



Research article

Systemic stockout risk in e-commerce fulfillment networks: A fractional stochastic control approach

Hui Sun*

School of Economics and Management, Shanghai Open University, Shanghai 200433, China

* **Correspondence:** Email: sunh@sou.edu.cn.

Abstract: E-commerce platforms increasingly rely on multi-warehouse fulfillment networks, which speed delivery but also expose them to systemic stockout risk, the network-wide loss of fulfillment capability caused by persistent demand shocks. Such persistence is common in digital commerce, where viral promotions and platform-wide campaigns produce demand surges with long memory that Markovian models cannot capture. We embedded long-memory demand pressure, modeled by fractional Brownian motion, in a continuous-time inventory network and obtained four results. First, we derived a closed-form expression for the systemic stockout risk of a benchmark network, defined as the probability that its average inventory of a critical item ends the period below its safety threshold, and identified the exact conditions under which this risk rises with demand persistence, the time horizon, and volatility. Its validity was proven rather than assumed; the result is exact under a common shock, asymptotically exact under correlated shocks, and with a quantified error under independent shocks. Companion measures capture the expected shortfall and the risk of breaching the threshold at any time, not only at the period's end. Second, the internal work of keeping the network balanced can be split into two quantities that should not be confused: a pressure index measuring how unevenly stock is spread and the actual volume of goods moved between warehouses. This distinction separates stock health from the logistics cost of maintaining it. Third, within a class of feedback policies based on the observed inventory gap, we obtained the optimal emergency-replenishment rule, which acts earlier and harder when demand is more persistent and is proved optimal by an elementary argument that sidesteps the difficulties associated with fractional noise. Fourth, when demand shocks are shared across warehouses, the network's overall risk does not fall to zero as it grows but settles at a strictly positive floor, so adding or diversifying warehouses removes location-specific risk but not the risk shared across all locations. Overall, the paper provides e-commerce platforms with a practical framework for quantifying systemic stockout risk under long-memory demand, calibrating that memory from data, and dynamically deploying emergency fulfillment capacity through a computable optimal feedback rule.

Keywords: e-commerce fulfillment networks; systemic stockout risk; long-memory demand; fractional brownian motion; linear-quadratic optimization

Mathematics Subject Classification: 49K15, 90B60, 49N90

1. Introduction

The rapid growth of e-commerce has reshaped consumers' expectations of delivery speed and product availability. Global retail e-commerce sales reached about USD 4.4 trillion in 2023, roughly 20% of total retail, and are projected to reach USD 6.8 trillion by 2028 [1]. To honor next-day and same-day delivery promises, platforms such as Amazon, JD.com, and Alibaba have moved from centralized warehouses to networks of fulfillment centers positioned near customers [2, 3]. Amazon alone expanded its U.S. same-day network by over 60% in 2024, covering more than 140 metropolitan areas, with about 60% of Prime orders in the top 60 U.S. cities delivered the same or next day [4, 5]. These multi-warehouse fulfillment networks have become the operational backbone of modern e-commerce.

The distributed architecture that speeds delivery also introduces a new and more complex vulnerability. Stockouts are costly: they are estimated to cost global retailers nearly USD 1 trillion in annual sales [6], and nearly one-third of consumers who encounter a stockout abandon the purchase or switch retailers [7], so that a single stockout episode erodes long-term customer lifetime value [8]. Crucially, social-media-driven demand is not only volatile but persistent: TikTok Shop's viral-content-plus-instant-purchase model grew U.S. sales by roughly 120% year-over-year in 2025 [9, 10], and live-streaming commerce exhibits herd dynamics in which past demand surges predict future ones [11, 12]. When a network-wide, persistent surge exhausts both local inventories and transshipment capacity, the central vulnerability is no longer the isolated stockout of a single warehouse but the joint failure of fulfillment capability across the network at once, which is called systemic stockout risk.

The literature leaves three gaps in the analysis of this risk. First, quantitative supply chain risk research concentrates on supply-side disruptions, such as supplier failures, port congestion, and natural disasters, and on how they propagate [13, 14], whereas systemic risk originating from demand-side persistence, transmitted through inventory-balancing linkages, remains unmodeled. Second, analytical inventory theory relies predominantly on memoryless stochastic primitives such as Poisson demand, geometric Brownian motion, and martingale or Lévy formulations [15–17], while empirical work documents long-range dependence in demand driven by herd behavior and social diffusion [11, 12, 18, 19] and in retail sales series [20]; models with short memory understate both the duration and the severity of demand episodes. Third, the workhorse control tools of operations management, dynamic programming and Hamilton–Jacobi–Bellman equations, require the Markov property [21, 22]; outside the classical memoryless boundary, fractional Brownian motion is neither Markovian nor a semimartingale, so these tools do not apply [23].

We address the three gaps with a single continuous-time model. Inventory at each warehouse evolves under five separately specified forces: routine inbound replenishment, lateral transshipment that redistributes stock toward the network mean, centralized emergency replenishment chosen by the platform, nondecreasing baseline sales demand, and a signed persistent net-flow mismatch that captures deviations around baseline demand pressure. All stochastic analysis is carried out under the Wick–Itô–Skorokhod integral of fractional stochastic calculus [24–26], stated explicitly as a modeling assumption; this convention preserves zero-mean stochastic integrals, yields closed-form lognormal-type solutions, and approaches the classical Itô framework at the memoryless boundary.

The paper makes four contributions. First, at the modeling level, it formulates systemic stockout risk as a property of network-average inventory under long-memory net-flow pressure and makes the

aggregation step explicit: transshipment cancels exactly in the network average; independent warehouse shocks obey an inverse-square-root pooling rule; and the closed-form risk expression is stated for a mean-field benchmark whose relation to the full finite network is established rather than assumed, holding exactly for common shocks, asymptotically for correlated shocks, and with quantified Monte Carlo error for independent shocks. Second, at the risk-measurement level, it complements terminal exposure with conditional comparative statics, a closed-form expected-shortfall expression of Black–Scholes put type, and a numerically characterized first-passage exposure that the terminal measure bounds from below. In particular, greater persistence raises stockout exposure when the evaluation period exceeds the model’s unit time scale and the drift-adjusted inventory buffer remains nonnegative. Third, at the decision level, it solves the platform’s emergency replenishment problem as a fractional linear-quadratic control problem over linear feedback policies and obtains an explicit time-varying feedback rule whose Riccati coefficient acts as a shadow price of inventory deviation; a second-moment reduction makes the optimality argument exact and self-contained. Fourth, at the network-design level, a large-network theorem for correlated shocks shows that aggregate risk converges to a strictly positive floor governed by the common demand component, qualifying the classical risk-pooling logic of warehouse diversification: pooling eliminates idiosyncratic demand risk but cannot remove the common component.

The remainder of the paper is organized as follows: Section 2 reviews the related literature. Section 3 presents the model, the stochastic-calculus convention, the systemic stockout exposure results, and the transshipment-activity results. Section 4 solves the emergency replenishment problem. Section 5 develops the mixed-shock and correlated-shock extensions, including the large-network theorem. Section 6 reports the numerical validation and calibration guidance. Section 7 concludes the work. All proofs are collected in Appendix A.

2. Literature review

This research draws on three streams: inventory and fulfillment management in e-commerce networks, long-memory demand modeling, and stochastic optimal control in operations management.

2.1. E-commerce inventory networks and systemic risk

E-commerce has transformed inventory operations. Boysen et al. [27] reviewed warehousing systems for online retail and documented the shift from full-pallet restocking to frequent small-batch picking; Risberg [28] developed a decision framework for e-commerce and omnichannel operations covering site selection, inventory, and distribution, highlighting the trade-off between fulfillment speed and operating cost. To expedite delivery, platforms operate many distributed fulfillment centers: Yang et al. [29] studied network planning for quick commerce with small local fulfillment centers, and Wang and Minner [30] optimized inventory placement and order splitting using online sales and website activity data. Research on multi-warehouse networks focuses on positioning inventory near customers [30], balancing online and offline channels [28, 31], and lateral transshipment to mitigate local shortages [32, 33]. Paterson et al. [32] reviewed lateral transshipment models and documented their effectiveness in reducing local stockouts and safety stock; Ekren et al. [33] embedded inventory sharing in an IoT-enabled fresh-food e-commerce model.

As networks grow more complex, risk management becomes central. A substantial literature has examined disruption propagation [34–36]: Li and Zobel [34] assessed supply chain resilience under

ripple effects, Habibi et al. [36] combined network structure and operations to stress-test multi-tier chains, Habibi et al. [37] analyzed how network structure and connection strength shape risk propagation, and Colon and Hochrainer-Stigler [35] argued for system-level governance of supply chain risk by an analogy with financial systems. Ma et al. [38] reviewed network-based resilience modeling and noted the field's focus on identifying critical nodes and testing redundancy and flexibility strategies [39]. This body of work, however, addressed predominantly supply-side disruptions such as factory shutdowns and port delays, together with their propagation; recent optimization-based network design work extended the toolbox to energy-aware dual-channel agricultural networks [40], robust biodiesel supply networks under supply and demand uncertainty [41], and circular supply chain network design [42], but the risk primitive remains a discrete disruption rather than persistent demand pressure.

On the demand side, empirical studies show that stockouts damage customer sentiment and platform reputation: attribution of blame shapes post-purchase behavior [43], shopping mindset and site familiarity moderate switching [44], stockout justifications affect satisfaction and repurchase [45], and unexpected stockouts raise consumer anxiety [46]. In platform ecosystems such as Amazon FBA, where the platform centrally controls storage and shipping across locations [27, 30], localized shortages can therefore escalate into network-level reputational damage.

A gap remains between these two bodies of work. In finance, systemic risk denotes large system-wide losses and is quantified by measures such as CoVaR and marginal expected shortfall [47, 48]. E-commerce inventory risk models, by contrast, concentrate on discrete supply shocks and do not define or quantify the systemic risk that arises when persistent demand amplification propagates through inventory-balancing linkages. This paper formalizes that risk as the probability that network-average inventory falls below a safety threshold, characterizes how it depends on demand persistence, and provides closed-form metrics for the internal logistics burden of network rebalancing.

2.2. Long memory in demand modeling

Stochastic inventory models depend on the demand primitive. Classical formulations assume memoryless demand: Poisson processes and their compound variants [15, 16], geometric Brownian motion for SKU life cycles [17, 49], martingale demand [50, 51], and Lévy demand [52]. In all of these, future demand depends on the past only through the current state.

For e-commerce demand, this memoryless assumption is empirically fragile. Herd behavior in online book purchases [18], informational cascades in review adoption [19], real-time interaction effects in live-streaming commerce [11], and reviewer-badge dynamics [12] all generate demand whose shocks persist: a successful campaign or viral event lifts demand for weeks [8, 18]. Econometrics measures such as persistence through the Hurst exponent: values above one half indicate long memory [53–55], and the fractional Brownian motion is the canonical continuous-time process exhibiting long-range dependence with persistent, positively correlated increments in that regime [56–58]. Direct evidence of long memory in retail sales series was provided by [20], who documented seasonal long memory in G7 retail sales, and ARFIMA-type long-memory specifications have proven useful in forecasting economic time series more broadly [59]. On the calibration side, Yerlikaya-Özkurt et al. [60] proposed a conic multivariate adaptive regression splines approach to estimate the Hurst parameter of fractional-Brownian-motion-driven stochastic differential equations; their focus was estimation, while ours is optimization and control once persistence has been estimated. Section 6.4 reviews the standard estimation procedures [61–63] that connect the two.

Data-driven operations research has, meanwhile, improved demand modeling through machine learning: cross-selling and substitution-aware demand estimation for category management [64], sequence models for day-store-product sales forecasting [65], recurrent networks for material demand in manufacturing [66], and demand estimation combined with robust optimization for product design [67]. Hybrid machine-learning-plus-metaheuristic approaches have also advanced adjacent operational optimization problems, such as robust project scheduling in Industry 4.0 environments [68], fairness-constrained operating-room assignment [69], and renewable energy portfolio selection [70]. These methods improve forecast accuracy and decision robustness, but they treat demand as a short-memory or finite-horizon series and do not embed long-range dependence in an inventory control framework [64–66].

The resulting gap is specific: empirical work documents long memory in e-commerce demand, and mathematical finance has long used fractional Brownian motion for long-range dependent prices and volatility [22, 71, 72], yet analytical inventory control retains Markovian primitives largely because they permit dynamic programming and standard Itô calculus. Fractional Brownian motion is neither Markovian nor a semimartingale, so importing it requires different machinery. We use Wick–Itô–Skorokhod fractional calculus [24–26] and linear-quadratic control theory to close this gap: fractional Brownian motion enters a multi-warehouse inventory network as the driver of persistent net-flow pressure, closed-form systemic risk expressions are derived and validated, and the control problem is solved without Markovian assumptions.

2.3. Stochastic optimal control in operations management

Operations management relies on dynamic programming (DP) for discrete-time decisions and on HJB equations for continuous-time ones. De Souza et al. [73] reviewed 278 inventory studies built on Bellman equations; Zhang et al. [74] applied DP to inventory relocation in disaster relief. In continuous time, Wu and Chao [75] used HJB methods for average-cost production-inventory control, [76] for production planning under output and price uncertainty, and [77] for high-dimensional newsvendor-type problems; joint pricing-inventory studies used HJB equations or the maximum principle for perishable products [78], fashion retail [79, 80], and risk-constrained settings [81]. All of these require the current state to summarize history, which is precisely the Markov property.

Linear-quadratic control is the structured alternative: with linear dynamics and quadratic costs, optimal policies are linear feedback rules computed from Riccati equations. Başar et al. [82] surveyed nonstandard LQ decision problems; Taskesen et al. [83] developed distributionally robust LQ control; Abi Jaber et al. [84] extended LQ theory to stochastic Volterra systems with memory kernels; and Zimbidis [85] applied LQ methods with fractional noise to insurance pricing. For fBm-driven systems specifically, Kleptsyna et al. [86] established the solvability of the LQ regulator under fractional Brownian perturbation, and Duncan et al. [87] and Han and Sun [88] developed LQ control for fBm-driven systems in Hilbert-space and finite-dimensional settings; Liu et al. [89] treated two-person zero-sum LQ games with Markov switching and fBm. Related work on regime-switching stochastic control and anticipation [90, 91] similarly treated uncertainty structure as integral to the optimization model. These advances remain concentrated in finance, insurance, and engineering.

Within operations management, memory effects have been introduced mainly through fractional calculus in deterministic or simulation-based inventory models: Caputo-derivative inventory models with partial backlogging [92], reliability- and memory-dependent deterioration [93], learning-and-memory

effects under uncertainty [94], and memory-dependent inventory-pricing models [95, 96]. These studies confirm that memory matters for inventory outcomes but stop short of optimal control under long-range dependent stochastic demand.

We connect the two literatures. Using the fractional linear-quadratic machinery of [86–88] under the Wick–Itô–Skorokhod convention, we formulate fractional-Brownian-motion-driven linear state equations and quadratic costs for emergency replenishment in a fulfillment network, prove the existence and uniqueness of the optimal policy within the class of linear feedback rules, and characterize the optimal feedback coefficient through a Riccati equation with a time-varying memory term determined by volatility, persistence, and elapsed time. The managerial output is a computable rule that links the observed inventory gap, the emergency logistics cost, and the persistence of demand shocks; the optimality condition, feedback coefficient, and minimum cost are all in closed form up to a scalar ordinary differential equation.

