mathematics

Special Issue Reprint

Theoretical and Applied Mathematics in Supply Chain Management

Edited by
Xiang Song

mdpi.com/journal/mathematics

MDPI

Theoretical and Applied Mathematics in Supply Chain Management

Theoretical and Applied Mathematics in Supply Chain Management

Guest Editor

Xiang Song

MDPI

Basel • Beijing • Wuhan • Barcelona • Belgrade • Novi Sad • Cluj • Manchester

Guest Editor
Xiang Song
School of Mathematics
and Physics
University of Portsmouth
Portsmouth
UK

Editorial Office
MDPI AG
Grosspeteranlage 5
4052 Basel, Switzerland

This is a reprint of the Special Issue, published open access by the journal *Mathematics* (ISSN 2227-7390), freely accessible at: https://www.mdpi.com/si/mathematics/2RZ4GAM2O8.

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

Lastname, A.A.; Lastname, B.B. Article Title. *Journal Name* **Year**, *Volume Number*, Page Range.

ISBN 978-3-7258-6832-2 (Hbk)
ISBN 978-3-7258-6833-9 (PDF)
https://doi.org/10.3390/books978-3-7258-6833-9

Contents

About the Editor

Xiang Song

Xiang Song is an Associate Professor in Operational Research and Operations Management at the University of Portsmouth, United Kingdom. Her research focuses on optimisation modelling, simulation, and data analytics for supply chain management, logistics, and transportation systems. She has led and co-led several research projects funded by the Royal Society, UK Research and Innovation (UKRI), and the Department for Transport (DfT), with a particular emphasis on resilient supply chain design, drone logistics, and future mobility systems. Dr. Song has published widely in operations research and quantitative logistics, contributing to the development of analytical models that enhance the efficiency, sustainability, and robustness of complex systems. Her interdisciplinary work bridges theory and practice, combining mathematical optimisation with real-world applications in both commercial and humanitarian contexts. She is also actively engaged in academic leadership, research supervision, and knowledge exchange with industry partners, promoting the use of quantitative methods to solve emerging operational challenges in the digital era.

Preface

This Reprint on *Quantitative Supply Chain Management* brings together research that applies quantitative, optimisation, and analytical methods to address complex supply chain challenges, including procurement, inventory management, facility location, sustainability, and maintenance planning. It is intended for researchers, practitioners, and postgraduate students seeking rigorous, data-driven approaches to improve efficiency, resilience, and decision-making in modern supply chains.

This Reprint is intended for researchers, academics, and practitioners in operations research, industrial engineering, and supply chain management who seek to understand and apply quantitative methods to real-world challenges. It also serves as a useful reference for postgraduate students who aspire to develop rigorous analytical skills for tackling emerging supply chain problems.

Xiang Song
Guest Editor

mathematics

MDPI

Article

A Swap-Integrated Procurement Model for Supply Chains: Coordinating with Long-Term Wholesale Contracts

Min-Yeong Ryu [1] and Pyung-Hoi Koo [2,*]

[1] Department of Industrial and Data Engineering, Pukyong National University, Busan 48513, Republic of Korea; alsdud35@pukyong.ac.kr
[2] Department of Systems Management and Safety Engineering, Pukyong National University, Busan 48513, Republic of Korea
* Correspondence: phkoo@pknu.ac.kr; Tel.: +82-51-629-6485

Abstract

In today's volatile supply chain environment, organizations require flexible and collaborative procurement strategies. Swap contracts, originally developed as financial instruments, have recently been adopted to address inventory imbalances—such as the 2021 COVID-19 vaccine swap between South Korea and Israel. Despite its increasing adoption in the real world, theoretical studies on swap-based procurement remain limited. This study proposes an integrated model that combines buyer-to-buyer swap agreements with long-term wholesale contracts under demand uncertainty. The model quantifies the expected swap quantity between parties and embeds it into the profit function to derive optimal order quantities. Numerical experiments are conducted to compare the performance of the proposed strategy with that of a baseline wholesale contract. Sensitivity analyses are performed on key parameters, including demand asymmetry and swap prices. The numerical analysis indicates that the swap-integrated procurement strategy consistently outperforms procurement based on long-term wholesale contracts. Moreover, the results reveal that under the swap-integrated strategy, the optimal order quantity must be adjusted—either increased or decreased—depending on the demand scale of the counterpart and the specified swap price, deviating from the optimal quantity under traditional long-term contracts. These findings highlight the potential of swap-integrated procurement strategies as practical coordination mechanisms across both private and public sectors, offering strategic value in contexts such as vaccine distribution, fresh produce, and other critical products.

Keywords: procurement strategy; swap contract; demand uncertainty; inventory coordination; supply chain risk management

MSC: 90B06

1. Introduction

Global supply chains have been increasingly exposed to uncertainty due to external shocks such as the COVID-19 pandemic, geopolitical conflicts, and rapid demand fluctuations. In this context, procurement has emerged as a critical challenge for maintaining supply chain continuity and resilience. Ineffective procurement not only reduces operational efficiency and customer service levels but also undermines trust in supply reliability and overall supply chain resilience [1]. To mitigate these risks, the supply chain management (SCM) literature has proposed various contract-based procurement strategies. Notable examples include buyback contracts [2], revenue-sharing contracts [3], and option

contracts [4]. However, these strategies primarily focus on vertical coordination between suppliers and retailers.

This paper introduces a swap-integrated procurement model that facilitates lateral cooperation between buyers. A swap is a financial derivative contract in which two parties agree to exchange streams of cash flow over a specified period, typically to manage interest rate, currency, or commodity price risks. Common types include interest rate swaps (exchanging fixed and floating payments) and currency swaps (exchanging payments in different currencies). Swaps are widely used in financial engineering as a tool for hedging against market volatility and aligning financial exposures with risk preferences. For example, during the 2008 Global Financial Crisis, South Korea entered into a USD 30 billion currency swap agreement with Japan to stabilize its financial markets by supplying short-term liquidity in foreign currencies. Under this agreement, Japan provided Japanese yen to South Korea in exchange for Korean won, with the option to swap the yen back into won at a later date. This swap agreement played a vital role in helping the two economies weather the financial storm during the 2008 crisis [5].

Recently, swap mechanisms have been adapted for use in supply chain procurement strategies to mitigate the risks associated with demand volatility. Unlike financial swaps, which are based on predefined cash flow exchanges over time, supply chain swaps refer to the post-demand reallocation of physical goods between buyers. In this context, a swap refers to an arrangement where buyers exchange surplus and shortage quantities of goods after demand uncertainty is realized. For instance, in 2021, Australia and Singapore executed a vaccine swap agreement: Australia received 500,000 near-expiry doses from Singapore to accelerate its domestic vaccination program, with the promise to return an equal quantity at a later date [6]. Similarly, India implemented an LNG swap tender in 2025 to exchange cargoes from the U.S. for future deliveries to meet seasonal demand fluctuations, thereby enhancing supply flexibility and reliability [7].

These examples highlight the practical value of swap contracts as flexible and collaborative mechanisms for resolving short-term supply–demand mismatches. However, despite their increasing adoption in supply chain practice, theoretical research on swap contracts remains at an early stage. In addition, most existing studies on swap contracts have primarily focused on logistics optimization from the supplier's perspective—such as minimizing transportation distances, reducing logistics costs, and shortening lead times—by facilitating resource exchanges between inventory hubs prior to demand realization. While these approaches enhance operational efficiency, the application of swap strategies in broader supply chain contexts, particularly from the buyer's perspective, has received limited attention [8]. In particular, there is a significant gap in theoretical and quantitative research on how buyers should optimally determine order quantities and evaluate procurement performance when swap-based sourcing options are available. This highlights the need for rigorous analytical investigation into the role and effectiveness of swaps as a collaborative procurement strategy under post-demand realization uncertainty.

This study proposes a swap-integrated procurement model that combines swap contracts with wholesale contracts and analyzes its performance in a single-period environment with uncertain demand. The analysis focuses on a setting in which two buyers, each having secured a long-term commitment contract—either with the same supplier or with different suppliers—determine their order quantities, and subsequently adjust surplus and shortage inventories through a swap contract after demand is realized. To the best of our knowledge, this study is the first to explore analytical models for the swap-integrated procurement strategy in supply chains, offering a novel contribution to the literature in three key aspects:

(1) Formulating the expected swap inflows and outflows under the swap-integrated strategy and embedding them into the expected profit function to derive optimal order quantities;
(2) Comparing the performance of the proposed strategy with a baseline wholesale contract to assess the structural benefits of swap integration;
(3) Conducting sensitivity analyses on key parameters—such as demand asymmetry and swap prices—to identify the conditions under which the proposed strategy yields its greatest performance benefits.

This paper is organized as follows: Section 2 reviews the related literature on swap-based procurement contracts and identifies research gaps. Section 3 presents the supply chain structure and mathematical model for the swap-integrated procurement strategy. Section 4 presents the numerical results, including sensitivity analyses. Finally, Section 5 concludes with implications, limitations, and future research directions.

2. Literature Review

Demand uncertainty serves as a major procurement risk across the entire supply chain, and the information asymmetry between the procurement stage and the demand realization stage exposes buyers to two opposing risks: excess inventory and stockouts. Various contract-based procurement strategies have been proposed in the SCM literature to mitigate risks associated with highly volatile demand. Common examples include buyback contracts [2], revenue-sharing contracts [3], and option contracts [4], most of which are characterized as forms of vertical cooperation between suppliers and retailers.

Our literature study focuses on swap contracts, which can be characterized as lateral cooperation between buyers. Swap practice refers to the mutual exchange of goods or inventory between firms, typically under pre-agreed terms, to mitigate supply–demand mismatches. Unlike traditional procurement, it emphasizes flexibility and collaboration, allowing firms to reduce holding costs or address shortages without relying solely on upstream suppliers. Swap practices were initially adopted in the petroleum and energy industries, where they offered significant advantages due to the commodity-based nature of products [9]. In these sectors, customers typically prioritize product specifications and on-time delivery over the origin of the goods. By collaborating, they can reduce transportation and inventory costs while enhancing customer service. These alliances are particularly effective in improving overall supply chain efficiency. The transportation cost savings achieved through such cooperation are shared among the participating firms. Bidyarthi and Deshmukh [10] highlight the role of swap mechanisms as a strategic inventory handling tool in downstream petroleum supply chains. They illustrate how swap practices can be applied as an alternative to conventional inventory management in addressing regional supply–demand mismatches. Al-Husain and Khorramshahgol [11] conduct a comprehensive analysis of the determinants of swap problems in the petroleum industry. The study explores contract terms, inventory capacity, and operational flexibility as key drivers of effective swap execution, emphasizing the importance of operational structure in swap-based systems. Farahani and Rahmani [12] analyze the impact of swap strategies on the financial performance of supply chains in large-scale distribution networks, such as those found in the petrochemical industry. They develop a mixed integer linear programming model to capture the complexities of production and transportation decisions within an oil supply chain and design scenarios in which long-distance transportation is replaced with swaps between nearby production sites. Their experimental results show that implementing swap strategies can simultaneously reduce transportation costs and lead time risks by shortening delivery distances, thereby significantly improving the supply chain's overall net present value.

Dizbay and Ozturkoglu [13] explore a product swapping mechanism among suppliers operating in a balanced supply network. The study investigates how transfer sales and swap agreements can help suppliers reallocate inventory to meet customer demand more effectively. By modeling supplier interactions, the research shows that swap-based coordination improves service levels and reduces total costs. Wang et al. [14] investigate the use of swap contracts in bunker fuel procurement for liner shipping companies. They develop a short-term procurement model that integrates fuel price uncertainty and the strategic use of swaps to hedge against price volatility. The study demonstrates that incorporating swap contracts can reduce procurement costs and financial risk under fluctuating market conditions. Park [15] explores the strategic inventory swapping between competing firms facing uncertain demand. The study models a two-stage game in which firms first determine inventory levels and then decide whether to swap inventory after demand is realized. The analysis shows that swap agreements can improve supply chain efficiency and profitability for both parties, even in competitive settings.

Kemper et al. [16] investigate the pricing of swaps and options in electricity markets by focusing on the market price of risk across different delivery periods. The study develops a risk-adjusted valuation model that captures the stochastic behavior of electricity prices and accounts for delivery-specific risk premia. Their findings show that incorporating delivery period-specific risks significantly improves the pricing accuracy of both swaps and options in volatile energy markets. Zhang and Thomspon [17] examine how product swaps can be optimized within an urban retail network under demand uncertainty. They propose a stochastic optimization model to determine when and where product swaps should occur between stores to minimize overall logistics costs and improve service levels. The study demonstrates that incorporating swap decisions into network planning can significantly enhance operational efficiency, especially under highly volatile demand. The existing literature has primarily focused on swap mechanisms from a logistics-oriented perspective, such as optimizing transportation or reallocating resources between production sites, while mathematical analyses of contract design and profit structures based on inventory adjustment between buyers are scarce. Accordingly, there is a clear need for a structural analysis of how swap-based strategies can mitigate procurement risks, improve performance, and inform the design of optimal ordering policies.

Swap-based procurement shares conceptual similarities with lateral transshipment, a widely used inventory management strategy designed to address stock imbalances across decentralized locations. Lateral transshipment aims to alleviate inventory discrepancies at local stores by enabling the transfer of goods between retail outlets, particularly in urban areas, to fulfill customer demand in regions facing stockouts. This approach seeks to reduce overall inventory costs and maintain desired service levels by reallocating surplus inventory from locations with excess stock to those experiencing shortages, typically within the same organization or integrated supply network. Diks and de Kok [18] find that such transshipments can reduce losses associated with excess inventory and stockouts, while Ekren and Arslan [8] emphasize that pre-planned transshipment policies are particularly effective in environments with high demand volatility. Paterson et al. [19] and Kumari et al. [20] provide a comprehensive review of lateral transshipments in inventory systems.

Existing studies on lateral transshipment generally focus on the internal transfer of resources or goods within the same organization or integrated system; they do not presuppose external cooperation or the exchange of resources between different stakeholders based on contractual agreements. In contrast, the swap-based procurement strategy proposed in this study is structured to enable the exchange of surplus and shortage inventories between different buyers based on contractual agreements after demand realization. This suggests that the swap strategy, as a post-realization collaboration mechanism, possesses structural

characteristics that distinguish it from lateral transshipment and offers the potential to enhance supply chain flexibility through autonomous agreements among buyers.

This study addresses gaps in the existing literature by proposing a swap-integrated procurement strategy aimed at mitigating supply–demand mismatches that emerge after demand realization. The proposed strategy attempts to achieve both procurement flexibility and risk diversification—outcomes that are difficult to secure through a single contract structure—by facilitating the exchange of surplus and shortage inventories among buyers through contractual agreements. This approach not only enhances the practical efficiency of supply chain operations but also serves as a new collaborative procurement model for responding to demand uncertainty. Table 1 summarizes the key distinctions between this study and prior research.

Table 1. Comparison between this study and existing studies.

	Supply Swap in Prior Studies	**Lateral Transshipment**	**Swap-Integrated Strategy**
Cooperation Entities	Competing suppliers	Logistics facilities within a firm facilities within a firm	Independent buyers
Decisions	Swap quantity and routing	Initial order quantity by location and point-to-point transshipment quantity	Long-term order quantity
Objective	Reduce transportation costs, improve delivery and logistics	Balance inventory levels, minimize transportation costs	Mitigate procurement risk and maximize profit
References	[9–17]	[8,18–20]	This study

3. Model Description

This study investigates a collaborative strategy between two buyers operating in a supply chain under demand uncertainty. Initially, both buyers independently determine their order quantities under a common procurement framework based on long-term wholesale contracts. While the demand for each buyer follows the same probability distribution, it differs in scale and is modeled as an independent random variable. The supplier is assumed to have unlimited production capacity, thereby imposing no constraints on order quantities, and procurement occurs only once within a single-period setting. Any excess inventory resulting from over-ordering is discarded without salvage value, whereas unmet demand incurs a penalty cost per unit. The retail price, wholesale price, and penalty cost are assumed to be exogenously given.

In this section, we first outline the procurement model based on long-term wholesale contracts, which serves as a benchmark. We then introduce the system configuration and mathematical formulation of the swap-integrated procurement strategy, which aims to enhance supply chain performance through buyer-to-buyer cooperation. Table 2 summarizes the key variables and notations used throughout this paper. The units of demand, order quantity, price, and cost are context-dependent and may, for instance, represent either individual units or thousands (for quantities), or dollars or thousands of dollars (for monetary values).

Table 2. Notations and descriptions.

Notations	**Descriptions**
w	per-unit wholesale price
p	per-unit retail price at the buyer
r	per-unit swap price

Table 2. *Cont.*

Notations	Descriptions
g	per-unit penalty cost per unit of unmet demand
c	demand scale factor of buyer 2
X	market demand of buyer 1, with pdf $f(x)$, and cdf $F(x)$ and average μ_X
Y	market demand of buyer 2, with pdf $f_y(y)$ and cdf $F_y(y)$
Q	order quantity in a long-term wholesale contract (decision variable)
q_1	swap-inflow quantity from the collaborating buyer
q_2	swap-outflow quantity to the collaborating buyer
π	buyer's profit under a wholesale contract
$S(Q)$	expected sales volume, given Q under a wholesale-only contract
$L(Q)$	expected shortage, given Q under a wholesale-only contract

3.1. Baseline Model: Long-Term Wholesale (WH) Procurement Model

This section describes the procurement structure through a long-term wholesale contract that serves as a benchmark. In this structure, the buyer purchases the products at a predetermined wholesale price and sells them in the market. The model is similar to the newsvendor framework. When the procurement quantity is Q and the demand is X, the sales quantity equals $min(X, Q)$, while the unmet demand is represented by $(X - Q)^+$ where $(X - Q)^+ = max(X - Q, 0)$. The buyer's expected profit function, $\pi^{\mathrm{WH}}(Q)$, is obtained as follows:

$$E\left[\pi^{\mathrm{WH}}(Q)\right] = pE[\min(X, Q)] - wQ - gE\left[(X - Q)^+\right] = (p + g - w)Q - (p + g)\int_0^Q F(x)dx - g\mu_X \quad (1)$$

The optimal order quantity that maximizes expected profit is obtained from the first-order condition, as follows:

$$Q^* = F^{-1}\left(\frac{p + g - w}{p + g}\right) \quad (2)$$

While the long-term wholesale contract is simple and easy to operate from the buyer's perspective, it may result in frequent losses due to excess inventory or stockouts in environments where demand is highly uncertain.

3.2. Swap-Integrated (SW) Procurement Model

This section presents an integrated procurement strategy that combines swap and wholesale contracts to address potential inventory imbalances inherent in long-term wholesale contract structures. The model is presented primarily from the perspective of Buyer 1, with a focus on the procurement process and inventory exchange mechanism. For clarity, Buyer 1 is referred to as the focal buyer, while Buyer 2 is referred to as the collaborating buyer (or counterpart) throughout the paper. A schematic illustration of the proposed strategy is shown in Figure 1.

The primary objective of this strategy is to mitigate inventory risk by enabling the mutual adjustment of excess inventory between the two buyers once demand is realized—the risk that each buyer would otherwise bear. The two buyers procure identical goods from arbitrary—and potentially different—suppliers. The term "identical goods" is adopted as a simplifying assumption for theoretical purposes; in practice, this may encompass functionally equivalent substitutes. For example, products such as the Pfizer and AstraZeneca vaccines are considered equivalent goods, as they serve the same function despite differences in formulation.

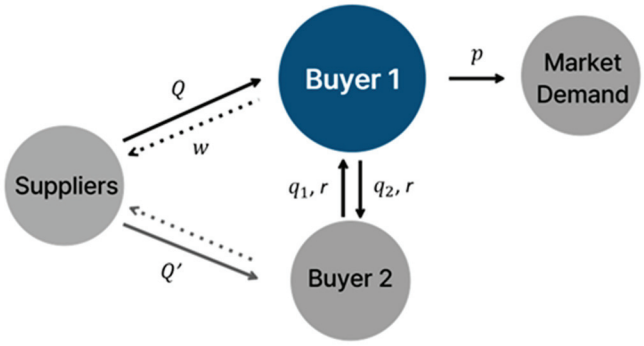


Figure 1. Schematic of the swap-integrated procurement strategy. Solid lines represent the physical flows of products, while dotted lines indicate monetary transfers.

Each buyer enters into a long-term wholesale-based contract with their respective supplier prior to demand realization, committing to quantities Q and Q', respectively, and subsequently sells the products on the market. Due to demand uncertainty, mismatches between ordered and realized demand may result in either excess inventory or stockouts. For instance, if Buyer 1 holds excess inventory while Buyer 2 experiences a shortage, the surplus is transferred from Buyer 1 to Buyer 2 at a predetermined swap price. The swap transaction is triggered only after demand is realized and proceeds according to the conditions agreed upon in advance. This mechanism serves as a structural risk-sharing tool, reducing each firm's exposure to individual inventory risks.

To facilitate the analysis of this swap-integrated procurement strategy, the following assumptions are introduced:

(1) The two buyers are fully cooperative under the swap contract. If one buyer holds excess inventory and the other faces a shortage after demand is realized, the surplus is mandatorily transferred at the pre-agreed swap price, r.
(2) The demand of Buyer 1 is represented by the random variable X, and Buyer 2's demand is represented by Y, where Y follows the same distributional form as X but is scaled by a factor of c. For instance, if Buyer 1 represents the Korean market and Buyer 2 represents the Japanese market, $c = 2.5$ may reflect the population size ratio. The scaling factor, c, plays a central role in the sensitivity analysis conducted later in this study.
(3) The swap price, r, is given exogenously. Sensitivity analyses with respect to varying values of r will be conducted in a following section.

3.2.1. Expected Profit Model for the Swap-Integrated Procurement Strategy

The swap-integrated procurement model aims to identify the optimal order quantity that maximizes the expected profit of the buyer: $E\left[\pi^{SW}(Q)\right]$. (To follow the model development more easily, it may be helpful to refer to Figure 2 at this point.) As described previously, the demands of Buyers 1 and 2 are given as independent random variables, X and Y, respectively. Buyer 1 procures quantity Q at unit price w in advance and sells the product at price p upon demand realization. Throughout this process, excess demand $(X-Q)^+$ and excess inventory $(Q-X)^+$ arise, which are adjusted through the swap contract. Buyer 2 operates under the same structure, with an order quantity set as cQ reflecting the demand scaling factor. Accordingly, Buyer 2's excess demand and excess inventory are defined as functions of $(Y-cQ)^+$, and $(cQ-Y)^+$. The swap contract is executed after demand realization when one buyer holds surplus inventory and the other experiences a shortage. The actual quantity exchanged is determined as the minimum of

these two values. Accordingly, the swap quantity procured by Buyer 1 from Buyer 2 is expressed as follows:

$$q_1 = min\Big((X-Q)^+, (cQ-Y)^+\Big) \tag{3}$$

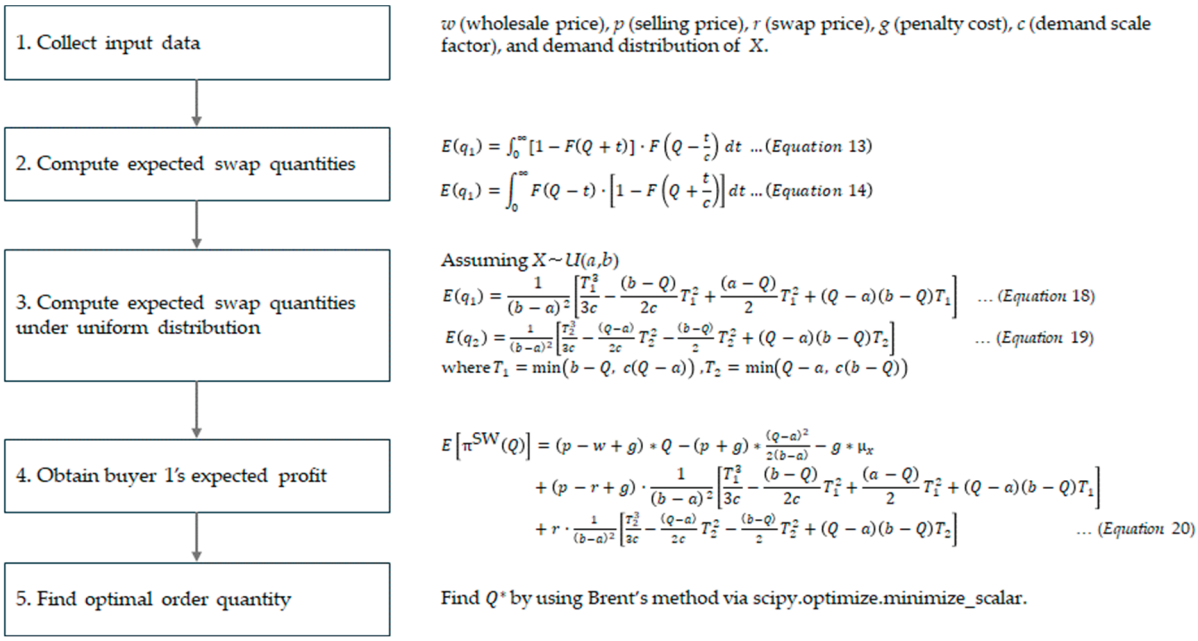


Figure 2. Procedure for deriving the optimal order quantity from a long-term supplier under uniform demand distribution conditions.

Conversely, the swap quantity supplied by Buyer 1 to Buyer 2 is expressed as follows:

$$q_2 = min\Big((Q-X)^+, (Y-cQ)^+\Big) \tag{4}$$

The expected profit function for Buyer 1 is expressed as follows:

$$E\left[\pi^{SW}(Q)\right] = p \cdot S(Q) - wQ + (p-r) \cdot E(q_1) - g \cdot (L(Q) - E(q_1)) + r \cdot E(q_2) \tag{5}$$

where $S(Q) = Q - \int_0^Q F(x)$ and $L(Q) = \mu_X - Q + \int_0^Q F(x)$, with a given order quantity, Q.

Each term corresponds to revenue from market sales, cost of procurement under the wholesale contract, additional revenue from sales through the swap, penalty costs from unmet demand, and revenue from supplying surplus inventory via the swap. The expected profit is restated as follows:

$$E\left[\pi^{SW}(Q)\right] = (p-w+g) \cdot Q - (p+g) \cdot \int_0^Q F(x)\,dx - g \cdot \mu_X + (p-r+g) \cdot E(q_1) + r \cdot E(q_2) \tag{6}$$

Buyer 1 needs to determine the optimal order quantity that maximizes the expected profit above. To compute the expected profit, it is necessary to determine $E(q_1)$ and $E(q_2)$, which is not trivial. The following section presents the procedure for calculating $E(q_1)$ and $E(q_2)$:

3.2.2. Obtaining the Expected Swap Quantity

To obtain $E(q_1)$ and $E(q_2)$, we apply the tail expectation formula, which states that $E(z) = \int_0^\infty P(z > t)\,dt$ for a non-negative random variable, z. The expected swap quantity $E(q_1)$ is then derived as follows:

$$E(q_1) = E\Big(min\Big((X-Q)^+, (cQ-Y)^+\Big)\Big) = \int_0^\infty P\Big[min\Big((X-Q)^+, (cQ-Y)^+\Big) > t\Big]\,dt \tag{7}$$

Expression (7) can be restated as follows:

$$E(q_1) = \int_0^\infty P\left[(X-Q)^+ > t, (cQ-Y)^+ > t\right]dt \quad = \int_0^\infty P(X > Q+t, Y < cQ-t)dt \tag{8}$$

Since we assume that X and Y are independent, we obtain the following equation:

$$E(q_1) = \int_0^\infty P(\,X > Q+t) \cdot P(Y < cQ-t\,)\,dt \tag{9}$$

This can be transformed in terms of the distribution function as follows:

$$P(X > Q+t) = [1-F(Q+t)];\; P(Y < cQ-t) = F_Y(cQ-t) \tag{10}$$

Finally, we obtain the following function:

$$E(q_1) = \int_0^\infty [1-F(Q+t)] \cdot\; F_Y(cQ-t)dt \tag{11}$$

Given that Y follows the same distributional structure as X but scaled by factor c, the *pdf* and *cdf* of Y satisfy the following relationship:

$$f_y(y) = \tfrac{1}{c} f\left(\tfrac{y}{c}\right),\; F_y(y) = F\left(\tfrac{y}{c}\right) \tag{12}$$

By applying this relationship, the expected swap transfer quantity from Buyer 2 to Buyer 1 can be summarized as follows:

$$E(q_1) = \int_0^\infty [1-F(Q+t)] \cdot F\left(Q-\tfrac{t}{c}\right)dt \tag{13}$$

By applying a similar procedure, the expected swap transfer quantity from Buyer 1 to Buyer 2 can also be obtained as follows:

$$E(q_2) = \int_0^\infty F(Q-t) \cdot \left[1-F\left(Q+\tfrac{t}{c}\right)\right]dt \tag{14}$$

By embedding expressions (13) and (14) into expression (6), $E\left[\pi^{SW}(Q)\right]$ can be restarted as follows:

$$\begin{aligned} E\left[\pi^{SW}(Q)\right] &= (p-w+g)\cdot Q \; - (p+g)\cdot \int_0^Q F(x)\,dx - g\cdot\mu_X + (p-r+g) \\ &\cdot \int_0^\infty (1-F(Q+t))\cdot F_Y(cQ-t)\,dt + r\cdot \int_0^\infty F(Q-t)\cdot(1-F_Y(cQ+t))\,dt \end{aligned} \tag{15}$$

Finally, the order quantity, Q, is determined to maximize $E\left[\pi^{SW}(Q)\right]$ in expression (15). The expected profit function described above has a nonlinear structure with respect to order quantity Q and consists of integrals involving multiple probability distribution functions and joint probabilities, making it highly complex to compute derivatives and derive the optimal solution analytically.

3.2.3. Optimal Order Quantity Model Under Uniform Distribution Conditions

In this section, we develop the model by assuming that demand follows a uniform distribution, which facilitates the derivation of closed-form solutions for evaluating the buyer's profitability. In the supply chain literature, this assumption is a well-established modeling convention that offers analytical simplicity while enabling insights into more complex and realistic demand settings [21]. Importantly, adopting a uniform distribution does not compromise the generality or validity of the results, as the model remains effective in capturing the structural characteristics and comparative performance of alternative procurement strategies. Let demand X and Y follow uniform distributions, $U(a, b)$ and

$U(ca, cb)$, respectively. For the uniform distribution, the *cdf* is given by $F(x) = \frac{x-a}{b-a}$, and accordingly, $F(Q+t)$ and $F\left(Q-\frac{t}{c}\right)$ are expressed as follows:

$$1 - F(Q+t) = \frac{b-Q-t}{b-a},\ F\left(Q-\frac{t}{c}\right) = \frac{Q-\frac{t}{c}-a}{b-a} \tag{16}$$

Since the *cdf* in expression (13) are defined only within specific domains, $Q + t \leq b$ and $Q - \frac{t}{c} \geq a$, the integral is valid only over the range that simultaneously satisfies these conditions. Let T_1 be the effective integration interval, $T_1 = min(b - Q, c(Q - a))$, for $E(q_1)$. Expression (13) can be restated as follows:

$$E(q_1) = \int_0^{T_1} \left(\frac{b-t-Q}{b-a}\right) \cdot \left(\frac{Q-\frac{t}{c}-a}{b-a}\right) dt \tag{17}$$

By computing the integral above, $E(q_1)$ is expressed in polynomial form as follows:

$$E(q_1) = \frac{1}{(b-a)^2}\left[\frac{T_1^3}{3c} - \frac{(b-Q)}{2c}T_1^2 + \frac{(a-Q)}{2}T_1^2 + (Q-a)(b-Q)T_1\right] \tag{18}$$

Using the same procedure as that used for obtaining $E(q_1)$, $E(q_2)$ can be derived. In Equation (14), the integral is valid when the conditions $Q - t \geq a$ and $Q + \frac{t}{c} \leq b$ are simultaneously satisfied. Let T_2 be the effective integration interval, $T_2 = min(Q - a, c(b - Q))$, for $E(q_2)$. Then, the following expression holds:

$$E(q_2) = \frac{1}{(b-a)^2}\left[\frac{T_2^3}{3c} - \frac{(Q-a)}{2c}T_2^2 - \frac{(b-Q)}{2}T_2^2 + (Q-a)(b-Q)T_2\right] \tag{19}$$

By utilizing the obtained values of $E(q_1)$ and $E(q_2)$, the expected profit under the swap-integrated procurement strategy can be expressed as follows:

$$\begin{aligned} E\left[\pi^{SW}(Q)\right] &= (p-w+g)\cdot Q - (p+g)\cdot\frac{(Q-a)^2}{2(b-a)} - g\cdot\mu_x \\ &+(p-r+g)\cdot\frac{1}{(b-a)^2}\left[\frac{T_1^3}{3c} - \frac{(b-Q)}{2c}T_1^2 + \frac{(a-Q)}{2}T_1^2 + (Q-a)(b-Q)T_1\right] \\ &+r\cdot\frac{1}{(b-a)^2}\left[\frac{T_2^3}{3c} - \frac{(Q-a)}{2c}T_2^2 - \frac{(b-Q)}{2}T_2^2 + (Q-a)(b-Q)T_2\right] \end{aligned} \tag{20}$$

(Note that $\int_a^Q F(x)\,dx = \frac{(Q-a)^2}{2(b-a)}$ under U(a, b) distribution conditions.)

When making procurement decisions, Buyer 1 determines the optimal order quantity, Q^*, that maximizes $E\left[\pi^{SW}(Q)\right]$ in expression (20). However, the complexity of the expected profit function makes it too challenging to derive the optimal order quantity in a closed form. This study adopts a numerical method, determining the optimal order quantity using the minimize_scalar function from Python's scipy.optimize package (Python version 3.10.0). This function is based on Brent's method, which combines bisection, linear interpolation, and quadratic interpolation. Brent's method is particularly well-suited for this type of problem, as it efficiently locates the extremum of a continuous, nonlinear function of a single variable without requiring derivative information [22]. Moreover, it provides high numerical precision with no fixed time complexity, making it a robust and practical choice for solving the optimal ordering problem in this study.

Figure 2 illustrates the procedure for determining the optimal order quantity to be placed with the long-term contract supplier under the swap-integrated procurement strategy.

4. Numerical Studies

This section evaluates the performance of the procurement model proposed in the previous section by comparing it with a long-term wholesale contract. In addition, sensitivity analyses are conducted to examine the impact of key parameter variations on model performance.

The baseline demand scenario follows the setting of Koo [23], with some modification for swap settings. Buyer 1's demand, denoted by X, is modeled as a uniform distribution $X \sim U(100, 300)$, while Buyer 2's demand, Y, is defined as $Y \sim U(100c, 300c)$, reflecting the demand scaling factor, c. Initially, the value of c is fixed at 1. The main cost parameters are as follows: retail price $p = 60$, wholesale price $w = 40$, swap price $r = 50$, and penalty cost for unmet demand $g = 30$. Numerical experiments were conducted using Python.

4.1. Performance Comparison: Wholesale-Based vs. Swap-Integrated Procurement Strategies

This section compares the procurement performance between wholesale-based and swap-integrated procurement strategies under the same demand and cost structure. Figure 3 illustrates how the expected profit of the two procurement strategies changes with order quantity, identifies the profit-maximizing order quantity for each strategy, and shows the corresponding expected profit at that point. It shows that the SW strategy achieves higher expected profit while selecting a lower order quantity compared to the WH strategy. This suggests that the complementary mechanism of addressing excess and shortage demand through the swap mitigates the buyer's procurement risk and ultimately enables a more efficient profit structure.

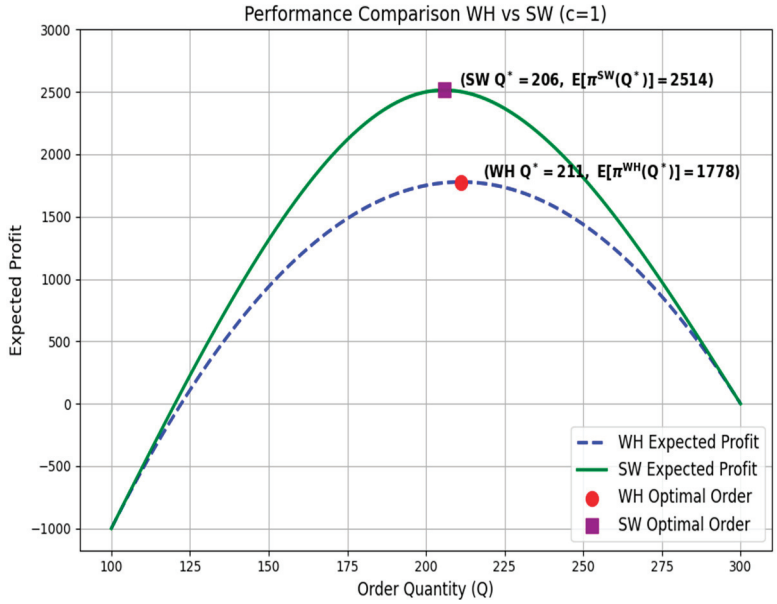


Figure 3. Comparison of the optimal order quantity and expected profit under the wholesale-only (WH) and swap-integrated (SW) procurement strategies in the baseline scenario. The SW strategy yields a 41.4% increase in expected profit with a 2.4% reduction in order quantity by enabling buyer-to-buyer inventory swaps, thereby mitigating both shortage and surplus costs.

Specifically, in the WH strategy, the optimal order quantity is 211 units, with an expected profit of 1,778. In contrast, under the SW strategy, the order quantity is slightly decreased to 206 units (−2.4%), while the expected profit is significantly increased to 2514 (+41.4%). The relatively lower optimal order quantity under the SW strategy is primarily attributed to the flexibility provided by the swap mechanism, which enables effective responses to both excess inventory and shortages. The SW strategy allows a buyer to procure surplus inventory from another buyer when demand exceeds expectations, and

conversely, to transfer surplus inventory to another buyer and earn additional revenue when demand falls short of expectations. This structure enables stable fulfillment of demand without excessive pre-ordering, thereby allowing for efficient operations even with a lower order quantity. However, it should be noted that a lower order quantity in the SW strategy should not be generalized, as it may vary depending on the demand structure, inter-buyer relationships, and the terms of the swap contract. Changes in the optimal order quantity may vary depending on factors such as demand structure, relationships between buyers, and swap conditions; these aspects are discussed in greater detail in the sensitivity analysis in Sections 4.2 and 4.3.

Meanwhile, the notable improvement in expected profit under the SW strategy stems from the swap mechanism effectively mitigating costs associated with supply-demand mismatches. When demand exceeds the pre-ordered quantity, the WH strategy obliges the buyer to bear the full penalty cost for the shortage. In contrast, the SW strategy enables the buyer to cover the shortage through swaps, thereby reducing the financial burden. Conversely, when demand falls below the ordered quantity, the WH strategy results in the buyer absorbing the full cost of surplus inventory, whereas the SW strategy allows excess inventory to be transferred to another buyer and converted into revenue.

These results demonstrate that the SW strategy provides a structural advantage by enabling efficient resource reallocation through inter-buyer cooperation and by allowing for adaptive responses to fluctuating demand. Consequently, the strategy mitigates procurement risks and enhances supply chain stability in environments characterized by high demand uncertainty.

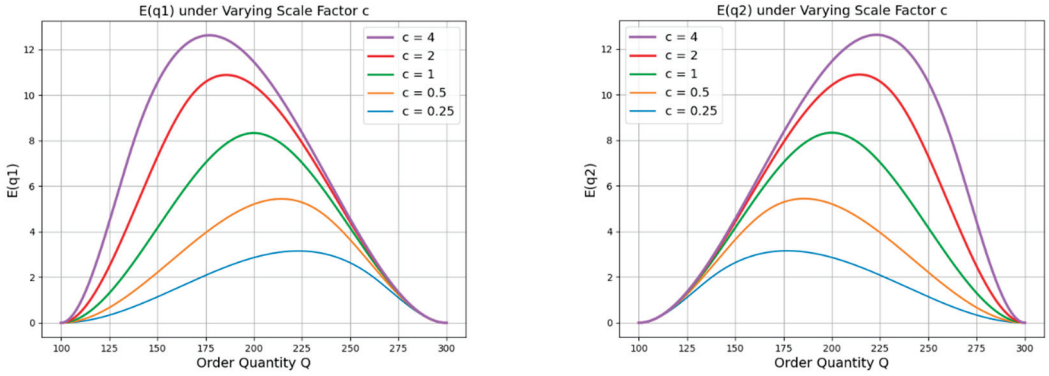
4.2. Effect of Demand Scaling Factor

This section examines the influence of demand asymmetry on the operational efficiency and performance of the SW strategy. To this end, the demand scaling factor, c, is designated as the key experimental variable. Sensitivity analyses are conducted with respect to three performance dimensions: the expected swap quantity, the optimal order quantity, and the expected profit. The analysis is performed by setting five scenarios in which c = 0.25, 0.5, 1, 2, and 4.

Figure 4 illustrates the expected swap quantities for Buyer 1 and Buyer 2 under the SW procurement strategy. As expected, the expected swap quantity reaches its minimum value of zero when the order quantity is equal to either the lower or upper bound of the demand. This result is intuitive, as either the inventory surplus or the inventory shortage becomes zero in such cases. It is also seen that both $E(q_1)$ and $E(q_2)$ exhibit a clear increasing trend as the demand scaling factor, c, increases. This result is also expected, as a larger demand scale of the collaborating partner implies greater potential for both surplus and shortage on their side. Consequently, when the focal buyer faces a shortage, the partner is more likely to have sufficient surplus to supply; conversely, when the focal buyer has excess inventory, the partner is more likely to absorb it due to higher unmet demand.

An interesting observation is that the order quantity, Q, that maximizes $E(q_1)$ and $E(q_2)$ varies with the value of c. Specifically, as c increases, the value of Q that maximizes $E(q_1)$ decreases, while the value that maximizes $E(q_2)$ increases. This outcome can be explained as follows: The expected swap inflow, $E(q_1)$, represents the amount of shortage that the focal buyer can replenish from the collaborating buyer's surplus inventory. The swap inflow is determined by the minimum of the focal buyer's shortage and the collaborating buyer's surplus. The peak of the $E(q_1)$ curve occurs when these two quantities are equal. As c increases, the collaborating buyer's demand becomes relatively larger, leading to more surplus inventory at a given order quantity. Consequently, the limiting factor for $E(q_1)$ gradually shifts from the collaborating buyer's surplus to the focal buyer's own shortage.

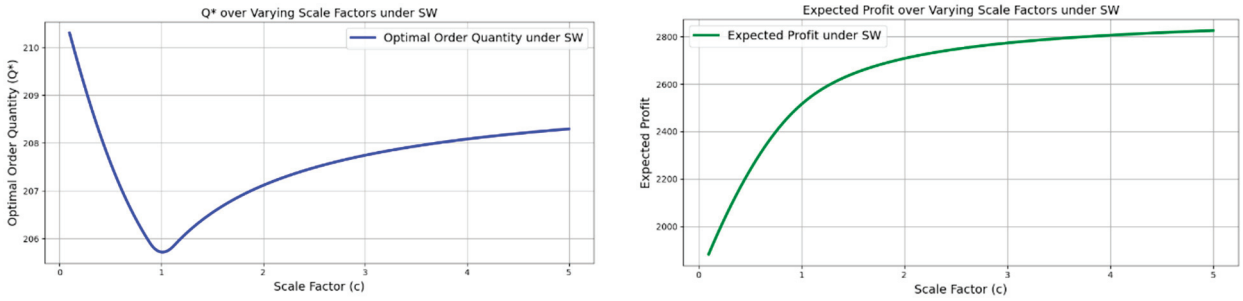
As a result, the point of equality between shortage and surplus occurs at a smaller Q, causing the peak of $E(q_1)$ to shift leftward as c increases. Similarly, the peak of $E(q_2)$ shifts rightward as c increases, since higher demand from the collaborating buyer leads to greater shortages on their side. To support more swap outflows, the focal buyer must order more to generate sufficient surplus. These opposing shifts intuitively illustrate how ordering decisions are structurally adjusted under the swap-integrated procurement strategy.



(**a**) Expected swap inflow, $E(q_1)$ (**b**) Expected swap outflow, $E(q_2)$

Figure 4. Expected swap transfer quantity under varying sale factor c. As c increases, the expected swap quantity rises, and the peaks of $E(q_1)$ and $E(q_2)$ shift in opposite directions due to changes in swap feasibility.

Figure 5 presents the optimal order quantity and expected profit under the SW strategy as the demand scaling factor, c, varies. Figure 5a shows that the optimal order quantity reaches a minimum near $c = 1$ and increases as the value of c deviates from this point in either direction. This pattern can be interpreted as the result of a combined effect of the structural feasibility of the swap mechanism and the associated economic incentives. When $c < 1$, the collaborating buyer, Buyer 2, has relatively small demand, which provides little opportunity for inventory contribution through swaps. Consequently, Buyer 1 adopts a more conservative strategy by procuring additional inventory in advance to hedge against potential shortages. In contrast, when $c > 1$, the collaborating partner is more likely to engage in substantial swap transactions, thereby reducing inventory risk for Buyer 1. As a result, it becomes advantageous for Buyer 1 to adopt a more relaxed ordering policy.



(**a**) Optimal order quantity (**b**) Expected profit

Figure 5. Optimal order quantity and expected profit under the SW strategy across varying scale factors. (**a**) A pattern in which the optimal order quantity reaches its minimum around $c = 1$ and increases as c becomes either smaller or larger. (**b**) A gradual increase in expected profit with increasing values of c.

The expected profit results shown in Figure 5b exhibit an upward trend as c increases, with growth being particularly pronounced when $c < 1$. When the demand scaling factor, c, is low, the inventory surplus or shortage of the counterpart with lower demand is relatively small, resulting in a limited volume of swap transactions. Consequently, the focal buyer's ability to compensate for inventory imbalances through swaps is constrained, leading to lower expected profits. On the other hand, as c increases, the inventory surplus or shortage of the counterpart becomes more significant, enhancing the potential for mitigating the focal buyer's inventory imbalance through swap transactions. This leads to a steady increase in expected profit, although the rate of increase slows down. Note that the results presented here are based on an assumption of a swap price of 50. The next section investigates how variations in the swap price influence the model's performance.

4.3. Effect of Swap Price

According to the results from the baseline scenario analysis, the SW strategy tends to yield a higher expected profit and a relatively lower optimal order quantity compared to the WH strategy. This is attributed to the fact that the swap mechanism enables buyers to complement excess and shortage demand through mutual adjustment, thereby encouraging a more conservative ordering strategy. However, whether this tendency holds consistently across different procurement environments requires further examination. In particular, if the swap terms are sufficiently favorable, buyers may instead be incentivized to adopt a more aggressive ordering strategy. In other words, it is uncertain whether the swap strategy always leads to a lower order quantity than the wholesale-only strategy.

To address this question, this section examines how changes in the swap price affect the decision-making behavior of swap participants and the performance of the proposed model, with particular focus on how performance varies according to the size of the counterpart. In the experiment, the demand scale factor, c, varies from 0.1 to 5, while the swap price, r, is set at five discrete levels: 30, 40, 50, 60, and 70.

Figure 6 illustrates how the optimal order quantity under the SW strategy is adjusted in response to changes in the demand scaling factor, c, for different levels of the swap price, r. Each solid line represents the ordering curve corresponding to a specific value of r under the SW strategy while the dotted line indicates the optimal order quantity under the WH strategy $\left(Q_{FW}^{*} = 211\right)$.

The experimental results reveal several noteworthy insights:

(1) When c = 1, the focal and collaborating buyers have identical demand distributions, resulting in a symmetric setting. In this case, the probabilities of surplus and shortage are evenly balanced between the two parties, making the expected gains from swap transactions invariant to the swap price, r. As a result, the optimal order quantity under the SW strategy remains constant regardless of the value of r. Furthermore, this order quantity is consistently lower than that of the WH strategy, reflecting the risk-mitigating role of the swap mechanism. By enabling post-demand adjustments, the swap contract reduces the need for inventory buffers, thereby encouraging more conservative upfront procurement.

(2) When $c \neq 1$, the effect of the swap price, r, on the optimal ordering behavior diverges significantly depending on whether $c < 1$ or $c > 1$. This outcome is driven by the directional behavior of the expected swap quantities and their influence on the profit function. Recall that the r-dependent part of the expected profit in expression (5) is stated as $-r \cdot E(q_1) + r \cdot E(q_2)$. Accordingly, the relative magnitude of $E(q_1)$ and $E(q_2)$ plays a critical role in shaping the impact of swap price changes. As shown in Figure 4, when $c < 1$, a lower order quantity, Q, tends to generate $E(q_2) > E(q_1)$, implying that an increase in r yields a net gain in the term $-r \cdot E(q_1) + r \cdot E(q_2)$. Therefore, in this

case, higher swap prices make lower Q values more desirable, leading to a decline in optimal order quantity as r increases. Conversely, applying a similar line of analysis reveals that when $c > 1$, the optimal order quantity tends to increase as the swap price, r, rises.

(3) In most cases, the SW strategy yields a lower optimal order quantity than the WH strategy. However, when both c and r are sufficiently high, the SW strategy may induce a higher order quantity than the WH benchmark. For instance, when $r = 70$, the SW strategy begins to induce higher ordering quantities than the WH strategy for $c \geq 2.6$. This is because the incentive to dispose of excess inventory through swaps becomes stronger when the swap price is high. Also, a similar pattern emerges when $c < 1$ and the swap price r is very low. In this case, even if surplus inventory occurs, it may be offloaded to the counterpart, which reduces the burden of holding excess stock. At the same time, the low swap price strengthens the collaborating buyer's incentive to participate in swaps, thereby increasing the likelihood of successful transactions. As a result, it may be more profitable for the focal buyer to adopt a more aggressive ordering strategy. Accordingly, when the swap price is either sufficiently high or very low—depending on the demand structure—the SW strategy may actually lead to more aggressive ordering than the WH strategy.

(4) When the swap price, r, is low, increasing the scale factor, c, does not lead to a higher order quantity under the SW strategy; instead, the focal buyer tends to adopt a more conservative ordering approach. This outcome reflects the interplay of two reinforcing factors. First, a low swap price provides the focal buyer with a cost-effective means of addressing potential shortages, as inventory can be secured post hoc at a price lower than the wholesale cost, w. This reduces the incentive for large initial orders. Second, the financial benefit from offloading surplus inventory through swaps is diminished at lower r values, limiting the upside of over-ordering. Together, these effects weaken the buyer's motivation to place aggressive orders in advance and instead encourage strategic reliance on the swap mechanism as a flexible, lower-cost buffer against demand uncertainty.

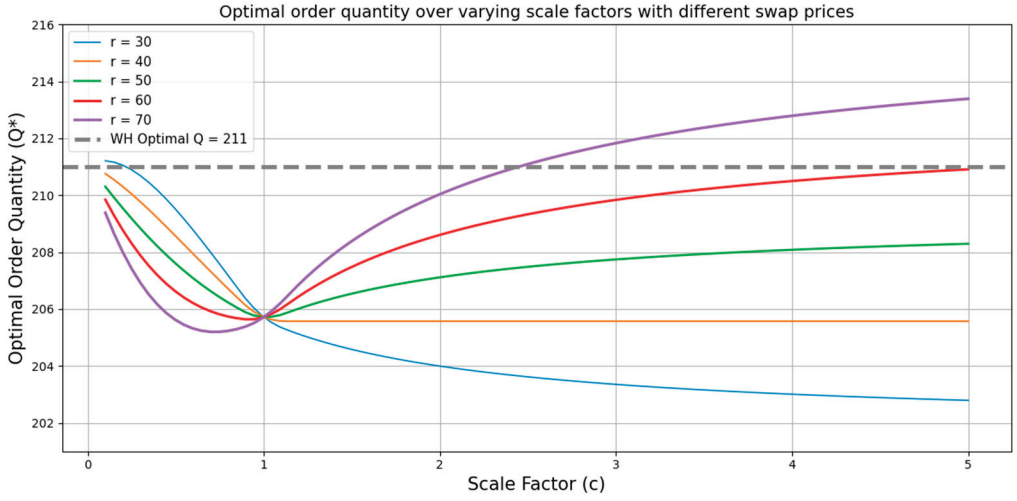


Figure 6. Optimal order quantity across varying scale factors under different swap prices. The solid lines illustrate how optimal order decisions adjust in response to changes in demand asymmetry and swap pricing. Notably, the order quantity remains unchanged when $c = 1$ and shows opposing trends depending on whether swap price, r, is high or low. The dashed line represents the optimal order quantity under the WH strategy and is provided for comparison.

The numerical findings, supported by Figure 6, show that the effectiveness of the SW strategy is highly sensitive to the interaction between the demand scaling factor, c, and

the swap price, r. When demand asymmetry is high and swap conditions are favorable, the SW strategy can induce even more aggressive ordering than the traditional wholesale approach, transforming it from a risk-hedging mechanism into a tool for strategic profit optimization. However, under low swap prices or limited counterpart demand, the strategy encourages conservative procurement, highlighting the need for careful contract design and coordination. These results underscore the practical value of swap-integrated procurement, not only for balancing inventory risks but also for enabling adaptable sourcing decisions in both private and public sector supply chains.

The subsequent analysis examines how variations in the swap price, r, and the demand scaling factor, c, affect the expected profit under the SW strategy. Figure 7 presents a heatmap that visualizes the expected profit across different combinations of r and c, thereby illustrating how the profitability of the SW strategy varies depending on the procurement environment. The results reveal that the expected profit under the SW strategy generally increases with the demand scaling factor, c. A higher value of c implies greater demand from the collaborating buyer, which in turn increases the likelihood of successful swap transactions and enhances the opportunity to offload excess inventory under favorable conditions.

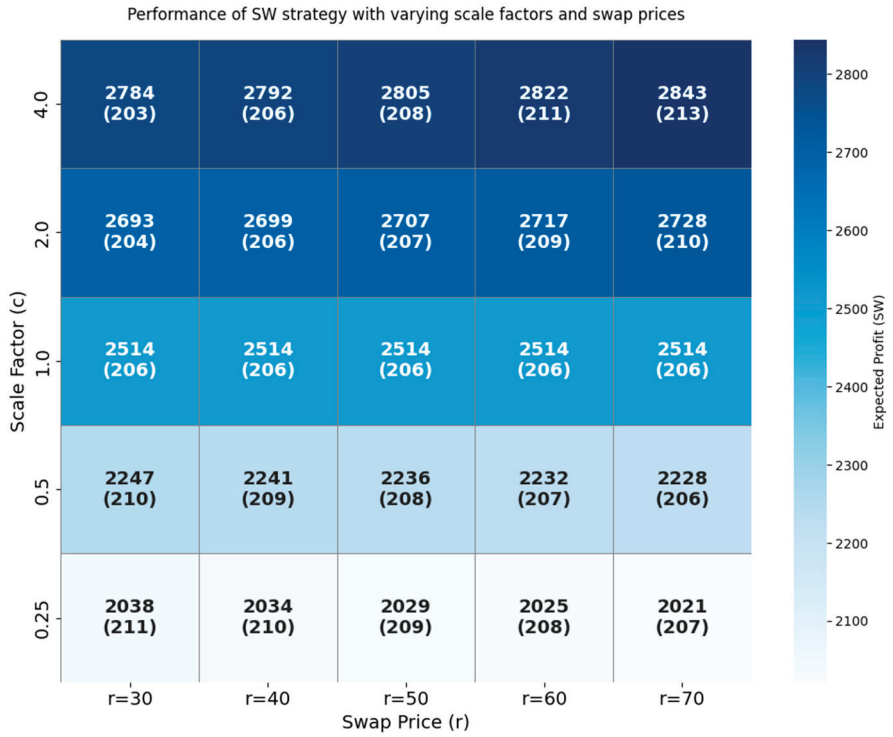


Figure 7. Performance of the SW procurement strategy under varying scale factors and swap prices. Each cell presents the expected profit (upper value) and the corresponding optimal order quantity Q^* (in parentheses). As the demand scale factor, c, increases, the expected profits generally improve due to enhanced swap flexibility, while the optimal order quantity adjust accordingly based on the swap price, r.

However, the impact of the swap price, r, is not uniform and depends on the demand asymmetry. When $c < 1$, the expected swap inflow is larger than the expected swap outflow with order quantities above 200 (see Figure 4). Therefore, from the perspective of the focal buyer, a higher swap price leads to a lower expected profit compared to a lower swap price. This is because although a higher swap price increases the revenue gained from supplying surplus inventory through swaps, it also increases the cost of procuring inventory via swaps. The reduction in swap procurement cost outweighs the loss in swap

revenue, resulting in a net gain in profit. However, since the counterpart's surplus or shortage is relatively small, its contribution to the profit is limited. Conversely, when $c > 1$, the expected swap inflow is smaller than the expected swap outflow, which contrasts with the case of $c < 1$. One interesting observation is that the impact of r on the focal buyer's expected profit is less pronounced than that of c. This can be attributed to the bidirectional nature of swap transactions, which involve both procuring shortages and supplying surpluses. As a result, the effect of the swap price tends to be largely offset.

Overall, these findings underscore that the expected profit from the SW strategy is shaped not by isolated parameters but by the interaction between demand structure and contractual terms. When demand asymmetry is high and swap prices are favorable, the SW strategy serves not only as a risk-mitigation mechanism but also as a tool for profit maximization. In contrast, when the collaborating buyer has limited demand and swap prices are high, the profitability of the strategy may decline. Therefore, aligning contract terms with the underlying demand structure is essential for the effective implementation of swap-integrated procurement.

5. Conclusions

This study proposed a swap-integrated procurement strategy that combines a long-term wholesale-based contract to enhance flexibility and risk mitigation under demand uncertainty. By enabling post-demand inventory exchanges between buyers, the strategy addresses limitations inherent in single-contract structures. The theoretical contributions of this study can be summarized in the following key points:

(1) This study presents a novel application of swap mechanisms, traditionally used in financial sectors, to the context of buyer-to-buyer inventory reallocation. By modeling swaps as contractual exchanges of surplus and shortage inventory between buyers, this research reconceptualizes swaps as a post-demand coordination mechanism in supply chain operations.
(2) A model is proposed to estimate the expected swap inflows and outflows resulting from swap transactions.
(3) An analytical model is presented for the expected profit in a system that adopts a swap-integrated procurement policy.
(4) Under the assumption that demand is uniformly distributed, a closed-form expression for the expected profit is derived, and a method for determining the optimal order quantity is presented.
(5) Experimental analyses are conducted to quantify the strategic benefits of swap integration by comparing it to traditional wholesale-based procurement. The results highlight the influence of key parameters, such as demand asymmetry and swap price, on the strategy's effectiveness.

The findings demonstrate that the proposed strategy consistently outperforms long-term wholesale contracts in terms of profitability and stability; under the baseline scenario, it achieves a 41.4% increase in expected profit compared to the wholesale-based strategy, which clearly supports the value of swap integration.

From a practical perspective, the study offers a strategic approach to mitigating inventory risk in supply chains with highly unpredictable demand, such as vaccines, fresh produce, or electronic components, by utilizing swap contracts. By analyzing performance variations across different levels of swap price and demand asymmetry, this study provides practical guidelines for contract design. For instance, the results indicate that swap contracts are particularly advantageous for the focal buyer when the counterpart has a relatively larger demand scale and the swap price is high. These results can guide decision-makers in structuring contract terms in various procurement environments and offer valuable insights

for designing cooperative structures, including participant selection and prioritization in swap strategies. However, it should be noted that these results are derived from the specific parameter settings assumed in this study. In practice, different supply chain environments may yield different outcomes; therefore, careful consideration of the actual procurement context is essential when applying these insights in practice.

This study is subject to several limitations that offer directions for future research. First, the model adopts a single-period setting, which does not capture multi-period dynamics such as inventory rollover or learning effects. Second, it assumes that swaps are always executed when conditions are met, overlooking real-world frictions such as information asymmetry or negotiation failures. Third, the analysis is based on uniform demand; future studies may examine the model's robustness under more realistic demand patterns. Fourth, the supply capacity is sufficient with fully reliable suppliers. Fifth, the current framework involves only two buyers in the swap contract. These assumptions may deviate from real-world supply chain conditions, highlighting the need for further research to relax or generalize them. In addition, future extensions could explore portfolio-based strategies that combine swap contracts with other mechanisms such as options or spot markets. Finally, in this study, the numerical analysis was conducted under a specific setting, in which parameters such as the selling price, wholesale price, and stockout penalty were fixed. Variations in these parameter values are expected to influence the performance of the proposed swap-integrated procurement strategy. Investigating how the model behaves across a broader set of parameter configurations remains a promising direction for future research.

Author Contributions: Conceptualization, P.-H.K.; methodology, M.-Y.R. and P.-H.K.; software, M.-Y.R.; validation, M.-Y.R. and P.-H.K.; formal analysis, M.-Y.R. and P.-H.K.; resources, P.-H.K.; data curation, M.-Y.R.; writing—original draft preparation, M.-Y.R.; writing—review and editing, M.-Y.R. and P.-H.K.; supervision, P.-H.K.; funding acquisition, P.-H.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Korea (No. 2022R1I1A3070919).

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; the collection, analysis, or interpretation of the data; the writing of the manuscript; or the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

WH	Long-term wholesale-based procurement strategy
SW	Swap-integrated procurement strategy

References

1. Alkhatib, S.F.; Momani, R.A. Supply Chain Resilience and Operational Performance: The Role of Digital Technologies in Jordanian Manufacturing Firms. *Adm. Sci.* **2023**, *13*, 40. [CrossRef]
2. Choi, T.-M.; Li, D.; Yan, H. Mean–variance analysis of a single supplier and retailer supply chain under a returns policy. *Eur. J. Oper. Res.* **2008**, *184*, 356–376. [CrossRef]
3. Cachon, G.P.; Lariviere, M.A. Supply Chain Coordination with Revenue-Sharing Contracts: Strengths and Limitations. *Manag. Sci.* **2005**, *51*, 30–44. [CrossRef]

4. Zhao, Y.; Choi, T.-M.; Cheng, T.C.E.; Wang, S. Supply option contracts with spot market and demand information updating. *Eur. J. Oper. Res.* **2018**, *266*, 1062–1071. [CrossRef]
5. The Bank of Korea. The Bank of Korea and the Bank of Japan Announce an Increase in the Size of the Won-Yen Swap Arrangement, 2008, Press Release. Available online: https://www.bok.or.kr/eng/bbs/E0000634/view.do?menuNo=400423&nttId=146770 (accessed on 15 July 2025).
6. Australian Government Department of Health and Aged Care. COVID-19 Vaccine Rollout Update: International Dose Swap Arrangements. 2024. Available online: https://www.health.gov.au/our-work/covid-19-vaccines/about-rollout/vaccine-agreements (accessed on 26 June 2025).
7. Reuters. India's GAIL Issues Swap Tender for 12 LNG Cargoes. Available online: https://www.reuters.com/business/energy/indias-gail-issues-swap-tender-12-lng-cargoes-sources-say-2025-02-04/ (accessed on 26 June 2025).
8. Ekren, B.Y.; Arslan, M.C. Simulation-based lateral transshipment policy optimization for s, S inventory control problem in a single-echelon supply chain network. *Int. J. Optim. Control. Theor. Appl.* **2020**, *10*, 9–16. [CrossRef]
9. Al Husain, R.; Assavapokee, T.; Khumawala, B. Modelling the supply chain swap problem in the petroleum industry. *Int. J. Appl. Decis. Sci.* **2008**, *1*, 261. [CrossRef]
10. Bidyarthi, H.M.J.; Deshmukh, L.B. Swap in downstream petroleum supply chain: An effective inventory handling tool. In *Driving the Economy Through Innovation and Entrepreneurship*; Mukhopadhyay, C., Akhilesh, K.B., Srinivasan, R., Gurtoo, A., Ramachandran, P., Eds.; Springer: Berlin/Heidelberg, Germany, 2013; pp. 363–370. [CrossRef]
11. Al-Husain, R.; Khorramshahgol, R. A comprehensive analysis of the determinants of swap problem in the supply chain of the petroleum industry. *Int. J. Stat. Probab.* **2016**, *5*, 97–108. [CrossRef]
12. Farahani, M.; Rahmani, D. Production and distribution planning in petroleum supply chains regarding the impacts of gas injection and swap. *Energy* **2017**, *141*, 991–1003. [CrossRef]
13. Dizbay, I.E.; Ozturkoglu, O. Product swapping and transfer sales between suppliers in a balanced network. In Proceedings of the 2013 Federated Conference on Computer Science and Information Systems, Krakow, Poland, 8–11 September 2013; pp. 1191–1194.
14. Wang, Y.; Meng, Q.; Tan, Z. Short-term liner shipping bunker procurement with swap contracts. *Marit. Policy Manag.* **2017**, *45*, 211–238. [CrossRef]
15. Park, S.J. Swapping inventory between competing firms. *Int. J. Prod. Econ.* **2018**, *199*, 26–46. [CrossRef]
16. Kemper, A.; Schmeck, M.D.; Balci, A.K. The market price of risk for delivery periods: Pricing swaps and options in electricity markets. *Energy Econ.* **2022**, *113*, 106221. [CrossRef]
17. Zhang, L.; Thomspon, R.G. Optimising product swaps in an urban retail network with uncertainty. *Transp. Res. Procedia* **2024**, *79*, 210–217. [CrossRef]
18. Diks, E.B.; de Kok, A.G. Controlling a divergent two-echelon network with transshipments using linear programming. *Eur. J. Oper. Res.* **1996**, *45*, 369–379. [CrossRef]
19. Paterson, C.; Kiesmüller, G.; Teunter, R.; Glazebrook, K. Inventory models with lateral transshipments: A review. *Eur. J. Oper. Res.* **2011**, *210*, 125–136. [CrossRef]
20. Kumari, A.G.; Wijayanayake, A.; Niwunhella, H. Lateral Transshipment Inventory Models: A Systematic Literature Review of Models and Solution Approaches. *Proc. Int. Conf. Bus. Manag.* **2022**, *18*, 54–80. [CrossRef]
21. Aggarwal, P.; Ganeshan, R. Using risk-management tools on B2Bs: An exploratory investigation. *Int. J. Prod. Econ.* **2007**, *108*, 2–7. [CrossRef]
22. Zhang, Z. An Improvement to the Brent's Method. *Int. J. Exp. Algorithms* **2011**, *2*, 21–26.
23. Koo, P.-H. Exploring the relationship between buyback and revenue sharing contracts in channel coordination. *J. Korean Inst. Ind. Eng.* **2024**, *50*, 211–221. [CrossRef]

mathematics

MDPI

Article

Wasserstein Distributionally Robust Optimization for Chance Constrained Facility Location Under Uncertain Demand

Iman Seyedi [1], Antonio Candelieri [2,*], Enza Messina [1] and Francesco Archetti [1]

[1] Department of Computer Science Systems and Communication, University of Milano-Bicocca, 20126 Milan, Italy; seyediman.seyedi@unimib.it (I.S.); enza.messina@unimib.it (E.M.); francesco.archetti@unimib.it (F.A.)
[2] Department of Economics Management and Statistics, University of Milano-Bicocca, 20126 Milan, Italy
* Correspondence: antonio.candelieri@unimib.it

Abstract

The purpose of this paper is to present a novel optimization framework that enhances Wasserstein Distributionally Robust Optimization (WDRO) for chance-constrained facility location problems under demand uncertainty. Traditional methods often rely on predefined probability distributions, limiting their flexibility in adapting to real-world demand fluctuations. To overcome this limitation, the proposed approach integrates two methodologies, specifically a Genetic Algorithm to search for the optimal decision about facility opening, inventory, and allocation, and a constrained Jordan–Kinderlehrer–Otto (cJKO) scheme for dealing with robustness in the objective function and chance-constraint with respect to possible unknown fluctuations in demand. Precisely, cJKO is used to construct Wasserstein ambiguity sets around empirical demand distributions (historical data) to achieve robustness. As a result, computational experiments demonstrate that the proposed hybrid approach achieves over 90% demand satisfaction with limited violations of probabilistic constraints across various demand scenarios. The method effectively balances operational cost efficiency with robustness, showing superior performance in handling demand uncertainty compared to traditional approaches.

Keywords: constrained JKO (cJKO); chance-constrained optimization; Wasserstein distance; facility location; Genetic Algorithm (GA)

MSC: 90C17

1. Introduction

Supply chain management (SCM) is critical in modern logistics and resource allocation. With the rise of the sharing economy, supply chains are becoming more decentralized, flexible, and demand driven. In the era of global uncertainty and volatile markets, supply chain networks are increasingly exposed to risks stemming from unpredictable factors such as demand fluctuations, global disruptions (e.g., COVID-19), geopolitical risks, and operational variability [1]. If not properly addressed, these uncertainties lead to inventory imbalances, excess operational costs, and reduced service levels.

The location of a facility (FLP) is a classical subject area in combinatorial optimization with wide applications ranging from communication systems to supply chain planning, public services, and economic development. FLP occurs in various real applications like the location of logistical centers, industrial estates, gas filling stations, communication base stations, and other essential infrastructures such as hospitals, schools, supermarkets, and

warehouses [2]. Despite its extensive use, classical FLP models usually assume deterministic conditions that do not reflect the intrinsic uncertainties characterizing real-world systems. Consequently, the obtained solutions are brittle and not optimal when subjected to real-time variations. To build robust and affordable supply chains, uncertainty must be integrated directly within both the framework of modeling and optimization. Among all the different kinds of uncertainty, demand uncertainty is especially significant since it has direct effects on capacity planning, inventory, and service fulfillments. Tackling this uncertainty has become a major challenge in the fields of operations research, decision science, and machine learning. Several methodological paradigms have emerged to address this issue, each grounded in different assumptions about the availability and structure of data—ranging from classical stochastic optimization to robust and distributionally robust approaches.

A standard approach to addressing uncertainty in optimization is Stochastic Optimization (SO), where decision-making is based on an assumed known probability distribution of demand. However, the true demand distribution is unknown, and only historical data (empirical distribution) are available. To model this uncertainty, distributionally robust optimization (DRO) has emerged as a powerful technique. Unlike traditional stochastic methods based on sample average approximation (SAA) optimization, DRO considers an ambiguity set of possible distributions instead of assuming a single known distribution. Table 1 presents three key optimization paradigms—SO, Worst-Case Robust Optimization (RO), and DRO—and their respective mathematical formulations.

Table 1. Different optimization objectives considered in Bayesian optimization.

Optimization Type	Mathematical Formulation
Stochastic Optimization (SO)	$\max_x \mathbb{E}_{c\sim P}[f(x,c)]$
Worst-Case Robust Optimization (RO)	$\max_x \min_{c\in\Delta} f(x,c)$
Distributionally Robust Optimization (DRO)	$\max_x \inf_{Q\in U} \mathbb{E}_{c\sim Q}\left[f(x,c)\right]$

where x represents the decision variable to be optimized and c denotes the uncertain parameter or *context* influencing the outcome, such as demand or price. P is the probability distribution assumed known or estimated (in SO), Δ is a predefined set of scenarios (in RO), and Q is a probability distribution within the ambiguity set U in DRO, where U is a set of plausible distributions for c. DRO optimizes the decision x considering the worst-case scenario across all distributions in U.

In this paper, the uncertain parameter is the customer's demand, that is, a I-dimensional vector with I the number of customers. Thus, a scenario is represented by one demand vector and a set of scenarios—for instance, a historical dataset—can be viewed as a point clouds, namely, an empirical multivariate distribution capturing the inherent fluctuations in customer behavior due to external factors such as market trends and disruptions.

Given these fluctuations, it is crucial to develop robust facility location and allocation strategies that remain effective under a wide range of possible future demands. To this end, we address the distributionally robust facility location problem under demand uncertainty, where the true demand distribution is unknown and only sample-based approximations are available. Instead of relying on a single known distribution, we propose a Wasserstein Distributionally Robust Optimization (WDRO) framework, which defines an ambiguity set (aka Wasserstein ball) of plausible distributions around historical data.

The Wasserstein distance is rooted in optimal transport theory [3–5] and provides a mathematically rigorous and flexible measure of difference between distributions, making it particularly well-suited for constructing ambiguity sets in data-driven settings. Its usefulness has been recently reported, for instance, in the optimization of composite functions [6].

We formulate a joint optimization problem that determines

(i) facility opening decisions,
(ii) inventory levels, and
(iii) customer allocations,

with the goal of minimizing the total supply chain cost, including fixed facility costs, inventory holding costs, transportation costs, and penalties for unmet demand, while satisfying demand with high probability.

To efficiently solve this complex and high-dimensional problem, we use a Genetic Algorithm (GA) to heuristically generate solutions for both discrete and continuous decision variables. Furthermore, we integrate a constrained Jordan–Kinderlehrer–Otto (cJKO) scheme, originally presented in [7]. Unlike standard JKO methods, cJKO provides a more effective and efficient mechanism for solving optimization problems over a space of probability distributions equipped with the Wasserstein distance.

The resulting hybrid metaheuristic-variational framework enables the effective exploration of the combinatorial decision space and robust handling of distributional uncertainty. By leveraging the Wasserstein distance, our model balances adaptability and robustness, avoiding overfitting to historical data while remaining responsive to dynamic demand shifts. Overall, this integrated approach ensures that facility placement and resource allocation remain cost-effective and resilient. The iterative cJKO updates allow for the dynamic refinement of uncertainty sets, improving the solution quality over time. Our method provides a scalable and theoretically grounded solution for optimizing supply chain operations under deep uncertainty.

The remainder of this paper is structured as follows: Section 2 reviews the relevant literature, highlighting the applications of stochastic optimization, robust optimization, and distributionally robust optimization in facility location problems. Section 3 presents the problem statement, formally defining the optimization problem. In Section 4, we introduce the approach of optimizing over probability distributions, discussing how uncertainty is incorporated into the decision-making process. Also, this section describes the metaheuristic-based approach used to efficiently solve the problem, highlighting its advantages. In Section 5, we provide computational results, demonstrating the performance of our proposed methodology through extensive numerical experiments. Finally, Section 6 concludes the paper by summarizing the key findings and outlining potential directions for future research.

2. Literature Review

Facility location problems under uncertainty have long been a central theme in operations research and supply chain design [8]. Classical approaches often rely on deterministic formulations, assuming perfect knowledge of input parameters such as demand and transportation costs [9]. However, real-world supply chains are increasingly impacted by volatile customer demand, supply disruptions, and economic shifts. To address these challenges, several optimization paradigms have been proposed, each with distinct assumptions and trade-offs. In this section, we review the literature surrounding these approaches and their application to facility location problems, with a focus on recent advances incorporating Wasserstein-based DRO models and metaheuristic strategies.

2.1. Stochastic Optimization in Facility Location

SO aims to maximize the expected performance of a decision x under a fixed probability distribution P of the uncertain parameter c (Formula (1)). This approach assumes that P is known in advance or can be estimated reliably from historical data.

$$\max_x \mathbb{E}_{c\sim P}[f(x,c)], \tag{1}$$

SO has been extensively applied to facility location problems where demand or cost parameters are modelled as random variables with known probability distributions. Classical SO methods aim to minimize the expected cost by solving a two-stage or multi-stage model. In the first stage, strategic decisions such as facility openings are made, while in the second stage, recourse actions, including allocations or inventory replenishments, are optimized based on the realization of uncertainty. Works such as those by Snyder [10] and Laporte et al. [11] provide foundational methodologies for modeling uncertainty in facility location through scenario-based planning.

However, one major limitation of SO lies in its reliance on the accuracy of the assumed probability distribution, which is often derived from historical or simulated data. In real-world applications, these distributions may not accurately reflect future conditions due to dynamic market trends, policy changes, or unforeseen disruptions. As a result, solutions based on SO can become suboptimal or even infeasible when the actual demand distribution deviates from the expected one. This sensitivity to distributional misspecification can undermine the robustness of the decisions derived from SO models. Nevertheless, SO remains a widely used and valuable benchmark approach in the facility location literature and has been successfully applied to various settings, including capacitated facility location problems, multi-period planning, and emergency service deployment.

2.2. Robust Optimization in Facility Location

RO takes a conservative approach by optimizing for the worst possible realization of the uncertain parameter c within a predefined set Δ (Formula (2)). The goal is to ensure that the decision x performs well even under the most adverse conditions. This approach is particularly useful in safety-critical applications, such as finance and healthcare, where extreme scenarios must be accounted for.

$$\max_{x} \min_{c \in \Delta} f(x, c), \tag{2}$$

RO offers an alternative approach by eschewing probabilistic information in favor of worst-case guarantees. The underlying philosophy of RO is to optimize decisions that perform acceptably under all realizations of uncertainty within a predefined uncertainty set. Kouvelis and Yu [12] were among the early pioneers in introducing RO into supply chain planning, and their work was followed by many others who sought to build resilient networks without requiring probabilistic information.

RO has been extensively applied in the facility location literature to ensure system performance under extreme or adversarial demand scenarios. For instance, Jabbarzadeh et al. [13] proposed a robust location allocation model for humanitarian supply chains that guarantees network resilience even under the most disruptive conditions. Despite its effectiveness in providing high levels of reliability, RO tends to be overly conservative. By focusing solely on worst-case scenarios, it often neglects probabilistic information about more likely demand realizations. This leads to overly cautious solutions that can result in excessive costs under normal operating conditions. Furthermore, RO does not fully exploit valuable insights derived from empirical data, potentially missing opportunities to achieve a more balanced trade-off between cost-efficiency and robustness.

2.3. DRO in Facility Location

DRO provides a middle ground between SO and RO by considering uncertainty in the probability distribution itself. Instead of assuming that the context c follows a fixed distribution P, DRO optimizes for the worst-case expectation over all distributions Q within an uncertainty set U [14] (Formula (3)). This makes DRO highly effective in environments where the true distribution is unknown or subject to shift. The uncertainty set U is often

defined using divergence measures, such as the Wasserstein distance, φ-divergences, or Maximum Mean Discrepancy (MMD) [15]. DRO has gained significant attention due to its ability to account for distributional shifts while avoiding the extreme conservatism of RO [16]. It has applications in robust machine learning models, adversarial training, reinforcement learning, and optimization problems under uncertainty. In the context of Bayesian optimization, adopting a DRO-based approach can significantly improve model robustness [17], particularly in non-stationary environments where data distributions change over time [18]. By selecting the appropriate optimization paradigm based on the problem's uncertainty characteristics, decision-makers can better balance performance and robustness in their optimization tasks.

$$\max_{x} \inf_{Q \in U} \mathbb{E}_{c \sim Q}[f(x, c)], \tag{3}$$

Many studies have explored DRO in facility location, often using moment-based or φ-divergence-based ambiguity sets. For example, Liu et al. [19] formulated a DRO model for emergency medical service station location using joint chance constraints. Their model ensured service reliability by optimizing over all distributions sharing known first- and second-order moments. More flexible are Wasserstein-based ambiguity sets, which define the uncertainty set using optimal transport metrics. Ji and Lejeune [20] applied Wasserstein DRO to chance-constrained facility location, demonstrating superior robustness and computational tractability. These models can incorporate empirical distributions (i.e., point clouds) and adapt to various demand scenarios, making them particularly effective for supply chain applications.

Wasserstein-based DRO has gained prominence due to its intuitive geometric interpretation, tractable reformulations, and strong theoretical guarantees. It measures the cost of transporting probability mass between distributions, naturally aligning with the logistics and spatial nature of facility location problems. Wang et al. [21] applied Wasserstein DRO to disaster relief logistics, modeling joint chance constraints under data-driven uncertainty.

Their results showed improved service levels and robustness under distributional shifts. Recent developments in Wasserstein DRO have focused on computational efficiency and scalability challenges: Ref. [22] provided a comprehensive duality framework for Wasserstein DRO that offers more efficient reformulations, while advances in statistical distance-based DRO algorithms have addressed computational complexity in large-scale scenarios through decomposition and approximation techniques.

2.4. Types of Ambiguity Sets in DRO

Before constructing a DRO model, it is crucial to define how the uncertainty about the true distribution is represented. While SO assumes a known distribution and RO operates without any probabilistic assumptions, DRO acknowledges distributional uncertainty—the fact that the true probability distribution is only partially known. This uncertainty is formalized through an ambiguity set, a collection of distributions deemed plausible based on the available data, prior knowledge, or statistical properties. The DRO model then seeks to optimize performance under the worst-case distribution within this ambiguity set. As such, the construction of the ambiguity set plays a central role in balancing robustness and conservatism, influencing both the theoretical properties and empirical outcomes of the solution.

Ambiguity sets can be constructed based on different principles:

- Moment-based ambiguity sets—constraints on mean, variance, and higher moments [16];

- Discrepancy-based ambiguity sets—define a neighborhood of plausible distributions around a reference measure (e.g., Kullback–Leibler divergence and total variation distance) [23];
- Wasserstein-based ambiguity sets—a natural and flexible framework using the Wasserstein metric to measure the distance between distributions [24,25].

Among these, the Wasserstein-based ambiguity set has recently gained dominance due to its mathematical/statistical background. The Wasserstein distance measures the cost of transporting probability mass between distributions. This makes it highly suitable for supply chain applications, where demand fluctuations can be viewed as shifts of empirical distributions (i.e., set of historical demand data) within a neighborhood of a certain Wasserstein radius.

2.5. Computational Trade-Offs in Distributionally Robust Facility Location

The computational complexity of DRO-based facility location models presents significant challenges that must be carefully balanced against solution quality and robustness benefits. Unlike traditional stochastic optimization, which deals with a fixed probability distribution, DRO requires optimization over an entire ambiguity set, fundamentally increasing the problem's computational burden. The primary computational challenges arise from the need to solve min-max optimization problems, where the inner maximization over the ambiguity set often leads to semi-definite programming reformulations or requires sophisticated duality theory [26]. In facility location contexts, this complexity is compounded by the discrete nature of location decisions and the continuous nature of allocation variables, creating mixed-integer optimization problems that are inherently difficult to solve. While Wasserstein-based ambiguity sets offer intuitive geometric interpretations and strong theoretical guarantees, they come with higher computational costs compared to moment-based or φ-divergence approaches. The state-of-the-art methods for Wasserstein DRO rely on global optimization techniques, which quickly become computationally prohibitive for large-scale problems [25]. However, recent algorithmic advances have improved tractability through efficient reformulations and decomposition strategies. Practitioners must balance between problem size, solution time, and robustness level. Larger ambiguity sets provide greater distributional robustness but exponentially increase the computational requirements. Effective scenario identification techniques can reduce this computational burden by focusing on scenarios that most significantly impact the optimal solution [26].

2.6. Metaheuristic Approaches and Genetic Algorithms in Facility Location

Given the combinatorial complexity of facility location problems, especially under uncertainty, exact algorithms often struggle to scale. Metaheuristic methods such as Genetic Algorithms (GAs), Simulated Annealing (SA), and Particle Swarm Optimization (PSO) have been widely adopted to overcome these limitations. GAs are particularly well-suited for mixed-integer optimization problems thanks to its ability to explore large and nonlinear search spaces efficiently [27].

In recent years, hybrid models combining GAs with SO or RO have emerged. Saeedi et al. [28] developed a two-stage stochastic programming model for a closed-loop Electric Vehicle (EV) battery supply chain under demand uncertainty. They used meta-heuristic algorithms to optimize facility location and inventory decisions, encoding key variables into chromosomes.

