) q_i - (p_i - v_i + h_i) \int_0^{q_i} F_i(x)dx \quad (2)$$

Total expected profit is therefore:

$$\Pi(\mathbf{q}, B) = \sum_{i=1}^{n} \Pi_i(q_i) - G(B). \quad (3)$$

Production quantities must satisfy the bottleneck constraint:

$$\sum_{i=1}^{n} a_i q_i \leq B \quad (4)$$

The firm should solve the following optimization problem:

$$\max_{\{q_i\}_{i=1}^n, B} \sum_{i=1}^{n} \left[(p_i - c_i) q_i - (p_i - v_i + h_i) \int_0^{q_i} F_i(x)dx \right] - G(B) \quad (5)$$

Subject to

$$\sum_{i=1}^{n} a_i q_i \leq B \tag{6}$$

$$q_i \geq 0, B \geq 0 \qquad i = 1, \ldots, n \tag{7}$$

The objective function is concave in q_i for each product, as $F_i(\cdot)$ is non-decreasing and $G(\cdot)$ is convex. The feasible region is convex, ensuring existence of an optimal solution. The Lagrangian approach is therefore applicable and yields economically interpretable first-order conditions. The differentiability of $F_i(\cdot)$ ensures that the first-order conditions derived from the Lagrangian are valid. If the demand distributions were not differentiable, the KKT conditions would need to be formulated using subgradients, which would complicate the analysis without changing the qualitative insights.

The objective function in (5) is bounded above under the following conditions. For each product (i), the expected profit $\Pi_i(q_i)$ is concave and bounded above because $F_i(\cdot)$ is a cumulative distribution function and $p_i - v_i + h_i > 0$. For the capacity investment cost, the assumption $G(B) \to \infty$ as $B \to \infty$ ensures that the objective does not grow without bound as capacity increases. Thus, a finite optimal solution exists.

3.3 Results and Analysis

Let $\lambda \geq 0$ denote the Lagrangian multiplier associated with the bottleneck capacity constraint. The Lagrangian is given by

$$\mathcal{L}(\mathbf{q}, B, \lambda) = \sum_{i=1}^{n} \left[(p_i - c_i) q_i - (p_i - v_i + h_i) \int_0^{q_i} F_i(x) dx \right] - G(B) + \lambda (B - \sum_{i=1}^{n} a_i q_i) \tag{8}$$

Differentiating the Lagrangian with respect to q_i yields

$$\frac{\partial \mathcal{L}}{\partial q_i} = (p_i - c_i) - (p_i - v_i + h_i) F_i(q_i) - \lambda a_i \tag{9}$$

The KKT conditions for optimality are:

$$(p_i - c_i) - (p_i - v_i + h_i) F_i(q_i^*) - \lambda^* a_i \leq 0, \quad i = 1, \ldots, n \tag{10}$$

With complementary slackness:

$$q_i^* [(p_i - c_i) - (p_i - v_i + h_i) F_i(q_i^*) - \lambda^* a_i] = 0$$

Equivalently, for each product i:

$$F_i(q_i^*) \geq \frac{p_i - c_i - \lambda^* a_i}{p_i - v_i + h_i} \tag{11}$$

With equality if $q_i^* > 0$, and strict inequality if $q_i^* = 0$.

Thus, the optimal production quantity for product i is characterized as follows:

- If $p_i - c_i - \lambda^* a_i > 0$, then $q_i^* > 0$ and:

$$F_i(q_i^*) = \frac{p_i - c_i - \lambda^* a_i}{p_i - v_i + h_i} \tag{12}$$

With $0 < F_i(q_i^*) < 1$ provided that $0 < p_i - c_i - \lambda^* a_i < p_i - v_i + h_i$.

- If $p_i - c_i - \lambda^* a_i \leq 0$, then $q_i^* = 0$. In this case, the product is not profitable to produce because its marginal contribution margin is non-positive after accounting for the opportunity cost of bottleneck capacity.

- If $p_i - c_i - \lambda^* a_i \geq p_i - v_i + h_i$, then $F_i(q_i^*) \geq 1$, which would imply $q_i^* \to \infty$. This case is excluded by the boundedness condition $p_i - c_i - \lambda^* a_i < p_i - v_i + h_i$, which is equivalent to $c_i + v_i > 0$ and ensures finite production quantities.

Differentiating the Lagrangian with respect to B gives

$$\frac{\partial \mathcal{L}}{\partial B} = \lambda - G'(B) \tag{13}$$

Thus, the optimality condition for capacity investment is

$$\lambda = G'(B) \tag{14}$$

With complementary slackness condition $\lambda \geq 0, B \geq 0$.

The capacity constraint satisfies $\lambda(B - \sum_{i=1}^{n} a_i q_i) = 0$; hence, whenever capacity is scarce (i.e. $B = \sum_{i=1}^{n} a_i q_i$), the multiplier λ is strictly positive and measures the marginal value of bottleneck capacity. As the multiplier λ represents the marginal increase in expected profit from a marginal increase in bottleneck capacity. Equation (11) shows that capacity scarcity effectively increases the marginal cost of production by λa_i, reducing optimal production quantities relative to the unconstrained case.

Proposition 1 (General Critical Fractile). For each product i, the optimal production quantity satisfies

If $p_i - c_i - \lambda^* a_i > 0$, then:

$$q_i^* = F_i^{-1}\left(\frac{p_i - c_i - \lambda^* a_i}{p_i - v_i + h_i}\right)$$

where λ^* is the optimal shadow price of bottleneck capacity, and $F_i^{-1}(\cdot)$ denotes the inverse cumulative distribution function.