3. Systemic stockout risk and inter-warehouse transshipment in e-commerce platforms

3.1. Model setting

We consider an e-commerce platform operating an inventory network of N fulfillment centers, indexed by $\mathcal{N} := \{1, 2, \dots, N\}$, on a complete probability space $(\Omega, \mathcal{F}, \mathbb{F} := (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ carrying N independent standard fractional Brownian motions $(B^{H,i})_{i \in \mathcal{N}}$ with common Hurst exponent $H \in (1/2, 1)$; $\mathbb{F}$ is the natural filtration they generate, augmented in the usual way. Let $Z_t^{(i)} > 0$ denote the inventory position of a critical SKU at warehouse i at time t . Initial inventories $Z_0^{(i)} = \xi^{(i)} > 0$ are deterministic. The risk assessment is made at time 0 by a platform that observes its own warehouse levels, so all probabilities below are conditional on the observed initial state. (If initial inventories were random and independent of the shocks, every expression below would hold conditionally on ξ , and unconditional stockout exposures would follow by averaging over the distribution of ξ .)

Inventory flow decomposition. Inventory changes at warehouse i arise from five separately specified components. They comprise routine inbound replenishment $R_t^{(i)}$, lateral transshipment within the network, centralized emergency replenishment $u_t^{(i)}$, nondecreasing baseline sales demand $D_t^{(i)}$, and a signed persistent net-flow mismatch $M_t^{(i)}$. Their combined effect is represented by

$$dZ_t^{(i)} = \underbrace{\frac{\kappa}{N} \sum_{j=1}^N (Z_t^{(j)} - Z_t^{(i)}) dt}_{\text{lateral transshipment}} + \underbrace{u_t^{(i)} dt}_{\text{emergency control}} + \underbrace{dR_t^{(i)}}_{\text{routine inbound}} - \underbrace{dD_t^{(i)}}_{\text{baseline sales}} - \underbrace{dM_t^{(i)}}_{\text{signed net-flow mismatch}}. \quad (3.1)$$

Routine inbound replenishment tracks the sales velocity of the warehouse, as automatic reorder policies do in practice, so we specify $dR_t^{(i)} = \mu_r Z_t^{(i)} dt$ with $\mu_r \geq 0$. Baseline cumulative sales demand is nondecreasing, while persistent deviations from the baseline are represented separately by the following signed net-flow mismatch.

$$dD_t^{(i)} = \mu_d Z_t^{(i)} dt, \quad dM_t^{(i)} = \sigma Z_t^{(i)} d\tilde{B}_t^{H,i}, \quad \mu_d > 0, \sigma > 0. \quad (3.2)$$

Writing $\nu := \mu_r - \mu_d$ for the net baseline imbalance rate and $B^{H,i} := -\tilde{B}^{H,i}$ (also a standard fBm), (3.1) becomes the state equation used throughout.

$$dZ_t^{(i)} = \kappa(\bar{Z}_t - Z_t^{(i)})dt + u_t^{(i)}dt + \nu Z_t^{(i)}dt + \sigma Z_t^{(i)}dB_t^{H,i}, \quad \bar{Z}_t := \frac{1}{N} \sum_{i=1}^N Z_t^{(i)}. \quad (3.3)$$

The flow-balanced case $\nu = 0$, in which routine replenishment is calibrated to offset baseline demand, serves as our benchmark regime and is the case illustrated in all figures; all formal results are stated for general ν .

We comment on each component of (3.3) in turn.

(i) Lateral transshipment. The term $\kappa(\bar{Z}_t - Z_t^{(i)})dt$ represents the platform's routine automatic rebalancing. Warehouses below the network average receive inventory, warehouses above it ship inventory out, and $\kappa \geq 0$ measures the speed of this internal redistribution. This mean-field form of the interaction is what makes the network tractable, and its operational preconditions are discussed under "Scope" below.

(ii) Emergency replenishment control. The control $u_t^{(i)}$ captures extraordinary, costly interventions such as cross-regional air freight and premium procurement, which the platform activates deliberately. Section 4 derives the optimal policy; until then, we study the uncontrolled network, $u^{(i)} \equiv 0$, to isolate its inherent risk profile.

(iii) Persistent net-flow mismatch shocks. The multiplicative noise $\sigma Z_t^{(i)}dB_t^{H,i}$ embodies three modeling choices. First, the absolute size of deviations from baseline demand pressure scales with the inventory position, so high-velocity SKUs with large stock footprints experience proportionally larger absolute swings while sales remain availability-constrained. Second, the driver is fBm with $H > 1/2$, so net-flow mismatch increments are positively correlated over time. A viral surge or persistent forecast error today, therefore, makes elevated demand pressure tomorrow more likely, consistent with the herd and diffusion behavior documented empirically [11, 12, 18, 19]. Third, the multiplicative-geometric structure keeps $Z_t^{(i)} > 0$ automatically (see Assumption 1 below), so the nonnegativity of inventory holds by construction rather than by an added constraint.

Three boundary features of the specification deserve explicit statement. (a) Baseline cumulative sales demand $D^{(i)}$ is pathwise nondecreasing because $dD_t^{(i)} = \mu_d Z_t^{(i)}dt > 0$. The fractional term $M^{(i)}$ is deliberately not interpreted as cumulative physical sales. Instead, it is a signed diffusion approximation to deviations around baseline net inventory pressure, including forecast errors, returns and cancellations, and demand–replenishment timing mismatches. A negative increment therefore represents an inventory-relieving deviation from baseline, not negative customer purchases. (b) The model is a diffusion-type approximation of order-flow pressure, appropriate at the aggregate SKU-network level rather than at the level of individual orders; jump-type formulations (Poisson, Hawkes) are compared in Section 6.4. (c) Replenishment lead times are compressed into the flow rates. Routine lead times are folded into μ_r , and emergency replenishment is modeled as a rate control, which is appropriate when expedited lead times are short relative to the evaluation horizon T . Explicit stochastic lead times, lost-sales accounting below the stockout threshold, and backlogging are outside the scope of this model and are flagged as limitations in Section 7.

Scope and modeling philosophy. The framework is macroscopic and describes the aggregate inventory dynamics of a fulfillment network under persistent demand pressure rather than the microscopic mechanics of individual purchases. This is also why fBm is the appropriate driver, since it is the canonical parsimonious model of macroscopic temporal persistence [56–58], while micro-level purchase mechanisms (reviews, recommendations, livestream interactions) enter only through the

persistence they induce. The mean-field transshipment term presumes (a) centralized inventory control with network-wide visibility, (b) regionally comparable warehouses serving statistically similar demand, and (c) rebalancing that is frequent and small relative to stock levels. These conditions are met by large centrally coordinated platforms (e.g., platform-fulfilled networks of the FBA type [27, 30]) but not necessarily by federations of independent sellers or by sparse two-echelon networks with rare bulk transfers. Similarly, the long-memory driver is an empirical hypothesis to be tested before deployment. Section 6.4 describes how to estimate H from demand data and when a short-memory or jump model should be preferred. These primitives are shaped by platform-coordinated e-commerce fulfillment. The long-memory driver reflects the viral, social-media, and campaign-driven persistence specific to digital-commerce demand; the mean-field transshipment term encodes the continuous cross-warehouse rebalancing that platform-fulfilled networks perform; and the network-average state is the systemic, fulfillment-wide stockout risk that sets a distributed e-commerce network apart from single-facility inventory control. Equations (3.1)–(3.3) thus embody the operational features of the e-commerce setting. At the same time, because that structure rests on mean-field balancing under persistent aggregate shocks rather than on anything specific to retail, the same equations describe other centrally coordinated resource networks as well, so e-commerce is the motivating application rather than the exclusive domain of validity.

Stochastic-calculus convention. Because fBm with $H \neq 1/2$ is not a semimartingale, the stochastic integral in (3.3) must be specified before any solution formula is meaningful.

Assumption 1 (WIS convention). *Stochastic integrals with respect to $B^{H,i}$ are Wick–Itô–Skorokhod (divergence-type) integrals in the sense of the fractional stochastic calculus of [24] and [25, 26]. Associated with H is the kernel $\varphi(t, s) := H(2H - 1)|t - s|^{2H-2}$, for which $\int_0^t \varphi(t, s) ds = Ht^{2H-1}$. Under this convention, (a) WIS integrals of admissible integrands have zero expectation, and (b) the linear equation $dZ_t = \nu Z_t dt + \sigma Z_t dB_t^H$ has the unique WIS solution*

$$Z_t = Z_0 \exp\left(\nu t + \sigma B_t^H - \frac{1}{2} \sigma^2 t^{2H}\right), \quad (3.4)$$

so that $\mathbb{E}[Z_t] = Z_0 e^{\nu t}$; and (c) for $H > 1/2$, (3.4) coincides pathwise with the Young (pathwise) solution of the corrected equation $dZ_t = (\nu - \sigma^2 H t^{2H-1}) Z_t dt + \sigma Z_t dB_t^H$, which is the representation used by our simulation schemes.

Three reasons motivate the WIS choice over the pathwise (Young/Stratonovich-type) alternative, under which the solution would be $Z_0 \exp(\nu t + \sigma B_t^H)$ without the correction term. First, the zero-expectation property (a) is what makes probabilities and moments computable in closed form. Second, as $H \rightarrow 1/2$, the WIS solution (3.4) converges to the classical Itô/geometric-Brownian solution, so all results nest the standard Markovian benchmark continuously. Third, under the pathwise convention, $\mathbb{E}[Z_t] = Z_0 e^{\nu t + \frac{1}{2} \sigma^2 t^{2H}}$ grows with the noise even in the flow-balanced regime, an economically spurious “volatility creates inventory” effect. Throughout the paper, “unique solution” means unique solution in the WIS sense of Assumption 1, and every proof states where the convention is used.

Systemic stockout exposure. Let $Z_{\min} > 0$ be the critical threshold below which the network can no longer sustain its fulfillment promises. Individual warehouse stockouts are routine; the event of interest is the network-level one.

Definition 1 (terminal systemic stockout exposure). *The terminal systemic stockout exposure at horizon T is $\text{SSE}(T) := \mathbb{P}\{\bar{Z}_T \leq Z_{\min}\}$.*

The terminal exposure is a deliberately conservative, closed-form benchmark because it registers whether the network ends the evaluation period below the critical level, not whether it visited that level earlier. The corresponding first-passage exposure $\mathbb{P}\{\min_{0 \leq t \leq T} \bar{Z}_t \leq Z_{\min}\}$ is at least as large (Remark 1), does not admit a comparable closed form under fBm, and is quantified numerically in Section 6.2. We accordingly interpret $\text{SSE}(T)$ as a lower bound on network distress over $[0, T]$ and avoid describing it as a “collapse” probability.

3.2. Systemic stockout risk

The first step is exact and does not involve any approximation.

Lemma 1 (aggregation). *Let $(Z^{(i)})_{i \in N}$ solve (3.3) with $u^{(i)} \equiv 0$. Then the network average satisfies, exactly and for every $\kappa \geq 0$,*

$$d\bar{Z}_t = v\bar{Z}_t dt + \sigma \frac{1}{N} \sum_{i=1}^N Z_t^{(i)} dB_t^{H,i}, \quad (3.5)$$

because the transshipment terms cancel identically, as shown by $\sum_{i=1}^N (\bar{Z}_t - Z_t^{(i)}) = 0$. Moreover (i) $\bar{B}_t^H := N^{-1} \sum_{i=1}^N B_t^{H,i}$ is a centered Gaussian process with $\text{Var}(\bar{B}_t^H) = t^{2H}/N$, hence equal in law to $N^{-1/2} \hat{B}_t^H$ for a standard fBm $\hat{B}^H$, and (ii) under Assumption 1, $\mathbb{E}[\bar{Z}_t] = \bar{Z}_0 e^{vt}$ for every $\kappa \geq 0$.

Lemma 1 already delivers two structural conclusions. Internal transshipment redistributes but never creates inventory, so κ is absent from the exact aggregate equation (3.5), which places the systemic level of the network beyond the reach of rebalancing. And averaging N independent shocks compresses volatility by exactly $\sqrt{N}$, which motivates the following explicit definition.

Definition 2 (mean-field benchmark and network-effective volatility). *The network-effective volatility is $\sigma_N := \sigma / \sqrt{N}$. The mean-field benchmark process $\hat{Z}$ is the unique WIS solution of*

$$d\hat{Z}_t = v\hat{Z}_t dt + \sigma_N \hat{Z}_t d\hat{B}_t^H, \quad \hat{Z}_0 = \bar{Z}_0. \quad (3.6)$$

The benchmark replaces the exact diffusion term of (3.5) by $\sigma \bar{Z}_t d\bar{B}_t^H$, which is the classical mean-field closure. The closure is *not* an identity. The replaced term differs from the true one by the remainder $\mathcal{R}_t := \sigma N^{-1} \sum_i \int_0^t (Z_s^{(i)} - \bar{Z}_s) dB_s^{H,i}$, which has zero expectation under Assumption 1 and whose magnitude is governed by the cross-warehouse dispersion of inventories (the same dispersion index $D_N(t)$ that drives transshipment activity in Section 3.3). Three regimes make the closure quality precise. (a) If all warehouses are driven by a common shock ($B^{H,i} \equiv B^H$ for all i , the fully systemic case of Section 5.2 with $\rho = 1$), the closure is exact for every N and every κ . Appendix A.2 exhibits the explicit solution of the coupled system, whose average is exactly geometric fBm. (b) Under partially correlated shocks, the benchmark error vanishes as $N \rightarrow \infty$ at rate $N^{-1/2}$ (Theorem 1 in Section 5.2). (c) Under fully independent shocks, the benchmark is accurate when per-warehouse cumulative volatility $\sigma^2 T^{2H}$ is moderate and initial inventories are approximately balanced, and it understates aggregate variance otherwise; Appendix A.2 gives the exact aggregate-variance identity behind this statement, and Section 6.1 quantifies the resulting probability error by Monte Carlo against the full N -warehouse system. All statements about the closed form below are statements about the benchmark process, with the exactness scope just described.

Proposition 1 (terminal systemic stockout exposure). *Under Assumption 1, for the mean-field benchmark (3.6) with deterministic $\bar{Z}_0 > 0$,*

$$\mathbb{P}\{\hat{Z}_T \leq Z_{\min}\} = \Phi\left(\frac{\ln(Z_{\min}/\bar{Z}_0) - \nu T + \frac{1}{2}\sigma_N^2 T^{2H}}{\sigma_N T^H}\right), \quad (3.7)$$

where Φ is the standard normal distribution function. The same expression with σ_N replaced by σ gives the exact terminal exposure of the full coupled network, for every N and κ , when all warehouses share a common demand shock.

The proof is given in Appendix A.2.

