ShortageSim: Simulating Drug Shortages Under Information Asymmetry

Mingxuan Cui*, Yilan Jiang*, Duo Zhou*, Cheng Qian, Yuji Zhang†, Qiong Wang†

University of Illinois Urbana-Champaign
{mc96, yilanj2, duozhou2, chengq9, yujiz, qwang04}@illinois.edu

Abstract

Drug shortages pose critical risks to patient care and healthcare systems worldwide, yet the effectiveness of regulatory interventions remains poorly understood due to information asymmetries in pharmaceutical supply chains. We propose **ShortageSim**, addresses this challenge by providing the first simulation framework that evaluates the impact of regulatory interventions on competition dynamics under information asymmetry. Using Large Language Model (LLM)-based agents, the framework models the strategic decisions of drug manufacturers and institutional buyers, in response to shortage alerts given by the regulatory agency. Unlike traditional game theory models that assume perfect rationality and complete information, ShortageSim simulates heterogeneous interpretations on regulatory announcements and the resulting decisions. Experiments on self-processed dataset of historical shortage events show that ShortageSim reduces the resolution lag for production disruption cases by up to 84%, achieving closer alignment to real-world trajectories than the zero-shot baseline. Our framework confirms the effect of regulatory alert in addressing shortages and introduces a new method for understanding competition in multi-stage environments under uncertainty. We open-source ShortageSim and a dataset of 2,925 FDA shortage events, providing a novel framework for future research on policy design and testing in supply chains under information asymmetry.

Code — https://github.com/Lemutisme/ShortageSim
Extended version — https://arxiv.org/pdf/2509.01813

Introduction

Every day, clinicians across the globe face an painful decision: which patient receives the last piece of a life-saving medication? Drug shortages have escalated from occasional supply disruptions to a persistent global crisis threatening the foundation of modern healthcare. Before the COVID-19 pandemic exposed these vulnerabilities to the public, the global pharmaceutical supply chain has already experienced multiple crises. Canada (Videau, Lebel, and Bussières 2019), Finland (Sarnola and Linnolahti 2019), and France (Benhabib et al. 2020) all reported unprecedented shortage levels, while the United States has averaged over 130 new shortage cases annually for the past decade (American Society of Health-System Pharmacists 2025). More alarmingly, the average shortage duration has surged from 9 months in 2011 (Government Accountability Office 2011) to 14 months in 2016 (Government Accountability Office 2016), with some critical medications remaining unavailable for over 8 years. Patients face delayed surgeries, substitution with less effective alternatives, or no treatment at all. Beyond the immeasurable costs of compromised patient outcomes, these shortages also impose substantial financial burdens. U.S. hospitals alone spend at least $359 million annually managing shortage related logistics (Vizient 2019).

While drug shortages stem from many causes, this work focuses on supply disruption driven shortages, a main trigger of drug shortages in the U.S. These shortages often result from manufacturing capacity reductions due to quality failures, production line breakdowns, or regulatory compliance issues (Hopp, Brown, and Shore 2022). For example, two major manufacturers of propofol recalled their products and halted production in 2009, and the sole remaining supplier could not scale to meet the demand, triggering a national shortage (Woodcock and Wosinska 2013). The propofol case reveals information asymmetry as a major factor that aggravates the drug shortage problem. Manufacturers guard production data as trade secrets, regulators operate with incomplete market visibility, and buyers cannot distinguish temporary disruptions from permanent exits. This information asymmetry transforms manageable supply shocks into prolonged crises, as stakeholders make decisions in isolation that collectively amplify shortages.

Information asymmetry in the U.S. drug supply chain stems primarily from two causes. First, there exists a cost of information sharing. Production capacity and quality control data are commercially sensitive, and disclosing them can expose competitive weaknesses or invite regulatory examinations. Second, there is a persistent lack of effective communication channels between manufacturers, buyers, and regulators. (Hopp, Brown, and Shore 2022) stated manufacturers are reluctant to invest without accurate demand information and regulatory agencies are receives limited information from manufacturers either. Comparative analyses in (Tadrous et al. 2024) further highlight that, unlike Canada's centralized shortage registry, U.S. reporting re-

*These authors contributed equally.
†Corresponding authors.

mains decentralized and voluntary, delaying signal transmission to healthcare systems and making similar policies ineffective. The resulting information asymmetry means the regulator only sees aggregate shortage levels instead of individual capacity losses or buyer inventories.

Recognizing the need for national coordination, the U.S. Food and Drug Administration (FDA) expanded its regulatory authority in 2012 by mandating manufacturers to report potential shortage risks (U.S. Food and Drug Administration 2013). The agency now maintains a public shortage database and issues alerts to inform the market about current or expected shortages. However, these interventions carry unintended risk of triggering hoarding behaviors to build safety stock and further exacerbate shortages. Traditional analytical frameworks using game-theoretic models struggle to capture how real-world decision-makers interpret ambiguous regulatory language, update beliefs dynamically, and respond to sequential market signals. Moreover, the subjective and sometimes irrational stockpiling behavior is often ignored. The emerging Large Language Model (LLM) provides a novel toolset for alleviating the aforementioned limitations (Wei et al. 2024; Ai et al. 2025; Ning et al. 2025). To fill the scientific research gap of evaluating regulatory interventions, this work presents **ShortageSim**, an LLM-powered multi-agent framework for analyzing the role of information sharing through public announcement in addressing drug shortage problems. The framework models FDA regulators, pharmaceutical manufacturers, and healthcare buyers as autonomous agents navigating information-constrained environments. Each stakeholder interprets FDA announcements through its own analysis, forms beliefs about competitors' actions, and makes decisions that ripple through the supply chain.

By creating a controlled experimental environment to compare reactive and proactive policies, we confirm that proactive policies lead to increased stockpiling behavior and worse effect in resolving shortages. ShortageSim enables counterfactual analysis of various FDA communication strategies and serves as a testbed for future research on policy design and evaluation. Our contributions are:

- We introduce the first LLM-based multi-agent simulation framework for drug shortage management under information asymmetry. Beyond traditional mathematical models, our agents interpret regulatory announcements, infer market states from partial information and make context-dependent decisions under uncertainty.
- We collect, process and publicly release a dataset on 2,925 FDA shortage reports and 51 resolved historical event trajectories, addressing the current reliance of drug shortage research on synthetic or proprietary data.
- Our framework achieves consistently closer alignment with historical shortage trajectories than the zero-shot baseline across four LLM providers, reducing resolution lag for production disruption cases by up to 84%.
- We provide a flexible testbed for counterfactual policy evaluation, enabling controlled evaluation of regulatory strategies, communication designs, and agent decision protocols in information constrained environments.