By integrating a GA with WDRO, this work combines the global search strengths of meta-heuristics with the rigor of modern DRO. The incorporation of the cJKO scheme

enables the dynamic evolution of Wasserstein ambiguity sets alongside decision updates, resulting in adaptive and resilient solutions under demand uncertainty.

2.7. Optimization of Probability Distributions

The standard JKO scheme is a variational formulation originally developed to solve gradient flows in the space of probability measures [7]. Recent research has adapted this scheme to solve optimization problems over distributions, notably in Bayesian optimization and physics-informed learning [29]. The constrained JKO (cJKO) variant enforces an upper bound on the Wasserstein distance between iterations, offering a principled way to control distributional shifts and update empirical distributions while preserving proximity to the observed data [30]. Although applications of cJKO in supply chain optimization are nascent, its potential is significant. Our work is among the first to apply the cJKO scheme to distributionally robust facility location under chance constraints.

3. Problem Statement

In supply chain management, uncertainty plays a crucial role in decision-making. Classical deterministic models assume that the demand is perfectly known, but real-world scenarios involve an uncertain demand due to market fluctuations, economic factors, and unpredictable customer behavior. To tackle demand uncertainty, we adopt a WDRO framework, that is, DRO with a Wasserstein-based ambiguity set. The goal is to minimize supply chain costs while ensuring robust facility, storage, and allocation decisions. Table 2 presents the notations used throughout this paper, defining the indices, parameters, and decision variables essential for formulating the optimization model.

Table 2. Notations.

Notation	Description
I	Index for customers (demand points) $i = 1, \ldots, I$
J	Index for facilities (potential locations) $j = 1, \ldots, J$
T	Index for time periods $t = 1, \ldots, T$
f_{jt}	Facility opening cost for facility j in period t
a_{jt}	Storage cost at facility j in period t
c_{ij}	Transportation cost from facility j to customer i
q_t	Total supply capacity available in period t
d_i	Demand at customer i
ρ_i	Penalty cost for unmet demand at customer i
X_{jt}	Binary variable indicating whether facility j is open in period t
S_{jt}	Inventory level at facility j in period t
Y_{jit}	Quantity transported from facility j to customer i in period t

Problem Formulation

In the supply chain optimization model, the primary decisions to be made include (1) facility location decisions, which determine whether facility j should be opened in period t, represented by the binary variable X_{jt}; (2) inventory management decisions, where the inventory level at facility j in period t, denoted by S_{jt}, is optimized to adjust storage levels in response to fluctuating demand while minimizing holding costs; and (3) transportation and allocation decisions, where the quantity Y_{jit} is transported from facility j to customer i in period t, aiming to minimize transportation costs while fulfilling customer demand d_i. Without loss of generality, the demand of each customer is considered fixed over time (i.e., it does not depend on t).

Decisions must be taken under demand uncertainty and are designed to minimize total supply chain costs while maintaining operational efficiency. The objective function of the reference problem consists in minimizing the total cost associated with facility location and allocation under demand uncertainty. It consists of two terms:

(i) **Facility operation costs**, represented by fixed opening costs $f_{jt}X_{jt}$ and inventory holding costs $a_{jt}(S_{jt} - S_{j,t-1})$, ensure that storage levels adjust optimally over time. This term is "deterministic" in the sense that its computation does not depend directly on the demand and its uncertainty.

(ii) **Transportation and allocation costs** are captured through expected costs over the unknown probability distribution P, including demand-dependent shipping costs $c_{ij}Y_{jit}d_i$ and penalties related to supply shortages. This term depends directly on the demand and its uncertainty; indeed, it is a penalization with respect to unmatched demand due to fluctuations. Its computation requires finding the worst-demand possible P, meaning to solve an optimization problem over probability distributions, specifically a Wasserstein neighborhood of the historical data.

$$\min \sum_{t\in[T]} \sum_{j\in[J]} \left(f_{jt}X_{jt} + a_{jt}\left(S_{jt} - S_{j,t-1}\right)\right) + \sup_{P\in\mathscr{P}_\epsilon} \mathrm{E}_{d\sim P}\left[\sum_{j,i,t} C_{jit}Y_{jit}d_i + \sum_{j,i,t} \rho_i \max\left(0, Y_{jit}d_i - S_{jt}\right)\right] \quad (4)$$

The optimization of the objective function (4) is subject to the following constraints:

Chance Constraint for Demand Satisfaction

$$\inf_{d\sim P} P\left\{\max_{j\in[J],t\in[T]}\left\{Y_{jt}^T d - S_{jt}\right\} \leq 0\right\} \geq 1 - \eta \quad (5)$$

Facility and Allocation Constraints

$$X_{jt} \in \{0,1\}, S_{jt} \in F^+, \quad Y_{jit} \geq 0, \quad \forall i \in [I], j \in [J], t \in [T] \quad (6)$$

Allocation Constraint

$$\alpha \leq \sum_{t\in[T]} \sum_{j\in[J]} Y_{jit} \leq 1, \forall i \in [I] \quad (7)$$

Facility and Capacity Constraints

$$\sum_{\tau=1}^{I} X_{j\tau} \leq 1, S_{jt} \leq q_t \sum_{\tau=1}^{I} X_{j\tau}, S_{jt} \geq S_{j,t-1}, \qquad \forall\, j \in [J],\ t \in [T] \quad (8)$$

The chance constraint for demand satisfaction (5) guarantees that the probability of supply shortages remains below a predefined threshold $(1 - \eta)$, improving reliability. Analogously to the objective function, the computation of (5) requires solving an optimization problem over probability distributions, meaning that cJKO will be used to compute the left-hand side term of the inequality in (5).

Facility operation constraints (6) enforce binary decisions for facility openings and nonnegative allocations. The allocation constraint (7) ensures that each demand node receives an appropriate share of the supply without exceeding the available capacity. Finally, capacity constraints (8) limit inventory levels based on facility availability, ensuring that stored goods do not exceed practical limits. These constraints collectively shape an efficient, resilient, and mathematically rigorous facility location model that effectively manages uncertainty and distributional shifts in demand.

Another important constraint is related to cJKO itself, rather the supply chain problem. Specifically, we must set $W_2^2(P, P_k) \leq \varepsilon$ at each iteration $k = 1, \ldots, K$ of the cJKO algorithm, where P_k is the probability distribution at step k, P is the decision variable (i.e., probability

distribution), and ε is the radius—in Wasserstein terms—of the ambiguity set (aka Wasserstein ball) around the current candidate solution P_k. All the details about cJKO and ε are detailed in the next section.

4. Solution Approach

4.1. Optimizing over Probability Distributions

The computation of the objective function (4) and the probabilistic constraint (5) requires solving two optimization problems over distributions separately, that is, searching for the worst demand distribution affecting the objective value and the worst demand distribution affecting the demand satisfaction constraint.

Without loss of generality, we can consider the reference problem

$$P^* \in \underset{P \in \mathscr{P}}{\operatorname{argmin}} \mathscr{F}(P), \tag{9}$$

where $\mathscr{F}(P)$ refers to a generic functional to be minimized over the space of probability measures $\mathscr{P}$. Solving problem (9) means searching for a sequence of probability distributions starting from an initial random guess P_0 (typically a multivariate Gaussian) and converging to P^*. A well-known example for problem (9) is given in [7], showing that solving the Fokker–Planck PDE is equivalent to minimize an entropy functional over the space of probability measures, equipped with the so-called Wasserstein distance.

The Wasserstein distance is defined as the minimum cost for transporting probability density mass from a source probability density to match a target probability density, and where the cost is computed in terms of a distance (also known as the *ground metric*) between points of the support. In this paper, we consider the case that the cost is the Euclidean distance, leading to the so-called 2-Wasserstein distance:

$$\mathscr{W}_2^2(P_0, P^*) = \min_{T_\# P^0 = P^*} \int_{\mathscr{D}} \|T(d) - d\|_2^2 \, dP(d) \tag{10}$$

where $T_\# P_0 = P^*$ means that applying $T : \mathscr{D} \to \mathscr{D}$ to P_0 (i.e., our source distribution) will obtain P^* (i.e., our target distribution) as result. $T_\#$ denotes the so-called push-forward operator; while $T(d)$ is interpreted as a function moving a single data point (i.e., a customers demand scenario, in our case) over the support $\mathscr{D}$, $T_\#$ represents its extension to an entire probability measure.

The Brenier Theorem [3,4] guarantees that the problem (10) has a unique optimal solution T^*, named the optimal transport map. Furthermore, the Brenier theorem also states that the optimal transport is equal to the gradient of a convex function $\varphi(d)$, such that $T^*(d) = \nabla\varphi(d) = d - \nabla f(d)$.

By introducing a coefficient $\tau \in [0, 1]$, we can define a *continuous-time* representation of the optimal transport map T^*, that is,

$$T_\tau^*(d) = d - \tau \nabla f(d) \tag{11}$$

Easily, the following discrete-time representation can be derived:

$$T_k^*(d) = d - \frac{k}{K} \nabla f(d) \tag{12}$$

with $k = 0, \ldots, K$ and where k/K is a discretization of τ. As a result, the sequence $P_0 = T_{0\#}^* P_0 \to \ldots \to T_{K\#}^* P_0 = P^*$ solves (10), with T^* as the solution of (10).

Since P^* is not known a priori, T^* cannot be directly computed. A widely adopted method to solve (10) is the JKO (Jordan–Kinderlehrer–Otto) scheme [7] that is a proximal

point method with respect to the 2-Wasserstein distance or, from another point of view, a backward Euler discretization. At a generic iteration k, JKO transports the current probability density, P_k, into the next P_{k+1} according to

$$P_{k+1} = \underset{P \in \mathscr{P}}{argmin} \frac{1}{2h} \mathscr{W}_2^2(P, P_k) + \mathscr{F}(P) \tag{13}$$

with h as the JKO's step size.

Solving problem (14) translates into searching for the next probability density P_{k+1} that minimizes the functional $\mathscr{F}$ within a 2-Wasserstein neighbourhood of the current P_k, that is, a Wasserstein ball (aka ambiguity set) around P_k. Iteratively solving problem (13) leads to a sequence $P_0, \ldots, P_K \simeq P^*$ whose rate of convergence to P^* depends on h: with $h \to \infty$, the first term goes to zero, meaning that we are minimizing $\mathscr{F}$ but at the expense of obtaining a transport far away from the optimal one. On the contrary, if $h \to 0$, then the first term becomes more relevant so that the generated sequence will be close to T^* but it will converge significantly slowly to P^*.

Finally, problem (9) is formulated in terms of $P \in \mathscr{P}$, that is, having probability densities as decision variables. State-of-the-art approaches recast problem (9) as an optimization problem in terms of a parametrized transport map from P_k to P^*. For instance, the convex function $\varphi(d)$ is usually approximated through an Input Convex Neural Network (ICNN) [31–33], a specific type of Deep Neural Network (DNN), and then $T(d)$ is obtained by following the Brenier Theorem, namely, $T(d) = \nabla\varphi(d)$. DNNs are also used in [34,35] to approximate $f(d)$ through a DNN, whose output is then used as an input—along with d—of another DNN estimating $T(d)$. Contrary to the ICNN-based approaches, this method allows for skipping the computation of $\nabla\varphi(d)$. More recently, [36] proposed to parametrize $T(d)$ through a Residual Neural Network (ResNet) and a variational formulation of the functional $\mathscr{F}(d)$ formulated as maximization over a parametric class of functions. This allows for reducing the computational burden with respect to the previous methods using small data samples and scaling well with the dimensionality of the support $\mathscr{D}$. The small data regime setting has been also recently investigated in [37], who proposed to combine OT solvers and Gaussian process regression to efficiently learn the transportation map. The most relevant issue for all the neural approaches is that at least one neural network must be trained, leading to relevant computational costs at each JKO step. Since getting close to $T^*(d)$ requires $h \to 0$ and consequently a number of iterations $K \to \infty$, this issue becomes even more critical. Moreover, there are not specific guidelines on how to choose the most suitable neural network to use. Moreover, [38] has recently reported that normalizing flow models based on neural networks can only generate planar flows, which are proven to be expressive only in the case of univariate probability densities (i.e., 1-dimensional support). Another relevant study on the convergence and the self-consistency of normalizing flows in the space of probability density equipped with the Wasserstein distance is given in [39].

The Constrained JKO Schema

The parametrization of the transport plan considered in this paper has been originally proposed in [30], and is defined as follows:

$$\widetilde{T}_d = d - \lambda(d) v R_{\alpha(d)} \tag{14}$$

with $\lambda : \mathbb{R}^d \to \mathbb{R}_0^+, \alpha : \mathbb{R}^d \to [0, 2\pi]^{d-1}$, and $R_{\alpha(d)}$ is a rotation matrix consisting of a sequence of matrix multiplications, $R_{\alpha(d)} = \Pi_{\ell=1}^{\mathrm{d}-1} R^{[\ell]}(d)$, where $R^{[\ell]}(d)$ is defined as

$$R_{i,j}^{[\ell]}(d) = \begin{cases} cos\Big(\alpha_{[\ell]}(d)\Big) & if\ (i,j) = (\ell,\ell) \\ -sin\Big(\alpha_{[\ell]}(d)\Big) & if\ (i,j) = (\ell,\kappa) \\ sin\Big(\alpha_{[\ell]}(d)\Big) & if\ (i,j) = (\kappa,\ \ell) \\ cos\Big(\alpha_{[\ell]}(d)\Big) & if\ (i,j) = (\kappa,\kappa) \\ 1 & if\ i = j\ \wedge\ i \neq \ell,\ \kappa \\ 0 & otherwise \end{cases} \tag{15}$$

where $\alpha_{[\ell]}(d)$ denotes the ℓth component of the vector-valued function $\alpha(d)$.

Finally, v is a reference versor; for simplicity, we consider $v = \left(\frac{1}{\sqrt{m}}\right)1^m$, where 1^m denotes the all-ones vector and, therefore, $\parallel\ v\ \parallel = 1$.

From the Brenier theorem, we know that $T\ (d) = d - \nabla f\ (d)$, and so we are posing $\nabla f\ (d)\ =\ \lambda(d) v R_{\alpha(d)}$. The parameters to be learned in the proposed parametrization are the two functions $\lambda(d)$ and $\alpha(d)$. Contrary to the ICNN-based parametrizations, but analogously to [34–36], the proposed parametrization does not require any assumptions on the two functions to be learned.

According to the proposed parametrization, the JKO schema can be recast into

$$\lambda_k(d), \alpha_k(d) = argmin\left\{ \frac{1}{2h}\ W_2^2(P_{k+1}, P_k) + \mathcal{F}(P_{k+1}) \right\} \tag{16}$$

with $P_{k+1} = \widetilde{T}_{k-1\#} P_k$.

Like all the other approaches, we learn the functions underlying our parametrization by accessing to point clouds (i.e., empirical distributions) sampled from probability densities $P_0, \ldots, P_k$ [30]. Analogously to other neural networks-based approaches, denote with $D_k \sim\ P_k$ the point cloud at iteration k obtained as

$$D_k = \widetilde{T}_{k-1\#} \widetilde{T}_{k-2\#}\ \ \ldots_{\#}\ \widetilde{T}_{0\#} D_0 = D_0 \tag{17}$$

with $D_0 \sim P_0$ such that $D_0 = \left\{ d_0^{(i)} \right\}_{i=1}^N$. According to our parametrized transport $\widetilde{T}(d)$, every point of the current cloud D_k is transported as follows:

$$d_{k+1}^{(i)} = d_k^{(i)} - \lambda_k^{(i)}\ v R_{\alpha_k^{(i)}}\ ,\ \forall i\ \in\ \{1, \ldots, N\ \} \tag{18}$$

with $\lambda_k^{(i)}$ and $\alpha_k^{(i)}$ representing shorthand for $\lambda\left(d_k^{(i)}\right)$ and $\alpha\left(d_k^{(i)}\right)$.

Since we are working with point clouds, and according to our parametrization, we can rewrite the 2-Wasserstein distance as follows:

$$\mathcal{W}_2^2(D_{k+1}, D_k) = \frac{1}{N}\sum_{i=1}^N \left\| d_k^{(i)} - \lambda_k^{(i)} v R_{\alpha_k^{(i)}} - d_k^{(i)} \right\|_2^2 = \frac{1}{N}\sum_{i=1}^N \sqrt{\left|\lambda_k^{(i)}\right|} \left\| v R_{\alpha_k^{(i)}} \right\|_2^2 \tag{19}$$

Since rotation does not modify the module of the rotated v, and according to our choice for v, we have $\left\| v R_{\alpha_k^{(i)}} \right\|_2^2 = 1$. Finally, we can write

$$\mathcal{W}_2^2(D_{k+1}, D_k) = \frac{1}{N}\sum_{i=1}^N \sqrt{\lambda_k^{(i)}} \tag{20}$$

with $\lambda_k^{(i)} > 0$.

This allows us to set a threshold ε on the term $\mathscr{W}_2^2(D_{k+1}, D_k)$ to explicitly quantify the Wasserstein ball around the current D_k, instead of weighting $\mathscr{W}_2^2(D_{k+1}, D_k)$ with respect to $\mathscr{F}(d)$ through the value of h in the standard JKO schema.

Finally, the constrained-JKO (cJKO) scheme can be formalized:

$$\lambda_k, \alpha_k = \underset{\substack{\lambda \in \mathbb{R}_0^{+N}, \\ \alpha \in [0, 2\pi]^N}}{argmin} \; \mathscr{F}(P_{k+1}) s.t. \qquad \frac{1}{N}\sum_{i=1}^{N} \sqrt{\lambda_k^{(i)}} \leq \varepsilon \tag{21}$$

with ε as a small positive quantity, $\lambda_k = \left(\lambda_k^{(1)}, \ldots, \lambda_k^{(N)}\right)$, $\alpha_k = \left(\alpha_k^{(1)}, \ldots, \alpha_k^{(N)}\right)$, and $D_{k+1} = \left\{ d_{k+1}^{(i)} = d_k^{(i)} - \lambda_k^{(i)} v R_{\alpha_k^{(i)}} \right\}_{i=:N}$.

In simpler terms, we are searching for $D_{k+1} \sim P_{k+1}$, which minimizes $\mathscr{F}(P_{k+1})$ within a 2-Wasserstein neighbourhood of D_k.

Indeed, $\varepsilon \to 0$ in cJKO operates analogously to $h \to 0$ in JKO, requiring a longer sequence to converge to P^*, but providing a transport close to the optimal one. However, it is important to clarify that, from a quantitative perspective, they are completely different. Since h is just a scalarization weight into a scalarized bi-objective problem, it is quite impossible to establish a suitable value without any prior knowledge about $\mathscr{F}(P)$. Furthermore, using a constant value for h, which is the common choice, leads to a significantly different relevance of the two objectives over the JKO iteration

The cJKO scheme overcomes these limitations thanks to ε: it is chosen in advance and kept fixed along the overall iterative optimization process, independently on the values of $\mathscr{F}(P)$ over the iterations. Figure 1 shows an example of the trajectories from D_0 to D_K depending on three different values of ε. The three presented cases converge to the same value of $\mathscr{F}(D_K)$ (and same final point cloud) but within a different number of cJKO steps: the smaller the value of ε, the larger the number of iterations K, and the closer the final transport to the optimal one (i.e., trajectories are closer to the actual T^*, because they are more straight and do not overlap).

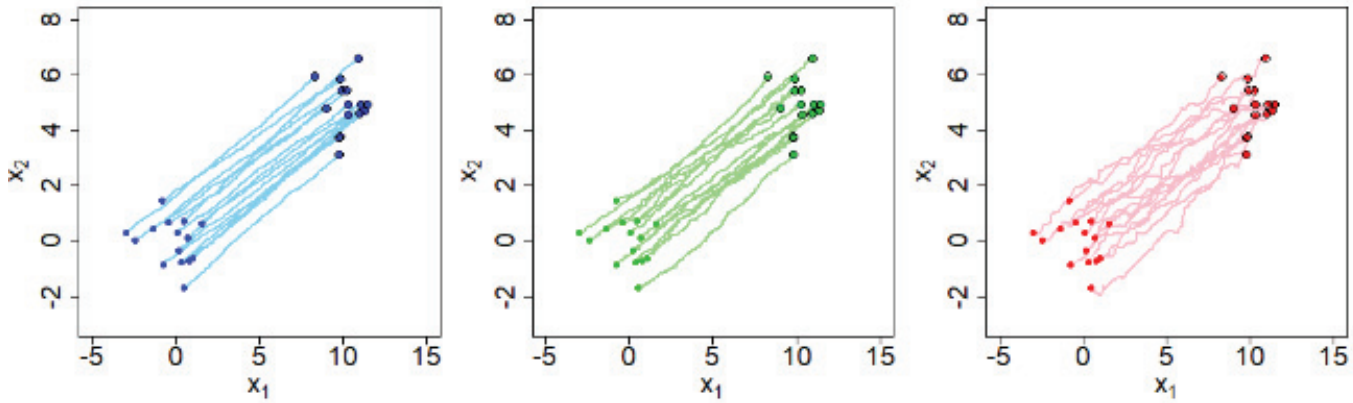


Figure 1. Effects of ε on cJKO convergence and Wasserstein flow behavior [29].

As far as the application problem considered in this paper is concerned, cJKO is used to compute the second (stochastic) term of the objective function (4) and the chance constraint (5) separately, given a candidate solution for the facility location problem.

For completeness, we report the generic algorithm for the cJKO scheme, as follows. It is important to clarify that it is separately applied to identify the demand distributions—within an ε-radius Wasserstein ball of the historical demand data D_0—providing (a) the

maximum penalization in the objective function, that is, the second term in (4), and (b) the minimum probability in the chance constraint (5).

cJKO Algorithm for distributionally robust optimization

- Input:
  - [X, S, Y] candidate solution (provided by the GA meta-heuristic);
  - D_0 historical demand data;
  - $\varepsilon > 0$ (i.e., cJKO's parameter);
  - $K \geq 1$ (i.e., cJKO's iterations);
  - $\mathscr{F}$ functional to be optimized (i.e., objective function's term to be minimized in (4) and left-hand side term of the chance constraint to be maximized in (5), separately);
- Step 0: $k \leftarrow 0$;
- Step 1: solve the following problem;

$$D_{k+1} \in \begin{array}{c} \text{optimize} \\ \lambda \in \mathbb{R}^N_{\geq 0} \\ \alpha \in [-\pi,\ \pi]^{N\times(d-1)} \\ s.t. \end{array} \begin{array}{l} \mathscr{F}(T_{\lambda,\alpha\ \#}D_k; X, S, Y) \\ \\ \\ \mathscr{F}(T_{\lambda,\alpha\ \#}D_k; X, S, Y) \end{array}$$

- Step 2: $k \leftarrow k + 1$;
- Step 3: if $k < K$, go to Step 1;
- Return D_k.

Since the true probability distribution of the demand is unknown, as well as changes due to unpredictable events, it is difficult to define suitable values for ε and K a priori. On the other way round, their values are easy to interpret: larger values increase robustness against events that could lead to demand values that are significantly different from the historical ones, but this means that the final decision $[X, S, Y]$ might be too conservative. On the contrary, smallest values assume that the future demand should not change too much with respect to historical data, leading to less conservative decisions but difficulty in dealing with possible significant changes. The suggestion is to perform different runs with different values of ε and K, with the aim to observe differences both in term of the final solution $[X, S, Y]$ and generated demand distributions. Finally, the user can select the most suitable decisions (and scenario) for the specific goals of the target setting.

4.2. Metaheuristic-Based Approach

In this section, we introduce a metaheuristic optimization framework for distributionally robust supply chain design under demand uncertainty. Our approach integrates

- A Genetic Algorithm (GA) for the heuristic initialization of facility location, storage levels, and allocation decisions;
- The cJKO scheme to compute the stochastic term of the objective function (4) and the chance constraint (5) separately;
- A robust evaluation function that balances operational cost and risk-averse decision-making.

This hybrid method provides an efficient alternative to exact methods, particularly in high-dimensional settings, in which classical optimization techniques struggle with combinatorial complexity.

4.2.1. Genetic Algorithm for Heuristic Initialization

In step ①. to enhance the efficiency of the optimization process, we employ a GA for heuristic initialization. A GA is an evolutionary-based meta-heuristic that iteratively improves candidate solutions by mimicking natural selection principles, including selection, crossover, and mutation [40]. In our approach, the GA generates an initial population of facility location and allocation decisions, ensuring diversity in solutions while providing a high-quality starting point for subsequent optimization. By leveraging a GA, we improve the convergence speed and enhance the feasibility of the optimization model by reducing the likelihood of poor initializations. This approach is particularly beneficial for large-scale problems where a purely random initialization could lead to suboptimal or infeasible solutions.

Chromosome Representation

Every individual in the genetic population represents a possible supply chain configuration as a chromosome, encoding

- Facility location decisions $\boldsymbol{X}$ ($\boldsymbol{J} \times \boldsymbol{T}$ binary matrix);
- Storage level decisions $\boldsymbol{S}$ ($\boldsymbol{J} \times \boldsymbol{T}$ continuous matrix);
- Allocation decisions $\boldsymbol{Y}$ ($\boldsymbol{J} \times \boldsymbol{I} \times \boldsymbol{T}$ continuous matrix).

Formally, a chromosome is structured as chromosome = $[X, S, Y]$, where

- $\boldsymbol{X_{jt}} \in \{\mathbf{0}, \mathbf{1}\}$ denotes whether facility j is open at time t;
- $\boldsymbol{S_{jt}} \in \mathbb{R}^+$ represents the storage level at facility j at time t;
- $\boldsymbol{Y_{jit}} \in \mathbb{R}^+$ captures how much demand from the customer i is allocated from facility j at time t.

Fitness Function

The objective function in the genetic algorithm evaluates the total cost of a candidate supply chain configuration with Formula (4). It minimizes the sum of three key components: (*i*) fixed facility opening costs, (*ii*) inventory holding costs, and (*iii*) expected transportation costs and shortage penalties costs under stochastic demand while satisfying demand constraints.

Genetic Operators

The GA employs three main evolutionary operators:

- Selection (Tournament Selection);
  - Randomly selects k-tournament competitors and chooses the best candidate.
- Crossover (Uniform Crossover);
  - Swaps facility decisions and allocations between parent solutions to create diverse offspring. In this strategy, each gene (i.e., element of the chromosome representing facility decisions or allocations) is independently chosen from one of the two parent solutions with equal probability. This encourages greater diversity and exploration of the solution space by recombining building blocks from both parents.
- Mutation (Storage and Allocation Adjustment);
  - Perturbs storage levels and allocations to introduce new feasible solutions. The swap mutation randomly selects two positions in the chromosome and exchanges their values. This perturbation helps the algorithm escape local optima by introducing structural variation into the solution while preserving feasibility.
- The GA runs for a predefined number of generations, yielding a near-optimal initial solution.

4.2.2. Wasserstein Distributionally Robustness via cJKO

Once an initial solution is generated in step ②, we use cJKO to compute the stochastic term of the objective function (4) and the chance constraint (5) separately.

Distributional Robustness via Wasserstein Distance. Starting from the point cloud associated with to historical demand, namely, D_0, we iteratively search for D_k to finally compute right term of the chance constraint (5).

cJKO Iterative Update Rule. At each iteration k, we solve the constrained Wasserstein optimization problem in Formulas (4) and (5). The process is repeated until a maximum number of iterations is reached. This ensures that our demand forecasts adapt dynamically based on historical and simulated uncertainty data.

4.2.3. Chance Constraint Verification ③

In supply chain and facility location problems, ensuring demand satisfaction under uncertainty is critical for maintaining service levels and system robustness. Traditional chance constraints provide a probabilistic guarantee that demand will be met with high probability, despite stochastic fluctuations. In our approach, the chance constraint is reformulated within the cJKO variational framework to account for distributional uncertainty. Specifically, we enforce the constraint (5). This formulation ensures that demand satisfaction holds with probability at least $1 - \eta$ under the worst-case distribution within the Wasserstein uncertainty set. Using a cJKO-based gradient flow, we iteratively evolve the distribution P_k to maximize constraint violation while remaining within $\mathscr{P}_\epsilon$. If the worst-case distribution still satisfies the demand constraint, robustness is achieved. This approach allows us to dynamically verify and enforce probabilistic demand satisfaction in a data-driven and distributionally robust manner.

4.2.4. Integrated Iterative Optimization Framework

The GA-initialized solution and cJKO-updated demand are combined in a metaheuristic-based iterative optimization loop.

Iterative Algorithm

Step 1: Initialize supply chain decisions using GA.
Step 2: Solve the cJKO Wasserstein update to refine demand scenarios.
Step 3: Evaluate feasibility via chance constraints. Step ④.
Step 4: Adjust storage and allocations based on updated demands.
Step 5: Repeat until convergence or max iterations. Step ⑤.

In this case, η is the risk level (e.g., 95% confidence).

If the constraint is violated, we adjust the ambiguity radius ε and re-optimize. The optimization process follows the iterative procedure illustrated in Figure 2. This framework begins with the initialization of supply chain decisions using GA, followed by the cJKO Wasserstein update to refine the demand distributions. Chance constraint verification ensures robustness, and necessary adjustments to storage and allocations are made until convergence or the maximum iteration limit is reached.

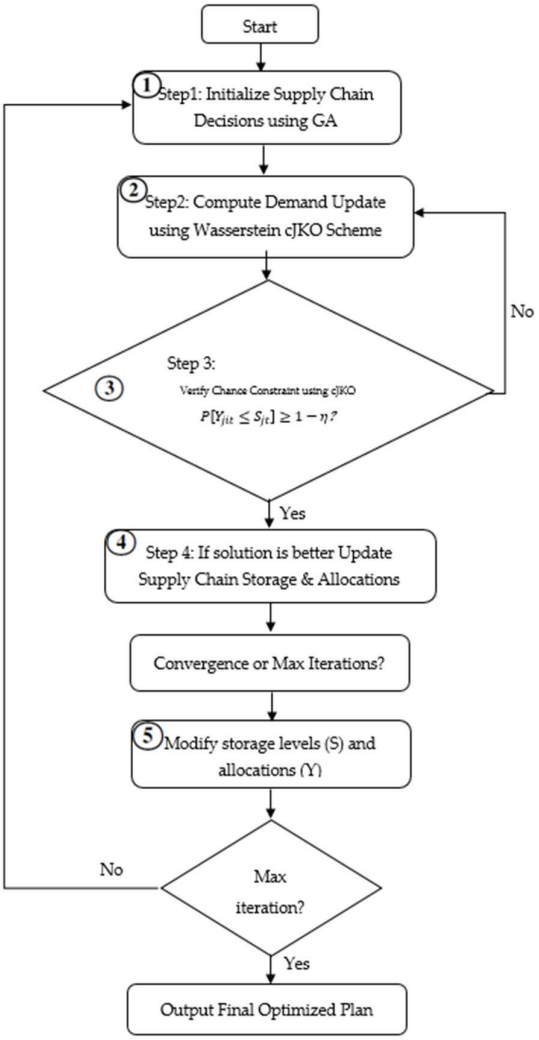


Figure 2. Iterative metaheuristic optimization framework using cJKO for a distributionally robust supply chain design.

5. Computational Results

5.1. Experimental Setup

The optimization was performed on a machine with an Intel i7-9700 CPU and 16 GB of RAM using Python 3.12 (NumPy, SciPy, DEAP for genetic algorithms, and Matplotlib 3.10.0 for visualization) [41–44]. The inner optimization problem in the cJKO scheme was solved using COBYLA (Constrained Optimization BY Linear Approximation) from the SciPy library to efficiently handle the Wasserstein-constrained subproblem. The code has been developed in Python and is freely available at the following repository: https://github.com/iman-ie/FacilityLocation_cJKO.git (accessed on 1 May 2025).

To evaluate the performance of the proposed stochastic facility location model under Wasserstein ambiguity and chance constraints, we conduct a comprehensive series of computational experiments using synthetically generated data. The testbed consists of six problem instances with varying scales, including configurations of 15, 30, and 40 customers and candidate facilities, tested over planning horizons of both three and five time periods to assess scalability and temporal complexity handling. The parameter settings follow data generation strategies adapted from well-established studies in the literature on robust and stochastic logistics [19,45]. The synthetic data are generated to simulate a wide range of realistic operating environments with uncertainty in demand and cost. Parameter values are drawn from uniform distributions to ensure heterogeneity across instances.

Furthermore, the model is tested under varying values of the Wasserstein radius ε, allowing a detailed sensitivity analysis on the trade-off between robustness and conservatism. The experimental parameters are detailed in Table 3.

Table 3. Experimental setup parameters.

Parameter	Description	Value/Range
Facility opening cost f	Cost of opening facilities per period	$U[100(T-t),\ 100+100(T-t)]$
Capacity level q	Capacity level per period	$20t$
Storage cost a	Cost of holding inventory at facilities	$U\ [0,2]$
Transportation cost c	Cost of transporting goods from facilities to customers	$U\ [0,5]$
Demand weight d	Weight assigned to demand at each customer location	$U\ [0,30]$
Penalty cost ρ	Penalty for unmet demand	$U\ [5,15]$
Reliability level η	Minimum probability of satisfying demand constraints	0.8
α	Allocation threshold for each customer	0.8

5.2. Parameter Setting

The effectiveness of metaheuristic algorithms is significantly impacted by their parameter settings. Consequently, this section focuses on optimizing the parameters of the GA to enhance the robustness of the solution strategy. While traditional studies have often used a full factorial design for parameter selection, this method becomes less practical as the number of parameters increases. To address this issue, the Taguchi method is adopted to streamline the experimental process by reducing the number of required tests and overall complexity [46]. Initially, the test problem is executed ten times. The outcomes of the objective functions are then converted into relative percentage deviation (*RPD*) values to standardize performance comparison. The average *RPD* is subsequently utilized to compute signal-to-noise (S/N) ratios, which help identify the most effective parameter levels. Table 4 outlines the selected parameters and their respective levels.

Based on the defined parameters and their corresponding levels, the L18 orthogonal array from the Taguchi method is applied to the GA. To assess the outcomes of each experimental run, the *RPD* is calculated using Equation (23):

$$RPD = \frac{Alg_{sol} - Min_{sol}}{Min_{sol}} \tag{22}$$

In this context, Min_{sol} denotes the lowest observed value of the cost function, while Alg_{sol} represents the solution produced by the algorithm [47,48]. Table 5 presents the L18 orthogonal array used in the experiments, along with the average *RPD* calculated over ten independent runs.

Taguchi's method aims to enhance the influences of controllable factors while reducing the effects of noise variables. The S/N ratio serves as a key metric to achieve both objectives. This approach is categorized into three types: "larger is better", "smaller is better", and "nominal is best". In the current study, the *RPD* is employed as the response

variable. Therefore, the "smaller is better" criterion is selected for parameter tuning, and Equation (23) is used to compute the corresponding S/N ratio.

$$S/N = -10\, log_{10}\left(\sum\left(Y^2\right)/n\right) \tag{23}$$

Table 4. The parameters level of GA.

Factor	Level	Symbol	Type
MaxIt (Maximum number of iterations)	A(1)-100 A(2)-200 A(3)-300	A	Numeric
nPop (Population size)	B(1)-80 B(2)-120 B(3)-160	B	Numeric
PC (Probability of crossover)	C(1)-0.3 C(2)-0.6 C(3)-0.9	C	Numeric
PM (Probability of mutation)	D(1)-0.05 D(2)-0.15 D(3)-0.25	D	Numeric
Type of crossover	E(1)-One-point E(2)-Two-point E(3)-Uniform	E	Categorical
Type of mutation	F(1)-Pairwise F(2)-Swap F(3)-Inversion	F	Categorical

In this context, Y represents the response value for each test instance, and n denotes the total number of experiments based on the orthogonal array. The analysis of the response (*RPD*) was conducted using Minitab 21.1.0 software. Figure 3 illustrates the average S/N ratios corresponding to each parameter level. Based on the results, the optimal levels for the GA parameters are identified as three, two, one, three, one, and two. Table 6 summarizes the best values for each parameter.

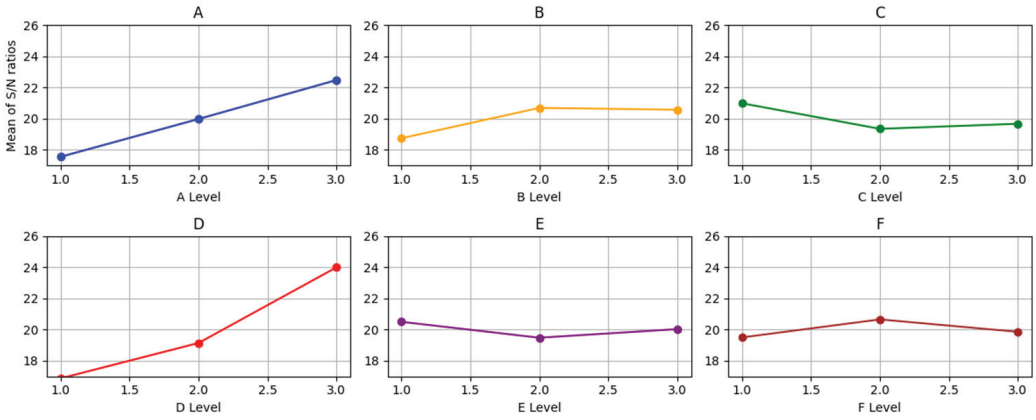


Figure 3. The GA factors' mean SN ratio plot: each chart refers to a specific parameter in Table 5.

Table 5. The orthogonal array L18 for the GA.

Trial	A	B	C	D	E	F	Mean of RPD	S/N
1	1	1	1	1	1	1	0.202	13.89
2	1	2	2	2	2	2	0.145	16.77
3	1	3	3	3	3	3	0.078	22.15
4	2	1	1	2	2	3	0.114	18.85
5	2	2	2	3	3	1	0.067	23.47
6	2	3	3	1	1	2	0.121	18.35
7	3	1	2	1	3	2	0.133	17.52
8	3	2	3	2	1	3	0.089	21.01
9	3	3	1	3	2	1	0.051	25.85
10	1	1	3	3	2	2	0.103	19.74
11	1	2	1	1	3	3	0.162	15.80
12	1	3	2	2	1	1	0.141	16.99
13	2	1	2	3	1	3	0.079	22.04
14	2	2	3	1	2	1	0.152	16.37
15	2	3	1	2	3	2	0.091	20.82
16	3	1	3	2	3	1	0.095	20.44
17	3	2	1	3	1	2	0.029	30.75
18	3	3	2	1	2	3	0.108	19.32

Table 6. Optimal GA parameter settings based on the Taguchi S/N ratio analysis.

Parameters	Symbol	Best Level
Maximum number of iterations	A(3)	120
Population size	B(2)	300
Probability of crossover	C(3)	0.25
Probability of mutation	D(3)	0.9
Type of crossover	E(3)	Uniform
Type of mutation	F(2)	Swap

5.3. Sensitivity Analysis of the Wasserstein Ambiguity Radius

To evaluate the sensitivity of the proposed distributionally robust facility location model with respect to the Wasserstein ambiguity radius ε, we solve the problem under a range of values. Specifically, we considered $\varepsilon \in \{0.025, 0.05, 0.1, 0.5\}$. Each instance is solved using the cJKO-based optimization framework, and the corresponding total expected cost (objective value) is computed by aggregating facility setup costs, inventory holding costs, and allocation-based penalties under the resulting worst-case demand distribution. The results relative to the smallest test case (i.e., I = 15, J = 15, and T = 3), which are summarized in Table 7 and illustrated in Figure 4, reveal a non-monotonic trend in the objective value as a function of ε. Initially, increasing the radius leads to a reduction in the objective, indicating improved robustness to distributional shifts without overcompensating in the decision variables. However, beyond a critical value of ε, the objective begins to increase, suggesting excessive conservatism as the ambiguity set becomes too large. Figure 5 compares the prescribed Wasserstein radius (ε) with the actual Wasserstein distance achieved after optimization. The observed near-monotonic trend highlights how increasing ε allows the model to explore a broader set of distributions, thereby enabling more robust—but potentially more conservative—solutions. This behavior confirms that the optimization process effectively leverages the flexibility provided by the ambiguity set while maintaining alignment with the WDRO framework.

Table 7. Costs for different values of the cJKO parameter ε for the case I = 15, J = 15, and T = 3.

ε	Objective Value	Wasserstein Distance
0.025	1775.907	2.6890
0.05	1757.1674	4.1560
0.1	2180.4731	9.1143
0.5	2314.4775	49.6896

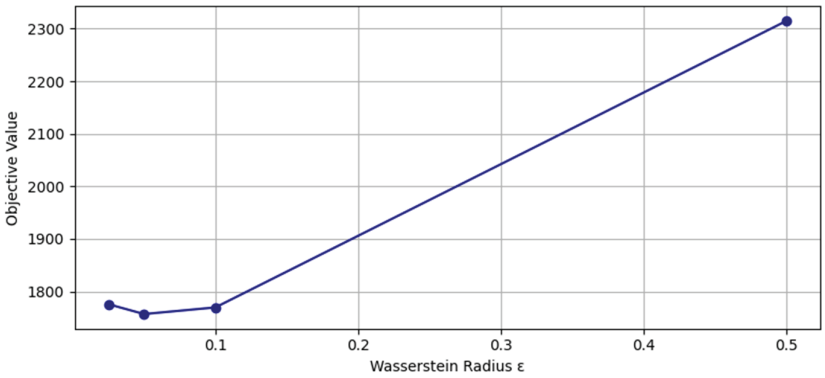


Figure 4. Effect of the Wasserstein ambiguity radius on the objective value (for the case I = 15, J = 15, and T = 3).

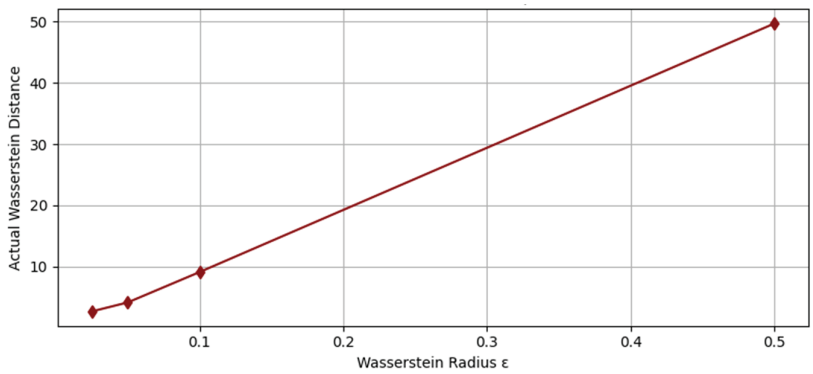


Figure 5. Actual Wasserstein distance achieved vs. specified radius ε (for the case I = 15, J = 15, and T = 3).

Figure 6 further details the behaviors of both the first-stage cost (on the left) and the overall cost (on the right) with respect to the value of ε and for 10 independent runs. Obviously, most of the total cost is given by the second-stage cost rather than the first-stage cost. While the first-stage cost monotonically increases with ε, the overall cost increases non-monotonically. Finally, the variability of the total costs significantly increases with ε increasing from 0.1 to 0.5 due to the fact that the demand scenarios generated are significantly different from the historical data.

The same behavior is also observed for the largest test case, that is, I = 40, J = 40, and T = 5, as reported in Table 8.

Table 8. Costs for different values of the cJKO parameter ε for the case I = 40, J = 40, and T = 5.

ε	Objective Value	Wasserstein Distance
0.025	5456.74	41.7554
0.05	5331.98	51.0905
0.1	6664.91	51.2219
0.5	7038.21	89.9972

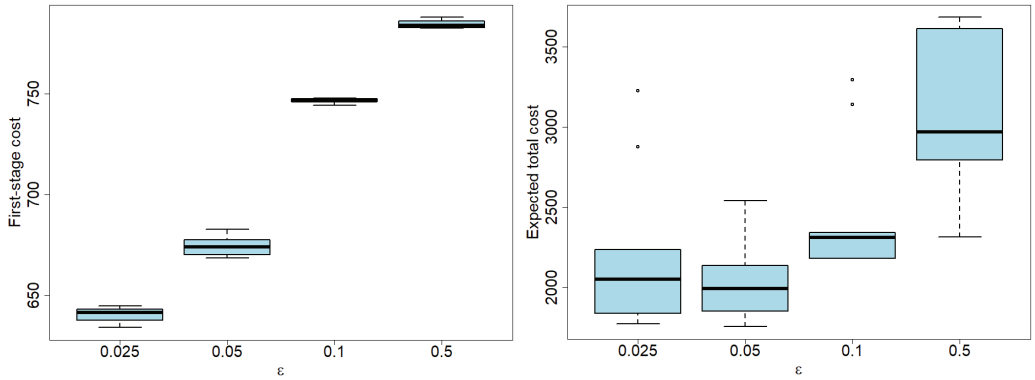


Figure 6. Effects of the Wasserstein ambiguity radius on the first-stage cost (**on the left**) and the overall cost (**on the right**) for the smallest test case (i.e., I = 15, J = 15, and T = 3). Box plots were obtained over 10 independent runs.

The observed behavior can be attributed to the interplay between robustness and conservatism inherent in WDRO. For small values of ε, the model remains vulnerable to a misestimation of demand, while at large values of ε, the solution tends to over-allocate resources to hedge against highly pessimistic demand realizations. This trade-off directly affects both cost efficiency and feasibility under chance constraints, as a larger ε generally improves satisfaction probability at the expense of increased operational cost.

Figure 7 illustrates the interplay between the out-of-sample objective value and the empirical satisfaction probability as a function of the Wasserstein radius ε, which governs the size of the ambiguity set in the distributionally robust optimization model. As shown, the objective value (blue line, left axis) initially increases with ε, reflecting the model's increasing conservatism in hedging against distributional shifts. Concurrently, the satisfaction probability (orange dashed line, right axis)—defined as the empirical frequency with which the chance constraints are met across test scenarios—monotonically increases with larger ε values, starting from 91%. This trend confirms the theoretical expectation that larger ambiguity sets provide more robust solutions by better encompassing the true demand distribution. However, the trade-off becomes evident as overly conservative solutions (large ε) yield diminishing returns in feasibility gains while incurring significantly higher costs.

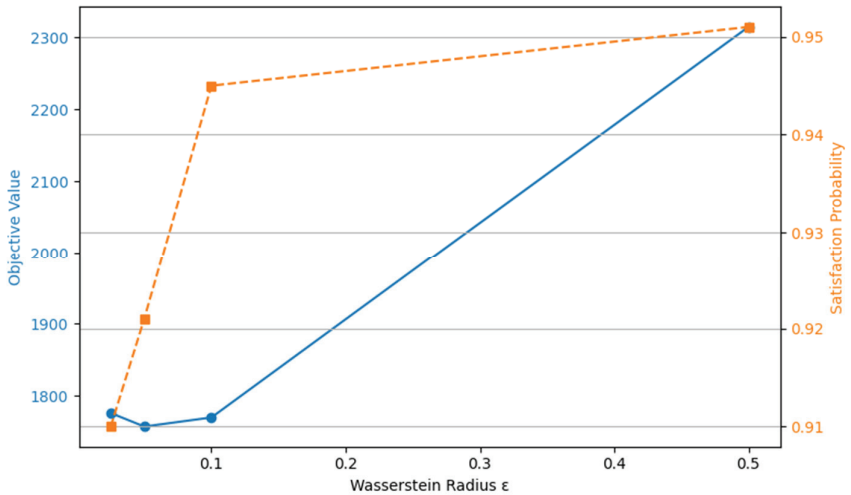


Figure 7. Effects of ε on performance and feasibility (for the case I = 15, J = 15, and T = 3).

5.4. Performance Analysis of the Proposed Algorithm

Table 9 demonstrates the superior performance of the proposed cJKO method compared to the Outer Approximation (OA) algorithm from reference [21] and the Robust Optimization (RO) baseline across six problem instances ranging from 15 to 40 facilities and customers over three to five time periods. The cJKO method consistently achieves the lowest objective values across all configurations, outperforming OA by 0.4% to 0.95% and showing even more significant improvements over RO, particularly in instances with longer planning horizons where RO performs poorly (e.g., achieving 6948.55 vs. cJKO's 5331.98 in the largest instance).

Table 9. Performance comparison: RO, approach presented in [21] (e.g., OA), and cJKO.

I	J	T	OA [21]	RO	cJKO
15	15	3	1764.67	1925.69	1757.1674
15	15	5	2048.83	3934.09	2028.83
30	30	3	3442.99	3998.26	3439.26
30	30	5	4045.88	5900.92	4025.88
40	40	3	4558.41	4797.00	4477.00
40	40	5	5351.98	6948.55	5331.98

The proposed cJKO approach leverages the Wasserstein distance and transport maps to balance robustness and solution quality more effectively than both benchmark methods, addressing the limitations of OA's worst-case joint chance constraints and RO's overly conservative uncertainty hedging. By modeling facility opening decisions as binary matrices over the planning horizon while dynamically managing storage levels under uncertain demand and capacity constraints, the cJKO method demonstrates superior handling of spatial-temporal complexity in facility location and inventory planning problems, maintaining its performance advantages as both the problem size and planning horizons increase.

5.5. Computational Time Analysis

The computational time comparison in Table 10 reveals distinct trade-offs between solution quality and computational efficiency across the three approaches. The RO baseline achieves the fastest execution times (0.46 to 36.87 s), while the OA algorithm exhibits moderate computational requirements (3.33 to 162.73 s). In contrast, the proposed cJKO method requires significantly higher computational resources, with execution times ranging from 224.42 s to 5346.68 s. This computational overhead stems from the iterative nature of the cJKO scheme, which requires solving multiple optimization subproblems at each iteration, computing Wasserstein distances between probability measures, and optimizing transport maps to characterize distributional ambiguity. With this approach, in order to obtain better solution quality through a more accurate uncertainty quantification, the method necessarily incurs substantial computational costs.

Table 10. Computational time analysis.

I	J	T	OA	RO	cJKO
15	15	3	3.33	0.46	224.42
15	15	5	7.99	0.93	453.33
30	30	3	35.25	5.25	5162.91
30	30	5	139.41	35.15	5162.91
40	40	3	21.95	26.87	4548.81
40	40	5	162.73	36.87	5346.68

The execution time scaling reveals that cJKO is particularly sensitive to increases in planning horizon length, with dramatic time increases when moving from three to five time periods (e.g., from 1561.33 to 5162.91 s for the 30 × 30 configuration). For practical applications, this computational profile makes the cJKO method most suitable for strategic planning scenarios where solution optimality is paramount, while simpler approaches like RO remain preferable for operational decisions requiring rapid response times.

Figure 8 presents the convergence behavior of the optimization algorithm, demonstrating the progressive improvement of the objective function. Initially, the solution explores a wider search space, leading to fluctuations due to the stochastic nature of the optimization process. As the iterations progress, the algorithm stabilizes around a near-optimal solution, reflecting the balance between exploration and exploitation. The sharp decrease in the objective value at a later stage suggests a structural shift in the solution space, possibly due to a critical update in decision variables or constraints. The final stabilization indicates convergence, confirming the algorithm's ability to efficiently navigate the solution landscape and identify an optimal or near-optimal solution within a finite number of iterations.

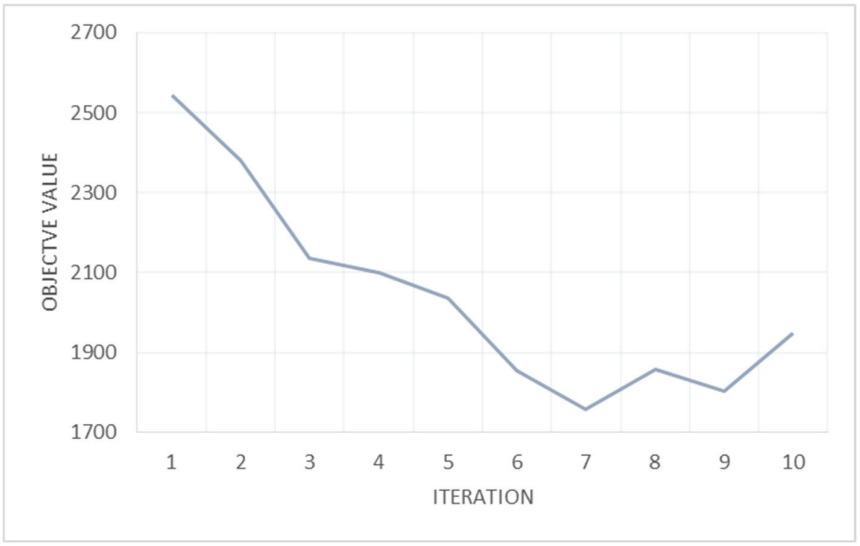


Figure 8. The convergence behavior of the optimization algorithm.

6. Conclusions

This study introduces an improved chance-constrained optimization framework for facility location under demand uncertainty, integrating a Wasserstein-based distributional approach to enhance robustness. By addressing the inherent variability in demand, the model ensures that probabilistic constraints are satisfied while optimizing facility placement and capacity allocation. Computational experiments across multiple problem instances demonstrate that the proposed cJKO approach consistently outperforms established benchmarks, achieving objective values 0.4% to 0.95% better than the Outer Approximation method and significantly superior results compared to traditional Robust Optimization, particularly in scenarios with extended planning horizons. Furthermore, the incorporation of Wasserstein distance constraints improves robustness against distributional shifts, ensuring that the model remains valid under different demand realizations. This highlights the significance of distributionally robust optimization in facility location problems, enabling more resilient decision-making under uncertainty.

Despite these advantages, some limitations remain. The model assumes to have historical data of demand. Additionally, the computational burden increases substantially with the problem size and planning horizon length, with execution times ranging from minutes to hours for larger instances, necessitating a consideration of the trade-off between

solution quality and computational efficiency. The cJKO method's iterative nature and complex Wasserstein distance computations make it most suitable for strategic planning scenarios where solution optimality justifies the computational investment, while simpler approaches may be preferred for operational decisions requiring a rapid response. Future research could explore adaptive uncertainty sets to refine demand estimates dynamically or integrate real-time learning mechanisms to enhance responsiveness. Extending the framework to multi-stage decision processes or multi-echelon networks would further improve its applicability in complex supply chain systems.

Author Contributions: Conceptualization, F.A. and E.M.; methodology, A.C.; software, I.S.; validation, I.S. and A.C.; writing—original draft preparation, I.S. and F.A.; writing—review and editing, all.; visualization, I.S.; supervision, A.C.; funding acquisition, E.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was co-funded by the National Research project ULTRA OPTYMAL (Urban Logistics and Sustainable Transportation: Optimization Under Uncertainty and Machine Learning) under the PRIN 2020 program, Italian Ministry for University and Research.

Data Availability Statement: The entire code, including the data generation, is freely available at the following repository: https://github.com/iman-ie/FacilityLocation_cJKO.git (accessed on 1 May 2025).
Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Meira, S. Everything Platforms. SSRN. 2025. Available online: https://ssrn.com/abstract=5160171 (accessed on 1 May 2025).
2. Xiao, H.; Zhang, J.; Zhang, Z.; Li, W. A Survey of Approximation Algorithms for the Universal Facility Location Problem. *Mathematics* **2025**, *13*, 1023. [CrossRef]
3. Peyré, G.; Cuturi, M. *Computational Optimal Transport*; CREST Working Papers; CREST: Fort Worth, TX, USA, 2017; pp. 2017–2086.
4. Peyré, G.; Cuturi, M. Computational optimal transport: With applications to data science. *Found. Trends Mach. Learn.* **2019**, *11*, 355–607. [CrossRef]
5. Uscidda, T.; Cuturi, M. The Monge Gap: A Regularizer to Learn All Transport Maps. In Proceedings of the International Conference on Machine Learning, Hawaii, HI, USA, 23–29 July 2023; PMLR 202, pp. 34709–34733.
6. Candelieri, A.; Ponti, A.; Archetti, F. Wasserstein enabled Bayesian optimization of composite functions. *J. Ambient Intell. Humaniz. Comput.* **2023**, *14*, 11263–11271. [CrossRef]
7. Jordan, R.; Kinderlehrer, D.; Otto, F. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.* **1998**, *29*, 1–17. [CrossRef]
8. Xu, S.; Govindan, K.; Wang, W.; Yang, W. Supply chain management under cap-and-trade regulation: A literature review and research opportunities. *Int. J. Prod. Econ.* **2024**, *271*, 109199. [CrossRef]
9. Peidro, D.; Martin, X.A.; Panadero, J.; Juan, A.A. Solving the uncapacitated facility location problem under uncertainty: A hybrid tabu search with path-relinking simheuristic approach. *Appl. Intell.* **2024**, *54*, 5617–5638. [CrossRef]
10. Snyder, L.V. Facility location under uncertainty: A review. *IIE Trans.* **2006**, *38*, 547–564. [CrossRef]
11. Laporte, G.; Nickel, S.; Saldanha-da-Gama, F. *Introduction to Location Science*; Springer International Publishing: Cham, Switzerland, 2019; pp. 1–21.
12. Kouvelis, P.; Yu, G. *Robust Discrete Optimization and Its Applications*; Springer Science & Business Media: Boston, MA, USA, 1996; Volume 14.
13. Jabbarzadeh, A.; Fahimnia, B.; Seuring, S. Dynamic supply chain network design for the supply of blood in disasters: A robust model with real world application. *Transp. Res. Part E Logist. Transp. Rev.* **2014**, *70*, 225–244. [CrossRef]
14. Shapiro, A.; Dentcheva, D.; Ruszczynski, A. *Lectures on Stochastic Programming: Modeling and Theory*, 3rd ed.; Shapiro, A., Dentcheva, D., Ruszczynski, A., Eds.; SIAM Publications Library: Philadelphia, PA, USA, 2021.
15. Ho, C.P.; Petrik, M.; Wiesemann, W. Robust ϕ-Divergence MDPs. *Adv. Neural Inf. Process. Syst.* **2022**, *35*, 32680–32693.
16. Delage, E.; Ye, Y. Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Oper. Res.* **2010**, *58*, 595–612. [CrossRef]
17. Kirschner, J.; Bogunovic, I.; Jegelka, S.; Krause, A. Distributionally Robust Bayesian Optimization. In Proceedings of the International Conference on Artificial Intelligence and Statistics (AISTATS), Online, 26–28 June 2020; PMLR 108, pp. 2174–2184.

18. Micheli, F.; Balta, E.C.; Tsiamis, A.; Lygeros, J. Wasserstein Distributionally Robust Bayesian Optimization with Continuous Context. *arXiv* **2025**, arXiv:2503.20341.
19. Liu, K.; Li, Q.; Zhang, Z.H. Distributionally robust optimization of an emergency medical service station location and sizing problem with joint chance constraints. *Transp. Res. Part B Methodol.* **2019**, *119*, 79–101. [CrossRef]
20. Ji, R.; Lejeune, M.A. Data-driven distributionally robust chance-constrained optimization with Wasserstein metric. *J. Glob. Optim.* **2021**, *79*, 779–811. [CrossRef]
21. Wang, Z.; Liu, K.; You, K.; Wang, Z. Wasserstein distributionally robust facility location and capacity planning for disaster relief. *Expert Syst. Appl.* **2025**, *272*, 126647. [CrossRef]
22. Zhang, L.; Yang, J.; Gao, R. A short and general duality proof for Wasserstein distributionally robust optimization. *Oper. Res.* **2024**. [CrossRef]
23. Bayraksan, G.; Love, D.K. Data-driven stochastic programming using phi-divergences. *Oper. Res. Revolut.* **2015**, 1–19. [CrossRef]
24. Mohajerin Esfahani, P.; Kuhn, D. Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations. *Math. Program.* **2018**, *171*, 115–166. [CrossRef]
25. Wang, Z.; Jiang, S.; Lin, Y.; Tao, Y.; Chen, C. A review of algorithms for distributionally robust optimization using statistical distances. *J. Ind. Manag. Optim.* **2025**, *21*, 1749–1770. [CrossRef]
26. Zhou, C.; Bayraksan, G. Effective Scenarios in Distributionally Robust Optimization with Wasserstein Distance. Optimization Online. Available online: https://optimization-online.org/wp-content/uploads/2024/12/Effective_Scen_s_in_DRO_W_paper.pdf (accessed on 20 June 2025).
27. Gen, M.; Cheng, R. *Genetic Algorithms and Engineering Optimization*; John Wiley & Sons: Hoboken, NJ, USA, 1999.
28. Saeedi, M.; Parhazeh, S.; Tavakkoli-Moghaddam, R.; Khalili-Fard, A. Designing a two-stage model for a sustainable closed-loop electric vehicle battery supply chain network: A scenario-based stochastic programming approach. *Comput. Ind. Eng.* **2024**, *190*, 110036. [CrossRef]
29. Lee, W.; Wang, L.; Li, W. Deep JKO: Time-implicit particle methods for general nonlinear gradient flows. *J. Comput. Phys.* **2024**, *514*, 113187. [CrossRef]
30. Candelieri, A.; Ponti, A.; Archetti, F. A Constrained-JKO Scheme for Effective and Efficient Wasserstein Gradient Flows. In Proceedings of the International Conference on Learning and Intelligent Optimization, Ischia Island, Italy, 9–13 June 2024; Springer Nature: Cham, Switzerland, 2024; pp. 66–80.
31. Korotin, A.; Egiazarian, V.; Asadulaev, A.; Safin, A.; Burnaev, E. Wasserstein-2 generative networks. *arXiv* **2019**, arXiv:1909.13082.
32. Korotin, A.; Li, L.; Solomon, J.; Burnaev, E. Continuous Wasserstein-2 barycenter estimation without minimax optimization. In Proceedings of the International Conference on Learning Representations, Vienna, Austria, 4 May 2021.
33. Mokrov, P.; Korotin, A.; Li, L.; Genevay, A.; Solomon, J.M.; Burnaev, E. Large-scale Wasserstein gradient flows. *Adv. Neural Inf. Process. Syst.* **2021**, *34*, 15243–15256.
34. Korotin, A.; Selikhanovych, D.; Burnaev, E. Kernel neural optimal transport. *arXiv* **2022**, arXiv:2205.15269.
35. Korotin, A.; Selikhanovych, D.; Burnaev, E. Neural optimal transport. *arXiv* **2022**, arXiv:2201.12220.
36. Fan, J.; Zhang, Q.; Taghvaei, A.; Chen, Y. Variational Wasserstein gradient flow. *Proc. Int. Conf. Mach. Learn.* **2022**, *PMLR 162*, 6185–6215.
37. Candelieri, A.; Ponti, A.; Archetti, F. Generative models via optimal transport and Gaussian processes. In *Learning and Intelligent Optimization*; Sellmann, M., Tierney, K., Eds.; Springer: Berlin/Heidelberg, Germany, 2023; pp. 135–149.
38. Kong, Z.; Chaudhuri, K. The expressive power of a class of normalizing flow models. *Proc. Int. Conf. Artif. Intell. Stat.* **2020**, *PMLR 108*, 3599–3609.
39. Li, L.; Hurault, S.; Solomon, J.M. Self-consistent velocity matching of probability flows. *Adv. Neural Inf. Process. Syst.* **2023**, *36*, 57038–57057.
40. Holland, J.H. Genetic Algorithms. *Sci. Am.* **1992**, *267*, 66–73. [CrossRef]
41. Harris, C.R.; Millman, K.J.; van der Walt, S.J.; Gommers, R.; Virtanen, P.; Counrnapeau, D.; Wieser, E.; Taylor, J.; Berg, S.; Smith, N.J.; et al. Array programming with NumPy. *Nature* **2020**, *585*, 357–362. [CrossRef]
42. Virtanen, P.; Gommers, R.; Oliphant, T.E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **2020**, *17*, 261–272. [CrossRef]
43. Fortin, F.-A.; Rainville, F.-M.; Gardner, M.-A.; Parizeau, M.; Gagné, C. DEAP: Evolutionary Algorithms Made Easy. *J. Mach. Learn. Res.* **2012**, *13*, 2171–2175.
44. Hunter, J.D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **2007**, *9*, 90–95. [CrossRef]
45. Zhang, Z.H.; Li, K. A novel probabilistic formulation for locating and sizing emergency medical service stations. *Ann. Oper. Res.* **2015**, *229*, 813–835. [CrossRef]
46. Taguchi, G. *Introduction to Quality Engineering: Designing Quality into Products and Processes*; Asian Productivity Organization: Tokyo, Japan, 1986.

47. Hosseini Baboli, S.A.; Arabkoohsar, A.; Seyedi, I. Numerical modeling and optimization of pressure drop and heat transfer rate in a polymer fuel cell parallel cooling channel. *J. Braz. Soc. Mech. Sci. Eng.* **2023**, *45*, 201. [CrossRef]
48. Rajabi-Kafshgar, A.; Seyedi, I.; Tirkolaee, E.B. Circular Closed-Loop Supply Chain Network Design Considering 3D Printing and PET Bottle Waste. *Environ. Dev. Sustain.* **2024**, 1–37. [CrossRef]

mathematics

MDPI

Article

Applying λ-Statistical Convergence in Fuzzy Paranormed Spaces to Supply Chain Inventory Management Under Demand Shocks ($\mathcal{DS}$)

Hasan Öğünmez [1] and Muhammed Recai Türkmen [2,*]

[1] Department of Mathematics, Faculty of Science and Literature, Afyon Kocatepe University, Afyonkarahisar 03200, Turkey; hogunmez@aku.edu.tr
[2] Department of Mathematics and Science Education, Faculty of Education, Afyon Kocatepe University, Afyonkarahisar 03200, Turkey
* Correspondence: mrtmath@gmail.com

Abstract: This paper introduces and analyzes the concept of λ-statistical convergence in fuzzy paranormed spaces, demonstrating its relevance to supply chain inventory management under demand shocks. We establish key relationships between generalized convergence methods and fuzzy convex analysis, showing how these results extend classical summability theory to uncertain demand environments. By exploring λ-statistical Cauchy sequences and (V, λ)-summability in fuzzy paranormed spaces, we provide new insights applicable to adaptive inventory optimization and decision-making in supply chains. Our findings bridge theoretical aspects of fuzzy convexity with practical convergence tools, advancing the robust modeling of demand uncertainty.

Keywords: fuzzy paranormed spaces; λ-statistical convergence; supply chain modelling; inventory management; demand forecasting

MSC: 40G15; 46S40; 90B05

1. Introduction and Background

The idea of convergence plays a crucial role in mathematical analysis and its applications. Conventional convergence, which studies how sequences approach a certain limit, is a well-established concept. Nevertheless, classical convergence can be limiting in some situations. To overcome these constraints, various alternative types of convergence have been proposed, with statistical convergence being particularly notable for its flexibility and wider applicability.

Statistical convergence was first introduced by Steinhaus and Fast in 1951 [1,2]. It broadens the concept of classical convergence by focusing on the density of terms in a sequence rather than their exact positions. A sequence is considered statistically convergent if the indices at which the sequence deviates significantly from the limit form a set with zero density. This method not only extends the scope of convergence but also finds applications in fields such as number theory, functional analysis, and approximation theory.

Over the years, the concept of statistical convergence has been extended to various frameworks, including normed spaces, metric spaces, fuzzy normed spaces, paranormed spaces, and fuzzy paranormed spaces [3–25].

Recent advances have significantly expanded these concepts, including higher-order statistical convergence in fuzzy difference sequence spaces with applications to fuzzy number theory [26], weighted statistical convergence of fractional order for double sequences

in paranormed spaces [27], λ-statistical convergence in fuzzy n-normed linear spaces [28], generalizations to triple sequences via $M\lambda_{m,n,p}$-statistical convergence [29], and Tauberian theory for statistically Cesàro summable triple sequences of fuzzy numbers [30].

These contributions underscore the ongoing evolution of statistical convergence, particularly in multidimensional settings (double and triple sequences), advanced fuzzy structures, and Tauberian theory, reinforcing its interdisciplinary relevance.

Building upon statistical convergence, the concept of λ-statistical convergence was introduced by Mursaleen [31] as a more generalized framework. Here, $\lambda = (\lambda_k)$ represents a sequence that determines the weighted density of terms in a sequence. This extension provides greater flexibility in analyzing convergence behavior, allowing the study of sequences under nonuniform density conditions. The λ-statistical convergence framework has been further explored in various mathematical settings, such as paranormed spaces and fuzzy spaces, offering new insights and tools for sequence analysis [32–41].

Despite these advancements, existing frameworks for λ-statistical convergence exhibit critical limitations in handling complex uncertainties. For instance, studies in fuzzy normed spaces (e.g., [25]) lack the flexibility to model non-homogeneous uncertainty distributions, while works in paranormed spaces (e.g., [8]) cannot capture gradual membership transitions inherent in imprecise data. Furthermore, recent extensions to n-normed spaces [28] or triple sequences [29] focus primarily on theoretical generalizations without providing adaptive convergence criteria for real-world volatility.

In contrast, our work bridges these gaps by unifying λ-statistical convergence with fuzzy paranormed spaces. This synthesis enables:

(i) Dynamic density-based weighting (via λ-sequences) for irregular demand patterns;
(ii) Fuzzy paranormed structures to quantify partial or transitional uncertainties; and
(iii) Robust convergence criteria for sequences with abrupt, nonuniform fluctuations—addressing rigidity in earlier models [5,15].

This framework thus overcomes the key constraint of prior approaches: their inability to jointly model stochastic volatility and fuzzy imprecision in a unified topology.

In practical supply chain inventory management, sudden demand shocks and imprecise demand forecasting often lead to significant stockouts or overstocking. Therefore, applying the framework of λ-statistical convergence in fuzzy paranormed spaces provides a robust tool for modeling such uncertainties and designing adaptive replenishment policies that maintain desired service levels under fluctuating demand.

The illustrative inventory case study serves to demonstrate the practical application of the theoretical framework; full-scale industrial implementations and empirical validations are postponed to future work.

This paper aims to extend the concept of λ-statistical convergence to fuzzy paranormed spaces, providing a comprehensive framework for analyzing convergence in these settings. The main results include new definitions, theorems, and illustrative examples that demonstrate the applicability and significance of this generalization.

The subsequent sections are organized as follows: Section 2 provides the necessary preliminaries, including key definitions and foundational concepts. Section 3 presents the fundamental definitions and propositions underpinning the theoretical framework of our study. In Section 4, we establish the core theorems that offer critical insights into λ-statistical convergence within this context. Finally, Section 5 summarizes the implications of these theorems and outlines directions for future research.

2. Preliminaries

This section introduces the foundational concepts of paranormed spaces, fuzzy normed spaces, fuzzy paranormed spaces, statistical convergence, and λ-statistical con-

vergence. We begin by defining paranormed and fuzzy normed spaces, followed by the notion of statistical convergence within these spaces. Finally, we extend these concepts to λ-statistical convergence in both classical and generalized settings, including fuzzy paranormed spaces.

2.1. Paranormed Spaces

Let $\wp : X \to \mathbb{R}$ be a function, and let X be a real or complex linear space. If all $p, q \in X$ satisfy the following requirements, then $\wp$ is a paranorm and the pair $(X, \wp)$ is called a paranormed space.

1. $\wp(\theta) = 0$;
2. $\wp(-p) = \wp(p)$;
3. $\wp(p+q) \leq \wp(p) + \wp(q)$;
4. $\wp(\gamma_k p_k - \gamma p) \to 0$ as $k \to \infty$ if (γ_k) is a sequence of scalars with $\gamma_k \to \gamma$ as $k \to \infty$ and (p_k) is a sequence in X with $\wp(p_k - p) \to 0$ as $k \to \infty$.

If $\wp(p) = 0$ implies $p = \theta$, then $\wp$ is said to be a total paranorm, where θ is the zero vector of X.

In paranormed spaces, the concept of convergence is defined similarly to normed spaces.

A sequence $p = (p_k)$ is considered convergent (or $\wp$-convergent) to the element ξ in $(X, \wp)$ if, for each $\varepsilon > 0$, there exists a positive integer k_0 such that $\wp(p_k - \xi) < \varepsilon$ whenever $k \geq k_0$. In this case, we write $\wp\text{-lim}\, p = \xi$, and ξ is referred to as the $\wp$-limit of p.

2.2. Fuzzy Normed Spaces

The fuzzy norm introduced by Felbin [42], further developed by Xiao and Zhu [43] and finalized by Şençimen and Pehlivan [17], is given as follows.

Here, $\tilde{0}$ denotes the zero fuzzy number, i.e., a fuzzy number whose membership function is fully concentrated at 0, typically represented as the degenerate fuzzy set

$$\mu_{\tilde{0}}(x) = \begin{cases} 1, & \text{if } x = 0, \\ 0, & \text{otherwise.} \end{cases}$$

Let X be a vector space over $\mathbb{R}$. Let $\|\cdot\| : X \to L^*(\mathbb{R})$ be a mapping, and let L and R be mappings (respectively, the left norm and right norm) from $[0,1] \times [0,1]$ to $[0,1]$, which are symmetric, nondecreasing in both arguments and satisfy $L(0,0) = 0$ and $R(1,1) = 1$.

The quadruple $(X, \|\cdot\|, L, R)$ is called a fuzzy normed linear space (briefly $(X, \|\cdot\|)$ FNS), and $\|\cdot\|$ is called a fuzzy norm if the following axioms are satisfied.

1. $\|x\| = \tilde{0}$ if and only if $x = \theta$;
2. $\|rx\| = |r| \cdot \|x\|$ for $x \in X$, $r \in \mathbb{R}$;
3. For all $x, y \in X$:
   a. $\|x+y\|(s+t) \geq L(\|x\|(s), \|y\|(t))$, whenever $s \leq \|x\|_1^-$, $t \leq \|y\|_1^-$, and $s + t \leq \|x+y\|_1^-$;
   b. $\|x+y\|(s+t) \leq R(\|x\|(s), \|y\|(t))$, whenever $s \geq \|x\|_1^-$, $t \geq \|y\|_1^-$, and $s + t \geq \|x+y\|_1^-$.

Write $[\|x\|]_\alpha = [\|x\|_\alpha^-, \|x\|_\alpha^+]$ for $x \in X$ and $0 \leq \alpha \leq 1$. Suppose that for all $x \in X$, $x \neq \theta$, we have $\inf_{\alpha \in [0,1]} \|x\|_\alpha^- > 0$, where θ is the zero vector of X.

Şençimen and Pehlivan [17] modified conditions 3-a and 3-b in the fuzzy norm definition as follows:

For all $x, y \in X$:

3-a. $\|x+y\|_0^- \leq \|x\|_0^- + \|y\|_0^-$;

3-b. $\|x+y\|_0^+ \le \|x\|_0^+ + \|y\|_0^+$.

Thus, in an FNS $(X, \|\cdot\|)$, the triangle inequality of the fuzzy norm definition (3) implies $\|x+y\| \preceq \|x\| \oplus \|y\|$. According to this definition, $x = \theta$ if and only if $\|x\|_\alpha^- = \|x\|_\alpha^+ = 0$ for all $\alpha \in [0,1]$. Furthermore, $\|x\|_0^- > 0$ whenever $x \neq \theta$.

Now, if $r = 0$, then

$$[\|rx\|]_\alpha = [\|\theta\|]_\alpha = [0,0] = [|r|\|x\|]_\alpha$$

for all $\alpha \in [0,1]$ and $x \in X$. For $r \neq 0$, we have

$$[\|rx\|]_\alpha = [|r|\|x\|]_\alpha$$

for each $\alpha \in [0,1]$, i.e.,

$$[\|rx\|]_\alpha^- = [|r|\|x\|]_\alpha^- \quad \text{and} \quad [\|rx\|]_\alpha^+ = [|r|\|x\|]_\alpha^+$$

for each $\alpha \in [0,1]$, where $L = \min$ and $R = \max$.

Along with this definition, Şençimen and Pehlivan [17] defined convergence in fuzzy normed spaces by using the distance framework developed by Kaleva [44] and Felbin [42]. Throughout this paper, we denote that distance by $\mathrm{D}(\cdot,\cdot)$. Specifically, if A and B are fuzzy numbers with α–cuts $[A]_\alpha = [A_\alpha^-, A_\alpha^+]$ and $[B]_\alpha = [B_\alpha^-, B_\alpha^+]$, we set

$$\mathrm{D}(A,B) = \sup_{0\le\alpha\le 1} \max\{|A_\alpha^- - B_\alpha^-|, |A_\alpha^+ - B_\alpha^+|\},$$

so that $\mathrm{D}(\|p_k - \ell\|, \tilde{0})$ measures the deviation of the fuzzy norm $\|p_k - \ell\|$ from the zero fuzzy number $\tilde{0}$.

A sequence (p_k) (in the fuzzy normed space $(X, \|\cdot\|)$) is convergent to $\ell \in X$ provided that

$$(\mathrm{D})\text{-}\lim_{k\to\infty} \|p_k - \ell\| = \tilde{0};$$

i.e., for all $\varepsilon > 0$, there exists $p_0 \in \mathbb{N}$ such that

$$\mathrm{D}(\|p_k - \ell\|, \tilde{0}) < \varepsilon$$

for all $k \ge p_0$. We denote this by $p_k \xrightarrow{FN} \ell$.

This means that for every $\varepsilon > 0$, there exists $k_0(\varepsilon) \in \mathbb{N}$ such that

$$\sup_{\alpha\in[0,1]} \|p_k - \ell\|_\alpha^+ = \|p_k - \ell\|_0^+ < \varepsilon$$

for all $k \ge k_0$. In terms of neighborhoods, we have $p_k \xrightarrow{FN} \ell$ provided that for each $\varepsilon > 0$, there exists $p_0(\varepsilon) \in \mathbb{N}$ such that $p_k \in \aleph_\ell(\varepsilon, 0)$ whenever $k \ge p_0$.

2.3. Fuzzy Paranormed Spaces

Çınar et al. [23] introduced the notion of a fuzzy paranorm—a paranorm-type extension of the classical fuzzy norms of Felbin [42] and Xiao–Zhu [43] that replaces the homogeneity axiom with a paranorm-continuity requirement—and their definition is recalled below as follows.

Let $\wp : X \to L^*(\mathbb{R})$ be a paranorm and let $X \subset \mathbb{R}$ be a vector space. Let the mappings L and $R : [0,1]\times[0,1] \to [0,1]$ be symmetric, non-decreasing, and satisfy $L(0,0) = 0$ and $R(1,1) = 1$. Write

$$[\wp(x)]_\alpha = [\wp(x)_\alpha^-, \wp(x)_\alpha^+]$$

for $p \in X$ and $\alpha \in (0,1]$. Suppose that for all $p \in X$, $p \neq \theta$ (where θ is the zero vector of X),

$$\inf_{\alpha \in [0,1]} \wp(p)_\alpha^- > 0, \quad \sup \wp(p)_\alpha^+ < \infty.$$

The quadruple $(X, \wp, L, R)$ is called a fuzzy paranormed space and $\wp$ is a fuzzy paranorm if the following conditions hold:

1. $\wp(p) = \tilde{0}$ if $p = \theta$;
2. $\wp(-p) = \wp(p)$ for all $p \in X$;
3. For all $p, q \in X$;

   (a) If $s \leq \wp(p)_1^-$, $t \leq \wp(q)_1^-$, and $s + t \leq \wp(p+q)_1^-$, then

   $$\wp(p+q)(s+t) \geq L(\wp(p)(s), \wp(q)(t));$$

   (b) If $s \geq \wp(p)_1^-$, $t \geq \wp(q)_1^-$, and $s + t \geq \wp(p+q)_1^-$, then

   $$\wp(p+q)(s+t) \leq R(\wp(p)(s), \wp(q)(t));$$

4. If (γ_k) is a sequence in $\mathbb{R}$ with $\gamma_k \to \gamma$ as $k \to \infty$ and $p_k, \ell \in X$ for all $k \in \mathbb{N}$ with $\wp(p_k - \ell) \to \tilde{0}$ as $k \to \infty$, then

   $$\wp(\gamma_k p_k - \gamma \ell) \to \tilde{0} \quad \text{as } k \to \infty.$$

If $\wp(p) = \tilde{0}$ implies $p = \theta$, then the fuzzy paranorm is referred to as a totally fuzzy paranorm.

2.4. Statistical Convergence

A set K of positive integers has a natural density defined by

$$\delta(K) = \lim_{n \to \infty} \frac{1}{n} |\{k \in K : k \leq n\}|,$$

where $|\{k \in K : k \leq n\}|$ represents the count of elements of K not exceeding n. It is evident that $\delta(K) = 0$ for any finite set K.

Statistical convergence using natural density was defined by Fast [2] as follows.

If the set

$$\{n \in \mathbb{N} : |p_k - \varsigma| > \varepsilon\}$$

has natural density zero for every $\varepsilon > 0$, then the real number sequence (p_k) is said to be statistically convergent to $\varsigma \in \mathbb{R}$. In this case, we write

$$st - \lim p_k = \varsigma.$$

The definition of statistical convergence in paranormed spaces is given by Alotaibi and Alroqi [13] as follows.

If, for each $\varepsilon > 0$, we have

$$\lim_{n \to \infty} \frac{1}{n} |\{k \leq n : \wp(p_k - \xi) > \varepsilon\}| = 0,$$

then the sequence $p = (p_k)$ is statistically convergent to ξ in $(X, \wp)$ (or $\wp(st)$-convergent). We write

$$\wp(st) - \lim p = \xi.$$

The set of all $\wp(st)$-convergent sequences is denoted by $S_\wp$.

Next, we give the definitions of statistical convergence in fuzzy normed spaces [17] and fuzzy paranormed spaces [23], which form the basis of our study.

Assume that $(X, \|\cdot\|)$ is a fuzzy normed space. If

$$st - \lim \|p_k - \varsigma\| = \tilde{0},$$

then a sequence (p_k) in X is statistically convergent to $\varsigma \in X$, and we write

$$p_k \overset{st(FN)}{\to} \varsigma,$$

i.e., for each $\varepsilon > 0$,

$$\delta\big(\{k \in \mathbb{N} : \mathrm{D}(\|p_k - \varsigma\|, \tilde{0}) \geq \varepsilon\}\big) = 0.$$

This means that for any $\varepsilon > 0$, the natural density of the set

$$\{k \in \mathbb{N} : \|p_k - \varsigma\|_0^+ \geq \varepsilon\}$$

is zero. In other words, for each $\varepsilon > 0$, the condition

$$\|p_k - \varsigma\|_0^+ < \varepsilon$$

holds for almost all k. In terms of neighborhoods, $p_k \overset{st(FN)}{\to} \varsigma$ means that for each $\varepsilon > 0$,

$$\delta\big(\{k \in \mathbb{N} : p_k \notin \aleph_\varsigma(\varepsilon, 0)\}\big) = 0,$$

i.e., $p_k \in \aleph_\varsigma(\varepsilon, 0)$ for almost all k.