Analysis of Proposition 1. The closed form in Proposition 1 reduces the network's systemic risk to a single normal-tail probability, and the four primitives that drive expression (3.7) each carry a distinct managerial reading. *Demand memory* operates when the horizon exceeds the unit time scale and the drift-adjusted initial buffer remains nonnegative. A larger Hurst exponent then inflates the effective standard deviation $\sigma_N T^H$ of cumulative shocks, so persistence raises systemic stockout exposure, as Figure 1 illustrates, and Corollary 1 states both conditions under which this monotonicity is a theorem rather than a numerical pattern. The comparison with the Markovian benchmark is quantitative. At $H = 1/2$, the variance of cumulative shocks grows linearly in T , while at $H > 1/2$, it grows like T^{2H} , so within that buffer regime, safety stocks calibrated under memoryless models understate tail exposure for viral, campaign-driven categories, and the understatement worsens with the horizon. *The horizon effect* makes stockout exposure accumulate with T under the conditions in Corollary 1(ii) (Figure 2), which rationalizes shorter risk-review cycles for long-memory categories. *The initial-buffer effect* makes stockout exposure fall with $\bar{Z}_0$ (Figure 3), quantifying the protective value of pre-positioned inventory before launch or campaign events. *The net-flow-balance effect* shifts the entire stockout exposure schedule upward through the $-\nu T$ term in a demand-dominant regime ($\nu < 0$), because persistent under-replenishment and persistent shocks compound. Relative to the disruption-propagation literature [34, 36, 37], the mechanism here is not the failure of a node but the slow, correlated depletion of the whole network. Risk is generated by the demand side and merely *redistributed* by transshipment, not reduced by it (Lemma 1).

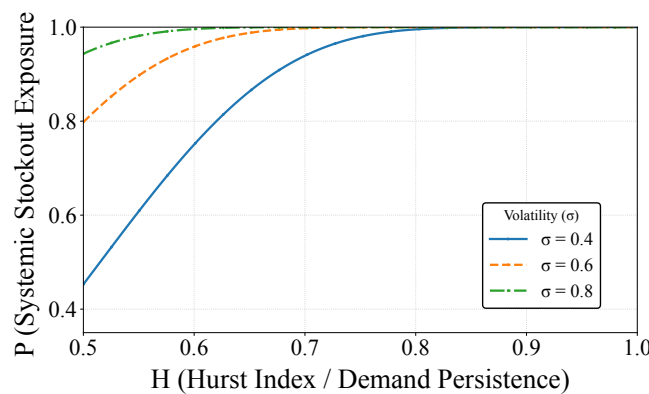


Figure 1. Terminal systemic stockout exposure P versus Hurst exponent H (baseline model, Proposition 1). The parameter setting uses $Z_{\min} = 0.7$, $\bar{Z}_0 = 10$, $T = 30$, $\nu = 0$, and network-effective volatility $\sigma_N \in \{0.4, 0.6, 0.8\}$. Since $\bar{Z}_0 > Z_{\min}$ and $T > 1$, all curves are strictly increasing in H (Corollary 1(iii)).

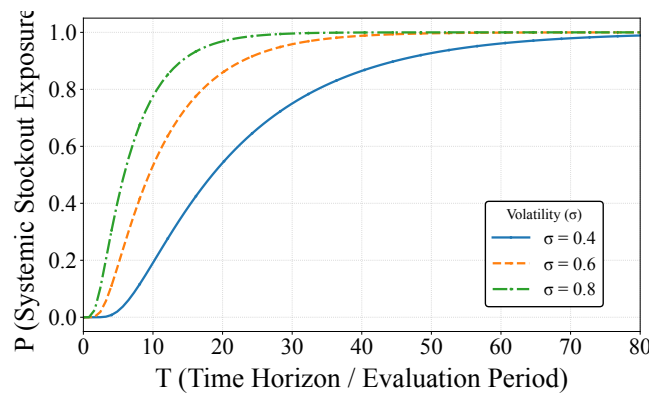


Figure 2. Terminal systemic stockout exposure P versus horizon T (baseline model). The parameter setting uses $Z_{\min} = 0.7$, $\bar{Z}_0 = 10$, $H = 0.6$, $\nu = 0$, and $\sigma_N \in \{0.4, 0.6, 0.8\}$. Since $\bar{Z}_0 > Z_{\min}$, stockout exposure is strictly increasing in T (Corollary 1(ii)).

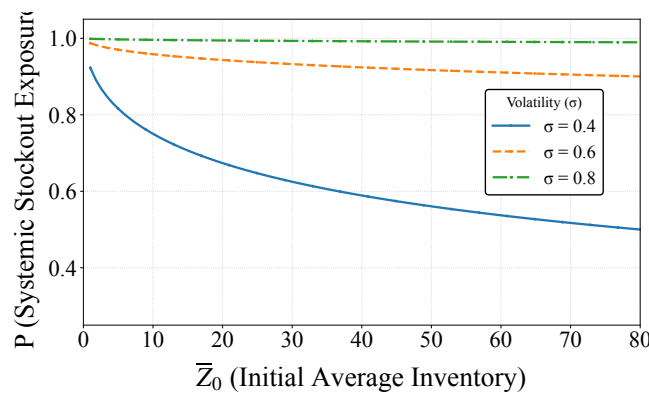


Figure 3. Terminal systemic stockout exposure P versus initial average inventory $\bar{Z}_0$ (baseline model). The parameter setting uses $Z_{\min} = 0.7$, $T = 30$, $H = 0.6$, $\nu = 0$, and $\sigma_N \in \{0.4, 0.6, 0.8\}$. Stockout exposure is strictly decreasing in $\bar{Z}_0$ (Corollary 1(i)).

The monotonicity patterns visible in Figures 1, 2, and 3 hold as theorems only under explicit parameter conditions, which we now state exactly; numerical illustrations are not a substitute for these conditions.

Corollary 1 (exact comparative statics). Let P denote the right-hand side of (3.7), write $a := \ln(\bar{Z}_0/Z_{\min})$ and $x := \sigma_N T^H$, and let ϕ be the standard normal density evaluated at the argument of Φ in (3.7). The following statements hold.

- (i) $\frac{\partial P}{\partial \bar{Z}_0} = -\frac{\phi}{\bar{Z}_0 x} < 0$ for all parameter values.
- (ii) $\frac{\partial P}{\partial T} = \frac{\phi}{x} \left[\frac{Ha}{T} + \frac{Hx^2}{2T} + \nu(H-1) \right]$; in particular, if $\bar{Z}_0 \geq Z_{\min}$ (i.e., $a \geq 0$) and $\nu \leq 0$, then $\partial P / \partial T > 0$ for all $T > 0$.
- (iii) $\frac{\partial P}{\partial H} = \phi \left[\frac{a + \nu T}{x^2} + \frac{1}{2} \right] x \ln T$; hence, whenever $a + \nu T \geq 0$, stockout exposure is strictly increasing in H if $T > 1$, strictly decreasing if $T < 1$, and locally invariant at $T = 1$.
- (iv) $\frac{\partial P}{\partial \sigma_N} = \phi \left[\frac{a + \nu T}{x^2} + \frac{1}{2} \right] T^H > 0$ whenever $a + \nu T \geq 0$.

Analysis of Corollary 1. The derivatives in Corollary 1 state the comparative statics exactly rather than as numerical patterns, and two of them temper any broad monotonicity claim. First, the memory effect (iii) is *not* universal because it reverses for sub-unit horizons, as T^{2H} decreases in H when $T < 1$. The condition $T > 1$ is stated relative to the time unit in which σ_N is calibrated. The managerially relevant case, in which campaign and season horizons run long relative to the daily or weekly calibration unit, satisfies it, but the condition is substantive and belongs in the statement. Second, condition $a + \nu T \geq 0$ has the transparent interpretation that the initial log buffer must cover the expected cumulative net drain over the horizon. When the network starts below threshold or is drained fast enough that $a + \nu T < 0$, stockout exposure is already near its ceiling and the derivative signs in (iii)–(iv) can flip. All derivatives above are exact; Appendix A.3 contains the calculations.

The terminal exposure records the probability of ending below threshold but not the size of the deficit. Because $\hat{Z}_T$ is lognormal, the expected deficit also has a closed form of Black–Scholes put type.

Corollary 2 (expected terminal shortfall). *Let $\nu := \sigma_N T^H$ and let d denote the argument of Φ in (3.7). Then*

$$\mathbb{E}[(Z_{\min} - \hat{Z}_T)^+] = Z_{\min} \Phi(d) - \bar{Z}_0 e^{\nu T} \Phi(d - \nu). \quad (3.8)$$

Analysis of Corollary 2. The closed form in Corollary 2 prices the *severity* of a systemic stockout in inventory units, where Proposition 1 prices only its likelihood. In expression (3.8), the first term is the threshold weighted by the stockout exposure probability, while the second subtracts the expected inventory retained in the shortfall states. The corollary therefore equips the platform with a loss-based risk budget rather than a pure event probability, since expected shortfall converts into expected lost revenue through unit margins, and this addresses the standard objection that probability metrics ignore magnitude. Stockout exposure (3.7) and shortfall (3.8) are computed from the same two sufficient statistics (d, ν) , so both can be monitored in real time as H , σ_N , or the inventory state are re-estimated.

Remark 1 (first-passage exposure). Since $\{\hat{Z}_T \leq Z_{\min}\} \subseteq \{\min_{0 \leq t \leq T} \hat{Z}_t \leq Z_{\min}\}$, the terminal exposure (3.7) is a lower bound for the first-passage exposure; no closed form comparable to (3.7) is available for the latter under fBm. Section 6.2 shows numerically that the first-passage exposure preserves all comparative statics of Corollary 1 and that the gap between the two metrics narrows as H grows, because persistent networks that dip below threshold tend to stay there, so the terminal benchmark is tightest precisely in the high-persistence regime this paper targets.

Remark 2 (small initial inventory). Expression (3.7) requires $\bar{Z}_0 > 0$ through $\ln \bar{Z}_0$ and is monotone. As $\bar{Z}_0 \downarrow Z_{\min}$, the stockout exposure rises to $\Phi(\frac{-\nu T + \frac{1}{2}\sigma_N^2 T^{2H}}{\sigma_N T^H}) \geq \Phi(\frac{1}{2}\sigma_N T^H) > 1/2$ when $\nu \leq 0$, and it tends to one as $\bar{Z}_0 \downarrow 0$. Numerical evaluations should therefore start from strictly positive initial inventories (all figures use $\bar{Z}_0 \geq 1$), and near-threshold networks should be treated as already-distressed cases for which the relevant question is the control response of Section 4, not the stockout exposure estimate.

3.3. Inter-warehouse transshipment, balancing pressure, and physical volume

Section 3.2 established that transshipment intensity κ leaves the aggregate inventory path unchanged, and with it systemic stockout exposure. This does not make κ irrelevant because it determines the *internal logistics activity* the network performs to keep inventories balanced, and that activity is costly. Care is needed, however, in defining what is being measured. Define the **net balancing-flow matrix** $F_t^{(N)}$ with entries

$$F_t^{(i,j)} = \frac{\kappa}{N}(Z_t^{(j)} - Z_t^{(i)}), \quad i, j \in \mathcal{N}, \quad (3.9)$$

the signed net flow rate associated with the pair (i, j) , where a positive value indicates that inventory flows from j toward i . Because entries are signed, $F^{(N)}$ measures *balancing pressure* rather than physical shipment volume, since a physical shipment cannot be negative. We therefore work with two complementary metrics. The first aggregates squared entries through the expectation-Frobenius norm,

$$\|F_t^{(N)}\| = \left(\mathbb{E} \left[\sum_{i,j=1}^N (F_t^{(i,j)})^2 \right] \right)^{1/2}, \quad (3.10)$$

which we call the **balancing-pressure index**. The second counts actual directed shipments through the instantaneous **physical shipment volume**

$$V_t := \sum_{i,j=1}^N \frac{\kappa}{N} (Z_t^{(j)} - Z_t^{(i)})^+, \quad (3.11)$$

the total rate at which units physically move between warehouses at time t ; the cumulative volume over the horizon is $\int_0^T V_t dt$, and with a per-unit-shipped cost c , the transshipment cost is $c \int_0^T V_t dt$. Both metrics are driven by cross-sectional imbalance, measured in expectation by the dispersion index $D_N(t) := (\mathbb{E}[\frac{1}{N} \sum_i (Z_t^{(i)} - \bar{Z}_t)^2])^{1/2}$ and pathwise by $d_N(t) := (\frac{1}{N} \sum_i (Z_t^{(i)} - \bar{Z}_t)^2)^{1/2}$; we also use the Gini mean difference $G_N(t) := N^{-2} \sum_{i,j} |Z_t^{(i)} - Z_t^{(j)}|$.

Proposition 2 (transshipment activity). *For each $t \geq 0$, provided second moments are finite, the following statements hold.*

- (i) (balancing pressure) $\|F_t^{(N)}\| = \sqrt{2} \kappa D_N(t)$;
- (ii) (physical volume) $V_t = \frac{\kappa N}{2} G_N(t)$, pathwise;
- (iii) (equivalence bounds) $\frac{\kappa}{\sqrt{2}} d_N(t) \leq V_t \leq \frac{\kappa N}{\sqrt{2}} d_N(t)$, pathwise.

Analysis of Proposition 2. The identities in Proposition 2 show that the rebalancing effort a platform expends admits two exact readings, and that the two must not be used interchangeably. Part (i) is the dispersion identity for balancing pressure, which scales with κ and with the standard-deviation-type imbalance D_N . Part (ii) supplies the physically meaningful counterpart, in which total shipment volume equals transshipment intensity times $N/2$ times the Gini mean difference of the inventory cross-section. Because the Gini mean difference is an L^1 rather than an L^2 imbalance measure, the two parts respond differently to outliers, and a single deeply depleted warehouse inflates D_N quadratically but G_N only linearly. Part (iii) shows that the two views remain equivalent up to the network size factor. For a fixed N , pressure and volume move together within explicit bounds, and for approximately Gaussian cross-sections $V_t \approx \kappa N d_N(t) / \sqrt{\pi}$ (Appendix A.4). Four objects must therefore stay distinct, namely signed balancing pressure (3.10), directed instantaneous volume (3.11), cumulative volume $\int_0^T V_t dt$, and transshipment cost $c \int_0^T V_t dt$; claims about “logistics volume” in what follows always refer to (3.11).

Separating these objects has a direct managerial payoff. Systemic stockout exposure and internal logistics burden are two distinct health indicators. The former is fixed by the aggregate path (Lemma 1), while the latter is controlled jointly by the rebalancing intensity κ chosen by the platform and by the inventory dispersion, which the platform shapes through pre-positioning, demand shaping, and regional allocation. Two networks with identical systemic stockout exposure can therefore carry very different

internal logistics costs, and monitoring (D_N, G_N) alongside SSE separates the “how much total stock” question from the “how well distributed” question.

4. Optimal emergency replenishment for e-commerce platforms

4.1. Active risk management and problem formulation

Section 3 characterized the network’s stockout exposure absent intervention ($u^{(i)} \equiv 0$), which corresponds to passive risk management through initial buffers and automatic rebalancing. We now let the platform intervene through costly emergency replenishment and derive the optimal intervention rule.