Related Works

Drug Shortage Management. Drug shortages in the United States are well documented in white papers, government reports, and case studies that examine their causes, impacts, and scope (Kweder and Dill 2013; Fox, Sweet, and Jensen 2014; Yurukoglu, Liebman, and Ridley 2017; McLaughlin et al. 2013; Phuong et al. 2019; Patel, Kesselheim, and Rome 2021; U.S. Food and Drug Administration. 2016; Edwards 2021). Most of these studies are descriptive, and root causes of the problem are yet to be investigated rigorously through analytical or empirical methods.

Operations management (OM) literature views drug markets through the lens of optimization and control (Jia and Zhao 2017; Tucker et al. 2020; Swinney, Wu, and Zhang 2024; Zhao, Jia, and Zhao 2025). The resulting static models assume rational decisions and overlook dynamic behaviors and information asymmetries (Zhang, Li, and Li 2023), thereby obscuring sparse but critical knowledge and limiting the accuracy of models trained on such data (Zhang et al. 2024, 2025a). Empirical evidence on policy effectiveness is likewise limited: although (Lee et al. 2021) show that mandatory disruption reporting mitigates shortages, the underlying mechanisms are still unclear. We address these gaps by examining how regulatory announcements influence supply chain decisions and shortage outcomes.

Agents for Pharmaceutical Supply Chain. Multi-agent systems have been extensively applied to pharmaceutical supply chains, though few has applied it to study the impact of regulations on mitigating drug shortages. Traditional frameworks demonstrate manufacturer-distributor interactions using platforms like MATLAB (Pourghahreman, Ghatari, and Moosivand 2018) and AnyLogic (Bozdoğan, Görkemli Aykut, and Demirel 2023), focusing on disruption response and risk management strategies (Tan et al. 2020). Recent drug shortage prediction models achieve high accuracy using machine learning approaches on historical data (Liu et al. 2021; Pall et al. 2023; Postma et al. 2024). However, these systems treat regulatory actions as external constraints rather than modeling FDA as an active strategic agent, and prediction systems remain disconnected from dynamic stakeholder response modeling.

The integration of Large Language Models into multi-agent systems represents a paradigm shift in complex system simulations (Li et al. 2023a; Wang et al. 2024; Li et al. 2024a; Qian et al. 2025a,b; Xu et al. 2025). Microsoft's OptiGuide framework (Li et al. 2023a) pioneered LLM integration for supply chain optimization. Recent frameworks demonstrate remarkable scaling and strategic reasoning capabilities (Li et al. 2023b; Hong et al. 2024; Zhang et al. 2025b; Song, Meng, and King 2024). Game-theoretic applications show that multi-agent LLM architectures significantly outperform single LLMs in simulating strategic behavior (Li et al. 2024b), while economic simulations like EconAgent (Li et al. 2023c) demonstrate the ability to produce realistic macro-level phenomena arising from micro-level interactions.

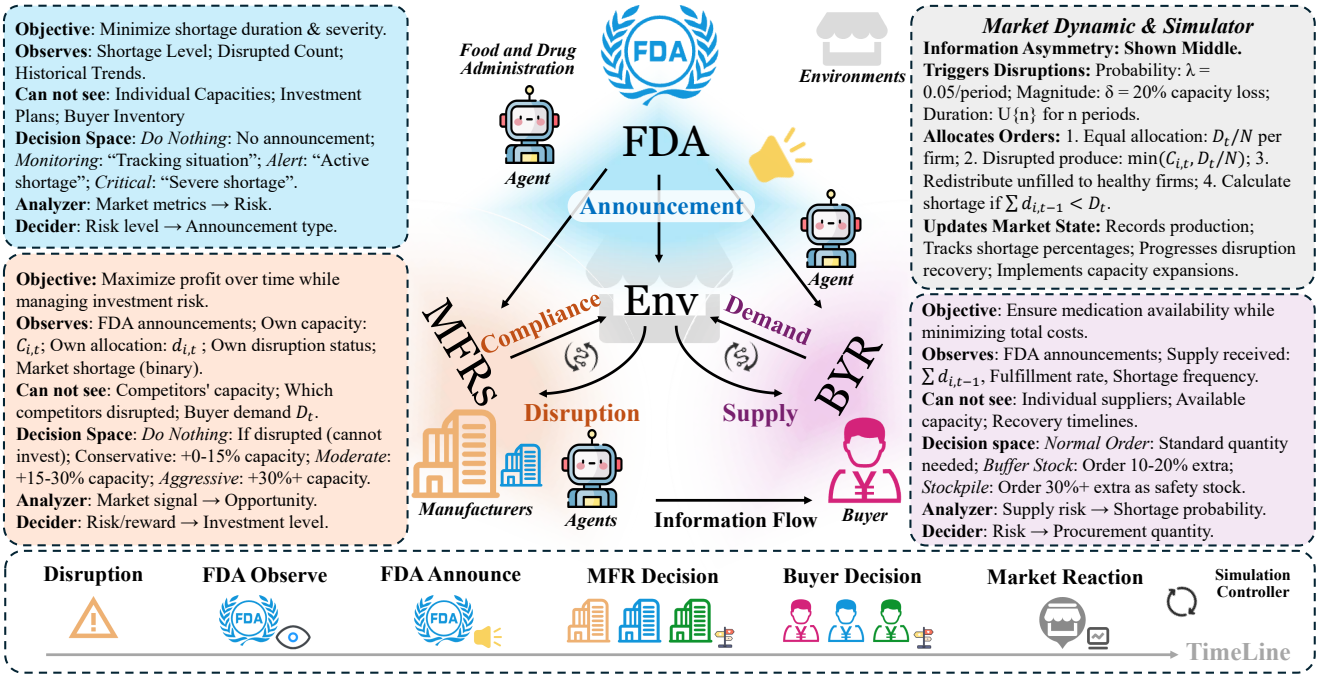

Figure 1: ShortageSim Architecture. The framework models the drug supply chain dynamics through three agent roles: FDA regulator, buyer consortium, and n competing manufacturers, coordinated by an Environment module. Each agent operates under information asymmetry (shown in **Information Availability**) and follows a two-stage LLM pipeline (Analyze → Decide). The timeline illustrates the sequential decision process from supply disruption through market reaction (details in **Sequential Decision Timeline**), capturing realistic stakeholder responses to regulatory announcements during drug shortage events.