A useful equivalent form is:

$$p_k \overset{st(FN)}{\to} \varsigma \iff st - \lim \|p_k - \varsigma\|_0^+ = 0.$$

When $st - \lim \|p_k - \varsigma\|_0^+ = 0$, it implies that for every $\alpha \in [0, 1]$,

$$st - \lim \|p_k - \varsigma\|_\alpha^- = st - \lim \|p_k - \varsigma\|_\alpha^+ = 0,$$

as for all $k \in \mathbb{N}$ and $\alpha \in [0, 1]$, we have

$$0 \leq \|p_k - \varsigma\|_\alpha^- \leq \|p_k - \varsigma\|_\alpha^+ \leq \|p_k - \varsigma\|_0^+.$$

Let $(X, \wp)$ be a fuzzy paranormed space. A sequence $p = (p_k)$ is statistically convergent to $\varsigma \in X$ if, for each $\varepsilon > 0$,

$$\delta\big(\{k \in \mathbb{N} : \mathrm{D}(\wp(p_k - \varsigma), \tilde{0}) \geq \varepsilon\}\big) = 0.$$

In this case, we write

$$p_k \overset{st(FP)}{\to} \varsigma \quad \text{or} \quad st^{FP} - \lim_{k \to \infty} \wp(p_k - \varsigma) = \tilde{0}.$$

This implies that for each $\varepsilon > 0$, the set

$$K(\varepsilon) = \{k \in \mathbb{N} : \wp(p_k - \varsigma)_0^+ \geq \varepsilon\}$$

has natural density zero. In other words, for each $\varepsilon > 0$, we have

$$\wp(p_k - \varsigma)_0^+ \leq \varepsilon$$

for almost all k, where

$$\wp(p_k - \varsigma)_0^+ = \sup_{\alpha \in [0,1]} \wp(p_k - \varsigma)_\alpha^+.$$

According to the above definition,

$$p_k \overset{st(FP)}{\to} \varsigma \iff st - \lim_{k\to\infty} \wp(p_k - \varsigma)_0^+ = 0.$$

Note that

$$st - \lim_{k\to\infty} \wp(p_k - \varsigma)_0^+ = 0$$

implies

$$st - \lim_{k\to\infty} \wp(p_k - \varsigma)_\alpha^- = st - \lim_{k\to\infty} \wp(p_k - \varsigma)_\alpha^+ = 0$$

for each $\alpha \in (0, 1]$ as

$$0 \leq \wp(p_k - \varsigma)_\alpha^- \leq \wp(p_k - \varsigma)_\alpha^+ \leq \wp(p_k - \varsigma)_0^+$$

holds for every $k \in \mathbb{N}$ and each $\alpha \in (0, 1]$. We denote the set of all statistically convergent sequences by $S(FP)$.

2.5. λ-Statistical Convergence

The concept of λ-statistical convergence generalizes statistical convergence by introducing a weight sequence $\lambda = (\lambda_n)$.

Let $\lambda = (\lambda_n)$ be a sequence of positive real numbers that is non-decreasing and tends to infinity, with the condition that $\lambda_{n+1} \leq \lambda_n + 1$ and $\lambda_1 = 1$. The collection of all such sequences is denoted by Λ.

The idea of λ-statistical convergence was proposed by Mursaleen [31] as follows.

A sequence $p = (p_k)$ is said to be λ-statistically convergent or S_λ-convergent to ς if for every $\varepsilon > 0$,

$$\lim_{n\to\infty} \frac{1}{\lambda_n} |\{k \in I_n : |p_k - \varsigma| \geq \varepsilon\}| = 0,$$

where $I_n = [n - \lambda_n + 1, n]$. In this case, we write $S_\lambda - \lim p = \varsigma$ or $p_k \to \varsigma(S_\lambda)$ and define

$$S_\lambda := \{p : \exists \varsigma \in \mathbb{R},\ S_\lambda - \lim p = \varsigma\}.$$

Alghamdi and Mursaleen [36] defined λ-statistical convergence in paranormed spaces as follows.

A sequence $p = (p_k)$ is said to be λ-statistically convergent to the number ξ in the paranormed space $(X, \wp)$ if, for each $\varepsilon > 0$,

$$\lim_{n\to\infty} \frac{1}{\lambda_n} |\{k \in I_n : \wp(p_k - \xi) \geq \varepsilon\}| = 0.$$

In this case, we write $st_\lambda(\wp) - \lim p = \xi$.

Türkmen and Çınar [34] defined the concept of λ-statistical convergence within fuzzy normed spaces as follows.

Let $(X, \|\cdot\|)$ be a fuzzy normed space (FNS) and $\lambda \in \Lambda$. A sequence $p = (p_k)$ in X is considered λ-statistically convergent to $\varsigma \in X$ with respect to the fuzzy norm on X, or FS_λ-convergent, if for every $\varepsilon > 0$,

$$\lim_{n\to\infty} \frac{1}{\lambda_n} \left|\{k \in I_n : \mathrm{D}(\|p_k - \varsigma\|, \tilde{0}) \geq \varepsilon\}\right| = 0,$$

and we write

$$p_k \overset{FS_\lambda}{\to} \varsigma,$$

or

$$p_k \to \varsigma(FS_\lambda),$$

or

$$S_\lambda \overset{FN}{-} \lim p_k = \varsigma,$$

where $I_n = [n - \lambda_n + 1, n]$.

This indicates that for every $\varepsilon > 0$, the set

$$K(\varepsilon) = \{k \in I_n : \|p_k - \varsigma\|_0^+ \geq \varepsilon\}$$

has natural density zero, which means that for each $\varepsilon > 0$, $\|p_k - \varsigma\|_0^+ < \varepsilon$ holds for almost all k.

In this case, we express $S_\lambda \overset{FN}{-} \lim p_k$ as ς. The collection of all sequences that converge statistically with respect to the fuzzy norm on X is denoted by FS_λ.

The element $\varsigma \in X$ serves as the FS_λ-limit of the sequence (p_k). In terms of neighborhoods, we say (p_k) converges to $\varsigma(FS_\lambda)$ if for every $\varepsilon > 0$, p_k lies in $\aleph_\varsigma(\varepsilon, 0)$ for almost all k.

An equivalent way to express this is:

$$p_k \overset{FS_\lambda}{\to} \varsigma \quad \text{if and only if} \quad S_\lambda \overset{FN}{-} \lim \|p_k - \varsigma\|_0^+ = 0.$$

Note that

$$S_\lambda \overset{FN}{-} \lim \|p_k - \varsigma\|_0^+ = 0$$

implies

$$S_\lambda \overset{FN}{-} \lim \|p_k - \varsigma\|_\alpha^- = S_\lambda \overset{FN}{-} \lim \|p_k - \varsigma\|_\alpha^+ = 0$$

for each $\alpha \in [0, 1]$, as

$$0 \leq \|p_k - \varsigma\|_\alpha^- \leq \|p_k - \varsigma\|_\alpha^+ \leq \|p_k - \varsigma\|_0^+$$

holds for every $k \in I_n$ and each $\alpha \in [0, 1]$.

3. Main Definitions and Propositions

In this section, we present the concept of λ-statistical convergence for sequences in fuzzy paranormed spaces. Furthermore, we explore its relationship with λ-statistically Cauchy sequences and λ-summability within this framework.

Definition 1. *Let $(X, \wp)$ be a fuzzy paranormed space and $\lambda \in \Lambda$. A sequence $p = (p_k)$ in X is said to be λ-statistically convergent to $\xi \in X$ with respect to the fuzzy paranorm on X, or FPS_λ-convergent, if for each $\varepsilon > 0$,*

$$\lim_{n\to\infty} \frac{1}{\lambda_n} \left|\{k \in I_n : \mathrm{D}(\wp(p_k - \xi), \tilde{0}) \geq \varepsilon\}\right| = 0,$$

and we write

$$p_k \overset{FPS_\lambda}{\to} \xi, \quad p_k \to \xi(FPS_\lambda), \quad or \quad S_\lambda \overset{FP}{-} \lim p_k = \xi,$$

where $I_n = [n - \lambda_n + 1, n]$.

For every $\varepsilon > 0$, the natural density of the set

$$\{k \in I_n : \wp(p_k - \xi)_0^+ \geq \varepsilon\}$$

is zero, meaning that for each $\varepsilon > 0$, $\wp(p_k - \xi)_0^+ < \varepsilon$ holds for almost all k.

In this case, we denote it as $S_\lambda \overset{FP}{-} \lim p_k = \xi$. The set of all statistically convergent sequences with respect to the fuzzy paranorm on X is denoted by FPS_λ.

The element $\xi \in X$ is the FPS_λ-limit of (p_k). In terms of neighborhoods, we have

$$p_k \overset{FPS_\lambda}{\to} \xi$$

provided that for each $\varepsilon > 0$, $p_k \in \aleph_\xi(\varepsilon, 0)$ for almost all k.

A useful way to interpret the above definition is:

$$p_k \overset{FPS_\lambda}{\to} \xi \quad \text{if and only if} \quad S_\lambda \overset{FP}{-} \lim \wp(p_k - \xi)_0^+ = 0.$$

Note that

$$S_\lambda \overset{FP}{-} \lim \wp(p_k - \xi)_0^+ = 0$$

implies

$$S_\lambda \overset{FP}{-} \lim \wp(p_k - \xi)_\alpha^- = S_\lambda \overset{FP}{-} \lim \wp(p_k - \xi)_\alpha^+ = 0$$

for each $\alpha \in [0, 1]$, as

$$0 \leq \wp(p_k - \xi)_\alpha^- \leq \wp(p_k - \xi)_\alpha^+ \leq \wp(p_k - \xi)_0^+$$

holds for every $k \in I_n$ and each $\alpha \in [0, 1]$.

Throughout this paper, we say that (p_k) is statistically convergent to $\xi \in X$ with respect to the fuzzy paranorm on X if it is FPS_λ-convergent to ξ.

Because the natural density of a finite set is zero, every convergent sequence in FPS is also λ-statistically convergent. However, the converse is not always true. Çınar et al. [23] demonstrated this for the case of statistical convergence in Example 3.4. If we take $\lambda = (\lambda_n) = (n) \in \Lambda$, we show that not every λ-statistically convergent sequence is convergent.

Proposition 1. *Let* (p_k) *and* (q_k) *be sequences in a fuzzy paranormed space* $(X, \wp)$ *such that*

$$p_k \overset{FPS_\lambda}{\to} \xi \quad and \quad q_k \overset{FPS_\lambda}{\to} \varsigma,$$

where $\xi, \varsigma \in X$. *Then, the following holds:*

(i) $(p_k + q_k) \overset{FPS_\lambda}{\to} \xi + \varsigma$;

(ii) $tp_k \overset{FPS_\lambda}{\to} t\xi$ *for all* $t \in \mathbb{R}$;

(iii) $S_\lambda \overset{FP}{-} \lim \wp(p_k) = \wp(\xi)$.

Definition 2. *Let $(X, \wp)$ be a fuzzy paranormed space (FPS). A sequence (p_k) in X is called λ-statistically Cauchy with respect to the fuzzy paranorm on X if for every $\varepsilon > 0$ there exists a number $p_0 = p_0(\varepsilon) \in \mathbb{N}$ such that*

$$\lim_{n\to\infty} \frac{1}{\lambda_n} \left| \{k \in I_n : \wp(p_k - p_{p_0})_0^+ \geq \varepsilon\} \right| = 0.$$

In the fuzzy paranorm on X, the fact that (p_k) is FPS_λ-Cauchy implies that it is λ-statistically Cauchy.

Proposition 2. *All FPS_λ-convergent sequences in $(X, \wp)$ are also FPS_λ-Cauchy sequences.*

Proof. Let $p_k \overset{FPS_\lambda}{\to} \xi$ and $\varepsilon > 0$. Then, we have $\wp(p_k - \xi)_0^+ < \varepsilon/2$ for almost all k. Choose $n_0 \in \mathbb{N}$ such that $\wp(p_{n_0} - \xi)_0^+ < \varepsilon/2$.

Because $\wp(\cdot)_0^+$ is a paranorm in the conventional sense, we have

$$\wp(p_k - p_{n_0})_0^+ = \wp\big((p_k - \xi) + (\xi - p_{n_0})\big)_0^+ \leq \wp(p_k - \xi)_0^+ + \wp(p_{n_0} - \xi)_0^+ < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

for almost all k. This shows that (p_k) is FPS_λ-Cauchy. □

Definition 3. *Let $(X, \wp)$ be an FPS and let $\lambda = (\lambda_n)$ be a non-decreasing sequence of positive numbers tending to ∞, satisfying $\lambda_{n+1} \leq \lambda_n + 1$ and $\lambda_1 = 1$. Let $p = (p_k)$ be a sequence in the set X.*

This sequence p is said to be strongly λ-summable with respect to the fuzzy paranorm on X if there exists $\xi \in X$ such that

$$\lim_{n\to\infty} \frac{1}{\lambda_n} \sum_{k\in I_n} \mathrm{D}\big(\wp(p_k - \xi), \tilde{0}\big) = 0,$$

where $I_n = [n - \lambda_n + 1, n]$.

In this context, we express $[V, \lambda]_{FP} - \lim p_k = \xi$. The collection of all fuzzy paranorms on X that are strongly (V, λ) summable is referred to as $[V, \lambda]_{FP}$.

We say that X is strongly λ-summable to ξ with respect to the fuzzy paranorm on X. If $\lambda_n = n$, then strong λ-summability reduces to strong Cesàro summability with respect to the fuzzy paranorm on X, which is defined as follows:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \mathrm{D}\big(\wp(p_k - \xi), \tilde{0}\big) = 0.$$

In this case, we write

$$[C, 1]_{FP} - \lim p_k = \xi.$$

The collection of all strongly $(C, 1)$ summable fuzzy paranorm sequences on X is denoted by $[C, 1]_{FP}$.

Therefore, we express them as follows:

$$[V, \lambda]_{FP} = \left\{ p = (p_k) : \lim_{n\to\infty} \frac{1}{\lambda_n} \sum_{k\in I_n} \mathrm{D}\big(\wp(p_k - \xi), \tilde{0}\big) = 0 \text{ for some } \xi \right\},$$

$$[C, 1]_{FP} = \left\{ p = (p_k) : \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \mathrm{D}\big(\wp(p_k - \xi), \tilde{0}\big) = 0 \text{ for some } \xi \right\}.$$

4. Core Theorems

In this section, we will give and prove important theorems using the definitions and propositions given in the previous section.

Theorem 1. *If a sequence $p = (p_k)$ is $\left[V, \lambda\right]_{FP}$-summable to ξ, then*

$$p_k \xrightarrow{FPS_\lambda} \xi.$$

Proof. Let $\varepsilon > 0$. From $\left[V, \lambda\right]_{FP}$-summability, we have

$$\lim_{n\to\infty} \frac{1}{\lambda_n} \sum_{k\in I_n} \mathrm{D}\big(\wp(p_k - \xi), \tilde{0}\big) = 0.$$

On the other hand, by positivity of D,

$$\sum_{k\in I_n} \mathrm{D}\big(\wp(p_k - \xi), \tilde{0}\big) \geq \sum_{\substack{k\in I_n \\ \mathrm{D}(\wp(p_k-\xi),\tilde{0})\geq\varepsilon}} \mathrm{D}\big(\wp(p_k - \xi), \tilde{0}\big) \geq \varepsilon \cdot \big|\{k \in I_n : \mathrm{D}(\wp(p_k - \xi), \tilde{0}) \geq \varepsilon\}\big|.$$

Dividing both sides by λ_n and taking the limit as $n \to \infty$,

$$0 = \lim_{n\to\infty} \frac{1}{\lambda_n} \sum_{k\in I_n} \mathrm{D}\big(\wp(p_k - \xi), \tilde{0}\big) \geq \varepsilon \cdot \lim_{n\to\infty} \frac{1}{\lambda_n} \big|\{k \in I_n : \mathrm{D}(\wp(p_k - \xi), \tilde{0}) \geq \varepsilon\}\big|.$$

Because $\varepsilon > 0$ was arbitrary, it follows that

$$\lim_{n\to\infty} \frac{1}{\lambda_n} \big|\{k \in I_n : \mathrm{D}(\wp(p_k - \xi), \tilde{0}) \geq \varepsilon\}\big| = 0,$$

i.e., p_k converges to ξ in the FPS_λ sense. □

Theorem 2. *If a bounded sequence $p = (p_k)$ converges λ-statistically to ξ in a fuzzy paranormed space, and $\lambda = (\lambda_n)$ satisfies $\lambda_n \to \infty$ and $\lambda_n/n \to 0$, then it is $\left[V, \lambda\right]_{FP}$-summable to ξ, which also implies that p is $\left[C, 1\right]_{FP}$-summable to ξ.*

Proof. Boundedness implies that there exists $M > 0$ such that

$$\mathrm{D}\big(\wp(p_k - \xi), \tilde{0}\big) < M$$

for all k. For $\varepsilon > 0$, we have

$$\frac{1}{\lambda_n} \sum_{k\in I_n} \mathrm{D}\big(\wp(p_k - \xi), \tilde{0}\big) \leq \frac{M}{\lambda_n} \big|\{k \in I_n : \mathrm{D}\big(\wp(p_k - \xi), \tilde{0}\big) \geq \varepsilon\}\big| + \varepsilon.$$

The first term vanishes by λ-statistical convergence, proving $\left[V, \lambda\right]_{FP}$-summability. For $\left[C, 1\right]_{FP}$, observe that

$$\frac{1}{n} \sum_{k=1}^{n} \mathrm{D}\big(\wp(p_k - \xi), \tilde{0}\big) = \underbrace{\frac{1}{n} \sum_{k=1}^{n-\lambda_n} \mathrm{D}\big(\wp(p_k - \xi), \tilde{0}\big)}_{\leq \frac{n-\lambda_n}{n} M \to 0} + \underbrace{\frac{1}{n} \sum_{k\in I_n} \mathrm{D}\big(\wp(p_k - \xi), \tilde{0}\big)}_{\leq \frac{\lambda_n}{n} M \to 0}.$$

Both terms vanish, completing the proof. □

Theorem 3. *If a sequence $p = (p_k)$ converges statistically to ξ with respect to the fuzzy paranorm on X, and if*

$$\liminf_{n\to\infty}\left(\frac{\lambda_n}{n}\right) > 0,$$

then p is FPS_λ-convergent to ξ.

Proof. For the given $\varepsilon > 0$, note that the set

$$\{k \le n : \mathrm{D}(\wp(p_k - \xi), \tilde{0}) \ge \varepsilon\}$$

contains the set

$$\{k \in I_n : \mathrm{D}(\wp(p_k - \xi), \tilde{0}) \ge \varepsilon\}.$$

Therefore,

$$\frac{1}{n}\left|\{k \le n : \mathrm{D}(\wp(p_k - \xi), \tilde{0}) \ge \varepsilon\}\right| \ge \frac{1}{n}\left|\{k \in I_n : \mathrm{D}(\wp(p_k - \xi), \tilde{0}) \ge \varepsilon\}\right| \ge \frac{\lambda_n}{n} \cdot \frac{1}{\lambda_n}\left|\{k \in I_n : \mathrm{D}(\wp(p_k - \xi), \tilde{0}) \ge \varepsilon\}\right|.$$

Taking the limit as $n \to \infty$, and using the fact that

$$\liminf_{n\to\infty} \frac{\lambda_n}{n} > 0,$$

we conclude that p converges to ξ in the FPS_λ sense. □

In this paper, unless otherwise specified, the phrase "for all $n \in \mathbb{N}_{n_0}$" means "for all $n \in \mathbb{N}$ except for finitely many positive integers," where

$$\mathbb{N}_{n_0} = \{n_0, n_0 + 1, n_0 + 2, \ldots\}$$

for some fixed $n_0 \in \mathbb{N} = \{1, 2, 3, \ldots\}$.

Theorem 4. *Let $\lambda = (\lambda_n)$ and $\mu = (\mu_n)$ be two sequences in Λ such that for every $n \in \mathbb{N}_{n_0}$, it holds that $\lambda_n \le \mu_n$.*

i. If

$$\liminf_{n\to\infty} \frac{\lambda_n}{\mu_n} > 0, \tag{1}$$

then

$$FPS_\mu \subseteq FPS_\lambda;$$

ii. If

$$\lim_{n\to\infty} \frac{\lambda_n}{\mu_n} = 1, \tag{2}$$

then

$$FPS_\lambda \subseteq FPS_\mu.$$

Proof. (i) Assume that $\lambda_n \le \mu_n$ for every $n \in \mathbb{N}_{n_0}$ and condition (1) holds. Consequently, $I_n \subseteq J_n$, where $I_n = [n - \lambda_n + 1, n]$ and $J_n = [n - \mu_n + 1, n]$.

Thus, for any $\varepsilon > 0$, we have

$$\left|\{k \in J_n : \mathrm{D}(\wp(p_k - \xi), \tilde{0}) \ge \varepsilon\}\right| \ge \left|\{k \in I_n : \mathrm{D}(\wp(p_k - \xi), \tilde{0}) \ge \varepsilon\}\right|.$$

Dividing both sides by μ_n and using $\lambda_n \leq \mu_n$, we get

$$\frac{1}{\mu_n}\left|\{k \in J_n : \mathrm{D}(\wp(p_k - \xi), \tilde{0}) \geq \varepsilon\}\right| \geq \frac{\lambda_n}{\mu_n} \cdot \frac{1}{\lambda_n}\left|\{k \in I_n : \mathrm{D}(\wp(p_k - \xi), \tilde{0}) \geq \varepsilon\}\right|.$$

Taking the limit inferior as $n \to \infty$ and applying (1), it follows that

$$FPS_\mu \subseteq FPS_\lambda.$$

(ii) Let $(p_k) \in FPS_\lambda$ and suppose condition (2) holds. As $I_n \subseteq J_n$, for any $\varepsilon > 0$ we can write

$$\begin{aligned}
&\frac{1}{\mu_n}\left|\{k \in J_n : \mathrm{D}(\wp(p_k - \xi), \tilde{0}) \geq \varepsilon\}\right| \\
&= \frac{1}{\mu_n}\left|\{n - \mu_n + 1 \leq k \leq n - \lambda_n : \mathrm{D}(\wp(p_k - \xi), \tilde{0}) \geq \varepsilon\}\right| \\
&\quad + \frac{1}{\mu_n}\left|\{k \in I_n : \mathrm{D}(\wp(p_k - \xi), \tilde{0}) \geq \varepsilon\}\right| \\
&\leq \frac{\mu_n - \lambda_n}{\mu_n} + \frac{1}{\lambda_n}\left|\{k \in I_n : \mathrm{D}(\wp(p_k - \xi), \tilde{0}) \geq \varepsilon\}\right| \\
&= \left(1 - \frac{\lambda_n}{\mu_n}\right) + \frac{1}{\lambda_n}\left|\{k \in I_n : \mathrm{D}(\wp(p_k - \xi), \tilde{0}) \geq \varepsilon\}\right|.
\end{aligned}$$

Because $\lim_{n\to\infty} \frac{\lambda_n}{\mu_n} = 1$ by (2) and $(p_k) \in FPS_\lambda$, the right-hand side tends to zero as $n \to \infty$.

Therefore,

$$\lim_{n\to\infty} \frac{1}{\mu_n}\left|\{k \in J_n : \mathrm{D}(\wp(p_k - \xi), \tilde{0}) \geq \varepsilon\}\right| = 0,$$

which means $(p_k) \in FPS_\mu$, and so

$$FPS_\lambda \subseteq FPS_\mu.$$

□

5. Conclusions

In this section, we summarize the main results from the theorems presented in the previous section. Additionally, we state important corollaries derived from these results.

Corollary 1. *Let $\lambda = (\lambda_n)$ and $\mu = (\mu_n)$ be two sequences in Λ such that $\lambda_n \leq \mu_n$ for all sufficiently large n. If*

$$\lim_{n\to\infty} \frac{\lambda_n}{\mu_n} = 1, \tag{3}$$

then

$$FPS_\lambda = FPS_\mu.$$

Corollary 2. *Let $\lambda = (\lambda_n) \in \Lambda$ satisfy $\lambda_n \leq n$ for all sufficiently large n and*

$$\lim_{n\to\infty} \frac{\lambda_n}{n} = 1.$$

Then, $FPS_\lambda = FPS$.
This result follows immediately from Corollary 1 by taking $\mu_n = n$.

Theorem 5. *Let $\lambda = (\lambda_n)$ and $\mu = (\mu_n)$ be elements of Λ such that $\lambda_n \leq \mu_n$ for all sufficiently large n.*

1. *If condition* (1) *holds, then*
$$[V,\mu]_{FP} \subseteq [V,\lambda]_{FP};$$
2. *If condition* (2) *holds, then*
$$\ell_\infty \cap [V,\lambda]_{FP} \subseteq [V,\mu]_{FP}.$$

Proof. (i) Because $\lambda_n \leq \mu_n$ for all sufficiently large n, it follows that $I_n = [n-\lambda_n+1, n]$ is a subset of $J_n = [n-\mu_n+1, n]$. Therefore, for all such n,

$$\frac{1}{\mu_n}\sum_{k\in J_n} \mathrm{D}\big(\wp(p_k-\xi),\tilde{0}\big) \geq \frac{1}{\mu_n}\sum_{k\in I_n} \mathrm{D}\big(\wp(p_k-\xi),\tilde{0}\big).$$

Taking the limit as $n\to\infty$ and applying condition (1), we conclude that

$$[V,\mu]_{FP} \subseteq [V,\lambda]_{FP}.$$

(ii) Let $p=(p_k)\in \ell_\infty \cap [V,\lambda]_{FP}$ so there exists $M>0$ such that

$$\mathrm{D}\big(\wp(p_k-\xi),\tilde{0}\big) \leq M \quad \text{for all } k.$$

Because $\lambda_n \leq \mu_n$, we have $\frac{1}{\mu_n} \leq \frac{1}{\lambda_n}$, and $I_n \subseteq J_n$ for all large n. Hence,

$$\begin{aligned}\frac{1}{\mu_n}\sum_{k\in J_n} \mathrm{D}\big(\wp(p_k-\xi),\tilde{0}\big) &= \frac{1}{\mu_n}\sum_{k\in J_n\setminus I_n} \mathrm{D}\big(\wp(p_k-\xi),\tilde{0}\big) + \frac{1}{\mu_n}\sum_{k\in I_n} \mathrm{D}\big(\wp(p_k-\xi),\tilde{0}\big)\\ &\leq \frac{\mu_n-\lambda_n}{\mu_n}\cdot M + \frac{1}{\mu_n}\sum_{k\in I_n} \mathrm{D}\big(\wp(p_k-\xi),\tilde{0}\big)\\ &\leq \left(1-\frac{\lambda_n}{\mu_n}\right)\cdot M + \frac{1}{\lambda_n}\sum_{k\in I_n} \mathrm{D}\big(\wp(p_k-\xi),\tilde{0}\big).\end{aligned}$$

By condition (2), the first term tends to zero as $n\to\infty$, and because $p\in[V,\lambda]_{FP}$, the second term also tends to zero. (Note that $1-\frac{\lambda_n}{\mu_n}\geq 0$ for all sufficiently large n.)

Therefore,
$$\ell_\infty \cap [V,\lambda]_{FP} \subseteq [V,\mu]_{FP},$$
which implies
$$\ell_\infty \cap [V,\lambda]_{FP} \subseteq \ell_\infty \cap [V,\mu]_{FP}.$$

□

Corollary 3. *Let $\lambda,\mu\in\Lambda$ be such that $\lambda_n\leq\mu_n$ for all sufficiently large $n\in\mathbb{N}_{n_0}$. If condition* (2) *holds, then*
$$\ell_\infty \cap [V,\lambda]_{FP} = \ell_\infty \cap [V,\mu]_{FP}.$$

Theorem 6. *Let $\lambda,\mu\in\Lambda$ such that $\lambda_n\leq\mu_n$ for all $n\in\mathbb{N}_{n_0}$.*

1. *If condition* (1) *holds, then*
$$p_k\to\xi \text{ in } [V,\mu]_{FP} \implies p_k\to\xi \text{ in } FPS_\lambda,$$
and the inclusion $[V,\mu]_{FP}\subset FPS_\lambda$ is strict for some $\lambda,\mu\in\Lambda$;
2. *If $(p_k)\in\ell_\infty$ and $p_k\to\xi$ in FPS_λ, then*
$$p_k\to\xi \text{ in } [V,\mu]_{FP},$$
whenever condition (2) *holds.*

Proof. (i) Let $\varepsilon > 0$ and suppose p_k converges to ξ in $[V, \mu]_{FP}$. For every $n \in \mathbb{N}_{n_0}$, we have

$$\sum_{k \in J_n} \mathrm{D}\big(\wp(p_k - \xi), \tilde{0}\big) \geq \sum_{k \in I_n} \mathrm{D}\big(\wp(p_k - \xi), \tilde{0}\big) \geq \varepsilon \big| \{k \in I_n : \mathrm{D}(\wp(p_k - \xi), \tilde{0}) \geq \varepsilon\} \big|.$$

Taking limits as $n \to \infty$ and applying (1), it follows that $p_k \to \xi$ in FPS_λ.
To show strictness of the inclusion, take $\lambda_n = \sqrt{n}$, $\mu_n = n$ and define

$$p_k = \begin{cases} k, & k = n^2, \\ 0, & \text{otherwise.} \end{cases}$$

Then, clearly $p_k \to 0$ in FPS_λ, but

$$\frac{1}{\mu_n} \sum_{j \in J_n} \mathrm{D}\big(\wp(p_k - 0), \tilde{0}\big) = \frac{1}{n} \sum_{k=1}^{n} \wp(p_k)_0^+ = \frac{1}{n}\Big(1 + 0 + 0 + 4 + \cdots + [\sqrt{n}]^2\Big).$$

Using the formula for sum of squares,

$$1^2 + 2^2 + \cdots + m^2 = \frac{m(m+1)(2m+1)}{6},$$

we get

$$\frac{1}{n} \sum_{k=1}^{n} \wp(p_k)_0^+ = \frac{1}{n} \cdot \frac{[\sqrt{n}]([\sqrt{n}]+1)(2[\sqrt{n}]+1)}{6}.$$

Because $\sqrt{n} < [\sqrt{n}] + 1$ and $\sqrt{n} < 2[\sqrt{n}] + 1$, it follows that

$$\frac{1}{n} > \frac{1}{([\sqrt{n}]+1)(2[\sqrt{n}]+1)},$$

hence

$$\frac{1}{\mu_n} \sum_{j \in J_n} \mathrm{D}\big(\wp(p_k - 0), \tilde{0}\big) > \frac{[\sqrt{n}]}{6} \to \infty,$$

so $p \notin [V, \mu]_{FP}$.

(ii) Suppose $p_k \to \xi$ in FPS_λ and $p = (p_k) \in \ell_\infty$, so there exists $M > 0$ such that $\mathrm{D}(\wp(p_k - \xi), \tilde{0}) \leq M$ for all k. Because $\frac{1}{\mu_n} \leq \frac{1}{\lambda_n}$ and $I_n \subseteq J_n$, for each $\varepsilon > 0$ and every $n \in \mathbb{N}_{n_0}$,

$$\begin{aligned} \frac{1}{\mu_n} \sum_{k \in J_n} \mathrm{D}(\wp(p_k - \xi), \tilde{0}) &= \frac{1}{\mu_n} \sum_{k \in J_n \setminus I_n} \mathrm{D}(\wp(p_k - \xi), \tilde{0}) + \frac{1}{\mu_n} \sum_{k \in I_n} \mathrm{D}(\wp(p_k - \xi), \tilde{0}) \\ &\leq \frac{\mu_n - \lambda_n}{\mu_n} M + \frac{1}{\mu_n} \sum_{k \in I_n} \mathrm{D}(\wp(p_k - \xi), \tilde{0}) \\ &\leq \left(1 - \frac{\lambda_n}{\mu_n}\right) M + \frac{1}{\lambda_n} \sum_{k \in I_n} \mathrm{D}(\wp(p_k - \xi), \tilde{0}) \\ &\leq \left(1 - \frac{\lambda_n}{\mu_n}\right) M + \frac{1}{\lambda_n} \sum_{\substack{k \in I_n \\ \mathrm{D}(\wp(p_k - \xi), \tilde{0}) \geq \varepsilon}} \mathrm{D}(\wp(p_k - \xi), \tilde{0}) + \varepsilon. \end{aligned}$$

Taking the limit as $n \to \infty$ and applying (2), the right side tends to zero as $p_k \to \xi$ in FPS_λ.

Thus,

$$\ell_\infty \cap FPS_\lambda \subseteq [V, \mu]_{FP}.$$

By setting $\mu_n = n$ for all $n \in \mathbb{N}_{n_0}$, we obtain subsequent results. Note that

$$\lim_{n\to\infty} \frac{\lambda_n}{\mu_n} = 1 \implies \lim_{n\to\infty} \inf \frac{\lambda_n}{\mu_n} = 1 > 0,$$

that is, (2) implies (1). □

Corollary 4. *Let* $\lim_{n\to\infty} \frac{\lambda_n}{\mu_n} = 1$. *Then:*

1. *If* $(p_k) \in \ell_\infty$ *and* $p_k \to \xi$ *in* FPS_λ, *then* $p_k \to \xi$ *in* $[C,1]_{FP}$;
2. *If* $p_k \to \xi$ *in* $[C,1]_{FP}$, *then* $p_k \to \xi$ *in* FPS_λ.

Example 1. *In Çınar et al.'s [23] definition, let* $X = \mathbb{R}$ *with zero vector* $\theta = 0$ *and define*

$$L(a,b) = \min\{a,b\}, \quad R(a,b) = \max\{a,b\}.$$

For each $x \in \mathbb{R}$, *set its* α*-cut by*

$$[\wp(x)]_\alpha = \left[\frac{|x|}{1+\alpha}, \frac{|x|}{1-\alpha}\right], \quad \alpha \in (0,1).$$

Moreover, define the 0-cut as the limit

$$[\wp(x)]_0 := \lim_{\alpha\to 0^+} [\wp(x)]_\alpha = [|x|,|x|].$$

Then, $(\mathbb{R}, \wp, L, R)$ *is a fuzzy paranormed space.*
Consider the sequence

$$p_k = \begin{cases} 1, & k = n^2 \text{ for some } n \in \mathbb{N}, \\ 0, & \text{otherwise.} \end{cases}$$

The 0-level function of the sequence (p_k) *is given by*

$$\wp_0^+(p_k) = \sup\{t \geq 0 : t \in [\wp(p_k)]_0\} = |p_k| = \begin{cases} 1, & k = n^2 \text{ for some } n \in \mathbb{N}, \\ 0, & \text{otherwise.} \end{cases}$$

Let $\lambda_n = \lfloor \sqrt{n} \rfloor$ *and*

$$I_n = \{k \in \mathbb{N} : n - \lambda_n < k \leq n\}.$$

Then, for any $\varepsilon > 0$,

$$\left|\{k \in I_n : \wp_0^+(p_k) \geq \varepsilon\}\right| = \left|\{k \in I_n : p_k = 1\}\right| \leq 1,$$

so

$$\frac{\left|\{k \in I_n : \wp_0^+(p_k) \geq \varepsilon\}\right|}{\lambda_n} \leq \frac{1}{\lambda_n} \to 0 \quad \text{as } n \to \infty.$$

Hence, (p_k) *is* λ*-statistically convergent to 0 in* $(\mathbb{R}, \wp, L, R)$, *but it does not converge to 0 in the classical sense.*

As illustrated in Figure 1, the sequence p_k takes the value 1 exactly at perfect square indices and 0 elsewhere. This highlights the "sparse shock" character of the example: infinitely many isolated peaks separated by arbitrarily large gaps.

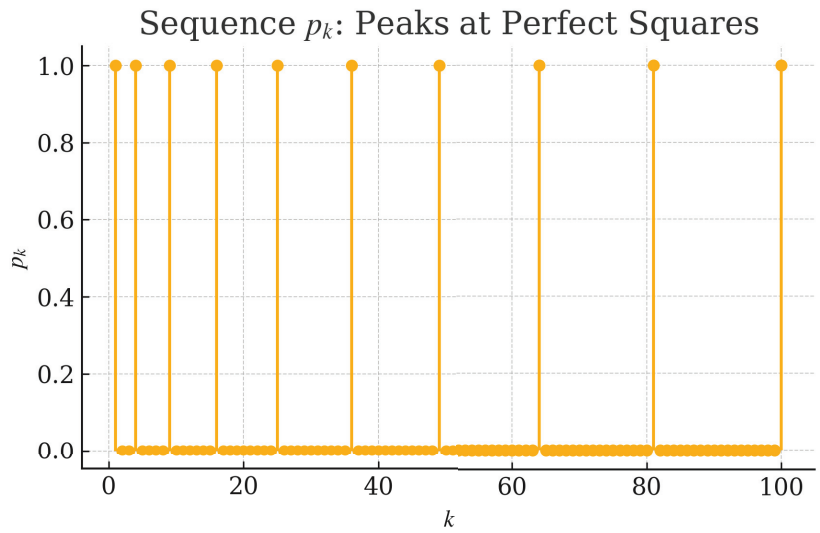


Figure 1. Sequence p_k with peaks at perfect squares.

Figure 2 depicts the theoretical upper bound $1/\lambda_n$ as a function of n. This monotone decrease to zero confirms the theoretical upper bound property, ensuring that

$$\frac{1}{\lambda_n}\left|\{k \le n : p_k \ge \varepsilon\}\right| \le \frac{1}{\lambda_n} \to 0,$$

for every fixed $\varepsilon > 0$.

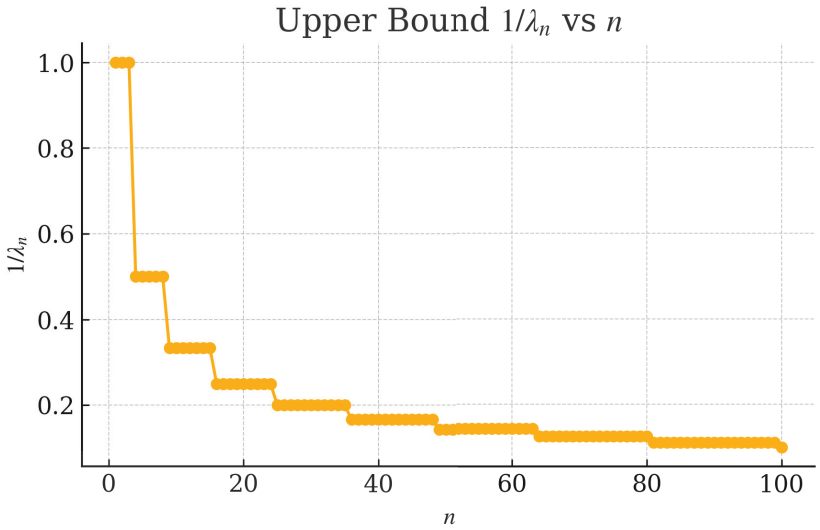


Figure 2. Upper bound $1/\lambda_n$ versus n.

Example 2 (Fuzzy inventory planning)**.** *Consider a retailer forecasting period-by-period demand modeled as a triangular fuzzy number:*

$$D_k = (80, 100, 120) \text{ units}.$$

The adaptive inventory policy is defined as follows:

1. *Initial stock:* $Q_1 = 100$;
2. *For each period* k*, observe actual demand* $D_k^{(\text{act})}$ *and compute the shortfall:*

$$p_k = \max\{0, D_k^{(\text{act})} - Q_k\};$$

3. *Update inventory with learning rate* $t = 0.5$*:*

$$Q_{k+1} = Q_k + 0.5\, p_k.$$

A 10-period simulation reveals only two shortfalls exceeding 5 units. Defining $\lambda_n = \lfloor n/2 \rfloor$*, we observe*

$$\frac{1}{\lambda_n}\left|\{k \le \lambda_n : p_k > 5\}\right| \to 0 \quad \text{as } n \to \infty,$$

which implies that the shortfall sequence $\{p_k\}$ *λ-statistically converges to zero.*

Figure 3 shows period-by-period shortfalls (orange circles) alongside the 5-unit threshold (dashed line). The shaded region indicates the fuzzy demand bounds (80–120 units). Notice that only periods 1, 4, and 7 exceed the threshold, with rapid correction thereafter.

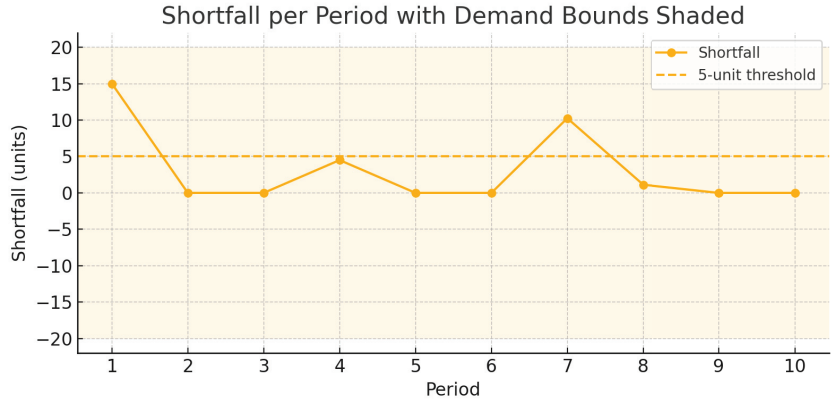


Figure 3. Period-by-period shortfalls with 5-unit threshold (dashed) and fuzzy demand bounds (shaded).

Figure 4 presents the shortfall trajectory (orange line, left axis) and cumulative holding cost (green dashed line, right axis; unit cost = 2). The inset zooms in on periods 13–20, highlighting the impact of a demand shock around $n = 15$*.*

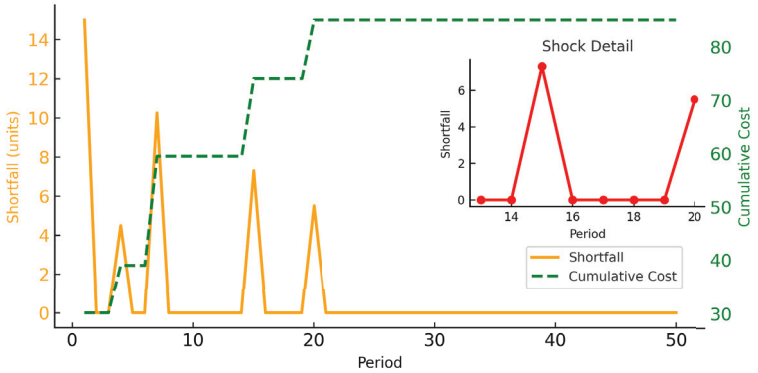


Figure 4. Shortfall vs. cumulative cost over 50 periods. Inset focuses on demand shock near $n = 15$.

Verification of Convergence Conditions:

Extending the simulation to 50 periods, the λ-statistical convergence condition, expressed probabilistically, is:

$$\lim_{n\to\infty} P\left(\frac{1}{\lambda_n}\sum_{k=1}^{\lambda_n} \mathbf{1}_{\{p_k>5\}} \ge \varepsilon\right) = 0, \quad \forall \varepsilon > 0.$$

In the simulation,

$$\max_{n>30} s_n \leq 2.1, \quad \frac{1}{25}\left|\{k \leq 25 : s_k > 5\}\right| = 0.08.$$

Figure 5 illustrates the long-term shortfall behavior over 50 periods. The red dashed line marks the 5-unit threshold; the shaded band shows fuzzy demand bounds. The dotted vertical line at $n = 15$ indicates a significant demand shock. After period 20, all shortfalls remain below the threshold, confirming λ-statistical convergence in practice.

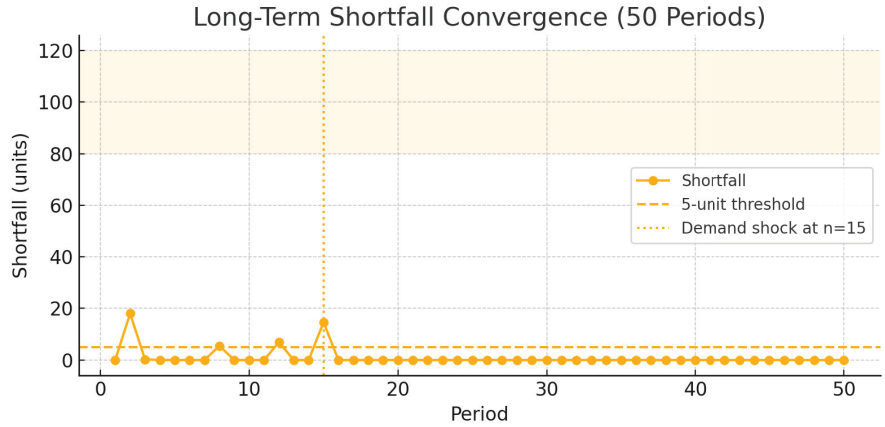


Figure 5. Long-term shortfall convergence over 50 periods. Red dashed line: 5-unit threshold; shaded band: fuzzy demand bounds (80–120). Vertical dotted line at $n = 15$ marks a demand shock.

In addition, the adaptive system satisfies standard convergence properties:

1. *Contraction property:* $\mathbb{E}[|s_{k+1}| \mid \mathcal{F}_k] \leq |s_k| - \eta$ *for some* $\eta > 0$;
2. *Martingale stability:* $\sum_{k=1}^{\infty} P(s_k > \epsilon) < \infty$ *for each* $\epsilon > 5$;
3. *Parameter robustness: Convergence holds for learning rates* $t \in (0.3, 0.7)$.

In practice, serial correlation is removed by pre-processing, and λ-statistical convergence is analyzed on the residuals.

Remark 1. *The λ-statistical convergence of the shortfall sequence depends solely on the finiteness of $D(\wp(\cdot), \tilde{0})$; hence, it remains valid if the triangular demand numbers are replaced by any bounded fuzzy profile, such as trapezoidal or Gaussian-shaped membership functions.*

5.1. Key Contributions and Practical Implications

Beyond theoretical advances, our work demonstrates practical value for adaptive inventory optimization under uncertain, shock-driven demand. By applying λ-statistical convergence within a fuzzy paranorm framework, inventory policies can dynamically adjust safety stocks based on observed demand fluctuations, thereby improving service levels while controlling holding costs. This bridges rigorous convergence theory with real-world replenishment strategies and lays the groundwork for further empirical validation.

5.2. Future Work

Classical EOQ (economic order quantity) and bullwhip mitigation rules typically rely on deterministic or linear forecasting assumptions. In contrast, the λ-statistical approach specifically addresses fuzzy, shock-driven demand. Consequently, a quantitative comparison of costs and service levels necessitates a dedicated simulation study with fully specified cost parameters and lead time structures, which we plan to undertake in future research.

Building on the current λ-statistical framework, we will explore two broader convergence concepts in fuzzy paranormed spaces:

(i) Lacunary statistical convergence, allowing the gaps between index blocks to grow super-linearly, thus isolating highly clustered, sporadic demand shocks that may be obscured by uniform windowing;
(ii) Ideal convergence, permitting the exclusion of a pre-specified negligible subset of indices (e.g., holiday blackouts or data outages), thereby enhancing robustness against irregular disruptions.

These extensions promise a more flexible toolkit for modeling real-world demand volatility and will be pursued in future empirical work.

Author Contributions: Conceptualization, H.Ö. and M.R.T.; formal analysis, H.Ö. and M.R.T.; writing—original draft preparation, H.Ö. and M.R.T.; writing—review and editing, H.Ö. and M.R.T.; visualization, H.Ö. and M.R.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: The authors are grateful to the responsible editor and the anonymous reviewers for their valuable comments and suggestions, which have greatly improved this paper.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Steinhaus, H. Sur la convergence ordinaire et la convergence asymptotique. *Colloq. Math.* **1951**, *2*, 73–74.
2. Fast, H. Sur la convergence statistique. *Colloq. Math.* **1951**, *2*, 241–244. [CrossRef]
3. Fridy, J.A. On statistical convergence. *Analysis* **1985**, *5*, 301–314. [CrossRef]
4. Karakaya, V.; Chishti, T.A. Weighted statistical convergence. *Iran. J. Sci. Technol. Trans. A Sci.* **2009**, *33*, 210–223.
5. Cakalli, H. A study on statistical convergence. *Funct. Anal. Approx. Comput.* **2009**, *1*, 19–24.
6. Aytar, S. Rough statistical convergence. *Numer. Funct. Anal. Optim.* **2008**, *29*, 291–303. [CrossRef]
7. Maddox, I.J. Statistical convergence in a locally convex space. *Math. Proc. Camb. Philos. Soc.* **1988**, *104*, 141–145. [CrossRef]
8. Belen, C.; Mohiuddine, S.A. Generalized weighted statistical convergence and application. *Appl. Math. Comput.* **2013**, *219*, 9821–9826. [CrossRef]
9. Et, M.; Nuray, F. Delta (m)-Statistical convergence. *Indian J. Pure Appl. Math.* **2001**, *32*, 935–939.
10. Nuray, F.; Rhoades, B.E. Statistical convergence of sequences of sets. *Fasc. Math.* **2012**, *49*, 87–99.
11. Karakus, S. Statistical convergence on probalistic normed spaces. *Math. Commun.* **2007**, *12*, 11–23.
12. Nuray, F.; Savaş, E. Statistical convergence of sequences of fuzzy numbers. *Math. Slovaca* **1995**, *45*, 269–273. [CrossRef]
13. Alotaibi, A.; Alroqi, A.M. Statistical convergence in a paranormed space. *J. Inequalities Appl.* **2012**, *2012*, 1–6. [CrossRef]
14. Karakus, S.; Demirci, K.; Duman, O. Statistical convergence on intuitionistic fuzzy normed spaces. *Chaos Solitons Fractals* **2008**, *35*, 763–769. [CrossRef]
15. Mursaleen, M.; Edely, O.H. Generalized statistical convergence. *Inf. Sci.* **2004**, *162*, 287–294. [CrossRef]
16. Bilalov, B.; Nazarova, T. On statistical convergence in metric spaces. *J. Math. Res.* **2015**, *7*, 37–44. [CrossRef]
17. Şençimen, C.; Pehlivan, S. Statistical convergence in fuzzy normed linear spaces. *Fuzzy Sets Syst.* **2008**, *159*, 361–370. [CrossRef]
18. Savas, E.; Gürdal, M. A generalized statistical convergence in intuitionistic fuzzy normed spaces. *Sci. Asia* **2015**, *41*, 289–294. [CrossRef]
19. Yıldırım, E.N. Statistical Convergence of Matrix Sequences. *Konuralp J. Math.* **2024**, *12*, 74–79.
20. Yapalı, R.; Çoşkun, H.; Gürdal, U. Statistical convergence on L-fuzzy normed space. *Filomat* **2023**, *37*, 2077–2085. [CrossRef]
21. Alotaibi, A.; Mursaleen, M. Statistical convergence in random paranormed space. *J. Comput. Anal. Appl.* **2014**, *17*, 297–304.
22. Çınar, M.; Karakaş, M.; Et, M. On the statistical convergence of type in paranormed spaces. *J. Interdiscip. Math.* **2022**, *25*, 323–334. [CrossRef]
23. Çınar, M.; Et, M.; Karakaş, M. On fuzzy paranormed spaces. *Int. J. Gen. Syst.* **2023**, *52*, 61–71. [CrossRef]
24. Connor, J.S. The statistical and strong Cesaro convergence of sequences. *Analysis* **1988**, *8*, 47–63. [CrossRef]
25. Türkmen, M.R. On I_{θ_2}-convergence in fuzzy normed spaces. *J. Inequalities Appl.* **2020**, *2020*, 127. [CrossRef]
26. Jasrotia, S.; Singh, U.; Raj, K. Applications of statistical convergence of order $(\eta, \delta + \gamma)$ in difference sequence spaces of fuzzy numbers. *J. Intell. Fuzzy Syst.* **2021**, *40*, 4695–4703. [CrossRef]

27. Solomon, J.; Chawla, M. Weighted statistical convergence of order $\tilde{\alpha}$ for double sequences in paranormed spaces. *J. Interdiscip. Math.* **2024**, *27*, 1875–1885. [CrossRef]
28. Türkmen, M.R.; Akbay, H.B. λ-Statistical Convergence in Fuzzy n-Normed Linear Spaces. *J. Math. Anal.* **2024**, *15*, 1–14. [CrossRef]
29. Granados, C.; Das, A.K.; Osu, B.O. $M\lambda m, n, p$-statistical convergence for triple sequences. *J. Anal.* **2022**, *30*, 451–468. [CrossRef]
30. Granados, C.; Das, A.K.; Das, S. New Tauberian theorem for statistical Cesáro summable triple sequences of fuzzy numbers. *Kragujev. J. Math.* **2021**, *48*, 787–802. [CrossRef]
31. Mursaleen, M. λ-statistical convergence. *Math. Slovaca* **2000**, *50*, 111–115.
32. Saranya, N.; Suja, K. λ-Statistical Convergence in Paranormed Spaces over Non-Archimedean Fields. *IAENG Int. J. Appl. Math.* **2023**, *53*, 257–262.
33. Karakaş, A.; Altın, Y.; Altınok, H. On generalized statistical convergence of order β of sequences of fuzzy numbers. *J. Intell. Fuzzy Syst.* **2014**, *26*, 1909–1917. [CrossRef]
34. Türkmen, M.R.; Çınar, M. λ-statistical convergence in fuzzy normed linear spaces. *J. Intell. Fuzzy Syst.* **2018**, *34*, 4023–4030. [CrossRef]
35. Karakaya, V.; Şimşek, N.; Ertürk, M.; Gürsoy, F. λ-Statistical Convergence of Sequences of Functions in Intuitionistic Fuzzy Normed Spaces. *J. Funct. Spaces* **2012**, *2012*, 926193.
36. Alghamdi, M.A.; Mursaleen, M. λ-Statistical Convergence in Paranormed Space. *Abstr. Appl. Anal.* **2013**, *2013*, 264520. [CrossRef]
37. Et, M.; Çınar, M.; Karakaş, M. On λ-statistical convergence of order β of sequences of function. *J. Inequalities Appl.* **2013**, *2013*, 204. [CrossRef]
38. Çolak, R. On λ-statistical convergence. In Proceedings of the Conference on Summability and Applications, Istanbul, Turkey, 18–22 May 2011.
39. Yaying, T.; Başar, F. A Study on Some Paranormed Sequence Spaces Due to Lambda–Pascal Matrix. *Acta Sci. Math.* **2024**, *91*, 161–180. [CrossRef]
40. Karakuş, M.; Başar, F. Vector Valued Closed Subspaces and Characterizations of Normed Spaces through J'-Summability. *Indian J. Math.* **2024**, *66*, 85–105.
41. Yeşilkayagil, M.; Başar, F. On the Paranormed Space of Bounded Variation Double Sequences. *Bull. Malays. Math. Sci. Soc.* **2020**, *43*, 2701–2712. [CrossRef]
42. Felbin, C. Finite-dimensional fuzzy normed linear space. *Fuzzy Sets Syst.* **1992**, *48*, 239–248. [CrossRef]
43. Xiao, J.; Zhu, X. On linearly topological structure and property of fuzzy normed linear space. *Fuzzy Sets Syst.* **2002**, *125*, 153–161. [CrossRef]
44. Kaleva, O.; Seikkala, S. On fuzzy metric spaces. *Fuzzy Sets Syst.* **1984**, *12*, 215–229. [CrossRef]

mathematics

MDPI

Article

Optimizing Inventory for Imperfect and Gradually Deteriorating Items Under Multi-Level Trade Credit in a Sustainable Supply Chain

Abhay Bansal [1], Aastha Panwar [2], Bhuvan Unhelkar [3] and Mandeep Mittal [1,*]

[1] School of Computer Science Engineering and Technology, Bennett University, Greater Noida 201310, India; abhay.bansal@bennett.edu.in
[2] School of Basic Sciences, Galgotias University, Greater Noida 203201, India; aasthapanwar7417@gmail.com
[3] Muma College of Business, University of South Florida, N. Tamiami Trail, Sarasota, FL 34243, USA; bunhelkar@usf.edu
* Correspondence: mittal_mandeep@yahoo.com

Abstract: Reducing carbon emissions is of immense interest to most modern organizations striving for sustainability. Effective inventory management is crucial for achieving resource optimization and minimizing environmental impact. Very little work has been conducted up to this point on slowly declining, low-quality products with multi-level trade credit rules under the influence of carbon emissions. In this study, an inventory model is tailored specifically for imperfect and gradually deteriorating products with a multi-level trade credit policy. Further, the impact of carbon emissions on the retailer's ordering strategies is also considered. To determine the optimal policy for supply chain partners, three trade credit instances with seven subcases are taken into consideration. To choose the best scenario out of ten cases, an algorithm is also developed. The model's validity is illustrated through a numerical experiment and sensitivity analysis. This study is an innovative approach to balancing economic trade credit policy in sustainable supply chain management.

Keywords: sustainable supply chain; carbon emission reduction; inventory optimization; non-instantaneous deterioration; trade credit policies; defective product management introduction

MSC: 90B06

1. Introduction

Balancing economic and environmental goals is a significant challenge in practice. Business organizations are continuously undertaking this balancing act to offer competitive products while ensuring lower carbon emissions. Incorporating carbon emission costs and trade credit policies in inventory management results in efficiency.

In the contemporary business landscape, stakeholders must collaborate to enhance processes to reduce carbon emissions while providing financial support to achieve environmental and economic sustainability. While holding lower inventory levels is often perceived as a strategy for economic sustainability, it can lead to stockouts and dissatisfied customers. Inventory management strategies need to carefully balance these dynamics, particularly when dealing with goods categorized by their shelf life. Extended usability allows non-perishable items to satisfy demand over longer periods of time without suffering appreciable quality deterioration. Perishable commodities, on the other hand, pose difficulties because their quality and value decline with time.

Perishable goods can be further classified into two categories: those with constant utility throughout their lifespan, such as blood (e.g., a static 21-day usability period) and certain medications, and those with utility that decreases exponentially, such as fruits and fish. This distinction is critical for devising effective inventory management strategies that align with sustainability goals while ensuring customer satisfaction [1]. As a result, research is shifting focus toward enhancing the inventory models for products with deteriorating quality [2,3].

Interestingly though, the collaborative environment of modern business is not conducive for it to approach sustainability as a single entity. Cooperation between different business players and their mutual financial understanding is essential for economic sustainability. Such collaboration can also improve profits and financial benefits. Marketing practices include suppliers granting retailers an acceptable payment delay and retailers granting customers an acceptable payment delay, along with the settlement of any outstanding balances within a predetermined time frame known as a trade credit period [4]. The ecological dimensions with returnable transport items and remanufacturing were evaluated by Ref. [5]. This policy ensures an efficient trade balance, maximizing profits for both trade credit policy partners. The trade credit also makes money in interest and revenue from items sold while being able to protect the buyer in case of insolvency from the retailer. It is in the seller's best interest not to receive prepaid payments or to pay interest within a given period. The final flow of inventory payment can be illustrated as follows: the supplier receives full value for their products, for example, the supplier is a business that provides inventory to the buyer in exchange for cash.

The supplier frequently collects interest on their cash balance [6]. The supplier pays out some of their interest to intermediate sellers on their cash balance. The intermediate sellers then pay out. If a person only makes minimum payments on their balance and does not pay off the entire account balance every month, there will be interest charges for all unpaid interest. If a person pays off 90% of their balance before the expiration date and only owes 10% inside the boundaries of the allowable holdup period, then they will not be billed with any interest. An increasing focus on sustainable development and environmental concerns make inventory management a major research topic. In order to provide retailers with the best ordering methods under multi-level trade credit regulations while lowering the cost of carbon emissions, an inventory model is therefore suggested. With the use of this modeling, businesspeople can reduce their carbon footprint and increase their profits.

A model has been created that focuses on the inventory management of a retailer by optimizing the ordering quantity and time, enabling the retailer to smoothly run their business. This model focuses on trade credit policy between the supply chain partners. By this method, the total holding cost per unit item will get reduced; this will maximize the total profit of the retailers.

This study introduces a novel inventory model integrating multi-level trade credit policies for imperfect and deteriorating products to optimize profitability and sustainability. Unlike traditional models, it incorporates flexible credit structures based on product quality, improving cash flow and reducing financial risk. By minimizing excess inventory and waste, the model aligns economic and environmental goals, lowering carbon emissions. It provides a quantitative framework to balance trade credit, deterioration, and profitability. Validated through numerical experiments and sensitivity analysis, it offers a practical solution for sustainable inventory management. This innovative approach enhances decision making in modern supply chains.

All of this basic foundation is briefly covered in Section 1, and then the literature of the model based on the current model is presented in Section 2. Notations and assumptions that are used in the model are presented in Section 3. Section 4 completes the mathematical

portion, and Section 5 deals with the solution procedure. Section 6 provides an explanation of the numerical examples used to support the correctness of this concept. A sensitivity analysis presented in Section 7, and Section 8 includes the final reflections.

2. Literature Review

The literature is related to the following points: carbon emission in the inventory mode trade credit effect on the inventory model and imperfect products in the inventory models.

2.1. Carbon Emission in Inventory Models

Carbon emission happens because of the storage and transporting of items. Carbon emissions occur because of the process of storing and transporting deteriorating products. Emissions in the warehouse are based on the inventory, as well as how much energy is used per unit. The disposal of items that are no longer needed is also a major contributor to carbon emission. For this problem, an inventory model is generated to balance global effects to meet the requirements without affecting the environment. Refs. [7,8] generated a model in which the lead time is taken as a stochastic incorporation with the production of defective products with a credit financing scheme. Ref. [9] developed a sustainability issue with a trade credit policy in their model. They have also worked on minimizing the greenhouse gas emission. Ref. [10] collaborated upon the carbon emission of defective products in their inventory model. Ref. [11] optimized the shipment amount in their model. They have also worked on the defective production rate. Ref. [12] generated a multi-stage formation model for defective items with a reducing carbon emission. Ref. [13] worked on reducing carbon output with the impact of credit financing on a multi-echelon inventory model. Ref. [14] generated a model using environmental and social strategies for the consumption of future energy. A sustainable inventory model was generated by [15]. This model worked on reducing carbon emission and. This model is also worked on non-instantaneous deteriorating products governed by a multi-level credit term policy.

2.2. Non-Instantaneous Degrading Items in Inventory Management Models

In traditional inventory models, the instantaneous deterioration of products upon delivery is often assumed. However, real-world scenarios reveal a delay between receipt and decay. This lag, known as non-instantaneous deterioration, has gained attention since 2006. Ref. [16] designed the structure of the inventory system allowing delayed payments to be available to benefit such products. Ref. [17] proposed a price-dependent demand model under trade credit facilities for non-instantaneous decaying commodities. Recent studies expanded on this concept [18], creating a production model incorporating region-specific, population-driven, and price-based demand. Ref. [19] introduced a multilevel trade credit policy for non-instantaneous deterioration, considering promotional and selling price influences on demand. In the case of non-instantaneously degrading items with a cost and time demand based on unpredictable conditions, Ref. [20] developed a two-warehouse EOQ model.

2.3. Trade Credit in Inventory Models

Delayed payments serve companies with a better access to the capital, and in such a process, they can empower the suppliers and incentivize a greater order size. Therefore, in this way, this policy ensures continuous financing. The theory of credit financing is different from business to business.

Research on credit financing inventory models has evolved significantly. Ref. [21] laid the groundwork by developing a model that ignored the distinction between purchase cost and selling price. Subsequent studies expanded on this concept. Ref. [22], for

example, presented a two-level credit financing strategy for non-immediately degrading commodities. A model for degrading items with a two-level credit timeframe was created by [23], taking into account finite time horizons and stock-dependent demand. Ref. [24] made further progress by establishing the best policies under allowable payment delays for degrading products with multivariate demand rates. Optimal trade credit policies for perishable goods were investigated by [25], taking into consideration an imperfect supply and demand that is dependent on stocks. Additionally, they presented the idea of a twofold wait time as a way to reward retailers. Ref. [26] looked at situations in which suppliers are paid a portion of the whole order amount right away and the remaining amount after a predetermined amount of time. This idea was expanded upon by Ref. [27] to include multi-level trade credit, in which suppliers provide merchants grace periods, which retailers subsequently give to customers. A model with variable degradation under credit financing was developed by Ref. [28]. The joint trade credit problem for deteriorating inventory systems presented by Ref. [29] and the model for determining the economic order quantity with flexible trade credit put forth by Ref. [30] are recent contributions. A model for degrading items using advance payment methods was proposed by Ref. [31]. A model for credit financing policy was created by Ref. [32] in a two-stockroom setting with an ambiguous deterioration and demand modeled by the Weibull distribution.

2.4. Imperfect Within Inventory Models

During the manufacturing process, most items are manufactured with certain imperfections. The retailer has to deliver a good quality of items; therefore, it is important to identify this imperfection via an inspection process. In fact, researchers should focus on the imperfect quality of items as it is not possible to produce all goods of a perfect quality. Furthermore, this problem has been focused upon by many academicians in the past years.

Ref. [33] laid the groundwork with an inventory system aimed at establishing optimal ordering guidelines for defective items. Imperfect production processes were later accounted for by Ref. [34] through a generated model. Ref. [35] addressed a problem where item demand depends on production units, incorporating a screening process to eliminate defective items. Ref. [36] modified this model to include shortage situations and backordering costs. As the area continues to evolve, Ref. [37] developed an inventory model that included learning effects for damaged products. Ref. [38] examined the usage of two storage facilities for damaged commodities after developing an inventory model for credit financing schemes in [39]. A multiple-production stock management model designed especially for damaged products was developed by Ref. [40]. Ref. [41] presented an inventory model for low-quality products that took shortages into consideration, leading to the creation of an inventory model for defective objects. A multi-repository model with shortages and discount plans for defective commodities was most recently introduced by Ref. [42].

3. Problem Description, Notation, and Assumptions

3.1. Problem Description

This research addresses the issue of multi-level trade credit policies in the context of non-instantaneously degrading products, with the goal of reducing the carbon emissions associated with storage. Recognizing the inherent challenges in achieving 100% defect-free production, the proposed model accommodates imperfect items and effectively explores strategies for managing such scenarios. The study investigates the interrelations among suppliers, retailers, and customers, aiming to optimize the retailer's profit and order quantities. By integrating environmental considerations and operational efficiency, this

research offers a comprehensive framework for sustainable inventory management in supply chains.

3.2. Notations

This section contains all the notations which are used in this model (Table 1).

Table 1. Notations.

Symbols	Description
Parameters	
$I_i(t),\ i = 1,2,3$	Inventory level for time t
y	Order level per cycle (units)
D	Demand rate (unit/years)
A	Ordering cost per order (in dollars)
A'	Cost of carbon emissions while placing an order
c	Purchase price (in dollars) of an item from the retailer
c'	Cost of carbon emissions when a store purchases a unit
v	Pricing each item for sale (in dollars) v > c
β	Deterioration rate $(0 \leq \beta \leq 1)$ (units)
h	Holding costs, excluding interest charges, for objects held per cycle (per unit annually)
h'	Cost of carbon emissions per item for each cycle of storing items
I_e	Interest earned (per dollar per year)
I_p	Interest charge (per dollar per year)
s	Inspection cost per unit (in dollars)
p	Percentage of defective items in y (units)
x	Inspection rate (per unit per unit time)
M	Delayed payment offered by the supplier
N	Delayed payment offered by the retailer
t_1	Screening time (in years)
t_d	No deterioration time (in years)
Functions:	
$f(p)$	Probability density function of p
$TC(T)$	Total cost (in dollars)
$TP(T)$	Total profit (in dollars)
$TPU(T)$	Total profit per unit of time (in dollars)
$E[TPU]$	Expected total profit per unit of time (in dollars)
Decision variable:	
T	Cycle time (in years)

3.3. Assumptions

This section contains all the assumption which are considered in the model.

i. There are carbon emissions and energy consumption per unit associated with storing inventory in warehouses. As a result, h' represents the holding cost related to carbon emissions per unit item.
ii. D is the rate of annual demand, which is known, uniform, and constant.
iii. The distribution of the proportion of items of defective quality (p) is uniform in [α, β] where [$0 \leq \alpha \leq \beta \leq 1$].
iv. There is a carbon emission cost c' on purchasing units from the supplier, which is paid by the retailer, as per government policy.
v. The cost of carbon emissions when an item is ordered is A'.
vi. Lead time is never changing. The replenishment rate occurs at an infinite rate.
vii. A single item is used in the inventory model.
viii. No shortages occur in the inventory model.
ix. The supplier's (M) credit period is always greater than or equal to the retailer's (N) credit duration for consumers. This relationship ensures M $\geq$ N.
x. During case 1: $T \leq M$, there is no interest charged. But in case 2: $N < T < M$ and case 3: $T \geq M$, I_p, the retailer will charge for those items that are left in stock.
xi. Between periods N and M, the retailer earns a profit I_e through credit financing. Earnings begin when customers start making payments and continue until the designated payment period for the goods expires.

4. Model Description and Analysis

This research investigates a multi-level trade credit policy for inventory comprising defective and non-instantaneously deteriorating items. Initially, a quantity related to y units enters the system. Within the time frame $[0,\ t_1]$, the inventory diminishes due to demand and imperfect item screening. After the screening time, items are separated as defective and non-defective items. The screening time can be calculated by $t_1(=y/x)$. Following the screening, the inventory decreases solely due to the demand between $[t_1, t_d]$. Subsequently, deterioration commences, and the inventory is rejected due to the demand and item degeneration from $[t_d,\ T]$, depleting entirely at $t = T$. Let $I_1(t)$ and $I_2(t)$ represent the inventory levels during $t \in [0,\ t_1]$ and $t \in [t_1, t_d]$, respectively. The equations from Figure 1, which is differential of inventory level at any time t, are mentioned below:

$$\frac{dI_1(t)}{dt} = -D, \quad (0 \leq t \leq t_1) \tag{1}$$

$$\frac{dI_2(t)}{dt} = -D, \quad (t_1 \leq t \leq t_d) \tag{2}$$

The solution of the differential equations using $I_1(t_1) = y - Dt_1, I_2(t_d) = y - py - Dt_1 - Dt_d$ is

$$I_1(t) = y - Dt \tag{3}$$

$$I_2(t) = (1-p)y - Dt_1 - Dt \tag{4}$$

Likewise, let $I_3(t)$ denote the inventory level at time $t \in [t_d,\ T]$

$$\frac{dI_3(t)}{dt} + \beta t = -D, \quad (t_d \leq t \leq T) \tag{5}$$

Using $I_3(T) = 0$, the solution of the equation is given by

$$I_3(t) = \frac{D}{\beta}\left(e^{\beta(T-t)} - 1\right) \tag{6}$$

For continuity of I(t) at $t = t_d$, $I_2(t_d) = I_3(t_d)$

$$(1-p)y - Dt_1 - Dt_d = \frac{D}{\beta}\left(e^{\beta(T-t_d)} - 1\right) \tag{7}$$

which implies that the green inventory per cycle is given by

$$y = \frac{D\left(e^{\beta(T-t_d)} + \beta t_d - 1\right)}{\beta\left(1 - p - \frac{D}{x}\right)} \tag{8}$$

Further, to prevent shortages during the screening period (t_1), the effective percentage, p (a random variable uniformly distributed in the range $[a, b]$, where $0 < a < b < 1$), is limited to $(1-p)y \geq Dt_1$.

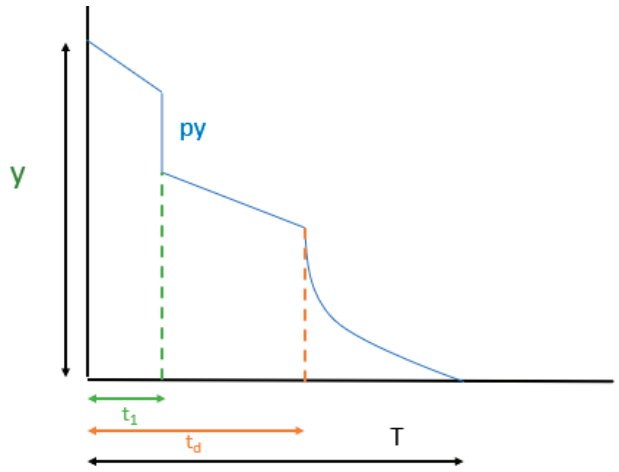


Figure 1. Inventory level at any time.

The expenses spent in the inventory model, which are used to compute the overall profit, are as follows:

(a) Sale revenue

$$(SR) = vDT \tag{9}$$

(b) The cost of order cost

$$(OC) = A + A' \tag{10}$$

(c) The cost of purchase

$$(PC) = (c + c')y$$

(d) The cost to hold the inventory

$$HC = (h + h')\left[\int_0^{t_1} I_1(t)dt + \int_{t_1}^{t_d} I_2(t)dt + \int_{t_d}^{T} I_3(t)dt\right] \tag{11}$$

$$= \left(h + h'\right)\left[(1-p)yt_d + \frac{py^2}{x} - \frac{Dyt_d}{x} + \frac{Dy^2}{x^2} - \frac{Dt_d{}^2}{2} - \frac{D}{\beta}(T - t_d) + \frac{D}{\beta^2}\left(e^{\beta(T-t_d)} - 1\right)\right] \quad (12)$$

(e) The cost of deterioration

$$(DC) = \left(c + c'\right)((1-p)y - DT) \quad (13)$$

(f) Screening cost

$$(SC) = sy \quad (14)$$

(g) Further, on the basis of the credit periods M and N, the interest paid (IP) and interest earned (IE) will be calculated as per the cases and subcases.

From the above costs, the total annual profit (TP) can be determined by total profit:

$$(TP) = \frac{1}{T}[SR - OC - PC - HC - DC - SC - IP - IE]$$

Additionally, three situations emerge based on the credit periods M and N, which are given below:

(i) $T \geq M$
(ii) $N \leq T \leq M$
(iii) $T \leq M$

Further, subcases also arise, which are given below in Table 2.

Table 2. Cases and subcases that arise for I_p and I_e.

Case I: $T \geq M$	Case II: $N \leq T \leq M$	Case III: $T \leq M$
(i) $0 \leq t_1 \leq t_d \leq N \leq M \leq T$	(i) $0 \leq t_1 \leq t_d \leq N \leq T \leq M$	(i) $0 \leq t_1 \leq t_d \leq T \leq N \leq M$
(ii) $0 \leq t_1 \leq N \leq t_d \leq M \leq T$	(ii) $0 \leq t_1 \leq N \leq t_d \leq T \leq M$	
(iii) $0 \leq N \leq t_1 \leq t_d \leq M \leq T$	(iii) $0 \leq N \leq t_1 \leq t_d \leq T \leq M$	
(iv) $0 \leq t_1 \leq N \leq M \leq t_d \leq T$		
(v) $0 \leq N \leq t_1 \leq M \leq t_d \leq T$		
(vi) $0 \leq N \leq M \leq t_1 \leq t_d \leq T$		

Case I: $T \geq M$.
There are six subcases that arise based on the values of t_1, t_d, N, M, and T:

$$(1.1)\ 0 \leq t_1 \leq t_d \leq N \leq M \leq T, \quad (1.2)\ 0 \leq t_1 \leq N \leq t_d \leq M \leq T,$$
$$(1.3)\ 0 \leq N \leq t_1 \leq t_d \leq M \leq T, \quad (1.4)\ 0 \leq t_1 \leq N \leq M \leq t_d \leq T,$$
$$(1.5)\ 0 \leq N \leq t_1 \leq M \leq t_d \leq T, \quad (1.6)\ 0 \leq N \leq M \leq t_1 \leq t_d \leq T$$

The formulation of these cases is discussed below:
Subcase I:

$$0 \leq t_1 \leq t_d \leq N \leq M \leq T$$

The retailer must pay interest I_p for the left stock in $M \leq T$ shown in Figure 2. The interest paid per cycle is as follows:

$$IP = cI_p \int_M^T I_3(t)dt = cI_p \frac{D}{\beta^2}\left(e^{\beta(T-M)} - (T-M)\beta - \beta\right) \tag{15}$$

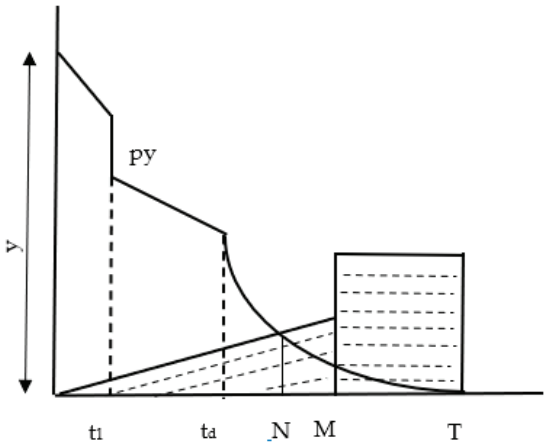


Figure 2. Interest earned and interest paid for $0 \leq t_1 \leq t_d \leq N \leq M \leq T$.

During the time period N to M, the retailer can use the sale revenue I_e shown in Figure 2. The interest earned is as follows:

$$IE = vI_e \int_N^M Dtdt = vI_e D\left(\frac{M^2}{2} - \frac{N^2}{2}\right) \tag{16}$$

Subcase II:

$$0 \leq t_1 \leq N \leq t_d \leq M \leq T$$

The retailer has some inventory left on hand as $M \leq T$. As a result, for the remaining stock depicted in Figure 3, the merchant must pay interest at the rate of I_p. The amount of interest paid is as follows:

$$IP = cI_p \int_M^T I_3(t)dt = cI_p \frac{D}{\beta^2}\left(e^{\beta(T-M)} - (T-M)\beta - \beta\right) \tag{17}$$

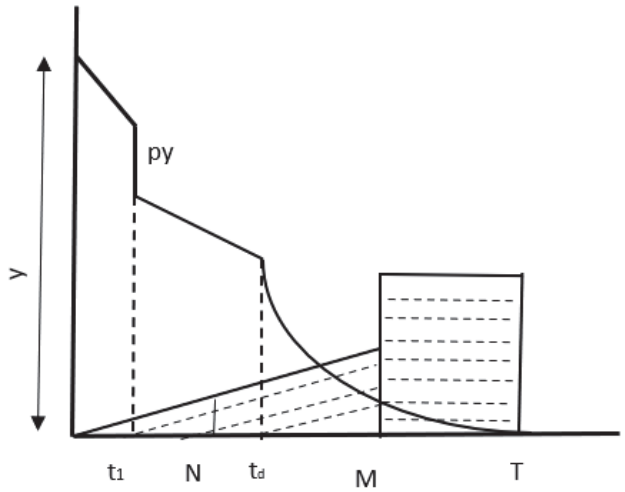


Figure 3. Interest earned and interest paid for $0 \leq t_1 \leq N \leq t_d \leq M \leq T$.

The retailer can use selling income, as depicted in Figure 3, to earn interest at the rate of I_e from N to M. The interest received per cycle is as follows:

$$IE = pI_e \int_N^M Dtdt = pI_e D\left(\frac{M^2}{2} - \frac{N^2}{2}\right) \tag{18}$$

Subcase III:

$$0 \le N \le t_1 \le t_d \le M \le T$$

The retailer has some inventory left on hand as $M \le T$. As a result, for the left stock depicted in Figure 4, the merchant must pay interest at the rate of I_p (Figure 4). The amount of interest each cycle paid is as follows:

$$IP = cI_p \int_M^T I_3(t)dt = cI_p \frac{D}{\beta^2}\left(e^{\beta(T-M)} - (T-M)\beta - \beta\right) \tag{19}$$

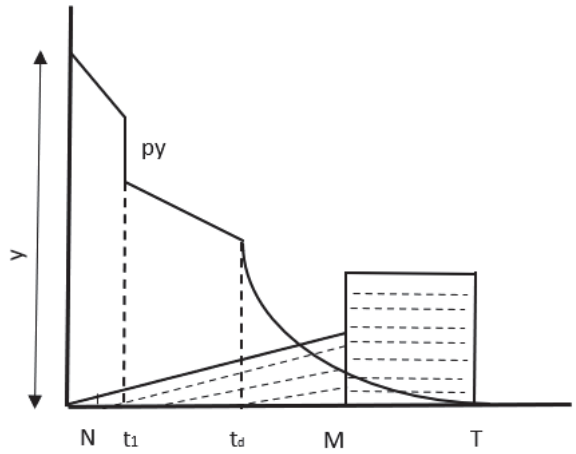


Figure 4. Interest earned (*IE*), and interest paid (*IP*) for $0 \le N \le t_1 \le t_d \le M \le T$.

The retailer can use selling income, as depicted in Figure 4, to earn interest at the rate of I_e from N to M. The interest received per cycle is as follows:

$$IE = vI_e \int_N^M Dtdt = vI_e D\left(\frac{M^2}{2} - \frac{N^2}{2}\right) \tag{20}$$

Subcase IV:

$$0 \le t_1 \le N \le M \le t_d \le T$$

Some inventory on hand is left by the retailer as $M \le T$. As a result, for the left stock depicted in Figure 5, the merchant must pay interest at the rate of I_p. The amount of interest each cycle paid is as follows:

$$\begin{gathered} IP = cI_p\left(\int_M^{t_d} I_2(t)dt + \int_{t_d}^T I_3(t)dt\right) \\ = cI_p\left(\frac{D}{2}\left(M^2 - t_d^2\right) + (1-p)y(t_d - M) + \frac{Dy}{x}(M - t_d) + \frac{D}{\beta}(t_d - T) + \frac{D}{\beta^2}\left(e^{\beta(T-t_d)} - 1\right)\right) \end{gathered} \tag{21}$$

The retailer can use selling income, as depicted in Figure 5, to earn interest at the rate of I_e from N to M. The interest received per cycle is as follows:

$$IE = vI_e \int_N^M Dtdt = vI_e D\left(\frac{M^2}{2} - \frac{N^2}{2}\right) \tag{22}$$

Subcase V:

$$0 \le N \le t_1 \le M \le t_d \le T$$

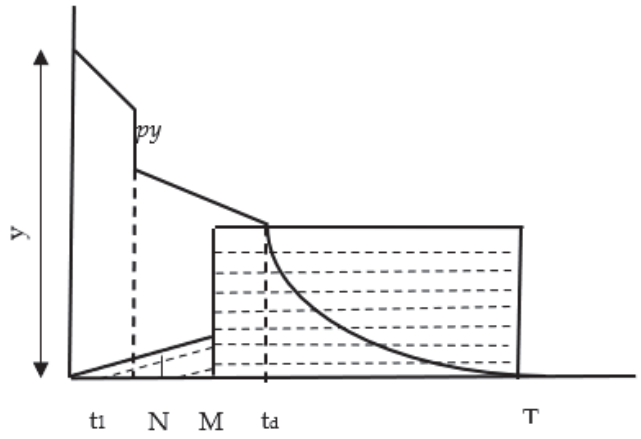


Figure 5. Interest earned and interest paid for $0 \le t_1 \le N \le M \le t_d \le T$.

Some inventory on hand is left by the retailer as $M \le T$. As a result, for the left stock depicted in Figure 6, the merchant must pay interest at the rate of I_p. The amount of interest each cycle paid is as follows:

$$\begin{gathered} IP = cI_p\left(\int_M^{t_d} I_2(t)dt + \int_{t_d}^{T} I_3(t)dt\right) \\ = cI_p\left(\frac{D}{2}(M^2 - t_d^2) + (1-p)y(t_d - M) + \frac{Dy}{x}(M - t_d) + \frac{D}{\beta}(t_d - T) + \frac{D}{\beta^2}\left(e^{\beta(T-t_d)} - 1\right)\right) \end{gathered} \tag{23}$$

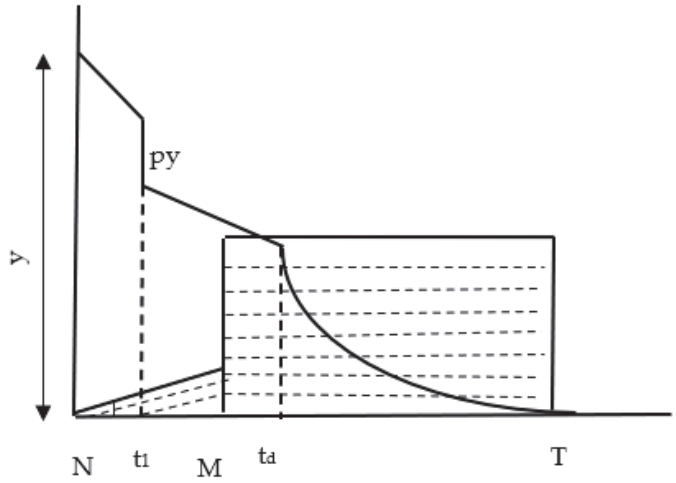


Figure 6. The interest earned and interest paid for $0 \le N \le t_1 \le M \le t_d \le T$.

