Rebalancing and Clearance Pricing of Near-Expiry Inventory in Online Grocery Retail

Ziyuan Zhou[1], Long He[1], Zhengling Qi[1], Guangrui Ma[2], Yuanbo Peng[2], Yushu Yao[2], Zhongyi Zha[2], Mingyong Zhao[2], and Zhenyu Zhuo[2]

[1]George Washington University
[2]Meituan Group

Abstract

Perishable inventory management remains a critical challenge for online grocery retailers, where mismatches between demand and shelf life can lead to substantial financial losses and food waste. This article describes the development and deployment of an integrated, data-driven system for managing near-expiry ("red-line") inventory across a nationwide network of micro-fulfillment centers (MFCs) operated by Meituan's Little Elephant Supermarket. The system jointly optimizes nighttime inventory rebalancing and daytime clearance pricing using a closed-loop learning-and-optimization framework. It combines causal machine learning to estimate price elasticities, a linearized mixed-integer model for inventory rebalancing, and a Markov decision process for dynamic pricing. Since its rollout, the solution has reduced spoilage rates by 20%, increased revenue from red-line inventory by 14%, and improved annual profit by approximately 50 million CNY ($\sim$ 7 million USD) as well as saved approximately 18 million CNY ($\sim$ 2.5 million USD) in food waste, as measured across the entire operational network. Our work illustrates how aligning OR insights with adaptive ML tools can deliver scalable, interpretable, and uncertainty-aware decision systems, generating both economic and sustainability gains in complex retail operations.

1 Introduction

Online grocery retail has grown rapidly, driven by demand for convenience and speed, with platforms building dense networks of micro-fulfillment centers (MFCs) for near-instant delivery [Belavina et al., 2017, Ahumada and Villalobos, 2009]. Meituan's Little Elephant Supermarket, one of China's largest online grocery brands, exemplifies this shift. It offers 30-minute deliveries from near thousands of MFCs, processing millions of orders daily under stringent freshness and timeliness constraints. Yet this fast-paced fulfillment model poses acute challenges in managing perishables. As items approach expiration, supply–demand mismatches become severe, creating what we define as "red-line" inventory—products with shelf life below a critical threshold, which in our practice narrows to less than 24 hours. These items mark the final opportunity for revenue recovery before spoilage and demand immediate reallocation or pricing action [Akkas et al., 2019, Nahmias, 1982]. Failure to intervene not only erodes profitability but also contributes to the broader food waste crisis [Akkaş and Gaur, 2022, Birkmaier et al., 2024].

Managing red-line inventory in a decentralized grocery system presents acute challenges: the clearance window is extremely short (often less than a day), inventory is unevenly distributed across MFCs due to stochastic demand and fixed service zones, and the scale of operations makes real-time coordination difficult. Traditional approaches rely on manual or rule-based heuristics, such as fixed

39th Conference on Neural Information Processing Systems (NeurIPS 2025) Workshop: MLxOR: Mathematical Foundations and Operational Integration of Machine Learning for Uncertainty-Aware Decision-Making.

markdown schedules (e.g., 20% off after 6 p.m.), while near-expiry transfers are often excluded for policy or execution reasons [Belavina et al., 2017]. This rigidity leaves no room to react to real-time demand fluctuations or spatial imbalances, leading to unnecessary spoilage and frequent use of steep markdowns that erode profitability.

To address the intertwined challenges of minimizing spoilage and recovering value from near-expiry inventory, we developed an integrated, data-driven framework that jointly optimizes nighttime inventory rebalancing and daytime dynamic pricing, the methodological details of which we present in the next section. Our solution replaces traditional rule-based operations with a closed-loop learning-and-optimization system that coordinates decisions across both space and time. By incorporating structural insights from optimization into the learning process, the framework produces elasticity estimates that remain interpretable and operationally meaningful. At the same time, it explicitly manages uncertainty at the data–decision interface, ensuring that pricing and allocation strategies are both adaptive and robust. Embedded in Meituan's live fulfillment infrastructure, it demonstrates how analytics and machine learning can simultaneously improve margins and reduce food waste in one of the most complex areas of modern retail.