Methodology: Shortage Simulation Agents

System Architecture Overview

To address the challenge of modeling decision-making under information asymmetry, ShortageSim employs a multi-agent simulation framework where LLM-powered agents interact within a drug supply chain. This framework design captures the nuanced interpretation of regulatory signals and market conditions, characterizing real-world responses to shortage events. ShortageSim models the market dynamics between manufacturers and buyers, under the regulatory interventions by the FDA during drug shortage events. All agents make decisions under partial information, reflecting the actual information asymmetry in drug supply chains.

As shown in Figure 1, the framework consists of four core components: (1) the **Environment** module manages market dynamics and state transitions while maintaining information barriers among agents, (2) the **Agents** system implementing role-specific decision-making through LLMs that interpret partial information and form strategies, (3) the **Information Flow** that controls inter-agent communication to simulate the information asymmetries, and (4) the **Simulation Controller** that operates the overall execution flow and records all decisions for subsequent analysis.

Market Mechanics and Dynamics

We consider a drug market with n competing *manufacturers*, a consortium of health care providers as the *buyer* and FDA as the *regulator*. Each manufacturer i possesses a *production capacity* c_i. The drug in question is medically necessary, with *patient demand* modeled deterministically as D_0. Initially, the market is assumed to be in equilibrium, meaning the total production exactly meets the patient demand D_0 until disruptions happen.

We choose not to include intermediary entities like distributors and group purchasing organizations (GPOs) in the model, as they primarily facilitate logistics and negotiate prices respectively without affecting production capacity or demand dynamics. Their detailed operational data are not publicly available either. A detailed discussion of this modeling choice is provided in the Appendix.

Disruption Modeling. Our model accounts for internal supply disruptions on manufacturers' production capacities. At the start of each period, each manufacturer faces an independent probability λ of experiencing disruption. When a disruption occurs to a manufacturer, its production capacity is reduced by a fixed fraction δ for t periods and returns to the regular level afterwards, where t is sampled from discrete uniform distributions.

Demand Allocation Mechanism. In the drug supply chain, buyers are observed to stockpile when they anticipate the cumulative orders will exceed the total production capacity (Sterman and Dogan 2015), driven by aversion to stockouts. To differentiate from the actual patient demand D_0, we define buyers' order quantity as *demand*, denoted by D. Ex-

cess orders ($D - D_0$) are maintained as buyers' inventory for future periods.

In the nominal model, we assume all manufacturers are symmetric and the total demand D is evenly allocated in equilibrium. If disruption affects a subset of manufacturers $M_D \subset [n] = \{1, 2, \cdots, n\}$, their production becomes

$$q_i = \min\left\{\frac{D}{n}, c_i(1-\delta)\right\}, i \in M_D. \quad (1)$$

The resulting unfilled demand is

$$D_{\text{unfill}} = \sum_{i \in M_D} \left(D/n - c_i(1-\delta)\right)^+,$$

where $x^+ = \max\{x, 0\}$. This unfilled demand is reallocated evenly among other undisputed manufacturers $M_U = [n] \backslash M_D$, whose production is updated accordingly as

$$q_i = \min\left\{\frac{D}{n} + \frac{D_{\text{unfill}}}{|M_U|}, c_i\right\}, i \in M_U. \quad (2)$$

The total *supply* is defined as $Q = \sum_{i \in [n]} q_i$ and *shortage* is calculated as $(D-Q)^+$. We relax the symmetry assumption in the **Experiment** section to analyze different agent behavior under unequal market shares.

Financial Model The simulation's economic framework emphasizes the deviation between individual incentives and collective outcomes, highlighting why market coordination often fails during shortages.

Each period undisputed manufacturers could ramp up their capacities at a unit *cost* Ct to fulfill the unmet patient demand. Capacity expansion typically involves constructing new production lines and is assumed to be available one time period after the investment decision, requiring manufacturers to anticipate future market conditions from regulatory signals. Moreover, if all manufacturers expand simultaneously based on the same signal, potential profits will be diluted by overcapacity, creating a coordination challenge.

On the other side, buyers pay the *price* p to purchase a unit of drug, incurring a *holding cost* h for unit leftover and a *penalty* s for unit unmet patient demand. As health providers, buyers are adverse to stockouts, but maintaining large safety stock imposes inventory costs and forces buyers to carefully calibrate their response to shortage signals.

Agent Design and Two-Stage Decision Pipeline

Agent Architecture. Each agent in ShortageSim follows a two-stage decision pipeline of information processing followed by strategic decision-making. This architectural choice directly simulates the dynamics of decision making based on incomplete and ambiguous market signals among stakeholders. Firstly, **Collector & Analyst** receives unstructured context, including FDA announcement text, aggregate market metrics, and historical demand trend, then produces a structured representation capturing the agent's interpretation of the current situation. Secondly, **Decision Maker** takes this structured state representation and generates decisions with detailed reasoning. The two-stage design examines the effects of FDA communications through reasoning output and captures the discrepancies in the reasoning process across different agents. More importantly, this stage does not require the output decisions to be optimal, an important assumption for realism. As (Croson et al. 2014) observed, some stakeholders mistrust the rationality of their competitors' decisions and respond with suboptimal, even unreasonable choices.

Role-Specific Agents Design

Manufacturer Agents represent pharmaceutical companies making capacity investment decisions under uncertainty about competitor status and future market conditions. The objective of each manufacturer is to maximize its own profit, by balancing the opportunity of profiting from competitors' disruptions and the financial losses from overcapacity when anticipated disruptions do not occur. To support effective decision-making, manufacturer agents are instructed to: 1) understand the overall market structure, demand allocation mechanisms investment opportunity and risk of overcapacity; 2) infer market conditions from observable signals, regulatory announcements and the amount of demand it gets; 3) identify profit opportunities, assess market risks and anticipate competitors' potential decisions; 4) determine whether and how much to invest in capacity expansion.