The retailer can use selling income, as depicted in Figure 6, to earn interest at the rate of I_e from N to M. The interest received per cycle is as follows:

$$IE = vI_e\int_N^M Dtdt = vI_eD\left(\frac{M^2}{2} - \frac{N^2}{2}\right) \tag{24}$$

Subcase VI:

$$0 \le N \le M \le t_1 \le t_d \le T$$

Some inventory on hand is left by the retailer as $M \le T$. As a result, for the left stock depicted in Figure 7, the merchant must pay interest at the rate of I_p. The amount of interest each cycle cleared is as follows:

$$\begin{gathered} IP = cI_p\left(\int_M^{t_1} I_1(t)dt + \int_{t_1}^{t_d} I_2(t)dt + \int_{t_d}^{T} I_3(t)dt\right) \\ = cI_p\left(\frac{DM^2}{2} + y(\frac{y}{x} - M) + \frac{D}{2}\left(\frac{y^2}{x^2} - M\right) - \frac{Dy^2}{2x^2} + (1-p)y(t_d - \frac{y}{x}) + \frac{D*y}{x}(\frac{y}{x} - t_d) + \frac{D}{\beta}(t_d - T) + \frac{D}{\beta^2}\left(e^{\beta(T-t_d)} - 1\right)\right) \end{gathered} \tag{25}$$

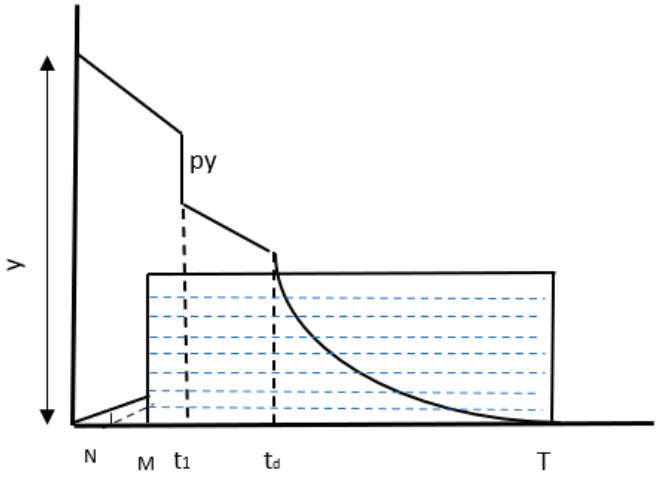


Figure 7. Interest earned and interest paid for $0 \le N \le M \le t_1 \le t_d \le T$.

The retailer can use selling income, as depicted in Figure 7, to earn interest at the rate of I_e from N to M. The interest received per cycle is as follows:

$$IE = vI_e \int_N^M Dtdt = vD\left(\frac{M^2}{2} - \frac{N^2}{2}\right) \tag{26}$$

Case II:

$$N \le T \le M$$

Based on the values of $t_1,\ t_d,\ N,\ M,$ and T, there are three subcases:

(1.1) $0 \le t_1 \le t_d \le N \le T \le M$
(1.2) $0 \le t_1 \le N \le t_d \le T \le M$
(1.3) $0 \le N \le t_1 \le t_d \le T \le M$

The formulation of these cases is discussed below:

Subcase I:

$$0 \le t_1 \le t_d \le N \le T \le M$$

Figure 8 illustrates how the store generates income and interest (at rate I_e) from the inventory sales from N to T. Every revenue is subject to interest from T to M, accumulating interest from N to M each cycle.

$$IE = vI_e\left(\int_N^T Dtdt + DT(M-T)\right) = vI_eD\left(TM - \frac{N^2}{2} - \frac{T^2}{2}\right) \tag{27}$$

Subcase II:

$$0 \le t_1 \le N \le t_d \le T \le M$$

Figure 9 illustrates how the store generates income and interest (at rate I_e) from inventory sales from N to T. Every revenue is subject to interest from T to M, accumulating interest from N to M each cycle.

$$IE = vI_e\left(\int_N^T Dtdt + DT(M-T)\right) = vI_eD\left(TM - \frac{T^2}{2} - \frac{N^2}{2}\right) \tag{28}$$

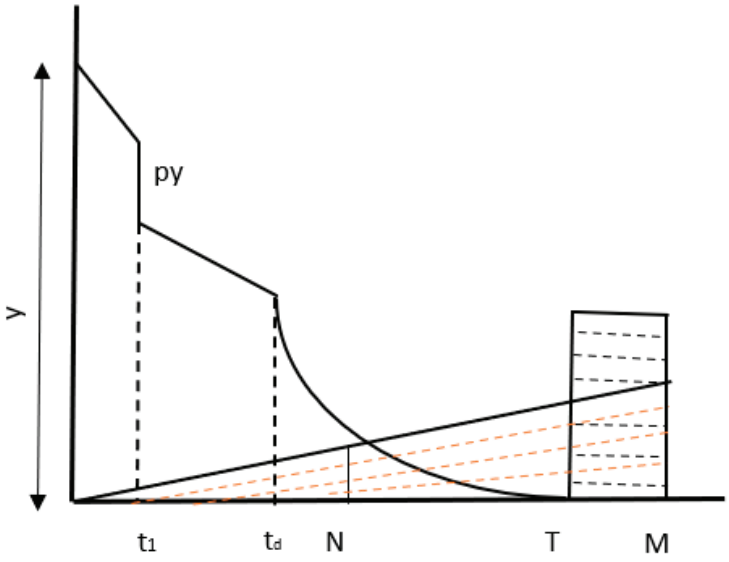


Figure 8. Interest earned for $0 \leq t_1 \leq t_d \leq N \leq T \leq M$.

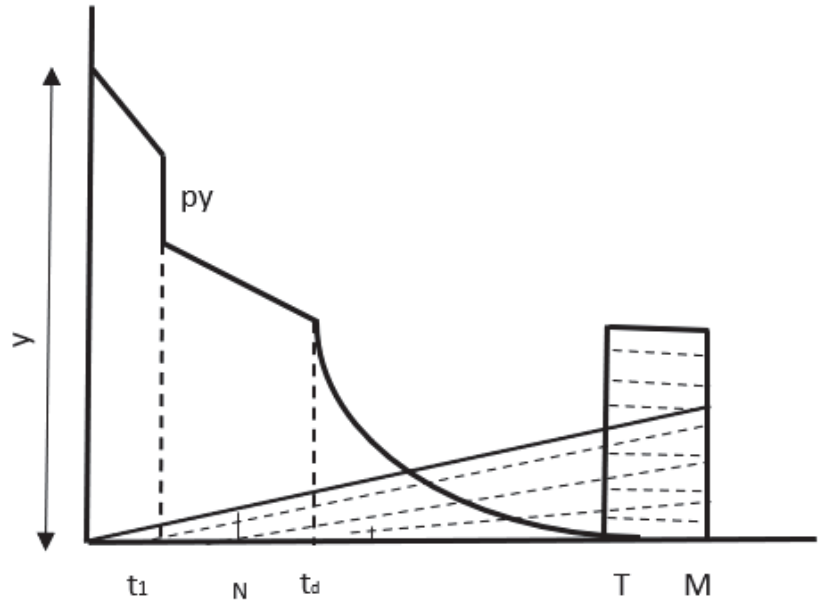


Figure 9. Interest earned for $0 \leq t_1 \leq N \leq t_d \leq T \leq M$.

Subcase III:

$$0 \leq N \leq t_1 \leq t_d \leq T \leq M$$

Figure 10 illustrates how the store generates income and interest (at rate I_e) from inventory sales from N to T. Every revenue is subject to interest from T to M, accumulating interest from N to M each cycle.

$$IE = vI_e\left(\int_N^T Dtdt + DT(M-T)\right) = vI_eD\left(TM - \frac{T^2}{2} - \frac{N^2}{2}\right) \tag{29}$$

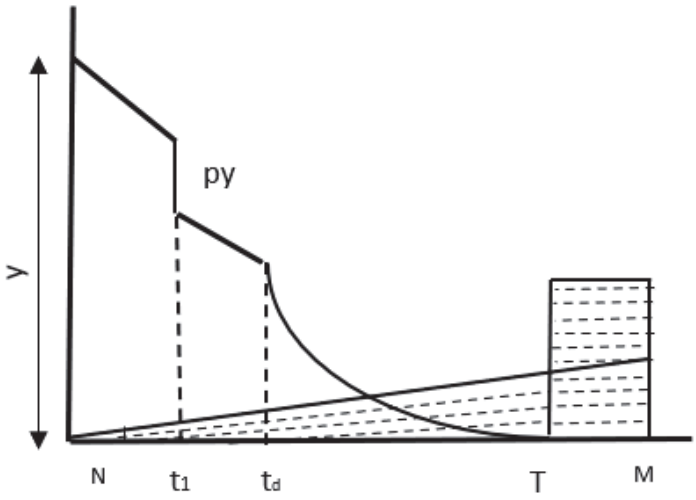


Figure 10. Interest earned for $0 \leq N \leq t_1 \leq t_d \leq T \leq M$.

Case III:

$$T \leq M$$

Based on the values of t_1, t_d, N, M, and T, only $0 \leq t_1 \leq t_d \leq T \leq N \leq M$ is the subcase. The formulation of this subcase is discussed below.

The store sells inventory from N to M and earns interest at a rate of I_e on the total sales income, as shown in Figure 11, which causes interest to accumulate throughout this time in each cycle.

$$IE = pvDT(M - N) \tag{30}$$

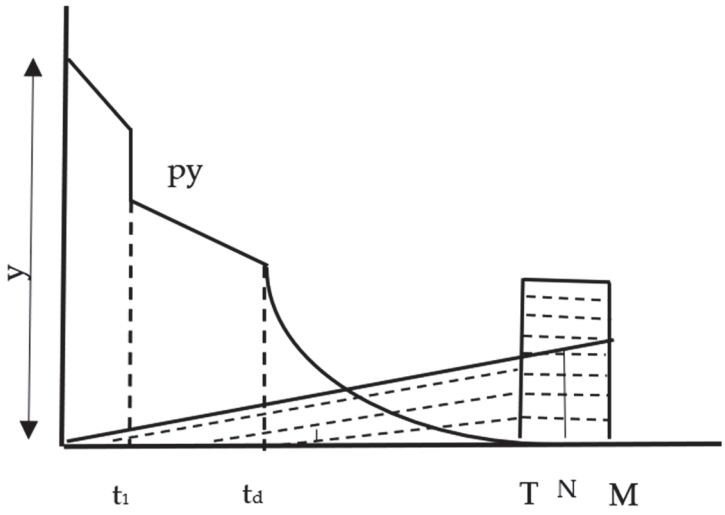


Figure 11. Interest earned for $0 \leq t_1 \leq t_d \leq T \leq N \leq M$.

From the above cases, since p is a random variable with a known probability density function $f(p)$, then the expected total profit per unit of time, $E[TPU]_{i.j}$, *where* $i = 1, 2, 3$ *and* $j = 1, 2, 3, 4, 5, 6$, can be expressed as follows:

Case I: When $T \geq M$ then

$$E[TPU]_{i.j} = \begin{cases} E[TPU]_{1.1} = E[TPU]_{1.2} = E[TPU]_{1.3}, & 0 \leq t_d \leq N \leq M \leq T \\ E[TPU]_{1.4} = E[TPU]_{1.5}, & 0 \leq t_1 \leq N \leq M \leq t_d \leq T \\ E[TPU]_{1.6}, & 0 \leq N \leq M \leq t_1 \leq t_d \leq T \end{cases} \tag{31}$$

Case II: When $N \le T \le M$ then

$$E[TPU]_{i,j} = E[TPU]_{2.1} = E[TPU]_{2.2}(T) = E[TPU]_{2.3,}\ 0 \ \le t_d \le N \le T \le M \quad (32)$$

Case III: When $T \ge M$ then

$$E[TPU]_{i,j} = E[TPU]_{3.1},\ 0 \le t_1 \le t_d \le T \le N \le M \quad (33)$$

where

$$\begin{aligned} E[TPU]_{1.1} = \tfrac{1}{T}[\ & vDT - (A + A') - (c + c')y \\ & -(h + h')\Big((1 - E[p])yt_d + \tfrac{E[p]y^2}{x} - \tfrac{Dyt_d}{x} + \tfrac{Dy^2}{x^2} - \tfrac{Dt_d{}^2}{2} - \tfrac{D}{\beta}(T - t_d) \\ & + \tfrac{D}{\beta^2}\Big(e^{\beta(T-t_d)} - 1\Big)\Big) - (c + c')((1 - E[p])y - DT) - sy - cI_p\tfrac{D}{\beta^2}\Big(e^{\beta(T-M)} - (T - M)\beta - \beta\Big) \\ & + vI_eD\Big(\tfrac{M^2}{2} - \tfrac{N^2}{2}\Big)\Big] = E[TPU]_{1.2} = E[TPU]_{1.3} \end{aligned} \quad (34)$$

$$\begin{aligned} E[TPU]_{1.4} = \tfrac{1}{T}[\ & vDT - (A + A') - (c + c')y \\ & -(h + h')\Big((1 - E[p])yt_d + \tfrac{E[p]y^2}{x} - \tfrac{Dyt_d}{x} + \tfrac{Dy^2}{x^2} - \tfrac{Dt_d{}^2}{2} - \tfrac{D}{\beta}(T - t_d) \\ & + \tfrac{D}{\beta^2}\Big(e^{\beta(T-t_d)} - 1\Big)\Big) - (c + c')((1 - E[p])y - DT) - sy \\ & - cI_p\Big(\tfrac{D}{2}(M^2 - t_d^2) + (1 - E[p])y(t_d - M) + \tfrac{Dy}{x}(M - t_d) + \tfrac{D}{\beta}(t_d - T) + \tfrac{D}{\beta^2}\Big(e^{\beta(T-t_d)} - 1\Big)\Big) \\ & + vI_eD\Big(\tfrac{M^2}{2} - \tfrac{N^2}{2}\Big)\Big] = E[TPU]_{1.5} \end{aligned} \quad (35)$$

$$\begin{aligned} E[TPU]_{1.6} = \tfrac{1}{T}[\ & vDT - (A + A') - (c + c')y \\ & -(h + h')\Big((1 - E[p])yt_d + \tfrac{E[p]y^2}{x} - \tfrac{Dyt_d}{x} + \tfrac{Dy^2}{x^2} - \tfrac{Dt_d{}^2}{2} - \tfrac{D}{\beta}(T - t_d) \\ & + \tfrac{D}{\beta^2}\Big(e^{\beta(T-t_d)} - 1\Big)\Big) - (c + c')((1 - E[p])y - DT) - sy \\ & - cI_p\Big(\tfrac{DM^2}{2} + y\big(\tfrac{y}{x} - M\big) + \tfrac{D}{2}\Big(\tfrac{y^2}{x^2} - M\Big) - \tfrac{Dy^2}{2x^2} + (1 - E[p])y\big(t_d - \tfrac{y}{x}\big) + \tfrac{Dy}{x}\big(\tfrac{y}{x} - t_d\big) \\ & + \tfrac{D}{\beta}(t_d - T) + \tfrac{D}{\beta^2}\Big(e^{\beta(T-t_d)} - 1\Big)\Big] + vI_eD\Big(\tfrac{M^2}{2} - \tfrac{N^2}{2}\Big)\Big] \end{aligned} \quad (36)$$

$$\begin{aligned} E[TPU]_{2.1} = \tfrac{1}{T}[\ & vDT - (A + A') - (c + c')y \\ & -(h + h')\Big((1 - E[p])yt_d + \tfrac{E[p]y^2}{x} - \tfrac{Dyt_d}{x} + \tfrac{Dy^2}{x^2} - \tfrac{Dt_d{}^2}{2} - \tfrac{D}{\beta}(T - t_d) \\ & + \tfrac{D}{\beta^2}\Big(e^{\beta(T-t_d)} - 1\Big)\Big) - (c + c')((1 - E[p])y - DT) - sy + vI_eD\Big(TM - \tfrac{N^2}{2} - \tfrac{T^2}{2}\Big)\Big] \end{aligned} \quad (37)$$

$$\begin{aligned} E[TPU]_{3.1} = \tfrac{1}{T}[\ & vDT - (A + A') - (c + c')y \\ & -(h + h')\Big((1 - E[p])yt_d + \tfrac{E[p]y^2}{x} - \tfrac{Dyt_d}{x} + \tfrac{Dy^2}{x^2} - \tfrac{Dt_d{}^2}{2} - \tfrac{D}{\beta}(T - t_d) \\ & + \tfrac{D}{\beta^2}\Big(e^{\beta(T-t_d)} - 1\Big)\Big) - (c + c')((1 - E[p])y - DT) - sy + vI_eDT(M - N)] \end{aligned} \quad (38)$$

5. Solution Procedure

Furthermore, concavity is extremely difficult to show theoretically because the derivatives of the expected total profit functions are complex. Additionally, Figure 12 illustrates the nature of the functions for each of the ten cases, as explained below:

The necessary condition required for maximization of total profit is defined by $\frac{d\,E[TPU]}{dT} = 0$, which provides the optimal values for T. This study aims to optimize the cycle time to maximize total profit, $E[TPU]$. Hence, the profit function's behaviour is analyzed for all cases using the derivative method.

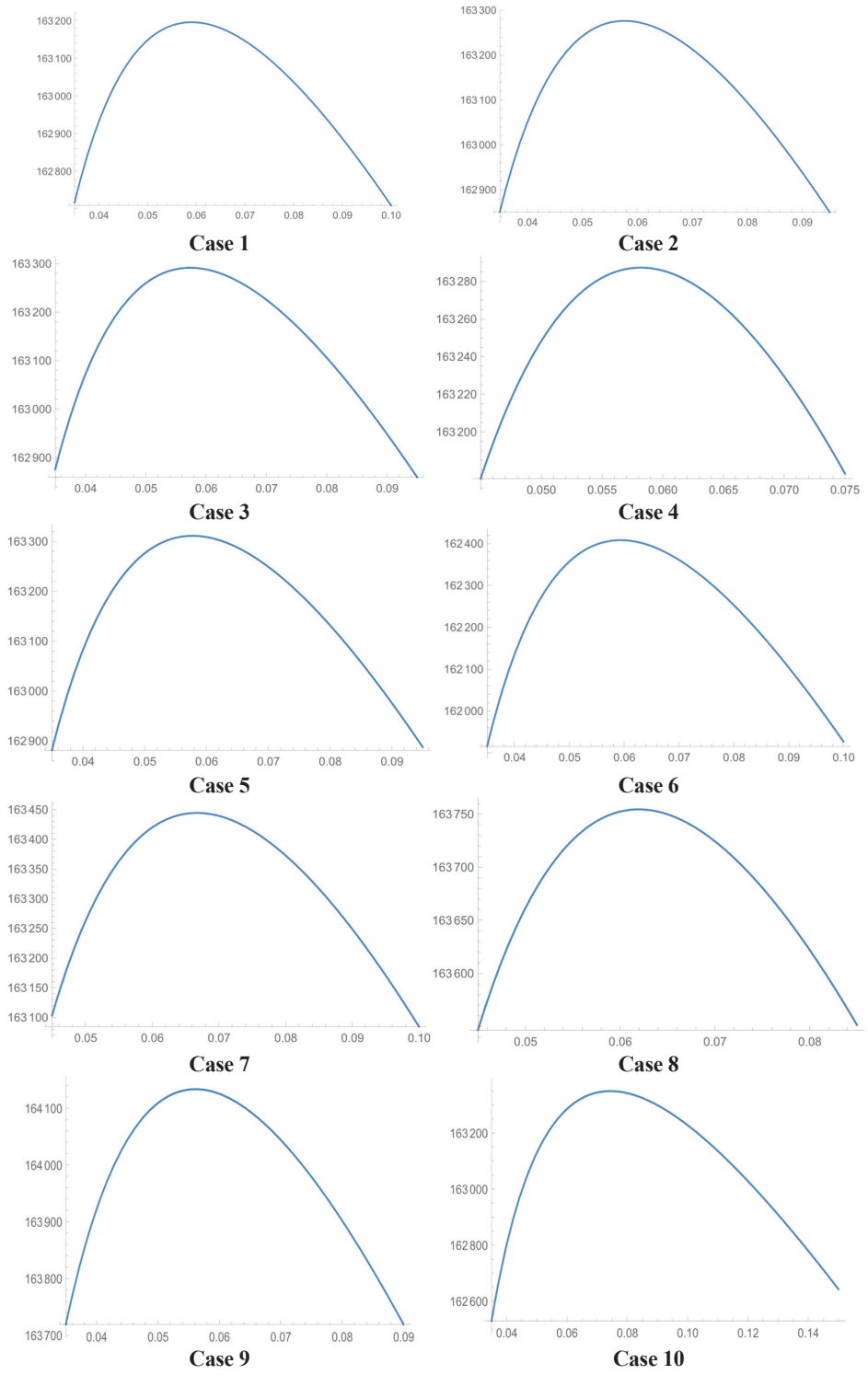


Figure 12. Concavity of the expected total profit function for all the cases.

Additionally, the following adequate conditions must be met for the predicted total profit $E[TPU]_{i,j}$ to be concave: $\left(\frac{d^2E[TPU]_{i,j}}{dT^2}\right) \leq 0$ (where $i = 1, 2, 3$ and $j = 1, 2, 3, 4, 5, 6$).

Appendix A contains the first and second derivatives of the expected total profit function, which are quite big equations.

In order to find the optimal value T^*, the following Algorithm 1 is proposed:

Algorithm 1. Finding the optimal value T^*

Step 1: Put values of all the parameters in Equation (A1)
Step 2: Determine the value of cycle length, T. Now, using the value of T, calculate the values of y from Equation (8).
If $(T \geq M$ & $0 \leq t_1 \leq t_d \leq N \leq M \leq T$ (case 1: subcase 1.1 satisfy))
then the cycle length will be T and the value of expected total profit, E[TPU], can be obtained by putting T in Equation (A1),
else
this case is not feasible and set E[TPU] = 0.
If $(0 \leq t_1 \leq N \leq t_d \leq M \leq T)$
then find T and E[TPU] for this case from Equation (A1),
else
case is not possible, set E[TPU] = 0.
If $(0 \leq N \leq t_1 \leq t_d \leq M \leq T)$
Then determine T and E[TPU] from Equation (A1),
else
put E[TPU] = 0, in case of not feasible
If $(0 \leq t_1 \leq N \leq M \leq t_d \leq T)$
then calculate T and E[TPU] from Equation (A3),
else
E[TPU] = 0.
If $(0 \leq N \leq t_1 \leq M \leq t_d \leq T)$
then find T and predicted total profit E[TPU] from (A3)
else
put E[TPU] = 0.
If $(0 \leq N \leq M \leq t_1 \leq t_d \leq T)$
then determine T and E[TPU] from Equation (A5)
else
place E[TPU] = 0.
If $(N \leq T \leq M$ & $0 \leq t_1 \leq t_d \leq N \leq T \leq M)$
then compute T and E[TPU] from (A6)
else
set E[TPU] = 0.
If $(N \leq T \leq M$ & $0 \leq t_1 \leq N \leq t_d \leq T \leq M)$
then evaluate T and $E[TPU]$ from (A6)
else
put E[TPU]= 0.
If $(N \leq T \leq M$ & $0 \leq N \leq t_1 \leq t_d \leq T \leq M)$
then determine T and E[TPU] from Equation (A6)
else
put E[TPU] = 0.
If $(T \leq M$ & $0 \leq t_1 \leq t_d \leq T \leq N \leq M)$
then evaluate T and E[TPU] from (A8)
else
in the case of not feasible, put E[TPU] = 0.
Step 3. Compare the expected total profit for all the 10 cases, i.e.,
Max {for all cases $E[\text{TPU}]_{\text{i.j}}$, (i = 1, 2, 3 and j = 1, 2, 3, 4, 5, 6}

Choose the ideal scenario, which has the optimal value of T and the largest predicted total profit, E[TPU].

6. Numerical Example

This section presents a numerical example to illustrate and validate the theoretical findings detailed in the examples. Most of the parameter values are considered from Ref. [38].

Example 1. $D = 7000\ units/year$, $A = 100\ \$/order$, $h = 5\ \$/unit/year$, $\beta = 0.06$, $t_d = 0.04\ year$, $c = 25\$$, $s = 0.3\$/unit$, $A' = 1\$/order$, $h' = 0.1\$/unit/year$, $c' = 1\$$, $x = 175200\ units\ /year$, $v = 50\$/units$, $I_e = 0.08\ \$/year$, $I_p = 0.12\$/year$. *N and M values are considered as per the cases 1 to 10.* $f(p) = \begin{cases} 25, & 0 \le p \le 0.04 \\ 0 & otherwise \end{cases}$, *then the expected value of defective items will be* $E[p] = 0.02$.

From Table 3, The first five cases have relatively consistent values of T around 0.0573 yr to 0.0581 yr, showing minimal variation. A noticeable increase occurs in cases 7 (0.0667 yr) and 10 (0.0743 yr), indicating some deviation in the values. The E[TPU] values show a stable pattern, though case 6 has a notably lower E[TPU] of 162408\$. Case 9 has the highest total cost at 164134\$, and the optimal order quantity (y) is 418 units, implying better performance or increased computational efficiency.

Table 3. Optimal values of cycle length and expected total profit for all the cases.

Cases	*T*(Yr)	*E* [*TPU*] ($)
1	0.0590	163195
2	0.0576	163276
3	0.0573	163291
4	0.0581	163287
5	0.0577	163311
6	0.0593	162408
7	0.0667	163444
8	0.0619	163754
9	0.0561	164134
10	0.0743	163350

7. Sensitivity Analysis

To determine the robustness of the model, a sensitivity analysis is carried out by adjusting a number of effective parameters (Tables 4–7 and Figures 13–15). The effect of the effective parameters is considered in the best case, i.e., 9.

- As I_e decreases, the processing time T gradually increases. This implies that higher interest rates may incentivize quicker or more efficient operations, likely due to better resource utilization.
- E[TPU] steadily declines as I_e decreases. This decline indicates that lower interest earnings directly affect profitability, possibly due to reduced capital reinvestment benefits.

Table 4. Effect of changes in $E[p]$ on *ETPU*.

E[*p*]	*T*(Yr)	*E* [*TP*U] ($)
0.01	0.056147	163863
0.02	0.056133	164134
0.03	0.056118	164410
0.04	0.056103	164692
0.05	0.056087	164980

Table 5. Effect of changes in $\mathbf{I}_e$ on ***TP***.

$\mathbf{I}_e$	***T*(Yr)**	**E[TPU] ($)**
0.096	0.053861	164372
0.088	0.054961	164252
0.080	0.056133	164134
0.072	0.057382	164017
0.064	0.058718	163902

Table 6. Effect of changes in h on $E[\boldsymbol{TPU}]$.

h	**T(Yr)**	**E[TPU] ($)**
6.0	0.053319	163941
5.5	0.054671	164036
5.0	0.056133	164134
4.5	0.057718	164234
4.0	0.059447	164337

Table 7. Effect of changes in β on $E[TPU]$.

β	***T*(Yr)**	**E[TPU] ($)**
0.02	0.056197	164135
0.04	0.056165	164134
0.06	0.056133	164134
0.08	0.056100	164133
0.10	0.056068	164132

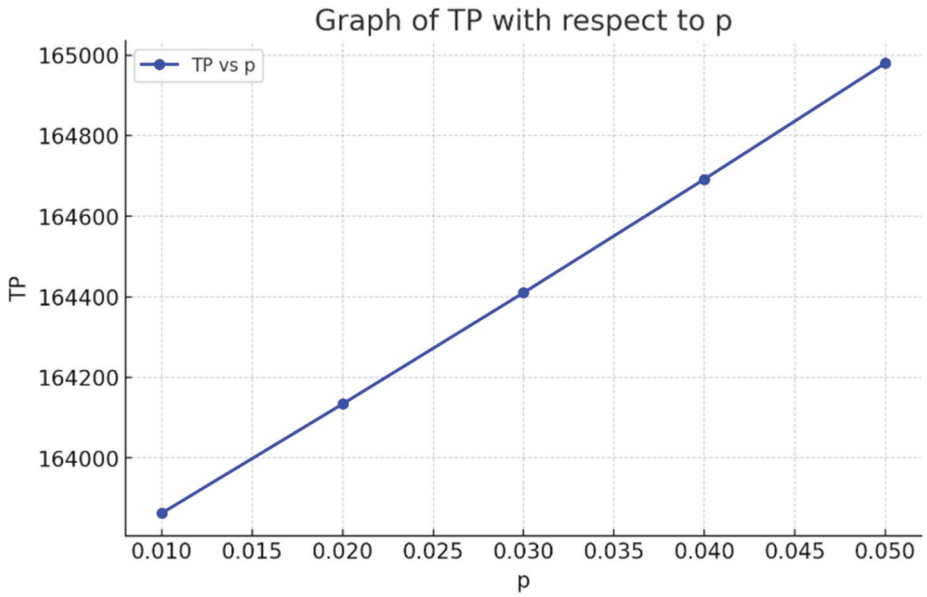


Figure 13. Graph of Defective percentage vs. Expected total profit.

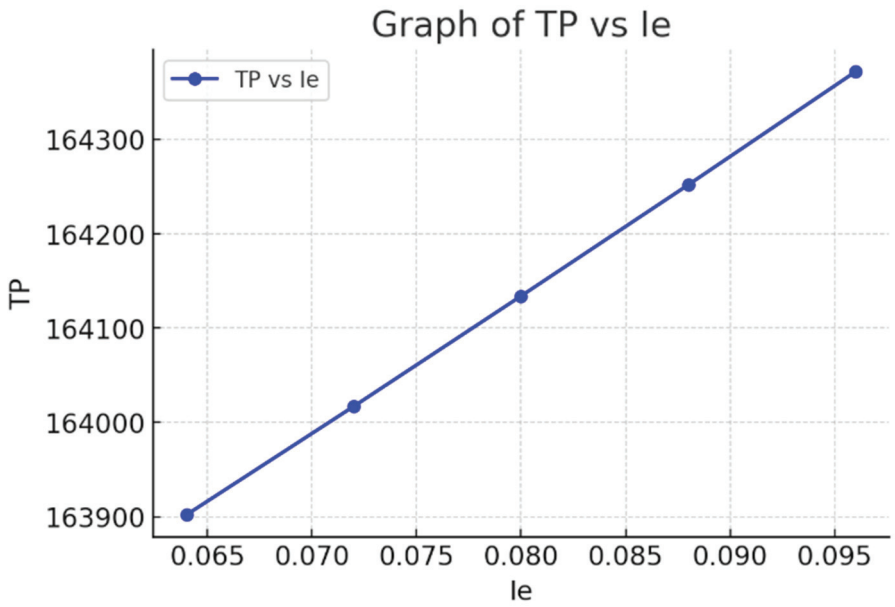


Figure 14. Graph of Ie vs. Expected total profit.

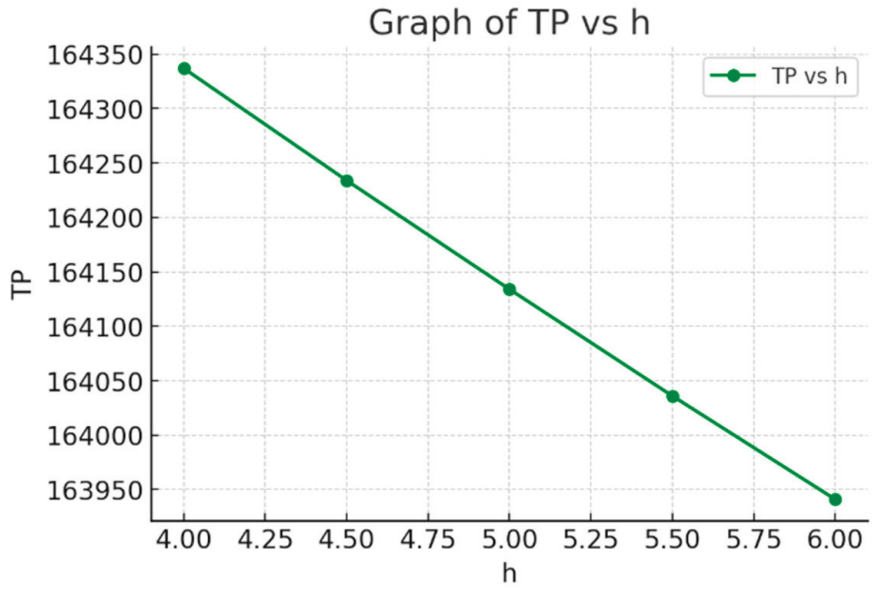


Figure 15. Graph of holding charges vs. Expected total profit.

As the holding cost decreases from 6 to 4, the E[**TP**U] steadily increases. This suggests that lower holding costs may introduce inefficiencies or require more time to process or execute certain operations. Higher E[**TP**U] values typically indicate better or more efficient system output, suggesting that lower holding costs may promote higher performance despite the increased time.

Table 7 shows that the overall profit would decline as the rate of deterioration rises. An increase in the rate of deterioration will raise the order quantity to make up for the loss from deterioration. As the deterioration rate rises, the cycle length decreases as well, slowing down the deterioration process and raising overall profit.

8. Managerial Implications

The results derived from the numerical examples highlight the strategic approaches inventory managers should adopt when handling inventories under a multi-level trade credit program that incorporates carbon emissions. The study examines three scenarios based on multi-level trade credit. The most advantageous case occurs when the retailer's and customer's credit periods fall between the screening time and the non-deteriorating time duration. This scenario enables inventory managers to achieve higher profitability. Managerial decisions must account for ordering costs, holding costs, interest payments, and carbon emissions, emphasizing the importance of a holistic approach to inventory management. Inventory managers must carefully evaluate all options before procurement, as decisions impact both overall profitability and the quantity of orders placed. Larger order quantities increase holding costs and carbon emission costs, highlighting the need for managers to optimize total order quantities. Incorporating carbon emissions into inventory management strategies helps minimize order quantities, directly reducing holding- and emission-related costs. Furthermore, profitability is enhanced when managers focus on trade credit terms and the management of imperfect-quality items. Managers achieve higher profits when the interest earned exceeds the interest paid, underlining the critical role of financial strategies in inventory management. This research aims to address carbon emissions within a multi-level trade credit framework for imperfect-quality items. By integrating carbon emission considerations, the proposed model supports more sustainable practices, balancing environmental and economic objectives while maximizing profitability. By effectively managing a perishable and deteriorating inventory, this model can help industries including FMCG, pharmaceuticals, and food and beverage. It can be used to optimize stock levels and reduce waste in the chemical and electronic industries. Through better trade credit regulations, the model aids in balancing sustainability and profitability.

9. Concluding Remarks and Future Scope

This study aims to reduce carbon emissions by implementing multi-level trade credit policies. An inventory model has been developed for items with non-instantaneous deterioration. In practical production scenarios, achieving a perfect quality rate for all units is unrealistic, and this issue is addressed within the study. The research focuses on items with a gradual deterioration rate, such as food products, electronic components, and fashion items, where defects during production can be addressed through an inspection process. The model examines imperfect-quality items within the framework of a multi-level trade credit policy, integrating an economic order quantity (EOQ) model. This approach emphasizes the optimization of the retailer's total profit while accounting for items with production defects that are rectifiable through inspection. Multi-level trade credit policies emerge as the most effective strategy for maximizing profits in this context. The study explores various scenarios involving suppliers, retailers, and customers, using numerical examples to validate the results. The optimal scenario for retailers occurs when the credit periods for both retailers and customers fall between the screening time and the non-instantaneous deterioration period. This configuration ensures enhanced profitability. The model's applicability can be further extended by incorporating finite replenishment rates, variable demand rates, or other dynamic parameters, offering a robust foundation for addressing complex inventory management challenges. Additionally, this model can be expanded to include closed-loop supply chains and warehouses.

Author Contributions: A.B.: conceptualization, visualization, data curation, funding acquisition, review, methodology; A.P.: writing—original draft, software; data curation, methodology, writing—review and editing, investigation, software; B.U.: supervision, writing—review and editing; M.M.: conceptualization, visualization, investigation, supervision, writing–review and editing, software. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data will be provided as and when required.

Conflicts of Interest: The authors declare that there were no conflicts of interest.

Appendix A

$$
\begin{aligned}
\boldsymbol{E[TPU]_{1.1}}' = &-\frac{1}{T^2}\Bigg(-\boldsymbol{A}-\mathbf{A}'+\mathbf{I}_{\boldsymbol{e}}\Big(-\frac{N^2}{2}+\frac{M^2}{2}\Big)\boldsymbol{vD}+\boldsymbol{DTv}-\frac{\boldsymbol{cI_pD}\big(e^{(-M+T)\beta}-\beta-(-M+T)\beta\big)}{\beta^2}-\frac{(c+\mathbf{c}')D\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathbf{t}_d\boldsymbol{\beta}\Big)}{\big(1-\boldsymbol{E[p]}-\frac{D}{x}\big)\boldsymbol{\beta}}\\
&-\frac{\boldsymbol{Ds}\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathbf{t}_d\boldsymbol{\beta}\Big)}{\big(1-\boldsymbol{E[p]}-\frac{D}{x}\big)\boldsymbol{\beta}}-\big(\boldsymbol{c}+\mathbf{c}'\big)\Bigg(-\boldsymbol{DT}+\frac{(1-\boldsymbol{E[p]})\boldsymbol{D}\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathbf{t}_d\boldsymbol{\beta}\Big)}{\big(1-\boldsymbol{E[p]}-\frac{D}{x}\big)\boldsymbol{\beta}}\Bigg)-\\
&(\boldsymbol{h}+\mathbf{h}')\Bigg(-\frac{D\mathrm{t}_d{}^2}{2}-\frac{D}{\beta^2}+\frac{e^{(T-\mathrm{t}_d)\beta}D}{\beta^2}-\frac{D(T-\mathrm{t}_d)}{\beta}+\frac{(1-p)D\mathrm{t}_d\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathbf{t}_d\boldsymbol{\beta}\Big)}{\big(1-\boldsymbol{E[p]}-\frac{D}{x}\big)\boldsymbol{\beta}}-\frac{\boldsymbol{D^2\mathrm{t}_d}\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathbf{t}_d\boldsymbol{\beta}\Big)}{\big(1-\boldsymbol{E[p]}-\frac{D}{x}\big)\boldsymbol{x\beta}}+\\
&\frac{\boldsymbol{D^3}\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathbf{t}_d\boldsymbol{\beta}\Big)^2}{\big(1-\boldsymbol{E[p]}-\frac{D}{x}\big)^2\boldsymbol{x^2\beta^2}}+\frac{\boldsymbol{D^2}\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathbf{t}_d\boldsymbol{\beta}\Big)^2}{\big(1-\boldsymbol{E[p]}-\frac{D}{x}\big)^2\boldsymbol{x\beta^2}}-\frac{(1-p)D^2\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathbf{t}_d\boldsymbol{\beta}\Big)^2}{\big(1-\boldsymbol{E[p]}-\frac{D}{x}\big)^2\boldsymbol{x\beta^2}}\Bigg)\Bigg)+\\
&\frac{1}{\boldsymbol{T}}\Bigg(\boldsymbol{Dv}-\frac{(c+\mathbf{c}')e^{(T-\mathrm{t}_d)\beta}D}{\big(1-\boldsymbol{E[p]}-\frac{D}{x}\big)}-\frac{e^{(T-\mathrm{t}_d)\beta}\boldsymbol{Ds}}{\big(1-\boldsymbol{E[p]}-\frac{D}{x}\big)}-\big(\boldsymbol{c}+\mathbf{c}'\big)\Bigg(-\boldsymbol{D}+\frac{e^{(T-\mathrm{t}_d)\beta}(1-\boldsymbol{E[p]})\boldsymbol{D}}{\big(1-\boldsymbol{E[p]}-\frac{D}{x}\big)}\Bigg)-\frac{\boldsymbol{cI_PD}\big(-\boldsymbol{\beta}+e^{(-M+T)\beta}\boldsymbol{\beta}\big)}{\boldsymbol{\beta^2}}\\
&-(\boldsymbol{h}+\mathbf{h}')\Bigg(-\frac{D}{\beta}+\frac{e^{(T-\mathrm{t}_d)\beta}(1-\boldsymbol{E[p]})D\mathrm{t}_d}{\big(1-\boldsymbol{E[p]}-\frac{D}{x}\big)}-\frac{e^{(T-\mathrm{t}_d)\beta}D^2\mathrm{t}_d}{\big(1-\boldsymbol{E[p]}-\frac{D}{x}\big)\boldsymbol{x}}+\frac{e^{(T-\mathrm{t}_d)\beta}\boldsymbol{D}}{\boldsymbol{\beta}}+\frac{2e^{(T-\mathrm{t}_d)\beta}D^3\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathbf{t}_d\boldsymbol{\beta}\Big)}{\big(1-\boldsymbol{E[p]}-\frac{D}{x}\big)^2\boldsymbol{x^2\beta}}+\\
&\frac{2e^{(T-\mathrm{t}_d)\beta}D^2\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathbf{t}_d\boldsymbol{\beta}\Big)}{\big(1-\boldsymbol{E[p]}-\frac{D}{x}\big)^2\boldsymbol{x\beta}}-\frac{2e^{(T-\mathrm{t}_d)\beta}(1-p)D^2\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathbf{t}_d\boldsymbol{\beta}\Big)}{\big(1-\boldsymbol{E[p]}-\frac{D}{x}\big)^2\boldsymbol{x\beta}}\Bigg)\Bigg)
\end{aligned}
\tag{A1}
$$

$$
\begin{aligned}
E[TPU]_{1.1}''(T) = &-\frac{1}{T^2}\Bigg(Dv-(c+\mathrm{c}')\Bigg(-D+\frac{e^{(T-\mathrm{t}_d)\beta}(1-p)D}{\big(1-E[p]-\frac{D}{x}\big)}\Bigg)-\frac{(c+\mathrm{c}')e^{(T-\mathrm{t}_d)\beta}D}{\big(1-E[p]-\frac{D}{x}\big)}-\frac{e^{(T-\mathrm{t}_d)\beta}Ds}{\big(1-E[p]-\frac{D}{x}\big)}-\frac{cI_pD\big(-\beta+e^{(-M+T)\beta}\beta\big)}{\beta^2}-(h+\mathrm{h}')\\
&\Bigg(\frac{e^{(T-\mathrm{t}_d)\beta}(1-p)D\mathrm{t}_d}{\big(1-E[p]-\frac{D}{x}\big)}-\frac{e^{(T-\mathrm{t}_d)\beta}D^2\mathrm{t}_d}{\big(1-E[p]-\frac{D}{x}\big)x}-\frac{D}{\beta}+\frac{e^{(T-\mathrm{t}_d)\beta}\mathrm{D}}{\beta}+\frac{2e^{(T-\mathrm{t}_d)\beta}D^3\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\big(1-E[p]-\frac{D}{x}\big)^2x^2\beta}+\frac{2e^{(T-\mathrm{t}_d)\beta}D^2\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\big(1-E[p]-\frac{D}{x}\big)^2x\beta}-\frac{2e^{(T-\mathrm{t}_d)\beta}(1-p)D^2\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\big(1-E[p]-\frac{D}{x}\big)^2x\beta}\Bigg)\Bigg)-\\
&\frac{1}{T^3}2\Bigg(A+\mathrm{A}'+\mathrm{I}_e\Big(\frac{M^2}{2}-\frac{N^2}{2}\Big)vD-DTv+\frac{cI_pD\big(e^{(-M+T)\beta}-\beta-(-M+T)\beta\big)}{\beta^2}+\frac{(c+\mathrm{c}')D\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\big(1-E[p]-\frac{D}{x}\big)\beta}+\frac{Ds\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\big(1-E[p]-\frac{D}{x}\big)\beta}+\\
&(c+\mathrm{c}')\Bigg(-DT+\frac{(1-E[p])D\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\big(1-E[p]-\frac{D}{x}\big)\beta}\Bigg)+(h+\mathrm{h}')\Bigg(-\frac{D\mathrm{t}_d{}^2}{2}-\frac{D}{\beta^2}+\frac{e^{(T-\mathrm{t}_d)\beta}D}{\beta^2}-\frac{D(T-\mathrm{t}_d)}{\beta}+\frac{(1-E[p])\mathrm{t}_dD\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{td}\beta\Big)}{\big(1-E[p]-\frac{D}{x}\big)\beta}-\frac{D^2\mathrm{t}_d\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\big(1-E[p]-\frac{D}{x}\big)x\beta}+\\
&\frac{D^3\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)^2}{\big(1-E[p]-\frac{D}{x}\big)^2x^2\beta^2}+\frac{D^2\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)^2}{\big(1-E[p]-\frac{D}{x}\big)^2x\beta^2}-\frac{(1-E[p])D^2\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)^2}{\big(1-E[p]-\frac{D}{x}\big)^2x\beta^2}\Bigg))+\frac{1}{T}\Bigg(-ce^{(-M+T)\beta}\mathrm{I}_pD-\frac{(c+\mathrm{c}')e^{(T-\mathrm{t}_d)\beta}D\beta}{\big(1-E[p]-\frac{D}{x}\big)}-\frac{(c+\mathrm{c}')e^{(T-\mathrm{t}_d)\beta}(1-p)D\beta}{\big(1-E[p]-\frac{D}{x}\big)}-\\
&\frac{e^{(T-\mathrm{t}_d)\beta}Ds\beta}{\big(1-E[p]-\frac{D}{x}\big)}-(h+\mathrm{h}')\Bigg(e^{(T-\mathrm{t}_d)\beta}D+\frac{2e^{2(T-\mathrm{t}_d)\beta}D^3}{\big(1-E[p]-\frac{D}{x}\big)^2x^2}+\frac{2e^{2(T-\mathrm{t}_d)\beta}D^2}{\big(1-E[p]-\frac{D}{x}\big)^2x}-\frac{2e^{2(T-\mathrm{t}_d)\beta}(1-p)\mathrm{t}_d{}^2}{\big(1-E[p]-\frac{D}{x}\big)^2x}+\frac{e^{(T-\mathrm{t}_d)\beta}(1-p)\mathrm{t}_dD\beta}{\big(1-E[p]-\frac{D}{x}\big)}-\frac{e^{(T-\mathrm{t}_d)\beta}D^2\mathrm{t}_d\beta}{\big(1-E[p]-\frac{D}{x}\big)x}+\\
&\frac{2e^{(T-\mathrm{t}_d)\beta}D^3\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\big(1-E[p]-\frac{D}{x}\big)^2x^2}+\frac{2e^{(T-\mathrm{t}_d)\beta}D^2\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\big(1-E[p]-\frac{D}{x}\big)^2x}-\frac{2e^{(T-\mathrm{t}_d)\beta}(1-E[p])D^2\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\big(1-E[p]-\frac{D}{x}\big)^2x}\Bigg))-\frac{1}{T^2}\Bigg(Dv-\frac{(c+\mathrm{c}')e^{(T-\mathrm{t}_d)\beta}D}{\big(1-E[p]-\frac{D}{x}\big)}-\frac{e^{(T-\mathrm{t}_d)\beta}Ds}{\big(1-pE[p]\frac{D}{x}\big)}-\\
&(c+\mathrm{c}')\Bigg(-D+\frac{e^{(T-\mathrm{t}_d)\beta}(1-p)D}{\big(1-E[p]-\frac{D}{x}\big)}\Bigg)-\frac{cI_pD\big(-\beta+e^{(-M+T)\beta}\beta\big)}{\beta^2}-(h+\mathrm{h}')\Bigg(-\frac{D}{\beta}+\frac{e^{(T-\mathrm{t}_d)\beta}(1-E[p])\mathrm{t}_dD}{\big(1-E[p]-\frac{D}{x}\big)}-\frac{e^{(T-\mathrm{t}_d)\beta}D^2\mathrm{t}_d}{\big(1-E[p]-\frac{D}{x}\big)x}+\frac{e^{(T-\mathrm{t}_d)\beta}D}{\beta}+\frac{2e^{(T-\mathrm{t}_d)\beta}\mathrm{D}^3\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\big(1-E[p]-\frac{D}{x}\big)^2x^2\beta}+\\
&\frac{2e^{(T-\mathrm{t}_d)\beta}D^2\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\big(1-E[p]-\frac{D}{x}\big)^2x\beta}-\frac{2e^{(T-\mathrm{t}_d)\beta}(1-E[p])D^2\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\big(1-E[p]-\frac{D}{x}\big)^2x\beta}\Bigg))
\end{aligned}
\tag{A2}
$$

$$
\begin{aligned}
\mathrm{E}[TPU]_{1.1}{}''(T) = &-\tfrac{1}{T^2}\Bigg(Dv-(c+\mathrm{c}')\Bigg(-D+\tfrac{e^{(T-\mathrm{t}_d)\beta}(1-p)D}{\left(1-E[p]-\frac{D}{x}\right)}\Bigg)-\tfrac{(c+\mathrm{c}')e^{(T-\mathrm{t}_d)\beta}D}{\left(1-E[p]-\frac{D}{x}\right)}-\tfrac{e^{(T-\mathrm{t}_d)\beta}Ds}{\left(1-E[p]-\frac{D}{x}\right)}-\tfrac{\mathrm{cI}_pD\left(-\beta+e^{(-M+T)\beta}\beta\right)}{\beta^2}-(h+\mathrm{h}')\Bigg(\tfrac{e^{(T-\mathrm{t}_d)\beta}(1-p)D\mathrm{t}_d}{\left(1-E[p]-\frac{D}{x}\right)}\\
&-\tfrac{e^{(T-\mathrm{t}_d)\beta}D^2\mathrm{t}_d}{\left(1-E[p]-\frac{D}{x}\right)x}-\tfrac{D}{\beta}+\tfrac{e^{(T-\mathrm{t}_d)\beta}\mathrm{D}}{\beta}+\tfrac{2e^{(T-\mathrm{t}_d)\beta}D^3\left(-1+e^{(T-\mathrm{t}_d)\beta+\mathrm{t}_d\beta}\right)}{\left(1-E[p]-\frac{D}{x}\right)^2x^2\beta}+\tfrac{2e^{(T-\mathrm{t}_d)\beta}D^2\left(-1+e^{(T-\mathrm{t}_d)\beta+\mathrm{t}_d\beta}\right)}{\left(1-E[p]-\frac{D}{x}\right)^2x\beta}-\tfrac{2e^{(T-\mathrm{t}_d)\beta}(1-p)D^2\left(-1+e^{(T-\mathrm{t}_d)\beta+\mathrm{t}_d\beta}\right)}{\left(1-E[p]-\frac{D}{x}\right)^2x\beta}\Bigg))-\\
&\tfrac{1}{T^3}2\Bigg(A+\mathrm{A}'+\mathrm{I}_e\left(\tfrac{M^2}{2}-\tfrac{N^2}{2}\right)vD-DTv+\tfrac{\mathrm{cI}_pD\left(e^{(-M+T)\beta}-\beta-(-M+T)\beta\right)}{\beta^2}+\tfrac{(c+\mathrm{c}')D\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)\beta}+\tfrac{Ds\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)\beta}+\\
&(c+\mathrm{c}')\Bigg(-DT+\tfrac{(1-E[p])D\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)\beta}\Bigg)+(h+\mathrm{h}')\Bigg(-\tfrac{D\mathrm{t}_d{}^2}{2}-\tfrac{D}{\beta^2}+\tfrac{e^{(T-\mathrm{t}_d)\beta}D}{\beta^2}-\tfrac{D(T-\mathrm{t}_d)}{\beta}+\tfrac{(1-E[p])\mathrm{t}_dD\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{td}\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)\beta}-\tfrac{D^2\mathrm{t}_d\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)x\beta}+\\
&\tfrac{D^3\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)^2}{\left(1-E[p]-\frac{D}{x}\right)^2x^2\beta^2}+\tfrac{D^2\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)^2}{\left(1-E[p]-\frac{D}{x}\right)^2x\beta^2}-\tfrac{(1-E[p])D^2\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)^2}{\left(1-E[p]-\frac{D}{x}\right)^2x\beta^2}\Bigg))+\tfrac{1}{T}\Bigg(-ce^{(-M+T)\beta}\mathrm{I}_pD-\tfrac{(c+\mathrm{c}')e^{(T-\mathrm{t}_d)\beta}D\beta}{\left(1-E[p]-\frac{D}{x}\right)}-\tfrac{(c+\mathrm{c}')e^{(T-\mathrm{t}_d)\beta}(1-p)D\beta}{\left(1-E[p]-\frac{D}{x}\right)}-\\
&\tfrac{e^{(T-\mathrm{t}_d)\beta}Ds\beta}{\left(1-E[p]-\frac{D}{x}\right)}-(h+\mathrm{h}')\Bigg(e^{(T-\mathrm{t}_d)\beta}D+\tfrac{2e^{2(T-\mathrm{t}_d)\beta}D^3}{\left(1-E[p]-\frac{D}{x}\right)^2x^2}+\tfrac{2e^{2(T-\mathrm{t}_d)\beta}D^2}{\left(1-E[p]-\frac{D}{x}\right)^2x}-\tfrac{2e^{2(T-\mathrm{t}_d)\beta}(1-p)\mathrm{t}_d{}^2}{\left(1-E[p]-\frac{D}{x}\right)^2x}+\tfrac{e^{(T-\mathrm{t}_d)\beta}(1-p)\mathrm{t}_dD\beta}{\left(1-E[p]-\frac{D}{x}\right)}-\tfrac{e^{(T-\mathrm{t}_d)\beta}D^2\mathrm{t}_d\beta}{\left(1-E[p]-\frac{D}{x}\right)x}+\\
&\tfrac{2e^{(T-\mathrm{t}_d)\beta}D^3\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)^2x^2}+\tfrac{2e^{(T-\mathrm{t}_d)\beta}D^2\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)^2x}-\tfrac{2e^{(T-\mathrm{t}_d)\beta}(1-E[p])D^2\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)^2x}\Bigg))-\tfrac{1}{T^2}\Bigg(Dv-\tfrac{(c+\mathrm{c}')e^{(T-\mathrm{t}_d)\beta}D}{\left(1-E[p]-\frac{D}{x}\right)}-\tfrac{e^{(T-\mathrm{t}_d)\beta}Ds}{\left(1-pE[p]\frac{D}{x}\right)}-\\
&(c+\mathrm{c}')\Bigg(-D+\tfrac{e^{(T-\mathrm{t}_d)\beta}(1-p)D}{\left(1-E[p]-\frac{D}{x}\right)}\Bigg)-\tfrac{\mathrm{cI}_pD\left(-\beta+e^{(-M+T)\beta}\beta\right)}{\beta^2}-(h+\mathrm{h}')\Bigg(-\tfrac{D}{\beta}+\tfrac{e^{(T-\mathrm{t}_d)\beta}(1-E[p])\mathrm{t}_dD}{\left(1-E[p]-\frac{D}{x}\right)}-\tfrac{e^{(T-\mathrm{t}_d)\beta}D^2\mathrm{t}_d}{\left(1-E[p]-\frac{D}{x}\right)x}+\tfrac{e^{(T-\mathrm{t}_d)\beta}D}{\beta}+\tfrac{2e^{(T-\mathrm{t}_d)\beta}\mathrm{D}^3\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)^2x^2\beta}+\\
&\tfrac{2e^{(T-\mathrm{t}_d)\beta}D^2\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)^2x\beta}-\tfrac{2e^{(T-\mathrm{t}_d)\beta}(1-E[p])D^2\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)^2x\beta}\Bigg))
\end{aligned}
\tag{A3}
$$

$$
\begin{aligned}
\mathrm{E}[TPU]_{1.1}{}''(T) = &-\tfrac{1}{T^2}\Bigg(Dv-(c+\mathrm{c}')\Bigg(-D+\tfrac{e^{(T-\mathrm{t}_d)\beta}(1-p)D}{\left(1-E[p]-\frac{D}{x}\right)}\Bigg)-\tfrac{(c+\mathrm{c}')e^{(T-\mathrm{t}_d)\beta}D}{\left(1-E[p]-\frac{D}{x}\right)}-\tfrac{e^{(T-\mathrm{t}_d)\beta}Ds}{\left(1-E[p]-\frac{D}{x}\right)}-\tfrac{\mathrm{cI}_pD\left(-\beta+e^{(-M+T)\beta}\beta\right)}{\beta^2}-(h+\mathrm{h}')\Bigg(\tfrac{e^{(T-\mathrm{t}_d)\beta}(1-p)D\mathrm{t}_d}{\left(1-E[p]-\frac{D}{x}\right)}-\\
&\tfrac{e^{(T-\mathrm{t}_d)\beta}D^2\mathrm{t}_d}{\left(1-E[p]-\frac{D}{x}\right)x}-\tfrac{D}{\beta}+\tfrac{e^{(T-\mathrm{t}_d)\beta}\mathrm{D}}{\beta}+\tfrac{2e^{(T-\mathrm{t}_d)\beta}D^3\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)^2x^2\beta}+\tfrac{2e^{(T-\mathrm{t}_d)\beta}D^2\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)^2x\beta}-\tfrac{2e^{(T-\mathrm{t}_d)\beta}(1-p)D^2\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)^2x\beta}\Bigg))-\\
&\tfrac{1}{T^3}2\Bigg(A+\mathrm{A}'+\mathrm{I}_e\left(\tfrac{M^2}{2}-\tfrac{N^2}{2}\right)vD-DTv+\tfrac{\mathrm{cI}_pD\left(e^{(-M+T)\beta}-\beta-(-M+T)\beta\right)}{\beta^2}+\tfrac{(c+\mathrm{c}')D\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)\beta}+\tfrac{Ds\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)\beta}+\\
&(c+\mathrm{c}')\Bigg(-DT+\tfrac{(1-E[p])D\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)\beta}\Bigg)+(h+\mathrm{h}')\Bigg(-\tfrac{D\mathrm{t}_d{}^2}{2}-\tfrac{D}{\beta^2}+\tfrac{e^{(T-\mathrm{t}_d)\beta}D}{\beta^2}-\tfrac{D(T-\mathrm{t}_d)}{\beta}+\tfrac{(1-E[p])\mathrm{t}_dD\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{td}\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)\beta}-\tfrac{D^2\mathrm{t}_d\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)x\beta}+\\
&\tfrac{D^3\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)^2}{\left(1-E[p]-\frac{D}{x}\right)^2x^2\beta^2}+\tfrac{D^2\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)^2}{\left(1-E[p]-\frac{D}{x}\right)^2x\beta^2}-\tfrac{(1-E[p])D^2\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)^2}{\left(1-E[p]-\frac{D}{x}\right)^2x\beta^2}\Bigg))+\tfrac{1}{T}\Bigg(-ce^{(-M+T)\beta}\mathrm{I}_pD-\tfrac{(c+\mathrm{c}')e^{(T-\mathrm{t}_d)\beta}D\beta}{\left(1-E[p]-\frac{D}{x}\right)}-\tfrac{(c+\mathrm{c}')e^{(T-\mathrm{t}_d)\beta}(1-p)D\beta}{\left(1-E[p]-\frac{D}{x}\right)}-\\
&\tfrac{e^{(T-\mathrm{t}_d)\beta}Ds\beta}{\left(1-E[p]-\frac{D}{x}\right)}-(h+\mathrm{h}')\Bigg(e^{(T-\mathrm{t}_d)\beta}D+\tfrac{2e^{2(T-\mathrm{t}_d)\beta}D^3}{\left(1-E[p]-\frac{D}{x}\right)^2x^2}+\tfrac{2e^{2(T-\mathrm{t}_d)\beta}D^2}{\left(1-E[p]-\frac{D}{x}\right)^2x}-\tfrac{2e^{2(T-\mathrm{t}_d)\beta}(1-p)\mathrm{t}_d{}^2}{\left(1-E[p]-\frac{D}{x}\right)^2x}+\tfrac{e^{(T-\mathrm{t}_d)\beta}(1-p)\mathrm{t}_dD\beta}{\left(1-E[p]-\frac{D}{x}\right)}-\tfrac{e^{(T-\mathrm{t}_d)\beta}D^2\mathrm{t}_d\beta}{\left(1-E[p]-\frac{D}{x}\right)x}+\\
&\tfrac{2e^{(T-\mathrm{t}_d)\beta}D^3\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)^2x^2}+\tfrac{2e^{(T-\mathrm{t}_d)\beta}D^2\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)^2x}-\tfrac{2e^{(T-\mathrm{t}_d)\beta}(1-E[p])D^2\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)^2x}\Bigg))-\tfrac{1}{T^2}\Bigg(Dv-\tfrac{(c+\mathrm{c}')e^{(T-\mathrm{t}_d)\beta}D}{\left(1-E[p]-\frac{D}{x}\right)}-\tfrac{e^{(T-\mathrm{t}_d)\beta}Ds}{\left(1-pE[p]\frac{D}{x}\right)}-\\
&(c+\mathrm{c}')\Bigg(-D+\tfrac{e^{(T-\mathrm{t}_d)\beta}(1-p)D}{\left(1-E[p]-\frac{D}{x}\right)}\Bigg)-\tfrac{\mathrm{cI}_pD\left(-\beta+e^{(-M+T)\beta}\beta\right)}{\beta^2}-(h+\mathrm{h}')\Bigg(-\tfrac{D}{\beta}+\tfrac{e^{(T-\mathrm{t}_d)\beta}(1-E[p])\mathrm{t}_dD}{\left(1-E[p]-\frac{D}{x}\right)}-\tfrac{e^{(T-\mathrm{t}_d)\beta}D^2\mathrm{t}_d}{\left(1-E[p]-\frac{D}{x}\right)x}+\tfrac{e^{(T-\mathrm{t}_d)\beta}D}{\beta}+\tfrac{2e^{(T-\mathrm{t}_d)\beta}\mathrm{D}^3\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)^2x^2\beta}+\\
&\tfrac{2e^{(T-\mathrm{t}_d)\beta}D^2\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)^2x\beta}-\tfrac{2e^{(T-\mathrm{t}_d)\beta}(1-E[p])D^2\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)^2x\beta}\Bigg))
\end{aligned}
\tag{A4}
$$

$$E[TPU]_{1.6}{}' = -\frac{1}{T^3}2\Bigg(A + \mathrm{A}' + \mathrm{I}_e\Big(\frac{M^2}{2} - \frac{N^2}{2}\Big)vD - DTv + c\mathrm{I}_P\Bigg(\frac{M^2D}{2} + \frac{\Big(-1+e^{(T-\mathrm{t}_d)\beta}\Big)D}{\beta^2} + \frac{D(-T+\mathrm{t}_d)}{\beta} - \frac{D^3\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)^2}{2\Big(1-E[p]-\frac{D}{x}\Big)^2x^2\beta^2} + \frac{(1-E[p])D\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)\Bigg(\mathrm{t}_d-\frac{D\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\Big(1-E[p]-\frac{D}{x}\Big)x\beta}\Bigg)}{\Big(1-E[p]-\frac{D}{x}\Big)\beta} + \frac{D\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)\Bigg(-M+\frac{D\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\Big(1-E[p]-\frac{D}{x}\Big)x\beta}\Bigg)}{\Big(1-E[p]-\frac{D}{x}\Big)\beta} + \frac{D^2\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)\Bigg(-\mathrm{t}_d+\frac{D\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\Big(1-E[p]-\frac{D}{x}\Big)x\beta}\Bigg)}{\Big(1-E[p]-\frac{D}{x}\Big)x\beta} + \frac{1}{2}D\Bigg(-\mathrm{t}_d{}^2 + \frac{D^2\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)^2}{\Big(1-E[p]-\frac{D}{x}\Big)^2x^2\beta^2}\Bigg)\Bigg) + \frac{(c+\mathrm{c}')D\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\Big(1-E[p]-\frac{D}{x}\Big)\beta} + \frac{Ds\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\Big(1-E[p]-\frac{D}{x}\Big)\beta} + (c+\mathrm{c}')\Bigg(-DT + \frac{(1-E[p])D\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\Big(1-E[p]-\frac{D}{x}\Big)\beta}\Bigg) + (h+\mathrm{h}')\Bigg(-\frac{D\mathrm{t}_d{}^2}{2} - \frac{D}{\beta^2} + \frac{e^{(T-\mathrm{t}_d)\beta}D}{\beta^2} - \frac{D(T-\mathrm{t}_d)}{\beta} + -\frac{D^2\mathrm{t}_d\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\Big(1-E[p]-\frac{D}{x}\Big)x\beta} + \frac{D^3\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)^2}{\Big(1-E[p]-\frac{D}{x}\Big)^2x^2\beta^2} + \frac{D^2\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)^2}{\Big(1-E[p]-\frac{D}{x}\Big)^2x\beta^2} - \frac{(1-p)D^2\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)^2}{\Big(1-E[p]-\frac{D}{x}\Big)^2x\beta^2}\Bigg)\Bigg) + \frac{1}{T}\Bigg(-c\mathrm{Ip}\Bigg(e^{(T-\mathrm{t}_d)\beta}D + \frac{e^{(T-\mathrm{t}_d)\beta}(1-p)D(-M+\mathrm{t}_d)\beta}{\Big(1-E[p]-\frac{D}{x}\Big)} + \frac{e^{(T-\mathrm{t}_d)\beta}D^2(M-\mathrm{t}_d)\beta}{\Big(1-E[p]-\frac{D}{x}\Big)x}\Bigg) - \frac{(c+\mathrm{c}')e^{(T-\mathrm{t}_d)\beta}D\beta}{\Big(1-E[p]-\frac{D}{x}\Big)} - \frac{(c+\mathrm{c}')e^{(T-\mathrm{t}_d)\beta}(1-p)D\beta}{\Big(1-E[p]-\frac{D}{x}\Big)} - \frac{e^{(T-\mathrm{t}_d)\beta}Ds\beta}{\Big(1-E[p]-\frac{D}{x}\Big)} - (h+\mathrm{h}')\Bigg(e^{(T-\mathrm{t}_d)\beta}D + \frac{2e^{2(T-\mathrm{t}_d)\beta}D^3}{\Big(1-E[p]-\frac{D}{x}\Big)^2x^2} + \frac{2e^{2(T-\mathrm{t}_d)\beta}D^2}{\Big(1-E[p]-\frac{D}{x}\Big)^2x} - \frac{2e^{2(T-\mathrm{t}_d)\beta}(1-p)\mathrm{t}_d{}^2}{\Big(1-E[p]-\frac{D}{x}\Big)^2x} + \frac{e^{(T-\mathrm{t}_d)\beta}(1-p)\mathrm{t}_dD\beta}{\Big(1-E[p]-\frac{D}{x}\Big)} - \frac{e^{(T-\mathrm{t}_d)\beta}D^2\mathrm{t}_d\beta}{\Big(1-E[p]-\frac{D}{x}\Big)x} + \frac{2e^{(T-\mathrm{t}_d)\beta}D^3\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\Big(1-E[p]-\frac{D}{x}\Big)^2x^2} + \frac{2e^{(T-\mathrm{t}_d)\beta}D^2\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\Big(1-E[p]-\frac{D}{x}\Big)^2x} - \frac{2e^{(T-\mathrm{t}_d)\beta}(1-p)D^2\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\Big(1-E[p]-\frac{D}{x}\Big)^2x}\Bigg)\Bigg) - \frac{2}{T^2}\Bigg(Dv - \frac{(c+\mathrm{c}')e^{(T-\mathrm{t}_d)\beta}D}{\Big(1-E[p]-\frac{D}{x}\Big)} - \frac{e^{(T-\mathrm{t}_d)\beta}Ds}{\Big(1-E[p]-\frac{D}{x}\Big)} - (c+\mathrm{c}')\Bigg(-D + \frac{e^{(T-\mathrm{t}_d)\beta}(1-p)D}{\Big(1-E[p]-\frac{D}{x}\Big)}\Bigg) - c\mathrm{Ip}\Bigg(\frac{M^2D}{2} + \frac{\Big(-1+e^{(T-\mathrm{t}_d)\beta}\Big)D}{\beta^2} + \frac{D(-T+\mathrm{t}_d)}{\beta} - \frac{D^3\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)^2}{2\Big(1-E[p]-\frac{D}{x}\Big)^2x^2\beta^2} + \frac{(1-p)D\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)\Bigg(\mathrm{t}_d-\frac{D\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\Big(1-E[p]-\frac{D}{x}\Big)x\beta}\Bigg)}{\Big(1-E[p]-\frac{D}{x}\Big)\beta} + \frac{D\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)\Bigg(-M+\frac{D\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\Big(1-E[p]-\frac{D}{x}\Big)x\beta}\Bigg)}{\Big(1-E[p]-\frac{D}{x}\Big)\beta} + \frac{D^2\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)\Bigg(-\mathrm{t}_d+\frac{D\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\Big(1-E[p]-\frac{D}{x}\Big)x\beta}\Bigg)}{\Big(1-E[p]-\frac{D}{x}\Big)x\beta} + \frac{1}{2}D\Bigg(-\mathrm{t}_d{}^2 + \frac{D^2\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)^2}{\Big(1-E[p]-\frac{D}{x}\Big)^2x^2\beta^2}\Bigg)\Bigg) - (h+\mathrm{h}')\Bigg(-\frac{D}{\beta} + \frac{e^{(T-\mathrm{t}_d)\beta}(1-p)\mathrm{t}_dD}{\Big(1-E[p]-\frac{D}{x}\Big)} - \frac{e^{(T-\mathrm{t}_d)\beta}D^2\mathrm{t}_d}{\Big(1-E[p]-\frac{D}{x}\Big)x} + \frac{e^{(T-\mathrm{t}_d)\beta}D}{\beta} + \frac{2e^{(T-\mathrm{t}_d)\beta}D^3\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\Big(1-E[p]-\frac{D}{x}\Big)^2x^2\beta} + \frac{2e^{(T-\mathrm{t}_d)\beta}D^2\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\Big(1-E[p]-\frac{D}{x}\Big)^2x\beta} - \frac{2e^{(T-\mathrm{t}_d)\beta}(1-p)D^2\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\Big(1-E[p]-\frac{D}{x}\Big)^2x\beta}\Bigg)\Bigg) \tag{A5}$$

$$TP_{2.1}{}'(T) = -\frac{1}{T^2}\Bigg(-A - \mathrm{A}' + \mathrm{I}_e\Big(-\frac{N^2}{2} - \frac{T^2}{2} + MT\Big)vD + DTv - \frac{(c+\mathrm{c}')D\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\Big(1-E[p]-\frac{D}{x}\Big)\beta} - \frac{Ds\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\Big(1-E[p]-\frac{D}{x}\Big)\beta} - (c+\mathrm{c}')\Bigg(-DT + \frac{(1-E[p])D\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\Big(1-E[p]-\frac{D}{x}\Big)\beta}\Bigg) - (h+\mathrm{h}')\Bigg(-\frac{D\mathrm{t}_d{}^2}{2} - \frac{D}{\beta^2} + \frac{e^{(T-\mathrm{t}_d)\beta}D}{\beta^2} - \frac{D(T-\mathrm{t}_d)}{\beta} + \frac{(1-p)D\mathrm{t}_d\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\Big(1-E[p]-\frac{D}{x}\Big)\beta} - \frac{D^2\mathrm{t}_d\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\Big(1-E[p]-\frac{D}{x}\Big)x\beta} + \frac{D^3\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)^2}{\Big(1-E[p]-\frac{D}{x}\Big)^2x^2\beta^2} + \frac{D^2\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)^2}{\Big(1-E[p]-\frac{D}{x}\Big)^2x\beta^2} - \frac{(1-p)D^2\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)^2}{\Big(1-E[p]-\frac{D}{x}\Big)^2x\beta^2}\Bigg)\Bigg) + \frac{1}{T}\Bigg(Dv + \mathrm{I}_evD(M-T) - \frac{(c+\mathrm{c}')e^{(T-\mathrm{t}_d)\beta}D}{\Big(1-E[p]-\frac{D}{x}\Big)} - \frac{e^{(T-\mathrm{t}_d)\beta}Ds}{\Big(1-E[p]-\frac{D}{x}\Big)} - (c+\mathrm{c}')\Bigg(-D + \frac{e^{(T-\mathrm{t}_d)\beta}(1-p)D}{\Big(1-E[p]-\frac{D}{x}\Big)}\Bigg) - (h+\mathrm{h}')\Bigg(-\frac{D}{\beta} + \frac{e^{(T-\mathrm{t}_d)\beta}(1-p)D\mathrm{t}_d}{\Big(1-E[p]-\frac{D}{x}\Big)} - \frac{e^{(T-\mathrm{t}_d)\beta}D^2\mathrm{t}_d}{\Big(1-E[p]-\frac{D}{x}\Big)x} + \frac{e^{(T-\mathrm{t}_d)\beta}D}{\beta} + \frac{2e^{(T-\mathrm{t}_d)\beta}D^3\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\Big(1-E[p]-\frac{D}{x}\Big)^2x^2\beta} + \frac{2e^{(T-\mathrm{t}_d)\beta}D^2\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\Big(1-E[p]-\frac{D}{x}\Big)^2x\beta} - \frac{2e^{(T-\mathrm{t}_d)\beta}(1-p)D^2\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\Big(1-E[p]-\frac{D}{x}\Big)^2x\beta}\Bigg)\Bigg) \tag{A6}$$

$$E[TPU_{2.1}]''(T) = -\frac{1}{T^3}2\Bigg(A + \mathrm{A}' + \mathrm{I}_e\Big(-\frac{N^2}{2} + MT - \frac{T^2}{2}\Big)vD - DTv + \frac{(c+\mathrm{c}')D\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\Big(1-E[p]-\frac{D}{x}\Big)\beta} + \frac{Ds\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\Big(1-E[p]-\frac{D}{x}\Big)\beta} + (c+\mathrm{c}')\Bigg(-DT + \frac{(1-p)D\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\Big(1-E[p]-\frac{D}{x}\Big)\beta}\Bigg) + (h+\mathrm{h}')\Bigg(-\frac{D\mathrm{t}_d{}^2}{2} - \frac{D}{\beta^2} + \frac{e^{(T-\mathrm{t}_d)\beta}D}{\beta^2} - \frac{D(T-\mathrm{t}_d)}{\beta} + \frac{(1-p)\mathrm{t}_dD\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{td}\beta\Big)}{\Big(1-E[p]-\frac{D}{x}\Big)\beta} - \frac{D^2\mathrm{t}_d\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\Big(1-E[p]-\frac{D}{x}\Big)x\beta} + \frac{D^3\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)^2}{\Big(1-E[p]-\frac{D}{x}\Big)^2x^2\beta^2} + \frac{D^2\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)^2}{\Big(1-E[p]-\frac{D}{x}\Big)^2x\beta^2} - \frac{(1-p)D^2\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)^2}{\Big(1-E[p]-\frac{D}{x}\Big)^2x\beta^2}\Bigg)\Bigg) + \frac{1}{T}\Bigg(-\mathrm{I}_evD - \frac{(c+\mathrm{c}')e^{(T-\mathrm{t}_d)\beta}D\beta}{\Big(1-E[p]-\frac{D}{x}\Big)} - \frac{(c+\mathrm{c}')e^{(T-\mathrm{t}_d)\beta}(1-p)D\beta}{\Big(1-E[p]-\frac{D}{x}\Big)} - \frac{e^{(T-\mathrm{t}_d)\beta}Ds\beta}{\Big(1-E[p]-\frac{D}{x}\Big)} - (h+\mathrm{h}')\Bigg(e^{(T-\mathrm{t}_d)\beta}D + \frac{2e^{2(T-\mathrm{t}_d)\beta}D^3}{\Big(1-E[p]-\frac{D}{x}\Big)^2x^2} + \frac{2e^{2(T-\mathrm{t}_d)\beta}D^2}{\Big(1-E[p]-\frac{D}{x}\Big)^2x} - \frac{2e^{2(T-\mathrm{t}_d)\beta}(1-p)\mathrm{t}_d{}^2}{\Big(1-E[p]-\frac{D}{x}\Big)^2x} + \frac{e^{(T-\mathrm{t}_d)\beta}(1-p)\mathrm{t}_dD\beta}{\Big(1-E[p]-\frac{D}{x}\Big)} - \frac{e^{(T-\mathrm{t}_d)\beta}D^2\mathrm{t}_d\beta}{\Big(1-E[p]-\frac{D}{x}\Big)x} + \frac{2e^{(T-\mathrm{t}_d)\beta}D^3\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\Big(1-E[p]-\frac{D}{x}\Big)^2x^2} + \frac{2e^{(T-\mathrm{t}_d)\beta}D^2\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\Big(1-E[p]-\frac{D}{x}\Big)^2x} - \frac{2e^{(T-\mathrm{t}_d)\beta}(1-p)D^2\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\Big(1-E[p]-\frac{D}{x}\Big)^2x}\Bigg)\Bigg) - \frac{2}{T^2}\Bigg(Dv - \frac{(c+\mathrm{c}')e^{(T-\mathrm{t}_d)\beta}D}{\Big(1-E[p]-\frac{D}{x}\Big)} - \frac{e^{(T-\mathrm{t}_d)\beta}Ds}{\Big(1-E[p]-\frac{D}{x}\Big)} - (c+\mathrm{c}')\Bigg(-D + \frac{e^{(T-\mathrm{t}_d)\beta}(1-p)D}{\Big(1-E[p]-\frac{D}{x}\Big)}\Bigg) + \mathrm{I}_epD(M-T) - (h+\mathrm{h}')\Bigg(-\frac{D}{\beta} + \frac{e^{(T-\mathrm{t}_d)\beta}(1-p)\mathrm{t}_dD}{\Big(1-E[p]-\frac{D}{x}\Big)} - \frac{e^{(T-\mathrm{t}_d)\beta}D^2\mathrm{t}_d}{\Big(1-E[p]-\frac{D}{x}\Big)x} + \frac{e^{(T-\mathrm{t}_d)\beta}D}{\beta} + \frac{2e^{(T-\mathrm{t}_d)\beta}D^3\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\Big(1-E[p]-\frac{D}{x}\Big)^2x^2\beta} + \frac{2e^{(T-\mathrm{t}_d)\beta}D^2\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\Big(1-E[p]-\frac{D}{x}\Big)^2x\beta} - \frac{2e^{(T-\mathrm{t}_d)\beta}(1-p)D^2\Big(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\Big)}{\Big(1-E[p]-\frac{D}{x}\Big)^2x\beta}\Bigg)\Bigg) \tag{A7}$$

$$TP_{3.1}{}'(T) = -\frac{1}{T^2}\Bigg(-A - \mathrm{A}' + \mathrm{I}_e(M-N)pDT + DTv - \frac{(c+\mathrm{c}')D\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)\beta} - \frac{Ds\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)\beta} -$$
$$(c+\mathrm{c}')\left(-DT + \frac{(1-p)D\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)\beta}\right) -$$
$$(h+\mathrm{h}')\Bigg(-\frac{D\mathrm{t}_d{}^2}{2} - \frac{D}{\beta^2} + \frac{e^{(T-\mathrm{t}_d)\beta}D}{\beta^2} - \frac{D(T-\mathrm{t}_d)}{\beta} + \frac{(1-p)D\mathrm{t}_d\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)\beta} - \frac{D^2\mathrm{t}_d\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)x\beta} +$$
$$\frac{D^3\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)^2}{\left(1-E[p]-\frac{D}{x}\right)^2x^2\beta^2} + \frac{D^2\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)^2}{\left(1-E[p]-\frac{D}{x}\right)^2x\beta^2} - \frac{(1-p)D^2\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)^2}{\left(1-E[p]-\frac{D}{x}\right)^2x\beta^2}\Bigg)\Bigg) + \tag{A8}$$
$$\frac{1}{T}\Bigg(Dv + \mathrm{I}_e vD(M-N) - \frac{(c+\mathrm{c}')e^{(T-\mathrm{t}_d)\beta}D}{\left(1-E[p]-\frac{D}{x}\right)} - \frac{e^{(T-\mathrm{t}_d)\beta}Ds}{\left(1-E[p]-\frac{D}{x}\right)} - (c+\mathrm{c}')\left(-D + \frac{e^{(T-\mathrm{t}_d)\beta}(1-p)D}{\left(1-E[p]-\frac{D}{x}\right)}\right) -$$
$$(h+\mathrm{h}')\Bigg(-\frac{D}{\beta} + \frac{e^{(T-\mathrm{t}_d)\beta}(1-p)D\mathrm{t}_d}{\left(1-E[p]-\frac{D}{x}\right)} - \frac{e^{(T-\mathrm{t}_d)\beta}D^2\mathrm{t}_d}{\left(1-E[p]-\frac{D}{x}\right)x} + \frac{e^{(T-\mathrm{t}_d)\beta}D}{\beta} + \frac{2e^{(T-\mathrm{t}_d)\beta}D^3\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)^2x^2\beta} +$$
$$\frac{2e^{(T-\mathrm{t}_d)\beta}D^2\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)^2x\beta} - \frac{2e^{(T-\mathrm{t}_d)\beta}(1-p)D^2\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)^2x\beta}\Bigg)\Bigg)$$

$$TP_{3.1}{}''(T) = -\frac{1}{T^3}2\Bigg(A + \mathrm{A}' + \mathrm{I}_e(M-N)pDT - DTv + \frac{(c+\mathrm{c}')D\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)\beta} + \frac{Ds\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)\beta} +$$
$$(c+\mathrm{c}')\left(-DT + \frac{(1-p)D\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)\beta}\right) +$$
$$(h+\mathrm{h}')\Bigg(-\frac{D\mathrm{t}_d{}^2}{2} - \frac{D}{\beta^2} + \frac{e^{(T-\mathrm{t}_d)\beta}D}{\beta^2} - \frac{D(T-\mathrm{t}_d)}{\beta} + \frac{(1-p)\mathrm{t}_dD\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{td}\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)\beta} - \frac{D^2\mathrm{t}_d\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)x\beta} +$$
$$\frac{D^3\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)^2}{\left(1-E[p]-\frac{D}{x}\right)^2x^2\beta^2} + \frac{D^2\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)^2}{\left(1-E[p]-\frac{D}{x}\right)^2x\beta^2} - \frac{(1-p)D^2\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)^2}{\left(1-E[p]-\frac{D}{x}\right)^2x\beta^2}\Bigg)\Bigg) +$$
$$\frac{1}{T}\Bigg(-\frac{(c+\mathrm{c}')e^{(T-\mathrm{t}_d)\beta}D\beta}{\left(1-E[p]-\frac{D}{x}\right)} - \frac{(c+\mathrm{c}')e^{(T-\mathrm{t}_d)\beta}(1-p)D\beta}{\left(1-E[p]-\frac{D}{x}\right)} - \frac{e^{(T-\mathrm{t}_d)\beta}Ds\beta}{\left(1-E[p]-\frac{D}{x}\right)} -$$
$$(h+\mathrm{h}')\Bigg(e^{(T-\mathrm{t}_d)\beta}D + \frac{2e^{2(T-\mathrm{t}_d)\beta}D^3}{\left(1-E[p]-\frac{D}{x}\right)^2x^2} + \frac{2e^{2(T-\mathrm{t}_d)\beta}D^2}{\left(1-E[p]-\frac{D}{x}\right)^2x} - \frac{2e^{2(T-\mathrm{t}_d)\beta}(1-p)\mathrm{t}_d{}^2D}{\left(1-E[p]-\frac{D}{x}\right)^2x} + \frac{e^{(T-\mathrm{t}_d)\beta}(1-p)\mathrm{t}_dD\beta}{\left(1-E[p]-\frac{D}{x}\right)} -$$
$$\frac{e^{(T-\mathrm{t}_d)\beta}D^2\mathrm{t}_d\beta}{\left(1-E[p]-\frac{D}{x}\right)x} + \frac{2e^{(T-\mathrm{t}_d)\beta}D^3\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)^2x^2} + \frac{2e^{(T-\mathrm{t}_d)\beta}D^2\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)^2x} -$$
$$\frac{2e^{(T-\mathrm{t}_d)\beta}(1-p)D^2\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)^2x}\Bigg)\Bigg) - \tag{A9}$$
$$\frac{2}{T^2}\Bigg(Dv + \mathrm{I}_e(M-N)pD - \frac{(c+\mathrm{c}')e^{(T-\mathrm{t}_d)\beta}D}{\left(1-E[p]-\frac{D}{x}\right)} - \frac{e^{(T-\mathrm{t}_d)\beta}Ds}{\left(1-E[p]-\frac{D}{x}\right)} - (c+\mathrm{c}')\left(-D + \frac{e^{(T-\mathrm{t}_d)\beta}(1-p)D}{\left(1-E[p]-\frac{D}{x}\right)}\right) -$$
$$(h+\mathrm{h}')\Bigg(-\frac{D}{\beta} + \frac{e^{(T-\mathrm{t}_d)\beta}(1-p)\mathrm{t}_dD}{\left(1-E[p]-\frac{D}{x}\right)} - \frac{e^{(T-\mathrm{t}_d)\beta}D^2\mathrm{t}_d}{\left(1-E[p]-\frac{D}{x}\right)x} + \frac{e^{(T-\mathrm{t}_d)\beta}D}{\beta} + \frac{2e^{(T-\mathrm{t}_d)\beta}\mathrm{D}^3\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)^2x^2\beta} +$$
$$\frac{2e^{(T-\mathrm{t}_d)\beta}D^2\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)^2x\beta} - \frac{2e^{(T-\mathrm{t}_d)\beta}(1-p)D^2\left(-1+e^{(T-\mathrm{t}_d)\beta}+\mathrm{t}_d\beta\right)}{\left(1-E[p]-\frac{D}{x}\right)^2x\beta}\Bigg)\Bigg)$$