Equivalently, the KKT condition can be written as:

$$F_i(q_i^*) \geq \frac{p_i - c_i - \lambda^* a_i}{p_i - v_i + h_i}, \quad \text{with equality if } q_i^* > 0. \tag{15}$$

The condition $0 < F_i(q_i^*) < 1$ requires $0 < p_i - c_i - \lambda^* a_i < p_i - v_i + h_i$. The upper bound is equivalent to $c_i + v_i > 0$, which always holds under our assumptions $c_i > 0$ and $v_i \geq 0$.

Proof. The result follows directly from the KKT conditions in Equations (10)-(12). Since the objective function is strictly concave in each q_i, the KKT conditions are both necessary and sufficient for optimality. The boundary case $q_i^* = 0$ arises when the product's marginal contribution $p_i - c_i$ does not exceed the opportunity cost of capacity $\lambda^* a_i$, making production unprofitable.

The optimal capacity level B^* is unique under the assumption $G''(B) > 0$ (strict convexity

of the capacity investment cost). To see this, note that $\lambda^* = G'(B^*)$ is strictly increasing in B^* when $G''(B) > 0$. The production quantities q_i^* are strictly decreasing in λ^*, and the capacity constraint $\sum_{i=1}^{n} a_i q_i^*(\lambda^*) = B^*(\lambda^*)$ has a unique solution by the monotonicity of both sides. Thus, the first-order conditions determine a unique optimal solution.

The boundary case $q_i^* = 0$ has a clear economic interpretation. A product is not produced if its unit contribution margin $p_i - c_i$ is less than or equal to the shadow price of the bottleneck capacity multiplied by its resource consumption a_i. In other words, the product does not generate enough profit to justify the bottleneck capacity it would consume. This condition endogenously determines which products are included in the optimal product mix, reflecting the economic trade-off between margins and resource usage. Products with low margins or high resource consumption are optimally excluded from production when capacity is scarce and valuable.

The condition for an interior solution $q_i^* > 0$ is $p_i - c_i - \lambda^* a_i > 0$, which means the product's net contribution after accounting for capacity opportunity cost is positive. The fractile $(p_i - c_i - \lambda^* a_i)/(p_i - v_i + h_i)$ lies strictly between 0 and 1 when:

$$0 < p_i - c_i - \lambda^* a_i < p_i - v_i + h_i$$

The left inequality is the interior condition above. The right inequality is equivalent to $c_i + v_i > 0$, which always holds since production costs are positive and salvage values are non-negative.

Proposition 2 (Endogenous Bottleneck Condition). At the optimum, bottleneck capacity satisfies

$$\lambda^* = G'(B^*) \tag{16}$$

Proof. From the first-order condition with respect to B, we have $\lambda - G'(B) = 0$ whenever $B > 0$. Since $G(\cdot)$ is convex and increasing, this condition uniquely determines B^*.

The result $\lambda^* = G'(B^*)$ in Proposition 2 holds for any strictly convex and increasing capacity cost function $G(B)$, including the quadratic, power, and S-shaped forms. The key requirement is the strict convexity of $G(B)$, which ensures a unique optimal capacity level and a monotonic relationship between the shadow price and capacity. The specific functional form affects the magnitude of the comparative statics but not their qualitative direction. For example, a more convex $G(B)$ (larger α in the power form or larger k in the quadratic form) implies that capacity is more costly to expand, leading to smaller optimal capacity and a higher shadow price for any given production plan.

Proposition 3 (Binding Capacity). If $G'(0) < \max_i \frac{p_i - c_i}{a_i}$, then the optimal bottleneck constraint binds:

$$\sum_{i=1}^{n} a_i q_i^* = B^*, \lambda^* > 0 \tag{17}$$

Proof. If the constraint were slack, then $\lambda = 0$, implying $G'(B) = 0$, which contradicts the assumption $G'(0) < \max_i \frac{p_i - c_i}{a_i}$. Hence, capacity must bind at the optimum.

Proposition 4 (Decentralized Interpretation). The centralized solution is equivalent to decentralized production decisions in which each product faces an adjusted marginal cost

$$c_i^{eff} = c_i + \lambda^* a_i \tag{18}$$

Proof. Substituting c_i^{eff} into the standard newsvendor condition yields exactly Equation (11). Thus, the bottleneck shadow price can be interpreted as a per-unit congestion charge.

Equation (11) provides direct insights into how inventory parameters and capacity investment costs affect optimal production decisions. An increase in the holding cost h_i raises the marginal cost of overage, which increases the denominator of the generalized critical fractile and consequently reduces the optimal production quantity q_i^*. Intuitively, higher holding costs make excess inventory more expensive, leading the firm to adopt more conservative production levels.

Conversely, an increase in the salvage value v_i reduces the effective overage cost and mitigates downside inventory risk. As a result, the critical fractile increases, leading to a higher optimal production quantity. Salvage opportunities therefore partially offset the adverse effects of demand uncertainty and encourage more aggressive production.

Finally, higher bottleneck investment costs, reflected in an increase in the marginal capacity cost $G'(B)$, raise the shadow price λ^* of the constrained resource. Through Equation (11), this increase acts as an implicit marginal cost proportional to each product's resource consumption a_i, reducing optimal production quantities across all products. Products that are more bottleneck-intensive experience a larger reduction, highlighting the role of resource usage in shaping product mix decisions under capacity scarcity.

Corollary 1 (Comparative Statics). Under the assumptions of Propositions 1–4:

- $\frac{\partial q_i^*}{\partial h_i} < 0$,
- $\frac{\partial q_i^*}{\partial v_i} > 0$,
- $\frac{\partial q_i^*}{\partial G'(B)} < 0$, with magnitude proportional to a_i.