Buyer Agent represents a healthcare consortium facing a variant of the newsvendor problem under supply uncertainty. With patient safety as their top priority, the buyers can stockpile to build safety stock, but must balance the risk of stockout against the cost of purchasing and holding excess inventory. To simulate realistic behavior, the buyer agent is guided to: 1) understand the market structure, own cost parameters and utility objectives; 2) interpret regulatory signal and detect changes in supply availability; 3) analyze current inventory levels and determine risk of stockout; 4) make purchasing decisions that balance the additional cost and the potential impact of stockouts.

FDA Agent operates under a reactive policy framework, with a goal of mitigating drug shortage through public announcements without disclosing any individual stakeholder's private information. The announcements should provide sufficient warnings to induce manufacturers' investment during shortage while preventing buyers' panic hoarding. To issue effective announcements, the FDA agent is instructed to: 1) understand market structure and its regulatory role in mitigating drug shortage; 2) detect supply-demand imbalances and predict potential shortages; 3) anticipate stakeholders' response to public announcements; and 4) determine the content and severity level of public announcements to balance urgency with stability.

Information Asymmetry

Sequential Decision Timeline. As illustrated in Figure 1, each simulation period follows an organized sequence of events that mimic information asymmetry in practice, ensuring agents make decisions based on partial information available to them in reality. A period begins with a disruption phase in which affected firms learn their remaining capacities and recovery times. Next, after observing new disruptions and aggregate market outcomes from the previous

period, the FDA decides whether to issue a public announcement. Unaffected manufacturers then simultaneously decide on capacity investments, based on allocated demand in the last period and the latest FDA announcements. The buyer simultaneously determines procurement quantities based on the amount of demand served in the last period and the same FDA announcements. Finally, the market clears through the aforementioned allocation mechanism, with excess supply or shortage levels determined accordingly..

Information Availability. The framework enforces strict information separation that mirrors real pharmaceutical supply chains and complicates decision-making process.

Manufacturers possess complete knowledge of their own status, including current capacity, recovery time if being disrupted and past allocated demand. However, they cannot directly observe competitors status and must infer through regulatory signals, bringing challenges in building capacity investment strategies. For example, increase in its own demand is not sufficient for a manufacturer to determine whether the surge is caused by a disruption at another manufacturer or demand fluctuation on the buyer side.

Buyer observes the total supply received each period, but not the allocation across manufacturers or their production capacities. Specifically, when a disruption happens, the buyer has no information on the remaining recovery time, bringing challenges to the buyer's purchasing strategy.

FDA operates outside the supply chain, and thus has no access to private information of manufacturers and the buyer except for aggregate shortage levels and the mandated report of disruptions. Like the real-world agency, no information on capacity and investment of manufacturers and inventory levels of the buyer is visible to the FDA in our simulation .

Table 1 summarizes what variables are accessible to manufacturers (Mfrs), buyer, and the FDA, where ✓ indicates full availability, ✗ means no availability, and † stands for partial availability, that a manufacturer can only observe its own status, but not those of their competitors.

Experiments

We evaluate ShortageSim on a self-processed FDA drug shortage dataset to verify its alignment with real-world trajectories and evaluate the effect of communication by FDA.

Variable Privacy	Mfr.	Buyer	FDA
Disruption	†	✗	✓
Capacity	†	✗	✗
Investment	†	✗	✗
Allocated Demand	†	✗	✗
Total Supply	✗	✓	✓
Total Demand	✗	✓	✓
Shortage	✗	✓	✓
Buyer Inventory	✗	✓	✗
FDA Signal	✓	✓	✓

Table 1: Information availability for each agent. † = private/-partial information; ✓ = known; ✗ = not known.

The framework offers a new method for understanding competition in complex, multi-stage environments under uncertainty. It also provides a controlled experiment pipeline for systematically comparing regulatory policies and agent decision protocols in future research.

Experiment Settings

To evaluate the ability of ShortageSim reproducing historical shortage dynamics, we collect and preprocess 2,925 FDA reported drug shortage events and 51 resolved trajectories in 2023 and 2024. These cases are split into two ground-truth trajectory sets based on their reported shortage reasons. All reported drugs are grouped and identified by their National Drug Code (NDC), which uniquely links manufacturers to specific drug products. The **FDA-Disc** set comprises events that the FDA explicitly cites manufacturer production discontinuation as the primary cause, while the **FDA-NR** set contains cases with no specific cause reported. Each trajectory is truncated upon shortage resolution so its length represents the total resolution time in quarters. Data curation details are provided in Appendix.

Each simulation is specified by the number of manufacturers (n), shortage duration (T), disruption probability (λ), and disruption magnitude (δ). One period corresponds to a calendar quarter. Following the FDA dataset statistics, we simulate cases with $T \in \{4, \ldots, 12\}$ and $n \in \{2, \ldots, 10\}$. Monopoly scenarios ($n = 1$) are excluded because our model reduces to a trivial case with no decision to be made. For the FDA-Disc dataset, the disruption magnitude of each drug is estimated by the proportion of discontinued NDC packages reported by the FDA. Moreover, each trajectory starts with an enforced disruption and no separate disruption probability is specified. Details of parameter selection are described in the Appendix. At each period t, the framework records total demand, total supply, shortage amount, buyer's on-hand inventory, FDA announcements, and the set of disrupted manufacturers. These variables combined with all agent decisions generate a simulation trajectory.

We test GPT-4o (Hurst et al. 2024), Gemini 2.5 Flash (Comanici et al. 2025), Claude Sonnet 4.5 (Anthropic 2025) and Deepseek V3.2 Exp (Liu et al. 2024) with temperature 0.3 as LLM backbones for all agents in our simulation. An ablation study on different temperatures is provided in the Appendix to show performance robustness. For comparison, we introduce a zero-shot baseline model that generates trajectories of supply, demand, and shortage dynamics without iterative decision-making. The input to the zero-shot model includes the market structure, financial parameters, demand allocation mechanisms, investment options, and stockpiling behavior, serving as a reference to assess the value of our framework in addressing sequential decision-making under competition and information asymmetry.

Evaluation Metrics and Results

We evaluate the alignment between the ground-truth and simulated trajectories based on Resolution-Lag Percentage (RLP) and FDA Intervention Percentage (FIP).