2 Integrated learning-and-optimization framework

High-level overview During the night, our *inventory rebalancing* module determines optimal transshipment flows across micro-fulfillment centers (MFCs). These decisions are informed by current inventory levels, capacity constraints, and projected demand, and are solved using a linearized mixed-integer program that approximates revenue outcomes under elasticity-adjusted pricing forecasts. It builds on lateral pooling and perishables logistics [Tagaras, 1999, Paterson et al., 2011], adapted for real-time use with short shelf life. Once inventory has been repositioned, our daytime *dynamic pricing* module activates. This component uses real-time sales signals and causal demand models to generate bi-hourly clearance prices tailored to local conditions. Demand is modeled using a semi-parametric formulation that blends log-log price elasticity with modern machine learning methods [Besbes and Zeevi, 2009, den Boer and Keskin, 2022, Javanmard, 2017]. The resulting price recommendations are optimized using a customized dynamic programming algorithm that maximizes expected revenue while accounting for inventory constraints, temporal decay, and probabilistic demand realization [Chen and Simchi-Levi, 2004, Hu et al., 2016]. We next detail each stage, with full formulations in the Appendix.

Inventory rebalancing stage: nighttime transfer planning Overnight, products transitioning to red-line status are reallocated across micro-fulfillment centers (MFCs) in preparation for the following day's sales. Each product is tracked as a Stock Keeping Unit (SKU), and transfer decisions are naturally defined at the SKU–MFC level. The decision variables are the nightly transshipment quantities between MFCs in a cluster denoted by set $\mathcal{C}$. Let x_{ij} denote the quantity of near-expiry stock transferred from MFC i to MFC j for $i, j \in \mathcal{C}$. The objective is to maximize the expected clearance revenue net of transfer costs, subject to standard operational constraints such as inventory balance, non-negativity, and cluster feasibility.

A central challenge lies in the clearance revenue term. For each MFC i, clearance revenue is defined as $\pi(q_i) = q_i \times p_i^*(q_i)$, where q_i is the post-transfer inventory and $p_i^*(q_i)$ is the optimized clearance price. To compute $p_i^*(q_i)$, we leverage the semi-parametric demand–pricing model learned from historical data (introduced in the next subsection). Based on the log–log elasticity relationship with baseline sales $\bar{q}_i$, baseline price $\bar{p}_i$, and estimated elasticity α_i, the revenue function becomes nonlinear in q_i, creating computational challenges even for a single SKU.

To make the problem tractable, we approximate the nonlinear revenue term around the baseline inventory using a first-order Taylor expansion, yielding a linear function of post-transfer inventory. In practice, we also observe that clearance prices rarely exceed the original retail price. This motivates the introduction of a breakpoint quantity corresponding to the retail price. Inventories below the breakpoint are assumed sellable at the retail price, while larger inventories follow the elasticity-based approximation. This treatment converts the nonlinear optimization into a mixed-integer linear program (MILP), enabling fast and reliable solutions under operational deadlines. The resulting transfer plans are implemented overnight, and the updated inventory levels are then passed to the daytime dynamic pricing stage.

Clearance pricing stage: daytime dynamic markdown Following nighttime rebalancing, the challenge shifts to dynamically adjusting clearance prices for near-expiry stock during the day. The goal is to maximize expected revenue while minimizing spoilage under uncertainty and a short clearance horizon. To capture these dynamics, we formulate the process as a Markov Decision Process (MDP).

In our formulation, the system is observed at discrete decision epochs (e.g., every hour). The state at period t for MFC i is $s_t^i = (p_{t-1}^i, q_t^i)$, where p_{t-1}^i is the effective price from the previous period and q_t^i is the available inventory. The action p_t^i represents the clearance price to be set, chosen within feasible bounds and restricted to be non-increasing over time. Demand in each period is modeled as a Poisson random variable whose mean depends on p_t^i via the demand–pricing model, with realized sales equal to the minimum of available inventory and stochastic demand.