References

1. Ghare, P.M.; Schrader, G.F. An inventory model for exponentially deteriorating items. *J. Ind. Eng.* **1963**, *14*, 238–243.
2. Alsaedi, B.S.; Alamri, O.A.; Jayaswal, M.K.; Mittal, M. A sustainable green supply chain model with carbon emissions for defective items under learning in a fuzzy environment. *Mathematics* **2023**, *11*, 301. [CrossRef]
3. Alamri, O.A.; Lamba, N.K.; Jayaswal, M.K.; Mittal, M. A sustainable inventory model with advertisement effort for imperfect quality items under learning in fuzzy monsoon demand. *Mathematics* **2024**, *12*, 2432. [CrossRef]
4. Jaggi, C.K.; Goel, S.K.; Mittal, M. Credit financing in economic ordering policies for defective items with allowable shortages. *Appl. Math. Comput.* **2013**, *219*, 5268–5282. [CrossRef]
5. Sarkar, B.; Mahapatra, A.S. Periodic review fuzzy inventory model with variable lead time and fuzzy demand. *Int. Trans. Oper. Res.* **2017**, *24*, 1197–1227. [CrossRef]

6. Shah, N.H.; Keswani, M.; Khedlekar, U.K.; Prajapati, N.M. Non-instantaneous controlled deteriorating inventory model for stock-price-advertisement dependent probabilistic demand under trade credit financing. *Opsearch* **2024**, *61*, 421–459. [CrossRef]
7. Ahmed, W.; Sarkar, B. Impact of carbon emissions in a sustainable supply chain management for a second generation biofuel. *J. Clean. Prod.* **2018**, *186*, 807–820. [CrossRef]
8. Sarkar, B.; Gupta, H.; Chaudhuri, K.; Goyal, S.K. An integrated inventory model with variable lead time, defective units and delay in payments. *Appl. Math. Comput.* **2014**, *237*, 650–658. [CrossRef]
9. Dye, C.Y.; Yang, C.T. Sustainable trade credit and replenishment decisions with credit-linked demand under carbon emission constraints. *Eur. J. Oper. Res.* **2015**, *244*, 187–200. [CrossRef]
10. Kazemi, N.; Abdul-Rashid, S.H.; Ghazilla, R.A.R.; Shekarian, E.; Zanoni, S. Economic order quantity models for items with imperfect quality and emission considerations. *Int. J. Syst. Sci. Oper. Logist.* **2018**, *5*, 99–115. [CrossRef]
11. Tayyab, M.; Sarkar, B. Optimal batch quantity in a cleaner multi-stage lean production system with random defective rate. *J. Clean. Prod.* **2016**, *139*, 922–934. [CrossRef]
12. Kim, S.J.; Sarkar, B. Supply Chain Model with Stochastic Lead Time, Trade-Credit Financing, and Transportation Discounts. *Math. Probl. Eng.* **2017**, *2017*, 6465912–6465925. [CrossRef]
13. Sarkar, B.; Ahmed, W.; Kim, N. Joint effects of variable carbon emission cost and multi-delay-in-payments under single-setup-multiple-delivery policy in a global sustainable supply chain. *J. Clean. Prod.* **2018**, *185*, 421–445. [CrossRef]
14. Ahmed, W.; Sarkar, B. Management of next-generation energy using a triple bottom line approach under a supply chain framework. *Resour. Conserv. Recycl.* **2019**, *150*, 104431–104450. [CrossRef]
15. Tiwari, S.; Ahmed, W.; Sarkar, B. Sustainable ordering policies for non-instantaneous deteriorating items under carbon emission and multi-trade-credit-policies. *J. Clean. Prod.* **2019**, *240*, 118183–118201. [CrossRef]
16. Ouyang, L.Y.; Wu, K.S.; Yang, C.T. A study on an inventory model for non-instantaneous deteriorating items with permissible delay in payments. *Comput. Ind. Eng.* **2006**, *51*, 637–651. [CrossRef]
17. Jaggi, C.K.; Tiwari, S.; Shafi, A. Effect of deterioration on two-warehouse inventory model with imperfect quality. *Comput. Ind. Eng.* **2015**, *88*, 378–385. [CrossRef]
18. Biswas, A.K.; Islam, S. A fuzzy EPQ model for non-instantaneous deteriorating items where production depends on demand which is proportional to population, selling price as well as advertisement. *Indep. J. Manag. Prod.* **2019**, *10*, 1679–1703. [CrossRef]
19. Aastha; Pareek, S.; Mittal, M. Non instantaneous deteriorating inventory model under credit financing when demand depends on promotion and selling price. In Proceedings of the 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, India, 4–5 June 2020.
20. Nath, B.K.; Sen, N. A completely backlogged two-warehouse inventory model for non-instantaneous deteriorating items with time and selling price dependent demand. *Int. J. Appl. Comput. Math.* **2021**, *7*, 145–166. [CrossRef]
21. Goyal, S.K. Economic order quantity under conditions of permissible delay in payments. *J. Oper. Res. Soc.* **1985**, *36*, 335–338. [CrossRef]
22. Liao, J.J. An EOQ model with non-instantaneous receipt and exponentially deteriorating items under two-level trade credit. *Int. J. Prod. Econ.* **2008**, *113*, 852–861. [CrossRef]
23. Min, J.; Zhou, Y.W.; Zhao, J. An inventory model for deteriorating items under stock-dependent demand and two-level trade credit. *Appl. Math. Model.* **2010**, *34*, 3273–3285. [CrossRef]
24. Soni, H.N. Optimal replenishment policies for non-instantaneous deteriorating items with price and stock sensitive demand under permissible delay in payment. *Int. J. Prod. Econ.* **2013**, *146*, 259–268. [CrossRef]
25. Singh, S.; Sharma, S. Optimal trade-credit policy for perishable items deeming imperfect production and stock dependent demand. *Int. J. Ind. Eng. Comput.* **2014**, *5*, 151–168. [CrossRef]
26. Taleizadeh, A.A.; Pentico, D.W.; Jabalameli, M.S.; Aryanezhad, M. An EOQ model with partial delayed payment and partial backordering. *Omega* **2013**, *41*, 354–368. [CrossRef]
27. Glock, C.H.; Ries, J.M.; Schwindl, K. A note on: Optimal ordering policy for stock-dependent demand under progressive payment scheme. *Eur. J. Oper. Res.* **2014**, *232*, 423–426. [CrossRef]
28. Sarkar, B.; Saren, S.; Cárdenas-Barrón, L.E. An inventory model with trade-credit policy and variable deterioration for fixed lifetime products. *Ann. Oper. Res.* **2015**, *229*, 677–702. [CrossRef]
29. Yang, C.T.; Dye, C.Y.; Ding, J.F. Optimal dynamic trade credit and preservation technology allocation for a deteriorating inventory model. *Comput. Ind. Eng.* **2015**, *87*, 356–369. [CrossRef]
30. Tyagi, A. An inventory model with a new credit drift: Flexible trade credit policy. *Int. J. Ind. Eng. Comput.* **2016**, *7*, 67–82. [CrossRef]
31. Mittal, M.; Jayaswal, M.K.; Kumar, V. Effect of learning on the optimal ordering policy of inventory model for deteriorating items with shortages and trade-credit financing. *Int. J. Syst. Assur. Eng. Manag.* **2021**, *13*, 914–924. [CrossRef]

32. Aastha; Pareek, S.; Dhaka, V. Credit Financing in a Two-Warehouse Inventory Model with Fuzzy Deterioration and Weibull Demand. In *Soft Computing in Inventory Management*, 1st ed.; Shah, N.H., Mittal, M., Eds.; Springer: Singapore, 2021; Volume 2, pp. 83–109.
33. Rosenblatt, M.J.; Lee, H.L. Economic production cycles with imperfect production processes. *IIE Trans.* **1986**, *18*, 48–55. [CrossRef]
34. Cheng, T.C.E. An economic order quantity model with demand-dependent unit production cost and imperfect production processes. *IIE Trans.* **1991**, *23*, 23–28. [CrossRef]
35. Salameh, M.K.; Jaber, M.Y. Economic production quantity model for items with imperfect quality. *Int. J. Prod. Econ.* **2000**, *64*, 59–64. [CrossRef]
36. Wee, H.M.; Yu, J.; Chen, M.C. Optimal inventory model for items with imperfect quality and shortage backordering. *Omega* **2007**, *35*, 7–11. [CrossRef]
37. Jaber, M.Y.; Goyal, S.K.; Imran, M. Economic production quantity model for items with imperfect quality subject to learning effects. *Int. J. Prod. Econ.* **2008**, *115*, 143–150. [CrossRef]
38. Jaggi, C.; Sharma, A.; Tiwari, S. Credit financing in economic ordering policies for non-instantaneous deteriorating items with price dependent demand under permissible delay in payments: A new approach. *Int. J. Ind. Eng. Comput.* **2015**, *6*, 481–502. [CrossRef]
39. Jaggi, C.K.; Goel, S.K.; Mittal, M. Economic order quantity model for deteriorating items with imperfect quality and permissible delay on payment. *Int. J. Ind. Eng. Comput.* **2011**, *2*, 237–248. [CrossRef]
40. Uthayakumar, R.; Sekar, T.A. A multiple production setups inventory model for imperfect items considering salvage value and reducing environmental pollution. *Oper. Res. Appl. Int. J.* **2017**, *4*, 4101. [CrossRef]
41. Aastha; Pareek, S.; Cárdenas-Barrón, L.E.; Mittal, M. Impact of imperfect quality items on inventory management for two warehouses with shortages. *Int. J. Math. Eng. Manag. Sci.* **2020**, *5*, 869–885.
42. Panwar, A.; Pareek, S.; Dhaka, V.; Mittal, M. An Imperfect Inventory Model for Multi-Warehouses Under Different Discount Policies with Shortages. *Int. J. Decis. Support Syst. Technol. (IJDSST)* **2022**, *14*, 302646. [CrossRef]

mathematics

MDPI

Article

Evolutionary Game Analysis of Electric Vehicle Distribution Entities with Shared Charging Facilities

Guangcan Xu [1], Jieyu Chen [1], Dennis Z. Yu [2,*] and Yong Liu [1]

[1] Chongqing Key Laboratory of Green Logistics Intelligent Technology, Chongqing Jiaotong University, Chongqing 400074, China; xgc@cqjtu.edu.cn (G.X.); chenjy@mails.cqjtu.edu.cn (J.C.); liuyongcs@cqjtu.edu.cn (Y.L.)

[2] The David D. Reh School of Business, Clarkson University, Potsdam, NY 13699, USA

* Correspondence: dyu@clarkson.edu; Tel.: +1-(315)-268-6435

Abstract: This study investigates the evolutionary game dynamics among electric vehicle distribution entities in the context of shared charging facilities, addressing the critical issue of inadequate charging resources. To understand the behavior of different stakeholders under government incentive policies, we develop an evolutionary game model involving a government department and two logistics enterprises (A and B). Through stability analysis, we explore equilibrium conditions of evolutionarily stable strategies (ESSs) for the tripartite evolutionary game. To ensure the robustness of our findings, we conduct a MATLAB simulation analysis to validate the analytical results. Our findings highlight that government subsidies, the costs incurred by logistics enterprises to share charging facilities, and the additional distribution income derived from this sharing are critical in determining whether the evolutionary game can achieve a stable equilibrium state. This research enables logistics companies to optimize the use of charging resources, lower operating costs, and enhance delivery efficiency. Additionally, government subsidy policies play a crucial role in encouraging logistics enterprises to engage in charging facility sharing, thereby fostering the sustainable development of the entire logistics industry. Based on these insights, the paper offers practical recommendations to further promote the sharing of charging facilities in electric vehicle distribution.

Keywords: electric vehicle distribution; charging facility sharing; tripartite evolutionary game; government incentive

MSC: 91A22; 91A80; 65C99

1. Introduction

In recent years, the rapid growth of urban distribution and rising consumer demand have led to a significant increase in the number of delivery vehicles, contributing to escalating environmental pollution in cities [1]. Among China's "dual carbon" goals, which include carbon peaking and carbon neutrality, the transportation sector is actively advancing energy conservation and emission reduction efforts. Promoting and adopting new energy vehicles has become a critical strategic initiative to drive sustainable development of the Chinese green economy [2]. The low energy consumption and low pollution characteristics of electric vehicles offer a solution to the traffic-related pollution caused by the rise in urban logistics, helping to address environmental concerns while promoting sustainable development. Many logistics companies have implemented carbon reduction plans to support "dual carbon" objectives and are increasingly integrating electric vehicles into their operations.

By 2023, more than 320,000 new energy logistics vehicles have been added nationwide [3]. As the adoption of new energy vehicles, particularly electric vehicles, continues to accelerate, the logistics distribution sector may face challenges related to insufficient

or unevenly distributed charging facilities, which can disrupt distribution operations [4]. Currently, most charging stations are unable to meet the demands of new energy logistics vehicles due to issues such as inadequate transformer capacity, site limitations, and other constraints. Additionally, the limited availability of public fast chargers leads to extended charging times, negatively impacting the operational efficiency of logistics companies [5]. Moreover, the acquisition cost of new energy logistics vehicles remains relatively high, and the charging infrastructure is still underdeveloped [6]. There is a wide variety of vehicle models with differing design parameters, making existing public charging facilities often incompatible, which further exacerbates the shortage of effective charging options for logistics fleets [7]. This lack of charging resources has significantly hindered the operational efficiency and cost control of new energy logistics vehicles, becoming a major obstacle to their broader adoption and implementation. Since each logistics enterprise has unique infrastructure requirements, such as sites and equipment, optimizing vehicle allocation and the sharing of charging facilities can significantly improve overall efficiency. For effective electric vehicle distribution, it is crucial to identify the optimal number and locations of charging stations and ensure reliable operation when power is available.

Sharing charging facilities among different logistics enterprises offers dual benefits. It reduces the time spent searching for charging stations during distribution, thereby speeding up operations, and it activates idle charging resources, maximizing efficiency. This study examines the evolutionary process of charging facility sharing in electric vehicle logistics distribution with government involvement. We develop a tripartite evolutionary game model involving the government and two logistics enterprises to analyze the conditions of their ESS and evolutionary trajectories. The study explores the impact of various factors on the critical question of whether logistics enterprises should share charging facilities and provides a rational basis for logistics enterprises to make informed decisions on facility sharing.

2. Literature Review

The literature relevant to this study can be categorized into two main areas: research on charging facilities and studies on government subsidy mechanisms within the context of evolutionary game theory.

2.1. Research on Charging Facilities

In their study of electric vehicle charging facilities and sharing strategies, Chen et al. [8] proposed a two-level mathematical model to optimize the location and capacity of charging stations, deriving an optimal design that balances route selection with charging wait times. Luo et al. [9] addressed the charging challenges of free-floating shared electric vehicles by proposing a collaborative charging planning method that leverages the complementary advantages of fixed stations and mobile charging vehicles to enhance efficiency and reduce costs. Chen et al. [10] employed a multistage stochastic integer programming model to assist governments in developing public vehicle charging networks, optimizing the location and capacity of facilities within a constrained planning scope. Luo et al. [11] developed an electric vehicle charging station location model to improve resource utilization in sustainable cities. Lv et al. [12] proposed a planning scheme for electric vehicle charging facilities based on an enhanced gray wolf optimization algorithm under V2G mode to address the impact of charging loads on the secure and stable operation of the distribution network and to improve the layout of charging infrastructure. Woo et al. [13] explored the planning of electric vehicle charging facilities to provide high-quality charging services.

In the study of private charging pile sharing, Wu et al. [14] developed an asymmetric evolutionary game model based on the investment costs and benefits of charging service operators and private charging pile owners to analyze their sharing behavior. Huang et al. [15] investigated the issues that arise from the configuration of the charging facility and provided an in-depth analysis of the sharing of the charging pile. To address capacity shortages resulting from the mismatch between charging stations and electric vehi-

cle charging loads, Chen et al. [16] proposed a layered scheduling model that incorporates shared charging piles. Ji et al. [17] suggested strategies for sharing charging facilities to enhance the utilization of electric bus charging stations. Wang et al. [18] applied an improved Shapley value model to effectively distribute benefits in private charging pile-sharing projects, easing the public charging infrastructure shortage. Yang et al. [19] employed a data-driven microsimulation method to explore the potential of private household charging pile sharing in Beijing, finding that it increases charging opportunities and reduces dependence on public stations. Hu et al. [20] established a game theory model that allows consumers to choose between private, public, and shared options, demonstrating through numerical analysis that the sharing model offers a win–win solution for all stakeholders.

2.2. Government Subsidy Mechanism in Evolutionary Game

Evolutionary game theory analyzes the strategic behavior evolution of game participants using replicator dynamics, a differential equation widely applied across various fields. The replicator dynamic mechanism is extensively used in energy conservation, environmental protection, strategy optimization, and cooperative innovation research [21]. Existing studies have demonstrated substantial and practical research on evolutionary games, including government participation as a critical factor.

Zhao et al. [22] examined the influence of government subsidies on adopting new energy vehicles, developing a three-stage evolutionary game model to analyze this effect. Zhang et al. [23] explored a tripartite evolutionary game model involving the government, manufacturers, and integrators from a risk society perspective, finding that increasing government rewards and reducing costs can boost integrators' motivation within a flexible logistics service supply chain. Luo et al. [24] constructed a tripartite evolutionary game model incorporating the government, logistics enterprises, and environmental NGOs, highlighting the critical role of government in encouraging logistics enterprises' participation. Dong et al. [25] developed a tripartite evolutionary game model focusing on logistics greening, including government, logistics enterprises, and users. Their research analyzed how government subsidies and refinement measures influence the strategic choices of enterprises and users. Li et al. [26] investigated a tripartite game model of subsidy and penalty mechanisms involving government, enterprises, and consumers, revealing the impact of these mechanisms on participant behavior. Zhang et al. [27] analyzed the implementation of subsidy and penalty policies for sustainable transportation development, while Wang et al. [28] explored the pivotal role of government policy intervention in promoting the diffusion of low-carbon technologies through supply networks.

Yu et al. [29] constructed a tripartite evolutionary game model to assess the impact of differential carbon tax policies and subsidy strategies on manufacturers. Yuan et al. [30] developed an evolutionary game model involving government, enterprises, and consumers, showing how varying levels of government regulation, innovation subsidies, and carbon tax rates influence enterprise and consumer behavior. In a supply chain market context, Guo et al. [31] examined the effects of government subsidies and green technology investments on manufacturers, governments, and consumers. Liu et al. [32] used evolutionary game analysis to evaluate the impact of government subsidies on green suppliers and manufacturers. Wang et al. [33] demonstrated through evolutionary game theory that government subsidy strategies significantly influence participant selection, with subsidies proving to be more effective than non-subsidy approaches. Finally, Sun et al. [34] studied the crucial role of government subsidies in promoting the development and implementation of recyclable green logistics packaging by logistics enterprises.

In summary, a review of the relevant literature reveals that the widespread adoption and promotion of electric vehicles across various industries have been significantly driven by robust national policy support and continuous advancements in research and development. The government has implemented numerous incentives, including tax breaks, subsidies, and infrastructure development, which have substantially lowered the costs and risks for logistics enterprises adopting electric vehicles for distribution. Simultane-

ously, ongoing technological progress, particularly in battery technology and charging efficiency, has greatly enhanced electric vehicles' practicality and economic viability in logistics distribution.

Evolutionary game theory, a powerful tool for analyzing the selection and evolution of strategy agents in complex systems, has provided significant insights in various fields. Numerous scholars have applied evolutionary game models to study interactions between governments, enterprises, and other stakeholders, investigating optimal strategies under different reward and punishment mechanisms, policy interventions, and market conditions. However, research on applying evolutionary game theory to the sharing of charging facilities in the logistics distribution of electric vehicles remains limited. Most existing studies focus on electric vehicle technology, policy incentives, and the construction and management of charging facilities, with comparatively little attention given to how logistics enterprises can enhance distribution efficiency and economic benefits by sharing charging infrastructure.

To address this research gap and offer practical strategic recommendations for enterprises, this paper introduces an evolutionary game model to conduct a comprehensive analysis of the strategic choices involved in sharing charging facilities among logistics companies. Specifically, the study constructs an evolutionary game model involving the government and logistics enterprises A and B to analyze the influence of various factors on the choice of a charging facility-sharing strategy. Additionally, the paper conducts a simulation study to validate the model's effectiveness and proposes practical policy recommendations and management measures. This study not only deepens the understanding of charging facility sharing in electric vehicle distribution entities but also provides a theoretical foundation and practical guidance for governments to formulate more effective policies and for enterprises to optimize their operational strategies, ultimately promoting the sustainable development and widespread adoption of electric vehicle distribution models.

Through the research of this paper, we can not only deepen the understanding and understanding of the sharing of charging facilities in the distribution of electric vehicles in logistics enterprises, but also provide a solid theoretical basis and practical guidance for the government to formulate more scientific and effective policies and provide a solid theoretical basis and practical guidance for logistics enterprises to optimize their operation strategies. Finally, we hope that through this study, we can promote the sustainable development and wide application of an electric vehicle distribution mode and contribute to the construction of a green, low-carbon, and efficient logistics distribution system.

3. An Evolutionary Game Theoretical Model

3.1. Problem Description

In the context of sharing charging facilities for electric vehicle distribution, the key participants include a government department and logistics enterprises A and B. These logistics enterprises, which rely on electric vehicles for distribution operations, may collaborate to form a charging facility-sharing alliance, thereby pooling their resources to share charging infrastructure. The government department has the option to actively promote this strategy by providing subsidies and incentives to the participating logistics enterprises. The relationship between the three participants is illustrated in Figure 1.

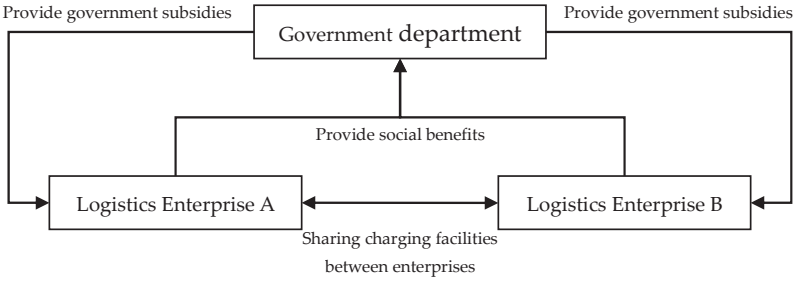


Figure 1. Relationship among the three participants.

3.2. Model Assumptions and Notations

Based on the actual benefits of the participants and the principles of evolutionary game theory, the following assumptions are proposed (with the understanding that all parameters are positive values), the related parameters are defined in Table 1.

Table 1. Notations of the tripartite evolutionary game.

Notations	Definition
$I_{a(i=a,b)}$	Distribution revenue when party *i* does not share
$T_{a(i=a,b)}$	The costs incurred by logistics enterprise *i* when charging facility sharing is not established among logistics enterprises (including the expenses for locating charging facilities, the operation and maintenance of those facilities, and the electricity costs for charging)
$P_{a(i=a,b)}$	The risk cost that a logistics enterprise *i* needs to pay when choosing a sharing strategy
C	Charging facility operation and maintenance costs incurred when sharing strategies are reached between logistics enterprises
Q	The power purchase cost of charging facilities must be paid when sharing strategies are reached between logistics enterprises
$r(0 \leq r \leq 1)$	Cost allocation coefficient of logistics enterprise A when choosing a sharing strategy
$1 - r$	Cost allocation coefficient of logistics enterprise B when choosing a sharing strategy
$S_{a(i=a,b)}$	The additional distribution income obtained by the logistics enterprise *i* after sharing charging facilities among logistics enterprises
G	The social benefits obtained when the government chooses to actively promote the strategy
g	The social benefits are obtained when the government chooses a negative promotion strategy. $G > g$ by default
D	The cost of government expenditures to implement surveillance measures
e	Enterprises share the environmental benefits they bring to government
L	When the government adopts an active promotion strategy, it must pay sharing incentive subsidies to both logistics enterprises
$\mu(0 \leq \mu \leq 1)$	Subsidy allocation coefficient of logistics enterprise A
$1 - \mu$	Subsidy allocation coefficient of logistics enterprise B

Assumption 1. *Game Participants and Their Strategy Choices: The model involves a tripartite game among logistics enterprises A, B, and the government. The strategy space for the logistics enterprises consists of two options: {sharing, not sharing}. They form a charging facility-sharing alliance to pool charging resources if they choose to share. If they decide not to share, each enterprise maintains its charging facilities independently. The strategy space for the government is {active promotion, negative promotion}, indicating whether or not to implement subsidy policies to encourage enterprises to form sharing alliances.*

Assumption 2. *In the decision-making model for charging facility sharing, all three game participants are characterized by bounded rationality and aim to maximize their individual benefits. They also possess learning abilities, making the game dynamic over time. Both logistics enterprises and the government will continuously adjust their decisions based on the strategies chosen by all parties and learn throughout the game process to arrive at their optimal equilibrium strategy.*

Assumption 3. *Assume that the probability of the government choosing the "positive promotion" strategy is x (where $0 \leq x \leq 1$), making the probability of choosing the "negative promotion" strategy $1 - x$. Regardless of whether the government opts to encourage enterprises to share charging facilities actively or takes a passive approach to promote the sharing model, it incurs a cost D for implementing supervisory measures. Sharing among logistics enterprises yields an environmental benefit e for the government. Suppose that the government chooses to actively promote the sharing model of charging facilities among logistics enterprises. In that case, it will gain social benefits G but must also provide subsidies and incentives to these enterprises. Conversely, if the government chooses to negatively promote the sharing model, it will gain social benefits g without incurring expenses, such as subsidy measures.*

Assumption 4. *Assume that the probability of logistics enterprise A choosing the sharing strategy is y (where* $0 \leq y \leq 1$*), meaning the probability of choosing the non-sharing strategy is* $1 - y$*. If logistics enterprise A opts to share charging facilities and joins the alliance, it can receive subsidies from the government, provided that the government has adopted the active guidance strategy. If neither party chooses the sharing strategy, there will be no charging facility-sharing alliance between logistics enterprises, and the final income for logistics enterprise A will be* $I_a - T_a + \mu L$*. If both parties choose to share, logistics enterprise A will gain additional distribution revenue* S_a *but must bear the cost* $r(C + Q)$ *associated with the sharing strategy. Conversely, if logistics enterprise A decides not to share charging facilities, it will not receive the additional distribution revenue or government subsidies available through the sharing model, and its income will remain at* $I_a - T_a$*.*

Assumption 5. *Assume that the probability of logistics enterprise B choosing the sharing strategy is z (where* $0 \leq z \leq 1$*), meaning the probability of choosing the non-sharing strategy is* $1 - z$*. Similar to logistics enterprise A, if logistics enterprise B opts to share charging facilities and join the alliance, it can receive subsidies from the government, provided the government has chosen the active guidance strategy, regardless of whether the other parties also decide to share. If no other parties choose the sharing strategy, a charging facility-sharing alliance will not be formed, and the final income for logistics enterprise B will be* $I_b - T_b + (1 - \mu L)$*. If both parties decide to share, logistics enterprise B will also gain additional benefits* S_b*. However, if logistics enterprise B decides not to share charging facilities, it will not receive the additional distribution revenue or government subsidies associated with the sharing model, and its income will remain at* $I_b - T_b$*.*

Additionally, the necessary notation of our model is introduced in Table 1.

3.3. Modeling Framework

Based on assumptions and notations, the payoff matrix of the government, logistics enterprise A, and logistics enterprise B can be obtained, as shown in Table 2.

Table 2. Payoff matrix for the government, logistics enterprises A and B.

Government Strategy Selection	Logistics Enterprise A Strategy Choice	Logistics Enterprise B Strategy Choice	
		Share z	**Not Shared** $1-z$
Active promotion x	Sharing y	$G + e - D - L$ $I_a + S_a - r(C + Q) - P_a + \mu L$ $I_b + S_b - (1 - r)\ (C + Q) -$ $P_b + (1 - \mu)L$	$G - D - L$ $I_a - T_a - P_a + L$ $I_b - T_b$
	Not sharing $1 - y$	$G - D - L$ $I_a - T_a$ $I_b - T_b - P_b + L$	$G - D$ $I_a - T_a$ $I_b - T_b$
Negative promotion $1 - x$	Sharing y	$g - D + e$ $I_a + S_a - r(C + Q) - P_a$ $I_b + S_b - (1 - r)\ (C + Q) - P_b$	$g - D$ $I_a - T_a - P_a$ $I_b - T_b$
	Not sharing $1 - y$	$g - D$ $I_a - T_a$ $I_b - T_b - P_b$	$g - D$ $I_a - T_a$ $I_b - T_b$

As shown in Table 2, when the government adopts a strategy of actively promoting charging facilities, and both logistics enterprises A and B choose to share these facilities, the government's social benefit is denoted by G. The costs incurred for implementation and supervision are represented by D, the environmental benefit by e, and the sharing incentive subsidy provided to enterprises by L. Thus, the government's net benefit is calculated as $G + e - D - L$. For logistics enterprise A, its daily distribution income is represented by I_a. Since both enterprises opt for sharing, logistics enterprise A avoids paying the non-sharing cost T_a but incurs a risk cost P_a, operational and maintenance

costs rC, and power purchase costs rQ. Additionally, the enterprise receives an additional distribution income of S_a and a government subsidy of μL. Therefore, the net income of logistics enterprise A is $I_a + S_a - r(C+Q) - P_a + \mu L$. Similarly, the daily distribution income for logistics enterprise B is I_b. Like enterprise A, it avoids the non-sharing cost T_b and instead incurs risk costs P_b, operational and maintenance costs $(1-r)C$, and power purchase costs $(1-r)Q$. Logistics enterprise B also gains an additional distribution income of S_b and a government subsidy of $(1-\mu)L$. Hence, the net income of logistics enterprise B is $I_b + S_b - (1-r)\ (C+Q) - P_b + (1-\mu)L$. In a similar manner, the benefits for the three participants under different strategies can be derived accordingly.

Thus, the expected returns, expected average returns, and replicator dynamics equation for participants can be determined from the game payoff matrix. Let U_{c1} represent the expected return for the government when it chooses the active promotion strategy. This value is the weighted average sum of the government's revenue when logistics enterprises A and B choose among the four possible combinations of sharing and not sharing strategies. U_{c1} is given by the following equation:

$$\begin{aligned} U_{c1} = yz(G+e-D-L) + (1-y)z(G-D-L) + y(1-z)\ (G-D-L) \\ +(1-y)\ (1-z)\ (G-D) \end{aligned} \tag{1}$$

The expected revenue when the government opts for the negative promotion strategy is denoted as U_{c2}. This represents the weighted average sum of the government's revenue when logistics enterprises A and B select among the four possible combinations of sharing and not sharing strategies. The calculation for this scenario is provided below.

$$\begin{aligned} U_{c2} = yz(g-D+e) + (1-y)z(g-D) + y(1-z)\ (g-D) \\ +(1-y)\ (1-z)\ (g-D) = yze + g - D \end{aligned} \tag{2}$$

The government's average expected revenue, U_c, is the weighted sum of its expected revenues when choosing either the positive or negative promotion strategy, as shown below.

$$U_c = xU_{c1} + (1-x)U_{c2} = g - D + Gx - gx - Lxy - Lxz + eyz \tag{3}$$

According to evolutionary game theory, the government's replicator dynamic can be obtained by the following differential equation:

$$\begin{aligned} F(x) = \tfrac{dx}{dt} = x(U_{c1} - U_c) = x(1-x)\ (U_{c1} - U_{c2}) \\ = x(1-x)(G - g - Ly - Lz + Lyz) \end{aligned} \tag{4}$$

The expected revenue for logistics enterprise A when it chooses the sharing strategy, denoted as U_{a1}, is the weighted average sum of revenues considering the various strategy combinations selected by the government and logistics enterprise B. This value is provided below.

$$\begin{aligned} U_{a1} = xz(I_a + S_a - r(C+Q) - P_a + \mu L) + x(1-z)(I_a - T_a - P_a + L) \\ +(1-x)z(I_a + S_a - r(C+Q) - P_a) + (1-x)\ (1-z) \\ (I_a - T_a - P_a) \end{aligned} \tag{5}$$

The expected revenue for logistics enterprise A when it opts not to share, denoted as U_{a2}, is the weighted average sum of revenues based on the strategy combinations chosen by the government and logistics enterprise B. This value is presented in the following equation:

$$\begin{aligned} U_{a2} = xz(I_a - T_a) + x(1-z)\,(I_a - T_a) + (1-x)z(I_a - T_a) + (1-x) \\ (1-z)\,(I_a - T_a) \end{aligned} \tag{6}$$

The logistics enterprise A's average expected return, U_a, is the weighted average of its expected returns when it chooses either the sharing or non-sharing strategy, as provided below.

$$\begin{aligned} U_a &= yU_{a1} + (1-y)U_{a2} \\ &= I_a - T_a - P_a y + S_a yz + T_a yz - rCyz - rQyz + Lxy - Lxyz \\ &+\mu Lxyz \end{aligned} \tag{7}$$

Therefore, the logistics enterprise A's replicator dynamic is given by the following differential equation:

$$\begin{aligned} F(y) = \frac{dy}{dt} &= y(U_{a1} - U_a) = y(1-y)\ (U_{a1} - U_{a2}) \\ &= y(1-y)(S_a z - P_a + T_a z - rCz - rQz + Lx - Lxz + \mu Lxz) \end{aligned} \tag{8}$$

The expected revenue for logistics enterprise B when it chooses the sharing strategy is denoted as U_{b1}. This represents the weighted average of revenues based on the combination of strategies chosen by the government and logistics enterprise A. U_{b1} can be given by

$$\begin{aligned} U_{b1} = xy(I_b + S_b \quad & -(1-r)(C+Q) - P_b + (1-\mu)L) \\ & +x(1-y)\ (I_b - T_b - P_b + L) \\ & +(1-x)y(I_b + S_b - (1-r)(C+Q) - P_b) \\ & +(1-x)\ (1-y)\ (I_b - T_b - P_b) \end{aligned} \tag{9}$$

The expected revenue for logistics enterprise B, when it opts for the non-sharing strategy, is denoted as U_{b2}. This value represents the weighted average of revenues based on the combination of strategies chosen by the government and the logistics enterprise A, as outlined below.

$$\begin{aligned} U_{b2} = xy(I_b - T_b) + x(1-y)(I_b - T_b) + (1-x)y(I_b - T_b) \\ +(1-x)\ (1-y)(I_b - T_b) \end{aligned} \tag{10}$$

The average expected return for logistics enterprise B is the weighted sum of its expected returns when choosing between the sharing and non-sharing strategies, as detailed below.

$$\begin{aligned} U_b = zU_{b1} + (1- \quad & z)U_{b2} \\ &= I_b - T_b - P_b z - Cyz - Qyz + S_b yz + T_b yz + rCyz + rQyz \\ &+Lxz - \mu Lxyz \end{aligned} \tag{11}$$

Thus, the following differential equation can obtain the logistics enterprise B's replicator dynamic:

$$\begin{aligned} F(z) = \frac{dz}{dt} = z(U_{b1} \quad & - U_b) = z(1-z)\ (U_{b1} - U_{b2}) \\ &= z(1-z)(S_b y + T_b y - P_b - Cy - Qy + rCy + rQy + Lx \\ &-\mu Lxy) \end{aligned} \tag{12}$$

4. Evolutionary Game Analysis

4.1. Solving ESS of the Tripartite Evolutionary Game

The evolutionary stable strategy (ESS) of the tripartite game involving the government and two logistics enterprises can be solved and analyzed using the above replicator dynamic equations, i.e., Equations (4), (8), and (12). Following Selten's [35] conclusion, in a tripartite evolutionary game, it is sufficient to focus on the analysis of pure strategies. By solving $F(x) = 0$, $F(y) = 0$, and $F(z) = 0$, eight local pure strategy equilibria can be given by $E_1(0,0,0)$, $E_2(1,0,0)$, $E_3(0,1,0)$, $E_4(0,0,1)$, $E_4(0,0,1)$, $E_5(1,1,0)$, $E_6(1,0,1)$, $E_7(0,1,1)$, and $E_8(1,1,1)$.

According to Friedman's [36] method for determining equilibrium stability, the evolutionarily stable strategy of the system can be derived through local stability analysis of the system's Jacobian matrix. By taking the partial derivatives of the replicator dynamic equations $F(x)$, $F(y)$, and $F(z)$ with respect to x, y, and z, the Jacobian matrix J of the model can be obtained as follows:

$$J = \begin{bmatrix} \frac{\partial F(x)}{\partial x} & \frac{\partial F(x)}{\partial y} & \frac{\partial F(x)}{\partial z} \\ \frac{\partial F(y)}{\partial x} & \frac{\partial F(y)}{\partial y} & \frac{\partial F(y)}{\partial z} \\ \frac{\partial F(z)}{\partial x} & \frac{\partial F(z)}{\partial y} & \frac{\partial F(z)}{\partial z} \end{bmatrix} \tag{13}$$

where

$$\frac{\partial F(x)}{\partial x} = (2x-1)\,(g - G + Ly + Lz - Lyz) \tag{14}$$

$$\frac{\partial F(x)}{\partial y} = Lx(x-1)(1-z) \tag{15}$$

$$\frac{\partial F(x)}{\partial z} = Lx(x-1)(1-y) \tag{16}$$

$$\frac{\partial F(y)}{\partial x} = Ly(1-y)(\mu z - z + 1) \tag{17}$$

$$\frac{\partial F(y)}{\partial y} = (2y-1)(P_a - Lx - S_a z - T_a z + rCz + rQz - \mu Lxz) \tag{18}$$

$$\frac{\partial F(y)}{\partial z} = y(1-y)(S_a + T_a - rC - rQ - Lx + \mu Lx) \tag{19}$$

$$\frac{\partial F(z)}{\partial x} = Lz(1-z)(1-\mu y) \tag{20}$$

$$\frac{\partial F(z)}{\partial y} = z(1-z)(S_b - Q - C + T_b + rC + rQ - \mu Lx) \tag{21}$$

$$\frac{\partial F(z)}{\partial z} = (2z-1)(P_b - Lx - S_b y - T_b y + Cy + Qy - rCy - rQy + \mu Lxy) \tag{22}$$

4.2. Equilibrium Stability Analysis

According to evolutionary game theory, an evolutionary stable strategy (ESS) exists only if the eigenvalues of each equilibrium are all negative. Therefore, substituting the eight equilibria into the Jacobian matrix (13), the corresponding eigenvalues are calculated, as shown in Table 3. The sign of the eigenvalue λ is then analyzed to determine the stability of each equilibrium. We follow the tabular presentation format used by Zhang et al. [27] in their evolutionary game study to present the equilibrium stability analysis results, as shown in Tables 3–8.

Table 3. Eigenvalues of each equilibrium.

Equilibrium	Eigenvalue λ_1	Eigenvalue λ_2	Eigenvalue λ_3	Stability
$E_1(0,0,0)$	$G-g$	$-P_a$	$-P_b$	$(+,-,-)$
$E_2(1,0,0)$	$g-G$	$L-P_a$	$L-P_b$	$(-,*,*)$
$E_3(0,1,0)$	$G-L-g$	P_a	$S_b-P_b-Q-C+T_b+rC+rQ$	$(*,+,*)$
$E_4(0,0,1)$	$G-L-g$	$S_a-P_a+T_a-rC-rQ$	P_b	$(*,*,+)$
$E_5(1,1,0)$	$L-G+g$	P_a-L	$L-C-P_b-Q+S_b+T_b+rC+rQ-\mu L$	$(*,*,*)$
$E_6(1,0,1)$	$L-G+g$	$\mu L-P_a+S_a+T_a-rC-rQ$	P_b-L	$(*,*,*)$
$E_7(0,1,1)$	$G-L-g$	$P_a-S_a-T_a+rC+rQ$	$P_b+C+Q-S_b-T_b-rC-rQ$	$(*,*,*)$
$E_8(1,1,1)$	$L-G+g$	$P_a-S_a-T_a+rC+rQ-\mu L$	$P_b+C+Q-L-S_b-T_b-rC-rQ+\mu L$	$(*,*,*)$

Note: The eigenvalue with "+" sign is positive; the eigenvalue with "–" is negative; the eigenvalue with "*" depends on the specific parameter values.

Table 4. Stability analysis of Scenario 1.

Equilibrium	Eigenvalue λ_1	Eigenvalue λ_2	Eigenvalue λ_3	Stability
$E_1(0,0,0)$	$+$	$-$	$-$	Unstable
$E_2(1,0,0)$	$-$	$-$	$-$	ESS
$E_3(0,1,0)$	$-,+$	$+$	$-,+$	Unstable
$E_4(0,0,1)$	$-,+$	$-,+$	$+$	Unstable
$E_5(1,1,0)$	$-,+$	$+$	$-,+$	Unstable
$E_6(1,0,1)$	$-,+$	$-,+$	$+$	Unstable
$E_7(0,1,1)$	$-,+$	$+$	$+$	Unstable
$E_8(1,1,1)$	$-,+$	$+$	$+$	Unstable

Table 5. Stability analysis of Scenario 2.

Equilibrium	Eigenvalue λ_1	Eigenvalue λ_2	Eigenvalue λ_3	Stability
$E_1(0,0,0)$	$+$	$-$	$-$	Unstable
$E_2(1,0,0)$	$-$	$+$	$-,+$	Unstable
$E_3(0,1,0)$	$-$	$+$	$-,+$	Unstable
$E_4(0,0,1)$	$-,+$	$-,+$	$+$	Unstable
$E_5(1,1,0)$	$-$	$-$	$-$	ESS
$E_6(1,0,1)$	$-,+$	$-,+$	$+$	Unstable
$E_7(0,1,1)$	$-,+$	$-,+$	$+$	Unstable
$E_8(1,1,1)$	$-,+$	$-,+$	$+$	Unstable

Table 6. Stability analysis of Scenario 3.

Equilibrium	Eigenvalue λ_1	Eigenvalue λ_2	Eigenvalue λ_3	Stability
$E_1(0,0,0)$	$+$	$-$	$-$	Unstable
$E_2(1,0,0)$	$-$	$-,+$	$+$	Unstable
$E_3(0,1,0)$	$-,+$	$+$	$-,+$	Unstable
$E_4(0,0,1)$	$+$	$-$	$+$	Unstable
$E_5(1,1,0)$	$-,+$	$+$	$-,+$	Unstable
$E_6(1,0,1)$	$-$	$-$	$-$	ESS
$E_7(0,1,1)$	$-,+$	$+$	$-,+$	Unstable
$E_8(1,1,1)$	$-,+$	$+$	$-,+$	Unstable

Table 7. Stability analysis of Scenario 4.

Equilibrium	Eigenvalue λ_1	Eigenvalue λ_2	Eigenvalue λ_3	Stability
$E_1(0,0,0)$	$+$	$-$	$-$	Unstable
$E_2(1,0,0)$	$-$	$-,+$	$-,+$	Unstable
$E_3(0,1,0)$	$-$	$+$	$+$	Unstable
$E_4(0,0,1)$	$-$	$+$	$+$	Unstable
$E_5(1,1,0)$	$-,+$	$-,+$	$+$	Unstable
$E_6(1,0,1)$	$-,+$	$+$	$-,+$	Unstable
$E_7(0,1,1)$	$-$	$-$	$-$	ESS
$E_8(1,1,1)$	$+$	$-$	$-$	Unstable

Based on the above analysis, the eigenvalues of each equilibrium are calculated according to the ESS equilibrium conditions, and the following scenarios can be obtained:

Scenario 1. *When $P_a > L$ and $P_b > L$, Table 4 shows that the eigenvalues corresponding to $E_2(1,\ 0,\ 0)$ are negative. This indicates that the model has a stable equilibrium at $E_2(1,\ 0,\ 0)$, where the corresponding stable strategy is (active promotion, not sharing, not sharing). The other equilibria are unstable in this scenario.*

In this scenario, the government will opt for a positive promotion strategy when the social benefits outweigh those from negative promotion measures. For logistics enterprises A and B, if the risk costs associated with the sharing strategy are too high and the additional distribution income and government subsidies are insufficient, both enterprises will choose not to share. At this point, the decisions of all three parties converge towards a stable equilibrium strategy.

Table 8. Stability analysis of Scenario 5.

Equilibrium	Eigenvalue λ_1	Eigenvalue λ_2	Eigenvalue λ_3	Stability
$E_1(0,0,0)$	$+$	$-$	$-$	Unstable
$E_2(1,0,0)$	$-$	$-,+$	$-,+$	Unstable
$E_3(0,1,0)$	$+$	$+$	$-,+$	Unstable
$E_4(0,0,1)$	$+$	$-,+$	$+$	Unstable
$E_5(1,1,0)$	$-$	$-,+$	$+$	Unstable
$E_6(1,0,1)$	$-$	$+$	$-,+$	Unstable
$E_7(0,1,1)$	$+$	$-,+$	$-,+$	Unstable
$E_8(1,1,1)$	$-$	$-$	$-$	ESS

Scenario 2. *When the conditions $G > L + g$, $P_a < L$ and $P_b + L + \mu L + C + Q > S_b + T_b + rC + rQ$ are satisfied, Table 5 shows that the eigenvalues corresponding to $E_5(1,\ 1,\ 0)$ are negative. This indicates that the model has a stable equilibrium at $E_5(1,\ 1,\ 0)$, where the stability strategy is (active promotion, sharing, not sharing). The other equilibria are unstable.*

In this scenario, the government will opt for a positive promotion strategy when its social benefits exceed those of the negative approach. Logistics enterprise A will choose the sharing strategy if the government's subsidy outweighs the associated risk costs. On the contrary, logistics enterprise B will refrain from sharing if the required risk costs are too high and the additional benefits from sharing are minimal. At this point, the choices made by all three parties converge towards a balanced strategy.

Scenario 3. *When $G > L + g$, $P_a + rC + rQ > \mu L + S_a + T_a$ and $P_b < L$, as shown in Table 6, the eigenvalues corresponding to $E_6(1,\ 0,\ 1)$ are negative. This indicates that the model has a stable equilibrium at $E_6(1,\ 0,\ 1)$, with the corresponding stability strategy being (active promotion, not sharing, sharing). The other equilibria are unstable.*

In this scenario, the government will opt for a positive promotion strategy if its social benefits outweigh those of the negative strategy. For logistics enterprise A, the high risk cost associated with sharing will lead to a decision not to share. Conversely, if the government subsidy outweighs the risk cost for logistics enterprise B, it will adopt the sharing strategy. At this point, the decisions of all three parties converge toward a balanced strategy.

Scenario 4. *When the conditions $G < L + g$, $P_a + rC + rQ < S_a + T_a$ and $P_b + C + Q < S_b + T_b + rC + rQ$ are satisfied, Table 7 shows that the eigenvalues corresponding to $E_7(0,\ 1,\ 1)$ are negative. This indicates that the model has a stable equilibrium at $E_7(0,\ 1,\ 1)$, where the corresponding stable strategy is (negative promotion, sharing, sharing). All other equilibria are unstable.*

In this scenario, if the social benefits of the government's positive promotion strategy are lower than those of the negative strategy, the government will opt for a negative promotion strategy. For logistics enterprises A and B, if the additional distribution income from choosing the sharing strategy is substantial, both will choose to share. At this point, the decisions of all three parties converge towards a balanced strategy.

Scenario 5. *When the conditions $G > L + g$, $P_a + rC + rQ < \mu L + S_a + T_a$ and $P_b + C + Q + \mu L < L + S_b + T_b + rC + rQ$ are met, it is evident from Table 8 that the eigenvalues corresponding*

to $E_8(1,\ 1,\ 1)$ *are negative. This indicates that the model has a stable equilibrium at* $E_8(1,\ 1,\ 1)$. *The corresponding stable strategy is (active promotion, sharing, sharing), while the other equilibria are unstable.*

In this scenario, if the social benefits derived from the positive promotion strategy exceed the combined total of the social benefits from the negative strategy and the incentive subsidies provided to enterprises, the government will opt for a positive promotion strategy. For logistics enterprises A and B, if the additional distribution income and government subsidies received from adopting the sharing strategy are sufficiently high, both enterprises will choose to participate in the sharing strategy. Consequently, the choices of all three parties will align towards a balanced strategy.

5. Simulation Study of the Tripartite Evolutionary Game

To provide a clearer understanding of the evolutionarily stable strategies (ESSs) among government and logistics enterprises under varying conditions, MATLAB 2020a is utilized to simulate the evolutionary trajectories of strategy selection for each party involved in the game.

5.1. Evolutionary Trajectories of Different Scenarios

We provide a simulation study of the evolutionary trajectories under each scenario of the analytical results obtained in Section 4.

Scenario 1. *The simulation parameters are assigned as follows.* $I_a = 50$, $I_b = 50$, $T_a = 5$, $T_b = 5$, $P_a = 20$, $P_b = 20$, $C = 5$, $Q = 5$, $r = 0.5$, $S_a = 10$, $S_b = 10$, $G = 50$, $g = 20$, $D = 8$, $e = 10$, $L = 16$, *and* $\mu = 0.5$, *while* $P_a > L$ *and* $P_b > L$ *are satisfied. The simulation results of the evolutionary trajectories are illustrated in Figure 2.*

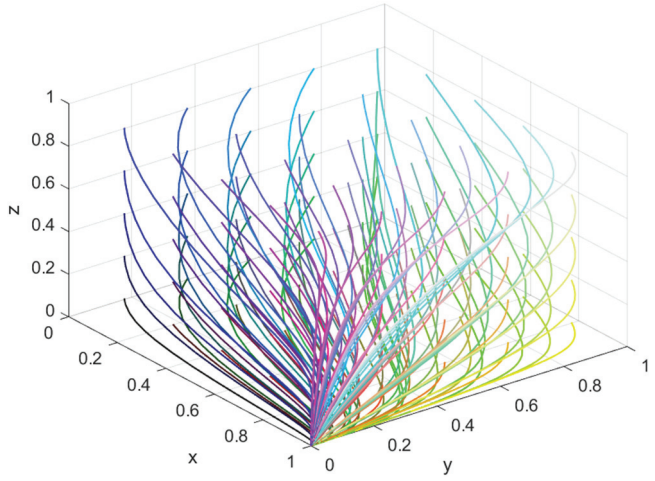


Figure 2. Evolutionary trajectories of Scenario 1.

As shown in Figure 2, the final decision probabilities for logistics enterprise A, logistics enterprise B and the government converge to $(1,\ 0,\ 0)$. This outcome indicates that when $P_a > L$ and $P_b > L$, the government ultimately opts for active promotion. At the same time, logistics enterprises A and B choose not to share charging facilities. Initially, both enterprises opt to share the facilities, but as the costs of sharing become too high and the additional income remains low, they gradually shift towards not sharing. At this point, the stable strategy in the tripartite evolutionary game is (active promotion, not sharing, not sharing), which confirms Scenario 1.

Scenario 2. *The simulation parameters are assigned as follows.* $I_a = 50$, $I_b = 50$, $T_a = 5$, $T_b = 5, P_a = 15, P_b = 20, C = 5, Q = 5, r = 0.5, S_a = 10, S_b = 10, G = 50, g = 20, D = 8, e =$

10, $L = 16$, $\mu = 0.5$, *when* $G > L + g$, $P_a < L$ *and* $P_b + L + \mu L + C + Q > S_b + T_b + rC + rQ$ *are met. The simulation results of the evolutionary trajectories are shown in Figure 3.*

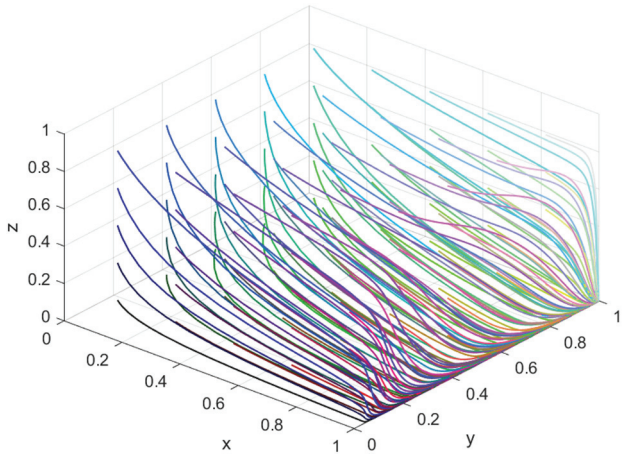


Figure 3. Evolutionary trajectories of Scenario 2.

As illustrated in Figure 3, the final decision probabilities for logistics enterprise A, logistics enterprise B, and the government converge to $(1,\ 1,\ 0)$. This indicates that when $G > L + g$, $P_a < L$ and $P_b + L + \mu L + C + Q > S_b + T_b + rC + rQ$, the government ultimately opts for active promotion, logistics enterprise A chooses to share, and logistics enterprise B opts not to share. As the game progresses, logistics enterprise A's profitability gradually increases, leading to enhanced government subsidies and reinforcing its decision to share charging facilities. In contrast, logistics enterprise B, facing lower income and minimal government subsidies, ultimately decides against sharing. At this point, the stable strategy in the tripartite evolutionary game is (active promotion, sharing, not sharing), confirming Scenario 2.

Scenario 3. *We assign the simulation parameters as follows.* $I_a = 50$, $I_b = 50$, $T_a = 5$, $T_b = 5$, $P_a = 20$, $P_b = 15$, $C = 5$, $Q = 5$, $r = 0.5$, $S_a = 10$, $S_b = 10$, $G = 50$, $g = 20$, $D = 8$, $e = 10$, $L = 16$, *and* $\mu = 0.5$, *while* $G > L_b + g$, $P_a + rC + rQ > \mu L + S_a + T_a$ *and* $P_b < L$ *are met. The simulation results of the evolutionary trajectories are illustrated in Figure 4.*

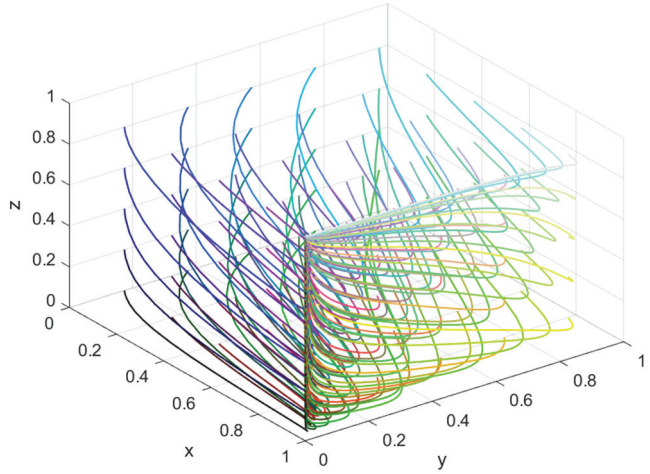


Figure 4. Evolutionary trajectories of Scenario 3.

As illustrated in Figure 4, the final decision probabilities of logistics enterprise A, logistics enterprise B, and the government converge to $(1,\ 0,\ 1)$. This occurs under conditions where $G > L_b + g$, $P_a + rC + rQ > \mu L + S_a + T_a$, and $P_b < L$. In this scenario, the

government ultimately chooses an active promotion strategy; logistics enterprise A opts not to share, while logistics enterprise B decides to share. Initially, logistics enterprises A and B may share charging facilities. However, logistics enterprise A shifts away from sharing as the game progresses due to low returns and limited government subsidies. On the contrary, as the profitability of logistics enterprise B increases and government subsidies grow, logistics enterprise B maintains its decision to share charging facilities. Thus, the stable strategy in the tripartite evolutionary game is (active promoting, not sharing, sharing), verifying Scenario 3.

Scenario 4. *The simulation parameters are set as follows.* $I_a = 50$, $I_b = 50$, $T_a = 5$, $T_b = 5$, $P_a = 8$, $P_b = 8$, $C = 5$, $Q = 5$, $r = 0.5$, $S_a = 15$, $S_b = 15$, $G = 50$, $g = 35$, $D = 8$, $e = 10$, $L = 20$, *and* $\mu = 0.5$, *while* $G < L + g$, $P_a + rC + rQ < S_a + T_a$, *and* $P_b + C + Q < S_b + T_b + rC + rQ$ *are satisfied. The simulation results of the evolutionary trajectories are shown in Figure 5.*

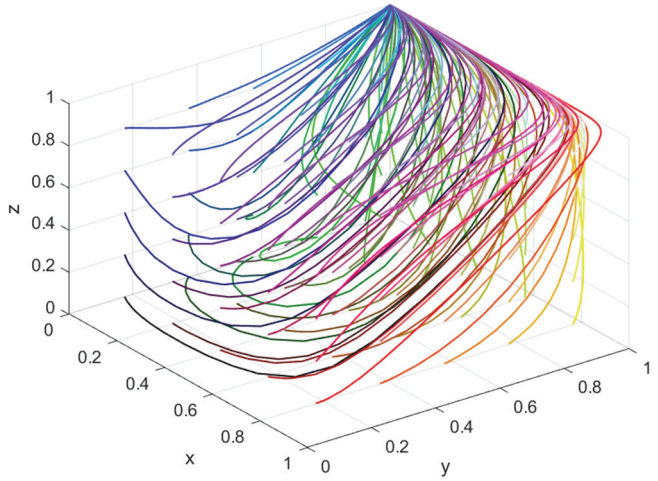


Figure 5. Evolutionary trajectories of Scenario 4.

As shown in Figure 5, the final decision probabilities of logistics enterprise A, logistics enterprise B, and the government converge to $(0, 1, 1)$. This outcome occurs when the conditions $G < L + g$, $P_a + rC + rQ < S_a + T_a$, and $P_b + C + Q < S_b + T_b + rC + rQ$ are satisfied. In this scenario, the government ultimately chooses a negative promotion strategy, while both logistics enterprises A and B opt for the sharing strategy. At this point, logistics enterprises A and B experience lower costs and higher income, whereas the government faces reduced income and increased subsidy expenditures. Consequently, as time progresses, the income of logistics enterprises A and B continues to rise, leading them to persist in sharing charging facilities. Meanwhile, the government maintains its choice of negative promotion for sharing charging facilities among logistics enterprises. In this case, the stable strategy in the tripartite evolutionary game (negative promotion, sharing, sharing) verifies Scenario 4.

Scenario 5. *We assign the following parameter values to the simulation.* $I_a = 50$, $I_b = 50$, $T_a = 5$, $T_b = 5$, $P_a = 8$, $P_b = 8$, $C = 5$, $Q = 5$, $r = 0.5$, $S_a = 15$, $S_b = 15$, $G = 50$, $g = 20$, $D = 8$, $e = 10$, $L = 20$, *and* $\mu = 0.5$, *when* $G > L + g$, $P_a + rC + rQ < \mu L + S_a + T_a$, *and* $P_b + C + Q + \mu L < L + S_b + T_b + rC + rQ$ *are met. The simulation results of the evolutionary trajectories are illustrated in Figure 6.*

As shown in Figure 6, the final decision probabilities of logistics enterprise A, logistics enterprise B, and the government converge to $(1, 1, 1)$. This outcome occurs when the conditions $G > L + g$, $P_a + rC + rQ < \mu L + S_a + T_a$, and $P_b + C + Q + \mu L < L + S_b + T_b + rC + rQ$ are satisfied. In this scenario, the government ultimately chooses to actively promote, while both logistics enterprises A and B adopt the sharing strategy. This situation

typically arises during the mature stage of shared charging facilities between logistics enterprises A and B. At this point, both enterprises incur lower costs and enjoy higher returns. As government subsidies increase, the returns for logistics enterprises A and B continue to rise, leading them to choose to share charging facilities consistently. In this case, the stable strategy in the tripartite evolutionary game (active promotion, sharing, sharing) verifies Scenario 5.

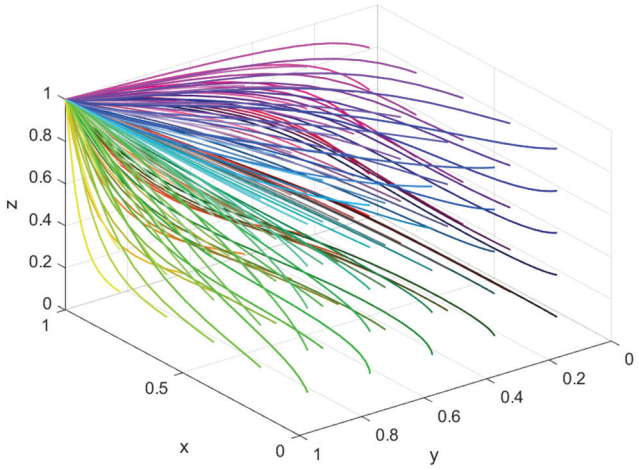


Figure 6. Evolutionary trajectories of Scenario 5.

5.2. Impact of Cost Allocation Coefficient on Evolutionary Strategies

The impact of the cost allocation coefficient on logistics enterprises' strategy selections can be analyzed by changing the value of r in our simulation study. We set r to be 0.2, 0.5, and 0.8 while assigning other parameters as follows: $I_a = 50$, $I_b = 50$, $T_a = 5$, $T_b = 5$, $P_a = 8$, $P_b = 8$, $C = 5$, $Q = 5$, $S_a = 15$, $S_b = 15$, $G = 50$, $g = 20$, $D = 8$, $e = 10$, $L = 20$, and $\mu = 0.5$. The resulting evolutionary strategies of the simulation study are depicted in Figure 7.

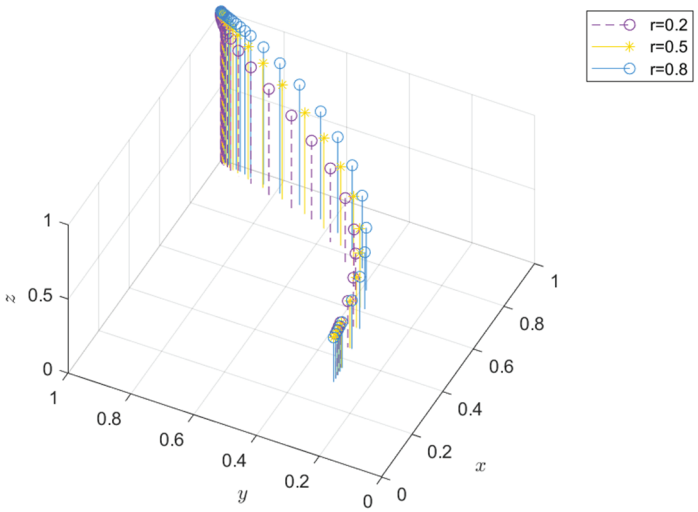


Figure 7. The impact of the cost allocation coefficient.

As shown in Figure 7, when the cost allocation coefficient is low, logistics enterprise A is more likely to choose cooperation, with its willingness to cooperate approaching 1 at a faster rate. This is primarily because a lower cost allocation coefficient means that logistics enterprise A bears less of the cost for sharing charging facilities, making it more inclined to reduce its own operating costs and enhance efficiency through sharing. The low distribution coefficient allows logistics enterprise A to access shared facilities with a smaller investment, effectively reducing its capital expenditure and operating costs for

charging infrastructure. This, in turn, enhances its market competitiveness. Conversely, for logistics enterprise B, the situation is the opposite. When the cost allocation coefficient is high, logistics enterprise B must shoulder a more significant portion of the sharing costs, which may exceed their financial capacity or expected returns, leading it to potentially opt out of sharing to avoid additional expenses. The high-cost sharing method may lead logistics enterprise B to incur expenses that outweigh the benefits it receives from sharing, ultimately diminishing its willingness to participate. According to the evolutionary game model, the cost constraints for logistics enterprises A and B can be expressed as $rC + rQ < \mu L + S_a + T_a - P_a$ and $(1-r)C + (1-r)Q < (1-\mu)L + S_a + T_a - P_a$. These inequalities represent the upper limits of costs that logistics enterprises A and B can bear, given different cost allocation coefficients.

Additionally, research on the dynamic stability of supply chain costs reveals that member firms have varying preferences for long-term versus short-term benefits, directly influencing their cost allocation decisions. Suppose that a firm prioritizes the advantages of long-term cooperation and sharing. In that case, it is more likely to opt for cooperation and form an alliance quickly when the cost allocation coefficient is lower. However, logistics enterprises A and B face different challenges depending on whether the cost allocation coefficient is too large or too small. Specifically, if the cost allocation coefficient is too high, logistics enterprise A may opt out of sharing the charging facilities, and logistics enterprise B might also refrain from sharing due to the significant cost burden. Conversely, when the cost allocation coefficient is too low, both parties may decide to share, but it may take longer to reach a consensus and establish a stable alliance.

In summary, the cost allocation coefficient directly impacts the willingness of logistics enterprises A and B to cooperate and their strategic choices. A lower cost allocation coefficient facilitates the quick formation of a shared alliance, whereas a higher coefficient may hinder both parties from reaching an agreement promptly.

5.3. Impact of Government Subsidies on Evolutionary Strategies

The impact of government subsidies on the strategy selections of logistics enterprises can be studied by changing the value of L to be 10, 20, and 30 in our simulation analysis while maintaining other parameters as follows: $I_a = 50$, $I_b = 50$, $T_a = 5$, $T_b = 5$, $P_a = 8$, $P_b = 8$, $C = 5$, $Q = 5$, $r = 0.5$, $S_a = 15$, $S_b = 15$, $G = 50$, $g = 20$, $D = 8$, $e = 10$, and $\mu = 0.5$. The evolutionary strategies of the simulation results are shown in Figure 8.

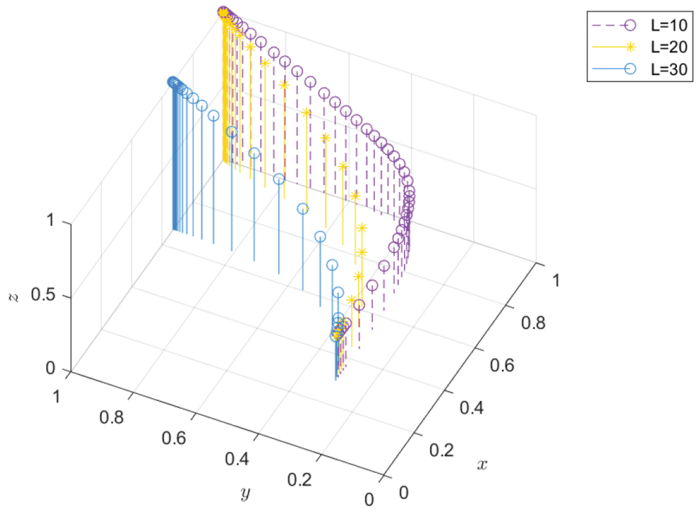


Figure 8. The impact of government subsidies.

As illustrated by the evolution results in Figure 8, government subsidies have a substantial positive impact on the willingness of the two logistics enterprises to cooperate. Specifically, when government subsidies are significant, the willingness to cooperate of

both enterprises approaches 1 more rapidly, indicating a stronger inclination to collaborate. This suggests that enterprises gain greater benefits and incentives with higher government subsidies, accelerating consensus and fostering cooperation in sharing charging facilities.

However, the impact of government subsidies varies depending on their size. When subsidies are too large, they might lead to a scenario in which enterprises receiving higher subsidies choose the sharing strategy. At the same time, the government, which bears the financial burden, may opt to reduce its promotion of charging facility sharing among logistics enterprises. Excessive subsidies can cause businesses to become overly dependent on government support, diminishing their motivation and initiative. On the other hand, when subsidies are too small or non-existent, the incentive for businesses to cooperate diminishes, as the economic benefits are insufficient to justify the collaboration. According to the evolutionary game model, the range of government subsidies should be $L < G - g$.

Furthermore, the government's approach to promote cooperation plays a crucial role. Active measures, such as clear policies and technical support, can foster collaboration between enterprises. However, if the government's promotion efforts are vague or poorly executed, the pace of cooperation may slow down. Thus, both government subsidies and promotion strategies significantly influence the willingness of logistics enterprises to cooperate. Government subsidies can mitigate the risks associated with participation in shared charging facilities. For logistics enterprises, charging facility sharing represents a new business model that entails certain uncertainties and risks. By providing financial support, government subsidies can help alleviate these concerns, encouraging companies to embrace this innovative approach to collaboration. Properly calibrated subsidies can encourage cooperation, while overly generous or insufficient subsidies and inadequate promotion strategies can impede the development of partnerships. Therefore, it is essential for the government to carefully consider multiple factors when crafting policy measures to ensure they are effective and adaptable.

Government subsidies are crucial in fostering cooperation between logistics enterprises, but their effectiveness is closely tied to the subsidy amount. Properly calibrated subsidies can significantly boost the willingness of enterprises to collaborate and accelerate their transition to cooperative arrangements. However, overly generous subsidies can lead to an overreliance on government support, which diminishes the motivation and initiative of the companies. Therefore, when designing subsidy policies, the government must carefully balance the subsidy amounts to ensure mutual benefits for both enterprises and the government, ultimately promoting the sustainable development of shared charging facilities.

5.4. Impact of Subsidy Allocation Coefficient on Evolutionary Strategies

We analyze the impact of the subsidy allocation coefficient on the logistics enterprises' strategy selections by changing μ value to be 0.2, 0.5, and 0.8, in the simulation when assigning other parameters as follows: $I_a = 50$, $I_b = 50$, $T_a = 5$, $T_b = 5$, $P_a = 8$, $P_b = 8$, $C = 5$, $Q = 5$, $r = 0.5$, $S_a = 15$, $S_b = 15$, $G = 50$, $g = 20$, $D = 8$, $e = 10$, and $L = 20$. The resulting evolutionary strategies of the game are illustrated in Figure 9.

As shown in Figure 9, an increase in the government's subsidy allocation coefficient significantly raises the likelihood of logistics enterprise A choosing to cooperate, with its willingness to cooperate reaching total commitment more rapidly. This indicates that as the government's subsidy allocation increases, logistics enterprise A becomes more inclined to collaborate, thereby speeding up the formation of a charging facility-sharing alliance. However, the situation is reversed for logistics enterprise B. When the government's subsidy allocation coefficient is either too high or too low, logistics enterprises that receive more substantial subsidies may opt to share charging facilities, while the other party, receiving smaller subsidies, might choose not to share. This disparity can prolong the time required for both logistics enterprises to establish a charging facility-sharing alliance. Therefore, the government's subsidy allocation coefficient plays a critical role in influencing the cooperation willingness and pace of logistics enterprises A and B. Specifically, a higher

subsidy allocation coefficient enhances logistics enterprise A's willingness and speed to cooperate. In contrast, logistics enterprise B demonstrates a stronger willingness to cooperate but at a slower pace.

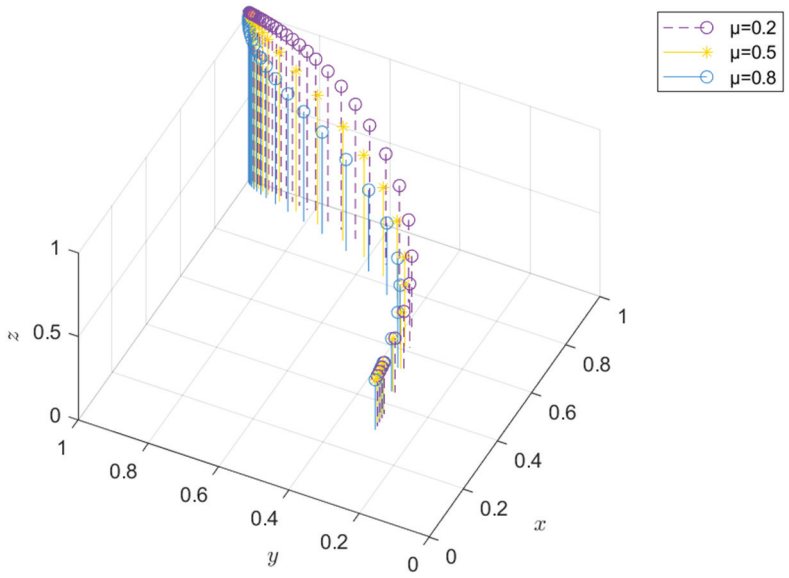


Figure 9. The impact of subsidy allocation coefficients on evolutionary strategies.

6. Conclusions

This research paper, grounded in evolutionary game theory, develops an evolutionary game model to explore the sharing of charging facilities in the logistics distribution of electric vehicles, incorporating the role of government as a key participant. Unlike previous studies that focused solely on the decision-making of logistics enterprises, this paper introduces the government's influence, emphasizing the impact of government policy support on the sharing of charging facilities. This approach aligns with sustainable development's current needs and realities in a green economy. The key findings are as follows:

1. Flexible Government Subsidy Measures: The numerical simulation analysis reveals that government subsidies can significantly foster collaboration among all parties involved. Furthermore, an appropriately designed subsidy mechanism can help sustain long-term evolutionary stability in cooperative efforts. Governments should implement adaptive subsidy strategies at different stages of development. In the initial phase of charging facility sharing, lenient subsidies are necessary to lower the initial costs for logistics enterprises, thereby facilitating the rapid growth of shared charging facilities. The government's subsidy policy should be less than the difference between the social benefits gained through the positive guidance strategy and those achieved through the negative guidance strategy. As the sharing system matures, the government should increase subsidies to encourage further investment and maintenance of shared infrastructure. Additionally, rewarding enterprises that actively participate in sharing can boost enthusiasm and accelerate the adoption of shared charging facilities in electric vehicle logistics.
2. Increased Investment by Logistics Enterprises: Numerical simulation analysis reveals that the lower the cost of sharing, the more willing logistics enterprises are to participate. Specifically, when charging facilities are shared among logistics companies, the additional revenue generated from distribution, the government subsidies and incentives, and the costs incurred if no sharing alliance is formed must outweigh the combined operation, maintenance, energy procurement, and risk-related costs associated with sharing. As the sharing of charging facilities progresses, logistics enterprises must invest more in the construction and upkeep of these facilities. Challenges such as insufficient facilities and poor location distribution may arise in the mature

stage. Logistics enterprises should optimize the construction and maintenance of charging facilities to ensure sufficient availability and strategic distribution. This will enhance the willingness of logistics companies to share charging resources, ultimately improving distribution efficiency and boosting customer satisfaction.

3. Enhanced Cooperation and Supervision: As the sharing of charging facilities advances and the volume of logistics activities increases, the frequency of facility use will also increase. To manage this, both the government and logistics enterprises must strengthen their cooperation and establish robust supervision mechanisms. This collaboration will ensure the rational use and maintenance of charging facilities, support the smooth operation of electric vehicle distribution, and further promote the widespread adoption of green distribution practices.

In conclusion, effective collaboration between government agencies and logistics enterprises is crucial for the sustainable development of shared charging facilities in electric vehicle distribution. This partnership not only contributes to achieving energy conservation and emission reduction goals and fostering a green economy but also provides essential support for logistics enterprises in enhancing operational efficiency and customer satisfaction.

7. Future Research Directions

In future research on the evolutionary game of charging facility sharing among logistics enterprises, with the government as a key participant, several factors can be explored to enrich the analysis:

1. Geographical Diversity: The geographical environment and infrastructure conditions vary significantly across different regions, impacting the sharing behavior of logistics enterprises. Future studies should examine how factors like urban versus rural energy distribution affect strategic decisions regarding charging facilities.
2. Regional Energy Price Discrepancies: Fluctuations in energy prices directly influence operating costs. Future research should account for regional variations in energy prices and their effect on charging facility sharing. For instance, regions benefiting from lower energy costs due to policy incentives might see increased participation in shared facilities.
3. Dynamic Demand Volatility: The logistics industry is characterized by fluctuating demand, driven by unexpected events or seasonal changes. Future studies should incorporate dynamic demand variability into models of charging facility sharing to better predict and respond to shifts in logistics needs.
4. Varying Company Sizes: Logistics enterprises differ in size and resources, which influences their strategic choices regarding charging facilities. While large companies may have greater access to funding and technology, smaller firms may rely more on government subsidies or partnerships. The impact of these differences should be analyzed.
5. Fleet Size Diversity: The size of vehicle fleets affects charging facility requirements. Larger fleets may necessitate more infrastructure, while smaller fleets may have more flexibility in utilizing existing facilities. Future research should explore the implications of fleet size on charging behavior.
6. Sources of Funding: Logistics enterprises access funds through various channels, including self-financing, bank loans, and government subsidies. Different funding sources can shape investment and operational strategies. For example, companies that depend on government subsidies may be more inclined to collaborate on the development of charging infrastructure.
7. Technological Innovation: Advancements in technology play a vital role in optimizing logistics operations. Future studies should investigate how innovations such as big data and artificial intelligence can enhance the efficiency of charging facility sharing and management.
8. Competition: The competitive behavior of logistics enterprises in charging facility sharing is another important factor. Competition could influence the layout and

efficiency of charging infrastructure. Future research should explore how competitive dynamics affect sharing strategies and outcomes.

9. Public–Private Partnerships (PPPs): Public–private partnerships are becoming increasingly common in the development and operation of charging infrastructure. Future studies should assess how PPP models can foster cooperation between governments and businesses to optimize charging facility sharing.

By considering these factors, future research can develop more sophisticated and realistic evolutionary game models, providing deeper insights into the dynamics of charging facility sharing among logistics enterprises and offering practical guidance for sustainable development.

Author Contributions: Conceptualization, G.X.; methodology, G.X. and D.Z.Y.; formal analysis, G.X. and J.C.; investigation, D.Z.Y. and Y.L.; writing—original draft preparation, J.C.; writing—review and editing, G.X. and D.Z.Y.; funding acquisition, G.X. All authors have read and agreed to the published version of the manuscript.

Funding: This research is funded by the Open Fund of Chongqing Key Laboratory of Green Logistics Intelligent Technology (Chongqing Jiaotong University) (Grant No. KLILN2023YB005) and the Chongqing Education Commission Humanities and Social Science Planning (Grant No. 22SKGH174).

Data Availability Statement: All data generated or analyzed during this study are available from the corresponding author upon request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Yang, K.; Sun, H.; Yang, L. Research on the optimization of Jingdong fresh food logistics distribution path based on electric logistics vehicle. In Proceedings of the Third International Conference on Intelligent Traffic Systems and Smart City (ITSSC 2023), Xi'an, China, 10–12 November 2024; pp. 9–15.
2. Wang, H.P.; Yan, X.C.; Zhao, D.; Ge, L.J. Research on production decision of automobile manufacturers considering consumers' low-carbon preference under government reward and punishment mechanism. *Syst. Eng. Theory Pract.* **2023**, *43*, 2669–2684.
3. National Road Transport Standardization Technical Committee. Notice on Soliciting Opinions on the National Standard "Technical Requirements for Urban Green Freight Distribution Evaluation (Draft for Comment)". Available online: https://jtst.mot.gov.cn/zxd/seekPublicAdvice/pagePublishAdviceStdList/819 (accessed on 3 August 2024).
4. Central Government of the People's Republic of China. Guiding Opinions of the General Office of the State Council on Further Building a High-Quality Charging Infrastructure System. Available online: https://www.gov.cn/zhengce/zhengceku/202306/content_6887168.htm (accessed on 3 August 2024).
5. China Journal of Logistics and Purchasing. The New Energy of Logistics Vehicles Faces Five Major Obstacles. Available online: https://new.qq.com/rain/a/20231014A065GL00 (accessed on 27 September 2024).
6. Floor Iron. New Energy Logistics Vehicle Development Status Analysis of Ground Rail New Energy Vehicles. Available online: https://www.dstcar.com/index.php/portal/article/index.html?id=437&cid=21 (accessed on 27 September 2024).
7. China Energy News. Three Major Problems Hinder the Promotion of New Energy Logistics Vehicles. Available online: http://paper.people.com.cn/zgnyb/html/2021-01/04/content_2027642.htm (accessed on 27 September 2024).
8. Chen, R.; Qian, X.W.; Miao, L.X.; Ukkusuri, S.V. Optimal charging facility location and capacity for electric vehicles considering route choice and charging time equilibrium. *Comput. Oper. Res.* **2020**, *113*, 104776. [CrossRef]
9. Luo, Q.Y.; Ye, Z.H.; Jia, H.F. A charging planning method for shared electric vehicles with the collaboration of mobile and fixed facilities. *Sustainability* **2023**, *15*, 16107. [CrossRef]
10. Chen, Q.Q.; Huang, K.; Ferguson, M.R. Capacity expansion strategies for electric vehicle charging networks: Model, algorithms, and case study. *Nav. Res. Logist.* **2022**, *69*, 442–460. [CrossRef]
11. Luo, X.Y.; Qiu, R. Electric vehicle charging station location towards sustainable cities. *Int. J. Environ. Res. Public Health* **2020**, *17*, 2785. [CrossRef]
12. Lv, Z.P.; Song, Z.H.; Li, L.S.; Liu, Y.; Zhou, S. Research on Electric Vehicle Charging Facility Planning Based on Improved Grey Wolf Optimization Algorithm in V2G Mode. In Proceedings of the 2023 3rd Power System and Green Energy Conference (PSGEC), Shanghai, China, 24–26 August 2023; pp. 325–331.
13. Woo, S.; Bae, S.; Moura, S.J. Pareto optimality in cost and service quality for an electric vehicle charging facility. *Appl. Energy* **2021**, *290*, 116779. [CrossRef]
14. Wu, W.J.; Deng, L.; Tao, Y.; Wang, X. Research on evolutionary game of value co-creation behavior of shared private charging piles of electric vehicles. *Discret. Dyn. Nat. Soc.* **2022**, *2022*, 9884962. [CrossRef]
15. Huang, X.; Li, L.; Xu, P. Electric vehicle charging pile sharing method based on multi-subject game and win-win. *Trans. China Electrotech. Soc.* **2023**, *38*, 2945–2961.