Two decision levels and the role of κ . The control problem is meaningful at two levels, and Lemma 1 links them exactly. At the *systemic* level, the platform steers the network-average inventory, the very state variable whose threshold crossing defines systemic stockout exposure. Averaging (3.3) over warehouses, the transshipment terms cancel identically for every $\kappa \geq 0$, so the network-average dynamics contain no transshipment term at all. Emergency control of the systemic state is therefore *exactly* a single-state control problem, not an approximation obtained by switching transshipment off. At the *decentralized* level, each warehouse manages its own deviation, and setting $\kappa = 0$ isolates the emergency-replenishment margin from the rebalancing margin studied in Section 3.3; this is the benchmark under which per-warehouse problems decouple. Both readings lead to the same mathematical problem below, which we state for a generic deviation state X_t , standing for the network-average gap in the systemic reading and for warehouse i ’s gap in the decentralized reading.

Centered state. Let $z^* > 0$ be the safety-stock target and define the inventory deviation $X_t := Z_t - z^*$, so $X_t < 0$ signals shortage pressure and $X_t > 0$ excess. The platform’s objective is to steer the deviation toward zero at minimal emergency cost.

Assumption 2 (controlled deviation dynamics and admissible policies). (a) *The deviation evolves as*

$$dX_t = u_t dt + \sigma X_t dB_t^H, \quad X_0 = x_0 \neq 0, \quad (4.1)$$

interpreted in the WIS sense of Assumption 1. (b) Admissible controls form the linear-feedback class

$$\mathcal{U}_{\text{lin}} := \{u_t = L_t X_t : L : [0, T] \rightarrow \mathbb{R} \text{ bounded and measurable, deterministic}\}. \quad (4.2)$$

(c) *The cost coefficients $h_t \geq 0$, $r_t \geq \delta$ for some $\delta > 0$, and $k \geq 0$ are bounded and measurable, with h, r deterministic functions of time.*

Part (a) is a modeling choice that deserves justification. The noise is proportional to the *deviation*, which means that persistent demand-supply imbalances feed on the size of the current gap. Larger shortfalls attract herd-driven scarcity effects, larger excesses reflect the aftermath of larger shocks, and the specification inherits the sign-preservation property of geometric dynamics, which proves operationally meaningful below. It is not the exact linearization of (3.3) around z^* , which would carry the affine noise $\sigma(X_t + z^*)dB_t^H$; the affine variant destroys the homogeneous structure that yields a scalar Riccati equation, and we leave it as an extension (Section 7). Part (b) restricts attention to policies that are *implementable* in fulfillment systems. The replenishment rate responds linearly to the observed gap through a precomputed, deterministic time-varying gain. Optimality below is claimed within $\mathcal{U}_{\text{lin}}$ and not over all adapted open-loop processes; for the wider classes, which require the machinery of

fractional backward equations, see [86–88]. This restriction is stated explicitly, and the practical loss is arguably small because the optimal member of $\mathcal{U}_{\text{lin}}$ already achieves the canonical LQ value $b_0 x_0^2$ characterized below.

The platform selects $u \in \mathcal{U}_{\text{lin}}$ to minimize the total expected cost

$$C(u) = \mathbb{E} \left[\int_0^T (h_t X_t^2 + r_t u_t^2) dt + k X_T^2 \right]. \quad (4.3)$$

The running penalty $h_t X_t^2$ prices deviations from the safety target, with shortage exposure on one side and capital occupation and handling pressure on the other; $r_t u_t^2$ prices the use of expensive emergency channels; the terminal penalty $k X_T^2$ prevents ending the horizon in a severely unbalanced state.

The objective is a quadratic tracking approximation. We state explicitly what (4.3) is and is not. It is *not* the minimization of systemic stockout exposure (3.7), of first-passage risk, or of expected shortfall (3.8). Quadratic costs penalize shortage and excess symmetrically, and no closed-form control minimizing a threshold-crossing probability is available under fBm, since the non-Markovian state rules out the HJB characterization such objectives would require. The LQ problem instead provides a tractable quadratic *tracking* surrogate that keeps the state near the safety target and thereby indirectly suppresses threshold crossings. Section 6.3 closes the loop empirically, showing that the optimal LQ policy reduces the terminal shortage probability by an order of magnitude relative to passive and naive benchmark policies at comparable or lower cost. The asymmetric-objective problem remains open and is flagged in Section 7.

Why classical tools fail. For $H \neq 1/2$, (4.1) is driven by a process that is neither Markovian nor a semimartingale. The future depends on the entire shock history, so dynamic programming and HJB equations do not apply [21, 23], and classical Itô calculus is unavailable. The WIS calculus of Assumption 1 provides the substitute, and the linear-quadratic structure preserves explicit solvability. Rather than manipulating fractional stochastic integrals directly, whose expectations require care to justify, our optimality proof reduces the problem to a *deterministic* control problem for the second moment of the state, for which every step is elementary.

4.2. Optimal replenishment policies

Under a linear feedback $u_t = L_t X_t \in \mathcal{U}_{\text{lin}}$, the controlled state (4.1) is the WIS solution

$$X_t = x_0 \exp \left(\int_0^t L_s ds + \sigma B_t^H - \frac{1}{2} \sigma^2 t^{2H} \right), \quad (4.4)$$

whose second moment $m_t := \mathbb{E}[X_t^2]$ is finite and, by the Gaussian moment formula $\mathbb{E}[e^{2\sigma B_t^H}] = e^{2\sigma^2 t^{2H}}$, equals $x_0^2 \exp(2 \int_0^t L_s ds + \sigma^2 t^{2H})$; hence, m solves the linear ordinary differential equation

$$\dot{m}_t = m_t(2L_t + 2\sigma^2 H t^{2H-1}), \quad m_0 = x_0^2. \quad (4.5)$$

The memory of the demand shocks enters through the coefficient $2\sigma^2 H t^{2H-1} = 2\sigma^2 \int_0^t \varphi(t, s) ds$, the integrated kernel of Assumption 1; at $H = 1/2$, it reduces to the constant Itô correction σ^2 . The associated Riccati equation is

$$\dot{b}_t + 2\sigma^2 H t^{2H-1} b_t + h_t - r_t^{-1} b_t^2 = 0, \quad b_T = k. \quad (4.6)$$

Lemma 2 (Riccati solvability). Under Assumption 2(c), equation (4.6) has a unique solution on $[0, T]$, and it satisfies $0 \leq b_t \leq (k + \bar{h}T)e^{2\bar{a}T}$, where $\bar{h} := \sup_t h_t$ and $\bar{a} := \sigma^2 H T^{2H-1}$.

Proposition 3 (optimal emergency replenishment). Under Assumptions 1 and 2, with b the solution of (4.6), the following statements hold.

(i) the policy $\hat{u}_t = -r_t^{-1}b_t X_t$ belongs to $\mathcal{U}_{\text{lin}}$ and is the unique optimal policy in $\mathcal{U}_{\text{lin}}$ (uniqueness up to modification on a Lebesgue-null set of times);

(ii) for every $u = LX \in \mathcal{U}_{\text{lin}}$,

$$C(u) = b_0 x_0^2 + \int_0^T m_t r_t (L_t + r_t^{-1}b_t)^2 dt \geq b_0 x_0^2, \quad (4.7)$$

with equality if and only if $L_t = -r_t^{-1}b_t$ for a.e. t ; hence, the minimal cost is $C(\hat{u}) = b_0 x_0^2$.

The proof is given in Appendix A.6.

One remark on the proof is useful. The argument requires no expectation of a fractional stochastic integral because Assumption 1 enters only through the explicit solution (4.4) and its second moment. The delicate zero-expectation properties of WIS integrals, which must otherwise be verified integrand by integrand, are therefore never invoked. The classical completing-the-square derivation via the fractional Itô formula of [24] delivers the same Riccati equation (4.6); we present the moment route because every step is verifiable by elementary calculus.

Analysis of Proposition 3. The optimal rule of Proposition 3, $\hat{u}_t = -r_t^{-1}b_t X_t$, is one the platform can actually run, and its managerial content has three layers. *Interpretability* follows because the platform needs only the observed gap X_t and the precomputed gain $-r_t^{-1}b_t$. No path history enters the policy even though the environment is non-Markovian, because all memory is absorbed into the deterministic coefficient b_t through the $2\sigma^2 H t^{2H-1}$ term of (4.6). The Riccati coefficient b_t is the shadow price of squared deviation, meaning the current marginal cost of carrying one more unit of squared gap from t to the horizon. Dividing by r_t converts that value into the economically justified intervention intensity. Higher demand persistence (larger H , at horizons beyond the unit scale) raises the memory coefficient and with it the shadow price early in the horizon, prescribing *earlier and stronger* intervention than a memoryless calibration would, which is the control-side counterpart of Proposition 1's warning that Markovian models understate risk. *The sign structure and nonnegativity constraint* follow from the geometric dynamics (4.4), which preserve the sign of X_t . If the state starts in shortage ($x_0 < 0$), then $X_t < 0$ throughout, so $\hat{u}_t = -r_t^{-1}b_t X_t > 0$ automatically. Along the entire shortage trajectory, the optimal emergency order rate is nonnegative and the constraint $u \geq 0$ is slack, so no projection is needed exactly where emergency replenishment is meant to operate. If the state starts in excess ($x_0 > 0$), the optimal $\hat{u}_t < 0$ is the cancellation, delay, or redirection of scheduled inbound supply; where such negative actions are institutionally unavailable, the projected policy $\max\{0, \hat{u}_t\}$ degenerates to passivity, the LQ optimality claim no longer applies to it, and the cost of that constraint is quantified in Table 2 of Section 6.3. *The value of initial positioning* follows because the minimal cost $b_0 x_0^2$ is quadratic in the initial gap, so halving the launch-time deviation quarters the expected control bill. Combined with Corollary 1(i), this gives initial inventory a dual role, buffering stockout exposure and economizing on future emergency intervention, and makes the calibration of z^* a first-order decision before campaign events.

5. Extensions

The two extensions that follow enrich the stochastic environment while keeping the decision architecture of Sections 3 and 4 unchanged; their purpose is to establish robustness and delimit generality, not to replace the baseline formulation. The first allows routine short-memory fluctuations to coexist with persistent shocks; the second drops the independence of shocks across warehouses, which is where the sharpest new conclusion emerges, namely an irreducible systemic risk floor.

5.1. Mixed demand shocks

In the baseline model, all demand variability is long-memory. In practice, aggregate demand mixes routine fluctuations from many quasi-independent consumers, which are well approximated by a memoryless process, with persistent amplification from campaigns and social diffusion. We therefore drive inventories by both a standard Brownian motion $W^{(i)}$ (short memory) and an independent fBm $B^{H,i}$ (long memory) as follows.

$$dZ_t^{(i)} = \kappa(\bar{Z}_t - Z_t^{(i)})dt + u_t^{(i)}dt + \nu Z_t^{(i)}dt + \sigma_1 Z_t^{(i)}dW_t^{(i)} + \sigma_2 Z_t^{(i)}dB_t^{H,i}, \quad (5.1)$$

with $\sigma_1, \sigma_2 > 0$ the short- and long-memory volatilities; Brownian integrals are Itô integrals, fBm integrals are WIS integrals, and the two families are independent. Aggregation proceeds exactly as in Lemma 1. Transshipment cancels, and averaging compresses each noise family by $\sqrt{N}$, yielding network-effective volatilities $\sigma_{1,N} := \sigma_1 / \sqrt{N}$ and $\sigma_{2,N} := \sigma_2 / \sqrt{N}$ for the mean-field benchmark, with the same closure caveats as in Section 3.2.

Proposition 4 (mixed shocks). *Under Assumptions 1 and 2 adapted to (5.1), let $v^2(T) := \sigma_{1,N}^2 T + \sigma_{2,N}^2 T^{2H}$. The following statements hold.*

(i) *the terminal systemic stockout exposure of the mean-field benchmark is*

$$\mathbb{P}\{\hat{Z}_T \leq Z_{\min}\} = \Phi\left(\frac{\ln(Z_{\min}/\bar{Z}_0) - \nu T + \frac{1}{2}v^2(T)}{v(T)}\right); \quad (5.2)$$

(ii) *the transshipment identities of Proposition 2 hold verbatim, because they are pathwise algebraic identities that do not involve the law of the noise;*

(iii) *the optimal emergency replenishment policy in $\mathcal{U}_{\text{lin}}$ retains the linear feedback form $\hat{u}_t = -r_t^{-1}b_t X_t$ with minimal cost $b_0 x_0^2$, where b now solves*

$$\dot{b}_t + (\sigma_1^2 + 2\sigma_2^2 H t^{2H-1})b_t + h_t - r_t^{-1}b_t^2 = 0, \quad b_T = k. \quad (5.3)$$

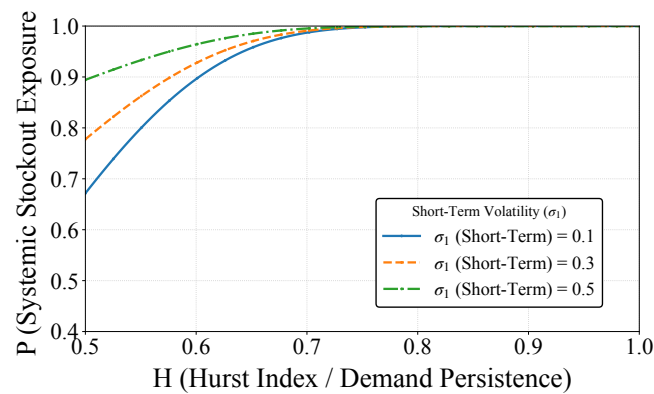


Figure 4. Terminal systemic stockout exposure P versus H under mixed shocks (Proposition 4). The parameter setting uses $Z_{\min} = 0.7$, $\bar{Z}_0 = 10$, $T = 30$, $\nu = 0$, fixed long-memory volatility $\sigma_{2,N} = 0.5$, and short-memory volatility $\sigma_{1,N} \in \{0.1, 0.3, 0.5\}$.

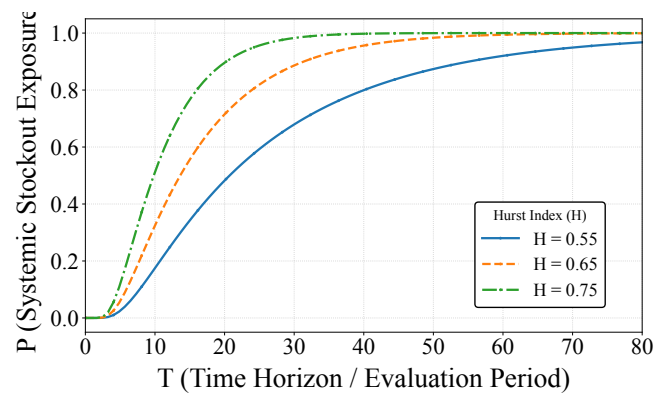


Figure 5. Terminal systemic stockout exposure P versus horizon T under mixed shocks. The parameter setting uses $Z_{\min} = 0.7$, $\bar{Z}_0 = 10$, $\nu = 0$, $\sigma_{1,N} = 0.2$, $\sigma_{2,N} = 0.4$, and $H \in \{0.55, 0.65, 0.75\}$.