To verify the comparative static with respect to salvage value, note that from Equation (11), the direct effect (holding λ^* fixed) gives:

$$\left.\frac{\partial q_i^*}{\partial v_i}\right|_{\lambda \text{ fixed}} = \frac{F_i(q_i^*)}{(p_i - v_i + h_i) f_i(q_i^*)} > 0 \tag{19}$$

However, since λ^* is endogenous and changes with v_i, the full comparative static is:

$$\frac{dq_i^*}{dv_i} = \left.\frac{\partial q_i^*}{\partial v_i}\right|_{\lambda \text{ fixed}} + \frac{\partial q_i^*}{\partial \lambda^*} \cdot \frac{d\lambda^*}{dv_i} \tag{20}$$

where $\frac{\partial q_i^*}{\partial \lambda^*} = -\frac{a_i}{(p_i - v_i + h_i) f_i(q_i^*)} < 0$

and $\frac{d\lambda^*}{dv_i} > 0$ (from the corrected Corollary 2).

Thus, the indirect effect through λ^* partially offsets the direct positive effect. The total effect ($\partial q_i^* / \partial v_i > 0$) when the direct effect dominates, which holds under standard regularity conditions. The managerial implication remains that higher salvage values encourage larger production quantities.

It is important to note that the above discussion describes the direct effect of salvage value on the optimal production quantity, holding the shadow price λ* fixed. In equilibrium, a change in v_i also affects the optimal shadow price λ* through its effect on the marginal profitability of capacity. This indirect effect partially offsets the direct effect. The corrected comparative statics presented in Corollary 1 below account for the net effect.

Corollary 2 (Monotonicity of the Shadow Price). An increase in holding costs h_i or a decrease in salvage values v_i weakly decreases the optimal shadow price of bottleneck capacity λ^*.

From Proposition 2, $\lambda^* = G'(B^*)$. Higher holding costs reduce optimal production quantities q_i^*, decreasing the marginal profitability of capacity. To restore optimality, capacity investment must decrease, which raises the marginal value of the remaining capacity, increasing λ^*. Corollary 2 establishes that greater inventory overage risk decreases the economic scarcity of the bottleneck resource. Inventory risk therefore affects not only production quantities but also the marginal value of capacity, as captured by the shadow price.

Corollary 3 (Capacity Response to Inventory Risk). The optimal bottleneck capacity B^* is decreasing in holding costs h_i and increasing in salvage values v_i.

From Proposition 2, the optimal bottleneck capacity satisfies $\lambda^* = G'(B^*)$. By Corollary 2, an increase in holding costs h_i or a decrease in salvage values v_i weakly decreases λ^*. Since $G(\cdot)$ is strictly convex and increasing, its derivative $G'(\cdot)$ is strictly increasing. Therefore, a decrease in λ^* implies a decrease in B^*, while an increase in λ^* implies an increase in B^*. Hence, B^* is decreasing in holding costs and increasing in salvage values.

To derive the corrected comparative statics, we analyze the system of optimality conditions. The optimal production quantities and capacity satisfy:

$$F_i(q_i^*) = \frac{p_i - c_i - \lambda^* a_i}{p_i - v_i + h_i}, \quad i = 1, \dots, n \tag{21}$$

$$\lambda^* = G'(B^*) \tag{22}$$

$$B^* = \sum_{i=1}^{n} a_i q_i^* \tag{23}$$

Consider the effect of a change in the holding cost (h_i). Implicit differentiation of the capacity binding condition gives:

$$\frac{dB^*}{dh_i} = a_i \frac{dq_i^*}{dh_i} \tag{24}$$

From the first-order condition, $(dq_i^*/dh_i < 0)$ (as shown in Corollary 1). Therefore:

$$\frac{dB^*}{dh_i} < 0 \tag{25}$$

Since $(\lambda^* = G'(B^*))$ and $(G''(B^*) > 0)$, we have:

$$\frac{d\lambda^*}{dh_i} = G''(B^*) \cdot \frac{dB^*}{dh_i} < 0 \tag{26}$$

Similarly, for the salvage value (v_i), $(dq_i^*/dv_i > 0)$ (from the corrected Corollary 1), hence:

$$\frac{dB^*}{dv_i} > 0, \quad \text{and} \quad \frac{d\lambda^*}{dv_i} > 0 \tag{27}$$

Thus, higher holding costs reduce both the optimal capacity (B^*) and the shadow price (λ^*), while higher salvage values increase both (B^*) and (λ^*). The economic intuition is that higher

holding costs reduce the profitability of production, lowering the marginal value of capacity and consequently reducing both capacity investment and its shadow price.

The economic interpretation of Proposition 3 is as follows. Part (a) establishes that capacity will never be left idle: since capacity is costly, the firm optimally installs exactly the amount needed to support the chosen production quantities. Part (b) provides the condition under which capacity investment is worthwhile: the marginal cost of the first unit of capacity must be less than the maximum marginal profit per unit of resource usage. If this condition fails, the firm would choose zero capacity and no production. Part (c) confirms that when capacity is positive, its shadow price is also positive, reflecting the economic scarcity of the bottleneck resource. This separation clarifies that the condition $(G'(0) < \max_i(p_i - c_i)/a_i)$ is primarily needed to ensure that the optimal solution involves positive capacity investment, not to establish that capacity binds.

4 Numerical Illustration

This section presents a numerical example to illustrate the analytical results derived in Section 3 and to demonstrate the comparative statics established in Corollaries 1-3. The example considers a firm producing two products that compete for a shared bottleneck resource.