The immediate reward in period t combines sales revenue with a penalty for unsold red-line inventory, ensuring that both profit generation and spoilage risk are reflected. Future rewards are captured by the value function $V_t(s_t^i)$, defined recursively through the Bellman equation. The full mathematical specification of the reward function and dynamic program is provided in the Appendix. The MDP concludes with natural boundary conditions: when the clearance horizon ends, no additional rewards are accrued ($V_T(\cdot) = 0$); initial inventory levels are set by the preceding rebalancing stage; and the starting price equals the product's original retail price.

A central component of this MDP is the demand–pricing model, which determines how sales respond to price interventions. Specifically, we use a semi-parametric double log model to capture the relationship between price and demand. The core idea is to model the relative change in sales volume as a function of the relative change in price, expressed as $\ln \frac{q}{\bar{q}} = \alpha \ln \frac{p}{\bar{p}}$, where q is the observed sales at price p, $\bar{q}$ is the baseline sales along with the baseline price $\bar{p}$, and α measures the price elasticity of demand.

To capture heterogeneity in α, we apply continuous causal forests, which extend tree-based methods to continuous treatments. In our work, we incorporate a broad set of features (X)—including product characteristics, time, weather, recent prices and sales, and promotional campaigns—to explain variations in demand. The price change for each observation, expressed as $T = \log(\frac{p}{\bar{p}})$, serves as the treatment variable, while demand change, $Y = \log(\frac{q}{\bar{q}})$, is the outcome of interest. Using the continuous causal forest method, we estimate the local price elasticity $\alpha(x)$ for each context x. Here, $\alpha(x)$ describes how a small price change is expected to affect demand, given the specific conditions represented by x. To ensure our estimates reflect recent consumer behavior, we train the model on data in rolling windows of the latest 21 days. The resulting elasticity $\alpha(x)$ provides a nuanced, data-driven measure of demand sensitivity for each product and scenario.

Daily sales forecasts are generated by combining baseline values—defined as the average sales and prices over the past 14 days ($\bar{q}$, $\bar{p}$)—with elasticity estimates from the demand model. To align predictions with reality, these forecasts are calibrated against observed sales using a group-based scaling strategy at the city–category level, with safeguards to avoid over-adjustment. The calibrated forecasts are then disaggregated into hourly demand using historical intraday proportions, assuming stable within-day patterns while preserving sensitivity to price. These elasticity-informed forecasts serve as direct inputs to the MDP-based clearance pricing model. Using dynamic programming, the model determines optimal price paths over the clearance horizon: starting from the final period and working backward, the value function recursively selects the price that maximizes the sum of immediate revenue and expected future gains.

3 Implementation and performance evaluation

Evaluation methods and metrics We evaluate our approach using business-aligned metrics that capture both clearance efficiency and value recovery. The *Clearance Discount Rate (CDR)* measures the average proportion of price reduction applied to near-expiry inventory relative to its standard price. The *Expiry Spoilage Rate (ESR)* is the share of red-line stock that remains unsold and is written off. The *Red-Line Inventory Revenue Conversion Rate (RIRCR)* quantifies the net revenue realized from red-line inventory relative to its cost. To assess impact, we adopt a controlled design in City A. MFCs were stratified by daily order volume, matched on baseline spoilage and discounting, and randomly assigned to experimental and control groups. Both baseline and test periods lasted one month. Impact is measured using a Difference-in-Differences (DiD) approach, which isolates the net effect by comparing changes between groups while controlling for temporal trends and external shocks.