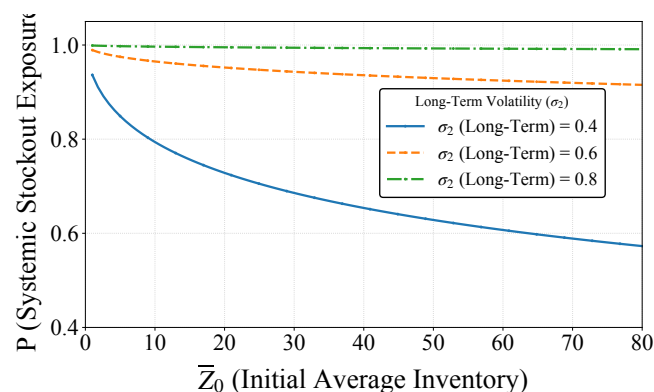


Figure 6. Terminal systemic stockout exposure P versus initial average inventory $\bar{Z}_0$ under mixed shocks. The parameter setting uses $Z_{\min} = 0.7$, $T = 30$, $H = 0.6$, $\nu = 0$, $\sigma_{1,N} = 0.2$, and $\sigma_{2,N} \in \{0.4, 0.6, 0.8\}$.

Analysis of Proposition 4. The two shock channels separate cleanly in Proposition 4, and expression (5.2) shows that the routine variance channel grows linearly in T , whereas the persistent channel grows as T^{2H} . Long memory therefore dominates tail exposure at long horizons for any fixed (σ_1, σ_2) , which is the analytical content behind Figures 4, 5, and 6. The comparative statics of Corollary 1 carry over with x replaced by $v(T)$; in particular, increasing memory raises stockout exposure only when the horizon exceeds the unit time scale and the corresponding drift-adjusted buffer condition holds (Appendix A.3). The Riccati equation (5.3) also exposes a structural asymmetry. The Brownian channel contributes the *constant* Itô correction σ_1^2 , while the fBm channel contributes the *time-growing* memory coefficient $2\sigma_2^2 H t^{2H-1}$. The two corrections have different origins (quadratic variation versus the WIS kernel) and are not interchangeable, which is why the proof in Appendix A.7 tracks them separately rather than “absorbing” one into the other. Managerially, Figure 4 illustrates persistence as the dominant risk driver across its reported routine-volatility levels; Figure 5 shows the nonlinear acceleration of risk accumulation over the plotted persistence values; and Figure 6 quantifies how much larger the initial buffer must be to hold a target stockout exposure as the persistent-shock intensity σ_2 rises.

5.2. Correlated demand shocks

The baseline assumes independent shocks across warehouses, under which pooling across N locations compresses aggregate volatility by $\sqrt{N}$ (Lemma 1). During platform-wide promotions, viral episodes, or sustained public-attention events, however, regional demands surge together. We therefore give the N -dimensional shock the following common-factor structure.

$$dB_t^{H,i} = \sqrt{\rho} dB_t^{H,\text{sys}} + \sqrt{1-\rho} dB_t^{H,\text{ind},i}, \quad \rho \in [0, 1], \quad (5.4)$$

where $B^{H,\text{sys}}$ is a systemic shock common to all warehouses, and $(B^{H,\text{ind},i})_i$ are warehouse-specific idiosyncratic shocks, all independent standard fBms. Each $B^{H,i}$ remains a standard fBm for every ρ ; $\rho = 0$ recovers the baseline, and $\rho = 1$ makes all warehouses move in lockstep.

Proposition 5 (correlated shocks). *Under Assumption 1 with the factor structure (5.4), the following statements hold.*

(i) *the terminal systemic stockout exposure of the Gaussian-aggregate benchmark is*

$$\mathbb{P}\{\hat{Z}_T \leq Z_{\min}\} = \Phi\left(\frac{\ln(Z_{\min}/\bar{Z}_0) - vT}{\sigma_{\text{agg}} T^H} + \frac{1}{2}\sigma_{\text{agg}} T^H\right), \quad \sigma_{\text{agg}} := \sigma \sqrt{\rho + \frac{1-\rho}{N}}; \quad (5.5)$$

at $\rho = 0$, $\sigma_{\text{agg}} = \sigma/\sqrt{N} = \sigma_N$, consistent with Definition 2, and at $\rho = 1$, the benchmark is exact for the full coupled network for every N and κ (Appendix A.2);

(ii) *the transshipment identities of Proposition 2 hold verbatim (pathwise identities);*

(iii) *each warehouse’s decentralized control problem is unchanged because its solution depends on the driving noise only through the marginal law of $B^{H,i}$, which is that of a standard fBm for every ρ . The Riccati equation (4.6), the feedback policy, and the minimal cost of Proposition 3 therefore apply without modification.*

Theorem 1 (large-network limit and the systemic risk floor). *Let $\kappa = 0$, $u \equiv 0$, let initial inventories be uniformly bounded with $\bar{Z}_0^{(N)} \rightarrow \bar{z}_0 > 0$, and let shocks follow (5.4) with $\rho \in [0, 1]$. Then, almost surely as $N \rightarrow \infty$,*

$$\bar{Z}_T \longrightarrow \bar{z}_0 \exp\left(vT + \sigma \sqrt{\rho} B_T^{H,\text{sys}} - \frac{1}{2}\rho\sigma^2 T^{2H}\right), \quad (5.6)$$

and consequently, for $\rho > 0$,

$$\lim_{N \rightarrow \infty} \mathbb{P}\{\bar{Z}_T \leq Z_{\min}\} = \Phi\left(\frac{\ln(Z_{\min}/\bar{z}_0) - \nu T}{\sigma \sqrt{\rho} T^H} + \frac{1}{2}\sigma \sqrt{\rho} T^H\right) > 0. \quad (5.7)$$

For $\rho = 0$, the limit (5.6) is the deterministic path $\bar{z}_0 e^{\nu T}$, and the limiting stockout exposure is 0 or 1 according to the sign of $\ln(Z_{\min}/\bar{z}_0) - \nu T$. The convergence rate is $O(N^{-1/2})$, with the constant governed by $(e^{(1-\rho)\sigma^2 T^{2H}} - 1)^{1/2}$ (Appendix A.8).

Analysis of Proposition 5 and Theorem 1. Together, Proposition 5 and Theorem 1 settle what network expansion can and cannot buy, and three conclusions follow. First, *the diversification logic of network expansion has a hard boundary*. Under independence, $\sigma_{\text{agg}} = \sigma/\sqrt{N} \rightarrow 0$, so enough warehouses drive systemic stockout exposure to zero, as predicted by classical risk pooling. With any common component $\rho > 0$, $\sigma_{\text{agg}} \downarrow \sigma \sqrt{\rho} > 0$, and Theorem 1 upgrades the finite- N intuition to an exact limit theorem. The idiosyncratic components average out at the explicit $N^{-1/2}$ rate, whereas the systemic component does not. The network's tail risk consequently converges to a strictly positive floor (5.7) that no expansion, reallocation, or transshipment policy can remove. The floor yields only to a reduction in ρ itself, through assortment diversification across demand communities or staggered campaigns, or to higher buffers. Second, the theorem also provides the *rigorous foundation for the closed forms of this paper*. It identifies exactly when the lognormal benchmark is the truth (in the limit, and exactly at $\rho = 1$ for any N, κ) rather than an unexamined closure, answering at the level of theory what Section 6.1 verifies numerically at finite N . Third, correlation is *invisible* at the decentralized level. Part (iii) shows that each warehouse's optimal policy is unchanged because individual optimization sees only its own marginal shock law. Systemic risk under correlation is thus an emergent network property that no collection of individually optimal policies internalizes, the operations analogue of the finance distinction between individually rational risk management and system-level externalities [47, 48]. This is why the platform-level stockout exposure metrics of Section 3 and the systemic-level control reading of Section 4 matter. Figures 7, 8, and 9 quantify the finite- N picture. Stockout exposure rises steeply with ρ at high H (Figure 7), risk accumulates on compressed time scales under high correlation (Figure 8), and the initial buffer required to hold a target stockout exposure grows sharply with ρ (Figure 9). Stocking more helps, but against a common persistent surge, it is a costly and ultimately bounded defense.

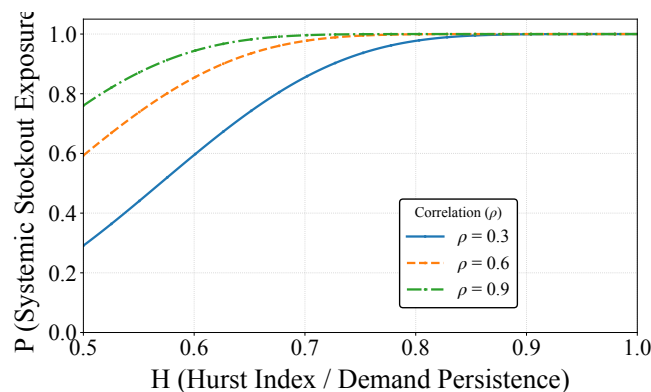


Figure 7. Terminal systemic stockout exposure P versus H under correlated shocks (Proposition 5). The parameter setting uses $Z_{\min} = 0.7$, $\bar{Z}_0 = 10$, $T = 30$, $\nu = 0$, per-warehouse volatility $\sigma = 0.6$, $N = 100$, and $\rho \in \{0.3, 0.6, 0.9\}$.

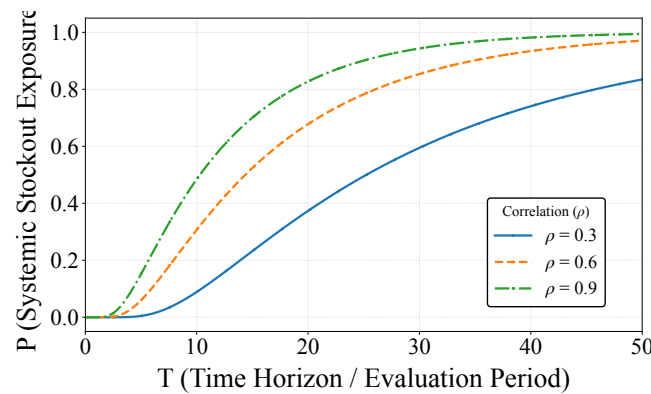


Figure 8. Terminal systemic stockout exposure P versus horizon T under correlated shocks. The parameter setting uses $Z_{\min} = 0.7$, $\bar{Z}_0 = 10$, $H = 0.6$, $\nu = 0$, $\sigma = 0.6$, $N = 100$, and $\rho \in \{0.3, 0.6, 0.9\}$.

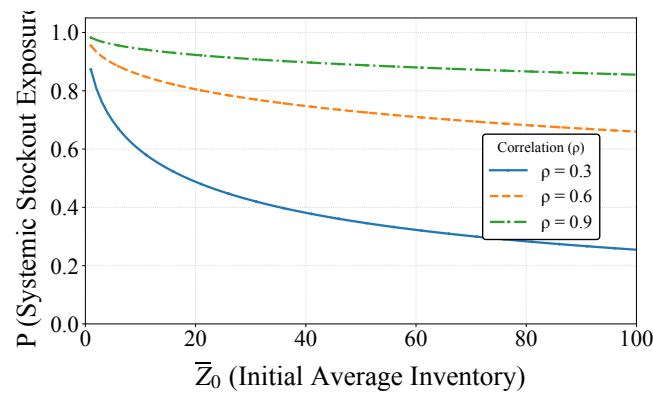


Figure 9. Terminal systemic stockout exposure P versus initial average inventory $\bar{Z}_0$ under correlated shocks. The parameter setting uses $Z_{\min} = 0.7$, $T = 30$, $H = 0.6$, $\nu = 0$, $\sigma = 0.6$, $N = 100$, and $\rho \in \{0.3, 0.6, 0.9\}$.

6. Numerical validation and calibration guidance

This section provides the numerical validation and implementation layer for the preceding analysis. It covers Monte Carlo accuracy checks of every closed form against the full N -warehouse system (Section 6.1), the relation between terminal and first-passage exposure (Section 6.2), a benchmark comparison for the optimal policy including the effect of the nonnegativity constraint (Section 6.3), and statistical guidance for calibrating the long-memory input before the framework is used (Section 6.4). All experiments report complete parameter sets and use exact-in-distribution simulation of fBm (circulant embedding with a Cholesky fallback [97]); the simulation representation of the WIS dynamics is the pathwise corrected form of Assumption 1(c).

6.1. Accuracy of the closed-form benchmark

This subsection establishes the accuracy envelope of the closed-form benchmark, identifying by direct comparison against exact full-system simulation the regimes in which the closed form can be

trusted and those in which the full simulator is needed. Table 1 reports the comparison, setting the closed-form stockout exposure (5.5), which nests (3.7) at $\rho = 0$, against the empirical stockout exposure of the *full* N -warehouse system. For $\rho < 1$, the full system is simulated exactly at $\kappa = 0$ through the per-warehouse WIS solutions under the factor structure (5.4); for $\rho = 1$, the exact solution of the *coupled* system with $\kappa > 0$ (Appendix A.2) is used, so no discretization error contaminates the comparison. The common parameter setting uses $T = 5$, $H = 0.6$, $\nu = 0$, $Z_{\min} = 8$, and $\bar{Z}_0 = 10$. “Balanced” initial inventories set all $\xi^{(i)} = 10$, whereas “dispersed” ones spread $\xi^{(i)}$ uniformly on a grid from 5 to 15 with mean 10.

Table 1. Closed-form versus full-system terminal systemic stockout exposure.

ρ	κ	N	σ	initial inventories	closed-form P	MC P	MC s.e.	abs. error
1.0	0	10	0.35	balanced	0.5859	0.5879	0.0025	0.0021
1.0	2	10	0.35	dispersed	0.5859	0.5867	0.0025	0.0009
0.5	0	5	0.35	dispersed	0.5170	0.5227	0.0029	0.0057
0.5	0	50	0.35	dispersed	0.4954	0.4953	0.0056	0.0001
0.0	0	10	0.15	balanced	0.0419	0.0482	0.0009	0.0063
0.0	0	10	0.15	dispersed	0.0419	0.0556	0.0009	0.0137
0.0	0	10	0.60	dispersed	0.4213	0.5551	0.0020	0.1338

Note: The common parameter setting uses $T = 5$, $H = 0.6$, $\nu = 0$, $Z_{\min} = 8$, and $\bar{Z}_0 = 10$, with 10,000–60,000 Monte Carlo paths per row (standard errors reported) and $n = 512$ time-grid points.

The pattern follows the theory of Sections 3.2 and 5.2 exactly. In the fully systemic rows ($\rho = 1$), the discrepancy is within Monte Carlo noise for both $\kappa = 0$ and $\kappa = 2$, as the exactness result predicts. In the correlated rows ($\rho = 0.5$), the benchmark error falls from 0.006 at $N = 5$ to below Monte Carlo resolution at $N = 50$, which is the $O(N^{-1/2})$ convergence of Theorem 1 at work. The independent rows ($\rho = 0$) delimit where the benchmark holds. With moderate per-warehouse cumulative volatility ($\sigma^2 T^{2H} \approx 0.16$) and balanced inventories, the error is below 0.007; initial dispersion roughly doubles it, consistent with the exact variance identity of Appendix A.2, in which cross-sectional heterogeneity inflates true aggregate variance by the factor $1 + s_{\xi}^2 / \bar{\xi}^2$; and the deliberately stressed final row ($\sigma = 0.6$, $\sigma^2 T^{2H} \approx 2.5$) shows the closure understating stockout exposure by 0.13. In heavy-tailed idiosyncratic regimes, the lognormal benchmark is *not* a substitute for simulation. In practical terms, the closed forms are reliable precisely in the systemic-shock regimes that motivate the paper (ρ substantial, or N large) and in the moderate-volatility independent regime. The full-system simulator that produced Table 1 should be used outside that envelope.