4.1 Parameter Settings

We consider two products with the following parameters:

Product 1:

- Selling price: $p_1 = 12$ per unit
- Production cost: $c_1 = 6$ per unit
- Holding cost: $h_1 = 1$ per unit
- Salvage value: $v_1 = 2$ per unit
- Capacity requirement: $a_1 = 2$ units of capacity per unit of product
- Demand: $D_1 \sim N(\mu_1 = 500, \sigma_1 = 100)$

Product 2:

- Selling price: $p_2 = 15$ per unit
- Production cost: $c_2 = 8$ per unit
- Holding cost: $h_2 = 1$ per unit
- Salvage value: $v_2 = 2$ per unit
- Capacity requirement: $a_2 = 3$ units of capacity per unit of product
- Demand: $D_2 \sim N(\mu_2 = 400, \sigma_2 = 80)$

The capacity investment cost function is quadratic:

$$G(B) = kB^2, \quad \text{with } k = 0.05$$

These parameter values are chosen to reflect realistic production settings where products have different margins, cost structures, and resource requirements. Product 2 has a higher margin but also requires more bottleneck capacity per unit.

4.2 Base Case Results

Solving the optimization problem in Equations (5)-(7) yields the following optimal solution:

Table 3. Base Case Optimal Solution

Variable	Product 1	Product 2	Total
q_i^*	485.6 units	391.2 units	-
$a_i q_i^*$	971.2	1,173.6	2,144.8
B*	-	-	2,144.8
λ*	-	-	$21.45

The shadow price $\lambda^* = 21.4$ is now much lower, and both products are profitable. The critical fractiles are:
For Product 1:

$$F_1(q_1^*) = \frac{120 - 60 - 21.45 \times 2}{120 - 20 + 10} = \frac{60 - 42.90}{110} = 0.1555$$

For Product 2:

$$F_2(q_2^*) = \frac{150 - 80 - 21.45 \times 3}{150 - 25 + 12} = \frac{70 - 64.35}{137} = 0.0412$$

These fractiles correspond to the optimal service levels for each product, accounting for the opportunity cost of capacity.

4.3 Comparative Statics Analysis

We now examine how changes in key parameters affect the optimal solution.

4.3.1 Effect of Holding Cost

We vary the holding cost of Product 1 from $h_1 = 5$ to $h_1 = 25$ while keeping all other parameters fixed at their base values.

Table 4. Effect of Holding Cost on Optimal Solution

h_1	q_1^*	q_2^*	B*	λ*
5	498.2	393.5	2,177.4	\$21.77
10	485.6	391.2	2,144.8	\$21.45
15	473.1	389.0	2,113.2	\$21.13
20	460.8	386.7	2,081.7	\$20.82
25	448.6	384.5	2,051.2	\$20.51

As predicted by Corollary 1, increasing the holding cost reduces the optimal production quantity of Product 1($\partial q_1^*/\partial h_1 < 0$). Interestingly, the production quantity of Product 2 also decreases slightly because the reduced profitability of Product 1 lowers the overall marginal value of capacity, reducing λ^* and B^*. This confirms Corollaries 2 and 3: higher holding costs reduce both the optimal capacity and the shadow price.

4.3.2 Effect of Salvage Value

We vary the salvage value of Product 1 from $v_1 = 5$ to $v_1 = 45$ while keeping all other parameters fixed.

Table 5. Effect of Salvage Value on Optimal Solution

v_1	q_1^*	q_2^*	B*	λ*
5	462.8	388.5	2,091.1	\$20.91
15	474.2	389.9	2,118.1	\$21.18
25	485.6	391.2	2,144.8	\$21.45
35	497.0	392.6	2,171.8	\$21.72
45	508.4	394.0	2,198.8	\$21.99

As predicted by the corrected Corollary 1, increasing the salvage value increases the optimal production quantity of Product 1 ($\partial q_1^*/\partial v_1 > 0$). The production quantity of Product 2 also increases slightly because the higher salvage value improves the profitability of Product 1, increasing the marginal value of capacity and thus (λ^*) and (B^*). This confirms that higher

salvage values increase both optimal capacity and the shadow price (Corollaries 2 and 3).

4.3.3 Effect of Capacity Cost Parameter

We vary the capacity cost parameter from $k = 0.001\ to\ k = 0.015$ while keeping all other parameters fixed.

Table 6. Effect of Capacity Cost on Optimal Solution

k	q_1^*	q_2^*	B*	λ*
0.001	510.2	398.7	2,306.5	$4.61
0.003	497.8	394.9	2,225.4	$13.35
0.005	485.6	391.2	2,144.8	$21.45
0.010	461.3	384.1	1,999.9	$40.00
0.015	438.5	377.3	1,866.9	$56.01

As expected, higher capacity costs reduce optimal capacity B^*, increase the shadow price λ^* since $\lambda^* = G'(B^*) = 2kB^*$, and reduce production quantities for both products. Products with higher capacity consumption Product 2, with $a_2 = 3$ experience a larger reduction in production quantity, consistent with the intuition that capacity-intensive products are more severely affected by capacity scarcity.

4.4 Managerial Insights

The numerical example yields several actionable insights:

1. Capacity investment is a strategic decision: The shadow price $\lambda^* = 21.4$ in the base case represents the marginal value of bottleneck capacity. Managers should invest in capacity expansion as long as the marginal benefit exceeds the marginal cost.
2. Holding costs discourage production and capacity investment: As shown in Table 2, higher holding costs reduce optimal production quantities and, in turn, the optimal bottleneck capacity and its shadow price. Firms with high holding costs should be cautious about capacity expansion.
3. Salvage value encourages production and capacity investment: As shown in Table 3, higher salvage values increase optimal production quantities, capacity, and the shadow price. Developing secondary markets or buy-back agreements can justify larger capacity investments.
4. Capacity-intensive products are more sensitive to capacity costs: Product 2, which requires more capacity per unit, experiences larger percentage reductions in production quantity when capacity costs increase (Table 4). This provides a formal foundation for throughput-based prioritization rules.
5. The capacity constraint binds optimally: In all scenarios, $B^* = \sum a_i q_i^*$, confirming Proposition 3(a). The firm never installs unused capacity.
6. The shadow price and capacity move together: Consistent with the corrected Corollaries 2 and 3, λ^* and B^* move in the same direction across all scenarios, confirming that $\lambda^* = G'(B^*)$ with $G''(B\hat{}\,*) > 0$.