(1) Resolution-Lag Percentage (RLP). The primary metric measures the accuracy of resolution time in simulation. Each ground-truth trajectory is truncated upon shortage resolution so the reference resolution time for shortage case j is the trajectory length: $t_j^{\text{GT}} = T_j$. Let $s_{j,t}^{(\cdot)}$ be the shortage level in period t , where $(\cdot) \in \{\text{GT}, \text{sim}\}$ denotes ground-truth or simulation. The resolution time in simulation is considered as the earliest sustained clearance time. With a shortage-tolerance threshold of $\varepsilon = 0.001$, define

$$t_j^{\text{sim}} = \min\Big\{ t \mid \max_{t \le u \le T_j} s_{j,u}^{\text{sim}} \le \varepsilon \Big\},$$

and set $t_j^{\text{sim}} = T_j + 1$ if no such t exists. The *resolution lag percentage* for case j is defined as

$$\text{RLP}_j = \frac{t_j^{\text{sim}} - t_j^{\text{GT}}}{t_j^{\text{GT}}} \times 100\%. \tag{3}$$

Positive values indicate slower simulated resolution than the ground truth, and the resolution is faster if RLP_j is negative.

(2) FDA Intervention Percentage (FIP). The secondary metric presents the frequency of FDA intervention. Suppose drug shortage trajectory j has T_j periods, let $I_{j,t}$ be the indicator function of FDA issuing an announcement in period $t = 1, 2, \ldots, T_j$. The *FDA Intervention Percentage* is the fraction of periods that FDA makes announcements:

$$\text{FIP}_j \;=\; \frac{1}{T_j} \sum_{t=1}^{T_j} I_{j,t} \;\times\; 100\%. \tag{4}$$

Results. We present model performance in both datasets using GPT-4o in Table 2. To demonstrate the robustness of performance across different LLM providers, results obtained from multiple backbones are reported in Table 3. All experiments are repeated three times, and RLP, FIP are reported as mean $\pm$ standard deviation.

Analysis. Tables 2 and 3 show that ShortageSim achieves higher accuracy in predicting resolution time than the zero-shot baseline across all four LLM providers for the FDA-Disc dataset. The mean RLP values of ShortageSim are consistently closer to zero, with GPT-4o achieving only a 4.5% mean resolution lag, demonstrating the framework's effectiveness in reproducing real-world dynamics. This improvement in resolution-time prediction accuracy is robust across the four tested LLM providers. Wilcoxon signed-rank tests

Provider: GPT-4o			
Model	**Dataset**	**RLP(%)**	**FIP(%)**
ShortageSim	FDA-Disc	**4.5±3.4**	82.6±3.0
	FDA-NR	-34.8±8.7	30.0±5.5
Zero-shot	FDA-Disc	-28.3±0.2	92.9±0.1
	FDA-NR	-23.1±1.6	90.4±18.6

Table 2: Evaluation metrics for two ground-truth datasets under ShortageSim and Baseline using GPT-4o.

Dataset: FDA-Disc			
Provider	**Model**	**RLP(%)**	**FIP(%)**
Gemini	ShortageSim	-35.8 ± 3.0	79.3 ± 4.5
	Zero-shot	-51.9 ± 2.1	88.3 ± 1.5
Claude	ShortageSim	-9.4 ± 1.4	69.1 ± 2.3
	Zero-shot	-32.7 ± 0.5	75.9 ± 0.7
Deepseek	ShortageSim	19.1 ± 0.1	84.4 ± 1.3
	Zero-shot	-26.6 ± 1.3	63.7 ± 2.1

Table 3: Evaluation metrics for the FDA-Disc dataset under ShortageSim and the zero-shot baseline across different LLM providers.

on absolute RLP values against the zero-shot baseline confirms the statistical significance of these gains ($p < 0.05$). Details of our statistical tests are provided in the Appendix.

The average FIP is higher for FDA-Disc than for FDA-NR, consistent with the dataset division criteria. Simulations of cases in the FDA-Disc dataset begin with forced disruption and trigger more frequent FDA announcements, while the FDA-NR dataset exhibits greater randomness in disruption appearance and duration. This mismatch between simulation setup and underlying characteristics of the FDA-NR dataset explains the observed deviation in model outputs.

To evaluate the effect of FDA communication under limited information, we also compare ShortageSim with a variant that disables FDA's announcement mechanism on the FDA-Disc dataset using GPT-4o. The no-announcement setting yields a significantly higher RLP (19.1% vs. 4.5%), highlighting the effect of FDA's announcements in resolving drug shortage under information asymmetry. Without the regulator's intervention, the buyer remains stockpiling behavior after detecting disruption from unsuccessful delivery, while manufacturer becomes more conservative in expanding capacity due to the lack of trusted public signals.

Effect of Market Competition

Different from the standard game theory model, our framework introduces a new paradigm for understanding competition in multi-stage environments under uncertainty.

We first analyze how manufacturer's investment decisions vary with market competition intensity, measured by the number of manufacturers. The 40 trajectories in the FDA-NR dataset are partitioned into four segments with $\{2\}$, $\{3\}$, $(3, 5]$, and $(5, 10]$ manufacturers, containing 15, 10, 10, and 5 trajectories, respectively. For trajectory j with M_j manufacturers and T_j periods, the *average supply per period* is defined as

$$\bar{q}_j = \frac{1}{T_j} \sum_{t=1}^{T_j} \sum_{m=1}^{M_j} q_{m,t},$$

where $q_{m,t}$ is the quantity produced by manufacturer m in period t. As Figure 2 shows, the mean of $\bar{q}_j$ decreases as the number of manufacturers increases, reflecting that more intense competition leads to lower profit margin, which discourages undisrupted manufacturers to invest. In duopoly

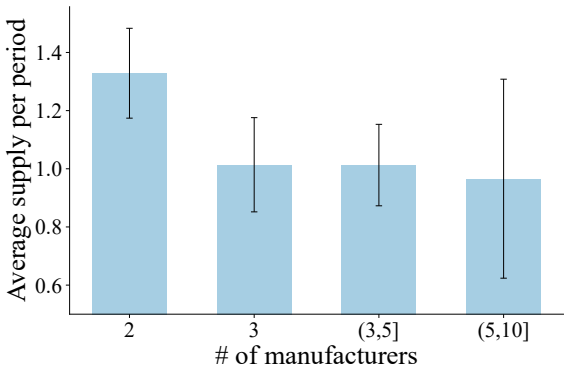

Figure 2: Effect of market competition on supplier investment (average supply per period)

markets, manufacturers invest more heavily in adding new capacities because they are guaranteed to capture unmet demand when the only competitor is disrupted. On the contrary, as competition intensifies, rationales generated by manufacturer agents increasingly emphasize terms such as *maintaining* and *competitors*, mirroring their conservative strategies. This observed decrease in supply as competition intensify aligns with empirical findings in (Lee et al. 2021). Instead of equilibrium analysis under the assumption of perfect information, our LLM-based simulation system reproduces these patterns. This system serves as a novel approach for understanding competitive dynamics and evaluating effects of regulatory interventions for drug shortage problems.