6.2. First-passage versus terminal exposure

This subsection establishes numerically how the closed-form terminal metric relates to the operationally richer first-passage metric, a relation anticipated in Remark 1. Figure 10 makes that relation precise by simulating the benchmark process exactly on a fine grid and recording both the terminal exposure and the running-minimum (first-passage) exposure. Two facts stand out. First, the ordering $P_{\text{fp}}(T) \geq P_{\text{term}}(T)$ holds everywhere, and both metrics share the comparative statics of Corollary 1, rising in T and ordered in H , so managerial conclusions drawn from the closed-form terminal metric carry over qualitatively to the operationally richer first-passage metric. Second, the gap

between the two shrinks as H grows because persistent networks that cross the threshold rarely recover within the horizon. The terminal measure is therefore tightest exactly in the high-persistence regime this paper targets and most conservative-biased in near-Markovian regimes. Grid-based minima slightly understate continuous-time first passage, so the plotted first-passage curves are themselves lower-bound estimates; this bias is second order at the plotted grid resolution.

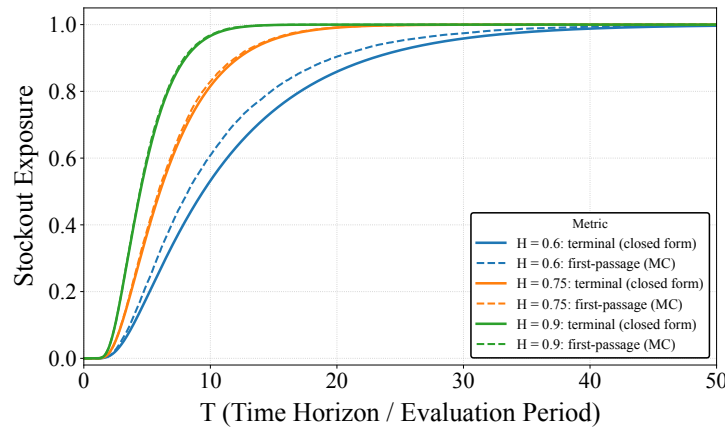


Figure 10. Terminal (closed form, solid) versus first-passage (Monte Carlo running minimum, dashed) systemic stockout exposure as functions of the horizon T . The parameter setting uses $Z_{\min} = 0.7$, $\bar{Z}_0 = 10$, $\nu = 0$, $\sigma_N = 0.6$, and $H \in \{0.6, 0.75, 0.9\}$, with 20,000 exact fBm paths on a 1,024-point grid.

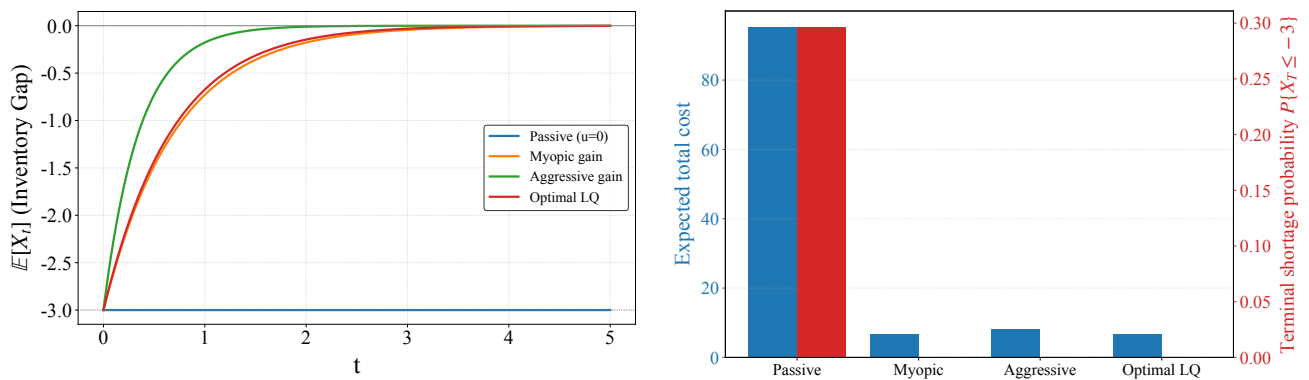
6.3. Benchmark replenishment policies

This subsection quantifies the operational value of the optimal policy of Proposition 3 and prices the nonnegativity constraint, benchmarking the policy against implementable alternatives in the two initial regimes whose sign structure the theory distinguishes. Table 2 and Figure 11 report this evaluation. The parameter setting uses $\sigma = 0.35$, $H = 0.7$, $T = 5$, $h = 1$, $r = 0.5$, $k = 0.6$, and safety target $z^* = 10$. Expected costs are computed *exactly* through the moment ODE (4.5) (no simulation error); adverse-tail probabilities use 30,000 exact simulated paths. In the shortage scenario, the adverse event is $X_T \leq -3$ (inventory still at or below its initial deficit, i.e., $Z_T \leq 7$); in the excess scenario, it is $X_T \geq +3$.

Table 2. Emergency replenishment policies under long-memory demand.

Scenario	Policy	Expected total cost	% above LQ optimum	Terminal adverse-tail probability
A (shortage, $X_0 = -3$)	Passive ($u = 0$)	95.1731	1343.3851	0.2961
A (shortage, $X_0 = -3$)	Myopic gain $-\sqrt{h/r}$	6.6029	0.1389	0.0000
A (shortage, $X_0 = -3$)	Aggressive gain $-2\sqrt{h/r}$	8.0642	22.3006	0.0000
A (shortage, $X_0 = -3$)	Optimal LQ	6.5937	0.0000	0.0000
A (shortage, $X_0 = -3$)	Projected LQ $\max\{0, u\}$	6.5937	0.0000	0.0000
B (excess, $X_0 = +3$)	Unconstrained LQ (cancel/redirect)	6.5937	0.0000	0.0000
B (excess, $X_0 = +3$)	Projected LQ (= passive, $u = 0$ binds)	95.1731	1343.3851	0.2961

Note: The parameter setting uses $\sigma = 0.35$, $H = 0.7$, $T = 5$, $h = 1$, $r = 0.5$, $k = 0.6$, and $z^* = 10$. Costs are exact via the moment ODE, and tail probabilities come from 30,000 exact paths on an $n = 800$ grid.



(a) Expected recovery paths

(b) Cost and terminal shortage risk

Figure 11. Emergency replenishment policies in the shortage scenario ($X_0 = -3$). (a) Deterministic expected recovery paths $\mathbb{E}[X_t] = X_0 e^{\int_0^t L_s ds}$ for the four policies of Table 2. (b) Expected total cost (left axis, exact) and terminal shortage probability $P\{X_T \leq -3\}$ (right axis, Monte Carlo). Parameters as in Table 2.

Four conclusions follow. First, *active control matters enormously*. The passive policy costs 14 times the optimum and leaves a 30% probability of ending at or below the initial deficit, while every feedback policy essentially eliminates that tail. This is the quantitative link between the LQ tracking objective and the stockout-risk objective established in Section 4.1. Second, *the optimal policy is inexpensive to approximate but easy to overdo*. A well-chosen constant-gain rule (the myopic steady-state gain $\sqrt{h/r}$) comes within 0.2% of the optimum in this calibration, whereas doubling the gain overshoots, recovering only slightly faster (panel (a) of Figure 11) at 22% higher cost. The value of Proposition 3 lies less in the raw cost savings over sensible heuristics than in its *guarantee*, the time-varying gain's automatic adjustment to (h_t, r_t, k, H) , and the shadow-price interpretation of b_t , which quantifies any deviation from optimality. Equation (4.7) provides the exact price of the constant-gain and aggressive rules. Third, *the nonnegativity constraint is free exactly where it matters*. In the shortage scenario, the projected policy coincides with the unconstrained optimum path by path (sign preservation), confirming that the LQ solution is implementable as a pure replenishment rule when the network is in deficit. Fourth, *the constraint is expensive where negative control is the whole point*. In the excess scenario, forbidding cancellation/redirection ($u \geq 0$ binding throughout) degenerates the projected policy to passivity at 14 times the unconstrained cost. Platforms whose contracts do not allow canceling or redirecting scheduled inbound flows should therefore read Table 2, Scenario B as the price of that contractual rigidity, not as a claim that the LQ policy remains optimal under the constraint.

6.4. Calibrating long-memory demand

This subsection is the deployment guide for the framework. It explains how a practitioner decides whether the fractional model applies to a given demand series and how to estimate the inputs the preceding results require. The key empirical input is the Hurst exponent, and using it responsibly requires answering three questions before deployment.

Table 3. Demand-shock model classes for network inventory risk.

Model	Long-range dependence	Jumps / burstiness	Continuous-time control	Closed-form network risk	Calibration
Poisson / compound Poisson	No (memoryless)	Yes (demand jumps)	Yes (Markovian DP/HJB)	Partial (renewal asymptotics)	Arrival-rate estimation
Brownian / GBM	No (independent increments)	No	Yes (Ito, HJB/LQ)	Yes (lognormal)	Volatility estimation
Levy processes	No (independent increments)	Yes (heavy tails)	Yes (jump HJB)	Partial (transform methods)	Levy-measure fitting
ARFIMA (discrete time)	Yes (fractional d)	No	No (discrete-time only)	No (simulation/forecast based)	d via Whittle/GPH
Hawkes processes	Quasi (self-excitation, short-to-mid memory)	Yes (clustered spikes)	Limited (intensity-state HJB)	No (intensity simulation)	Kernel + branching ratio
fBm, $H > 1/2$ (this paper)	Yes (Hurst exponent H)	No (continuous paths)	Yes (fractional LQ, Sec. 4)	Yes (Prop. 1, 4, 5)	H via R/S, DFA, local Whittle

Note: fBm is the unique entry combining genuine long-range dependence with closed-form network risk and solvable continuous-time control; ARFIMA is its discrete-time forecasting counterpart [59], and Hawkes/compound-Poisson classes should be preferred when the data are dominated by discrete bursts rather than persistent diffusion.

Is long memory present? The default diagnostic battery comprises Lo's modified rescaled-range test, which is robust to short-range dependence [62], detrended fluctuation analysis and aggregated-variance methods, whose comparative properties are documented in [63], and semiparametric frequency-domain estimators, namely the log-periodogram regression of [98] and the Gaussian local Whittle estimator of [99], the latter with well-understood asymptotics ($d = H - 1/2$ in the ARFIMA correspondence). Textbook treatments are given in [61]. Because retail series carry strong seasonality and promotions that can masquerade as spurious long memory [53, 54], demand series should be deseasonalized and campaign-dummied before estimation, and seasonal-fractional specifications of the type applied to G7 retail sales by [20] are the natural robustness check; entropy-based [55] and regression-splines [60] estimators provide further alternatives. Agreement of two or more estimators on $H > 1/2$ with confidence intervals excluding $1/2$ is a reasonable evidence bar for adopting the fBm layer.

Which model class fits the phenomenon? Table 3 answers this by positioning fBm against the main demand-process alternatives on the five dimensions that matter here. Unlike the numerical tables above, its cells are a qualitative synthesis of the process properties reviewed in Section 2.2 and the results established in this paper. Reading the fBm row against the others isolates why the fractional model is adopted, since fBm is the only class that combines genuine long-range dependence with closed-form network risk and solvable continuous-time control.

How do estimates enter the decision layer? The estimated $\hat{H}$ and volatility feed into Proposition 1 and Corollaries 1–2 for risk assessment (with the finite- N accuracy envelope of Table 1 in mind). The correlation ρ , which can be estimated from the cross-sectional correlation of warehouse-level demand innovations, feeds into Proposition 5 and the floor formula (5.7). Additionally, $(\hat{H}, \hat{\sigma})$ parameterize the Riccati equation (4.6), whose solution provides the deployment-ready feedback gain. When diagnostics reject long memory ($\hat{H} \approx 1/2$), the framework reduces gracefully to its classical Itô special case; when they detect burst-dominated demand, Table 3 points to jump-based alternatives instead. This places the model where it belongs: as an option exercised on statistical evidence, not a default.

7. Conclusion

This paper has studied systemic stockout risk in multi-warehouse e-commerce fulfillment networks driven by persistent demand pressure. Under an explicitly stated Wick–Itô–Skorokhod convention, the model couples a nondecreasing baseline sales process with a signed long-memory net-flow mismatch, linking persistent inventory pressure to a network-wide risk measure and then to an explicit emergency replenishment rule.

Main findings. The paper establishes four results. First, we obtain a formula for systemic stockout risk that a platform can compute directly, where this risk is the chance that the network’s average inventory of a critical item finishes the planning period below the level needed to sustain fulfillment. The formula applies to a mean-field benchmark of the network and needs only a few estimable inputs rather than simulation, and it comes with an exact account of how the risk moves with its drivers, rising as demand grows more persistent, as the evaluation period lengthens, and as volatility increases, with the condition for each effect stated precisely. It is more than an approximation, holding exactly when every warehouse faces the same network-wide shock, becoming exact as the network grows when the shocks are only partly shared, and carrying a simulation-measured error when the shocks are independent, so the range in which it can be trusted is proven rather than assumed. Two further measures complete the picture: a closed form for the expected size of a shortfall once one occurs and a stricter reading, evaluated numerically, that asks whether inventory ever dips below the threshold during the period rather than only at its end and that can never fall short of the end-of-period risk. Second, we show that the internal work of keeping the network balanced splits into two quantities that should not be confused, a pressure indicator reflecting how unevenly stock is spread across warehouses and how the directed volume of goods is actually moved between them, tied together by an exact relationship, so that a platform can read the network’s overall stock health separately from the internal logistics cost of maintaining it. Third, we solve the emergency replenishment problem within the class of feedback rules that set the order rate from the observed inventory gap, where the optimal rule follows from a Riccati equation that carries a memory term for demand persistence. We prove that this equation has a unique solution, a time-varying replenishment gain readable as the marginal value of closing the gap, established through the average squared inventory deviation by elementary steps that avoid the difficulties of handling fractional noise directly. Fourth, we prove that when demand shocks are correlated across warehouses, the network’s overall risk does not fall to zero as the network grows but settles at a strictly positive floor set by the shared part of demand, so that adding or diversifying warehouses removes the risk specific to individual locations but not the risk they all share.