5 Discussion and Managerial Insights

The analysis demonstrates that bottleneck capacity behaves as an endogenous economic resource whose shadow price arises from both technological scarcity and capital costs. Unlike prior studies that treat bottleneck capacity as fixed or exogenously priced, the model shows that optimal capacity emerges from a balance between expected marginal profits and the convex cost of capacity expansion. This endogenization yields a clear economic interpretation of bottleneck investment decisions under demand uncertainty.

The results highlight the critical role of inventory risk in shaping capacity decisions. Holding costs and salvage values enter the optimality conditions symmetrically with production costs and demand uncertainty, effectively altering the marginal value of capacity. As a consequence, firms facing high overage risk optimally restrain bottleneck expansion, while access to salvage or secondary markets encourages more aggressive capacity investment. These findings provide a formal theoretical foundation for several managerial heuristics observed in practice, including throughput-based prioritization and conservative capacity planning in high-uncertainty environments.

This study yields several actionable insights for managers operating production systems with binding bottlenecks and uncertain demand. A central implication of the model is that bottleneck capacity has an endogenous economic value, captured by the shadow price $\lambda^* = G'(B^*)$. Managers should therefore view bottleneck capacity as an investment decision rather than a fixed technological parameter. Instead of asking "*How can we schedule around the bottleneck?*", managers should ask "*At what marginal cost does expanding the bottleneck cease to be profitable?*" This reframing supports capital budgeting decisions that explicitly link demand uncertainty and inventory risk to capacity investment.

Both the analytical results and numerical experiments show that higher holding costs reduce both optimal capacity and the shadow price. When leftover inventory is costly, expanding bottleneck capacity increases exposure to overproduction risk. As a result, firms optimally restrain capacity investment even when demand is high on average. Firms with high obsolescence risk, short product life cycles, or expensive storage requirements should be cautious in expanding bottleneck capacity and may benefit more from flexibility-enhancing investments (e.g., faster changeovers or postponement) than from pure capacity expansion.

The model shows that higher salvage values increase both production quantities and optimal bottleneck capacity. Salvage opportunities reduce downside risk from excess inventory, effectively insuring capacity investments against demand uncertainty. Developing secondary markets, remanufacturing channels, or buy-back agreements can be as effective as reducing production costs when justifying bottleneck expansion. In some settings, market design decisions (creating liquidation options) may dominate process improvement efforts.

Products that consume more bottleneck capacity per unit are penalized more heavily when capacity is expensive. This leads to endogenous product prioritization even when margins are attractive. Product portfolio decisions should account for throughput per bottleneck unit, not just gross margins. High-margin products may be suboptimal if they consume disproportionate bottleneck resources under demand uncertainty. This insight provides a formal economic foundation for common heuristics used in practice, such as throughput accounting and Theory of Constraints prioritization rules.

A key contribution of this paper is showing that capacity and inventory decisions are tightly coupled through the bottleneck shadow price. Separating capacity planning (often strategic) from inventory decisions (often tactical) can lead to systematic inefficiencies. Cross-functional coordination between operations, finance, and marketing is essential when bottlenecks and demand uncertainty coexist.

5.1 Practical Implementation Framework

While the analytical model provides clear theoretical insights, its practical application requires estimation of several key parameters. This subsection provides a step-by-step framework for managers and practitioners to implement the model in real settings.

5.1.1 Estimating the Shadow Price

The shadow price λ^* represents the marginal value of one additional unit of bottleneck capacity. In practice, λ^* can be estimated using one of the following approaches:

(a) Marginal cost approach: From Proposition 2, $\lambda^* = G'(B^*)$. If the capacity investment cost function $G(B)$ is known (or can be estimated from historical data), the shadow price is simply the marginal cost of capacity at the optimal capacity level. For example, if $G(B) = kB^2$, then $\lambda^* = 2kB^*$.

(b) Dual variable from linear programming: In practice, λ^* can be obtained as the dual variable (shadow price) associated with the capacity constraint in a linear programming formulation of the production planning problem. Most commercial optimization solvers provide dual variable values directly.

(c) Sensitivity analysis: Managers can approximate λ^* by calculating the change in expected profit resulting from a small increase in capacity:

$$\lambda^* \approx \frac{\Delta\Pi}{\Delta B}$$

where $\Delta\Pi$ is the increase in expected profit and ΔB is the increase in capacity.

5.1.2 Estimating the Capacity Cost Function

The capacity investment cost function $G(B)$ can be estimated using:

(a) Historical project data: Collect data on past capacity expansion projects, including the total investment cost and the amount of capacity added. Fit a regression model to estimate the functional form. Common functional forms include:

- Linear: $G(B) = g_0 + g_1 B$
- Quadratic: $G(B) = g_0 + g_1 B + g_2 B^2$
- Power: $G(B) = g_0 + g_1 B^{\alpha}), with\ (\alpha > 1)$

(b) Vendor quotes and engineering estimates: For new capacity investments, obtain quotes from equipment suppliers or engineering firms. These estimates can be used to construct the cost function for the relevant range of capacity levels.

(c) Economies of scale consideration: In practice, capacity expansion costs often exhibit economies of scale (initially concave) followed by diminishing returns (convex). Managers should carefully select the functional form that best captures the cost structure for their specific setting. A piecewise linear or quadratic function often provides a good approximation.