Since competition intensity is also affected by the market share among manufacturers, we relax the symmetry assumption and analyze competition under unequal market shares. Specifically, we enable asymmetric demand allocation proportional to manufacturers' available capacities. A case study with three manufacturers holding 60%, 30%, and 10% of the market shows asymmetric responses: when one of the smaller manufacturers is disrupted, only the dominant one expands production capacity, reflecting the expectation that a larger market share enables it to capture the unmet demand. Conversely, when the largest manufacturer is disrupted, the two smaller competitors expand their capacities proportional to their respective market shares and as a result, the total supply of the drug remains stable. These results indicate that market share affect risk attitudes and coordination incentives of manufacturers during shortages. Future work can extend these findings to jointly model coordination and competition among manufacturers.

Counterfactual Policy Experiments

Our framework demonstrate substantial improvement in prediction accuracy of shortage resolution; it also provides flexibilities to test different policy designs. The FDA currently employs both proactive measures (e.g. mandating disruption reporting) and reactive interventions (e.g. expediting new drug approvals) to mitigate shortages. However, stockpiling behavior occurs when buyers anticipate future shortages, so early warning may inadvertently intensify shortages. This double-edged sword effect persist in other early-warning systems (Boettiger and Hastings 2012; Nof, Yagoda-Biran, and Zwebner 2025). Our framework is the first to quantitatively explore this trade-off in the context of drug shortage early detection, conducting counterfactual experiments in comparing two FDA policy regimes: a reactive policy that issues alerts only after confirmed shortages, and a proactive policy that warns of potential disruptions based on early indicators. Using proactive policy in the FDA-NR dataset, the average buyer stockpiling quantity increases from 7.7% to 16.9%, yet total supply is largely unchanged. When the FDA makes advisory announcements, uncertainty about future demand is the major reason that manufacturers not investing in additional capacity. In contrast, because holding costs are lower than stockout penalties, buyers face less downside risk and tend to increase stockpiling.

These results highlight that ShortageSim is not limited to reproducing historical trajectories, but can also perform counterfactual policy evaluations. By modifying agent prompts, the framework can simulate market response to different regulatory strategies, for example, specific content for manufactures and buyers, quantitative vs. qualitative details and the linguistic framing of FDA communications.

Conclusion

We develop **ShortageSim** , a multi-agent simulation framework that captures strategic interactions of regulators, manufacturers, and buyers under information asymmetry and random supply disruptions. Along with the framework, we collect, process and open-source a real-world drug shortage dataset from the FDA shortage list, as a public benchmark for future research. Experiment results on this dataset demonstrate that ShortageSim achieves better alignments with high-temporal resolution than the zero-shot baseline. Beyond reproducing historical trajectories, our framework also enables controlled comparisons of regulatory strategies and realistic agent decision protocols. By building this framework, we contribute a new methodology to study dynamic decision-making under competition and the effect of regulatory interventions, in the presence of random disruptions and information asymmetry.

Limitations and Future Work

Although ShortageSim provides a flexible testbed for policy and behavioral analysis, several modeling assumptions can be refined to improve external validity. First, the effect of disruption is represented by 20% reduction of a manufacturer's production capacity and the disruption duration is sampled from the same uniform distribution. In practice, these parameters vary across drugs. Future extensions could incorporate data-driven estimates to capture this heterogeneity. Second, supply disruption are modeled as independent events, but in reality, they may arise from shared causes like raw material shortages or natural disasters. Modeling correlated disruptions would more accurately reflect systemic risks. Lastly, although intermediary agencies like wholesalers do not directly change supply or demand, including them could provide deeper insights into coordination challenges within the drug supply chain.

Ethical Statement

All analyses in this work rely on anonymized, publicly available FDA shortage list and therefore raise no privacy concerns. We discuss the potential risks of hoarding behavior and show these effects can be reproduced in simulation for policy stress-testing. More broadly, our framework is intended as a research testbed for evaluating policy interventions rather than a prescriptive decision-making tool, thereby reducing the risk of misuse.

Acknowledgments

We thank the anonymous reviewers for their constructive feedback and Zichen Tian for editing the manuscript.

References

Ai, M.; Wei, T.; Chen, Y.; Zeng, Z.; Zhao, R.; Varatkar, G.; Rouhani, B. D.; Tang, X.; Tong, H.; and He, J. 2025. ResMoE: Space-efficient Compression of Mixture of Experts LLMs via Residual Restoration. In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 1*, 1–12.

American Society of Health-System Pharmacists. 2025. Drug Shortages Statistics. DrugShortagesStatistics. Last Accessed: 2025-07-27.

Anthropic. 2025. Introducing Claude Sonnet 4.5. https://www.anthropic.com/news/claude-sonnet-4-5.

Benhabib, A.; Ioughlissen, S.; Ratignier-Carbonneil, C.; and Maison, P. 2020. The French reporting system for drug shortages: description and trends from 2012 to 2018: an observational retrospective study. *BMJ open*, 10(3): e034033.

Boettiger, C.; and Hastings, A. 2012. Quantifying limits to detection of early warning for critical transitions. *Journal of the Royal Society Interface*, 9(75): 2527–2539.

Bozdoğan, A.; Görkemli Aykut, L.; and Demirel, N. 2023. An agent-based modeling framework for the design of a dynamic closed-loop supply chain network. *Complex & Intelligent Systems*, 9(1): 247–265.

Comanici, G.; Bieber, E.; Schaekermann, M.; Pasupat, I.; Sachdeva, N.; Dhillon, I.; Blistein, M.; Ram, O.; Zhang, D.; Rosen, E.; et al. 2025. Gemini 2.5: Pushing the frontier with advanced reasoning, multimodality, long context, and next generation agentic capabilities. *arXiv preprint arXiv:2507.06261*.