Managerial implications. These findings translate into concrete guidance for platforms that run centrally coordinated fulfillment networks. A platform can track network-wide stockout risk continuously from a small set of estimated demand inputs, rather than from full simulation, and can read it alongside a companion measure of the expected shortfall size, so that risk is managed by its likely magnitude and not only by the chance of an event. Because greater demand persistence raises this risk over evaluation horizons longer than the calibration unit, the most exposed categories are those propelled by viral promotion, social diffusion, or platform-wide campaigns. These warrant shorter risk-review cycles and larger pre-positioned inventory ahead of launches and campaign windows, since a larger opening buffer both lowers the risk and reduces the later cost of emergency intervention. This continuous reading is dependable only inside the region where it has been validated, so a platform

should fall back on full-system simulation when idiosyncratic volatility is high, when inventory is very unevenly distributed, or when the network is small. A platform should also watch two health signals side by side, the overall stock adequacy of the network and the internal logistics burden of keeping it balanced, because two networks that look equally healthy overall can incur very different rebalancing costs. For emergency replenishment, the platform can precompute a replenishment gain and apply it as a simple rule that responds to the observed inventory gap, intervening earlier and more forcefully for categories with stronger demand persistence than a memoryless calibration would advise. When the network is running short, this rule is a clean replenishment policy, whereas when the network is carrying excess, its value depends on whether scheduled inbound orders can be canceled or redirected, so platforms whose supply contracts forbid such actions should read the excess case as the cost of that contractual rigidity. Finally, because demand shocks shared across the network impose a floor that adding warehouses cannot remove, a platform should treat network growth as a defense against localized rather than system-wide surges, and should attack the residual exposure either by lowering the correlation of demand across regions, through diversifying assortments across distinct demand communities and staggering campaigns, or by holding larger buffers. Adopting the long-memory framework is itself a decision to be taken on statistical evidence, applied when demand data support persistent memory and set aside in favor of jump-based descriptions when demand is dominated by discrete bursts.

Limitations and future research. The boundaries of the results also point to the natural next steps. First, the closed-form risk measure describes a mean-field benchmark whose accuracy degrades when idiosyncratic shocks are very large, when inventories are strongly dispersed, when the network is small, or when cross-warehouse dependence is weak, so the full-system simulator should replace the formula outside the validated region; sharpening the first-passage bounds in the high-persistence regime remains open. Second, the control result is optimal only within the linear-feedback class under noise proportional to the inventory deviation and targets a tracking surrogate for the stockout probability, so extending it to additive noise, to asymmetric penalties for shortage and excess, and to a direct treatment of the nonnegative-replenishment constraint is the clear analytical direction. Third, the large-network floor assumes inactive transshipment and comparable warehouses, so establishing convergence with active rebalancing and heterogeneous warehouses, as well as endogenizing cross-warehouse dependence through assortment and campaign design, would broaden its reach. Fourth, several modeling simplifications invite relaxation, since lead times are folded into flow rates, lost sales below the threshold go uncounted, and a single critical item is modeled rather than a full assortment. Fifth, the study is theoretical, so the empirical priorities are to estimate persistence, correlation, and volatility from platform transactions across categories, to test the predicted ordering of risk for viral products, and to evaluate the replenishment rule against incumbent policies in the field. Sixth, the wider literature on stochastic optimal control and games points to natural enrichments of the decision environment studied here. Savku and coauthors developed regime-switching formulations with jumps, delays, and constrained stochastic differential games [90, 91, 100], machinery that would let the replenishment problem respond to discrete shifts in the demand regime and to strategic interactions among sellers sharing a fulfillment network. The stochastic control treatment of portfolio problems with stochastic interest rates by Korn and Kraft [101] and the portfolio insurance analysis for defined contribution pension plans by Temoçin, Korn, and Selcuk-Kestel [102] show how floor and guarantee constraints, close analogues of the nonnegative-replenishment constraint noted above, are handled in continuous-time financial optimization. On the

computational side, the Markov chain approximation of stochastic hybrid systems with jumps by Temoçin and Weber [103] and the Runge–Kutta schemes with minimal truncation error constants by Öz Bakan, Yılmaz, and Weber [104] offer discretization tools for computing replenishment policies beyond the linear-feedback class. From robust portfolio optimization, the parallelepiped-uncertainty framework of Kara, Özmen, and Weber [105], together with the robust conic regression methodology of Özmen et al. [106] that robustifies the spline estimator invoked for Hurst calibration in Section 6.4 [60], suggests replenishment rules that remain dependable when persistence, correlation, and volatility are estimated with error. The same mathematical structure may also extend to other centrally coordinated resource networks under persistent aggregate shocks, a possibility left to future work rather than claimed here.

Author contributions

Hui Sun: Writing–review & editing, Writing–original draft, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Use of Generative-AI tools declaration

During the revision stage, the authors used generative-AI tools to assist with language polishing, organization of revision responses, and clarity improvement. The authors carefully reviewed, verified, and edited all AI-assisted content and take full responsibility for the accuracy, integrity, and originality of the final manuscript.

Data Availability Statement

All relevant data are within the paper.

Funding

This research was supported by the Shanghai Open University Center for Research on Digital Management and Service Innovation (yjzx2403).

Acknowledgments

We thank the editor-in-chief and anonymous reviewers whose insights enhanced this manuscript's quality.

Conflicts of interest

The authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

References

1. J. Miglani, *Global retail e-commerce sales will reach \$ 6.8 trillion by 2028*, Forrester, 2024. Available from: <https://www.forrester.com/blogs/global-retail-e-commerce-sales-will-reach-6-8-trillion-by-2028/>.
2. N. Kim, B. Montreuil, W. Klibi, M. Z. Babai, Network inventory deployment for responsive fulfillment, *Int. J. Prod. Econ.*, **255** (2023), 108664. <https://doi.org/10.1016/j.ijpe.2022.108664>
3. M. Rodríguez-García, I. González-Romero, Á. Ortiz-Bas, J. C. Prado-Prado, E-fulfillment cost management in omnichannel retailing: An exploratory study, *Comput. Ind.*, **159** (2024), 104094. <https://doi.org/10.1016/j.compind.2024.104094>
4. Amazon Staff, Amazon Prime members enjoyed fastest-ever delivery speeds in 2024, and US members saved on average over \$500 on delivery fees, Amazon News, 2025. Available from: <https://www.aboutamazon.com/news/retail/prime-members-us-savings-fastest-delivery-2024>.
5. Capital One Shopping, Amazon logistics statistics, Capital One Shopping Research, 2026. Available from: <https://capitaloneshopping.com/research/amazon-logistics-statistics/>.
6. D. Y. Kim, B. S. Knight, D. Mitrofanov, How online retailers can avoid costly out-of-stock issues, Harvard business review, *Harvard Business Review*, 2024. Available from: <https://hbr.org/2024/10/how-online-retailers-can-avoid-costly-out-of-stock-issues>.
7. H. Ruben, 27% of shoppers will switch retailers, abandon purchase if they come across locked-up products: Report, *Retail Dive*, 2024. Available from: <https://www.retaildive.com/news/locked-up-merchandise-shrink-consumer-behavior-numerator/732111/>.
8. X. Jing, M. Lewis, Stockouts in online retailing, *J. Mark. Res.*, **48** (2011), 342–354. <https://doi.org/10.1509/jmkr.48.2.342>
9. TikTok, TikTok Shop is where shoppers come to discover, TikTok Newsroom, 2025. Available from: <https://newsroom.tiktok.com/en-us/tiktok-shop-is-where-shoppers-come-to-discover>.
10. Z. Yang, TikTok Shop is now the size of eBay, *Wired*, 2025. Available from: <https://www.wired.com/story/tiktok-shop-sales-global-ecommerce/>.
11. X. Chen, Y. Wang, S. Wei, J. Shen, The effect of herd behavior on consumer intention in live streaming e-commerce: The moderating role of interaction, *Int. J. Hum. Comput. Interact.*, **41** (2025), 1674–1687. <https://doi.org/10.1080/10447318.2024.2364464>
12. E. Yazdani, S. Gopinath, S. J. Carson, The role of reviewer badges in the dynamics of online reviews, *Int. J. Res. Mark.*, **41** (2024), 547–566. <https://doi.org/10.1016/j.ijresmar.2024.01.003>
13. S. Hosseini, D. Ivanov, A. Dolgui, Review of quantitative methods for supply chain resilience analysis, *Transp. Res. Part E Logist. Transp. Rev.*, **125** (2019), 285–307. <https://doi.org/10.1016/j.tre.2019.03.001>
14. R. C. Basole, M. A. Bellamy, Supply network structure, visibility, and risk diffusion: A computational approach, *Decis. Sci.*, **45** (2014), 753–789. <https://doi.org/10.1111/deci.12099>
15. P. H. Zipkin, *Foundations of inventory management*, McGraw-Hill, Boston, Springer, 2000.

16. M. Z. Babai, H. Chen, A. A. Syntetos, D. Lengu, A compound-Poisson Bayesian approach for spare parts inventory forecasting, *Int. J. Prod. Econ.*, **232** (2021), 107954. <https://doi.org/10.1016/j.ijpe.2020.107954>
17. R. Qin, D. A. Nembhard, Demand modeling of stochastic product diffusion over the life cycle, *Int. J. Prod. Econ.*, **137** (2012), 201–210. <https://doi.org/10.1016/j.ijpe.2012.01.027>
18. Y.-F. Chen, Herd behavior in purchasing books online, *Comput. Hum. Behav.*, **24** (2008), 1977–1992. <https://doi.org/10.1016/j.chb.2007.08.004>
19. X. L. Shen, K. Z. K. Zhang, S. J. Zhao, Herd behavior in consumers' adoption of online reviews, *J. Assoc. Inf. Sci. Technol.*, **67** (2016), 2754–2765. <https://doi.org/10.1002/asi.23602>
20. L. A. Gil-Alana, C. Poza, Seasonal long memory in retail sales in the G7 countries, *Fractal Fract.*, **8** (2024), 650. <https://doi.org/10.3390/fractalfract8110650>
21. M. L. Puterman, *Markov decision processes: Discrete stochastic dynamic programming*, John Wiley & Sons, New York, Springer, 1994. <https://doi.org/10.1002/9780470316887>
22. C. Guo, S. Fang, Y. He, Y. Zhang, Some empirical studies for the applications of fractional G-Brownian motion in finance, *Quant. Finance Econ.*, **9** (2025), 1–39. <https://doi.org/10.3934/qfe.2025001>
23. R. Cont, D.-A. Fournié, Functional Itô calculus and stochastic integral representation of martingales, *Ann. Probab.*, **41** (2013), 109–148. <https://doi.org/10.1214/11-AOP721>
24. T. E. Duncan, Y. Hu, B. Pasik-Duncan, Stochastic calculus for fractional Brownian motion I. Theory, *SIAM J. Control Optim.*, **38** (2000), 582–612. <https://doi.org/10.1137/S036301299834171X>
25. F. Biagini, Y. Hu, B. Øksendal, T. Zhang, *Stochastic calculus for fractional brownian motion and applications*, Springer, London, 2008. <https://doi.org/10.1007/978-1-84628-797-8>
26. Y. Hu, B. Øksendal, Fractional white noise calculus and applications to finance, *Infin. Dimens. Anal. Quantum Probab. Relat. Top.*, **6** (2003), 1–32. <https://doi.org/10.1142/S0219025703001110>
27. N. Boysen, R. de Koster, F. Weidinger, Warehousing in the e-commerce era: A survey, *Eur. J. Oper. Res.*, **277** (2019), 396–411. <https://doi.org/10.1016/j.ejor.2018.08.023>
28. A. Risberg, A systematic literature review on e-commerce logistics: towards an e-commerce and omni-channel decision framework, *Int. Rev. Retail Distrib. Consum. Res.*, **33** (2023), 67–91. <https://doi.org/10.1080/09593969.2022.2089903>
29. X. Yang, M. Ostermeier, A. Hübner, Winning the race to customers with micro-fulfillment centers: an approach for network planning in quick commerce, *Cent. Eur. J. Oper. Res.*, **32** (2024), 295–334. <https://doi.org/10.1007/s10100-023-00893-x>
30. Y. Wang, S. Minner, Data-driven multi-location inventory placement in digital commerce, *Comput. Ind. Eng.*, **200** (2025), 110842. <https://doi.org/10.1016/j.cie.2024.110842>
31. M. Melacini, S. Perotti, M. Rasini, E. Tappia, E-fulfilment and distribution in omni-channel retailing: a systematic literature review, *Int. J. Phys. Distrib. Logist. Manag.*, **48** (2018), 391–414. <https://doi.org/10.1108/IJPDLM-02-2017-0101>
32. C. Paterson, G. Kiesmüller, R. Teunter, K. Glazebrook, Inventory models with lateral transshipments: A review, *Eur. J. Oper. Res.*, **210** (2011), 125–136. <https://doi.org/10.1016/j.ejor.2010.05.048>

33. B. Y. Ekren, S. K. Mangla, E. E. Turhanlar, Y. Kazancoglu, G. Li, Lateral inventory share-based models for IoT-enabled E-commerce sustainable food supply networks, *Comput. Oper. Res.*, **130** (2021), 105237. <https://doi.org/10.1016/j.cor.2021.105237>
34. Y. Li, C. W. Zobel, Exploring supply chain network resilience in the presence of the ripple effect, *Int. J. Prod. Econ.*, **228** (2020), 107693. <https://doi.org/10.1016/j.ijpe.2020.107693>
35. C. Colon, S. Hochrainer-Stigler, Systemic risks in supply chains: a need for system-level governance, *Supply Chain Manag.*, **28** (2023), 682–694. <https://doi.org/10.1108/SCM-03-2022-0101>
36. F. Habibi, R. K. Chakraborty, A. Abbasi, Evaluating supply chain network resilience considering disruption propagation, *Comput. Ind. Eng.*, **183** (2023), 109531. <https://doi.org/10.1016/j.cie.2023.109531>
37. F. Habibi, R. K. Chakraborty, A. Abbasi, W. Ho, Investigating disruption propagation and resilience of supply chain networks: interplay of tiers and connections, *Int. J. Prod. Res.*, **63** (2025), 6229–6251. <https://doi.org/10.1080/00207543.2025.2470348>
38. C. Ma, L. Zhang, L. You, W. Tian, A review of supply chain resilience: A network modeling perspective, *Appl. Sci.*, **15** (2025), 265. <https://doi.org/10.3390/app15010265>
39. Y. Kim, Y.-S. Chen, K. Linderman, Supply network disruption and resilience: A network structural perspective, *J. Oper. Manag.*, **33–34** (2015), 43–59. <https://doi.org/10.1016/j.jom.2014.10.006>
40. A. Hosseini, A. Goli, Energy-aware dual-channel agricultural supply chain network design under uncertainty, *Comput. Ind. Eng.*, **207** (2025), 111294. <https://doi.org/10.1016/j.cie.2025.111294>
41. S. Rahmani, A. Goli, Robust sustainable canola oil-based biodiesel supply chain network design under supply and demand uncertainty, *Environ. Sci. Pollut. Res.*, **30** (2023), 86268–86299. <https://doi.org/10.1007/s11356-023-28044-4>
42. I. Shahsavani, A. Goli, A systematic literature review of circular supply chain network design: application of optimization models, *Environ. Dev. Sustain.*, **27** (2025), 23483–23514. <https://doi.org/10.1007/s10668-023-03362-2>
43. K. Ma, T. Chen, C. Zheng, Influence of thinking style and attribution on consumer response to online stockouts, *J. Retail. Consum. Serv.*, **43** (2018), 218–225. <https://doi.org/10.1016/j.jretconser.2018.03.014>
44. A. Gunness, H. Oppewal, How mindset and store familiarity impact online stockout responses, *Int. J. Retail Distrib. Manag.*, **48** (2020), 326–347. <https://doi.org/10.1108/IJRDM-02-2018-0034>
45. M. E. Kumar, D. P. Sharma, A. V. Tapar, Out-of-stock justifications and consumers' behavioral outcomes—exploring the role of product type and sales level information in out-of-stock situations, *J. Retail. Consum. Serv.*, **60** (2021), 102458. <https://doi.org/10.1016/j.jretconser.2021.102458>
46. D. M. Piehlmaier, U. Dávila Gamiño, Fear and loathing in the supermarket: the effect of incidental stockouts on consumer anxiety, *J. Mark. Anal.*, (2025), 1–10. <https://doi.org/10.1057/s41270-025-00418-2>
47. V. V. Acharya, L. H. Pedersen, T. Philippon, M. Richardson, Measuring systemic risk, *Rev. Financ. Stud.*, **30** (2017), 2–47. <https://doi.org/10.1093/rfs/hhw088>