16. Chen, J.; Huang, X.Q.; Cao, Y.J.; Li, L.Y.; Yan, K.; Wu, L.; Liang, K. Electric vehicle charging schedule considering shared charging pile based on generalized Nash game. *Int. J. Electr. Power Energy Syst.* **2022**, *136*, 107579. [CrossRef]
17. Ji, J.H.; Bie, Y.M.; Wang, L.H. Optimal electric bus fleet scheduling for a route with charging facility sharing. *Transp. Res. C Emerg. Technol.* **2023**, *147*, 104010. [CrossRef]
18. Wang, Y.X.; Zhao, Z.L.; Baležentis, T. Benefit distribution in shared private charging pile projects based on modified Shapley value. *Energy* **2023**, *263*, 125720. [CrossRef]
19. Yang, X.; Liu, J.X.; Zhuge, C.X.; Wong, A.T.C.; Wang, P.X. Exploring the potential of sharing private charging posts: A data-driven micro-simulation approach. *Sustain. Cities Soc.* **2024**, *100*, 105053. [CrossRef]
20. Hu, X.; Yang, Z.J.; Sun, J.; Zhang, Y.L. Sharing economy of electric vehicle private charge posts. *Transp. Res. B Methodol.* **2021**, *152*, 258–275. [CrossRef]
21. Zhao, D.; Yan, X.C.; Wang, H.P.; Li, Y. Under double integral policy auto enterprise cooperative innovation evolutionary game analysis. *Manag. Sci. China* **2024**, *2021*, 1443.
22. Zhao, D.; Ji, S.; Wang, H.; Jiang, L. How do government subsidies promote new energy vehicle diffusion in the complex network context? A three-stage evolutionary game mode. *Energy* **2021**, *230*, 120899. [CrossRef]
23. Zhang, G.S.; Wang, X.; Wang, Y.L.; Kang, J.Y. Research on the resilient evolutionary game of logistics service supply chain with government participation. *Mathematics* **2022**, *10*, 630. [CrossRef]
24. Luo, Y.M.; Zhang, Y.K.; Yang, L. How to promote logistics enterprises to participate in reverse emergency logistics: A tripartite evolutionary game analysis. *Sustainability* **2022**, *14*, 12132. [CrossRef]
25. Dong, Y.; Yang, T.T. Evolutionary game analysis of promoting the development of green logistics under government regulation. *JUSTC* **2022**, *52*, 4. [CrossRef]
26. Li, J.C.; Yang, S.H. Analysis of a multiparticipant game under a subsidy and punishment mechanism: An evolutionary theory perspective. *Math. Probl. Eng.* **2021**, *2021*, 1984676. [CrossRef]
27. Zhang, L.L.; Long, R.Y.; Huang, Z.; Li, W.B.; Wei, J. Evolutionary game analysis on the implementation of subsidy policy for sustainable transportation development. *J. Clean. Prod.* **2020**, *267*, 122159. [CrossRef]
28. Wang, Z.H.; Feng, Y.; Yao, N.N.; Lin, O.W.; Li, Q.Y.; Liu, B.W. Impact of policy interventions on low-carbon technology innovation diffusion in supply networks. *Manag. Decis. Econ.* **2024**, *45*, 5040–5053. [CrossRef]
29. Yu, L.Y.; Xu, H.; Zhang, Z.Y. Tripartite evolutionary game analysis of supply chain low-carbon transformation considering the coupling of carbon tax and subsidy regulations. *Manag. Decis. Econ.* **2024**, *45*, 702–725. [CrossRef]
30. Yuan, X.M.; Zheng, C.C. Evolutionary game and simulation analysis of low-carbon technology innovation with multi-agent participation. *IEEE Access* **2022**, *10*, 11284–11295. [CrossRef]
31. Guo, L.B.; Zhang, Q.Q.; Wu, J.; Santibanez Gonzalez, E.D.R. An evolutionary game model of manufacturers and consumers' behavior strategies for green technology and government subsidy in supply chain platform. *Comput. Ind. Eng.* **2024**, *189*, 109918. [CrossRef]
32. Liu, Z.; Qian, Q.S.; Hu, B.; Shang, W.L.; Li, L.L.; Zhao, Y.J.; Zhao, Z.; Han, C.J. Government regulation to promote coordinated emission reduction among enterprises in the green supply chain based on evolutionary game analysis. *Resour. Conserv. Recycl.* **2022**, *182*, 106290. [CrossRef]
33. Wang, J.T.; Liu, Y.; Li, Y.F. A game study on low-carbon transportation behavior considering government subsidies. In Proceedings of the 2023 11th International Conference on Traffic and Logistic Engineering (ICTLE), Macau, China, 25–27 August 2023; pp. 95–99.
34. Sun, H.X.; Li, J.L. Behavioural choice of governments, enterprises and consumers on recyclable green logistics packaging. *Sustain. Prod. Consum.* **2021**, *28*, 459–471. [CrossRef]
35. Selten, R. A note on evolutionarily stable strategies in asymmetric animal conflicts. *J. Theor. Biol.* **1980**, *84*, 93–101. [CrossRef]
36. Friedman, D. Evolutionary games in economics. *Econometrica* **1991**, *59*, 637–666. [CrossRef]

mathematics

MDPI

Article

Shortage Policies for a Jump Process with Positive and Negative Batch Arrivals in a Random Environment

Yonit Barron

Industrial Engineering and Management, Ariel University, Ariel 40700, Israel; ybarron@ariel.ac.il;
Tel.: +972-50-7749774

Abstract: We study a continuous-review stock management of a retailer for a single item in a limited storage (buffer) in a random environment. The stock level fluctuates according to two independent compound Poisson processes with discrete amounts of items (batches) that enter and leave the storage facility. The storage facility is controlled by a three-parameter base-stock replenishment policy. All items exceeding the storage capacity are transferred to an unlimited foreign facility. In addition, a restricted backlogging possibility is permitted; additional demands for items are lost sales. We further assume a random shelf life, the possibility of total inventory collapse, and a random lead time. Applying Markov theory, we derive the optimal control parameters minimizing the long-run expected total cost. A sensitivity analysis is conducted focusing on the comparison between the pure lost-sales policy and a partial backordering policy. Accordingly, we identify cases where one policy is cost effective compared to the other, particularly with respect to the batch patterns (sign, rate, average, and variability), and the associated costs.

Keywords: inventory; batch arrival; base-stock policy; backlog; lost sales

MSC: 90B05; 90B06; 60J28; 60G51

1. Introduction

The planning of a supply chain subject to market uncertainty is challenging and requires decisions on when and how many items should be ordered. This challenge is increased due to rapid changes in the economic environment, increasing both customer intolerance and market competitiveness. As a result, retailers have to anticipate increasing changes in customer consumption and returned items as e-commerce sales soar and, hence, must carefully manage their stock in order to reap the benefits of these changes. Despite the implementation of advanced inventory-management tools and advanced forecasting of customer demands and returns, shortages are still inevitable. With regard to managing stock, retailers must maintain a high stock level since shortages reduce customer satisfaction and service levels. On the other hand, maintaining a high stock level that is further increased by returned items reduces the profit. Faced with this trade-off, retailers must decide whether to allow a (perhaps limited) backlogged level or to direct the customer to a competing company, thereby losing his/her loyalty. In addition, the complexity of managing stock is exacerbated by other factors, such as random delivery times, random amount sizes, random shelf lives, and unexpected events.

Motivated by the above problem, we consider the control management of a retailer with a single-item two-sided bounded storage (buffer) that faces jump-pattern arrivals. The storage facility is continuously controlled according to a triple-parameter base-stock (S, s, B) policy for $0 \leq s < S$ and $B \geq 0$, which is an extension of the conventional (S, s)-type. Under the (S, s) policy, when the on-hand stock level drops to level s or below, an order is issued to bring the stock up to level $S > s$, which the supplier will carry out after an exponentially distributed lead time. We extend the (S, s) policy by assuming a combination

of backorders and lost sales. That is, in the case of zero on-hand stock, limited B backlogged items are allowed and additional demands for items are lost. Thus, S is the replenishment level, s is the reorder point, and B is the maximum backlogged items permitted during a stockout. The stock level process belongs to the class of jump models: it jumps upward and downward at random times, and stays constant in between jumps. When a positive (negative) batch arrives, the stock increases (decreases). A positive jump is caused by the entry of returned or loaded items into the system, or by the receipt of produced items from the manufacturer. A negative jump is caused by the demand or consumption of items. The batch arrivals are governed by two independent compound Poisson processes with positive batches, and each batch size follows a given discrete probability distribution. Since the storage facility has a limited capacity, any arriving amount exceeding it is transferred immediately to a second storage facility. We further assume two prototypes of shelf lives for the items held in the storage facility: each item has an i.i.d. exponentially distributed shelf life, after which it is useless and out of stock, and a total loss of all items in stock occurs at exponential times. The random shelf life of an item can be the result of spoilage, perishability, or failure. The possibility of losing all items may occur due to a malfunction of the storage facility, the bankruptcy of the retailer, an obsolescence event, an external disaster, or changing customer needs. In such cases, all on-hand items (if any) lose their value at once, and the stock collapses instantaneously to zero, after which the items are unusable.

The customer–retailer–supplier flow, operations processes, and outputs/inputs from/to the retailer are illustrated in Figure 1. We use blue and red arrows to indicate positive batches' inputs and negative batches' outputs of items, respectively.

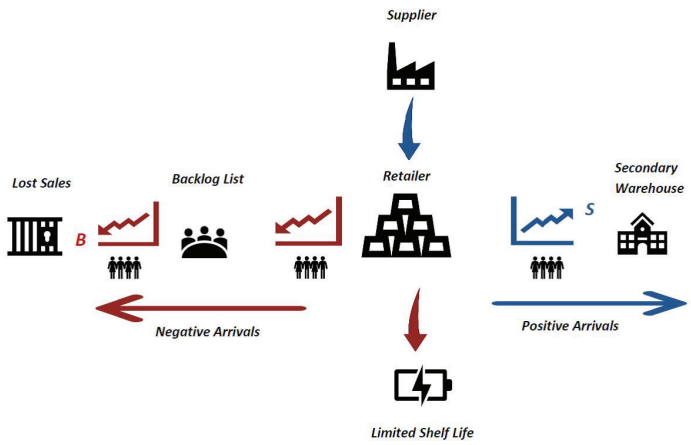


Figure 1. A general supply chain structure and operations process.

Our study assumes negative and positive jumps of random sizes occurring at random times. Positive jumps may be caused, e.g., by returned items from customers, returned items for recycling, used product collection, production packets received from the manufacturing plant [1], and canceled orders, such as booked rooms in hotels. With regard to returned items, customers increasingly opt to return products, and with the higher share of e-commerce channel sales, product return rates are soaring upwards. Therefore, it is no longer feasible for retailers to ignore returned items when determining the optimal inventory policy since it could result in overstocking and overestimation of the profit. Moreover, retail e-commerce has been growing rapidly and, with it, the number of returned items [2]. For example, the National Retail Federation found that, in the U.S., USD one billion in merchandise was returned, which is a 66% increase from five years earlier. The lack of power of touch at online stores further increases the number of returns; brick-and-mortar stores have a return rate as high as 10% of total sales [3]; U.S. consumers returned goods worth USD 261 billion, and the return rates for online sales sometimes exceeded 30% [4]; a California-based online clothing retailer disclosed that it made just under USD 400 million in net sales a year ago, but paid out USD 385 million for returned items. Buyers return a

huge number of packages they buy from Amazon and other e-commerce sites, so much that retailers are sometimes left with little choice but to get rid of large swaths of inventory at a cost https://www.cnbc.com/2018/12/13/returned-goods-are-a-problem-for-retailers-resellers-are-cashing-in.html, accessed on 1 January 2024. A recent data on sales, returns, and return rates can be found in [2,4,5].

Our model is motivated by several practical applications. Take, for example, Israeli organic vegetable marketing farm Aley Bodek (https://aleybodek.co.il, accessed on 1 January 2024). The farm, established in the early 2000s in Kibbutz Be'erot Yitzhak in Israel, uses hydroponic methods of growing vegetables and fruits, using substrates transplanted from the ground to greenhouses that ensure fresh, high-quality agricultural produce. Aley Bodek markets its produce in packed boxes to specific retailers, who, in turn, market the produce to industrial plants, hotels, and other institutions. Clearly, the supply is random and depends on weather and other conditions. When the stock at the retailer falls below a certain level, an order is placed and the product arrives from the farm after some random time. Due to the uniqueness of the product, the retailer maintains a limited waiting list for the product, while customers not waitlisted are directed to competing companies. Another common application can be found in health systems, especially the management of blood units in emergency and operating rooms. The blood units are necessary for all medical activities, and usually arrive in batches from a central blood bank. Here, positive jumps represent the arrival of units, and negative jumps represent patients' needs. Of course, a safe stock of blood units is required. Since the capacity of the storage facility is limited, the surplus blood units are disposed of or used for other needs, such as academic or laboratory research. The negative stock expresses the number of blood units required for the non-urgent patients [6]. Another example comes from reliability and maintenance services. Consider the case of parts delivered to maintenance sites in isolated locales (where the delivery requires some lead time). Repair activities increase the stock level, and using the parts lowers the stock level. Due to the geographical distance, the maintenance site allows backlogged parts when stock is out; when the next replenishment arrives, these backlogged parts are satisfied first, and then, the stock is filled up to its maximum capacity (see, e.g., [7]).

We note that, although our model is described in the context of the stock management of a retailer, it can be generalized to a wide variety of other contexts. For example, in financial- and cash-management models, a positive jump may represent a bailout from the government, and a negative jump may represent the effect of a financial crisis [8]. Take, for example, the Israeli contracting company Sharbat Brothers. Sharbat Brothers invests in the construction of buildings around the world. The company's cash flow includes rental fees (positive jumps) and maintenance costs and other transactions for the public sector (negative jumps). The bank requires the company to maintain a certain level of financial cash flow, below which the required amount of money is transferred from other sources with some overhead time. The negative cash level represents a deficit; any financial withdrawal beyond the maximum permitted level is not allowed and forces the company to turn to more expensive sources. For additional examples, see, e.g., [9,10].

This paper further assumes that the positive and negative jumps arrive according to independent compound Poisson processes. The Poisson assumption is a common one in inventory models; statistical analysis has shown that it models customers' needs to retailers in service systems, queueing systems, and inventory systems well. One of the specific advantages of the Poisson process is that it exactly captures a similar effect as the short-term non-stationarity encountered in some arrival processes [11–13].

We further consider a random batch size. Clearly, a random batch size complicates the analysis since each inventory level faces two-sided jumps from several levels and some levels are not necessarily occupied between two replenishments. In practice, retailers usually deal with batch pattern arrivals. Examples include spare parts and drugs [14] and supplier-constrained or discounted items. Other examples come from chemical processing, water purification systems [15], medical equipment, and out-of-fashion products such as

newspapers or personal computers [16]. As an excellent application for a batch type, we mention the tourism companies managed by travel agents. Travel agents reserve several hotel rooms in advance or allocate seats on planes (negative jumps) and may cancel the entire reservation due to a security situation or other unexpected event (positive jumps). Similarly, there are activities for teens that are given at clubs or at academic institutions and are open only through centralized registration at the schools.

Our model employs a triple-parameter (S, s, B) control policy, which extends the base-stock (S, s) replenishment policy by allowing limited shortages of up to B items. In general, the base-stock class includes two common policies: the constant-order policy (Q, r) and the variable-order up-to-level policy (S, s). In this paper, we assume the latter. As the above examples demonstrate, a variable-order size is a practical control policy in numerous scenarios [17].

As a result of the random arrivals (in terms of time and size) and uncertain shelf lives, shortages are common in most customer–retailer–supplier flows. Especially currently, with the quick flow of open information, any shortage of inventory causes dissatisfaction, lowers the service level, redirects customers to competing retailers, and increases unsupported reviews. There are two common policies for tackling shortages: the backordering policy, where unsatisfied demand is put on a waiting list until the next replenishment, and the lost-sales policy, where unsatisfied demand is redirected to other providers and is considered as lost sales. Both policies have a strong anchor in the literature and in practice. The backordering policy is commonly employed in the case of famous brand products, fashionable commodities, or critical products. For example, backordering is necessary to ensure the reliability and availability of critical products like military systems and medical supplies [18]. In monopolies, the absence of competitors may lead to a situation where customers may prefer to wait while their items are backordered if a shortage occurs [19]. On the other hand, for more common items, only 15% of customers will wait if a shortage occurs, while the remaining 85% will visit another store [20]; this percentage may even increase with the volatility and predictability of the customers' needs [21]. Examples include the apparel industry, the grocery industry, the steel industry, and the fast-moving consumer goods industry; more examples are given in [22].

From the retailer's perspective, the backordering policy improves the service level and customer satisfaction; however, it requires the payment of overhead costs and the building of appropriate infrastructure. On the other hand, the lost-sales policy results in a loss of customers, and less profitability. In should be noted that many retailers allow some limited backordering, so as not to lose customers and to take advantage of the fluctuations in customers' consumption habits. To address this trade-off, we propose a combination of the two policies whereby some of the items (up to B) are backordered and additional demands for items are lost. An excellent example of such a combination of backordering and lost-sales policies is found in the field of healthcare. Here, customers are waitlisted for treatment by a specialist and, thus, can be considered backlogged items. However, the waiting list is limited in the number of waits. Thus, when the specialist's calendar fills up, the customers are referred to other specialists, and are considered as lost sales. For more examples, we refer the reader to [23,24] and the references therein.

We seek to determine the optimal control levels s^*, S^*, and B^* that minimize the overall average cost of managing the stock. To that end, we assume the following cost structure: (i) a fixed cost per order and a purchasing cost per item, (ii) a cost for handling a positive batch, (iii) a linear cost for holding the on-hand stock and a backordering cost for not satisfying the customer on time (both costs are charged per item per unit time), (iv) a transfer cost for each item exceeding level S, (v) a lost cost for unsatisfied demand beyond level -B, and (vi) a loss cost for each useless item due to end of life or an unexpected event.

Using a Markovian formulation, we derive the steady-state probabilities of having i ($-B \leq i \leq S$) items in the storage facility. These probabilities enable us to derive, by simulation-based optimization, the optimal control parameters and other operating characteristics of the system to be used. Based on an extensive set of experiments, we can

obtain valuable insights, such as the interplay between the positive and negative batches (rate, mean, and variance) and the impact of this interplay on the system's performance, the economic benefit of backordering compared to lost sales, and how the best policy responds to changes in the system's characteristics (costs, lead times, and arrivals). It should be noted that, from a theoretical point of view, using the exponential distribution for the shelf life and lead time simplifies the analysis due to the Markovian property. If, however, a general or constant distribution variable is introduced for the lead time or for the shelf life, the system becomes significantly complicated as its performance is a function of many factors. In such cases, strong assumptions are required in order to obtain explicit results, and deriving the optimal or near-optimal control parameters is analytically complex. However, when either the shelf life or the lead time is not exponentially distributed, our results may still be used as an approximation.

The contribution of this paper is fourfold. First, we develop an easy-to-implement mathematical framework for studying processes, and formulate closed-form (albeit cumbersome) expressions for stationary distributions and costs. These expressions are then used to obtain, numerically, the optimal control parameters S^*, s^*, and B^*, so as to minimize the average total cost per time unit. The scope of the problem covers a wide range of real-world problems, such as the uncertainty of lead times, random arrivals (rate, mean, and variance), shelf lives, and unexpected disasters. To the best of our knowledge, no mathematical framework studying this combination of backordering and lost-sales policies has been explored in the literature; thus, our model improves the understanding of stock management systems.

Second, we coordinate positive and negative random jumps to the arrival processes; in doing so, we capture the reality of the changing patterns of customers/manufacturers. Studying the impact of the positive jumps on the system's performance reveals that the optimal control parameters are significantly impacted by these jumps, and integrating them into the derivation may lead to substantial cost savings. Our results further imply that the optimal parameters are more sensitive to the average batch size than to its variability. Interestingly, when the jumps are fixed to k, the optimal parameters and cost can be approximated by k times the corresponding values obtained for the unit-size system (the Poisson process), especially when more negative batches arrive and at high lost cost. In addition, our numerical analysis demonstrates the similarities between systems with the same average total arrival patterns (positive and negative).

Third, we introduce a replenishment policy that is a combination of backordering and lost-sales strategies, and focus on determining the optimal backlogged level that minimizes the total cost. From a practical point of view, a mixture of backordering and lost sales emphasizes the interplay between the loss cost due to stockout events and the overhead cost due to a high stock level maintained to retain customers. Although several papers have considered a mixture of backordering and lost-sales policies, to the best of our knowledge, none of them have studied the framework addressed here; thus, our model significantly expands the perspective on dealing with shortage and has a wide range of applications. For example, our study shows that allowing backordering may yield substantial cost savings. Surprisingly, the optimal backlogged level B^* is decreasing (increasing) in outflow (inflow). That is, the fewer the negative batches there are, the more economically profitable the backordering policy is. It is further shown that the benefit of backordering increases with the variability of the outflows; however, the impact of inflows' variability is more complicated.

Finally, using numerical examples, we compare the backordering policy ($B > 0$) and the pure lost-sales policy (corresponding to $B = 0$) in a variety of scenarios, and provide retailers with managerial insights and practical applications when considering changing their replenishment policy during a stockout. By that, the current study is adapted to the dynamic reality of rapid changes in order to maintain a high service level and to stay competitive. For example, our comparison shows that the maximum backlogged cost,

beyond which the policy is not worthwhile, decreases as more outflows arrive, but as the variability increases, the outflows' arrival rate is less significant.

The remainder of the paper is organized as follows. In Section 2, we review the relevant literature. In Section 3, we present the basic features, assumptions, and notation. In Section 4, we formulate the mathematical model. Using numerical examples and a sensitivity analysis, Section 5 investigates the impact of the parameters on the system's performance under the two policies; a comparison, conclusions, and managerial insights derived from the results are presented. Finally, Section 6 presents concluding remarks and some directions for future research.

2. Literature Review

The main innovative focus of this paper is on two major topics: (1) systems with positive and negative jumps under the base stock policy and (2) different attitudes to stockout: a full backordering policy, a pure lost-sales policy, and a combination of both. In this section, we discuss previous research on these topics.

The literature on inventory systems under base-stock policies with inflows and outflows can be broadly separated into periodic- and continuous-review inventory systems. In this section, we address only continuous-review inventory systems. Schrady [25] was the first to introduce the concept of reusable inventory and study the classical economic order quantity problem (known as the EOQ model) with returned items. Gajdalo [26] extended Schrady's work by assuming independent compound Poisson processes for the inflows and outflows. Since those pioneering works, several policies have been suggested for a single-item inventory system with bilateral movements, e.g., an optimal (a, b) policy with no backorders or lost sales [27], an (s, Q)-type policy with backorders and fixed or zero lead times [28], differently sized lot sizes over a planning horizon [29], band policies with lost sales [30], and band policies with emergency supply [21]. For periodic-review systems with returned items, see [31] and the references therein.

Studies on inventory systems under the backordering policy include linear demand rate models [32], exponential demand rate models [33], EOQ and EPQ models [7], discrete probability processes [34], lot-sizing systems and remanufacturing models [35], power demand models [23], and stock-dependent demand models [36,37].

Base-stock replenishment systems with lost sales are more challenging, especially when dealing with a short shelf life. Here, the inventory level is constant during a shortage, and thus, the treatment of mathematical formulae for lost sales is more difficult than for backorders [11]. Hence, only various heuristic policies have been developed for continuous-review systems with lost sales. Studies on perishable systems with lost sales include [38,39]. We further refer the reader to [22,40] for excellent reviews of base-stock inventory systems with shortages.

In this paper, we address a policy that is a mixture of backordering and lost sales. A considerable body of literature has been written on this mixture. Montgomery et al. [41] were the first to study such a policy where α is the fraction of the excess demand backordered. Rosenberg [42] and Moinzadeh [43] reformulated Montgomery et al.'s model by assuming deterministic demand so that the optimal solution can be easily obtained. Several papers have studied the (Q, r) inventory models with partial backorders and lost sales (e.g., [44] and, more recently, [18,45]). Chang and Lo [46] proposed an approach to overcome the drawback of traditional methods for improving the continuous and discrete lead time with a mixture of backorders and lost sales. Sicilia et al. [33] analyzed an inventory system with a mixture of backorders and lost sales, where the backordered demand rate is an exponential function of the customer's waiting time. Applying a Markov decision model, Wang and Tang [47] obtained the optimal dynamic rationing levels for multiple demand classes' priorities. Taleizadeh and Zamani-Dehkordi [48] presented an inventory system with partial backordering where some of the backlogged items become lost sales. Li et al. [49] presented a production–inventory (M, m) control model in which a reflected Brownian motion governs the inventory level variation and the total amount of stockout is

a mixture of backordering and lost sales. Taleizadeh et al. [24] developed four independent EPQ profit-maximization problems for four different shortage situations. More recently, Wang et al. [50] built a computational model using the response surface methodology to determine the levels of factors, including the quantity of backlogged items. San-José et al. [23] derived the economic order quantity for a time-dependent power demand rate where only a fixed proportion of the demand during a stockout is satisfied.

In our study, we consider an exponentially distributed lead time and shelf life [14]. A stochastic lead time is a factor that introduces uncertainty and challenges the determination of the optimal policy. When an inventory system is characterized by random inflows and outflows, any change in lead time has an immediate impact on every step of the supply chain, as well as on the safety stock, the out-of-stock loss, and the service level [46]. In practice, a stochastic lead time fits the case where the lead time depends on different logistics factors [51]. For example, consider a retailer that has several independent suppliers, each behaving as an M/M/1 system. The lead time for each supplier can be interpreted as the total time for handling and delivering the batch. Queueing theory implies that the total (sojourn) time in such a system is an exponentially distributed random variable. We further assume a random shelf life. In particular for models with a variable lead time, the life time of items from the time they reach the retailer until the time they perish may be uncertain even for items with a known expiration date [17]. Therefore, our assumption of an exponential life time is more applicable than it seems. Here, Markovian models with random arrivals, exponential shelf lives, and total losses are discussed in [11] and the references therein.

To outline our position, an overview of the most relevant literature studies concerning the continuous-review base-stock policy is given in Table A1 in Appendix A. It seems that the discussed studies are not as comprehensive as the one we present in this paper. As far as we know, the combination of a continuous-review base-stock policy for jump processes with a mixture of backorderings and lost sales has not been explored in the literature; hence, the model developed below significantly contributes to the existing literature.

3. Description, Notation, and Assumptions

3.1. Problem Statement

We considered a continuous-review stock-level process $\mathbb{I} = \{I(t), t \geq 0\}$ fluctuating due to two independent arrival processes of negative and positive batches of items. The stock level $I(t)$ has a double-sided bound: an upper bound S and a lower bound -B, where -$B \leq I(t) \leq S$. The stock is managed according to a triple-parameter base-stock policy $(S, s, B), 0 \leq s < S, B \geq 0$, with random lead times and random shelf lives. Specifically, we assume the following:

(i) *Arrival processes*. We assume that the on-hand stock process is fluctuating due to negative and positive batches of items (downward and upward jumps, respectively). A negative batch represents items on demand or items stored for specific needs. A positive batch represents returned items, refunds, or arriving items from the manufacturer. The negative batches arrive according to a compound Poisson process with an exponentially distributed inter-arrival time at rate λ. The positive batches arrive according to another independent compound Poisson process with an exponentially distributed inter-arrival time at rate η (there are no restrictions on the ratio between η and λ). The (absolute) size of a negative batch is a random variable D, where D is independent of the arrival process, and has a discrete probability distribution function (PDF) $P_D(d), d = 1, 2, \ldots$, a cumulative distribution function (CDF) $F_D(\cdot)$, and a coefficient of variation cv_D. Similarly, the size of a positive batch is denoted by R, where R is a random variable independent of the arrival process, and has a PDF $P_R(r), r = 1, 2, \ldots$, CDF $F_R(\cdot)$, and cv_R. We note that, although we assume fixed rates λ and η, it is easy to generalize the model to support state-dependent rates $\lambda(i)$ and $\eta(i)$, respectively, as a function of i items in stock (i may be positive or negative; see Part 2 of Remark 1).

(ii) *Upper bound*. We assume that the on-hand stock process is limited from above by level S. This upper bound can be considered as the storage capacity, or the safety or sanitary constraints on the storage capacity. Thus, all items in the batch exceeding capacity S are immediately transferred to an unlimited secondary storage or warehouse, at some cost.

(iii) *Lower bound*. The stock level is bounded from below by level -B ($B \geq 0$). The negative stock level indicates the number of permitted backlogged items.

(iv) *Control policy of the stock*. We introduce the triple-parameter base-stock (S, s, B) control policy, which generalizes the well-known (S, s) policy. Under the (S, s) policy, whenever the stock level decreases to or below s (either due to negative batches, short shelf lives, or a total loss of all on-hand items in stock), a replenishment is ordered to bring the stock level up to its maximum capacity S. The replenishment is carried out after an exponential lead time. During the lead time, new orders are not allowed; thus, at most one oncoming order exists at any given time. Our policy implements a mixture of backorders and lost sales, i.e., in the case of zero on-hand stock, B backlogged items are allowed. Each backlogged item is incurred with some financial compensation per time unit until it is satisfied. A negative batch (or portion thereof) that arrives when $I(t) =$ -B is considered as lost sales, at the price of lost cost and the decrease in the service level. Note that the special case $B = 0$ means no backorders are allowed, meaning that a negative batch (or portion thereof) that arrives when there is a zero stock is considered as lost sales. This policy is known as a *pure* lost-sales policy. Thus, under the (S, s, B) policy, the quantity replenished at time t is random, depending on $I(t)$, and equals $max(S - I(t), 0)$.

(v) *Lead time*. We assume an *exponential lead time with parameter* μ.

(vi) *Limited shelf life*. We assume that each stored item lasts a random amount of time before it becomes unusable. The shelf life of an item is an i.i.d. r.v. with an exponential distribution at rate θ. We further consider a limited shelf life of all items in stock due to a disaster, an unexpected event, or obsolescence. Here, all on-hand items lose their value at once after an exponential time at rate Ξ, whereupon the stock collapses instantaneously to zero and all items become unusable.

Based on the assumptions outlined above, we distinguish between two shortage policies for different domains of B, namely $B = 0$ and $B > 0$:

(i) **Pure lost-sales policy**. Here, we assume that $B = 0$; i.e., no backlogged items are allowed, and $I(t)$ lies in the interval $[0, S]$. Any negative batch (or portion thereof) that faces zero stock is lost. A typical sample path of $I(t)$ under the lost-sales policy is depicted in Figure 2. The black (blue) arrows present negative (positive) batches. The number of transferred items and of lost items are indicated by the dashed blue and red lines, respectively. When $I(t)$ drops below level s, a replenishment is ordered and arrives after some lead time (*LT*); the lead time duration is indicated by the dotted gray line. Finally, the actual amount replenished is represented by the yellow line. Figure 2 illustrates two scenarios. During the first lead time (*LT* 1), we see that $I(t) > 0$; i.e., all batches are satisfied. By contrast, during the second lead time (*LT* 2), there are two incidents of lost sales. In the first lost-sales event, just a portion of the batch is lost (the first red segment). In the second lost-sales event, the entire batch is lost (the second red segment). Note that, when the order arrives, $\mathbb{I}$ starts over with S items; thus, due to the zero bound, the actual quantity replenished is always less than or equal to S items.

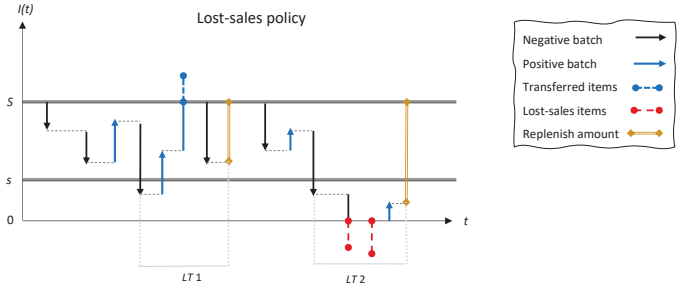


Figure 2. A typical sample path of $I(t)$ under the lost-sales policy.

(ii) **Mixture of backorders and lost-sales policy**. Here, we assume that $B > 0$; i.e., a maximum of B backlogged items are permitted, and all additional demand for items is considered as lost sales. Thus, $I(t)$ lies in the interval $[\text{-}B, S]$. For brevity, we will refer to this policy as a *backordering* policy. A typical sample path of $I(t)$ under the backordering policy is depicted on the right-hand side of Figure 3. We emphasize that, as long as $I(t) > 0$, the stock-level processes under the pure lost-sales and backordering polices have the same distribution. They are distributed differently when a negative batch arrives, causing the stock level to drop below level zero, until the replenishment order is carried out, whereupon $I(t)$ starts over at level S. This process is graphically represented under the magnifying glass in Figure 3. We see that, under the lost-sales policy, there are two incidents of lost sales, part of the first negative batch and the subsequent batch. Under the backordering policy, due to level $\text{-}B < 0$, the part that was lost in the first negative batch is backlogged and, regarding the subsequent batch, partly is backlogged and or partly is lost (the red segment on the right-hand side of Figure 3). Here, the actual quantity replenished can be even more than S items, as illustrated in Figure 3.

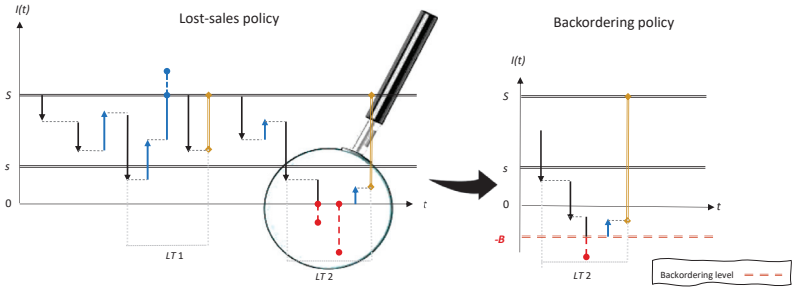


Figure 3. A typical sample path of $I(t)$ under the backordering policy.

3.2. Parameters and Costs

In this section, our notation and analysis are presented as a function of a general level B, $B \geq 0$. Substituting $B = 0$ leads to the pure lost-sales policy, while substituting $B > 0$ leads to the backordering policy. We start by summarizing the model's variables, parameters, and costs.

Decision variables

S	upper bound (the storage capacity and the replenishment level); $S > 0$ (items)
s	reorder level; $0 \leq s < S$ (items)
$\text{-}B$	lower bound (B is the maximum number of backlogged items); $B \geq 0$ (items)

Parameters and functions

μ	lead time rate (per time unit)
λ	arrival rate of negative batches (per time unit)
D	random (absolute) size of a negative batch (items)
$P_D(d), F_D(d)$	$P_D(d) = P(D = d)$ (PDF), $F_D(d) = P(D \leq d)$ (CDF)
$\widehat{F}_D(d)$	$\widehat{F}_D(d) = 1 - F_D(d) = P(D > d)$
cv_D	coefficient of variation of D
η	arrival rate of positive batches (per time unit)
R	random size of a positive batch (items)
$P_R(r), F_R(r)$	$P_R(r) = P(R = r)$ (PDF), $F_R(r) = P(R \leq r)$ (CDF)
$\widehat{F}_R(r)$	$\widehat{F}_R(r) = 1 - F_R(r) = P(R > r)$
cv_R	coefficient of variation of R
θ	expiration rate per item in stock, end of life (per time unit)
Ξ	loss rate of all items in stock (per time unit)

The costs

K_o, c_o	fixed and item purchasing cost per replenishment, respectively (USD/order)
c_θ	cost of each expired item (USD/item)
c_ϵ	cost of each lost item in case of total loss (USD/item)
c_l	penalty cost of an unsatisfied item (beyond the B backlogged items) (USD/item)
$\gamma(i)$	transfer cost of j items to a secondary storage facility (when the stock exceeds level S) (USD for j items)
c_r	overhead cost for each item of a positive batch (USD/item)
$h(i)$	holding cost for i on-hand items in stock per unit time ($0 < i \leq S$) (USD for i items)
$\beta(i)$	backordering cost for ($-i$) backlogged items per unit time ($-B \leq i < 0$) (USD for $-i$ items)
TC	long-run average total cost per unit time (USD/time unit)

Assumptions on the costs

(1) When the on-hand stock exceeds level S, a cost $\gamma(j)$ is charged for transferring the i excess items to a secondary storage facility. It is reasonable to assume that $\gamma(j)$ is an increasing and non-convex function of j. Formally, $\gamma'(j) > 0, \gamma''(j) \leq 0$; i.e., the growth rate of $\gamma(j)$ is decreasing (or constant) in j. That is, the greater the number of items that are transferred together in the same shipment, the lower the transfer cost of each item. This assumption is practical, since, usually, most of the transfer cost is for shipping and handling and is less affected by the number of items. Thus, we assume that $\gamma(j) = \Upsilon + c_\gamma \cdot (j)^\gamma, 0 < \gamma \leq 1$, where $\Upsilon, c_\gamma \geq 0$ are constant, and γ is the concavity coefficient. Note that a linear cost function is included, when $\gamma = 1$. It is worth pointing that our analysis is not limited to the above assumption, and also fits a general cost function.

(2) The holding cost $h(i)$ is linear in the number of the i on-hand items, and has the structure $h(i) = h \cdot i, 0 < i \leq S$. This assumption is reasonable due to the fixed capacity of the storage, and therefore, the significant increase in the holding cost is due to the cost of each item in stock.

(3) The backordering cost $\beta(i)$ is an increasing and non-concave function of -i, for -$B \leq i < 0$; i.e., $\beta'(\text{-}i) > 0$, $\beta''(\text{-}i) \geq 0$. That is, the growth rate of $\beta(i)$ is increasing (or constant) in -i. The convexity (non-concave) assumption means that, the more items that are backlogged (up to B items), the higher the compensation payment that the retailer pays. Thus, the backordering cost includes the additional cost due to the decrease in the service level.

(4) As in practice, we assume that $c_r < c_o$; i.e., the overhead cost of an arriving item is lower than the cost of ordering a new item.

(5) Finally, the cost of a lost-sales item is higher than the marginal cost of a backlogged item, i.e., $c_l > \beta'(\text{-}i)$ for -$B \leq i < 0$. This assumption is well established because, in today's competitive reality, lost-sales items result in the loss of customers. Conversely, in the case of backordering, customers remain loyal to the retailer, albeit with increasing reluctance.

As mentioned above, it may happen that, due to a low stock level, part of a negative batch is satisfied (decreasing the stock to level 0), part of it is backlogged (up to B items), and the remaining part is lost. Similarly, it may happen that a positive batch exceeds level S, in which case, all items above level S are transferred to a secondary storage facility. Specifically, assume that $i \geq 0$ items are held in stock. When a positive batch of D items arrives, we distinguish between three cases: (1) if $D \leq i$, then the whole batch is satisfied and the stock level continues with $(i - D)$ items (0 items possible); (2) if $i < D \leq i + B$, then i items are satisfied, and the rest of the $(D - i)$ items are backlogged at a cost of $\beta(i - D)$ (note that $D - i \leq B$); (3) if $D > i + B$, then B items are backlogged at a cost of $\beta(\text{-}B)$, and the remaining $(D - (i + B))$ items are lost, each at a cost of c_l. To summarize, when $D > i$, then i items are immediately satisfied, $\min(D - i, B)$ items will be satisfied later, and the remaining $\max(D - i - B, 0)$ items are lost (when $B = 0$, then i items are satisfied and $D - i$ items are lost). Accordingly, when a positive batch of R items arrives, if $R \leq S - i$, then the stock level increases by $i + R$ items; otherwise, if $R > S - i$, the stock increases to level S, and the remaining $R - (S - i)$ items are transferred to a secondary storage facility at a cost of $\gamma(R - (S - i))$.

4. Mathematical Description

4.1. The State Diagram of the Markov Chain

Let $\mathbb{I} = \{I(t) : t \geq 0\}$ denote the stock level at time t and state space Ω. Based on the exponential assumptions and the compound Poisson processes, it is clear that the stock process $\mathbb{I}$ is a continuous-time Markov chain (CTMC). Assume that $I(t) = i$ items are held in stock. When $i \in \{s+1, \ldots, S\}$, two different scenarios should be considered according to the replenishment placement: (1) with oncoming replenishment, when the stock increases from level s or below to level i by a positive batch, or (2) without oncoming replenishment, when the stock decreases from levels $i+1, \ldots, S$ to level i by a negative batch, or increases from levels $s+1 \ldots i-1$ to level i by a positive batch. By contrast, when $i \in \{\text{-}B, \ldots, s\}$, only the case with oncoming replenishment holds. Thus, systems with i items are split into two types, according to their replenishment placement. To distinguish between the two types, we mark states with oncoming replenishment by a superscript plus, and states without replenishment by a superscript minus. That is, the state space Ω is composed of two sets: *(i)* the set $\Omega^+ = \{\text{-}B, \ldots, S\}$ that includes states during lead time; *(ii)* the set marked by a minus $\Omega^- = \{(s+1)^-, (s+2)^-, \ldots, S^-\}$ that includes states without replenishment. The state space is, thus, Ω=$\Omega^+ \cup \Omega^-$. Figure 4 presents the state diagram of the Markov chain $\mathbb{I}$. For clarification, the states with oncoming replenishment are shaded in yellow (without a shortage) or red (with a shortage), and states without replenishment are shaded in in gray; the transition rates are marked by arrows.

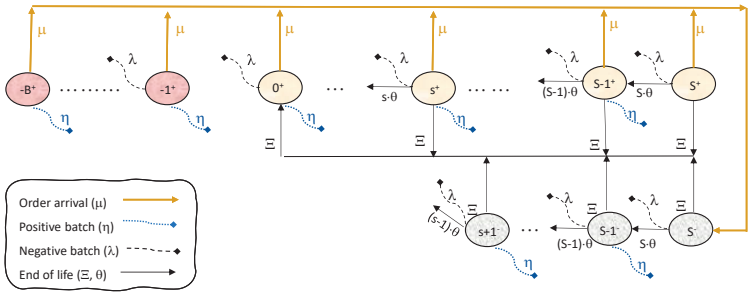


Figure 4. The state diagram of the Markov chain.

Let

$$\begin{aligned} P^+(i,j,t) &= P(I(t) = j \mid I(0) = i), i \in \Omega, j \in \Omega^+, \\ P^-(i,j,t) &= P(I(t) = j \mid I(0) = i), i \in \Omega, j \in \Omega^-. \end{aligned} \tag{1}$$

Let $\pi_j^+ = \lim_{t\to\infty} P(i,j,t)$, and $\pi_j^- = \lim_{t\to\infty} P^-(i,j,t)$. Then, π_j^+ $\left(\pi_j^-\right)$ denotes the steady-state probability of having $j, j \in \Omega^+$ $(j \in \Omega^-)$ items in stock with (without) oncoming replenishment. Applying Markov theory, we obtain that the steady-state distribution of the stock level exists as the state space is finite, and the Markov chain is irreducible. Denote by $\lambda_j = \lambda \cdot p_D(j), j = 1, 2, \ldots, \infty$ the arrival rate of a positive batch with j items, and by $\eta_j = \eta \cdot p_R(j), j = 1, 2, \ldots, \infty$ the arrival rate of a negative batch with j items. By the law of total probability, we have:

$$\lambda = \sum_{j=1}^{\infty} \lambda_j, \quad \eta = \sum_{j=1}^{\infty} \eta_j. \tag{2}$$

In addition, denote by $\widehat{F}_D(k) = 1 - F_D(k) = p(D > k)$ the probability of a negative batch being larger than k items, and by $\widehat{F}_R(k) = 1 - F_R(k) = 1 - P(R > k)$ the probability of a negative batch being larger than k items.

Claim 1. *Assume that $s > 0$ and $B > 0$. It is easy to verify that the steady-state probabilities $\left\{\pi_i^+, \pi_{i'}^-\right\}_{i\in\Omega^+, i'\in\Omega^-}$ for the stock process $\mathbb{I}$ satisfy the following balance equations.*
States during lead time:

$$(i)\ \pi_S^+(\lambda+S\cdot\theta+\Xi+\mu)=\eta\sum_{k=-B}^{S-1}\pi_k^+\cdot\widehat{F}_R(S-k-1), \qquad i=S,$$

$$(ii)\ \pi_i^+(\lambda+i\cdot\theta+\Xi+\eta+\mu)=(i+1)\cdot\theta\pi_{i+1}^+ +\sum_{k=i+1}^{S}\pi_k^+\cdot\lambda_{k-i}+\sum_{k=-B}^{i-1}\pi_k^+\cdot\eta_{i-k}, \qquad s+1\le i\le S-1,$$

$$(iii)\ \pi_s^+(\lambda+s\cdot\theta+\Xi+\eta+\mu)=\left(\pi_{s+1}^+ +\pi_{s+1}^-\right)(s+1)\cdot\theta+\sum_{k=s+1}^{S}\left(\pi_k^+ +\pi_k^-\right)\cdot\lambda_{k-s}+\sum_{k=-B}^{s-1}\pi_k^+\cdot\eta_{s-k}, \qquad i=s,$$

$$(iv)\ \pi_i^+(\lambda+i\cdot\theta+\Xi+\eta+\mu)=(i+1)\cdot\theta\pi_{i+1}^+ +\sum_{k=i+1}^{S}\pi_k^+\cdot\lambda_{k-i}+\sum_{k=s+1}^{S}\pi_k^-\cdot\lambda_{k-i}+\sum_{k=-B}^{i-1}\pi_k^+\cdot\eta_{i-k}, \qquad 0<i<s,$$

$$(v)\ \pi_0^+(\lambda+\eta+\mu)=\pi_1^+\cdot\theta+\sum_{k=1}^{S}\pi_k^+\cdot\lambda_k+\sum_{k=s+1}^{S}\pi_k^-\cdot\lambda_k+\sum_{k=-B}^{-1}\pi_k^+\cdot\eta_{-k}+\Xi\cdot\left(\sum_{k=1}^{S}\pi_k^+ +\sum_{k=s+1}^{S}\pi_k^-\right), \qquad i=0,$$

$$(vi)\ \pi_i^+(\lambda+\eta+\mu)=\sum_{k=i+1}^{S}\pi_k^+\cdot\lambda_{k-i}+\sum_{k=s+1}^{S}\pi_k^-\cdot\lambda_{k-i}+\sum_{k=-B}^{i-1}\pi_k^+\cdot\eta_{i-k}, \qquad -B<i<0,$$

$$(vii)\ \pi_{-B}^+(\eta+\mu)=\lambda\left[\sum_{k=-B+1}^{S}\pi_k^+\cdot\widehat{F}_D(k+B-1)+\sum_{k=s+1}^{S}\pi_k^-\cdot\widehat{F}_D(k+B-1)\right], \qquad i=-B.$$

States without replenishment:

$$(viii)\ \pi_S^-(\lambda+S\cdot\theta+\Xi)=\eta\sum_{k=s+1}^{S-1}\pi_k^-\cdot\widehat{F}_R(S-k-1)+\mu\sum_{k=-B}^{S}\pi_k^+, \qquad i=S,$$

$$(ix)\ \pi_i^-(\lambda+i\cdot\theta+\Xi+\eta)=(i+1)\cdot\theta\cdot\pi_{i+1}^- +\sum_{k=i+1}^{S}\pi_k^-\cdot\lambda_{k-i}+\sum_{k=s+1}^{i-1}\pi_k^-\cdot\eta_{i-k}, \qquad s+1<i\le S-1, \tag{3}$$

$$(x)\ \pi_{s+1}^-(\lambda+(s+1)\cdot\theta+\Xi+\eta)=(s+2)\cdot\theta\cdot\pi_{s+2}^- +\sum_{k=s+2}^{S}\pi_k^-\cdot\lambda_{k-s-1} \qquad i=s+1.$$

When $s=0$ and $B>0$, equations (iii) and (iv) of Claim 1 are omitted. When $s>0$ and $B=0$ (i.e., the pure lost-sales policy), then $\Omega^+=\{0,\ldots,S\}$, $\Omega^-=\{(s+1)^-,\ldots,S^-\}$ and Equations (v)–(vii) of Claim 1 are replaced by

$$\begin{aligned}\pi_0^+(\eta+\mu)=&\lambda\left[\sum_{k=1}^{S}\pi_k^+\cdot\widehat{F}_D(k-1)+\sum_{k=s+1}^{S}\pi_k^-\cdot\widehat{F}_D(k-1)\right]\\&+\pi_1^+\theta+\Xi\cdot\left(\sum_{k=1}^{S}\pi_k^+ +\sum_{k=s+1}^{S}\pi_k^-\right), \quad i=B=0.\end{aligned} \tag{4}$$

Otherwise, if $s=B=0$, then $\Omega^+=\{0,\ldots,S\}$, $\Omega^-=\{1^-,\ldots,S^-\}$, and Equations (iii)–(vii) of Claim 1 are replaced by Equation (4).

4.2. The Expected Total Cost

The cost structure includes a fixed cost and a variable cost for each replenishment, a cost for handling each arriving item of a positive batch, a cost for transferring each item exceeding level S, a lost cost for each unsatisfied item, a cost for each expired item, a cost for a total collapse of the stock, and a cost for maintaining the stock (i.e., a holding cost for each on-hand item and a backordering cost for each backlogged item, if allowed). Our aim is to minimize the average total cost TC per time unit. To that end, denote by I^+ the

average number of on-hand (positive) items in stock per time unit, and by I^- the average number of backlogged items per time unit. Formally,

$$I^+ = \sum_{i=1}^{S} i \cdot \pi_i^+ + \sum_{i=s+1}^{S} i \cdot \pi_i^-, \quad I^- = -\sum_{i=-B}^{-1} i \cdot \pi_i^+. \tag{5}$$

(Recall that, when $B = 0$, backordering is not allowed, and thus, $I^- = 0$.)

Applying Markov theory, the long-run expected cost components can be obtained in terms of the steady-state probabilities, and are given by

$$\begin{aligned} TC = & \underbrace{\sum_{i=-B}^{S} \mu \pi_i^+ \cdot (K_o + c_o(S-i))}_{\text{Replenishment cost}} + \underbrace{c_r \eta \cdot E(R)}_{\text{Overhead cost}} + \underbrace{\sum_{i=1}^{S} h(i) \cdot \pi_i^+ + \sum_{i=s+1}^{S} h(i) \cdot \pi_i^-}_{\text{Holding cost}} + \underbrace{\sum_{i=-B}^{-1} \beta(i) \cdot \pi_i^+}_{\text{Backordering cost}} \\ & + \underbrace{\sum_{i=-B}^{S} \sum_{k=S+1-i}^{\infty} \eta_k \cdot \gamma(i+k-S) \cdot \pi_i^+ + \sum_{i=s+1}^{S} \sum_{k=S+1-i}^{\infty} \eta_k \cdot \gamma(i+k-S) \cdot \pi_i^-}_{\text{Transfer cost}} \\ & + \underbrace{(c_\theta \theta + c_\epsilon \Xi) \cdot I^+}_{\text{End-of-life costs}} + \underbrace{c_l \left(\sum_{i=-B}^{S} \sum_{k=i+B+1}^{\infty} \lambda_k \cdot (k-i-B) \cdot \pi_i^+ + \sum_{i=s+1}^{S} \sum_{k=i+B+1}^{\infty} \lambda_k \cdot (k-i-B) \cdot \pi_i^- \right)}_{\text{Lost-sales cost}}. \end{aligned} \tag{6}$$

We emphasize that, although the costs incurred by random shelf life and total loss have similar expressions, they are derived in different paths. To be specific, when $i > 0$ items are held in stock during the lead time, each of these items has an exponentially distributed shelf life at rate θ. By the exponential properties, the time until the first item expires is exponential at rate $i \cdot \theta$. The cost incurred by random shelf life is, thus, $\left(c_\theta \cdot (i \cdot \theta) \cdot \pi_i^+\right)$ (here, the multiplication by i is due to the *rate* of the minimum of exponentially distributed random variables). By contrast, when a total loss event occurs at rate Ξ, all i items in stock become useless, each at cost c_ϵ. Thus, the loss cost is $\left((i \cdot c_\epsilon) \cdot \Xi \cdot \pi_i^+\right)$ (here, the multiplication by i is due to the *total* cost of the i items). Similar calculations lead to the case without oncoming replenishment.

Remark 1. ***1. Constant costs.*** *We assume that $h(i) = c_h \cdot i, i > 0$. It is easy to verify that the holding cost can be written as $\sum_{i=1}^{S} h(i) \cdot \pi_i^+ + \sum_{i=s+1}^{S} h(i) \cdot \pi_i^- = c_h \cdot I^+$. Similarly, if we assume that the backordering cost has the structure $\beta(i) = -c_\beta \cdot i, i < 0$, i.e., the backordering cost is constant per item, then it can be written as $\sum_{i=-B}^{-1} \beta(i) \cdot \pi_i^+ = c_\beta \cdot I^-$.*

2. State-dependent rates. *Although we assume fixed rates λ and η, it is easy to generalize the model to support state-dependent rates $\lambda(i)$ and $\eta(i)$. In this case, we make the following adjustments:*

- *(a) Replace λ by $\lambda(i)$ on the left-hand side of Claim 1 (1–6, 8–10). Similarly, insert λ into the sum in Claim 1 (7), and replace λ by $\lambda(k)$.*
- *(b) Replace $\pi_i^{\pm} \cdot \lambda_k$ by $\pi_i^{\pm} \cdot \lambda_k(i)$ and $\pi_i^{\pm} \eta_k$ by $\pi_i^{\pm} \cdot \eta_k(i)$ in Claim 1 (1–6, 8–10) and in Equation (6).*
- *(c) The overhead cost $c_r \cdot \eta \cdot E(R)$ in Equation (6) becomes $c_r \cdot \overline{\eta} \cdot E(R)$, where $\overline{\eta}$ is the average size of a positive batch per time unit, i.e., $\overline{\eta} = \left(\sum_{n=-B}^{S} \eta(i) \cdot \pi_i^+ + \sum_{n=s+1}^{S} \eta(i) \cdot \pi_i^- \right)$.*

5. Numerical Examples

In this section, we study, numerically, the sensitivity of the optimal controllers to the system's parameters and costs; we will focus mainly on the effect of the random arrivals

(rate, mean, and variance) and the lost vs. backlogged costs on the pure lost-sales and backordering policies. We start with the pure lost-sales policy (Section 5.1). Here, we perform a hierarchical two-step sensitivity analysis. In the first step, we assume a fixed batch size (for all arrivals); then, in the second step, we extend the analysis to include batch size variability and, similarly, for the corresponding backordering policy (Section 5.2). The analysis is completed by comparing the policies and identifying situations where one policy is more cost effective than another. We note that, although our mathematical analysis does not assume any restrictions on the ratio between η and λ, we nevertheless see that, typically, the output rate exceeds the input rate. Thus, we focus on the cases where $\lambda \geq \eta$.

5.1. Pure Lost-Sales Policy

Under the pure lost-sales policy, any unsatisfied item is lost. Setting $B = 0$ in (3)–(6) leads to the steady-state probabilities and the corresponding costs. We start with constant positive and negative batch sizes.

Fixed batch size:

In order to study the effect of the system's parameters on the optimal control parameters, we start with constant batch sizes $D, R = \{1, 2, 3\}$; i.e., $p_D(D) = 1$ and 0 otherwise, and $p_R(R) = 1$ and 0 otherwise. Here, $\lambda = \lambda_D$, and $\eta = \eta_R$. The rate λ varies in $\{5, 7.5, 10\}$, and we fix $\eta = 5$; thus, the percentage of positive arrivals is $\eta/(\eta + \lambda) = \{50\%, 40\%, 33.3\%\}$, respectively. The total loss rate is set to $\Xi = 0.025$; the perishability rate is $\theta = 0.1$; the lead time μ varies in $\{0.05, 0.1\}$. We assume that

$$K_o = 50,\ c_o = 2.5,\ c_\theta = 1,\ c_\epsilon = 1,\ c_r = 0.5,\ c_h = 1. \tag{7}$$

We further let $c_l \in \{10, 25, 50\}$, and $\gamma(i) = 10 + i$ (i.e., the transfer cost includes a fixed component Y = 10 and a cost per each transferred item $c_\gamma = 1$). Our aim is to derive the optimal S^* and s^* minimizing the total cost as given in Equation (6). Clearly, the optimal control parameters depend on all parameters and costs. However, we focus mainly on the impact of the arrival rate, lead time, batch size, and penalty cost. Accordingly, Table 1 presents the optimal S^*, s^* and the total cost $TC^*(S^*, s^*, B = 0)$ as functions of λ, μ, and c_l for the batch size combinations $(D, R) = \{(1,1), (2,1), (2,2), (3,1), (3,2),$ and $(3,3)\}$. We note that we assume $D \geq R$ since, when dealing with stock management, it only makes sense that the average amount of outflows exceeds the average amount of inflows.

Table 1. (S^*, s^*) and TC^* as functions of $\lambda,\ \mu,\ c_l$ for different combinations of (D, R).

$B = 0$		D, R											
S^*, s^*		D = 1, R = 1		D = 2, R = 1		D = 2, R = 2		D = 3, R = 1		D = 3, R = 2		D = 3, R = 3	
TC^*		μ = 0.05	μ = 0.1	μ = 0.05	μ = 0.1	μ = 0.05	μ = 0.1	μ = 0.05	μ = 0.1	μ = 0.05	μ = 0.1	μ = 0.05	μ = 0.1
λ = 5	c_l = 10	15.0 15.91	15.0 17.80	27.0 53.26	27.0 53.08	25.0 28.43	24.0 30.68	49.14 98.97	47.12 96.17	34.0 60.66	34.0 61.08	34.0 40.53	33.0 42.98
	c_l = 25	18.0 23.39	18.0 24.2	63.32 113.55	56.26 102.39	33.0 43.89	32.0 44.27	117.76 216.2	103.64 193.35	71.34 125.72	64.28 114.99	47.0 64.20	45.0 64.06
	c_l = 50	23.2 35.17	22.2 34.03	108.74 200.66	90.65 167.32	43.9 67.44	40.8 63.96	206.160 385.57	167.124 316.61	119.79 219.77	100.62 185.82	62.17 99.53	58.15 93.53
λ = 7.5	c_l = 10	17.0 30.42	17.0 31.41	48.14 97.97	46.13 94.98	31.0 58.28	31.0 58.47	80.38 167.54	77.33 160.49	62.20 124.67	59.18 120.77	46.3 86.09	45.2 85.42
	c_l = 25	35.12 62.21	32.10 57.66	117.77 214.53	102.63 189.45	68.33 121.86	62.27 110.95	201.145 370.13	173.130 324.36	148.101 272.19	130.84 240.63	101.56 181.43	91.47 164.04
	c_l = 50	59.33 108.22	50.25 92.28	204.135 383.02	165.122 313.79	115.77 213.72	97.60 179.92	353.280 662.84	284.228 540.03	258.200 485.34	210.159 398.37	172.123 319.16	143.97 267.39
λ = 10	c_l = 10	26.0 51.66	26.0 51.25	70.30 143.79	66.27 137.98	50.13 100.89	49.11 98.21	114.63 236.77	197.57 225.51	94.45 192.53	88.40 184.32	74.27 150.05	71.24 145.00
	c_l = 25	60. 32 110.94	54.26 99.23	172.12 317.32	149.103 278.29	120.78 219.36	105.64 194.55	286.219 525.14	250.190 458.48	231.174 424.95	199.144 372.40	179.126 327.73	156.104 289.43
	c_l = 50	105.72 196.59	87.55 163.03	298.247 568.06	243.192 462.98	209.160 390.52	170.130 321.33	392.367 962.26	350.319 767.24	360.316 762.18	326.264 619.97	312.252 584.41	253.196 479.47

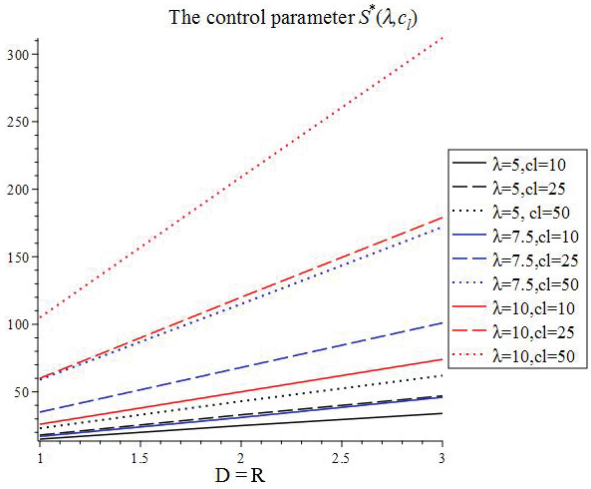
We see that, as expected, S^*, s^* and TC^* are increasing in λ, c_l, and D and decreasing in μ; clearly, as demand increases (rate and size) and as the costs of the shortage increase (i.e., c_l increases), the control parameters and the total cost increase. In addition, the effect of μ is significant as a function of c_l; that is, for high values of c_l, shortening the lead time

causes a more significant decrease in S^* and s^* (e.g., for $c_l = 10$, the results for $\mu = 0.05$ and $\mu = 0.1$ are almost the same, in contrast to the results for $c_l = 25, 50$). Table 1 further shows that the impact of the arriving batches on the system's performance is sharply pronounced. To emphasize this impact, we distinguish between batch size (D, R), where $D = R$, and batch size (D, R), where $D > R$.

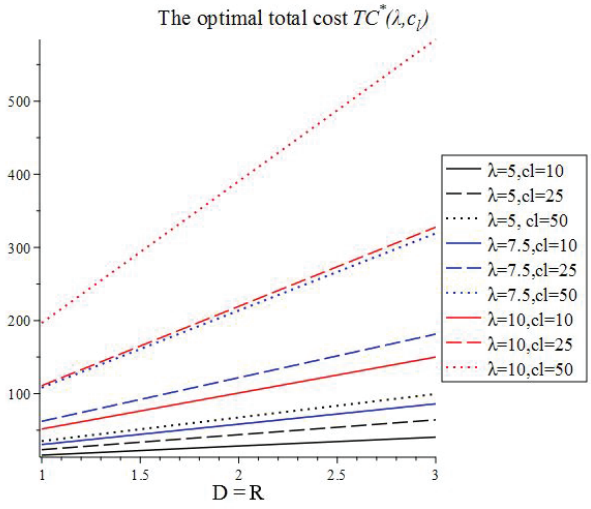
We start with batch size (D, R), where $D = R$. Figure 5a,b show the curves S^* and TC^*, respectively, as functions of λ and c_l, where $(D, R) = \{(1, 1), (2, 2), (3, 3)\}$. Similar behavior is observed for s^*. The black, blue, and red lines refer to $\lambda = \{5, 7.5, 10\}$, respectively. The solid, dashed, and dotted lines refer to $c_l = \{10, 25, 50\}$, respectively.

Figure 5 (and the results for s^*) imply that S^*, s^*, and TC^* have a similar pattern: the growth rate is increasing in λ and c_l; we see that the slopes of the red ($\lambda = 10$) and dotted ($c_l = 50$) lines increase more sharply than the other lines. Somewhat surprisingly, it further seems that S^* and TC^* are almost linear as a function of $D = R$. That is, when $D = R = k$, $k = 2, 3 \ldots$, the control parameter S^* and the total cost TC^* may be approximated by k times the results for the unit batch size; i.e.,

$$\begin{aligned} S^*(D = R = k) &\approx k \cdot S^*(D = R = 1), \\ TC^*(D = R = k) &\approx k \cdot TC^*(D = R = 1). \end{aligned} \tag{8}$$



(**a**)



(**b**)

Figure 5. The optimal S^* (**a**) and TC^* (**b**) as functions of λ, c_l for $D = R = \{1, 2, 3\}$.

(This approximation barely works for s^*.) To explore the implications of this result, let $\%error$ be the percentage of the error obtained by using (8), i.e.,

$$\%error_S = \frac{S^*(D=R=k) - k \cdot S^*(D=R=1)}{S^*(D=R=k)},$$
$$\%error_{TC} = \frac{TC(D=R=k) - k \cdot TC^*(D=R=)}{TC^*(D=R=k)}, k = 2,3. \quad (9)$$

Table 2 tabulates $\%error_S, \%error_{TC}$ for $\mu = 0.05, \lambda = \{5, 7.5, 10\}$ and $c_l = \{10, 25, 50\}$. We see that, approximating S^* and TC^* by using (8) is most effective for high λ and c_l. It should be noted that a similar approximation for s^* performs significantly worse and yields lower values. Thus, we recommend that (8) be applied as an upper bound for S^* and TC^*, especially for high λ and c_l and low k.

Table 2. The percentage $\%error_S, \%error_{TC}$ for $\lambda = \{5, 7.5, 10\}$ and $c_l = \{10, 25, 50\}$.

		$k=2$		$k=3$	
		$\%error_S$	$\%error_{TC}$	$\%error_S$	$\%error_{TC}$
$\lambda = 5$	$c_l = 10$	-20%	-11.9%	-32.3%	-17.7%
	$c_l = 25$	-9.0%	-6.58%	-14.89%	-9.29%
	$c_l = 50$	-6.97%	-4.30%	-11.29%	-6.0%
$\lambda = 7.5$	$c_l = 10$	-9.67%	-4.39%	-10.86%	-6.0%
	$c_l = 25$	-2.94%	-2.12%	-3.96%	-2.86%
	$c_l = 50$	-2.6%	-1.27%	-2.9%	-1.7%
$\lambda = 10$	$c_l = 10$	-4%	-2.4%	-5.4%	-3.28%
	$c_l = 25$	0%	-1.14%	-0.55%	-1.55%
	$c_l = 50$	-0.47%	-0.68%	-0.96%	-0.91%

We now proceed to batch size (D, R), where $D > R$. Here, increasing R while keeping D fixed decreases the control parameters; probably, since the stock level increases only with positive batches, the risk of a shortage decreases, and thus, S^* and s^* are set lower. In order to delve deeper into the impact of the inflows, we fix $\lambda = 7.5$ and $D = 3$. Figure 6 represents S^* (big circles) and s^* (small circles) for $R \in \{1, 2, 3\}$ and $c_l \in \{10, 25, 50\}$. The figure clearly shows the effect of R on lowering S^* and s^*, and the effect of c_l on increasing S^* and s^* (with a slight increase in the difference $(S^* - s^*)$). This means that, as the cost of the shortage increases, it is profitable to order earlier and in larger quantities. To complete the study, Figure 7 shows the curves of TC^* as a function of R, for $D = \{2, 3\}$ and $c_l = \{10, 25, 50\}$. Accordingly, we see that TC^* is increasing in c_l, and decreasing in R; the slope of that decrease seems to be independent of D (the solid, dashed, and dotted lines seem to be parallel). Figure 7 further shows that the slope of the decrease of TC^* as a function of R becomes sharper as c_l increases, which emphasizes the impact of inflows on the total cost and the need to integrate them when deriving the optimal control parameters, especially when shortages are costly.

To complete our study of the effect of the system's parameters on the optimal controllers when the batch size is fixed, we compare systems that have the same average number of arriving batches. Specifically, we focus on systems with the same λD and ηR. Our numerical study includes three such combinations of (D, R):

(*i*) Systems with $(\lambda = 5, \eta = 5, D = 2, R = 1)$ and $(\lambda = 10, \eta = 5, D = 1, R = 1)$, where $\lambda D = 10$ and $\eta R = 5$ (the submatrices shaded in blue in Table 1);
(*ii*) Systems with $(\lambda = 7.5, \eta = 5, D = 2, R = 1)$ and $(\lambda = 5, \eta = 5, D = 3, R = 1)$, where $\lambda D = 15$ and $\eta R = 5$ (the submatrices shaded in red in Table 1);
(*iii*) Systems with $(\lambda = 7.5, \eta = 5, D = 2, R = 2)$ and $(\lambda = 5, \eta = 5, D = 3, R = 2)$, where $\lambda D = 15$ and $\eta R = 10$ (the submatrices shaded in olive in Table 1).

The results show that, overall, systems with the same λD and ηR have similar optimal control parameters and total cost. However, it should be noted that the values are slightly more sensitive to changes in the batch size than to changes in the arrival rate. For

example, comparing the systems (i) shows that (recall that increasing λ and D increases the controllers and cost), although both have $\lambda D = 10$, the system with $\lambda = 5$, $D = 2$ yields slightly higher values ($S^* = 27, s^* = 0, TC^* = 53.26$) compared to the system with $\lambda = 10, D = 1$ (here, we obtain $S^* = 26, s^* = 0, TC^* = 51.66$). Thus, decreasing D is more impactful than increasing λ.

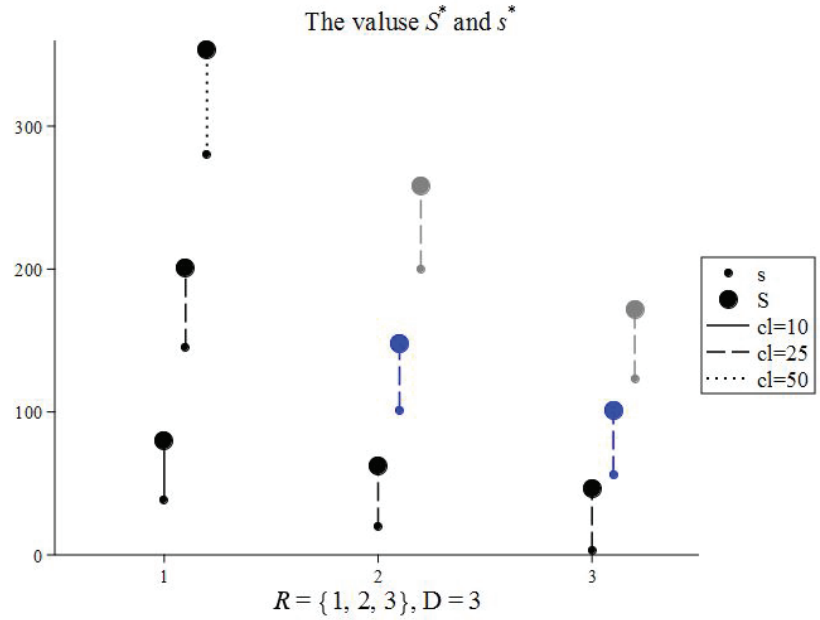


Figure 6. S^*, s^* for $D = 3$, $R = \{1, 2, 3\}$, $c_l = \{10, 25, 50\}$, $\lambda = 7.5$.

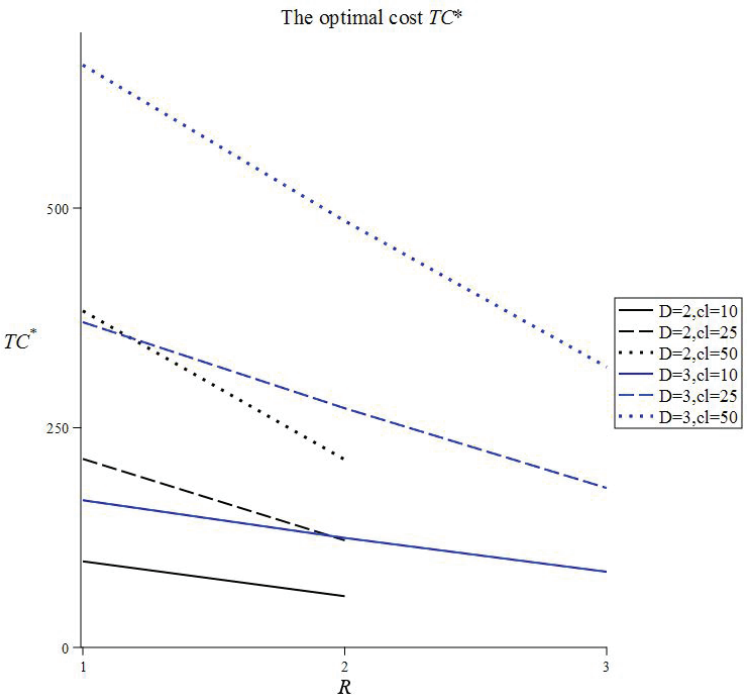


Figure 7. TC^* as a function of R, for $D = \{2, 3\}$, $c_l = \{10, 25, 50\}$.

Variable batch size:

Next, we allow D and R to vary. We fix $E(D) = 3$ and $E(R) = 2$, and assume the following four distributions of D with coefficient of variation $cv_D = \{0, 0.66, 1.15, 2.21\}$ (where $E(D) = 3$):

$$(i)\ \begin{matrix} D=3 \\ (cv=0) \end{matrix} \quad (ii)\ \begin{matrix} R=\left\{ \begin{matrix} 1 & 0.5 \\ 5 & 0.5 \end{matrix} \right\} \\ (cv=0.66) \end{matrix} \quad (iii)\ \begin{matrix} R=\left\{ \begin{matrix} 1 & 0.75 \\ 9 & 02.5 \end{matrix} \right\} \\ (cv=1.15) \end{matrix} \quad (iv)\ \begin{matrix} D=\left\{ \begin{matrix} 1 & 8/9 \\ 25 & 1/9 \end{matrix} \right\} \\ (cv=2.21) \end{matrix} \tag{10}$$

and five distributions of R with coefficient of variation $cv_R = \{0, 0.5, 0.86, 1.41, 2.39\}$ (where $E(R) = 2$):

$$(i)\ \begin{matrix} R=2 \\ (cv=0) \end{matrix} \quad (ii)\ \begin{matrix} R=\left\{ \begin{matrix} 1 & 0.5 \\ 3 & 0.5 \end{matrix} \right\} \\ (cv=0.5) \end{matrix} \quad (iii)\ \begin{matrix} R=\left\{ \begin{matrix} 1 & 0.75 \\ 5 & 02.5 \end{matrix} \right\} \\ (cv=0.86) \end{matrix} \quad (iv)\ \begin{matrix} R=\left\{ \begin{matrix} 1 & 8/9 \\ 10 & 1/9 \end{matrix} \right\} \\ (cv=1.41) \end{matrix}$$

$$(v)\ \begin{matrix} R=\left\{ \begin{matrix} 1 & 23/24 \\ 25 & 1/24 \end{matrix} \right\}. \\ (cv=2.39) \end{matrix} \tag{11}$$

Table 3 tabulates (S^*, s^*) and TC^* for each pair of distributions $D \in \{i, \ldots, iv\}$ and $R \in \{i, \ldots, v\}$ (for a total of 20 pairs). The table includes four submatrices, each corresponding to a different combination of $\lambda \in \{5, 7.5\}$ and $c_l \in \{10, 20\}$. Here, we fix $\mu = 0.05$; other values are given in (7). Note that the shaded cells with $cv_D = cv_R = 0$ correspond to the case of $D = 3, R = 2$, as given in Table 1.

Table 3. (S^*, s^*) and TC^* as functions of cv_D, cv_R (where $E(D) = 3,\ E(R) = 2$), for different values of $\lambda,\ c_l$.

S^*, s^* / TC^*		E(R) = 2				
E(D) = 3		***(i) cv_R = 0***	***(ii) cv_R = 0.5***	***(iii) cv_R = 0.86***	***(iv) cv_R = 1.41***	***(v) cv_R = 2.39***
λ = 5, c_l = 10	***(i) cv_D = 0***	34.0 60.66	34.0 61.35	36.0 63.04	40.0 68.95	51.0 74.61
	(ii) cv_D = 0.66	36.0 63.29	37.0 63.93	38.0 65.21	42.0 70.85	52.0 76.21
	(iii) cv_D = 1.15	40.0 67.82	40.0 68.35	42.0 69.40	44.0 74.10	54.0 78.97
	(iv) cv_D = 2.21	51.0 81.14	51.0 81.50	52.0 82.17	52.0 85.30	57.0 87.45
λ = 5, c_l = 25	***(i) cv_D = 0***	71.34 125.72	71.34 126.91	73.35 129.64	80.40 142.46	85.39 148.63
	(ii) cv_D = 0.66	73.34 130.03	74.35 131.15	75.35 133.31	82.40 145.65	88.40 151.34
	(iii) cv_D = 1.15	78.35 137.50	79.36 138.79	80.36 140.03	85.42 151.23	94.41 156.12
	(iv) cv_D = 2.21	90.40 159.81	90.40 160.56	91.40 161.88	96.45 170.41	99.43 171.32
λ = 7.5, c_l = 10	***(i) cv_D = 0***	62.20 124.67	62.20 125.00	63.20 125.99	66.21 131.93	70.19 133.55
	(ii) cv_D = 0.66	63.19 126.25	63.19 126.60	63.19 127.29	66.21 133.11	71.18 134.65
	(iii) cv_D = 1.15	64.17 128.94	64.17 129.29	65.17 129.94	67.19 135.14	72.17 136.54
	(iv) cv_D = 2.21	66.16 136.92	66.15 137.36	67.16 138.11	68.16 142.63	74.14 142.85
λ = 7.5, c_l = 25	***(i) cv_D = 0***	148.101 272.19	149.102 272.78	150.102 274.4	157.106 287.26	158.104 287.04
	(ii) cv_D = 0.66	150.102 274.83	150.102 275.45	151.102 276.66	158.108 289.3	159.105 288.95
	(iii) cv_D = 1.15	152.103 279.42	153.103 280.05	154.103 281.22	160.109 292.92	161.107 292.31
	(iv) cv_D = 2.21	159.107 293.57	160.107 294.35	165.11 295.7	166.112 306.27	166.11 303.77

Table 3 shows that letting D and R vary has a similar effect on S^*, s^*, and TC^*. In most cases, increasing the variability increases the controllers; in fact, it seems that S^* is more

sensitive to variability and increases faster than s^*. Surprisingly, we see that, for high cv_R, the control parameter s^* even decreases; see Column (v) of Table 3. Apparently, letting a positive batch vary has a similar effect as increasing its average size: in both cases, s^* is decreasing (as we already inferred from Table 1). This is probably to keep the stock level balanced and low and, thus, reduce costs.

Conclusion 1. *(pure lost-sales policy):*

1. *The control parameters S^*, s^* and the total cost TC^* are increasing in λ, c_l, and D, and decreasing in R and μ.*
2. *The increase of S^* and TC^* in D is relatively constant. Accordingly, when $D = R = k$, $k = 2, 3 \ldots$, approximating S^* and TC^* by k times the corresponding values of the Poisson process (i.e., the unit batch size) performs well, especially when λ and c_l are high.*
3. *Comparing systems with the same average inflow and outflow shows similarities between the controllers and cost, where the batch size has a slightly more significant effect than the arrival rate.*
4. *Increasing the variability of the batch size increases S^*, s^*, and TC^*, especially S^*. Quite surprisingly, we see that the impact of the variability of D and R on the optimal control parameters is similar.*
5. *Investigating the effect of the batch size distribution shows that, for both D and R, the average batch size has a more significant impact on the system's performance than the variability.*
6. *Integrating positive batches (with the focus on the average size) into the analysis significantly impacts the optimal control parameters, improves the system's performance, and may reduce costs.*

5.2. Backordering Policy

We next consider the policy in which the retailer allows backordering up to B items ($B > 0$) in order to improve the service level and retain customers. Each backlogged item incurs some financial compensation per time unit until it is satisfied in the next replenishment. A negative batch (or portion thereof) that exceeds level -B is lost. Here, the level -B is a lower bound for the stock level (clearly, $B = 0$ corresponds to the pure lost-sales policy). Setting $B > 0$ in (3)–(6) leads to the steady-state probabilities and costs under the backordering policy.

Clearly, the decision variable B is impacted by all parameters of the system. Accordingly, and further in Section 5.1, we focus on the impact of λ, c_l and the distribution of D and R on the level B. To do so, we assume a backordering cost $\beta(i) = 1.5 \cdot (\text{-}i)$, $i < 0$; all other values and costs are equal to those given in Table 1 and in Equation (7).

Fixed batch size:

We start with fixed batch sizes and consider the combinations $(D, R) = \{(1,1), (2,1), (2,2), (3,1), (3,2), (3,3)\}$. Let λ vary in $\{5, 7.5, 10\}$ and c_l vary in $\{10, 25, 50\}$. Here, we assume that the storage capacity is fixed to the optimal capacity of the pure lost-sales policy, i.e., $S = S^*(B = 0)$ (see Table 1). This assumption follows the reality that changing the shortage policy (from pure lost sales to backordering) usually changes the reorder point (s^*) without changing the storage capacity. For each set (λ, c_l, D, R), Table 4 presents the optimal order point s^*, the optimal backlogged level -B^*, and the cost $TC^*(S, s^*, \text{-}B^*)$ (we present the results for $\mu = 0.05$; the results for $\mu = 0.1$ vary accordingly). In the table (and similar to Table 1), systems that have the same average arriving size are shaded in the same color.

The parameters s^* and B^*. Similar to Table 1, Table 4 shows that s^* is increasing in λ, c_l, and D (for fixed R), and decreasing in R (for fixed D). We further see that B^* is increasing in c_l, and that it eventually approaches a pure backordering policy when c_l is extremely high (i.e., when $B = \infty$). However, B^* changes unexpectedly and even somewhat surprisingly. In contrast to s^*, B^* is decreasing in D and λ and increasing in R. The result is that the difference s^*-(-B^*) is increasing in D and λ. A possible explanation lies in the simultaneous vision of s^* and B^*. When D and λ are increasing and R is fixed, the stock level must be replenished fast in order to avoid shortage (i.e., s^* is increasing), so

backorders become less frequent and less profitable. Thus, B^* is increasing and the number backlogged items is decreasing. By contrast, when R is increasing and D is fixed, the stock level increases, and the risk of a shortage is already decreasing. Thus, B^* increases to allow a relatively large number of backorders while still maintaining profitability. It is interesting to see that increasing D and R simultaneously increases both s^* and B^* (see the pairs $(D,R) = \{(1,1),(2,2),(3,3)\}$). Similar to Equation (8), Table 4 shows that, when $D = R = k$, $k = 2,3$, then B^* can be approximated by k times the value of B^* for the unit batch size, i.e., $B^*(k) \approx k \cdot B^*(1)$.

Table 4. (S, s^*, B^*) and TC^* as functions of λ, c_l, for different combinations of (D,R).