Performance improvements In City A, a phased rollout across approximately 70 MFCs showed substantial performance gains from the clearance pricing (CP) module alone. Relative to matched controls, the experimental group achieved a 9.46% reduction in the CDR, a 22.85% reduction in the ESR, and a 38.89% increase in the RIRCR. These double-digit improvements demonstrate the module's effectiveness in simultaneously reducing waste and enhancing value recovery. Category-level results in Table 1 show consistent improvements across major product groups, with spoilage reductions greatest in highly perishable items and revenue gains strongest in high-turnover categories.

Table 1: Category-level improvements in City A (relative to matched controls)

Category	ESR reduction (%)	RIRCR increase (%)
Fruits	18.80	58.37
Vegetables	10.77	14.09
Seafood	57.00	1.99
Meat & Poultry	11.52	3.52
Bakery	22.12	21.97

The inventory rebalancing (IR) stage was then evaluated in cities where CP had already been fully deployed, isolating its incremental impact. Using CDR as the principal measure, IR consistently lowered average discount levels by reallocating stock more effectively across MFCs. This confirmed that better spatial balancing of at-risk inventory enables clearance with less reliance on steep markdowns, further improving overall efficiency. More detailed results are provided in the Appendix.

Economic and sustainability impact Following the validation of both Clearance Pricing and Inventory Rebalancing, the framework was progressively scaled and has now been deployed nationwide across all operational regions since early 2024. Beyond operational metrics, we evaluate two dimensions of high-level impact. From a *monetary standpoint*, the overall financial benefit is estimated by scaling the observed performance improvements in key business metrics against the aggregate Gross Merchandise Volume (GMV) during the analysis period. Applying this methodology, we observe a net profit improvement of approximately 50 million CNY ($\sim$ 7 million USD). For *sustainability*, we measure food waste reduction by the decline in unsold or written-off inventory, yielding a tangible improvement of approximately 18 million CNY ($\sim$2.5 million USD) during the implementation period. These concrete outcomes highlight the dual value of our solution—delivering substantial financial benefits while advancing measurable sustainability goals.

4 Conclusion

Our study demonstrates that a data-driven, end-to-end framework—combining nighttime inventory rebalancing and daytime clearance pricing—can deliver measurable improvements in both operational performance and economic value at scale. Methodologically, we advance the state of practice by integrating machine learning–based demand estimation and causal inference for elasticity modeling with optimization techniques that render dynamic allocation and pricing problems tractable. This tight coupling of machine learning and operations research enables adaptive, data-driven decisions that respond to real-time market conditions and operational constraints. Beyond technical innovation, the framework has generated business impact by reducing spoilage, improving revenue realization, and enabling automated decision-making across a complex fulfillment network, with lessons broadly applicable to other perishable sectors such as pharmaceuticals and floriculture.

Looking ahead, there are several directions to extend this work. Within Meituan, the framework can be expanded to holiday or event-driven scenarios, as well as slow-moving or newly launched items. More broadly, this work highlights the promise of ML–OR integration: OR insights provide interpretable, structured demand estimates, while ML enriches OR with real-time adaptability. Together, they offer a principled way to address uncertainty across data, models, and decisions, with applications extending beyond grocery to other high-uncertainty domains.

References

Omar Ahumada and J Rene Villalobos. Application of planning models in the agri-food supply chain: A review. *European journal of Operational research*, 196(1):1–20, 2009.

Arzum Akkaş and Vishal Gaur. Om forum—reducing food waste: An operations management research agenda. *Manufacturing & Service Operations Management*, 24(3):1261–1275, 2022.

Arzum Akkas, Vishal Gaur, and David Simchi-Levi. Drivers of product expiration in consumer packaged goods retailing. *Management Science*, 65(5):2179–2195, 2019.

Elena Belavina, Karan Girotra, and Ashish Kabra. Online grocery retail: Revenue models and environmental impact. *Management Science*, 63(6):1781–1799, 2017.

Omar Besbes and Assaf Zeevi. Dynamic pricing without knowing the demand function: Risk bounds and near-optimal algorithms. *Operations research*, 57(6):1407–1420, 2009.