5.1.3 Estimating Demand Distributions

The demand distributions $F_i(\cdot)$ for each product can be estimated using:

(a) Historical sales data: Collect historical sales or demand data for each product. Fit parametric distributions using maximum likelihood estimation. For example, if demand is normally distributed, estimate μ_i and σ_i from the sample mean and standard deviation.

(b) Forecasting models: Use time series forecasting models to generate demand forecasts and uncertainty measures. The forecast error distribution can serve as the demand distribution.

(c) Expert judgment: When historical data is limited, use expert judgment to specify the parameters of the demand distribution.

(d) Robust approach: When demand distributions are difficult to estimate with confidence, managers can use the robust optimization approach of Zhang et al. (2021), which does not require exact distributional knowledge.

5.1.4 Step-by-Step Implementation Procedure

The following step-by-step procedure provides a practical roadmap for implementing the model:

Step 1: Collect data

- Historical demand data for each product
- Production costs c_i, selling prices p_i, holding costs h_i, and salvage values v_i
- Capacity requirements a_i for each product
- Historical capacity investment cost data

Step 2: Estimate parameters
- Fit demand distributions $F_i(\cdot)$ using historical data
- Estimate the capacity cost function $G(B)$

Step 3: Solve the optimization problem
- Solve the optimization problem in Equations (5)-(7) using numerical optimization software
- Obtain optimal production quantities q_i^*, optimal capacity B^*, and the shadow price λ^*

Step 4: Validate the results
- Check that the capacity constraint binds: $B^* = \sum a_i q_i^*$
- Verify that $\lambda^* = G'(B^*)$
- Perform sensitivity analysis on key parameters (holding costs, salvage values, capacity cost) as shown in Section 4.3

Step 5: Interpret and act
- Use λ^* to evaluate the marginal value of capacity
- Use q_i^* to set production targets
- Use B^* to guide capacity investment decisions
- Re-evaluate periodically as market conditions change

Step 6: Monitor and update
- Track actual demand and compare with forecasts
- Update demand distribution estimates periodically
- Adjust production and capacity decisions as new information becomes available

5.2 Implications of Risk Preferences

Our analysis assumes a risk-neutral decision maker who maximizes expected profit. However, in practice, managers often exhibit risk aversion, particularly for capacity investment decisions that involve large, irreversible commitments. This subsection discusses how risk preferences would modify the shadow price interpretation and the optimal decisions.

5.2.1 Risk-Averse Objective Functions

A risk-averse manager would optimize a risk-adjusted objective function rather than expected profit. Common approaches include:

(a) **Mean-variance objective**: $\mathbb{E}[\Pi] - \rho \cdot \text{Var}[\Pi]$, where $\rho > 0$ is the risk aversion coefficient. This penalizes profit variability and captures the trade-off between expected return and risk.

(b) **CVaR constraint**: Maximize expected profit subject to a Conditional Value-at-Risk constraint $\text{CVaR}_\alpha(\Pi) \geq \Pi_0$, which ensures that the expected profit in the worst $(1-\alpha) \times 100\%$ of scenarios exceeds a threshold.

(c) **Utility maximization**: Maximize $\mathbb{E}[U(\Pi)]$ where $U(\cdot)$ is a concave utility function (e.g., $U(\Pi) = 1 - e - \rho\Pi$).

5.3.2. Effect on Optimal Decisions

With risk aversion, the effective cost of overage increases because excess inventory not only incurs holding costs but also adds to profit variability. As a result:

- **Optimal production quantities** would be lower than in the risk-neutral case, as the manager would adopt more conservative production targets to avoid downside risk.
- **Optimal capacity investment** would be lower, as capacity expansion increases exposure to demand uncertainty.
- The **shadow price** λ^* would still satisfy $\lambda^* = G'(B^*)$, but it would incorporate a **risk**

premium reflecting the manager's aversion to uncertainty. Specifically, the shadow price would be higher than in the risk-neutral case for any given capacity level, as the manager requires a higher return to justify capacity investment.

5.2.2 Shadow Price Interpretation with Risk Aversion

The generalized critical fractile in Equation (11) would become:

$$F_i(q_i^*) = \frac{p_i - c_i - \lambda^* a_i - \text{Risk Premium}_i}{p_i - v_i + h_i}$$

where Risk Premium $i > 0$ captures the additional cost of uncertainty due to risk aversion. The risk premium depends on the manager's risk aversion coefficient, the variance of demand, and the covariance between products.

For managers facing capacity investment decisions under uncertainty, the following guidelines are relevant:

1. **Recognize the risk premium**: The shadow price λ^* in a risk-averse setting includes a premium for uncertainty. Managers should account for this when evaluating capacity investment projects.
2. **Consider risk mitigation strategies**: Hedging, diversification, or flexible capacity options can reduce the risk premium and justify larger capacity investments.
3. **Use risk-adjusted metrics**: When evaluating capacity investments, use risk-adjusted performance metrics (e.g., CVaR) rather than expected profit alone.
4. **Match risk preferences to investment horizon**: Managers with shorter planning horizons or limited financial flexibility should adopt more conservative capacity investment policies.

These insights highlight the importance of incorporating risk preferences into capacity investment decisions and provide a foundation for extending our model to risk-averse settings.

6 Conclusion

This paper develops an economic model of production bottlenecks that integrates stochastic demand, inventory overage risk, and endogenous capacity investment within a single analytical framework. Extending the classical single-period bottleneck analysis, the model incorporates holding costs and salvage values and allows bottleneck capacity to be chosen optimally through costly investment. By doing so, the analysis unifies production planning, inventory management, and capacity investment decisions that are often studied in isolation.