S*, s*, -B* ***TC**** ***(%E_B)***		**D = 1, R = 1**	**D = 2, R = 1**	**D = 2, R = 2**	**D = 3, R = 1**	**D = 3, R = 2**	**D = 3, R = 3**
λ = 5	**c_l = 10**	15.0. −7 13.39 (15.83%)	27.0. −2 52.14 (2.10%)	25.0. −13 23.12 (18.67%)	49.13. −1 98.22 (0.75%)	34.0. −5 57.37 (5.42%)	34.0 −19 32.33 (20.23%)
	c_l = 25	18.0. −18 15.74 (32.7%)	63.30. −5 109.54 (3.53%)	33.0. −36 28.42 (35.24%)	117.76. −4 213.08 (1.44%)	71.30. −13 114.68 (8.78%)	47.0. −54 40.74 (36.40%)
	c_l = 50	23.0. −∞ 16.77 (52.30%)	108.51. −∞ 158.05 (21.23%)	43.0. −∞ 30.75 (54.40%)	206.132. −∞ 336.32 (12.77%)	119.48. −∞ 149.33 (32.05%)	62.0. −∞ 45.26 (59.09%)
λ = 7.5	**c_l = 10**	17.0. −2 29.17 (4.10%)	48.14. −1 97.47 (0.51%)	31.0. −4 55.75 (4.34%)	80.37. −1 167.15 (0.23%)	62.19. −2 123.34 (1.06%)	46.0. −6 82.29 (4.41%)
	c_l = 25	35.10. −5 57.78 (7.12%)	117. 76. −3 212.36 (1.01%)	68.28. −11 113.03 (7.24%)	201.148. −3 368.29 (0.33%)	148.98. −7 266.71 (2.01%)	101.49. −16 168.21 (7.28%)
	c_l = 50	59.21. −∞ 75.36 (30.36%)	204.132. −∞ 335.34 (12.44%)	115.47. −∞ 146.95 (31.24%)	353. 272. −∞ 621.3 (6.26%)	258.176. −∞ 433.04 (10.77%)	172.88. −∞ 218.4 (31.57%)
λ = 10	**c_l = 10**	26.0 −1 51.06 (1.16%)	70.30. −1 143.56 (0.16%)	50.12. −2 99.69 (1.18%)	114.63. −1 236.61 (0.06%)	94.45. −2 191.84 (0.35%)	74.26. −3 148.26 (1.19%)
	c_l = 25	60.31. −3 108.43 (2.26%)	172.124. −2 315.86 (0.47%)	120.75. −6 214.36 (2.27%)	286.223. −2 523.83 (0.24%)	231.172. −5 421.47 (0.81%)	179.122. −9 320.23 (2.28%)
	c_l = 50	105.50. −∞ 156.67 (20.40%)	298.221. −∞ 518.18 (8.78%)	209.130. −∞ 308.35 21.05%)	392.315. −∞ 936.05 (2.72%)	360.295. −∞ 713.24 (6.42%)	253. 145. −∞ 458.46 (21.55%)

The optimal total cost TC^*. As expected, TC^* is increasing in c_l, λ, and D (for fixed R), and decreasing in R (for fixed D). Here, we see that TC^* is more sensitive to D than to R, and thus, TC^* increases also when both D and R increase. Comparing Tables 1 and 4 reveals that $TC^*(B = 0) > TC^*(B)$ in many cases; i.e., backordering is economically viable. In order to better our understanding of these cases, let $\%E_B$ be the percentage of cost reduction obtained by applying the backordering policy instead of the lost-sales policy, i.e.,

$$\%E_B = \frac{TC^*(B = 0) - TC^*(B)}{TC^*(B = 0)}. \tag{12}$$

The percentage $\%E_B$ is indicated by the italic numbers in Table 4. We see that the backordering policy yields a significant cost reduction, in particular for high c_l and low λ (in fact, in most cases, the reduction is very significant). Clearly, the greater the cost of the shortage, the greater the economic benefit of the backordering policy. By contrast, the greater the outflow of items (by increasing $\lambda D - \eta R$), the level B^* is achieved quicker, which makes the backordering policy less efficient. For example, for $c_l = 10$, $\lambda = 5$, $D = 3$, and $R = 1$, the cost reduction is 0.75%, and when $\lambda = 10$, the cost reduction is less than 1.2%. We further see that systems that have the same average arriving size (i.e., the same λD and ηR) yield similar -B^*, s^*, and TC^* (see, respectively, the blue, red, and olive submatrices in Table 4).

Variable batch size:

Next, we allow the batch sizes to vary. To do so, we apply the distributions given in (10) for D, and those given in (11) for R (where $E(D) = 3, E(R) = 2$). For each pair (cv_D, cv_R), we set the storage capacity to the corresponding value for the lost-sales policy; i.e., we set $S = S^*_{B=0}(cv_D, cv_R)$, and derive the optimal control parameters s^*, -B^*, and total cost $TC^*(S, s^*, \text{-}B^*)$. Table 5 presents the suboptimal policy for the combinations $(\lambda, c_l) = \{(5, 10), (5, 25), (7.5, 10), (7.5, 25)\}$ as functions of cv_D and cv_R. Note that the shaded cells in Table 5 refer to $D = 3, R = 2$ (i.e., where $cv = 0$; see also Table 4). We further let $\%\widetilde{E}_B$ be the percentage of cost reduction obtained by applying the backordering policy instead of the lost-sales policy, i.e.,

$$\%\widetilde{E}_B = \frac{TC^*(B = 0, cv_D, cv_R) - TC^*(B, cv_D, cv_R)}{TC^*(B = 0, cv_D, cv_R)}. \tag{13}$$

The percentage $\%\widetilde{E}_B$ is indicated by the italic numbers in Table 5.

The optimal control parameters s^* and -B^*. Table 5 shows that s^* is less sensitive to the batch size variability; the changes are minimal (if they exist) and are not monotonic. By contrast, in most cases, B^* is significantly affected by both inflow variability (cv_R) and outflow variability (cv_D). Increasing cv (for both inflow and outflow) usually increases B^*; i.e., more backlogged items are permitted. It further seems that B^* is more sensitive to cv_D, probably due to the fact that the average outflow $\lambda E(D) = \{15, 22.5\}$ is higher than the average inflow $\eta E(R) = 10$. To better understand the effect of cv_D and cv_R on B^*, Figure 8 shows the curves B^* as a function of cv_R for two combinations: (1) lower outflow and higher lost cost, $\lambda = 5, c_l = 25$ (black lines), and (2) higher outflow and lower lost cost, $\lambda = 7.5, c_l = 10$ (blue lines). For each combination, the parameter B, for $cv_D = \{0, 0.66, 1.15, 2.21\}$, is indicated by the solid, dashed, dotted, and dashed–dotted lines, respectively. It is clearly shown that, when λ is relatively high and c_l is low, B^* is weakly affected by cv_R and by cv_D (see Figure 8, the blue lines). The low values of B^* can be explained by the fact that, when outflows are frequent, and the lost cost (i.e., the cost of losing customers) is relatively low, the economic benefit of backordering decreases. Thus, the variability of the inflows contributes to the prevention of a shortage, and further reduces the need for backordering (i.e., B^* decreases). However, when outflows are less frequent, and the lost cost is high, the economic benefit of backordering increases and B^* increases (see Figure 8, the black lines); this benefit is further emphasized by the increase in the variance of cv_R and cv_D. Note that, although only the combination $\lambda = 5, c_l = 25$ is presented, B^* varies accordingly also for other values.

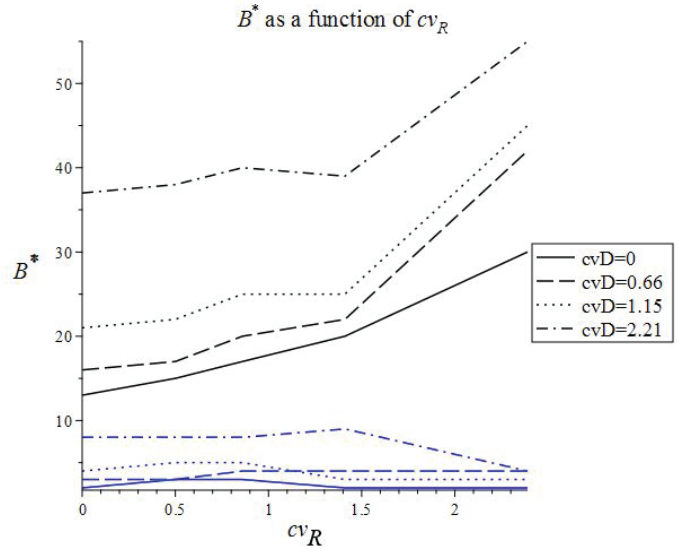


Figure 8. $B^*(cv_R)$ for (1) $(\lambda, c_l) = (5, 25)$ (black lines) and (2) $(\lambda, c_l) = (7.5, 10)$ (blue lines).

Table 5. $(S, s^*,\ -B^*)$ and TC^* as functions of cv_D, cv_R (where $E(D) = 3,\ E(R) = 2$), for different values of $\lambda,\ c_l$.

	$S^*, s^*, -B^*$ TC^* $\%\widetilde{E}_B$	E(R) = 2				
E(D) = 3		**(i) cv_R = 0**	**(ii) cv_R = 0.5**	**(iii) cv_R = 0.86**	**(iv) cv_R = 1.41**	**(v) cv_R = 2.39**
λ = 5 **c_l = 10**	**(i) cv_D = 0**	34.0, −5 57.37 (5.42%)	34.0, −6 58.03 (5.41%)	36.0, −6 59.5 (5.61%)	40.0, −7 65.98 (4.30%)	51.0, −10 71.89 (3.78%)
	(ii) cv_D = 0.66	36.0, −6 59.17 (6.50%)	37.0, −7 59.81 (6.44%)	38.0, −7 61.1 (6.30%)	42.0, −8 67.37 (4.91%)	52.0, −11 73.03 (4.17%)
	(iii) cv_D = 1.15	40.0, −9 62.4 (7.99%)	40.0, −9 62.93 (7.92%)	42.0, −10 64.13 (7.59%)	44.0, −10 69.77 (5.84%)	54.0, −12 75.19 (4.78%)
	(iv) cv_D = 2.21	51.0, −15 72.52 (10.62%)	51.0, −15 72.91 (10.53%)	52.0, −15 73.8 (10.18%)	52.0, −14 78.26 (8.25%)	57.0, −21 82.66 (5.47%)
λ = 5 **c_l = 25**	**(i) cv_D = 0**	71.30, −13 114.68 (8.78%)	71.27, −15 115.31 (9.14%)	73.27, −17 116.7 (9.98%)	80.33, −20 129.29 (9.24%)	85.27, −30 129.09 (13.14%)
	(ii) cv_D = 0.66	73.26, −16 116.65 (10.28%)	74.27, −17 117.26 (10.59%)	75.26, −20 118.55 (11.07%)	82.33, −22 130.91 (10.12%)	88.26, −42 130.22 (13.95%)
	(iii) cv_D = 1.15	78.25, −21 120.36 (12.46%)	79.25, −22 120.96 (12.84%)	80.25, −25 122.13 (12.78%)	85.31, −25 133.98 (11.40%)	94,26, −45 133.29 (14.62%)
	(iv) cv_D = 2.21	90.20, −37 132.84 (16.87%)	90.20, −38 133.32 (16.96%)	91.20, −40 134.33 (17.01%)	96.27, −39 145.01 (14.90%)	99.22, −55 143.48 (16.25%)
λ = 7.5 **c_l = 10**	**(i) cv_D = 0**	62.19, −2 123.34 (1.06%)	62.19, −3 123.74 (1.00%)	63,19, −3 124.77 (0.96%)	66.21, −2 130.95 (0.74%)	70.18, −2 132.59 (0.71%)
	(ii) cv_D = 0.66	63.18, −3 124.41 (1.45%)	63.18, −3 124.81 (1.41%)	63.18, −4 125.77 (1.19%)	66.20, −3 131.87 (0.93%)	71.17, −3 133.47 (0.87%)
	(iii) cv_D = 1.15	64.15, −4 126.19 (2.13%)	64.16, −5 126.65 (2.04%)	65.16, −5 127.5 (1.87%)	67.18, −3 133.67 (1.08%)	72.16, −3 135.07 (1.07%)
	(iv) cv_D = 2.21	66.8, −8 131.87 (3.68%)	66.8, −8 132.36 (3.64%)	67.9, −8 133.33 (3.46%)	68.12, −9 136.06 (4.60%)	74.11, −4 141.08 (1.23%)
λ = 7.5 **c_l = 25**	**(i) cv_D = 0**	148.98, −7 266.71 (2.01%)	149.99, −8 267.09 (2.08%)	150.98, −9 267.9 (2.36%)	157.104, −11 280.43 (2.37%)	158.98, −25 275.65 (3.96%)
	(ii) cv_D = 0.66	150.98, −8 267.86 (2.53%)	150.98, −9 268.23 (2.62%)	151.98, −11 269.01 (2.76%)	158.104, −13 281.49 (2.69%)	159.98, −26 276.67 (4.24%)
	(iii) cv_D = 1.15	152.98, −11 269.98 (3.37%)	153.98, −12 270.38 (3.45%)	154.98, −14 271.18 (3.57%)	160.04, −15 283.51 (3.21%)	161.98, −27 278.65 (4.67%)
	(iv) cv_D = 2.21	159.98, −19 277.58 (5.44%)	160.98, −19 278.04 (5.54%)	165.97, −21 278.98 (5.65%)	166.103, −22 290.79 (5.05%)	166.98, −31 286.33 (5.74%)

The total cost TC^*. Table 5 shows that TC^* is increasing in both cv_D and cv_R, with a more notable growth rate in cv_D. Comparing the costs $TC^*(B > 0)$ and $TC^*(B = 0)$ (as measured by $\%\widetilde{E}_B$) shows that backordering yields a significant cost reduction, especially for low λ and high c_l. To better understand the impact of randomness on this cost reduction, Figure 9 shows the curves $\%\widetilde{E}_B$ as a function of cv_R for the combinations $(\lambda, c_l) = \{(5, 25), (5, 10), (7.5, 10)\}$ as indicated by the black, red, and blue lines, respectively. Here, also, for each combination, the results for $cv_D = \{0, 0.66, 1.15, 2.21\}$ are indicated by the solid,dashed, dotted, and dashed–dotted lines, respectively. Similar to above, Table 5 and Figure 9 show that, when the lost cost is low and the benefit of backordering is in doubt, adding variability to inflows only strengthens this doubt, and therefore, the economic benefit decreases, i.e., $\%\widetilde{E}_B$ decreases (see, e.g., the red and blue lines for $c_l = 10$).

By contrast, when the lost cost is high, the financial benefit of backordering increases, and this increase is boosted by the increase in the variability of the inflows (see the black lines for $c_l = 25$). To summarize, the measure $\%\widetilde{E}_B$ is decreasing in λ, and increasing in c_l and cv_D. The impact of cv_R is more challenging as the randomness of the inflows intensifies the economic benefit of the existing policy, i.e., when $\%\widetilde{E}_B$ is low (high), increasing vc_R decreases (increases) it even more.

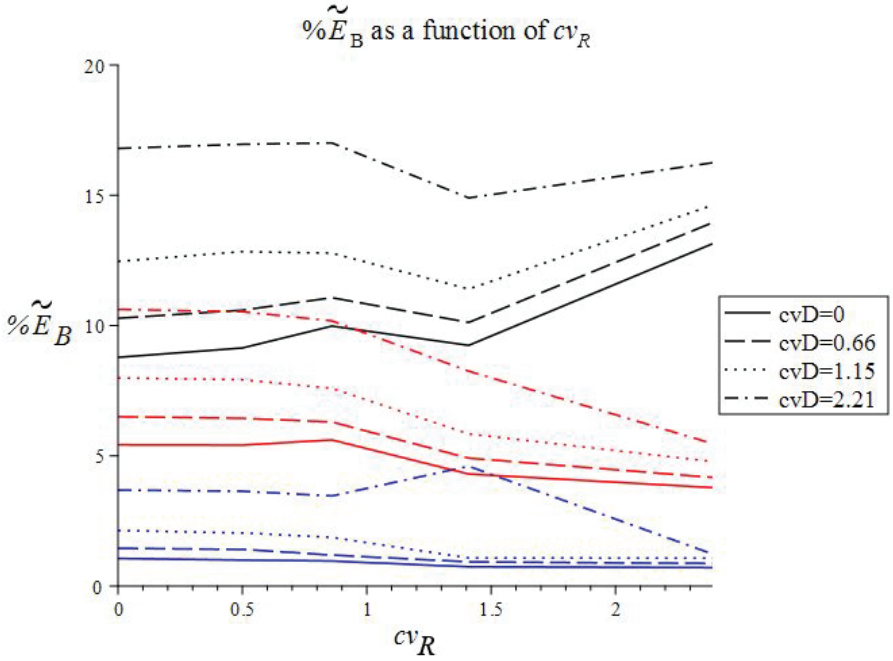


Figure 9. $\%\widetilde{E}_B(cv_R)$ for $(\lambda, c_l) = \{(5,25), (5,10), (7.5,10)\}$ as indicated by the black, red, and blue lines, respectively.

The effect of c_β. In the above discussion, we assumed that the retailer considers a backordering policy instead of a lost-sales policy. We show that backordering yields a significant cost reduction (in addition to raising the service level and maintaining customer satisfaction). To complete the discussion, it would be interesting to identify situations where the lost-sales policy is preferable. Clearly, the financial benefit of one policy or the other depends on the costs and, in particular, the interplay between the backordering cost and the lost cost, c_β and c_l, respectively. Increasing c_β (or, alternatively, decreasing c_l) decreases the economic benefit of backordering. To better understand this interplay, we assume that a backordering policy is being implemented, and that the retailer examines the policy's profitability as a function of the cost c_β. First, we fix $cv_R = 0.86$, and calculate for each set (λ, c_l, cv_D) the total cost $TC^*(S, s^*, B^*)$ as a function of c_β (the values S, s^*, and B^* are given in Table 5). For example, for $\lambda = 5, c_l = 10, cv_D = 0$, and $cv_R = 0.86$, we obtain $TC^*(S = 36, s^* = 0, \text{-}B^* = -6) = 105.420502 + 7.519728 \cdot c_\beta$. Figure 10 indicates the variable component of $TC^*(c_\beta)$ for $(\lambda, c_l) = \{(5,10), (5,25), (7.5,10), (7.5,25)\}$ by the gray, black, blue, and red lines, respectively. For each pair, the solid (dashed) line refers to $cv_D = 0$ ($cv_D = 1.15$). For example, when $(\lambda, c_l, cv_D) = (5, 10, 0)$, we obtain $7.519728 \cdot c_\beta$ (as indicated by the gray line). We see that the growth rate of $TC^*(c_\beta)$ is sharper when c_l and cv_D are high and λ is low. That is, when the lost cost is high, and fewer (but more variable) outflows arrive, the backlogged cost has a greater impact on the total cost. In particular, $TC^*(c_\beta)$ is more sensitive to changes in c_β when c_l is high (as indicated by the black and red lines for $c_l = 25$ in Figure 10).

Furthermore, let c_{β}_max be the maximum cost for which the backordering policy is more profitable compared to the lost-sales policy. In order to derive c_{β}_max, we compare the total cost $TC^*(c_\beta)$ for $B > 0$ (given in Table 5) with the corresponding $TC^*(B = 0)$ of the lost-sales policy (given in Table 3). Solving $TC^*(c_\beta) = TC^*(B = 0)$ yields c_{β}_max.

Let $cv_R = 0.86$, and let cv_D vary in $\{0, 0.66, 1.15, 2.21\}$. Figure 11 points out c_β_max for the pairs $(\lambda, c_l) = \{(5,10), (5,25), (7.5,10), (7.5,25)\}$ as indicated by the gray and black circles and the blue and red squares, respectively (i.e., each pair is assigned four values of c_β_max corresponding to cv_D). Note that, for $c_\beta > c_\beta\_max$, the backordering policy is less profitable. Figure 11 implies that the backordering policy has the highest economic benefit when λ is low and c_l is high. Similar to Table 5, Figure 11 shows that increasing cv_D increases the economic benefit of backordering, and eventually, the values c_β_max become closer.

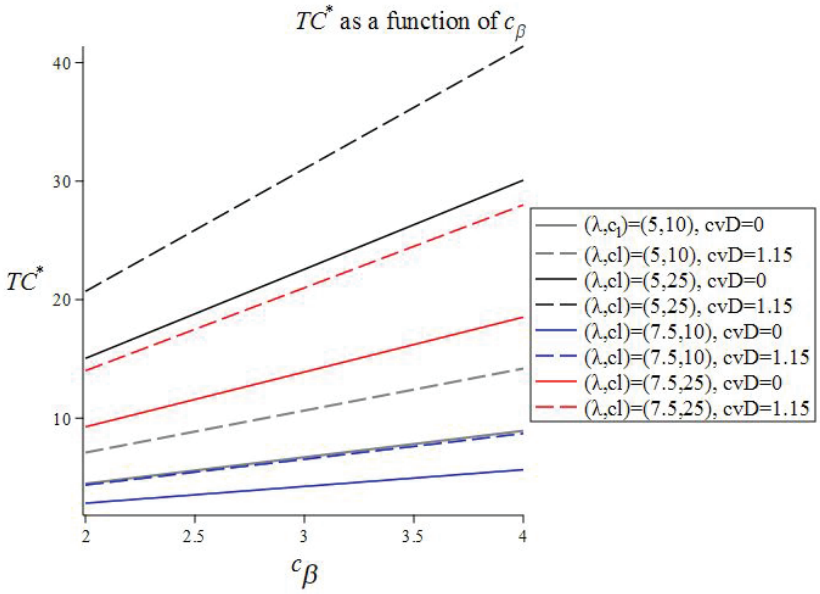


Figure 10. $TC^*(c_\beta)$ for $(\lambda, c_l) = \{(5,10), (5,25), (7.5,10), (7.5,25)\}$ as indicated by the gray, black, blue, red lines, respectively.

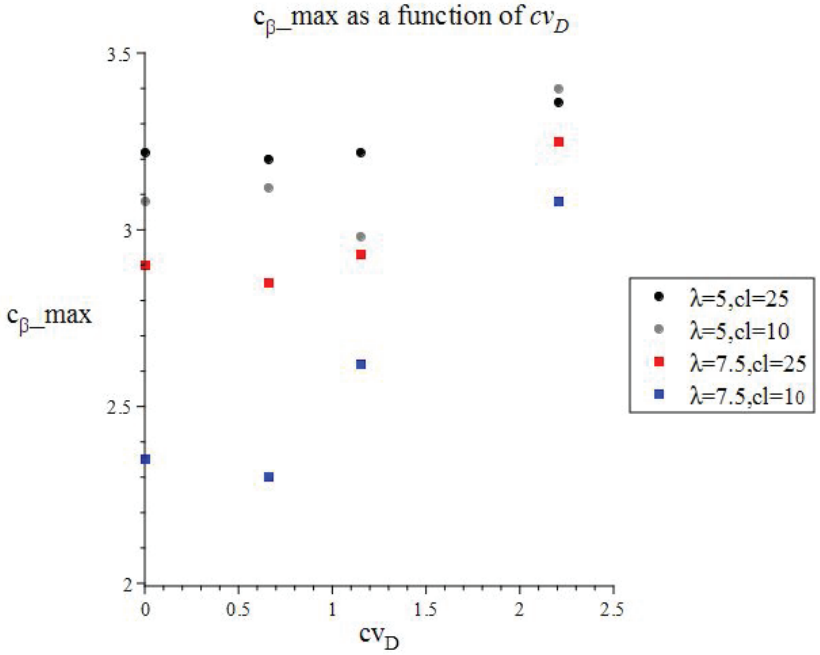


Figure 11. c_β_max for $(\lambda, c_l) = \{(5,10), (5,25), (7.5,10), (7.5,25)\}$ as indicated by the gray, black, blue, and red lines, respectively.

Conclusion 2. *(Backordering policy).*

1. *Increasing c_l and (in most cases) cv_R increases both s^* and B^*. However, the negative batch D (average and variance), the rate λ, and the average positive batch $E(R)$ have an inverse effect on s^* and B^*. The total cost TC^* is increasing in $\lambda E(D)$ and decreasing in $E(R)$; letting the batch size vary increases TC^* (except for the high cv_R).*
2. *Systems that have the same average batch size yield similar optimal control parameters, s^* and B^*, and a similar total cost TC^*.*
3. *When the batch size satisfies $D = R = k$, then B^* can be approximated by using the unit Poisson system, i.e., $B^*(k) \approx k \cdot B^*(1)$.*
4. *Comparing the backordering policy and the pure lost-sales policy shows that the former yields a significant cost reduction, particularly for high c_l and low λ. Thus, increasing the randomness of outflows increases the economic benefit of the backordering policy; however, increasing the randomness of inflows intensifies the economic benefit of the existing policy.*
5. *Accordingly, the maximum backlogged cost c_{β}_max, at which the backordering policy is economically preferable, is decreasing in λ and is increasing in c_l.*

Managerial insights:
We summarize the main managerial insights and rules of thumb of our study:

- The total cost increases with the demands (arrival rate and size) and decreases with the returns.
- The batch size distribution has a significant impact on the efficiency of the backordering policy. As a rule of thumb, we can point out that the fewer the demands (rate and quantity) and the higher the variance, the backordering policy is more recommended.
- Our study shows that integrating returns (with the focus on the average size) into the analysis significantly impacts the optimal control parameters, improves the system's performance, and may reduce costs. Specifically, when more returns arrive, we recommend increasing the backordering level (i.e., more backordered items are permitted).
- The mean batch size has a more significant impact on the system's performance than the variability. Additionally, systems that have the same average batch size yield similar optimal control parameters.

6. Concluding Remarks and Further Research

This paper studies a continuous-review triple-parameter (S, s, B) stock replenishment model for a jump process with positive and negative batch arrivals that combines backordering and lost-sales policies during a stockout. The model is characterized by random lead times, random life times (for each and all items in stock), and random arrivals (rate, mean, and variance). Using a Markovian framework, we derived the steady-state probabilities of the stock level, and constructed closed-form expressions for the average cost functions. Using a numerical analysis, we studied the impact of the different parameters, such as arrival rates, jump size distributions (mean and variance), lead times, and the costs, on the system's performance. Specifically, we focused on the impact of the arrival rates and batch sizes on the economic benefit of a backordering versus a lost-sales policy, and identified cases where one policy is more cost effective than the other. It is shown that the fewer the outflows (rate and quantity) and the higher the variance, the more economically beneficial the backordering policy. Surprisingly, more inflows increase the optimal number of backordered items and, thus, increase the economic benefit of the backordering policy.

There are several avenues for future research. In this paper, we assumed that the item's life time is an exponentially distributed random variable. A natural expansion is to assume an m-Erlang (or phase-type)-distributed life time. Under an m-Erlang distribution, the life time of each item consists of m phases, each of which is an exponentially distributed random variable. The family of m-Erlang distributions provides flexibility in model design that enables covering more cases in practice (note that as m approaches ∞, the life time becomes deterministic). Another interesting direction would be to consider an emergency supply during a stockout. This supply can be an appropriate, albeit costly, response to a

stockout under both policies, $B = 0$ and $B > 0$. Here, it would be interesting to derive the optimal control parameters under the discounted cost criterion. As in many real-world relationships, the fixed order cost is paid when the order is placed, whereas the purchasing cost is paid when the stock is replenished. This difference in the timing of the payments and the cost of the backorders clearly has a significant effect on the profitability of one policy versus the other and, thus, is worth further investigation. Finally, investigating other replenishment policies, such as the (Q, r) policy with a mixture of backordering and lost sales, is an obvious extension. Under the (Q, r) policy, when the stock drops to level r, a fixed order of size Q is placed. Thus, the stock level is no longer known at the replenishment time, which makes the analysis more challenging, but worthwhile.

Funding: This research received no external funding.

Data Availability Statement: Data are contained within the article.

Acknowledgments: We sincerely thank Michael Dreyfuss for his help in designing parts of the simulation program.

Conflicts of Interest: The author declares no conflicts of interest.

Appendix A

In Table A1, we summarize recent and relevant literature studies concerning the continuous-review base-stock policy.

Table A1. Relevant literature studies concerning the continuous-review base-stock policy.

Continuous-Review Base-Stock Models		Model Features						Optimization Approach	
-Returns **-Backorder+ Lost Sales** **-Limited Shelf Life** **-Total Loss**	**Paper**	**Demand/Return**	**Shelf Life/Loss**	**Lead Time**	**Base-Stock Inventory Policy**	**Stockout Policy**	**Additional Features**		
	[1]	Linear, Comp. Pois.	Infinite	Zero	Multi-targets(s, S)	Not allowed	• Disc. Criteria • Cash management	■	Hamilton-Jacobi-Bellman equation
	[9]	Brownian	Decay rate	Det., Stoch.	(Q, r)	Backorder, lost sales	• Ave. criteria	■ ■	Kuhn-Tucker conditions numerical
	[25]	Pois.	General/ no loss	General	(S, $S-1$)	Lost sales	• Extension to backorder • Ave. criteria	■ ■	Queueing modelling Algorithmic approach
	[21]	Fixed rates	Fixed rate	Fixed	(Q, r)	Backorder	• Ave. criteria • Correlated demands/returns	■	Derivation
	[13]	Linear	Decay rate	Det., Stoch.	(Q, r)	Backorder	• Ave. criteria	■ ■	Kuhn-Tucker conditions numerical
	[48]	Pois.	exp	exp	(s, Q)	Lost sales	• Ave. criteria • Two types of customers	■	Markov equations
	[6]	State-dep. Pois.	exp/ exp	exp	(S, s)	Lost sales	• Ave. criteria • Cancelations	■ ■	Recursive formulas numerical
	[11]	General	Three-period	Zero	(S, s, $S(x)$)	Lost sales or backorder	• Emergency option • Disc. criteria	■ ■	Closed-form optimum Structural properties
	[4]	Linear + MAP	no	exp	(M, S, s)	Lost sales	• Disc. criteria	■ ■	Martingales First passage times analysis
	This study	**Comp. Pois.**	**exp /exp**	**exp**	**(S, s,B)**	**Lost sales, Partial Backorder**	• **Ave. criteria** • **Limited storage capacity**	■	Markov chain

References

1. Huang, S.; Lu, H.; Lin, J.; Ponte, B. On the dynamics of return collection in closed-loop supply chains. *Int. J. Prod. Res.* **2023**, *62*, 909–932. [CrossRef]
2. Jeganathan, K.; Selvakumar, S.; Saravanan, S.; Anbazhagan, N.; Amutha, S.; Cho, W.; Joshi, G.P.; Ryoo, J. Performance of stochastic inventory system with a fresh item, returned Item, refurbished item, and multi-class customers. *Mathematics* **2022**, *10*, 1137. [CrossRef]
3. Smith, E.G. How online retailers predict your perfect outfit. *Wall Str. J.* **2015**, *5*.
4. Cui, H.; Rajagopalan, S.; Ward, A.R. Predicting product return volume using machine learning methods. *Eur. J. Oper. Res.* **2020**, *281*, 612–627. [CrossRef]
5. Ambilkar, P.; Dohale, V.; Gunasekaran, A.; Bilolikar, V. Product returns management: A comprehensive review and future research agenda. *Int. J. Prod.* **2022**, *60*, 3920–3944. [CrossRef]
6. Meneses, M.; Marques, I.; Cárdenas-Barrón, A. Blood inventory management: Ordering policies for hospital blood banks under uncertainty. *Int. Oper. Res.* **2023**, *30*, 273–301. [CrossRef]
7. Gharaei, A.; Hoseini Shekarabi, S.A.; Karimi, M. Optimal lot-sizing of an integrated EPQ model with partial backorders and re-workable products: An outer approximation. *Int. J. Syst. Sci. Oper. Logist.* **2023**, *10*, 2015007. [CrossRef]
8. Li, R.; Liu, Y.; Teng, J.T.; Tsao, Y.C. Optimal pricing, lot-sizing and backordering decisions when a seller demands an advance-cash-credit payment scheme. *Eur. J. Oper. Res.* **2019**, *278*, 283–295. [CrossRef]
9. Azcue, P.; Muler, N. Optimal cash management problem for compound Poisson processes with two-sided jumps. *Appl. Math. Optim.* **2019**, *80*, 331–368. [CrossRef]
10. Barron, Y. A stochastic card balance management problem with continuous and batch-type bilateral transactions. *Oper. Res. Perspect.* **2023**, *10*, 100274. [CrossRef]
11. Kouki, C.; Legros, B.; Babai, M.Z.; Jouini, O. Analysis of base-stock perishable inventory systems with general lifetime and lead-time. *Eur. J. Oper. Res.* **2020**, *287*, 901–915. [CrossRef]
12. Presman, E.; Sethi, S.P. Inventory models with continuous and Poisson demands and discounted and average costs. *Prod. Oper. Manag.* **2006**, *15*, 279–293. [CrossRef]
13. Sahin, I. On the continuous-review (s, S) inventory model under compound renewal demand and random lead times. *J. Appl. Probab.* **1983**, *20*, 213–219. [CrossRef]
14. Pacheco-Velázquez, E.A.; Robles-Cárdenas, M.; Juárez Ordóñez, S.; Damy Solís, A.E.; Cárdenas-Barrón, L.E. A heuristic model for spare parts stocking based on Markov chains. *Mathematics* **2023**, *11*, 3550. [CrossRef]
15. Barron, Y. An order-revenue inventory model with returns and sudden obsolescence. *Oper. Lett.* **2017**, *46*, 88–92. [CrossRef]
16. Barron, Y. Integrating replenishment policy and maintenance services in a stochastic inventory system with bilateral movements. *Mathematics* **2023**, *11*, 864. [CrossRef]
17. Chen, K.; Song, J.S.; Shang, J.; Xiao, T. Managing hospital platelet inventory with mid-cycle expedited replenishments and returns. *Prod. Oper. Manag.* **2022**, *31*, 2015–2037. [CrossRef]
18. Sermkulpat, V.; Buddhakulsomsiri, J. An (R, Q) Inventory Policy for Medicine under Limited Storage with a Backlog. Doctoral Dissertation, Thammasat University, Bangkok, Thailand, 2022.
19. Rabinowitz, G.; Mehrez, A.; Chu, C.W.; Patuwo, B.E. A partial backorder control for continuous review (r, Q) inventory system with Poisson demand and constant lead time. *Comput. Oper. Res.* **1995**, *22*, 689–700. [CrossRef]
20. Barron, Y.; Hermel, D. Shortage decision policies for a fluid production model with MAP arrivals. *Int. J. Prod. Res.* **2017**, *55*, 3946–3969. [CrossRef]
21. Poormoaied, S.; Atan, Z.; de Kok, T.; Van Woensel, T. Optimal inventory and timing decisions for emergency shipments. *IISE Trans.* **2020**, *52*, 904–925. [CrossRef]
22. Bijvank, M.; Huh, W.T.; Janakiraman, G. 1. Lost-sales inventory systems. In *Research Handbook on Inventory Management*; Edward Elgar Publishing: Cheltenham Glos, UK, 2023; pp. 2–26.
23. San-José, L.A.; Sicilia, J.; Pando, V.; Alcaide-López-de-Pablo, D. An inventory system with time-dependent demand and partial backordering under return on inventory investme. *Comput. Oper. Res.* **2022**, *145*, 105861. [CrossRef]
24. Taleizadeh, A.A.; Soleymanfar, V.R.; Govindan, K. Sustainable economic production quantity models for inventory systems with shortage. *J. Clean. Prod.* **2018**, *174*, 1011–1020. [CrossRef]
25. Schrady, D.A. A deterministic inventory model for reparable items. *Nav. Res. Logist. Q.* **1967**, *14*, 391–398. [CrossRef]
26. Gajdalo, S. *Heuristics for Computing Variable Safety Levels/Economic Order Quantities for Repairable Items*; AMC Inventory Research Office, Institute of Logistics Research, US Army Logistics Management Center: Fort Lee, VA, USA, 1973.
27. Heyman, D.P. Return policies for an inventory system with positive and negative demands. *Nav. Res. Q.* **1978**, *25*, 581–596. [CrossRef]
28. Yuan, X.M.; Cheung, K.L. Modeling returns of merchandise in an inventory system. *Oper. Res. Spektrum* **1998**, *20*, 147–154. [CrossRef]
29. Schulz, T.; Voigt, G. A flexibly structured lot sizing heuristic for a static remanufacturing system. *Omega* **2014**, *44*, 21–31. [CrossRef]
30. Barron, Y.; Dreyfuss, M. A triple (S, s, ℓ)-thresholds base-stock policy subject to uncertainty environment, returns and order cancellations. *Comput. Oper. Res.* **2021**, *134*, 105320. [CrossRef]

31. Fleischmann, M.; Kuik, R. On optimal inventory control with independent stochastic item returns. *Eur. J. Oper. Res.* **2003**, *151*, 25–37. [CrossRef]
32. Choné, P.; Linnemer, L. Linear demand systems for differentiated goods: Overview and user's guide. *Int. J. Ind. Organ.* **2020**, *73*, 102663. [CrossRef]
33. Sicilia, J.; San-José, L.A.; García-Laguna, J. An inventory model where backordered demand ratio is exponentially decreasing with the waiting time. *Ann. Oper. Res.* **2012**, *199*, 137–155. [CrossRef]
34. Goedhart, J.; Haijema, R.; Akkerman, R. Modelling the influence of returns for an omni-channel retailer. *Eur. J. Oper. Res.* **2023**, *306*, 1248–1263. [CrossRef]
35. Ho, T.F.; Lin, C.C.; Lin, C.L. Determining the optimal inventory and number of shipments for a two-resource supply chain with correlated demands and remanufacturing products allowing backorder. *Mathematics* **2020**, *8*, 548. [CrossRef]
36. Jani, M.Y.; Patel, H.A.; Bhadoriya, A.; Chaudhari, U.; Abbas, M.; Alqahtani, M.S. Deterioration control decision support system for the retailer during availability of trade credit and shortages. *Mathematics* **2023**, *11*, 580. [CrossRef]
37. Thilagavathi, R.; Viswanath, J.; Cepova, L.; Schindlerova, V. Effect of inflation and permitted three-slot payment on two-warehouse inventory system with stock-dependent demand and partial backlogging. *Mathematics* **2022**, *10*, 3943. [CrossRef]
38. Chen, K.; Xiao, T.; Wang, S.; Lei, D. Inventory strategies for perishable products with two-period shelf-life and lost sales. *Int. J. Prod.* **2021**, *59*, 5301–5320. [CrossRef]
39. Kouki, C.; Arts, J.; Babai, M.Z. Performance evaluation of a two-echelon inventory system with network lost sales. *Eur. J. Oper. Res.* **2023**, *314*, 647–664. [CrossRef]
40. Perera, S.C.; Sethi, S.P. A survey of stochastic inventory models with fixed costs: Optimality of (s, S) and (s, S)-type policies-Continuous-time case. *Prod. Oper. Manag.* **2023**, *32*, 154–169. [CrossRef]
41. Montgomery, D.C.; Bazaraa, M.S.; Keswani, A.K. Inventory models with a mixture of backorders and lost sales. *Nav. Res. Logist. Q.* **1973**, *20*, 255–263. [CrossRef]
42. Rosenberg, D. A new analysis of a lot-size model with partial backlogging. *Nav. Res. Logist. Q.* **1979**, *26*, 349–353. [CrossRef]
43. Moinzadeh, K. Operating characteristics of the $(S-1, S)$ inventory system with partial backorders and constant resupply times. *Manag. Sci.* **1989**, *35*, 472–477. [CrossRef]
44. Braglia, M.; Castellano, D.; Marrazzini, L.; Song, D. A continuous review (Q, r) inventory model for a deteriorating item with random demand and positive lead time. *Comput. Oper. Res.* **2019**, *109*, 102–121. [CrossRef]
45. Guo, Z.; Chen, H. Cost evaluation of a (Q, r, K) inventory model with two demand classes of lost sales and backorders. In Proceedings of the IFIP International Conference on Advances in Production Management Systems, Trondheim, Norway, 17–21 September 2023; Springer Nature: Cham, Switzerland, 2023; pp. 257–270.
46. Chang, C.T.; Lo, T.Y. On the inventory model with continuous and discrete lead time, backorders and lost sales. *Appl. Math. Model.* **2009**, *33*, 2196–2206. [CrossRef]
47. Wang, D.; Tang, O. Dynamic inventory rationing with mixed backorders and lost sales. *Int. J. Prod. Econ.* **2014**, *149*, 56–67. [CrossRef]
48. Taleizadeh, A.A.; Zamani-Dehkordi, N. A stochastic lot sizing model with partial backordering and imperfect production processes. *Int. J. Inventory Res.* **2017**, *4*, 75–96. [CrossRef]
49. Lin, H.J. Two-critical-number control policy for a stochastic production inventory system with partial backlogging. *Int. J. Prod. Res.* **2017**, *55*, 4123–4135. [CrossRef]
50. Wang, C.N.; Dang, T.T.; Nguyen, N.A.T. A computational model for determining levels of factors in inventory management using response surface methodology. *Mathematics* **2020**, *8*, 1210. [CrossRef]
51. Uzunoglu Kocer, U.; Yalcin, B. Continuous review (s, Q) inventory system with random lifetime and two demand classes. *Opsearch* **2020**, *57*, 104–118. [CrossRef]

mathematics

MDPI

Article

The Emission Reduction Technology Decision of the Port Supply Chain

Yan Zhou and Haiying Zhou *

Department of Port and Shipping Management, Guangzhou Maritime University, Guangzhou 510725, China; zhouyan@gzmtu.edu.cn

* Correspondence: zhouhaiying@gzmtu.edu.cn; Tel.: +86-135-8033-1480

Abstract: The technology options for sustainable development are explored with customer low-carbon preference in a port supply chain consisting of one ship and one port. Port supply chains can opt for either shower power or low-sulfur fuel oil to cut down emissions. We set game models considering three power structures: the port dominant (port-led Stackelberg game), the ship dominant (ship-led Stackelberg game), and the port and ship on the same footing (Nash game). We compare the performances of different technologies. It is shown that, when customer low-carbon preference and carbon tax are both low, LSFO is the appropriate choice from the supply chain's profit perspective, SP is preferred from the emission control perspective, and LSFO is preferred from the social welfare perspective. However, when customers' low-carbon preferences, carbon tax, and environmental concerns are all low or all high, LSFO should be adopted from the view of social welfare. The profits and carbon emissions of the supply chain in the Nash game are higher than those in the Stackelberg game. While the environmental concern is low, the social welfare of the supply chain in the Nash game is greater than that in the Stackelberg game. Otherwise, it is less than that in the Stackelberg game. The obtained results can help governments formulate policies and ships make emission reduction technology decisions according to their own interests.

Keywords: port supply chain; customers' low-carbon preference; emission reduction technologies

MSC: 91-10

1. Introduction

The port industry is an important link to the modern logistics supply chain. However, the port is also a large carbon emitter [1] because it relies on the consumption of petrochemical energy. Statistics show that the carbon emissions emitted by the ports and ships yearly are more than 2.7% of total global emissions. Without appropriate steps, the proportion will more than double by 2050 [2]. A 2018 study by the International Transport Forum emphasizes the crucial role of ports in decreasing the carbon emissions of worldwide marine transport. Implementing low-carbon initiatives at ports may greatly enhance the process of lowering carbon emissions in maritime transport [3]. Ships are usually powered by diesel auxiliary engines when berthing at the ports, and the diesel auxiliary engines emit large amounts of pollutants such as CO, NOx, and SOx into the air during operation, which account for 55.0–77.0% of port emissions [4]. The 70th session of the IMO in October 2016 introduced amendments, guidelines, and circulars of the International Convention for the Prevention of Pollution from Ships. The focus was on amending Annex VI "Rules for the Prevention of Air Pollution from Ships" and establishing a global mandate effective from 1 January 2020, limiting the sulfur content of marine fuels to 0.50%.

In recent years, China's port development has achieved remarkable results, and world-class port clusters such as the Yangtze River Delta, the Bohai Sea Rim, Guangdong, Hong Kong, and Macao have been gradually built. The port's cargo throughput and container

throughput have ranked first in the world for many consecutive years. Among the top 10 ports in terms of cargo throughput and container throughput, China accounts for eight and seven seats, respectively. To promote the emission reduction of air pollutants from ships, the Chinese government has implemented a series of measures. For example, in December 2015, the Ministry of Transport of China established a domestic emission control area (ECA) in the waters of the Pearl River Delta, the Yangtze River Delta, and the Bohai Rim (Beijing, Tianjin, and Hebei). At the end of 2018, the scope of the ECA in China was expanded to include coastal areas and major inland waters. In September 2021, the "Measures for the Administration of Shore Power for Ports and Ships" stipulated that ships (except liquid cargo ships) with shore power (SP) facilities should use SP when berthing at a berth with SP supply capacity in a coastal port for more than 3 h, or at a berth with SP supply capacity in an inland port for more than 2 h, if the ships do not adopt effective alternative emission reduction measures [5].

SP and low-sulfur oil (LSFO) are frequently used as emission reduction technologies for ships when berthing at berths. SP serves as a land-to-ship electricity connection, enabling ships to switch off onboard diesel-powered engines when they are docked. In order to use shore power, ships have to install facilities to receive power, while ports build facilities on the shore to transmit power. Both ports and ships have to afford a significant initial investment to construct SP facilities, but SP can cut down carbon emissions effectively. Hall [6] proposed that SP could reduce all carbon emissions of ships at ports by 48–70%. LSFO is a clean energy source with less than 0.1% sulfur content. Adopting LSFO without retrofitting and upgrading means no initial investment. Nevertheless, LSFO is expensive and emits a greater amount of carbon emissions compared to SP.

Enterprises make emission reduction decisions based on their own factors and the supply chain's external factors which include government regulation policy and consumers' green preferences. The carbon trading system and carbon tax are two common regulation policies. A carbon tax is a tax on carbon emissions from the burning of fossil fuels. The carbon tax policy is more flexible, fairer, and broader in coverage, and it can effectively achieve the sustainable development of the economy and environment. Under the carbon tax policy, to avoid high taxes, ports and shipping companies have to take measures to reduce carbon emissions, and reducing emissions inevitably requires investment, such as upgrading and improving technology, replacing with cleaner fuels, and limiting speed. At present, there are many countries implementing carbon tax policies, such as Sweden, the Netherlands, Norway, and Finland [7]. To help government regulators formulate more reasonable carbon tax schemes, many scholars have studied the carbon tax policies in port areas [8–10].

On the other hand, consumers are becoming more inclined to protect the environment and are more likely to opt for low-carbon products when making purchases. They are prepared to shell out more money for low-carbon products. Wang and Zhao [11] investigated how manufacturers and retailers reduce their carbon emissions when customers have a preference for low-carbon products and concluded that consumers' preference for low-carbon products is a key factor in enterprise decision making. Therefore, under the carbon tax, considering the low-carbon preferences of customers, how do port and shipping enterprises make emission reduction technology decisions?

Based on this, we evaluate the two technologies from a port supply chain point of view considering the customers' low carbon preference. The port supply chain consists of one port and one ship. The port is an upstream member and is responsible for providing the service of loading/unloading to the ship, while the ship offers a range of services to customers. Furthermore, the regulators impose appropriate carbon taxes on the carbon emissions of a port to maximize social welfare [12]. The port supply chain chooses suitable emission reduction technologies with profit maximization considering the customers' low-carbon preference. We set models by the game method in three scenarios (port-led Stackelberg (PS) [13], ship-led Stackelberg (SS) [14], and Nash (NS) game). A comparison of profits, carbon emissions, and sustainability between the two technologies is conducted.

We offer the following three primary contributions. (1) We investigate the selection of technologies from a port supply chain point of view, a topic that has not been widely explored in the current research. (2) We compare SP and LSFO under carbon tax from the perspective of social benefits. (3) We consider the customers' low-carbon preference, which is rarely discussed in the field of port and shipping. The acquired outcomes assist the port and ship in selecting suitable technologies and offer valuable perspectives on government policy.

The remainder of this paper continues with some related literature. We then introduce the necessary notation and assumptions, which are followed by the models and solutions. We compare the equilibrium results and engage in a comprehensive discussion, subsequently proceeding with numerical experiments. Finally, the conclusions are given.

2. Literature Review

This research has a strong correlation with three areas of the literature: low-carbon technology choice in supply chains, marine transport logistics and green ports, and customers' low-carbon preference.

2.1. Low-Carbon Technology Choice in Supply Chains

In the background of the growing influence of environmental regulations, constructing low-carbon supply chains is an inevitable choice for enterprises' development. The adoption of low-carbon technology is very important in supply chains. Liu et al. [15] evaluated multi-stage low-carbon technology investment strategies by constructing an evolution model that consists of a manufacturer investing in low-carbon technology and a supplier providing low-carbon technology. Cao et al. [16] examined mode selection strategies of energy performance contracts considering carbon tax policy. Jiang et al. [17] constructed a differential game model based on social welfare maximization and discussed how the technology spillover effect and incremental cost of carbon transfer influence carbon quota allocation. Liu et al. [18] investigated how GT's investment strategy in a manufacturer–supplier supply chain affected investment and sustainable production decisions, as well as the most advantageous government subsidy incentive. Yu et al. [19] put forward a framework for investing in energy-efficient technologies to address the decision-making challenges faced by companies striving to achieve a harmonious equilibrium between profits and investments. Song et al. [20] considered subsidies, consumers' low-carbon preferences, and the low-carbon information trust to analyze the optimal operation decisions of two manufacturers in a green supply chain. Yang et al. [21] investigated the port emission reduction decision-making problem within a carbon trading mechanism. They compared the SP and LFSO from the view of sustainable development, but they did not consider the impact of customers' green preferences.

The above research mainly investigates the emission reduction investment strategy and technology selection in the production supply chain under government regulation policies, such as carbon trading systems, carbon tax, and subsidies. However, in the port and shipping supply chain, relevant research on emission reduction investment strategies and technology decision making is still lacking. In fact, the emission reduction decision-making process between the port supply chain and the general production supply chain is completely different. Taking shore power as an example, the use of SP not only requires facilities investment from ports but also requires ships to transform SP facilities, and the initial investment of both sides is huge. Ports can incentivize ships to use SP through measures such as granting priority berthing rights to ships and preferential SP service fees. Considering the characteristics of emission reduction technologies in the port industry, the study compares the two technologies from the supply chain perspective, which could enrich the current study.

2.2. Marine Transport Logistics and Green Ports

The emergence of international trade has made shipping a major means of transporting goods across the globe. The port has shifted its role from being a mere loading and unloading hub to becoming the hub of worldwide logistic services, making it a critical element of maritime transportation. Consequently, the port area is likely to generate more pollutants, which has piqued the interest of researchers in the development of eco-friendly harbors.

Ding et al. [22] examined the economic viability of the Northern Sea Route in comparison to the Suez Canal Route through the implementation of two proposed carbon tax schemes (fixed vs. progressive). Cariou et al. [23] analyzed the consequences of using a maritime bunker levy on the financial benefits, commercial activities, and ship owners' emissions. Gao et al. [24] investigated the architecture of a container ocean shipping system under a carbon tax. Song et al. [25] examined the effects of the carbon tax on the green shipping supply chain in the context of port competition. Xin et al. [26] developed a programming model based on integers to facilitate the green scheduling of shuttle tanker fleets.

Scholars also explored the impact of environmental regulation policies on port operations. Zhao et al. [27] established a carbon emission calculation model for container port operations based on energy consumption, using ships and containers as measurement units, aiming to enable port enterprises to develop reasonable carbon peaking plans. The support vector regression model considering outlier is used to predict the container throughput, then the time series of carbon emissions is obtained through the calculation model, and the judgment standard of carbon peak time based on the Mann–Kendall trend test method is designed to determine the carbon peak time. Li et al. [28] built a two-tiered maritime supply chain comprising a solitary port and shipping company to tackle the problem of collaborative emission reduction between ports and shipping companies while ensuring pollution control. They established four decision-making models of ports and ships: port-led Stackelberg game, vertical integration model, Nash bargaining of port and ship cooperation, and green port and ship cooperation. They also analyzed changes in port prices, emission reduction levels, freight volume of shipping companies, and profits. Chen et al. [29] studied the influence of ECA on international shipping and proposed that a great deal of collaboration and coordination should be implemented on a global scale to decrease ship emissions. Hu et al. [30] investigated the shipping container subsidy in the context of multimodal transportation that includes waterways within the regional transportation network. Liu et al. [31] examined the impact of dividing three distinct ECAs on the quality of port air in the Pearl River Delta region.

At present, in terms of green ports, traditional green shipping management issues have been combined with emission reduction requirements, and some green port technologies and their emission reduction effects have been studied in depth. Ports have various resources such as berths, shore bridges, and yard bridges. Most of the current research uses cost-oriented objectives to optimize the allocation or scheduling of these resources, and when the emission minimization or energy-saving maximization goals are taken into account, the decision model and solution algorithm need to be redefined and designed. This study systematically analyzes and compares port emission reduction technologies, taking into account not only their costs but also their emissions, as well as overall social welfare.

2.3. Customers' Low-Carbon Preference

With the gradual emphasis on ecosystems and the widespread dissemination of the low-carbon consumption concept, consumers' low-carbon preference is gradually increasing. Studies indicate that products with lower carbon emissions are more marketable, and consumers are inclined to spend more on them [32,33]. The customers' preference in the port and navigation field contributes positively to the reduction in port emissions, although research on this subject is scarce. Scholars have extensively debated the low-carbon preference of consumers in the past few years. This study primarily focuses on three

aspects: verifying whether consumers have such preferences, identifying key elements influencing their low-carbon preferences, and examining how low-carbon preferences affect decisions in low-carbon supply chains.

In terms of how low-carbon preferences affect the operational decisions in supply chains, Yu and Hou [34] established a differential game model based on cost sharing and coordination to study how consumers' low-carbon preferences affected market demand in the product supply chain. Xu et al. [35] proposed that the manufacturers' emission reduction efforts and the dealers' low-carbon promotion efforts were affected by the cost coefficient related to the promotion of low-carbon products by dealers, the low-carbon reputation sensitivity coefficient of consumers, and the impact coefficient of the emission reduction efforts of manufacturers on the low-carbon reputation. Gao et al. [36] studied the incentive strategy of the low-carbon supply chain through modeling and optimization methods based on the information update of low-carbon preferences. The results show that profit-driven cooperation among supply chain members could improve both their own profits and the carbon reduction efficiency of the entire supply chain. Effective information updates are more efficient at reducing carbon emissions than promotional allowances. Ding et al. [37] investigated manufacturers' decisions on encroachment and carbon emissions reduction, taking into account the carbon trading system and consumers' preferences for low-carbon production. The findings indicated that manufacturers invariably benefited from encroachment decisions if the government opted against cap-and-trade regulations, leading to consistent profit losses for retailers. Sun et al. [38] established a manufacturer-led Stackelberg game model considering the low-carbon preferences of consumers and the lag of emission reduction technologies. They proposed that the low-carbon preferences of consumers and the lag time of emission reduction technologies have a positive effect on the carbon emission transfer level of manufacturers but have no effect on the commitment level of suppliers. Only when the lag time of emission reduction technologies is kept within an appropriate range would the increase in consumers' low-carbon preference increase the supply chain profits. Zhu [39] concluded that in the case of low R&D difficulty, high consumer trust or R&D difficulty, high consumers' low-carbon preferences, and brand recognition, advantageous brand enterprises could improve CER levels and profits by using blockchain.

At present, the research on low-carbon preference mainly focuses on the field of low-carbon supply chain operation and management, including supply chain production decision making, emission reduction investment strategy, and cooperation among supply chain members. But, there are relatively few studies on the influence of customers' low-carbon preferences on emission reduction decision making within the port supply chain. This study compares port emission reduction technologies, considering the impact of customers' low-carbon preferences.

3. Notation and Assumptions

The port emission reduction game model considering customers' low-carbon preferences under carbon tax is composed of a port, a ship, and port customers, as shown in Figure 1. SP and LSFO could be chosen to cut down carbon emissions. There is a lack of correlation between demand and competition among vessels. The simple supply chain consisting of one port and one ship is put forward to expound the problem in the literature [21,23,40]. Of course, in future studies, we could also extend to numerous ports and ships.

3.1. Notation

The parameters and variables are expressed in Table 1. i ($i = E, L$) represents the emission reduction technologies, where E describes SP and L denotes LSFO. The subscript j ($j = S, P, N$) denotes the types of game (SS, PS, and NS). The supply chain adopts technology i ($i = E, L$) in j game, denoted by the subscript $i - j$. Port, ship, and supply chain are denoted by superscripts k ($k = p, s, sc$). p represents the total service price and satisfies $p = m + w$.

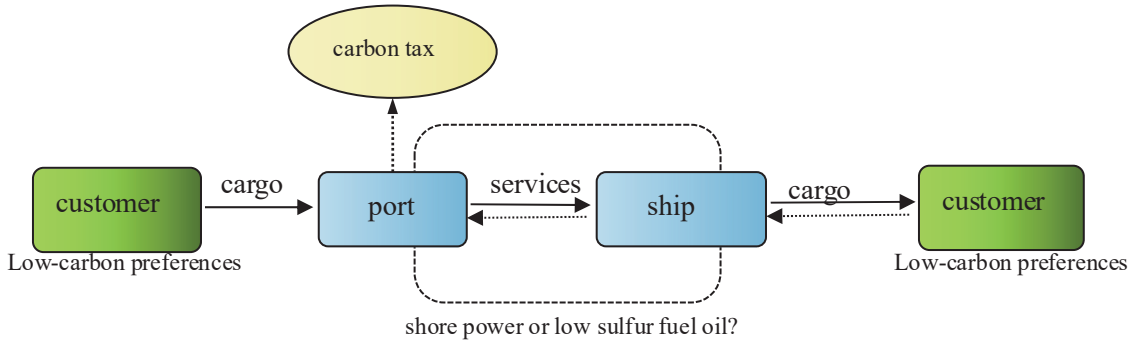


Figure 1. The process of emission reduction decision making in the port supply chain.

Table 1. Variables.

Notation	Description
Variables	
a	Market size, $a > 0$
b	Sensitivity coefficient of market demand to the price $(b > 0)$
q	Demand for cargo
c_t	The unit navigation transportation cost of ship
c_s	The unit service cost of ship when SP is implemented
c_L	The unit service cost of ship when LSFO is implemented
c_E	The unit service cost of port when SP is implemented
γ	Low-carbon preference of customers
θ	Carbon tax
e_i	Technology i's unit carbon emissions
p_c	Environmental concern
U^k_{i-j}	The profit of k in j game with adoption technology i
T_{i-j}	The carbon emissions in j game with adoption technology i
sw_{i-j}	The social welfare in j game with adoption technology i
Decision variables	
w_{i-j}	The ship's service price when adopting technology i in j game
m_{i-j}	The port's service price when adopting technology i in j game

3.2. Basic Assumptions

To make it easier to model and analyze, the following two assumptions are given.

Assumption 1. $c_L < c_E + c_s, e_L > e_E$.

In practice, the construction investment of shore power is huge; taking the power supply of 10,000 TEU container ships as an example, the construction cost for the port to transmit power is about CNY 6–10 million, and the cost for the ship to install facilities to receive power is about CNY 3–6 million. Especially under the premise that the current utilization rate of shore power is low in China, the total cost of using SP for ships is greater than the cost of using LSFO, which is about CNY 1 more per KWH [41].

Therefore, it is assumed that the costs of SP are higher, i.e., $c_L < c_E + c_s$. Moreover, SP produces fewer carbon emissions compared with LSFO [42], i.e., $e_E < e_L$.

Assumption 2. *The function of demand is* $q = a - b(m + w) + \gamma(e - e_i), i = E, L; a, b > 0$. *$m$ denotes the port's service price, while w indicates the ship's service price. $w + m$ denotes the service price faced by the customer, and $\gamma(e - e_i)$ represents the effect of low-carbon preferences on demand. In the research on port supply chains, the linear demand function is widely used, referring to the literature* [43,44].

4. Modeling and Solving

4.1. The Adoption of SP

Formulas (1)–(5) are used to calculate the profits, carbon emissions, and social welfare when SP is employed.

$$U^p = mq - c_E q - e_E \theta q \tag{1}$$

$$U^s = wq - c_t q - c_s q = (w - c_t - c_s)(a - b(m + w) + \gamma(e - e_E)) \tag{2}$$

$$U^{sc} = U^s + U^p \tag{3}$$

$$T = qe_i, i = E, L \tag{4}$$

$$sw = U^s + U^p - p_c T^2 \tag{5}$$

Equation (1) represents the earnings of the port, mq represents the revenues from shipping companies, $c_E q$ describes the cost to use SP, and $qe_L\theta$ describes the expenses of releasing carbon dioxide. In Equation (2), wq represents the ship's profits from customers, $c_t q$ describes the cost of transportation, and $c_L q$ is the expense associated with using SP. Equations (3) and (4) represent the total profits and the total carbon emissions, respectively. According to Krass [12], $p_c T^2$ refers to the negative disutility of carbon emissions in Equation (5), and the social welfare (sw) includes the ship's profits and the port's profits minus the negative disutility of emissions.

The results are given in Table 2 by standard backward induction, and the solution process is given in Appendix A.

Table 2. Results with adoption SP.

Cases	m_{E-j}	w_{E-j}	T_{E-j}
$j = N$	$\frac{a+\gamma e-\gamma e_E-bc_E+2bc_s-bc_t+2be_E\theta}{3b}$	$\frac{a+\gamma e-\gamma e_E+2bc_E-bc_s+2bc_t-be_E\theta}{3b}$	$\frac{e_E A}{3}$
$j = S$	$\frac{a+\gamma e-\gamma e_E-bc_E+3bc_s-bc_t+3be_E\theta}{4b}$	$\frac{a+\gamma e-\gamma e_E+bc_E-bc_s+bc_t-be_E\theta}{2b}$	$\frac{e_E A}{4}$
$j = P$	$\frac{a+\gamma e-\gamma e_E-bc_E+bc_s-bc_t+be_E\theta}{2b}$	$\frac{a+\gamma e-\gamma e_E+3bc_E-bc_s+3bc_t-be_E\theta}{4b}$	$\frac{e_E A}{4}$
Cases	U^p_{E-j}	U^s_{E-j}	U^{sc}_{E-j}
$j = N$	$\frac{(a+\gamma e-\gamma e_E-bc_E-bc_s-bc_t-be_E\theta)^2}{9b}$	$\frac{(a+\gamma e-\gamma e_E-bc_E-bc_s-bc_t-be_E\theta)^2}{9b}$	$\frac{2A^2}{9b}$
$j = S$	$\frac{(a+\gamma e-\gamma e_E-bc_E-bc_s-bc_t-be_E\theta)^2}{16b}$	$\frac{(a+\gamma e-\gamma e_E-bc_E-bc_s-bc_t-be_E\theta)^2}{8b}$	$\frac{3A^2}{16b}$
$j = P$	$\frac{(a+\gamma e-\gamma e_E-bc_E-bc_s-bc_t-be_E\theta)^2}{8b}$	$\frac{(a+\gamma e-\gamma e_E-bc_E-bc_s-bc_t-be_E\theta)^2}{16b}$	$\frac{3A^2}{16b}$
Cases	sw_{E-j}		
$j = N$	$\frac{(2-be_E^2 p_c)(a+\gamma e-\gamma e_E-bc_E-bc_s-bc_t-be_E\theta)^2}{9b}$		
$j = S$	$\frac{(3-be_E^2 p_c)(a+\gamma e-\gamma e_E-bc_E-bc_s-bc_t-be_E\theta)^2}{16b}$		
$j = P$	$\frac{(3-be_E^2 p_c)(a+\gamma e-\gamma e_E-bc_E-bc_s-bc_t-be_E\theta)^2}{16b}$		

Here, $A = a + \gamma e - \gamma e_E - bc_E - bc_s - bc_t - be_E\theta$.

4.2. Implementation of LSFO

Formulas (6) and (7) are used to calculate the profits when LSFO is employed.

$$U^p = mq - e_L \theta q \tag{6}$$

$$U^s = wq - c_t q - c_L q = (w - c_t - c_L)(a - b(m + w) + \gamma(e - e_L)) \tag{7}$$

Equation (1) represents the earnings of the port, mq represents the revenues from shipping companies, and $qe_L\theta$ describes the expenses of releasing carbon dioxide. Equation (2) represents the profits of the ship, wq represents the ship's profits, $c_t q$ describes the cost of transportation, and $c_L q$ is the expense associated with using LSFO.

By standard backward induction again, the results are given in Table 3. The solution process is shown in Appendix A.

Table 3. Results with adoption of LSFO.

Cases	m_{L-j}	w_{L-j}	T_{L-j}
$j = N$	$\frac{a+\gamma e+2bc_L+2be_L\theta-\gamma e_L-bc_t}{3b}$	$\frac{a+\gamma e+2bc_t-\gamma e_L-bc_L-be_L\theta}{3b}$	$\frac{e_L B}{3}$
$j = S$	$\frac{a+\gamma e+3bc_L+3be_L\theta-\gamma e_L-bc_t}{4b}$	$\frac{a+\gamma e+bc_t-\gamma e_L-bc_L-be_L\theta}{2b}$	$\frac{e_L B}{4}$
$j = P$	$\frac{a+\gamma e+bc_L+be_L\theta-\gamma e_L-bc_t}{2b}$	$\frac{a+\gamma e+3bc_t-\gamma e_L-bc_L-be_L\theta}{4b}$	$\frac{e_L B}{4}$
Cases	U^p_{L-j}	U^s_{L-j}	U^{sc}_{L-j}
$j = N$	$\frac{(a-bc_L-be_L\theta-bc_t+\gamma e-\gamma e_L)^2}{9b}$	$\frac{(a-bc_L-be_L\theta-bc_t+\gamma e-\gamma e_L)^2}{9b}$	$\frac{2B^2}{9b}$
$j = S$	$\frac{(a-bc_L-be_L\theta-bc_t+\gamma e-\gamma e_L)^2}{16b}$	$\frac{(a-bc_L-be_L\theta-bc_t+\gamma e-\gamma e_L)^2}{8b}$	$\frac{3B^2}{16b}$
$j = P$	$\frac{(a-bc_L-be_L\theta-bc_t+\gamma e-\gamma e_L)^2}{8b}$	$\frac{(a-bc_L-be_L\theta-bc_t+\gamma e-\gamma e_L)^2}{16b}$	$\frac{3B^2}{16b}$
Cases		sw_{L-j}	
$j = N$		$\frac{\left(2-be_L^2 p_c\right)(a+\gamma e-bc_t-bc_L-be_L\theta-\gamma e_L)^2}{9b}$	
$j = S$		$\frac{\left(3-be_L^2 p_c\right)(a+\gamma e-\gamma e_L-bc_L-bc_t-be_L\theta)^2}{16b}$	
$j = P$		$\frac{\left(3-be_L^2 p_c\right)(a+\gamma e-bc_t-bc_L-be_L\theta-\gamma e_L)^2}{16b}$	

Here, B = $a + \gamma e - \gamma e_L - bc_L - bc_t - be_L\theta$.

5. Analysis and Discussion

This section evaluates the impact of various variables on profits, emissions, and social welfare, which can help shipping companies decide on appropriate technologies and provide decision-making advice for governments to formulate corresponding policies.

For detailed proofs of lemmas and propositions, we refer the reader to Appendix A.

5.1. Profit Analysis

In this section, we investigate how profits are affected by the parameters and carbon tax in the port supply chain.

Lemma 1. *U^{sc}_{i-j} is a concave of $c_i, c_t, \gamma, \theta,$ and e_i. Furthermore, U^{sc}_{E-j} is also concave with respect to c_s.*

When the customers' low-carbon preferences are certain, as carbon taxes, emissions, and operational costs increase, the total profits of the port supply chain initially suffer as a result of higher costs. When operating costs and carbon taxes reach an exorbitant level, the supply chain may not be able to provide services, resulting in a complete cessation of profits. When the carbon tax is certain, with the low-carbon preference increasing, the market demand also increases; thus, the profits of the supply chain rise. However, when the consumers' low-carbon preference and carbon tax increase at the same time, although consumers' low-carbon preference leads to more market demand, the whole supply chain's profits still decline with the increase in carbon tax and production costs.

Proposition 1. *The port supply chain's profits satisfy*

$$U^p_{i-P} \geq U^p_{i-N} \geq U^p_{i-S}; U^s_{i-P} \leq U^s_{i-N} \leq U^s_{i-S}; U^{sc}_{i-N} \geq U^{sc}_{i-P} = U^{sc}_{i-S}$$

Proposition 1 shows that being a leader always earns more profit than acting as a follower. In the Nash game, the supply chain's profit is the highest; thus, equality of relationships should be encouraged, with the regulators' concern being the whole profits of the supply chain.

Proposition 2. *The port supply chain's total profits of different technologies satisfy the following: if* $\theta - \frac{\gamma}{b} < \frac{c_s + c_E - c_L}{e_L - e_E}$, *then*

$$U^p_{E-j} < U^p_{L-j}, U^s_{E-j} < U^s_{L-j} \text{ and } U^{sc}_{E-j} < U^{sc}_{L-j}$$

Proposition 2 indicates, given a certain customer's low-carbon preference, that LSFO is the most suitable option for a low carbon tax ($\theta < \frac{c_s + c_E - c_L}{e_L - e_E} + \frac{\gamma}{b}$) in terms of the supply chain's profitability. Otherwise, SP is the preferable option. It should be noted, however, that an exceptionally high carbon tax would lead to a lack of demand in the market. Consequently, regardless of the technology implemented, the gains from the supply chain are nearly negligible.

In the situation that the carbon tax is certain, LSFO is the most suitable option for a high customer low-carbon preference $\left(\gamma > b\theta - b\frac{(c_s + c_E - c_L)}{e_L - e_E}\right)$, in terms of the supply chain's profitability. Otherwise, SP is the preferable option.

5.2. Carbon Emissions Analysis

In this section, we investigate how carbon emissions are affected by operational costs and carbon taxes in the port supply chain.

Lemma 2. T_{E-j} *decreases in* $c_E, c_t, c_s; \theta$, *increases in* γ; T_{L-j} *decreases in* c_t, c_L; *and* θ *increases in* γ.

As costs and carbon taxes increase, so does the cost of services of the supply chain, leading to a reduction in market demand and total carbon emissions.

Proposition 3. *The port supply chain's carbon emissions satisfy*

$$T_{i-N} \geq T_{i-S} = T_{i-P}$$

Proposition 3 implies that when the port and shipping company are on an equal footing, they are likely to provide lower prices to draw in more customers and broaden the market, thus leading to an increase in emissions. The leader in a Stackelberg game always limits the motivation of the other to offer services, which in turn induces carbon emission reduction. Proposition 1 suggests that the Nash game yields the highest total supply chain profits. Nevertheless, when regulators aim for carbon emission reduction, the Stackelberg approach is favored.

Proposition 4. *The total carbon emissions of the supply chain with the adoption of different technologies satisfy*

If $\theta - \gamma\frac{(e + e_E + e_L)}{b(e_E + e_L)} < \frac{ae_E - bc_E e_E - bc_s e_E - bc_t e_E - ae_L + bc_L e_L + bc_t e_L}{be_E{}^2 - be_L{}^2}$, *then* $T_{E-j} < T_{L-j}$ *otherwise,* $T_{E-j} \geq T_{L-j}$.

In the situation that the customer's low-carbon preference is certain, in the port supply chain, if the carbon tax is minimal and the optimal profits are almost unaffected by carbon emissions, then LSFO is the more desirable option. In this situation, the low cost of services has caused a surge in demand, coupled with the higher unit carbon emissions, and the utilization of LSFO yields a greater amount of carbon emissions compared to the use of SP. As the carbon tax rises, Proposition 2 suggests that SP is the more cost-effective option due to its overall cost advantage. Even though using SP has lower carbon emissions per unit, if the carbon tax is increased to a certain level, the total carbon emissions from using SP could surpass those of using LSFO due to the availability of more services in the supply chain. Importantly, though, there exists a maximum limit on the carbon tax, guaranteeing favorable market demand. The upper limit may be surpassed by the threshold

in Proposition 4, leading to $T_{E-j} < T_{L-j}$ under all possible values of θ and γ. Yet even then, the gap between emissions of the two technologies is still decreasing in θ.

In the case that the carbon tax is certain, if the customers' low-carbon preferences are high, the market demand for greener SP increases, and the carbon emissions of implementing SP would exceed those of using LSFO.

In the situation of low carbon tax and low consumers' low-carbon preference, the supply chain's optimal profit is basically not affected by the two factors. From the perspective of overall cost, LSFO becomes the first choice. But LSFO with higher unit carbon emissions produces more carbon emissions than SP. As carbon tax increases, the advantage of using SP in terms of cost becomes more and more obvious. If the carbon tax and the customers' low-carbon preference attain certain conditions, the overall emissions of adopting SP may exceed those of using LSFO due to the fact that the port and ship may tend to provide more services. It should be noted that the value of the carbon tax cannot exceed a threshold to guarantee positive demand. This finding provides valuable insights to develop effective carbon tax policies and emission reduction management measures.

5.3. The Analysis of Social Welfare

We examine how channel power structures, operational costs, and carbon taxes affect social welfare in this section.

Proposition 5. *In different power structures, the supply chain's social benefits satisfy*

$$\begin{gathered} If\, p_c < \tfrac{5}{7be_i^2}, then\ sw_{i-P} = sw_{i-S} < sw_{i-N}, \\ otherwise\ sw_{i-P} = sw_{i-S} \geq sw_{i-N} \end{gathered}$$

Propositions 1 and 3 indicate that the Nash game yields the highest profits and produces the most carbon emissions. When the environmental concern is kept low, the detrimental consequences of emissions on social welfare are minimal. Consequently, there is the greatest level of social welfare in the Nash game. As the environmental concern rises, the detrimental consequences of emissions on social welfare become more pronounced. When environmental concerns rise to a threshold, the social benefits of the Stackelberg game exceed those of the Nash game. This suggests that regulators should incentivize certain channel power structures to maximize social welfare.

Proposition 6. *The social welfare under various technologies satisfies*

$$\begin{gathered} If\ p_c < \tfrac{2\rho}{\omega}, then sw_{E-j} < sw_{L-j} (j = P, S, N), \\ If\ \tfrac{2\rho}{\omega} < p_c < \tfrac{3\rho}{\omega}, then sw_{E-N} > sw_{L-N}, sw_{E-j} < sw_{L-j} (j = P, S), \\ If\ p_c > \tfrac{3\rho}{\omega}, then sw_{E-j} > sw_{L-j}. \end{gathered}$$

where

$$\begin{aligned} \rho = {} & 2b\gamma\theta c_L e_e + 2b\gamma\theta c_t e_e - 2b\gamma\theta e_L^2 + \gamma^2\theta^2 e_e^2 + 2a\gamma e_L - 2b\gamma c_t e_L + 2e\gamma^2 e_L \\ & + 2b\gamma\theta^2 e_e e_L - 2a\gamma\theta e_e - 2e\gamma^2\theta e_e - 2b\gamma c_L e_L - \gamma^2 e_L^2 \end{aligned}$$

$$\begin{aligned} \omega = {} & 2abc_L e_e^2 \quad + 2abc_t e_e^2 + 2be\gamma\theta e_e^2 e_L - a^2 e_e^2 - 2ae\gamma e_e^2 + 2be\gamma c_L e_e^2 - e^2\gamma^2 e_e^2 \\ & - b^2 c_L^2 e_e^2 + 2ab\theta e_e^2 e_L - 2b^2 c_L c_t e_e^2 - b^2 c_t^2 e_e^2 + 2be\gamma c_t e_e^2 + 2a\gamma\theta e_e^3 \\ & + 2e\gamma^2\theta e_e^3 - 2b\gamma\theta c_L e_e^3 - 2b\gamma\theta c_t e_e^3 - \gamma^2\theta^2 e_e^4 - 2b^2\theta c_L e_e^2 e_L \\ & - 2b^2\theta c_t e_e^2 e_L - 2b\gamma\theta^2 e_e^3 e_L + a^2 e_L^2 + 2ae\gamma e_L^2 + e^2\gamma^2 e_L^2 - 2abc_L e_L^2 \\ & - 2be\gamma c_L e_L^2 + b^2 c_L^2 e_L^2 - 2abc_t e_L^2 - 2be\gamma c_t e_L^2 + 2b^2 c_L c_t e_L^2 + b^2 c_t^2 e_L^2 \\ & - b^2\theta^2 e_e^2 e_L^2 - 2a\gamma e_L^3 - 2e\gamma^2 e_L^3 - 2ab\theta e_L^3 - 2be\gamma\theta e_L^3 + 2b\gamma c_L e_L^3 \\ & + 2b^2\theta c_L e_L^3 + 2b\gamma c_t e_L^3 + 2b^2\theta c_t e_L^3 + \gamma^2 e_L^4 + 2b\gamma\theta e_L^4 + b^2\theta^2 e_L^4 \end{aligned}$$

Proposition 6 examines the impact of two distinct technologies on social benefits. The findings indicate that when the carbon emission cost is reduced, the social benefits associated with SP are greater than those of LSFO. Nevertheless, due to the complexity of the threshold expression, we will illustrate this situation with a numerical example.

5.4. Managerial Insights and Practical Implications

In this section, some management insights and practical inspiration are proposed based on the above results to provide some countermeasures and suggestions to relevant stakeholders.

5.4.1. The Emission Reduction Technology Decision Making for the Port and Ship

According to Proposition 1, as to the market relationship between port and ship, the profit obtained by one part acting as a leader is more than that obtained being a follower in the game model. Meanwhile, the total profit of the port supply chain in the Nash game is greater than that in the Stackelberg game. Therefore, ports or ships either strive to be in a dominant position or strive to be in an equal position in the game. In practice, due to the characteristics of port resources, ports are often in the leading position, so shipping companies tend to carry out alliances. For example, the world's three largest alliances (THE Alliance, 2M, and Ocean Alliance) began operations in April 2017, and all eight of the world's largest container shipping companies are included. These three alliances carry about 80% of the total container trade and about 95% of the cargo on east–west trade routes [45]. Such alliances greatly increase market share and improve coordination with regional transportation departments and ports.

The choice of technologies depended on the carbon tax policy and the low-carbon preferences of customers. According to Proposition 2, if the carbon tax is low, using LSFO for the ship would be preferred. Under the high carbon taxes, ports and ships should choose to use SP when customers' low-carbon preferences are low. With the growth of customers' low-carbon preferences, the advantages of using SP do not exist, and more profits can be obtained by choosing to use LSFO.

5.4.2. The Policies and Management Measures for Government

At present, the Chinese government is vigorously promoting shore power. According to Proposition 2, levying a higher carbon tax is conducive to the promotion of shore power. When the carbon tax is high, the port and ship will be more active in using shore power, but the carbon tax has a ceiling; beyond this ceiling, port and shipping enterprises will not be willing to offer services, because there is no profit at this time. The customers' low-carbon preferences may lead to the expansion of market demand, which may lead to more carbon taxes. Therefore, a higher preference of customers for low carbon is not necessarily conducive to the promotion of shore power. However, according to Proposition 4, under a high carbon tax, although the unit carbon emission of SP is relatively low, the total emissions from the use of SP may exceed the total emissions from the use of LSFO, because the port and shipping enterprises may provide more services at this time. Therefore, in order to reduce the total emissions, the carbon tax should be in the appropriate middle range. At this time, enterprises will choose SP, and the total emissions of using SP will also be lower than those of using LSFO.

In terms of market competition relations, if the government aims to maximize the profits of enterprises, according to Propositions 1 and 3, then the government administration should support the Nash game that encourages ports and ships to be on an equal footing. If regulators are concerned about emissions control, then the Stackelberg game in which one party is in a dominant position should be encouraged. If the government focuses on the maximization of social welfare, then when the social attention is large, the Stackelberg game should be better, and the Nash game should be supported when the social attention is small.

6. Numerical Example Analysis

A few numerical illustrations are presented in this section to provide a more comprehensive explanation of the quotes and propositions obtained above. The parameters are [31] $a = 200, b = 3.5, c_E = 2.8, c_t = 3.6, c_s = 0.6, e_E = 4.2,\ e_L = 4.9,\ e = 5.4, c_L = 1.6$.

6.1. Influence of θ and γ on Profits

The profits of the port and ship are shown in Figures 2 and 3, respectively. In line with Proposition 1, being a leader always achieves more profits than acting as a follower, while the Nash game has the greatest overall earnings compared to the Stackelberg game, as depicted in Figure 4.

The comparison of profits between the two technologies is complicated. In cases where carbon taxes and customers' low-carbon preferences are minimal, the emission penalties are negligible, and using LSFO is the suitable option. As the carbon tax increases, SP is expected to become the preferred option. As the carbon tax rises to a certain level, the supply chain will be unable to generate any additional revenue due to the stringent emission penalties, resulting in the supply chain's financial loss.

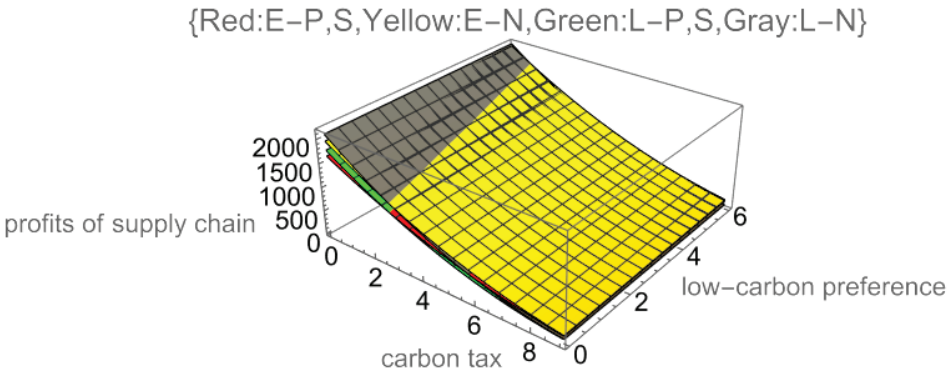


Figure 2. The port supply chain's total profits.

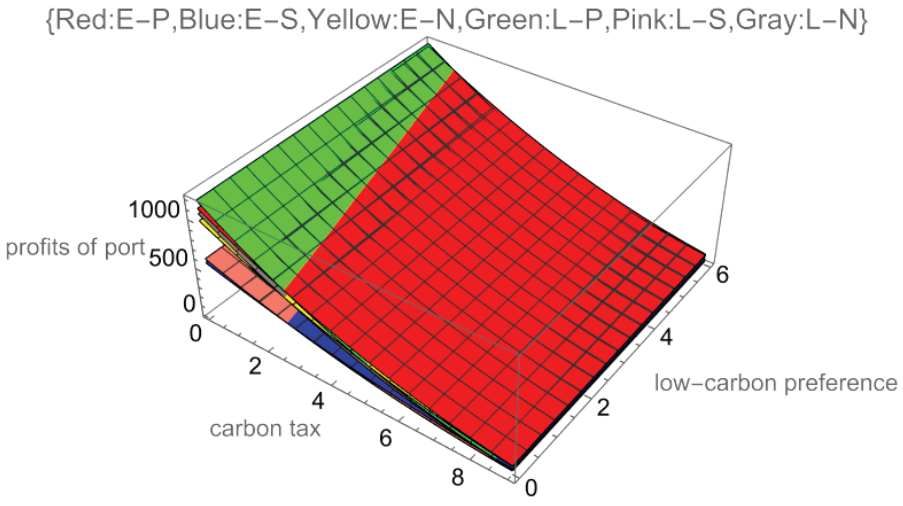


Figure 3. The profits of the port.