Alexandra Birkmaier, Adhurim Imeri, and Gerald Reiner. Improving supply chain planning for perishable food: Data-driven implications for waste prevention. *Journal of Business Economics*, 94(6):1–36, 2024.

Xin Chen and David Simchi-Levi. Coordinating inventory control and pricing strategies with random demand and fixed ordering cost: The finite horizon case. *Operations research*, 52(6):887–896, 2004.

Arnoud V den Boer and N Bora Keskin. Dynamic pricing with demand learning and reference effects. *Management Science*, 68(10):7112–7130, 2022.

Peng Hu, Stephen Shum, and Man Yu. Joint inventory and markdown management for perishable goods with strategic consumer behavior. *Operations Research*, 64(1):118–134, 2016.

Adel Javanmard. Perishability of data: dynamic pricing under varying-coefficient models. *Journal of Machine Learning Research*, 18(53):1–31, 2017.

Steven Nahmias. Perishable inventory theory: A review. *Operations research*, 30(4):680–708, 1982.

Colin Paterson, Gudrun Kiesmüller, Ruud Teunter, and Kevin Glazebrook. Inventory models with lateral transshipments: A review. *European Journal of Operational Research*, 210(2):125–136, 2011.

George Tagaras. Pooling in multi-location periodic inventory distribution systems. *Omega*, 27(1): 39–59, 1999.

A Modeling and algorithms

A.1 Inventory rebalancing formulation

For each SKU within a cluster $\mathcal{C}$, the optimization model is:

$$\max_{\{x_{ij},q_i\}} \quad \sum_{i\in\mathcal{C}} \pi(q_i) - \sum_{i,j\in\mathcal{C}} c_{ij}x_{ij} \tag{1}$$

$$\text{s.t.} \quad q_i \le q_i^{pre} - \sum_{j\in\mathcal{C}} x_{ij} + \sum_{j\in\mathcal{C}} x_{ji}, \qquad \forall i \in \mathcal{C} \tag{2}$$

$$\sum_{j\in\mathcal{C}} x_{ij} \le q_i^{pre}, \qquad \forall i \in \mathcal{C} \tag{3}$$

$$x_{ij} \ge 0, \qquad \forall i, j \in \mathcal{C}. \tag{4}$$

The clearance revenue term is defined as

$$\pi(q_i) = q_i \times p_i^*(q_i), \tag{5}$$

where q_i is the adjusted inventory after rebalancing and $p_i^*(q_i)$ is the optimized clearance price.

To compute $p_i^*(q_i)$, we leverage the semi-parametric demand–pricing model estimated from historical data:

$$\ln \frac{q_i}{\bar{q}_i} = \alpha_i \ln \frac{p_i}{\bar{p}_i}, \tag{6}$$

which yields

$$p_i^*(q_i) = \bar{p}_i \left(\frac{q_i}{\bar{q}_i}\right)^{1/\alpha_i}. \tag{7}$$

To simplify computation, we approximate $\pi(q_i)$ around the baseline inventory $\bar{q}_i$ using a first-order Taylor expansion:

$$\pi(q_i) \approx \bar{p}_i\bar{q}_i + \bar{p}_i\left(1 + \frac{1}{\alpha_i}\right)(q_i - \bar{q}_i). \tag{8}$$

We also impose an upper bound on prices by introducing the breakpoint

$$q_i' = \bar{q}_i \left(\frac{p_0^i}{\bar{p}_i}\right)^{\alpha_i}, \tag{9}$$

where p_0^i is the original retail price. The revenue function is then:

$$\pi(q_i) = \begin{cases} \bar{p}_i\bar{q}_i + \bar{p}_i\left(1 + \frac{1}{\alpha_i}\right)(q_i - \bar{q}_i), & \text{if } q_i > q_i', \\ p_0^i q_i, & \text{if } q_i \le q_i'. \end{cases}$$

A.2 Clearance pricing as a markov decision process

At each decision period t for MFC i, the state is

$$s_t^i = (p_{t-1}^i, q_t^i), \tag{10}$$

where p_{t-1}^i is the effective price from the previous period and q_t^i is the available inventory. The action is the clearance price p_t^i, chosen within feasible bounds and restricted to be non-increasing.