The results show that bottleneck capacity behaves as a strategic economic resource whose shadow price reflects both technological scarcity and inventory risk. Inventory considerations, traditionally confined to production and ordering decisions, directly influence the marginal value of capacity and thus the optimal level of bottleneck investment. As a consequence, capacity expansion is optimally restrained in environments with high overage risk, while improved salvage opportunities or secondary markets justify more aggressive investment. These findings provide a rigorous economic foundation for observed managerial practices in bottleneck-driven systems and clarify when conservative capacity planning is optimal under demand uncertainty.

By endogenizing bottleneck capacity and explicitly linking it to inventory economics, this paper extends the seminal insights of Lawrence (1995) and contributes to a more comprehensive understanding of constrained production systems. The proposed framework highlights the importance of coordinating production, inventory, and capacity decisions and offers a tractable foundation for future research on dynamic settings, decentralized decision-making, and empirical validation of bottleneck investment behavior under uncertainty.

6.1 Limitations

Despite its analytical tractability and economic clarity, the model relies on several simplifying assumptions. First, the analysis is restricted to a single-period setting, abstracting from intertemporal inventory dynamics and learning effects. Second, demand processes are assumed to be independent across products, excluding potential substitution or complementarity effects. Third, the firm is modeled as a centralized, risk-neutral decision maker, which may not fully capture behavioral biases or decentralized organizational structures commonly observed in practice. These assumptions are deliberately adopted to preserve analytical transparency and isolate the economic mechanisms of interest. Relaxing them offers promising directions for future research.

Our analysis assumes differentiable demand distributions to justify the derivative-based optimization approach. While this is standard in the newsvendor literature and covers a wide range of continuous distributions, it does not directly apply to discrete or empirical demand distributions. Extending the framework to handle non-differentiable distributions using subgradient methods would be a valuable direction for future research, particularly for applications where demand data is discrete or historical.

Our analysis is restricted to a single-period setting, abstracting from intertemporal inventory dynamics, capacity adjustment costs over time, and learning effects. While this simplification is standard in the analytical newsvendor literature and allows us to isolate the core economic mechanisms, it limits the model's applicability to settings where decisions are made repeatedly over time.

The single-period model can be viewed as a stationary stage of a multi-period dynamic program. In a multi-period setting with time-varying parameters, the shadow price interpretation would generalize to $\lambda_t = G_t'(B_t)$ at each period t, where $G_t(\cdot)$ represents the period-specific capacity cost function. The KKT conditions would become time-dependent, with production quantities q_{it} and capacity B_t chosen to maximize the expected net present value of profits over the planning horizon. These extensions are promising directions for future research that would build on the analytical foundation provided by our single-period model.

Our model assumes that capacity investment decisions are made once and that the firm commits to the investment level at the beginning of the period. In practice, capacity investments are often irreversible and involve significant sunk costs. The real options literature (Dixit & Pindyck, 1994; Trigeorgis, 1996) emphasizes that the ability to delay investment creates an option value, as firms can wait for better information about demand conditions before committing to capacity expansion. Incorporating irreversibility and the option to delay into our framework would require a multi-period or continuous-time model in which the firm can choose the optimal timing of investment. This extension would allow the analysis of how demand uncertainty, holding costs, and salvage values affect not only the optimal capacity level but also the optimal investment timing. We leave this extension for future research.

6.2 Future Directions

Several avenues for future research emerge from this study. A natural extension is to embed the model in a multi-period setting with inventory carryover and gradual capacity adjustment, which would allow the analysis of time-varying bottleneck shadow prices, investment timing, and capacity smoothing under persistent uncertainty. Another promising direction is to model bottleneck capacity as stochastic, capturing disruptions due to machine failures, labor shortages, or energy constraints, which may significantly affect both production and investment decisions. Extending the framework to decentralized organizational settings, where production and capacity decisions are made by different units, could yield insights into internal pricing, incentive alignment, and coordination in bottleneck-driven systems. Incorporating behavioral preferences, such as loss aversion or probability weighting, may further explain systematic

under- or overinvestment in capacity. Finally, the model's predictions invite empirical validation using firm- or plant-level data to assess how inventory risk is internalized in strategic bottleneck investment decisions.

While our numerical example illustrates the analytical results and demonstrates the comparative statics, this paper is primarily a theoretical study. Empirical validation of the model's predictions using firm-level or plant-level data is an important direction for future research. We hope that future empirical work will build on the analytical foundation provided here.

Disclosure Statement

No potential conflict of interest was reported by the author.

Data Availability Statement

This theoretical study does not involve the generation or analysis of empirical data. All results are derived analytically from the mathematical model presented in the paper. No datasets were generated or examined during this research. Accordingly, the 'Share Upon Reasonable Request' policy is not applicable to this manuscript.

Funding Statement

No funding was received for conducting this study.

References

Arnott, R., de Palma, A., & Lindsey, R. (1990). Economics of a bottleneck. *Journal of Urban Economics*, *27*(1), 111–130. https://doi.org/10.1016/0094-1190(90)90028-L

Buddas, H. (2014). A bottleneck analysis in the IFRC supply chain. *Journal of Humanitarian Logistics and Supply Chain Management*, *4*(2), 222–244. https://doi.org/10.1108/JHLSCM-10-2013-0036

Dixit, A. K., & Pindyck, R. S. (1994). Investment under Uncertainty. Princeton University Press.

Trigeorgis, L. (1996). Real Options: Managerial Flexibility and Strategy in Resource Allocation. MIT Press.