Croson, R.; Donohue, K.; Katok, E.; and Sterman, J. 2014. Order Stability in Supply Chains: Coordination Risk and the Role of Coordination Stock. *Production and Operations Management*, 23(2): 176–196.

Edwards, E. 2021. The U.S. Needs to Reimagine Its Pharma Supply Chain. https://hbr.org/2021/08/the-u-s-needs-to-reimagine-its-pharma-supply-chain. Last Accessed: 2025-07-27.

Fox, E. R.; Sweet, B. V.; and Jensen, V. 2014. Drug Shortages: A Complex Health Care Crisis. *Mayo Clinic Proceedings*, 89(3): 361–373.

Government Accountability Office. 2011. Drug shortages: FDA's ability to respond should bestrengthened. https://www.gao.gov/products/gao-12-116. Last Accessed: 2025-07-27.

Government Accountability Office. 2016. Drug Shortages: Certain Factors Are Strongly Associated with This Persistent Public Health Challenge. https://www.gao.gov/products/gao-16-595. Last Accessed: 2025-07-27.

Hong, S.; Zhuge, M.; Chen, J.; Zheng, X.; Cheng, Y.; Wang, J.; Zhang, C.; Wang, Z.; Yau, S. K. S.; Lin, Z.; Zhou, L.; Ran, C.; Xiao, L.; Wu, C.; and Schmidhuber, J. 2024. MetaGPT: Meta Programming for A Multi-Agent Collaborative Framework. In *The Twelfth International Conference on Learning Representations*.

Hopp, W. J.; Brown, L.; and Shore, C. 2022. *Building Resilience into the National Medical Product Supply Chains*. Washington, DC: The National Academies Press. ISBN 978-0-309-27469-2.

Hurst, A.; Lerer, A.; Goucher, A. P.; Perelman, A.; Ramesh, A.; Clark, A.; Ostrow, A.; Welihinda, A.; Hayes, A.; Radford, A.; et al. 2024. Gpt-4o system card. *arXiv preprint arXiv:2410.21276*.

Jia, J.; and Zhao, H. 2017. Mitigating the U.S. Drug Shortages Through Pareto-Improving Contracts. *Production and Operations Management*, 26(8): 1463–1480.

Kweder, S. L.; and Dill, S. 2013. Drug Shortages: The Cycle of Quantity and Quality. *Clinical Pharmacology & Therapeutics*, 93(3): 245–251.

Lee, J.; Lee, H. S. H.; Shin, H.; and Krishnan, V. 2021. Alleviating Drug Shortages: The Role of Mandated Reporting Induced Operational Transparency. *Management Science*, 67(4): 2326–2339.

Li, B.; Mellou, K.; Zhang, B.; Pathuri, J.; and Menache, I. 2023a. Large language models for supply chain optimization. *arXiv preprint arXiv:2307.03875*.

Li, G.; Hammoud, H. A. A. K.; Itani, H.; Khizbullin, D.; and Ghanem, B. 2023b. CAMEL: Communicative Agents for "Mind" Exploration of Large Language Model Society. In *Thirty-seventh Conference on Neural Information Processing Systems*.

Li, J.; Lai, Y.; Li, W.; Ren, J.; Zhang, M.; Kang, X.; Wang, S.; Li, P.; Zhang, Y.-Q.; Ma, W.; et al. 2024a. Agent hospital: A simulacrum of hospital with evolvable medical agents. *arXiv preprint arXiv:2405.02957*.

Li, J.; Zhang, Q.; Yu, Y.; Fu, Q.; and Ye, D. 2024b. More agents is all you need. *arXiv preprint arXiv:2402.05120*.

Li, N.; Gao, C.; Li, M.; Li, Y.; and Liao, Q. 2023c. Econagent: large language model-empowered agents for simulating macroeconomic activities. *arXiv preprint arXiv:2310.10436*.

Liu, A.; Feng, B.; Xue, B.; Wang, B.; Wu, B.; Lu, C.; Zhao, C.; Deng, C.; Zhang, C.; Ruan, C.; et al. 2024. Deepseek-v3 technical report. *arXiv preprint arXiv:2412.19437*.

Liu, I.; Colmenares, E.; Tak, C.; Vest, M.-H.; Clark, H.; Oertel, M.; and Pappas, A. 2021. Development and validation of a predictive model to predict and manage drug

shortages. *American Journal of Health-System Pharmacy*, 78(14): 1309–1316.

McLaughlin, M.; Kotis, D.; Thomson, K.; Harrison, M.; Fennessy, G.; Postelnick, M.; and Scheetz, M. H. 2013. Effects on patient care caused by drug shortages: a survey. *Journal of Managed Care Pharmacy*, 19(9): 783–788.

Ning, X.; Fu, D.; Wei, T.; Xu, W.; and He, J. 2025. Graph4MM: Weaving Multimodal Learning with Structural Information. In *Forty-second International Conference on Machine Learning*.

Nof, R. N.; Yagoda-Biran, G.; and Zwebner, Y. 2025. The urgency-necessity earthquake alert trade-off: considering the public response factor. *Natural Hazards: Journal of the International Society for the Prevention and Mitigation of Natural Hazards*, 121(8): 8951–8973.

Pall, R.; Gauthier, Y.; Auer, S.; and Mowaswes, W. 2023. Predicting drug shortages using pharmacy data and machine learning. *Health care management science*, 26(3): 395–411.

Patel, A. N.; Kesselheim, A. S.; and Rome, B. N. 2021. Frequency Of Generic Drug Price Spikes And Impact On Medicaid Spending. *Health Affairs*, 40(5): 779–785.

Phuong, J. M.; Penm, J.; Chaar, B.; Oldfield, L. D.; and Moles, R. 2019. The impacts of medication shortages on patient outcomes: a scoping review. *PloS one*, 14(5): e0215837.

Postma, D. J.; Heijkoop, M. L.; De Smet, P. A.; Notenboom, K.; Leufkens, H. G.; and Mantel-Teeuwisse, A. K. 2024. Identifying Medicine Shortages With the Twitter Social Network: Retrospective Observational Study. *Journal of Medical Internet Research*, 26: e51317.

Pourghahreman, N.; Ghatari, A. R.; and Moosivand, A. 2018. Agent based simulation of sale and manufacturing agents acting across a pharmaceutical supply chain. *Iranian journal of pharmaceutical research: IJPR*, 17(4): 1581.