Let $\tilde{d}_t^i$ denote random demand in period t, whose mean depends on p_t^i. Realized sales are

$$q_t^i \wedge \tilde{d}_t^i = \min(q_t^i, \tilde{d}_t^i), \tag{11}$$

and inventory evolves as

$$q_{t+1}^i = q_t^i - (q_t^i \wedge \tilde{d}_t^i). \tag{12}$$

The expected immediate reward is

$$R_t(s_t^i, p_t^i) = \mathbb{E}_{\tilde{d}_t^i}\left[p_t^i(q_t^i \wedge \tilde{d}_t^i) - c^h q_{t+1}^i \mid s_t^i, p_t^i\right], \tag{13}$$

where c^h is a per-unit cost penalizing unsold red-line inventory.

The value function $V_t(s_t^i)$ represents the maximum expected cumulative reward from period t onward, satisfying the Bellman recursion:

$$V_t(s_t^i) = \max_{p_t^i \le p_{t-1}^i} \left\{R_t(s_t^i, p_t^i) + \mathbb{E}\big[V_{t+1}(s_{t+1}^i) \mid s_t^i, p_t^i\big]\right\}. \tag{14}$$

Boundary conditions:

$$V_T(\cdot) = 0, \tag{15}$$

$$q_1^i = \text{inventory from rebalancing}, \tag{16}$$

$$p_0^i = \text{original retail price}. \tag{17}$$

B Experimental design and additional results

For completeness, this appendix provides further details that complement the main results reported in Section 3. Core performance outcomes and category-level effects for City A are already presented in the main text.

Experimental setup in City A For completeness, we provide additional details on the experimental setup in City A, where the clearance pricing stage was first evaluated. MFCs were stratified into three tiers by average daily order volume. Within each tier, we selected stores with the longest operational history to ensure data stability and process maturity. Candidate MFCs were further matched on baseline Clearance Discount Rate (CDR) and Expiry Spoilage Rate (ESR), before being randomly and evenly assigned to experimental and control groups.

Both baseline and test periods lasted one month. To identify causal impact, we employed a Difference-in-Differences (DiD) framework: for each performance index, we first calculated the change from baseline to test within each group, and then measured the net effect as the difference between experimental and control groups. This design effectively controls for temporal shocks and common trends, attributing improvements specifically to the intervention.

Inventory rebalancing results Table 2 summarizes the impact of the inventory rebalancing (IR) module on the Clearance Discount Rate (CDR) across five major cities. Consistently, IR reduced average discount levels by reallocating stock more effectively across MFCs, confirming that better spatial balancing of at-risk inventory enables clearance with less reliance on steep markdowns. These results complement the aggregate findings reported in the main text and highlight the robustness of the approach across geographies.

Table 2: Impact of inventory rebalancing on CDR (DiD estimates, five largest cities).

City	CDR reduction (%)
A	8.04
B	11.83
C	10.89
D	11.36
E	6.63

C Implementation notes

All computed prices are clipped to fall within predefined upper and lower bounds to ensure stability. The algorithms are seamlessly integrated into Meituan's backend operational systems to support automated decision-making and real-time execution. Each night, inventory rebalancing plans are generated through scheduled tasks, after which the system pulls the optimization results, issues transfer orders, and coordinates the workflow from sorting and logistics to shelf restocking. During the clearance phase, the backend periodically queries the algorithm for optimal markdowns at the SKU–MFC level, enabling dynamic, data-driven pricing updates. To retain flexibility, operation managers can make manual adjustments when necessary. This hybrid approach combines algorithmic efficiency with operational expertise, supporting robust, scalable, and adaptive deployment in a complex retail environment.