Fleuren, T., Merzifonluoglu, Y., Sotirov, R., & Hendriks, M. (2025). Production–inventory planning in high-tech low-volume manufacturing supply chains. *International Journal of Production Economics*, *288*, 109687. https://doi.org/10.1016/j.ijpe.2025.109687

Francas, D., Öner-Közen, M., & Minner, S. (2024). Is it beneficial to invest in flexible capacity for hybrid remanufacturing? *International Journal of Production Research*, *62*(21), 7836–7859. https://doi.org/10.1080/00207543.2024.2332640

Gharaei, A., Karimi, M., & Hoseini Shekarabi, S. A. (2020). Joint Economic Lot-sizing in Multi-product Multi-level Integrated Supply Chains: Generalized Benders Decomposition. *International Journal of Systems Science: Operations & Logistics*, *7*(4), 309–325. https://doi.org/10.1080/23302674.2019.1585595

Hamzaoui, A. F., Malek, M., Zied, H., & Sadok, T. (2023). Optimal production and non-periodic maintenance plan for a degrading serial production system with uncertain demand. *International Journal of Systems Science: Operations & Logistics*, *10*(1), 2235808. https://doi.org/10.1080/23302674.2023.2235808

Irma-Delia, R.-C., & Santiago-Omar, C.-M. (2025). Bottleneck Analysis. In J. L. García Alcaraz, G. C. Robles, & A. Realyvásquez Vargas (Eds.), *Lean Manufacturing in Latin America: Concepts, Methodologies and Applications* (pp. 233–260). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-70984-5_11

Lai, X., Shui, H., Ding, D., & Ni, J. (2021). Data-driven dynamic bottleneck detection in complex manufacturing systems. *Journal of Manufacturing Systems*, *60*, 662–675. https://doi.org/10.1016/j.jmsy.2021.07.016

Lawrence, S. R., & Buss, A. H. (1995). Economic analysis of production bottlenecks. *Mathematical Problems in Engineering*, *1*(4), 890420. https://doi.org/10.1155/S1024123X95000202

Mandal, A., & Pal, B. (2021). Optimizing profit for pricing and advertisement sensitive demand under unreliable production system. *International Journal of Systems Science: Operations & Logistics*, *8*(2), 99–118. https://doi.org/10.1080/23302674.2019.1646835

Modak, I., Bardhan, S., & Giri, B. C. (2025). Optimal dynamic pricing and preservation policies under quality-,

stock- and price-sensitive demand with partial backlogging and controllable lead time. *International Journal of Systems Science: Operations & Logistics*, *12*(1), 2462098. https://doi.org/10.1080/23302674.2025.2462098

Murray, C. C., Gosavi, A., & Talukdar, D. (2012). The multi-product price-setting newsvendor with resource capacity constraints. *International Journal of Production Economics*, *138*(1), 148–158. https://doi.org/10.1016/j.ijpe.2012.03.014

Ostovari, A., Bozorgi Amiri, A., Yousefi-Babadi, A., & Benyoucef, L. (2025). Multi-configuration modular capacity wheat supply chain network design under uncertainty: robust optimisation-based approach. *International Journal of Systems Science: Operations & Logistics*, *12*(1), 2472965. https://doi.org/10.1080/23302674.2025.2472965

Subramaniyan, M., Skoogh, A., Salomonsson, H., Bangalore, P., Gopalakrishnan, M., & Sheikh Muhammad, A. (2018). Data-driven algorithm for throughput bottleneck analysis of production systems. *Production & Manufacturing Research*, *6*(1), 225–246. https://doi.org/10.1080/21693277.2018.1496491

Tang, H. (2019). A new method of bottleneck analysis for manufacturing systems. *Manufacturing Letters*, *19*, 21–24. https://doi.org/10.1016/j.mfglet.2019.01.003

Thorstenson, A., & Ramani, V. (2020). Coordination in a supply chain with two manufacturers, two substitute products, and one retailer. *International Journal of Systems Science: Operations & Logistics*, *7*(2), 105–120. https://doi.org/10.1080/23302674.2018.1522015

Van Mieghem, J. A. (2007). Risk Mitigation in Newsvendor Networks: Resource Diversification, Flexibility, Sharing, and Hedging. *Management Science*, *53*(8), 1269–1288. https://doi.org/10.1287/mnsc.1070.0700

Wang, Y., Zhao, Q., & Zheng, D. (2005). Bottlenecks in production networks: An overview. *Journal of Systems Science and Systems Engineering*, *14*(3), 347–363. https://doi.org/10.1007/s11518-006-0198-3

Wu, K., Zheng, M., & Shen, Y. (2020). A generalization of the Theory of Constraints: Choosing the optimal improvement option with consideration of variability and costs. *IISE Transactions*, *52*(3), 276–287. https://doi.org/10.1080/24725854.2019.1632503

Yang, H.-L. (2021). Retailer's ordering policy for demand depending on the expiration date with limited storage capacity under supplier credits linked to order quantity and discounted cash flow. *International Journal of Systems Science: Operations & Logistics*, *8*(2), 136–153. https://doi.org/10.1080/23302674.2019.1701727

Yin, S., Nishi, T., & Zhang, G. (2016). A game theoretic model for coordination of single manufacturer and multiple suppliers with quality variations under uncertain demands. *International Journal of Systems Science: Operations & Logistics*, *3*(2), 79–91. https://doi.org/10.1080/23302674.2015.1050079

Yu, M., Yang, Y., Zhu, H., & Cao, T. (2026). The data-driven multi-item newsvendor problem with resource constraints. *OR Spectrum*. https://doi.org/10.1007/s00291-025-00846-7

Zhang, J., Xie, W., & Sarin, S. C. (2021). Robust multi-product newsvendor model with uncertain demand and substitution. *European Journal of Operational Research*, *293*(1), 190–202. https://doi.org/10.1016/j.ejor.2020.12.023



This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license.