Qian, C.; Du, H.; Wang, H.; Chen, X.; Zhang, Y.; Sil, A.; Zhai, C.; McKeown, K.; and Ji, H. 2025a. ModelingAgent: Bridging LLMs and Mathematical Modeling for Real-World Challenges. *arXiv preprint arXiv:2505.15068*.

Qian, C.; Liu, Z.; Prabhakar, A.; Liu, Z.; Zhang, J.; Chen, H.; Ji, H.; Yao, W.; Heinecke, S.; Savarese, S.; et al. 2025b. UserBench: An Interactive Gym Environment for User-Centric Agents. *arXiv preprint arXiv:2507.22034*.

Sarnola, K.; and Linnolahti, J. 2019. A regulatory perspective on the availability of medicines and medicine shortages in outpatient care: case Finland. *International Journal of Clinical Pharmacy*, 41(4): 825–830.

Song, Z.; Meng, Z.; and King, I. 2024. A Diffusion-Based Pre-training Framework for Crystal Property Prediction. In *AAAI*, 8993–9001. AAAI Press.

Sterman, J. D.; and Dogan, G. 2015. "I'm not hoarding, I'm just stocking up before the hoarders get here.": Behavioral causes of phantom ordering in supply chains. *Journal of Operations Management*, 39: 6–22.

Swinney, R.; Wu, X.; and Zhang, C. 2024. Mitigating shortages of generic drugs: The role of reliability certification. *Available at SSRN 4890135*.

Tadrous, M.; Kim, K. C.; Hernandez, I.; Rothenberger, S. D.; Devine, J. W.; Hershey, T. B.; Maillart, L. M.; Gellad, W. F.; and Suda, K. J. 2024. Differences in drug shortages in the US and Canada. *Jama*, 332(22): 1912–1922.

Tan, J.; Xu, R.; Chen, K.; Braubach, L.; Jander, K.; and Pokahr, A. 2020. Multi-agent System for Simulation of Response to Supply Chain Disruptions. In Kotenko, I.; Badica, C.; Desnitsky, V.; El Baz, D.; and Ivanovic, M., eds., *Intelligent Distributed Computing XIII*, 128–139. Cham: Springer International Publishing. ISBN 978-3-030-32258-8.

Tucker, E. L.; Daskin, M. S.; Sweet, B. V.; and Hopp, W. J. 2020. Incentivizing resilient supply chain design to prevent drug shortages: policy analysis using two- and multi-stage stochastic programs. *IISE Transactions*, 52(4): 394–412.

U.S. Food and Drug Administration. 2013. Strategic plan for preventing and mitigating drugshortages. https://www.fda.gov/media/86907/download. Last Accessed: 2025-07-27.

U.S. Food and Drug Administration. 2016. A review of FDA's approach to medical product shortages. https://www.ipqpubs.com/wp-content/uploads/2012/02/FDA_drug_shortages_report.pdf. Last Accessed: 2025-07-27.

Videau, M.; Lebel, D.; and Bussières, J.-F. 2019. Drug shortages in Canada: Data for 2016–2017 and perspectives on the problem. *Annales Pharmaceutiques Françaises*, 77(3): 205–211.

Vizient. 2019. Drug shortages and labor costs. https://vizientinc-delivery.sitecorecontenthub.cloud/api/public/content/2904ca0ddec449b793fd3b39ff4bdf8b?v=9a739094. Last Accessed: 2025-07-27.

Wang, L.; Ma, C.; Feng, X.; Zhang, Z.; Yang, H.; Zhang, J.; Chen, Z.; Tang, J.; Chen, X.; Lin, Y.; et al. 2024. A survey on large language model based autonomous agents. *Frontiers of Computer Science*, 18(6): 186345.

Wei, T.; Jin, B.; Li, R.; Zeng, H.; Wang, Z.; Sun, J.; Yin, Q.; Lu, H.; Wang, S.; He, J.; et al. 2024. TOWARDS UNIFIED MULTI-MODAL PERSONALIZATION: LARGE VISION-LANGUAGE MODELS FOR GENERATIVE RECOMMENDATION AND BEYOND. In *12th International Conference on Learning Representations, ICLR 2024*.

Woodcock, J.; and Wosinska, M. 2013. Economic and Technological Drivers of Generic Sterile Injectable Drug Shortages. *Clinical Pharmacology & Therapeutics*, 93(2): 170–176.

Xu, W.; Shi, Y.; Liang, Z.; Ning, X.; Mei, K.; Wang, K.; Zhu, X.; Xu, M.; and Zhang, Y. 2025. Instructagent: Building user controllable recommender via llm agent. *arXiv e-prints*, arXiv–2502.

Yurukoglu, A.; Liebman, E.; and Ridley, D. B. 2017. The Role of Government Reimbursement in Drug Shortages. *American Economic Journal: Economic Policy*, 9(2): 348–82.

Zhang, Y.; Li, J.; and Li, W. 2023. VIBE: Topic-Driven Temporal Adaptation for Twitter Classification. In Bouamor, H.;

Pino, J.; and Bali, K., eds., *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, 3340–3354. Singapore: Association for Computational Linguistics.

Zhang, Y.; Li, S.; Liu, J.; Yu, P.; Fung, Y. R.; Li, J.; Li, M.; and Ji, H. 2024. Knowledge overshadowing causes amalgamated hallucination in large language models. *arXiv preprint arXiv:2407.08039*.

Zhang, Y.; Li, S.; Qian, C.; Liu, J.; Yu, P.; Han, C.; Fung, Y. R.; McKeown, K.; Zhai, C.; Li, M.; and Ji, H. 2025a. The Law of Knowledge Overshadowing: Towards Understanding, Predicting and Preventing LLM Hallucination. In Che, W.; Nabende, J.; Shutova, E.; and Pilehvar, M. T., eds., *Findings of the Association for Computational Linguistics: ACL 2025*, 23340–23358. Vienna, Austria: Association for Computational Linguistics. ISBN 979-8-89176-256-5.

Zhang, Y.; Wang, Q.; Qian, C.; Liu, J.; Sun, C.; Zhang, D.; Abdelzaher, T.; Zhai, C.; Nakov, P.; and Ji, H. 2025b. Atomic Reasoning for Scientific Table Claim Verification. *arXiv preprint arXiv:2506.06972*.

Zhao, X.; Jia, J.; and Zhao, H. 2025. Reimbursement policy and drug shortages. *Management Science*.