48. T. Adrian, M. K. Brunnermeier, CoVaR, *Am. Econ. Rev.*, **106** (2016), 1705–1741. <https://doi.org/10.1257/aer.20120555>
49. S. Kharvi, T. P. M. Pakkala, An optimal inventory policy when purchase price follows geometric Brownian motion process, *OPSEARCH*, **58** (2021), 835–851. <https://doi.org/10.1007/s12597-020-00500-6>
50. D. C. Heath, P. L. Jackson, Modeling the evolution of demand forecasts with application to safety stock analysis in production/distribution systems, *IIE Trans.*, **26** (1994), 17–30. <https://doi.org/10.1080/07408179408966604>
51. L. Xin, D. A. Goldberg, Distributionally robust inventory control when demand is a martingale, *Math. Oper. Res.*, **47** (2022), 2387–2414. <https://doi.org/10.1287/moor.2021.1213>
52. K. Yamazaki, Inventory control for spectrally positive Lévy demand processes, *Math. Oper. Res.*, **42** (2017), 212–237. <https://doi.org/10.1287/moor.2016.0801>
53. R. T. Baillie, Long memory processes and fractional integration in econometrics, *J. Econom.*, **73** (1996), 5–59. [https://doi.org/10.1016/0304-4076\(95\)01732-1](https://doi.org/10.1016/0304-4076(95)01732-1)
54. J. E. Vera-Valdés, Nonfractional long-range dependence: Long memory, antipersistence, and aggregation, *Econometrics*, **9** (2021), 39. <https://doi.org/10.3390/econometrics9040039>
55. N. Zeinali, A. Pourdarvish, An entropy-based estimator of the Hurst exponent in fractional Brownian motion, *Physica A*, **591** (2022), 126690. <https://doi.org/10.1016/j.physa.2021.126690>
56. C. Bender, L. Viitasaari, Fractional brownian motion in financial modeling, *Wiley StatsRef: Statistics Reference Online*, (2016), 1–5. <https://doi.org/10.1002/9781118445112.stat07865>
57. P. Guasoni, Z. Nika, M. Rásonyi, Trading fractional Brownian motion, *SIAM J. Financial Math.*, **10** (2019), 769–789. <https://doi.org/10.1137/17M113592X>
58. B. B. Mandelbrot, J. W. V. Ness, Fractional Brownian motions, fractional noises and applications, *SIAM Rev.*, **10** (1968), 422–437. <https://doi.org/10.1137/1010093>
59. G. Bhardwaj, N. R. Swanson, An empirical investigation of the usefulness of ARFIMA models for predicting macroeconomic and financial time series, *J. Econom.*, **131** (2006), 539–578. <https://doi.org/10.1016/j.jeconom.2005.01.016>
60. F. Yerlikaya-Özkurt, C. Vardar-Acar, Y. Yolcu-Okur, G. W. Weber, Estimation of the Hurst parameter for fractional Brownian motion using the CMARS method, *J. Comput. Appl. Math.*, **259** (2014), 843–850. <https://doi.org/10.1016/j.cam.2013.08.001>
61. J. Beran, *Statistics for long-memory processes*, Chapman & Hall, New York, 1994. <https://doi.org/10.1201/9780203738481>
62. A. W. Lo, Long-term memory in stock market prices, *Econometrica*, **59** (1991), 1279–1313. <https://doi.org/10.2307/2938368>
63. M. S. Taqqu, V. Teverovsky, W. Willinger, Estimators for long-range dependence: An empirical study, *Fractals*, **3** (1995), 785–798. <https://doi.org/10.1142/S0218348X95000692>
64. Z. Y. Chen, Z. P. Fan, M. Sun, Machine learning methods for data-driven demand estimation and assortment planning considering cross-selling and substitutions, *INFORMS J. Comput.*, **35** (2023), 158–177. <https://doi.org/10.1287/ijoc.2022.1251>

65. I. Vallés-Pérez, E. Soria-Olivas, M. Martínez-Sober, A. J. Serrano-López, J. Gómez-Sanchís, F. Mateo, Approaching sales forecasting using recurrent neural networks and transformers, *Expert Syst. Appl.*, **201** (2022), 116993. <https://doi.org/10.1016/j.eswa.2022.116993>
66. J. A. O. Torres, A. M. Santiago, J. R. García-Martínez, B. Y. López-Zapata, J. A. M. López, O. J. R. Zapata, et al., A data-driven approach using recurrent neural networks for material demand forecasting in manufacturing, *Logistics*, **9** (2025), 130. <https://doi.org/10.3390/logistics9030130>
67. Y. Yu, B. Wang, S. Zheng, Data-driven product design and assortment optimization, *Transp. Res. Part E Logist. Transp. Rev.*, **182** (2024), 103413. <https://doi.org/10.1016/j.tre.2024.103413>
68. A. Goli, Efficient optimization of robust project scheduling for industry 4.0: A hybrid approach based on machine learning and meta-heuristic algorithms, *Int. J. Prod. Econ.*, **278** (2024), 109427. <https://doi.org/10.1016/j.ijpe.2024.109427>
69. A. Ala, A. Goli, Incorporating machine learning and optimization techniques for assigning patients to operating rooms by considering fairness policies, *Eng. Appl. Artif. Intell.*, **136** (2024), 108980. <https://doi.org/10.1016/j.engappai.2024.108980>
70. A. Goli, Optimization of renewable energy project portfolio selection using hybrid AIS-AFS algorithm in an international case study, *Sci. Rep.*, **14** (2024), 17388. <https://doi.org/10.1038/s41598-024-68449-w>
71. T. T. Dufera, Fractional Brownian motion in option pricing and dynamic delta hedging: Experimental simulations, *N. Am. J. Econ. Finance*, **69** (2024), 102017. <https://doi.org/10.1016/j.najef.2023.102017>
72. X. Wang, J. Yu, C. Zhang, On the optimal forecast with the fractional Brownian motion, *Quant. Finance*, **24** (2024), 337–346. <https://doi.org/10.1080/14697688.2023.2297730>
73. E. A. G. D. Souza, T. Mlinar, M. V. D. Broeke, S. Creemers, Dynamic programming in inventory management: A review, *Comput. Oper. Res.*, **183** (2025), 107164. <https://doi.org/10.1016/j.cor.2025.107164>
74. Y. Zhang, A. R. Richter, J. G. Shanthikumar, Z. J. M. Shen, Dynamic inventory relocation in disaster relief, *Prod. Oper. Manag.*, **31** (2022), 1052–1070. <https://doi.org/10.1111/poms.13594>
75. J. Wu, X. Chao, Optimal control of a Brownian production/inventory system with average cost criterion, *Math. Oper. Res.*, **39** (2014), 163–189. <https://doi.org/10.1287/moor.2013.0603>
76. X. Luo, Z. Liu, J. Wu, Dynamic pricing and optimal control for a stochastic inventory system with non-instantaneous deteriorating items and partial backlogging, *Mathematics*, **8** (2020), 906. <https://doi.org/10.3390/math8060906>
77. J. Ma, S. Yang, High-dimensional stochastic control models for newsvendor problems and deep learning resolution, *Ann. Oper. Res.*, **339** (2024), 789–811. <https://doi.org/10.1007/s10479-024-05872-2>
78. K. Sazvari, A. Mahmoodi, H. Sadeghi, Joint dynamic pricing and inventory decisions for duopoly retailers with deterioration items and reference price effect, *J. Ind. Prod. Eng.*, **43** (2026), 456–480. <https://doi.org/10.1080/21681015.2025.2580972>

79. X. Wang, Q. Chen, Stochastic optimal control for dynamic pricing and production in fashion retailing: An economic sustainability approach, *Discrete Dyn. Nat. Soc.*, **2024** (2024), 4854557. <https://doi.org/10.1155/ddns/4854557>
80. W. Lu, L. Yan, Dynamic pricing and inventory strategies for fashion products using stochastic fashion level function, *Axioms*, **13** (2024), 453. <https://doi.org/10.3390/axioms13070453>
81. A. Shapiro, Y. Li, Distributionally robust stochastic optimal control, *Oper. Res. Lett.*, **60** (2025), 107234. <https://doi.org/10.1016/j.orl.2024.107234>
82. T. Başar, J. Moon, J. Yong, H. Zhang, Nonstandard linear-quadratic decision making, *Proc. 62nd IEEE Conf. Decis. Control*, 2023, 8901–8920. <https://doi.org/10.1109/CDC49753.2023.10384310>
83. B. Taskesen, D. Iancu, Ç. Koçyigit, D. Kuhn, Distributionally robust linear quadratic control, *Adv. Neural Inf. Process. Syst.*, **36** (2023), 18613–18632. <https://doi.org/10.52202/075280-0817>
84. E. A. Jaber, E. Miller, H. Pham, Linear-quadratic control for a class of stochastic Volterra equations: Solvability and approximation, *Ann. Appl. Probab.*, **31** (2021), 2244–2274. <https://doi.org/10.1214/20-AAP1645>
85. A. A. Zimbidis, Pricing in a competitive insurance market driven by fractional noise, *Variance*, **5** (2011), 55–67. <https://doi.org/10.66573/001c.84251>
86. M. L. Kleptsyna, A. L. Breton, M. Viot, About the linear-quadratic regulator problem under a fractional Brownian perturbation, *ESAIM Probab. Stat.*, **7** (2003), 161–170. <https://doi.org/10.1051/ps:2003007>
87. T. E. Duncan, B. Maslowski, B. Pasik-Duncan, Linear-quadratic control for stochastic equations in a Hilbert space with fractional Brownian motions, *SIAM J. Control Optim.*, **50** (2012), 507–531. <https://doi.org/10.1137/110831416>
88. Y. Han, Y. Sun, Stochastic linear quadratic optimal control problem for systems driven by fractional Brownian motions, *Optim. Control Appl. Methods*, **40** (2019), 900–913. <https://doi.org/10.1002/oca.2523>
89. C. Liu, H. Fan, Y. Li, Two-person zero-sum stochastic linear quadratic control problems with Markov chains and fractional Brownian motion in infinite horizon, preprint, *arXiv preprint*, arXiv:2412.16538, 2024. <https://doi.org/10.48550/arXiv.2412.16538>
90. E. Savku, A stochastic control approach for constrained stochastic differential games with jumps and regimes, *Mathematics*, **11** (2023), 3043. <https://doi.org/10.3390/math11143043>
91. E. Savku, Memory and anticipation: Two main theorems for Markov regime-switching stochastic processes, *Stochastics*, **97** (2025), 1079–1096. <https://doi.org/10.1080/17442508.2024.2427733>
92. R. Pakhira, S. Sarkar, U. Ghosh, Study of memory effect in an inventory model for deteriorating items with partial backlogging, *Comput. Ind. Eng.*, **148** (2020), 106705. <https://doi.org/10.1016/j.cie.2020.106705>
93. P. K. Santra, G. S. Mahapatra, A. Kumar, Effect of reliability and memory on fractional inventory model incorporating promotional effort on demand, *RAIRO Oper. Res.*, **57** (2023), 1767–1784. <https://doi.org/10.1051/ro/2023095>

94. A. Dasgupta, A. S. Mahapatra, P. K. Santra, G. S. Mahapatra, A. K. Shaw, B. Sarkar, Learning and memory effect in a fractional order quantity model incorporating promotion-assisted demand under uncertainty, *J. Ind. Manag. Optim.*, **20** (2024), 3514–3551. <https://doi.org/10.3934/jimo.2024063>
95. M. Jain, H. Indoria, A. Chaudhary, P. Singh, Effects of memory on inventory control and pricing policy for imperfect production with rework process, *RAIRO Oper. Res.*, **59** (2025), 77–99. <https://doi.org/10.1051/ro/2024215>
96. A. F. Momena, R. Pakhira, R. Haque, A. Sobczak, S. P. Mondal, A memory-dependent inventory model with fuzzy price-dependent demand under backlogged shortages, *J. Uncertain Syst.*, **18** (2025), 2550003. <https://doi.org/10.1142/S1752890925500035>
97. R. B. Davies, D. S. Harte, Tests for Hurst effect, *Biometrika*, **74** (1987), 95–101. <https://doi.org/10.1093/biomet/74.1.95>
98. J. Geweke, S. Porter-Hudak, The estimation and application of long memory time series models, *J. Time Ser. Anal.*, **4** (1983), 221–238. <https://doi.org/10.1111/j.1467-9892.1983.tb00371.x>
99. P. M. Robinson, Gaussian semiparametric estimation of long range dependence, *Ann. Stat.*, **23** (1995), 1630–1661. <https://doi.org/10.1214/aos/1176324317>
100. E. Savku, G. W. Weber, A stochastic maximum principle for a Markov regime-switching jump-diffusion model with delay and an application to finance, *J. Optim. Theory Appl.*, **179** (2018), 696–721. <https://doi.org/10.1007/s10957-017-1159-3>
101. R. Korn, H. Kraft, A stochastic control approach to portfolio problems with stochastic interest rates, *SIAM J. Control Optim.*, **40** (2002), 1250–1269. <https://doi.org/10.1137/S0363012900377791>
102. B. Z. Temoçin, R. Korn, A. S. Selcuk-Kestel, Constant proportion portfolio insurance in defined contribution pension plan management, *Ann. Oper. Res.*, **266** (2018), 329–348. <https://doi.org/10.1007/s10479-017-2449-8>
103. B. Z. Temoçin, G. W. Weber, Optimal control of stochastic hybrid system with jumps: A numerical approximation, *J. Comput. Appl. Math.*, **259** (2014), 443–451. <https://doi.org/10.1016/j.cam.2013.10.021>
104. H. Ö. Bakan, F. Yılmaz, G. W. Weber, Minimal truncation error constants for Runge–Kutta method for stochastic optimal control problems, *J. Comput. Appl. Math.*, **331** (2018), 196–207. <https://doi.org/10.1016/j.cam.2017.10.011>
105. G. Kara, A. Özmen, G. W. Weber, Stability advances in robust portfolio optimization under parallelepiped uncertainty, *Cent. Eur. J. Oper. Res.*, **27** (2019), 241–261. <https://doi.org/10.1007/s10100-017-0508-5>
106. A. Özmen, G. W. Weber, İ. Batmaz, E. Kropat, RCMARS: Robustification of CMARS with different scenarios under polyhedral uncertainty set, *Commun. Nonlinear Sci. Numer. Simul.*, **16** (2011), 4780–4787. <https://doi.org/10.1016/j.cnsns.2011.04.001>



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>)