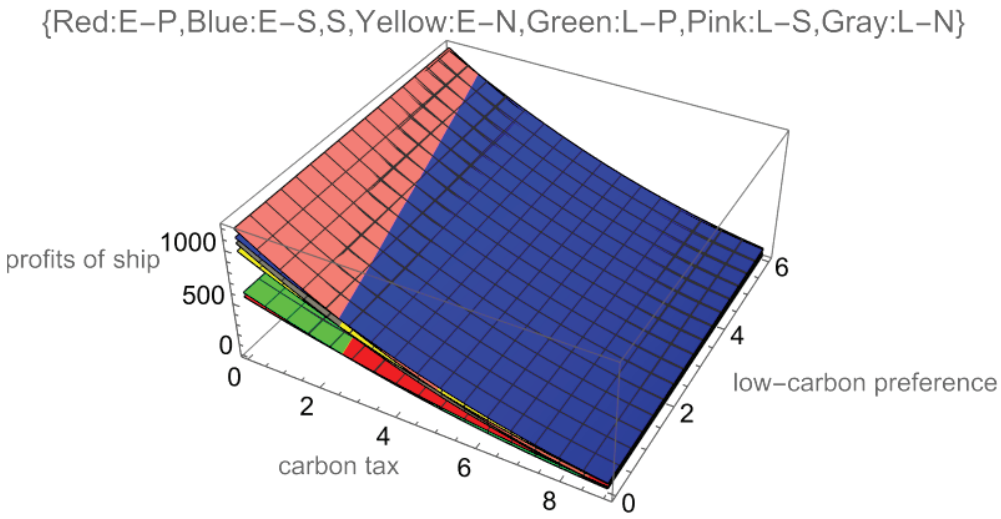


Figure 4. The ship's profits.

6.2. Influence of θ and γ on Carbon Emissions

As illustrated in Figure 5, the Nash game has the most carbon emissions as described in Proposition 4. When θ and γ are minimal, the total carbon emissions with the adoption of LSFO is more than those of using SP. With the carbon tax increasing, the total carbon emissions of SP are greater compared to LSFO. As the carbon tax rises to a considerable degree, the supply chain's services of all three power structures become significantly lower, and the emissions under different power structures also approach one another. Given the minimal carbon emission penalty associated with low carbon taxes, adopting LSFO can earn more profits due to its cost-effectiveness. Low service costs will lead to a surge in demand and, combined with high carbon emissions per unit, using LSFO produces more emissions compared to the adoption of SP.

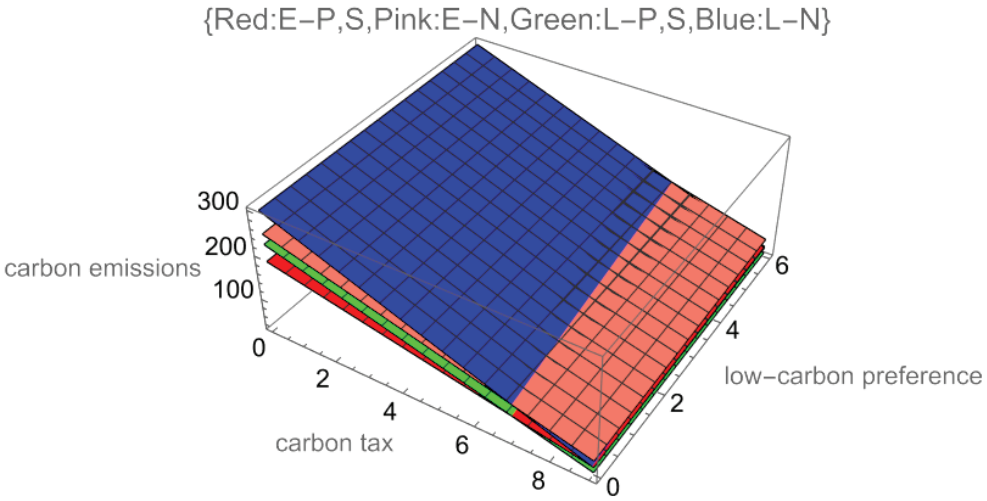


Figure 5. Total carbon emissions.

6.3. Influence of θ and γ on Social Welfare

Figure 6 illustrates how carbon tax, customers' low-carbon preferences, and environmental concerns impact social welfare. Given a certain technology, if environmental concern is low, the Nash game would be the most suitable option from a social welfare point of view. In any other cases, the Stackelberg games are expected to be superior, as indicated in Proposition 4.

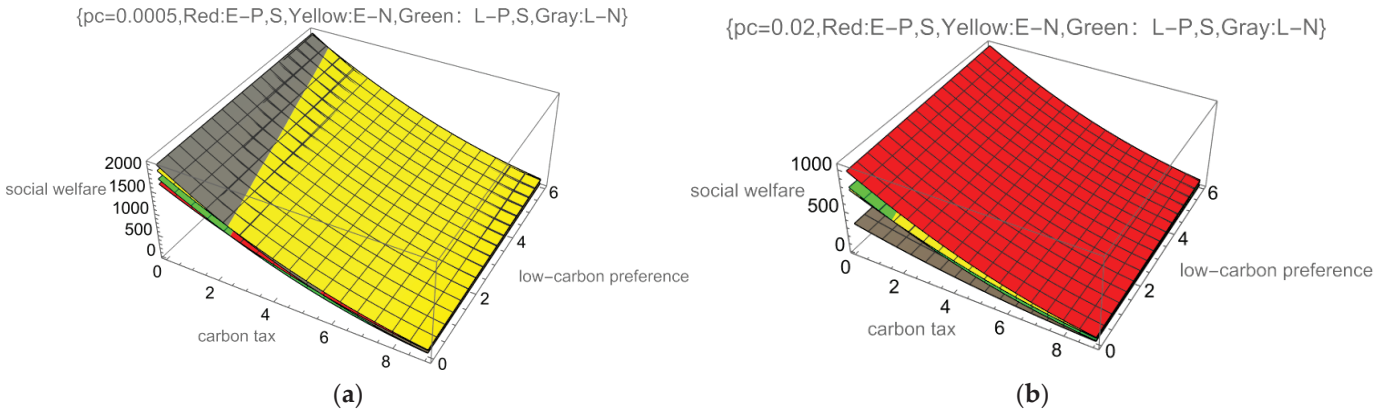


Figure 6. The social welfare.

Comparing the social welfare between the two technologies is a complex task, so the Nash game is regarded as a representative instance. Two different environmental concern cases are shown in Figure 6. The low-environmental consciousness situation (pc $= 0.0005$) is shown as Figure 6a, and Figure 6b describes the high-environmental consciousness situation (pc $= 0.02$). Proposition 2 suggests that when θ and γ are both low, the economic benefit of the supply chain implementing LSFO is greater than that of using SP. However, the total carbon emissions with the adoption of LSFO are more than that of SP according to Proposition 4. Under the low environmental concern, implementing LSFO is more suitable from the viewpoint of social welfare. As the environmental consciousness grows, the disutility of carbon emissions increases. With higher unit carbon emissions, the social welfare of the adoption of LSFO drops more significantly than that of SP. Consequently, SP with more social welfare is favored over LSFO when it comes to high environmental concerns.

7. Conclusions

This paper discusses the decision making of emission control technology (SP and LSFO) in the port supply chain. The model is constructed in three power structures using game theory. By analyzing the equilibrium results, some results are obtained.

Firstly, for maximizing the supply chain's profit, a balanced supply chain on both sides is better than the Stackelberg game. When the customers' low-carbon preferences and carbon tax are low, LSFO is the preferred option; otherwise, SP should be the more suitable choice.

Secondly, for controlling emissions, the Stackelberg game consistently produces fewer carbon emissions than the Nash game. When the customers' low-carbon preferences and carbon tax are low, LSFO is the better option than SP; otherwise, SP should be chosen.

Finally, from the point of social welfare, the Nash game has the highest social welfare when the environmental concern is low, whereas the Stackelberg games have more social welfare when the environmental concern is high. The technical comparison, however, is more complex. Generally speaking, when customers' low-carbon preferences, carbon tax, and environmental concerns are all low or high, LSFO is preferred. Otherwise, SP would be better.

Most of the existing studies on emission control in port areas focus on the emission reduction efforts and emission reduction investment of port and shipping enterprises under government policies, such as subsidies, carbon tax, and carbon trading mechanisms. However, there is little literature on quantitative analysis and comparison of the two technologies (SP and LSFO). Yang et al. [21] compared SP and LSFO from the perspective of sustainable development under the carbon trading mechanism, but they did not take into account the influence of customers' low-carbon preferences. This study shows that under customers' low-carbon preferences, a higher carbon tax may not necessarily make the supply chain choose the cleaner technology SP, because the expansion of market demand led by customers' low-carbon preferences may promote the supply chain to use general

clean technology (LSFO) to obtain more profits, which will inevitably lead to more carbon emissions. Also, in terms of social welfare, when customers' low-carbon preferences, carbon tax, and environmental concerns are all low or high, the social welfare of using LSFO is greater than that of using SP. Otherwise, the social benefits of using SP are more substantial. The results obtained can further enrich the existing research. Meanwhile, these can give ports and ships valuable information to select suitable emission reduction technologies, and it is also very useful for regulars to formulate appropriate policies from the view of their own interests.

The decisions made by ports and ships regarding emission reduction are influenced by various factors, including the ports' natural resources, ship characteristics, the competitive dynamics among ports, the vessels' competitive relationship, and market unpredictability. So the process is more intricate. Consequently, exploring a more realistic port and shipping supply chain that includes multiple ports and ships is the future research direction. In addition, the game-theoretical models could be expanded to include subsidies or additional strategic interactions, and the artificial neural network [46] could also be used as an optimization method in the next study.

Author Contributions: Conceptualization, H.Z.; methodology, H.Z.; software, Y.Z.; validation, H.Z.; formal analysis, H.Z. and Y.Z.; investigation, Y.Z.; resources, H.Z.; writing—original draft preparation, Y.Z.; writing—review and editing, Y.Z.; funding acquisition, H.Z. All authors have read and agreed to the published version of the manuscript.

Funding: The research is supported by the Natural Science Foundation of Guangdong Province of China, grant number 2021A1515110213, and the Key Construction Discipline Scientific Research Ability Enhancement Project of Guangdong Province, grant number 2022ZDJS094.

Data Availability Statement: All data generated or analyzed during this study are included in this article.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

1. Solutions with adoption of SP

(1) In ship-leading Stackelberg, the port's profit is $U^p = (m - c_E)q - qe_E\theta$

Solving $\frac{\partial U^p}{\partial m} = 0$, $m^* = \frac{a+bc_s-bw+\gamma(e-e_E)+be_E\theta}{2b}$, substituting m* into U^s, and solving $\frac{\partial U^s}{\partial w} = 0, w_{E-S} = \frac{a+\gamma(e-e_E)+b(c_E-c_s+c_t-e_E\theta)}{2b}$; thus,

$$m_{E-S} = \frac{a+\gamma(e-e_E)-b(c_E-3c_s+c_t-3e_E\theta)}{4b},$$

$$U^p_{E-S} = \frac{[a+\gamma(e-e_E)-b(c_E+c_s+c_t+e_E\theta)]^2}{16b}, U^s_{E-S} = \frac{[a+\gamma(e-e_E)-b(c_E+c_s+c_t+e_E\theta)]^2}{8b}$$

$$U^{sc}_{E-S} = \frac{3[a+\gamma(e-e_E)-b(c_E+c_s+c_t+e_E\theta)]^2}{16b},$$

$$T_{E-S} = \frac{e_E(a+\gamma(e-e_E)-b(c_E+c_s+c_t+e_E\theta))}{4}$$

$$sw_{E-S} = \frac{(3-be_E{}^2p_c)(a+\gamma(e-e_E)-b(c_E+c_s+c_t+e_E\theta))^2}{16b}$$

(2) In the Nash game, the ship's profit is $U^s = (w - c_t - c_s)[a - b(m + w)] - qe_E\theta$. Solving $\frac{\partial U^s}{\partial w} = 0, w^* = \frac{a+bc_E+bc_t-bm+\gamma e-\gamma e_E}{2b}$, the port's profit is $U^p = (m - c_E)q - qe_E\theta$. Solving $\frac{\partial U^p}{\partial m} = 0, m^* = \frac{a+bc_s-bw+\gamma(e-e_E)+be_E\theta}{2b}$

Now, solving above two equations simultaneously,

$$m_{E-N} = \frac{a+\gamma(e-e_E)-b(c_E-2c_s+c_t-2e_E\theta)}{3b}, w_{E-N} = \frac{a+\gamma(e-e_E)+b(2c_E-c_s+2c_t-e_E\theta)}{3b},$$

Thus,

$$U^p_{E-N} = \frac{(a+\gamma(e-e_E)-b(c_E+c_s+c_t+e_E\theta))^2}{9b},$$

$$U^s_{E-N} = \frac{(a+\gamma(e-e_E)-b(c_E+c_s+c_t+e_E\theta))^2}{9b}$$

$$U^{sc}_{E-N} = \frac{2(a+\gamma(e-e_E)-b(c_E+c_s+c_t+e_E\theta))^2}{9b},$$

$$T_{E-N} = \frac{e_E(a+\gamma(e-e_E)-b(c_E+c_s+c_t+e_E\theta))}{9},$$

$$sw_{E-N} = \frac{(2-be_E{}^2p_c)(a+\gamma(e-e_E)-b(c_E+c_s+c_t+e_E\theta))^2}{9b}.$$

(3) In the port-leader game, the ship's profit is $U^s = (w-c_t-c_s)[a-b(m+w)] - qe_E\theta$. Solving $\frac{\partial U^s}{\partial w} = 0, w^* = \frac{a+bc_E+bc_t-bm+\gamma e-\gamma e_E}{2b}$, substituting w^* into U^p, and solving $\frac{\partial U^p}{\partial m} = 0, m_{E-P} = \frac{a+\gamma(e-e_E)+b(c_E-c_s+c_t-e_E\theta)}{2b}$; thus,

$$w_{E-P} = \frac{a+\gamma(e-e_E)+b(3c_E-c_s+3c_t-e_E\theta)}{4b}$$

$$U^p_{E-P} = \frac{[a+\gamma(e-e_E)-b(c_E+c_s+c_t+e_E\theta)]^2}{8b}, U^s_{E-P} = \frac{[a+\gamma(e-e_E)-b(c_E+c_s+c_t+e_E\theta)]^2}{16b}$$

$$U^{sc}_{E-P} = \frac{3[a+\gamma(e-e_E)-b(c_E+c_s+c_t+e_E\theta)]^2}{16b}, T_{E-P} = \frac{e_E(a+\gamma(e-e_E)-b(c_E+c_s+c_t+e_E\theta))}{4}$$

$$sw_{E-P} = \frac{(3-be_E{}^2p_c)(a+\gamma(e-e_E)-b(c_E+c_s+c_t+e_E\theta))^2}{16b}.$$

2. Solutions with adoption of LSFO

When LSFO is implemented, the results can be easily obtained by the same method used in the situation of adoption of SP. In order to avoid repetition, the solving process is omitted.

3. Proofs

Proof of Lemma 1.

(1) when $j = P$: $\frac{\partial^2 U^{sc}_{i-P}}{\partial^2 c_i} = \frac{\partial^2 U^{sc}_{i-P}}{\partial^2 c_t} = \frac{3b}{8} > 0, \frac{\partial^2 U^{sc}_{E-P}}{\partial^2 c_s} = \frac{3b}{8} > 0, \frac{\partial^2 U^{sc}_{i-P}}{\partial^2 \gamma} = \frac{3(e_i-e)^2}{8b} > 0, and\ \frac{\partial^2 U^{sc}_{i-P}}{\partial^2 e_i} = \frac{3(b+\theta)^2}{8} > 0.$

(2) when $j = S$: $\frac{\partial^2 U^{sc}_{i-S}}{\partial^2 c_i} = \frac{\partial^2 U^{sc}_{i-S}}{\partial^2 c_t} = \frac{3b}{8} > 0, \frac{\partial^2 U^{sc}_{E-S}}{\partial^2 c_s} = \frac{3b}{8} > 0, \frac{\partial^2 U^{sc}_{i-S}}{\partial^2 \gamma} = \frac{3(e_i-e)^2}{8b} > 0, and\ \frac{\partial^2 U^{sc}_{i-S}}{\partial^2 e_i} = \frac{3(b+\theta)^2}{8} > 0.$

(3) when $j = N$: $\frac{\partial^2 U^{sc}_{i-N}}{\partial^2 c_i} = \frac{\partial^2 U^{sc}_{i-N}}{\partial^2 c_t} = \frac{4b}{9} > 0, \frac{\partial^2 U^{sc}_{E-N}}{\partial^2 c_s} = \frac{4b}{9} > 0, \frac{\partial^2 U^{sc}_{i-N}}{\partial^2 \gamma} = \frac{4(e_i-e)^2}{9} > 0, and\ \frac{\partial^2 U^{sc}_{i-N}}{\partial^2 e_i} = \frac{4(b+\theta)^2}{9} > 0.$

□

Proof of Proposition 1. For the port,

$$U^p_{E-P} - U^p_{E-S} = \frac{A^2}{16b} \geq 0, U^p_{E-P} - U^p_{E-N} = \frac{A^2}{72b} \geq 0, U^p_{E-S} - U^p_{E-N} = -\frac{7A^2}{144b} \leq 0$$

$$U^p_{L-P} - U^p_{L-S} = \frac{B^2}{16b} \geq 0, U^p_{L-P} - U^p_{L-N} = \frac{B^2}{72b} \geq 0, U^p_{L-S} - U^p_{L-N} = -\frac{7B^2}{144b} \leq 0$$

Therefore, $U^p_{i-P} \geq U^p_{i-N} \geq U^p_{i-S}$.
For the ship,

$$U^s_{E-P} - U^s_{E-S} = -\frac{A^2}{16b} \leq 0, U^s_{E-P} - U^s_{E-N} = -\frac{7A^2}{144b} \leq 0, U^s_{E-S} - U^s_{E-N} = \frac{A^2}{72b} \geq 0$$

$$U^s_{L-P} - U^s_{L-S} = -\frac{B^2}{16b} \leq 0, U^s_{L-P} - U^s_{L-N} = -\frac{7B^2}{144b} \leq 0, U^s_{L-S} - U^s_{L-N} = \frac{B^2}{72b} \geq 0$$

Therefore, $U^s_{i-P} \leq U^s_{i-N} \leq U^s_{i-S}$.
For the supply chain,

$$U^{sc}_{E-P} - U^{sc}_{E-S} = 0, U^{sc}_{E-P} - U^{sc}_{E-N} = \frac{5(a+\gamma(e-e_E)-b(c_E+c_s+c_t+e_E\theta))^2}{144b} \geq 0$$

$$U^{sc}_{L-P} - U^{sc}_{L-S} = 0, U^{sc}_{L-P} - U^{sc}_{L-N} = \frac{5(a+\gamma(e-e_L)-b(c_L+c_t+e_L\theta))^2}{144b} \geq 0$$

Therefore, $U^{sc}_{i-N} \geq U^{sc}_{i-P} = U^{sc}_{i-S}$. □

Proof of Proposition 2.

$$U^s_{E-j} - U^s_{L-j} = \beta(2a - 2\gamma e + \gamma(e_E + e_L) - b(c_E + c_L + c_s + 2c_t + e_L\theta + e_E\theta))(\gamma(e_E - e_L) + c_L - c_E - c_s + \theta e_L - \theta e_E)$$

$$U^p_{E-j} - U^p_{L-j} = \mu(2a - 2\gamma e + \gamma(e_E + e_L) - b(c_E + c_L + c_s + 2c_t + e_L\theta + e_E\theta))(\gamma(e_E - e_L) + c_L - c_E - c_s + \theta e_L - \theta e_E)$$

$$U^{sc}_{E-j} - U^{sc}_{L-j} = \varphi(2a - 2\gamma e + \gamma(e_E + e_L) - b(c_E + c_L + c_s + 2c_t + e_L\theta + e_E\theta))(\gamma(e_E - e_L) + c_L - c_E - c_s + \theta e_L - \theta e_E)$$

where β, μ, φ are constants.

Since $q > 0$, θ must satisfy $\theta < \frac{a+\gamma(e-e_E)-b(c_E+c_s+c_t)}{be_E}, \theta < \frac{a+\gamma(e-e_L)-b(c_L+c_t)}{be_L}$, if $0 < \theta < \frac{c_s+c_E-c_L}{e_L-e_E}$, then

$$U^p_{E-j} < U^p_{L-j}, U^s_{E-j} < U^s_{L-j} \text{ and } U^{sc}_{E-j} < U^{sc}_{L-j}$$

□

Proof of Lemma 2.

(1) when $j = P, S$: $\frac{\partial T_{i-j}}{\partial c_i} = \frac{\partial T_{i-j}}{\partial c_t} = -\frac{b}{4}e_i < 0, \frac{\partial T_{E-j}}{\partial c_s} = -\frac{b}{4}e_E < 0, \frac{\partial T_{i-j}}{\partial \theta} = -\frac{b}{4}e_i^2 < 0, \frac{\partial T_{i-j}}{\partial \gamma} = \frac{(e-e_i)e_i}{4} > 0$ *and* $\frac{\partial^2 T_{i-j}}{\partial^2 e_i} = -\frac{(\gamma+b\theta)}{2} < 0$.

(2) when $j = N$: $\frac{\partial T_{i-N}}{\partial c_i} = \frac{\partial T_{i-N}}{\partial c_t} = -\frac{be_i}{3} < 0, \frac{\partial T_{E-N}}{\partial c_s} = -\frac{be_E}{3} < 0, \frac{\partial T_{i-N}}{\partial \theta} = -\frac{be_i^2}{3} < 0, \frac{\partial T_{i-N}}{\partial \gamma} = \frac{(e-e_i)e_i}{3} > 0$ *and* $\frac{\partial^2 T_{i-N}}{\partial^2 e_i} = -\frac{2(\gamma+b\theta)}{3} < 0$.

□

Proof of Proposition 3. Comparing the supply chain's carbon emissions under different power structures,

$$\begin{aligned} T_{i-P} - T_{i-S} = 0, \quad & T_{L-P} - T_{L-N} \\ & = -\tfrac{e_L}{12}(a + \gamma(e - e_L) \\ & \quad -b(c_L + c_t + e_L\theta)) \end{aligned}$$

$$\begin{aligned} T_{E-P} - T_{E-N} = \quad & -\tfrac{e_E}{12}(a + \gamma(e - e_E) \\ & -b(c_E + c_t + c_s + e_E\theta)) \end{aligned}$$

Since $q > 0$, θ must satisfy $\theta < \frac{a+\gamma(e-e_L)-b(c_L+c_t)}{be_L}$, $\theta < \frac{a+\gamma(e-e_E)-b(c_E+c_s+c_t)}{be_E}$; therefore, $T_{L-N} > T_{i-P} = T_{i-S}$. □

Proof of Proposition 4. Comparing carbon emissions of two technologies,

$$T_{E-j} - T_{L-j} = \frac{e_E(a + \gamma(e - e_E) - b(c_E + c_t + c_s + e_E\theta)) - e_L(a + \gamma(e - e_L) - b(c_L + c_t + e_L\theta))}{4}$$

If $0 < \theta < \frac{ae_E - bc_e e_E - bc_s e_E - bc_t e_E - ae_L + bc_L e_L + bc_t e_L}{be_E{}^2 - be_L{}^2}$, then $T_{E-j} < T_{L-j}$ otherwise, $T_{E-j} \geq T_{L-j}$. □

Proof of Proposition 5. The social welfare of the supply chain is compared as follows:

$$sw_{i-P} - sw_{i-S} = 0 sw_{i-P} - sw_{i-N} = \frac{\left(3 - be_i{}^2 p_c\right)\left(a - b(c_i + c_s + c_t + e_i\theta)\right)^2}{16b} - \frac{\left(2 - be_i{}^2 p_c\right)\left(a - b(c_i + c_s + c_t + e_i\theta)\right)^2}{9b}$$

Therefore, if $p_c < \frac{5}{7be_i{}^2}$, then $sw_{i-P} = sw_{i-S} \leq sw_{i-N}$; otherwise, $sw_{i-N} \leq sw_{i-P} = sw_{i-S}$. □

Proof of Proposition 6.

$$\begin{aligned} sw_{E-j} - sw_{L-j} \quad & = \tfrac{\left(3-be_E{}^2\beta\right)(a+\gamma e-\gamma e_E-bc_E-bc_s-bc_t-be_E\theta)^2}{16b} \\ & \quad -\tfrac{\left(3-be_L{}^2\beta\right)(a+\gamma e-\gamma e_L-bc_L-bc_t-be_L\theta)^2}{16b} (j = P, S) \end{aligned}$$

$$\begin{aligned} sw_{E-N} - sw_{L-N} = \quad & \tfrac{\left(2-be_E{}^2\beta\right)(a+\gamma e-\gamma e_E-bc_E-bc_s-bc_t-be_E\theta)^2}{16b} \\ & -\tfrac{\left(2-be_L{}^2\beta\right)(a+\gamma e-\gamma e_L-bc_L-bc_t-be_L\theta)^2}{16b} \end{aligned}$$

Therefore,
If $p_c < \frac{2\rho}{\omega}$, then $sw_{E-j} < sw_{L-j} (j = P, S, N)$,
If $\frac{2\rho}{\omega} < p_c < \frac{3\rho}{\omega}$, then $sw_{E-N} > sw_{L-N}$, $sw_{E-j} < sw_{L-j} (j = P, S)$,
If $p_c > \frac{3\rho}{\omega}$, then $sw_{E-j} > sw_{L-j}$.

$$\begin{aligned} \rho = 2b\gamma\theta c_L e_e + 2b\gamma\theta c_t e_e \quad & -2a\gamma\theta e_e + \gamma^2\theta^2 e_e^2 - 2e\gamma^2\theta e_e + 2a\gamma e_L + 2e\gamma^2 e_L - 2b\gamma c_L e_L \\ & -2b\gamma c_t e_L + 2b\gamma\theta^2 e_e e_L - 2b\gamma\theta e_L^2 - \gamma^2 e_L^2 \end{aligned}$$

$$\begin{aligned} \omega = 2abc_L e_e^2 - a^2 e_e^2 \quad & -e^2\gamma^2 e_e^2 + 2be\gamma c_L e_e^2 - b^2 c_L^2 e_e^2 - 2ae\gamma e_e^2 + 2abc_t e_e^2 + 2be\gamma c_t e_e^2 \\ & -2b^2 c_L c_t e_e^2 - b^2 c_t^2 e_e^2 + 2a\gamma\theta e_e^3 + 2e\gamma^2\theta e_e^3 - 2b\gamma\theta c_L e_e^3 - 2b\gamma\theta c_t e_e^3 \\ & -\gamma^2\theta^2 e_e^4 + 2ab\theta e_e^2 e_L + 2be\gamma\theta e_e^2 e_L - 2b^2\theta c_L e_e^2 e_L - 2b^2\theta c_t e_e^2 e_L \\ & -2b\gamma\theta^2 e_e^3 e_L + a^2 e_L^2 + 2ae\gamma e_L^2 + e^2\gamma^2 e_L^2 - 2abc_L e_L^2 - 2be\gamma c_L e_L^2 \\ & +b^2 c_L^2 e_L^2 - 2abc_t e_L^2 - 2be\gamma c_t e_L^2 + 2b^2 c_L c_t e_L^2 + b^2 c_t^2 e_L^2 - b^2\theta^2 e_e^2 e_L^2 \\ & -2a\gamma e_L^3 - 2e\gamma^2 e_L^3 - 2ab\theta e_L^3 - 2be\gamma\theta e_L^3 + 2b\gamma c_L e_L^3 + 2b^2\theta c_L e_L^3 \\ & +2b\gamma c_t e_L^3 + 2b^2\theta c_t e_L^3 + \gamma^2 e_L^4 + 2b\gamma\theta e_L^4 + b^2\theta^2 e_L^4 \end{aligned}$$

□

References

1. Li, J.; Liu, X.; Jiang, B. An Exploratory study on low-carbon ports development strategy in China. *Asian J. Shipp. Logist.* **2011**, *27*, 91–111. [CrossRef]
2. Charlier, R.H. Life cycle of ports. *Int. J. Environ. Stud.* **2013**, *70*, 594–602. [CrossRef]
3. The Harbor Circle. How Can Ports Help the Shipping Industry Reduce Carbon Emissions? Available online: https://www.163.com/dy/article/DFS02DF10519CUIJ.html (accessed on 20 April 2018).
4. Hulskotte, J.; van der Gon, H.D. Fuel consumption and associated emissions from seagoing ships at berth derived from an on-board survey. *Atmos. Environ.* **2010**, *44*, 1229–1236. [CrossRef]
5. Ministry of Transport of the People's Republic of China. Decision of the Ministry of Transport on Amending the Measures for the Administration of Shore Power for Ports and Ships. Available online: https://www.gov.cn/gongbao/content/2022/content_5669432.htm (accessed on 13 September 2021).
6. Hall, W. Assessment of CO_2 and priority pollutant reduction by installation of shoreside power. *Resour. Conserv. Recycl.* **2010**, *54*, 462–467. [CrossRef]
7. Pian, F.; Xu, L.; Chen, Y.; Lee, S.-H. Global Emission Taxes and Port Privatization Policies under International Competition. *Sustainability* **2020**, *12*, 6595. [CrossRef]
8. Zhuge, D.; Wang, S.; Zhen, L.; Laporte, G. Subsidy design in a vessel speed reduction incentive program under government policies. *Nav. Res. Logist. (NRL)* **2021**, *68*, 344–358. [CrossRef]
9. Wang, T.; Wang, X.; Meng, Q. Joint berth allocation and quay crane assignment under different carbon taxation policies. *Transp. Res. Part B Methodol.* **2018**, *117*, 18–36. [CrossRef]
10. Cui, H.; Notteboom, T. Modelling emission control taxes in port areas and port privatization levels in port competition and co-operation sub-games. *Transp. Res. Part D Transp. Environ.* **2017**, *56*, 110–128. [CrossRef]
11. Wang, Q.; Zhao, D. Cooperative strategy of carbon emissions reduction and promotion in a two-echelon supply chain. *Control Decis.* **2014**, *29*, 307–314.
12. Krass, D.; Nedorezov, T.; Ovchinnikov, A. Environmental Taxes and the Choice of Green Technology. *Prod. Oper. Manag.* **2013**, *22*, 1035–1055. [CrossRef]
13. Asadabadi, A.; Miller-Hooks, E. Optimal transportation and shoreline infrastructure investment planning under a stochastic climate future. *Transp. Res. Part B Methodol.* **2017**, *100*, 156–174. [CrossRef]
14. Liu, W.; Xie, D.; Liu, Y.; Liu, X. Service capability procurement decision in logistics service supply chain: A research under demand updating and quality guarantee. *Int. J. Prod. Res.* **2015**, *53*, 488–510. [CrossRef]
15. Liu, L.; Wang, Z.; Li, X.; Liu, Y.; Zhang, Z. An evolutionary analysis of low-carbon technology investment strategies based on the manufacturer-supplier matching game under government regulations. *Environ. Sci. Pollut. Res.* **2022**, *29*, 44597–44617. [CrossRef]
16. Cao, Z.H.; Feng, Y.; Min, J.; Ou, J. Mode selection strategy of energy performance contracting under the regulation of carbon tax policy. *Energy Environ.* **2023**, 0958305X231189184. [CrossRef]
17. Jiang, X.; Sun, L.; Wang, Y. Technology spillover and incremental cost effects of carbon transfers on optimal carbon quotas in supply chains. *J. Clean. Prod.* **2023**, *419*, 138171. [CrossRef]
18. Liu, L.; Wang, Z.; Zhang, Z. Matching-game approach for green technology investment strategies in a supply chain under environmental regulations. *Sustain. Prod. Consum.* **2021**, *28*, 371–390. [CrossRef]
19. Yu, S.; Zheng, S.; Gao, S.; Yang, J. A multi-objective decision model for investment in energy savings and emission reductions in coal mining. *Eur. J. Oper. Res.* **2017**, *260*, 335–347. [CrossRef]
20. Song, J. Impact of Government Subsidies, Competition, and Blockchain on Green Supply Chain Decisions. *Sustainability* **2023**, *15*, 3633. [CrossRef]
21. Yang, L.; Cai, Y.; Wei, Y.; Huang, S. Choice of technology for emission control in port areas: A supply chain perspective. *J. Clean. Prod.* **2019**, *240*, 118105. [CrossRef]
22. Ding, W.; Wang, Y.; Dai, L.; Hu, H. Does a carbon tax affect the feasibility of Arctic shipping? *Transp. Res. Part D Transp. Environ.* **2020**, *80*, 102257. [CrossRef]
23. Cariou, P.; Halim, R.A.; Rickard, B.J. Ship-owner response to carbon taxes: Industry and environmental implications. *Ecol. Econ.* **2023**, *212*, 107917. [CrossRef]
24. Gao, S.; Xin, X.; Li, C.; Liu, Y.; Chen, K. Container ocean shipping network design considering carbon tax and choice inertia of cargo owners. *Ocean. Coast. Manag.* **2022**, *216*, 105986. [CrossRef]
25. Song, J.; Xu, C.; Wang, C. Impacts of the carbon tax on green shipping supply chain under the port competition. *Expert Syst.* **2023**, e13229. [CrossRef]
26. Xin, X.; Wang, X.; Tian, X.; Chen, Z.; Chen, K. Green scheduling model of shuttle tanker fleet considering carbon tax and variable speed factor. *J. Clean. Prod.* **2019**, *234*, 1134–1143. [CrossRef]
27. Zhao, R.J.; Xie, Y.; Su, B.H.; Xie, X.L. Carbon emission and carbon peak time calculation method of container port. *J. Dalian Marit. Univ.* **2021**, *47*, 56–64. [CrossRef]
28. Li, X.D.; Kuang, H.B.; Hu, Y. Cooperation strategy of port and shipping company under emission control. *Syst. Eng.-Theory Pract.* **2021**, *41*, 1750–1760.
29. Chen, L.; Yip, T.L.; Mou, J. Provision of Emission Control Area and the impact on shipping route choice and ship emissions. *Transp. Res. Part D Transp. Environ.* **2018**, *58*, 280–291. [CrossRef]

30. Hu, Q.; Gu, W.; Wang, S. Optimal subsidy scheme design for promoting intermodal freight transport. *Transp. Res. Part E Logist. Transp. Rev.* **2022**, *157*, 102561. [CrossRef]
31. Liu, H.; Jin, X.; Wu, L.; Wang, X.; Fu, M.; Lv, Z.; Morawska, L.; Huang, F.; He, K. The impact of marine shipping and its DECA control on air quality in the Pearl River Delta, China. *Sci. Total Environ.* **2018**, *625*, 1476–1485. [CrossRef]
32. Cohen, M.A.; Vandenbergh, M.P. The potential role of carbon labeling in a green economy. *Energy Econ.* **2012**, *34*, S53–S63. [CrossRef]
33. Michaud, C.; Llerena, D.; Joly, I. Willingness to pay for environmental attributes of non-food agricultural products: A real choice experiment. *Eur. Rev. Agric. Econ.* **2013**, *40*, 313–329. [CrossRef]
34. Yu, S.; Hou, Q. Supply chain investment in carbon emission-reducing technology based on stochasticity and low-carbon preferences. *Complexity* **2021**, *2021*, 8881605. [CrossRef]
35. Xu, C.; Jing, Y.; Shen, B.; Zhou, Y.; Zhao, Q.Q. Cost-sharing contract design between manufacturer and dealership considering the customer low-carbon preferences. *Expert Syst. Appl.* **2023**, *213*, 118877. [CrossRef]
36. Gao, M.; Xia, L.; Xiao, Q.; Goh, M. Incentive strategies for low-carbon supply chains with information updating of customer preferences. *J. Clean. Prod.* **2023**, *410*, 137162. [CrossRef]
37. Ding, F.; Lu, Z.; Jin, M.; Sun, L. Manufacturer's Encroachment and Carbon Emission Reduction Decisions Considering Cap-and-Trade Regulation and Consumers' Low-Carbon Preference. *Int. J. Environ. Res. Public Health* **2022**, *19*, 10407. [CrossRef] [PubMed]
38. Sun, L.; Cao, X.; Alharthi, M.; Zhang, J.; Taghizadeh-Hesary, F.; Mohsin, M. Carbon emission transfer strategies in supply chain with lag time of emission reduction technologies and low-carbon preference of consumers—ScienceDirect. *J. Clean. Prod.* **2020**, *264*, 121664. [CrossRef]
39. Zhu, J.; Feng, T.; Lu, Y.; Jiang, W. Using blockchain or not? A focal firm's blockchain strategy in the context of carbon emission reduction technology innovation. *Bus. Strategy Environ.* **2024**, 1–27. [CrossRef]
40. Liu, W.; Wang, D.; Tang, O.; Zhu, D. The impacts of logistics service integrator's overconfidence behaviour on supply chain decision under demand surge. *Eur. J. Ind. Eng.* **2018**, *12*, 558–597. [CrossRef]
41. Water Transport Science Research Institute of China. Analysis of the Benefits of Using Shore Power and Low-Sulphur Fuel for Ships Calling at Ports. Available online: https://www.wti.ac.cn/wti/TAfbdwzlhb/2521.jhtml (accessed on 29 July 2019).
42. China Port. Management Suggestions on the Development of China's Shore Power. Available online: http://www.escn.com.cn/news/show-430111.html (accessed on 12 June 2017).
43. Qian, W.Y.J. Game theory analysis of technology adoption timing and pricing decision in supply chain system under asymmetric nash equilibrium. *J. Intell. Fuzzy Syst. Appl. Eng. Technol.* **2018**, *35*, 3101–3111. [CrossRef]
44. Yang, L.; Cai, Y.; Zhong, X.; Shi, Y.; Zhang, Z. A Carbon Emission Evaluation for an Integrated Logistics System—A Case Study of the Port of Shenzhen. *Sustainability* **2017**, *9*, 462. [CrossRef]
45. Liu, J.; Wang, J. Carrier alliance incentive analysis and coordination in a maritime transport chain based on service competition. *Transp. Res. Part E Logist. Transp. Rev.* **2019**, *128*, 333–355. [CrossRef]
46. Balakrishnan, R.; Geetha, V.; Kumar, M.R.; Leung, M.F. Reduction in Residential Electricity Bill and Carbon Dioxide Emission through Renewable Energy Integration Using an Adaptive Feed-Forward Neural Network System and MPPT Technique. *Sustainability* **2023**, *15*, 14088. [CrossRef]

mathematics

MDPI

Article

A Cost Optimisation Model for Maintenance Planning in Offshore Wind Farms with Wind Speed Dependent Failure Rates

Xiaodong Li [1,*], Xiang Song [2] and Djamila Ouelhadj [2]

[1] School of Computing and Creative Technologies, University of the West of England, Bristol BS16 1QY, UK
[2] School of Mathematics and Physics, University of Portsmouth, Portsmouth PO1 2UP, UK; xiang.song@port.ac.uk (X.S.); djamila.ouelhadj@port.ac.uk (D.O.)
* Correspondence: xiaodong.li@uwe.ac.uk

Abstract: This paper presents an optimisation model for cost optimisation of maintenance at an offshore wind farm (OWF). The model is created for OWF project developers to optimise strategic resources to meet their maintenance demand. The model takes into account various maintenance categories on a full range of wind turbine components; the failure rate associated with each component is dependent on wind speed in order to consider weather uncertainty. Weibull distribution is used to predict the probability of wind speed occurring during a given period based on available historical data. The performance of the proposed optimisation model has been validated using reference cases and a UK OWF in operation. Various optimal solutions are investigated for the problems with increased and decreased mean turbine failure rates as a sensitivity test of the model.

Keywords: offshore wind; renewable energy; operations and maintenance (O&M); decision making; cost optimisation

MSC: 90-08

1. Introduction

For an offshore wind farm (OWF), the operations and maintenance (O&M) activities should be conducted throughout the project life. The O&M activities of offshore wind turbines contribute up to 30% of the energy production cost [1,2], although such costs dislike the huge amount of installation cost during the construction phase. In practice, a major proportion of the O&M costs occur from corrective maintenance activities scheduled to recover the failures on different wind turbine components [3]. Hence, an improvement in the costing performance on corrective maintenance may effectively reduce the energy production costs in OWFs.

During maintenance services in OWFs, one of the common challenges is the transport of technicians, equipment, spare parts and large components to wind turbines offshore [4]. An efficient fleet of transport is required for an offshore wind project, especially to recover wind turbine failures quickly in corrective maintenance. Hence, a large part of O&M costs is spent on purchasing or chartering-in transport, including vessels and helicopters. Transport efficiency plays a key role in determining transport demand in terms of working hours required for fixing different faults by considering vessel/helicopter compatibility and weather restriction. The most popular vessels used in OWF maintenance include crew transfer vessels, field support vessels and jack-up vessels; some other types of vessels might be requested for specialised tasks such as cable-laying [4]. Helicopters are usually considered to take emergency repairs or minor maintenance services in order to help wind turbines re-start work in order after a short breakdown period.

Apart from the purchase or charter-in costs of transportation, a number of other cost elements occurring in maintenance services in OWFs, such as labour costs caused by repair and replacement of turbine components, fuel cost resulting from transport, the repair cost of materials or spare parts and revenue loss due to production downtime [5]. In practice, corrective maintenance management is critical for maximising the availability of production systems and minimising the overall O&M cost [6]. The revenue income loss can be estimated by computing the required service time, the expected waiting time, and the productivity level associated with the probabilistic wind speed. The accessibility of the installed facilities by different transports under various sea states greatly affects the downtime length. Hence, the maintenance of any offshore wind turbines is not simple due to the restricted logistics and accessibility.

An optimal plan of both preventative and corrective maintenance is critical for reducing the O&M cost of an offshore wind farm. The key issue in developing the optimal plan is the decision of how to use the transport and labour to carry out maintenance jobs. A survey of OWF owners was conducted by Pahlke [7], with almost three-quarters of the respondents stating the need for a decision-aiding model/tool, whereas few had existing models for use [5,8]. The existing decision support approaches to date use mainly simulation techniques [8–10]. However, an optimisation solution cannot be derived directly through simulation. Hence, recently a variety of mathematical optimisation models have been developed for the cost minimisation of maintenance planning in offshore wind farms [3,4,11–17]. The most recent research combined mathematical optimisation modelling and simulation techniques [18,19].

The failure rate affects the activity time and costs of transport and labour, especially the corrective maintenance for turbine component breakdown [20,21]. The unscheduled repairs/replacement of failed wind turbine components result in a significant proportion of the maintenance actions, typically between 50–70% [22]. The maintenance practices in an OWF can be optimised with respect to the failure frequency and repairs/replacement costs of wind turbines in the offshore environment. An effectively optimised maintenance schedule for OWFs could potentially reduce the overall maintenance expedition costs to a minimum level in conjunction with the use of historical data on offshore wind turbine failure rates [23].

In this paper, we propose a mathematical optimisation model for OWF developers to improve the cost-effectiveness of conducting maintenance activities. The main objective is to achieve the minimum overall cost incurred in both preventative and corrective maintenance, including transportation, labour, fuel, repair and downtime costs. A variety of wind turbine components are considered under the classification of four categories of maintenance tasks [24]. The contribution of this paper resides in the determination of the failure rates of turbine components, which are expressed as a function of wind speed and the related wind speed probability. To the best of our knowledge, no such study exists in the literature that sets the failure rates of turbine components as a wind-speed-dependent parameter to estimate the maintenance demand. In contrast, a significant relationship exists between wind speed and wind turbine failure rate, according to Wilson and McMillan [25,26]. This relationship needs to be taken into account when the management team is scheduling the corrective maintenance activities for the offshore wind farm.

Although there are many existing decision support tools/systems for optimising OWF maintenance plans, it is not easy to see an algorithm considering wind speed via Weibull distribution. Weibull distribution has been recognised as an effective way to forecast wind speed on the basis of historical data [27]. In the mathematical optimisation model proposed in this paper, the failure rates are differentiated within a range of wind turbine components under four corrective maintenance categories. The wind speed forecasting formula is developed based on Weibull distribution, and the associated energy generation is calculated. The solution of the mathematical optimisation model provides an efficient decision-making approach for optimising and analysing maintenance activities.

The rest of this paper is expressed as follows: In Section 2, a review of existing decision models/tools/algorithms developed for offshore wind farm maintenance is presented. Section 3 introduces the background information on offshore wind farm maintenance, which also gives the essential assumptions of the developed optimisation model. Section 4 describes the proposed OWF maintenance model for optimal strategic planning. Model results and sensitivity analysis are illustrated in Section 5. Finally, some concluding remarks and further research suggestions are given in Section 6.

2. Existing Decision Aid Models for Offshore Wind Farm Maintenance

When modelling O&M practices in OWFs, the failure rate of the wind turbine components is a key parameter that will significantly affect the energy output and cost per unit of energy produced. Several models have been produced to forecast wind power revenue [28] or to predict O&M costs [29,30] by considering wind turbine reliability. Reliability models are utilised to estimate the failure frequency of offshore wind turbines and identify the repair time for each type of failure [31]. The revenue losses due to wind turbine failures and necessary maintenance actions are recognised as the main portion of maintenance costs. This literature review focuses mainly on the development of mathematical optimisation models for the cost minimisation of maintenance planning in offshore wind farms. Operational research (OR) techniques have been used widely in scheduling and capacity planning of renewable energy production [11,32]. For simulation tools to analyse the O&M costs in offshore wind farms, we refer the readers to Hofmann [10] for a survey of decision support models for offshore wind farms with a special emphasis on O&M strategies.

The first mathematical optimisation model that addresses the vessel fleet composition problem for maintenance operations at OWFs was proposed by Halvorsen–Weare et al. [4]. The solution of the model would be used by decision-makers when deciding which vessel type should be purchased or chartered in. The model also helps to determine which infrastructure, such as the maintenance base, should be used to minimise the total cost of the vessel fleet. The model considered uncertain weather parameters, including wind speed, wave heights, wave direction and current, to estimate the spot prices of charter-in contracts and the number of failures that lead to corrective maintenance operations. Finally, they indicated clearly that all these parameters are treated as known in their deterministic model. The work of Halvorsen–Weare et al. [4] has been extended to develop a three-stage stochastic programming model, in which the uncertainty in vessel spot rates, weather conditions, electricity prices and failures are considered. Gundegjerde et al. [12] claimed that these uncertainties are often considered in simulation models, whereas they are mainly handled as deterministic parameters in mathematical programming models. Stålhane et al. [16] applied a two-stage stochastic programming model to help decide the optimal vessel fleet to support maintenance operations at an offshore wind farm.

A number of researchers investigated the optimisation of vessel routes and schedules for maintenance tasks at an offshore wind farm. The problem is similar to a vehicle routing problem with pick-up and delivery [33]. In [34], the fleet of vessels is heterogeneous and located at a depot (base) at the beginning of the planning horizon. The goal is to create one route for each vessel so that the vessel travels from the depot (base) to a set of wind turbines, where it will deliver and pick up technicians and spare parts to perform the maintenance tasks at each turbine. In their problem, the cost function includes travel cost, downtime cost and penalty cost for not performing maintenance tasks in the current time period. The mathematical model is deterministic, and no uncertainties are considered. Their model was later extended to a two-stage stochastic programming model where uncertainty in demand and weather conditions were taken into account [13]. Irawan et al. [14] extended the model in [34] to resolve maintenance routing and scheduling issues within multiple wind farms and O&M bases. This case is relevant to when clusters of neighbouring wind farms are being developed, allowing maintenance resources to be shared between them. In the proposed model, they also took into account different skill types of technicians at each

O&M base, the availability of maintenance vessels and spare parts and the capacity of each type of vessel to transfer spare parts.

Although they addressed the weather window to reflect the uncertainty of the weather upon the solution, the weather window has been given as a known value by [34] in their deterministic model. Fan et al. [15] applied mixed particle swarm optimisation to identify a mapping relation between vessels and wind farms and explore the optimal vessel allocation scheme. Based on the scheme of vessel allocation, then, a discrete wolf pack search is introduced for the maintenance route optimisation under all constraints.

Most recently, to handle the uncertainties of weather conditions and turbine failure rate in offshore wind turbine maintenance, Irawan et al. [19] proposed a sim-metaheuristic algorithm which combines a metaheuristic with Monte Carlo simulation to solve the stochastic maintenance routing problem. The Monte Carlo simulation takes a number of uncertainties into consideration: weather conditions, the condition of the turbine, technicians' skills, vessel conditions and the weight of equipment/parts. Turan et al. [18] combined system dynamics and discrete event simulation approaches to model and solve a strategic problem of fleet renewal to match future requirements under uncertain conditions. The uncertainties considered are resource uncertainties. Li et al. [35] considered more uncertainties in OWF failure to generate a multi-objective OWF maintenance strategy optimisation framework by using a probabilistic method and the Monte Carlo method.

The most relevant paper to the proposed research is the one by Li et al. [3]. In the paper, the decisions need to be made on the maintenance strategies to select for OWF developers, the number of technicians for HR managers and the number of chartered vessels for O&M planners. The objective is to pursue a minimum total cost of personnel, transport and breakdown for O&M in offshore wind farms. Li et al. [3] developed both deterministic and stochastic optimisation models for this problem. The deterministic optimisation model is used when the failure rates of wind turbine components are given, whilst the stochastic model is utilised in case accurate failure data is unavailable.

From the review of the existing optimisation models for maintenance in OWFs, there is scarce research in the literature that sets the failure rates of turbine components as a weather-based parameter to estimate the maintenance demand. The main contribution is to link the failure rates of turbine components with wind speed and the related wind speed probability. The new mathematical optimisation model proposed in this paper concentrates on corrective maintenance activities in an offshore wind farm since they are more sensitive to weather variations. The objective of the optimisation model is to minimise the overall maintenance cost with a wider range of cost elements resulting from labour, transport, fuel, repair/replacement and downtime in practice. Different wind speed levels are considered with an occurrence probability based on the historical data; modified model constraints will correspond to estimated failure rates with the probabilistic wind speed.

3. OWF Maintenance Characteristics and Assumptions of the Optimisation Model

The proposed model focuses on minimising the expected maintenance costs of an offshore wind farm during a given period. A range of maintenance categories is specified technically on wind turbine components. Different kinds of transport from a base port, including vessels and helicopters, are used to execute the maintenance work at a single offshore wind farm. The travel distance is a return journey between the base port and the offshore wind farm, as shown in Figure 1. Maintenance technicians are hired on either a full-time or part-time basis. Wind speed probability, as a key parameter of sea state, is predicted by using Weibull distribution. This parameter is crucial to predict turbine failures and to determine the expected energy production.

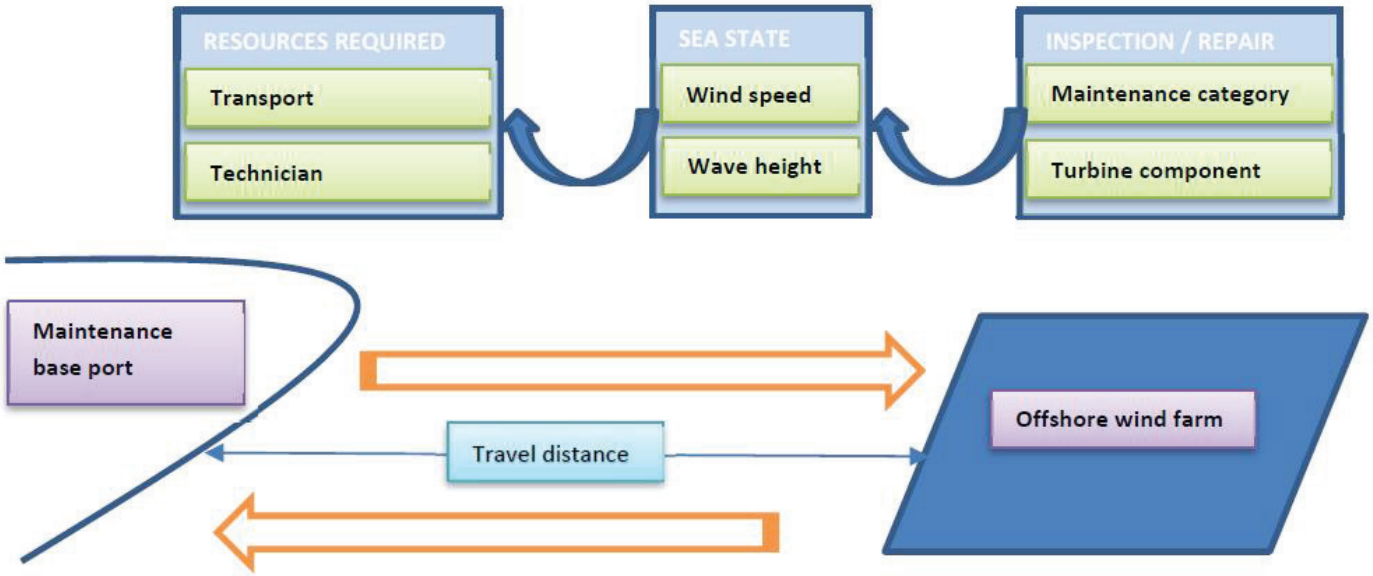


Figure 1. Offshore wind farm maintenance decision-making workflow.

Based on the weather-based failure rates on each turbine component in a given OWF size, the required number of transports and technicians would be balanced by the model with the minimised total amount of annual cost. Maintenance demand, vessel/helicopter compatibility and sea state restriction on transport are considered in the transport selection and cost minimisation. Figure 2 shows the workflow of the proposed optimisation model with its inputs and outputs. The overall workflow of the optimisation model considers two sets of input parameters, namely OWF data and maintenance technical data. The wind speed via Weibull distribution, with historical weather data, is used to determine the power generation and failure frequency. The input of OWF maintenance data, such as transport compatibility and deployability, will be picked to meet the maintenance requirement. On a given actual problem, the maintenance technical data could be altered by using its realistic values in the proposed model. After implementing the model, the demand for transport and technicians is estimated, and the minimised costs are computed.

3.1. Categorisation of Maintenance on Wind Turbine Components

The developed optimisation model takes into account both preventative and corrective maintenance. Different categories of maintenance activities, such as minor repair and major replacement, are allocated to a range of key turbine components. Index i denotes the maintenance category, and j represents the component. All key components of a wind turbine, such as a gearbox and rotor blade, are considered in the maintenance activities of the model.

For each maintenance category on every wind turbine component, the repair time, repair cost and the number of technicians required are determined (Carroll et al., 2016). Repair time (T_{ij}^{repair}) covers the time that the maintenance technicians use during inspection, repair or replacement. Repair cost (C_{ij}^{repair}) is the cost of materials, equipment and tools. In addition, the number of required technicians (Q_{ij}) is also pre-determined depending on the workload of each maintenance category on different wind turbine components. The travel time of a vessel or helicopter from the maintenance base port to the offshore wind farm is defined as $T_k^{travel} = \frac{2D}{S_k}$ by the travel distance of a returned trip over the transport speed.

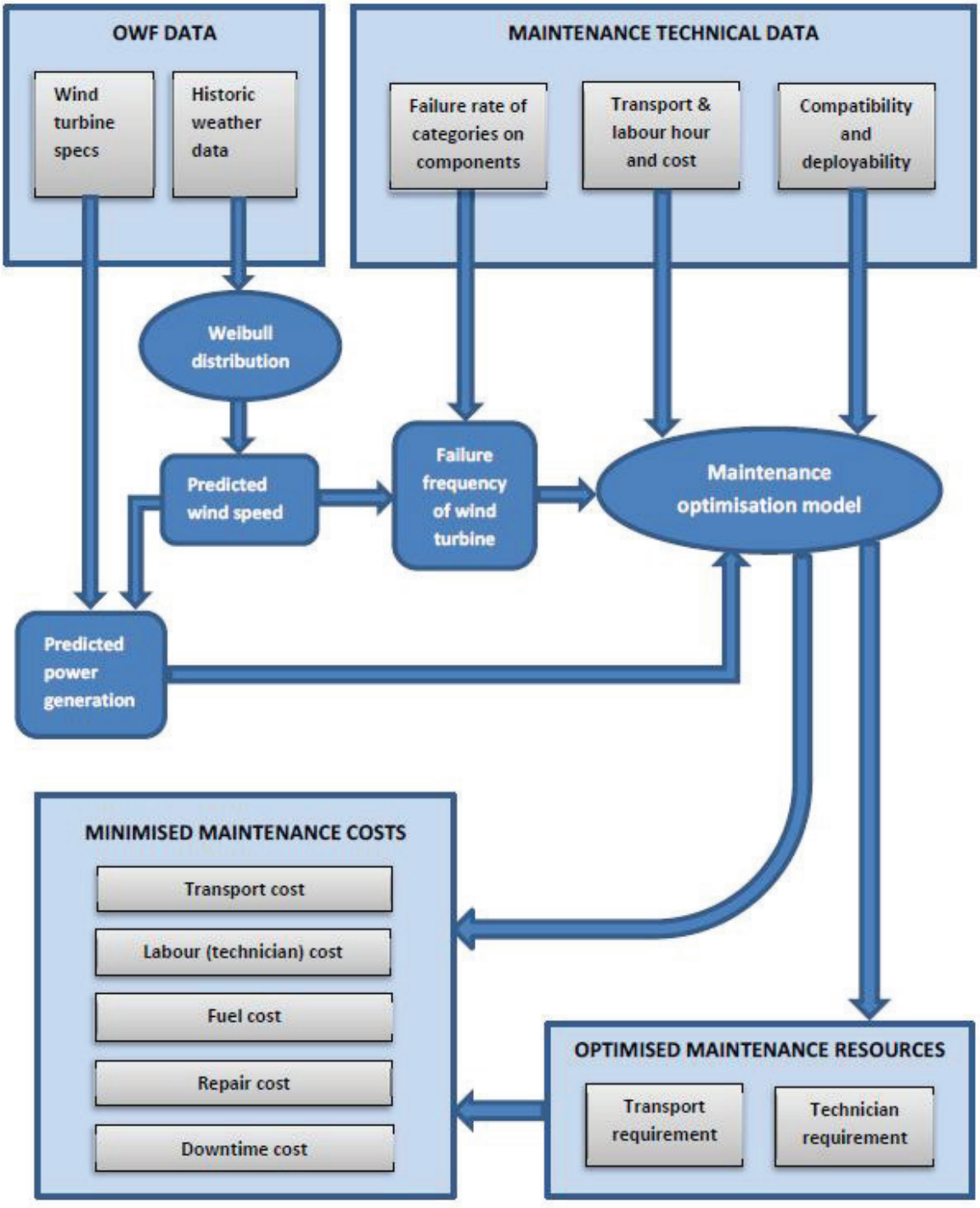


Figure 2. Workflow of the proposed maintenance optimisation model.

3.2. Compatibility and Deployability of Maintenance Transport

The main activities in offshore wind maintenance are the transport of the technicians, materials and spare parts and the execution of service, repair or replacement. The compatibility and deployability of maintenance transport are recognised with OWF practitioners in the 2OM (Offshore Operations & Management Mutualisation) project.

A range of transport means, including vessels and helicopters, are used to execute different maintenance tasks on the wind turbine components. Index k denotes the common transportation type used in offshore wind maintenance. The most suitable transport type should be selected to execute a maintenance job subject to the compatibility of transport

and weather restrictions. Different wind speed levels (w) are considered to investigate the impact of failure rate on energy productivity.

Vessels and/or helicopters can be chartered, on a short-term or long-term lease, to carry out maintenance tasks during the planning horizon. According to the response in interviews with OWF practitioners, each type of transport has a given fixed cost (C_k^{fixed}), charter cost ($C_k^{charter}$), fuel consumption (O_k), fuel cost (C_k^{fuel}) and transport speed (S_k). The length of the lease period in hours ($H_k^{transport}$) of each transportation type is pre-determined in the developed model. In addition, the labour hours of a full-time technician ($H_k^{labourFT}$) and a part-time technician ($H_k^{labourPT}$) are introduced as working time in one year and three months, respectively. Hence, the labour cost per full-time technician $\left(C_k^{labourFT}\right)$ and part-time technician $\left(C_k^{labourPT}\right)$ working on each transport type is defined as the annual and quarterly salaries.

A maintenance team is usually sent to execute a task; the number of technicians (Q_{ij}) in a team depends on the workload of maintenance category i on component j. Each maintenance category requires compatible transportation. For instance, a major replacement of large turbine components must be executed by a heavy vessel with a crane. The OWF practitioners also introduced the two binary parameters of compatibility and deployability. A binary parameter ($B_{ik}^{compatible}$) is used to express the compatibility of transport type k on maintenance category i. The use of transportation means also considers the sea state, wind speed acts as a key parameter to determine whether a transport type can go to execute maintenance work. If the wind speed reaches the operational limit of the suitable transports, the maintenance activities will be postponed. Hence, another binary ($B_{kw}^{deployable}$) is defined to show the deployability of each transport type k under wind speed w. The selection of transport type must be subject to the binary variable of both compatibility on maintenance categories and deployability on wind speeds.

3.3. Wind Speed Dependent Failure Rates

A significant relationship exists between the wind speed and the wind turbine failure rate. Wilson and McMillan [25] proposed that the failure rates could be computed as a function of wind speed, and they then developed the following model of wind speed-dependent failure rates to assess wind farm reliability.

$$\lambda_{ij,w}^{fail|wind} = \frac{Prob_{w,ij}^{wind|fail} \cdot F_{ij}}{Prob_w^{wind}} \tag{1}$$

$Prob_{w,ij}^{wind|fail}$ is the probability of wind speed w occurring, given a failure category i has occurred to component j. It could be calculated by taking a probability density function of average wind speed recorded on days when a failure occurred. F_{ij} is the mean failure rate of category i on component j. $Prob_w^{wind}$ represents the probability that the average wind speed is w. Therefore, Equation (1) is used to calculate the probability of a failure to category i of component j, given an average wind speed w. The key advantage of using this model is that the effect of seasonal changes on wind turbine operation can be accounted for.

3.4. Weibull Distribution to Predict Wind Speed Probabilities

As a common method to forecast wind speed probability, the Weibull distribution provides a close approximation and has been used to represent wind speed distribution for many applications of wind sources [27]. Its greater flexibility and simplicity make it ideal for experimental data [36]. The Weibull distribution function, as a two-parameter function for wind speed probability, is expressed in Equation (2).

$$Prob_w^{wind} = \left(\frac{m}{c}\right)\left(\frac{w}{c}\right)^{m-1} exp\left(-\left(\frac{w}{c}\right)^m\right) \tag{2}$$

where w is the wind speed, m is the shape parameter and c is the scale parameter. A range of methods can be applied to determine the value of the parameters, such as the empirical method, maximum likelihood method and graphical method [36,37].

3.5. Energy Generation with Different Wind Speeds

The available energy generation rate of an offshore wind turbine varies with different wind speeds. Cut-in (WS^{in}) and cut-out (WS^{out}) wind speeds specify the minimum and maximum wind speeds that the turbine can work to generate energy. If wind speed is less than the cut-in level or greater than the cut-out level of a specific wind turbine model, then the turbine terminates energy generation. The energy production rate keeps increasing with the strength of the wind between cut-in and rated wind speeds. The rated wind speed (WS^{rated}) provides sufficient wind power that the turbine works with the rated capacity (Cap^{rated}). The production rate is stable at the rated capacity level when the wind speed is over the technically rated wind speed until the cut-out level. A common formula, expressed in Equation (3), is used to calculate how much power could be generated in one hour by a given wind turbine under wind speed w.

$$G_w^{avail}(WS_w) = \begin{cases} 0,\ WS_w < WS^{in} \\ \frac{1}{2}\cdot\rho\cdot A\cdot(WS_w)^3\cdot Coe^p,\ WS^{in} \le WS_w < WS^{rated} \\ Cap^{rated},\ WS^{rated} \le WS_w \le WS^{out} \\ 0,\ WS_w > WS^{out} \end{cases} \tag{3}$$

Sweep area (A) is usually determined by the length of the turbine blade. Air density (ρ) and wind speed (WS_w) are core parameters in the energy production formula. The value of the power coefficient (Coe^p) is unique to each wind turbine and is a function of the wind speed of the turbine. The Betz Limit [38] specified that 0.59 is the theoretically maximum power coefficient of any design of a turbine. The realistic power coefficient is significantly below the Betz Limit; values between 0.35 and 0.45 are common in the best-designed wind turbines.

4. The Optimisation Model Formulation

The optimisation model aims to minimise the overall cost, which includes the following costs: transportation, labour, fuel consumption, repair cost and downtime.

- Transport fixed cost: Transport fixed cost includes one-off costs during an agreed charter period of a vessel or helicopter, such as insurance, maintenance, etc. The cost for each transport type k is computed by the unit fixed cost of a charter period and the number of lease charters on this transport type required for repair/replacement works.

$$Transport\ fixed\ cost\ =\ C_k^{fixed}\cdot x_k^{\alpha} \tag{4}$$

- Labour cost: Full-time technicians are charged on the basis of an annual salary; part-time technicians are assumed to take a short-term contract every quarter. Short-term temporary employment provides the flexibility to hire more technicians during the busy seasons. However, the salary of quarterly contracts comes with a 37.5% extra from the annual salary rate. Then, the total labour cost of each transportation k is determined by the full-time and part-time salary ($C_k^{labourFT}$ *and* $C_k^{labourPT}$) with the number of technicians employed ($y_k^{FT,\alpha}$ *and* $y_k^{PT,\alpha}$) for undertaking maintenance activities.

$$Labour\ cost\ =\ C_k^{labourFT}\cdot y_k^{FT,\alpha} + C_k^{labourPT}\cdot y_k^{PT,\alpha} \tag{5}$$

- Transport charter cost: The total charter cost of transport type k is determined in terms of a daily charter rate ($C_k^{charter}$) and the length of chartering period. As a popular transport type in OWF maintenance, crew transfer vessels are assumed to be chartered on an annual basis. Field support vessels and jack-up vessels are usually chartered weekly when a major repair or replacement is required. Helicopters are required in

the case of urgent maintenance demand; they are chartered on the basis of the number of hours. The number of charter periods for each transport type must be multiple of a work shift, L_k^{shift} is the working hours in a daily shift.

$$Transport\ charter\ cost\ =\ C_k^{charter}.\left\lceil \frac{\left(T_k^{travel}+T_{ij}^{repair}\right)\cdot z_{ijkw}}{L_k^{shift}} \right\rceil \tag{6}$$

where z_{ijkw} is a binary variable whether transport type k is selected for maintenance category i on component j under wind speed w

- Transport fuel cost: The fuel cost rate (C_k^{fuel}) is defined per m^3 for each transport type k. Fuel consumption of a specific transport (O_k) is estimated per hour of travel time.

$$Transport\ fuel\ cost\ =\ C_k^{fuel}\cdot O_k\cdot T_k^{travel}\cdot z_{ijkw} \tag{7}$$

- Repair cost: Repair cost is the direct maintenance cost of repair materials, spare parts and equipment. The total amount of repair cost should be determined by the unit cost (C_{ij}^{repair}) of category i on component j, with the maintenance demand that is dependent on the failure rate.

$$Repair\ cost\ =\ C_{ij}^{repair}\cdot\lambda_{ij,w}^{fail|wind}\cdot N \tag{8}$$

- Downtime cost: Any revenue loss due to the breakdown of turbines is defined as downtime cost, which is computed by the hourly income of wind power production (C^{down}) and the length of downtime, including travel time, repair time and waiting time of each maintenance task. A single trip travel time ($\frac{T_k^{travel}}{2}$) of the selected transport k is accounted-for downtime. The length of repair time (T_{ij}^{repair}) is given as a constant of the maintenance category i on turbine component j. It is not related to the type of transportation used.

$$downtime\ cost\ =\ C^{down}\cdot G_w^{avail}.\left(\frac{T_k^{travel}}{2}+T_{ij}^{repair}+T_{kw}^{wait}\right)\cdot z_{ijkw} \tag{9}$$

The objective function of the model is to minimise the sum of the two fixed costs and four expected variable costs.

$$\begin{aligned} Min\ Total\ Cost\ =\ & \sum_{k\in K} C_k^{fixed}\cdot x_k^{\alpha}+\sum_{k\in K} C_k^{labourFT}\cdot y_k^{FT,\alpha}+\sum_{k\in K} C_k^{labourPT}\cdot y_k^{PT,\alpha} \\ & +\sum_{i\in I}\sum_{j\in J}\sum_{k\in K}\sum_{w\in W} C_k^{charter}.\left\lceil \frac{\left(T_k^{travel}+T_{ij}^{repair}\right)\cdot z_{ijkw}\cdot\lambda_{ij,w}^{fail|wind}\cdot N}{L_k^{shift}} \right\rceil \\ & +\sum_{i\in I}\sum_{j\in J}\sum_{k\in K}\sum_{w\in W} C_k^{fuel}\cdot O_k\cdot T_k^{travel}\cdot z_{ijkw}\cdot\lambda_{ij,w}^{fail|wind}\cdot N \\ & +\sum_{i\in I}\sum_{j\in J}\sum_{w\in W} C_{ij}^{repair}\cdot\lambda_{ij,w}^{fail|wind}\cdot N \\ & +\sum_{i\in I}\sum_{j\in J}\sum_{k\in K}\sum_{w\in W} C^{down}\cdot G_w^{avail}.\left(\frac{T_k^{travel}}{2}+T_{ij}^{repair}+T_{kw}^{wait}\right)\cdot z_{ijkw}\cdot\lambda_{ij,w}^{fail|wind}\cdot N \end{aligned} \tag{10}$$

Constraints:

The above objective function of minimising the total cost should be achieved subject to a variety of constraints in O&M for the use of vessels/helicopters and technicians. In order to hire enough length of transportation and labour times to execute maintenance works, both transport hours (constraint set 1(a)–1(d) and labour hours (constraint set 2(a)–2(d)) should cover the requirement of different maintenance categories on turbine components.

The repair time (T_{ij}^{repair}) and the travel time (T_{k}^{travel}) are the two major portions to estimate the length of the required time of transport type k.

Constraint set 1(a): The total available time of each transport type k must be greater than the length of the working time required, including travel and repair/replacement, for undertaking maintenance.

$$x_k^{\alpha} \cdot H_k^{transport} \geq \sum_{i \in I} \sum_{j \in J} \sum_{w \in W} \left(T_{ij}^{repair} + T_k^{travel} \right) \cdot z_{ijkw} \cdot \lambda_{ij,w}^{fail|wind} \cdot N \; \forall \; k \in K \tag{11}$$

Constraint set 1(b): The available time of each transport type k used for maintenance category i must be greater than the length of the working time required, including travel and repair/replacement, for undertaking maintenance.

$$x_{ik}^{\beta} \cdot H_k^{transport} \geq \sum_{j \in J} \sum_{w \in W} \left(T_{ij}^{repair} + T_k^{travel} \right) \cdot z_{ijkw} \cdot \lambda_{ij,w}^{fail|wind} \cdot N \; \forall \; i \in I, k \in K \tag{12}$$

Constraint set 1(c): The available time of each transport type k used for maintenance category i on turbine component j must be greater than the length of the working time required, including travel and repair/replacement, for undertaking maintenance.

$$x_{ijk}^{\gamma} \cdot H_k^{transport} \geq \sum_{w \in W} \left(T_{ij}^{repair} + T_k^{travel} \right) \cdot z_{ijkw} \cdot \lambda_{ij,w}^{fail|wind} \cdot N \; \forall \; i \in I, j \in J, k \in K \tag{13}$$

Constraint set 1(d): The available time of each transport type k used for maintenance category i on turbine component j under wind speed w must be greater than the length of working time required, including travel and repair/replacement, for undertaking maintenance.

$$x_{ijkw}^{\delta} \cdot H_k^{transport} \geq \left(T_{ij}^{repair} + T_k^{travel} \right) \cdot z_{ijkw} \cdot \lambda_{ij,w}^{fail|wind} \cdot N \; \forall \; i \in I, j \in J, k \in K, w \in W \tag{14}$$

where

$$x_k^{\alpha} \geq x_{ik}^{\beta} \geq x_{ijk}^{\gamma} \geq x_{ijkw}^{\delta} \tag{15}$$

Constraint set 2(a): The available labour hours for both full-time and part-time technicians on transport k should cover the travel and repair/replacement of maintenance executed by the transport.

$$y_k^{FT,\alpha} \cdot H_k^{labourFT} + y_k^{PT,\alpha} \cdot H_k^{labourPT} \geq \sum_{i \in I} \sum_{j \in J} \sum_{w \in W} \left(T_{ij}^{repair} + T_k^{travel} \right) \cdot z_{ijkw} \cdot \lambda_{ij,w}^{fail|wind} \cdot N \quad \forall \; k \in K \tag{16}$$

Constraint set 2(b): The available labour hours for both full-time and part-time technicians on transport k should cover the travel and repair/replacement of maintenance category i executed by the transport.

$$y_{ik}^{FT,\beta} \cdot H_k^{labourFT} + y_{ik}^{PT,\beta} \cdot H_k^{labourPT} \geq \sum_{j \in J} \sum_{w \in W} \left(T_{ij}^{repair} + T_k^{travel} \right) \cdot z_{ijkw} \cdot \lambda_{ij,w}^{fail|wind} \cdot N \quad \forall \; i \; \in \; I, \; k \in K \tag{17}$$

Constraint set 2(c): The available labour hours for both full-time and part-time technicians on transport k should cover the travel and repair/replacement of maintenance category i on component j executed by the transport.

$$y_{ijk}^{FT,\gamma} \cdot H_k^{labourFT} + y_{ijk}^{PT,\gamma} \cdot H_k^{labourPT} \geq \sum_{w \in W} \left(T_{ij}^{repair} + T_k^{travel} \right) \cdot z_{ijkw} \cdot \lambda_{ij,w}^{fail|wind} \cdot N \quad \forall \; i \in I, \; j \in J, \; k \in K \tag{18}$$

Constraint set 2(d): The available labour hours for both full-time and part-time technicians on transport k should cover the travel and repair/replacement of maintenance category i on component j under wind speed w executed by the transport.

$$y_{ijkw}^{FT,\delta} \cdot H_k^{labourFT} + y_{ijkw}^{PT,\delta} \cdot H_k^{labourPT} \geq \left(T_{ij}^{repair} + T_k^{travel}\right) \cdot z_{ijkw} \cdot \lambda_{ij,w}^{fail|wind} \cdot N \quad \forall\, i \in I,\ j \in J,\ k \in K,\ w \in W \tag{19}$$

where

$$y_k^{FT,\alpha} \geq y_{ik}^{FT,\beta} \geq y_{ijk}^{FT,\gamma} \geq y_{ijkw}^{FT,\delta} \tag{20}$$

$$y_k^{PT,\alpha} \geq y_{ik}^{PT,\beta} \geq y_{ijk}^{PT,\gamma} \geq y_{ijkw}^{PT,\delta} \tag{21}$$

Constraint set 3: The total number of full-time and part-time technicians on transport k must be at least equal to the number required to carry any maintenance category i on wind turbine component j if the transport is selected to execute the work under wind speed w.

$$y_{ijkw}^{FT,\delta} + y_{ijkw}^{PT,\delta} \geq Q_{ij} \cdot z_{ijkw} \quad \forall i \in I,\ j \in J,\ k \in K,\ w \in W \tag{22}$$

Constraint set 4: Transport type k can be used to execute maintenance category i on component j under wind speed w only if the transport type is compatible with the maintenance category and deployable under the weather condition.

$$z_{ijkw} \leq B_{ik}^{compatible} \times B_{kw}^{deployable} \quad \forall i \in I,\ j \in J,\ k \in K,\ w \in W \tag{23}$$

Constraint set 5: A binary variable is used to indicate whether transport type k is selected to execute maintenance category i on component j under wind speed w. If transport type k is not selected for any category of maintenance on any wind turbine component under any wind speed, the number of the transport must be zero.

$$x_{ijkw}^{\delta} \leq M \cdot z_{ijkw} \text{ and } x_{ijkw}^{\delta} \geq z_{ijkw} \quad \forall i \in I,\ j \in J,\ k \in K,\ w \in W \tag{24}$$

where M is a large positive number.

Constraint set 6: Each maintenance job of category i on component j under wind speed w must be served by at least one type of transport if the failure rate is greater than zero and compatible transports are available.

$$\sum_{k \in K} z_{ijkw} \geq \sum_{k \in K} \lambda_{ij,w}^{fail|wind} \cdot B_{ik}^{compatible} \quad \forall i \in I,\ j \in J,\ w \in W \tag{25}$$

The optimisation model provides cost-effective planning for the maintenance operations of one offshore wind farm. It can be used to select the right type(s) of maintenance transport and technicians and to determine the optimal amount of transport charters and technicians for executing requested maintenance activities with the minimised total cost.

5. Experimental Results

A series of reference cases, initially published in [39], was applied to the developed model for validation. Rampion offshore wind farm is used, as a sample case study, to verify this optimisation model. The optimisation model has been solved by minimising multiple types of costs. Experimental results and sensitivity analysis are presented in this section.

5.1. Data Setting

According to Carroll et al.'s categorisation of corrective maintenance activities [24], four categories of corrective and preventative tasks (in Figure 3) are allocated to the maintenance of the nineteen wind turbine components (in Table 1). Maintenance frequency for the corrective maintenance depends significantly on the component failure rates. The

essential characteristics of the four corrective maintenance categories of the nineteen wind turbine components in the developed model are collected from the research work in [24], including mean failure rate, repair time, repair cost and the number of technicians required.

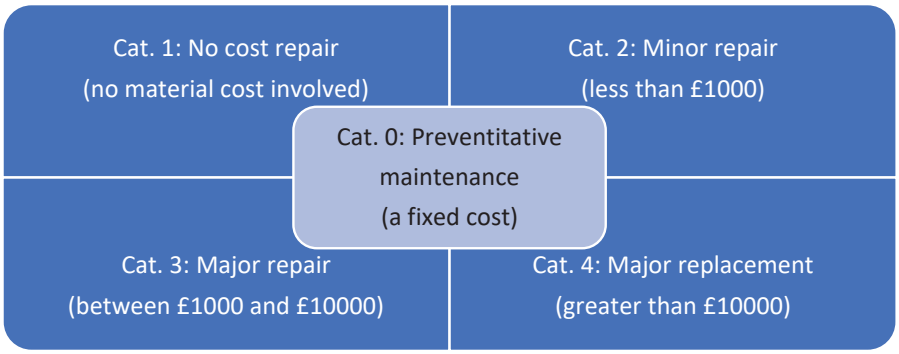


Figure 3. Categorisation of the maintenance on offshore wind turbines.

Table 1. List of key turbine components (Carroll et al., 2016).

Comp.1	Pitch system	Comp.11	Pumps/motors
Comp.2	Generator	Comp.12	Hub
Comp.3	Gearbox	Comp.13	Heating/cooling system
Comp.4	Rotor blades	Comp.14	Yaw system
Comp.5	Grease oil/cooling liquid	Comp.15	Tower/foundation
Comp.6	Electrical components	Comp.16	Power supply/converter
Comp.7	Contactor/circuit breaker	Comp.17	Transformer
Comp.8	Control system	Comp.18	Service items
Comp.9	Safety system	Comp.19	Other components
Comp.10	Sensors		

A range of transport types are used to execute different maintenance tasks on the wind turbine components; type $k = 1 \ldots 4$ denote the four common transportation means, namely crew transfer vessel (CTV), field support vessel (FSV), jack-up vessel (JUV) and helicopter (HEL). CTVs are popular for working in the offshore energy field, such as oil and gas. FSVs and JUVs are used to take large repair and/or heavy wind turbine components. HELs can support the transportation of technicians and small spare parts in emergencies and can significantly reduce the length of downtime. According to the data acquired from O&M practice in the sector, the binary variable ($B_{ik}^{compatible}$) of compatibility of transport k for maintenance category i is clarified in Table 2. The value of 1 indicates 'compatible', and 0 represents 'incompatible'. Based on the technical knowledge, in addition, the technicians working on field support vessels and jack-up vessels can be used on crew transfer vessels for minor repairs.

Table 2. Compatibility of each transport type on maintenance categories.

$B_{ik}^{comlpatible}$	*CTV*	*FSV*	*JUV*	*HEL*
CAT.0	1	1	1	1
CAT.1	1	1	1	1
CAT.2	1	1	1	1
CAT.3	0	1	1	0
CAT.4	0	0	1	0

The use of the maintenance transport is also subject to weather restrictions. The wind speeds $w = 1 \ldots 22$ are considered to investigate the impact of failure rate on energy productivity. The binary variable ($B_{kw}^{deployable}$) is shown in Table 3 to describe the deployability of each transport type under different wind speeds. By considering safety, for instance, a jack-up vessel is not allowed to operate a heavy lift for a major replacement if the wind speed is over 15 m/s.

Table 3. Deployability of each transport type on wind speeds (1–22 m/s).

$B_{kw}^{deployable}$	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*
CTV	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
FSV	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
JUV	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0
HEL	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0

The rest of the standard industrial data are summarised from the work of Dalgic et al. [40], such as transportation fixed cost, transport charter cost, transport fuel consumption, fuel unit cost, labour cost, transport speed, transport restriction by wind speed and usual transport charter period. The wind turbine specifications, such as rated capacity and rated wind speed, are available on several websites (i.e., http://4cOffshore.com, accessed on 15 May 2020). To achieve the expected solution in different cases by using the developed optimisation model, perhaps it is necessary to amend the model parameters with updated data on industrial operations and the market.

5.2. Validation of the Developed Model Based on a Reference Case

The proposed model has been evaluated, and its performance has been compared with other existing models published in [39]. The study uses a number of reference cases to verify four decision-making models for OWF maintenance as follows: the Strathclyde analysis tool, the NOWIcob decision support tool, the University of Stavanger (UiS) Simulation model and the ECUME model. A case study of an OWF that consists of eighty (80) 3.0 MW wind turbines, which is developed 50 km from an onshore maintenance port. Three types of vessels were considered to execute the annual preventative maintenance and four categories of corrective maintenance, including manual resets, minor repair, medium repair and major repair/replacement. Three CTVs, one FSV and one heavy-lift vessel are available in the maintenance base port.

A comparative analysis has been conducted to compare the proposed model (new model) with the models in the literature, as shown in Table 4. There are two main cost components, which are the annual loss of production and the annual direct O&M cost. The direct O&M cost contains vessel cost, repair cost and technician cost. By comparing the base case results, the annual loss of production from this new model, £19.27 million, is slightly higher than other models, which is based on the predicted power generation with stochastic wind speeds. The direct O&M cost indicates that the result (£19.35 million) from the newly developed model stays at the median level of these five models. Of the three elements, only the repair cost stays at the highest level.

Table 4. Comparison of cost results in the base case between models.

	New Model	***Strathclyde CDT***	***NOWIcob***	***UiS Sim Model***	***ECUME Model***
Annual loss of production	**£19.27m**	**£17.28m**	**£16.63m**	**£15.48m**	**£18.64m**
Annual direct O&M cost	**£19.35m**	**£22.44m**	**£25.17m**	**£17.92m**	**£14.48m**
Annual vessel cost	£13.25m	£17.84m	£19.18m	£12.24m	£9.30m
Annual repair cost	£4.50m	£3.00m	£4.39m	£4.08m	£3.58m
Annual technician cost	£1.60m	£1.60m	£1.60m	£1.60m	£1.60m

Several further cases were generated from the base case for investigating the quantitative sensitivity, including more (5) CTVs and fewer (1) CTVs, more (30) and fewer (10) technicians and failure rates down (50%) and up (200%). Figure 3 shows direct O&M costs for the base case and the other cases. By comparing with the results of the other four models presented by Dinwoodie et al. (2015), the quantitative trend is relatively consistent across the cases. The model presented in this paper gives the median level of direct O&M costs in most of the reference cases (Figure 4), regardless of the changes in vessel or technician. However, the new model generates the highest cost level within the higher failure case. By comparing to the other models, this highest cost is probably due to the probabilistic failure rates applied on each wind turbine component under various wind speeds. The cost performance in Figure 4 also indicates, in the aggregate, that these model results are placed in the range of other results.

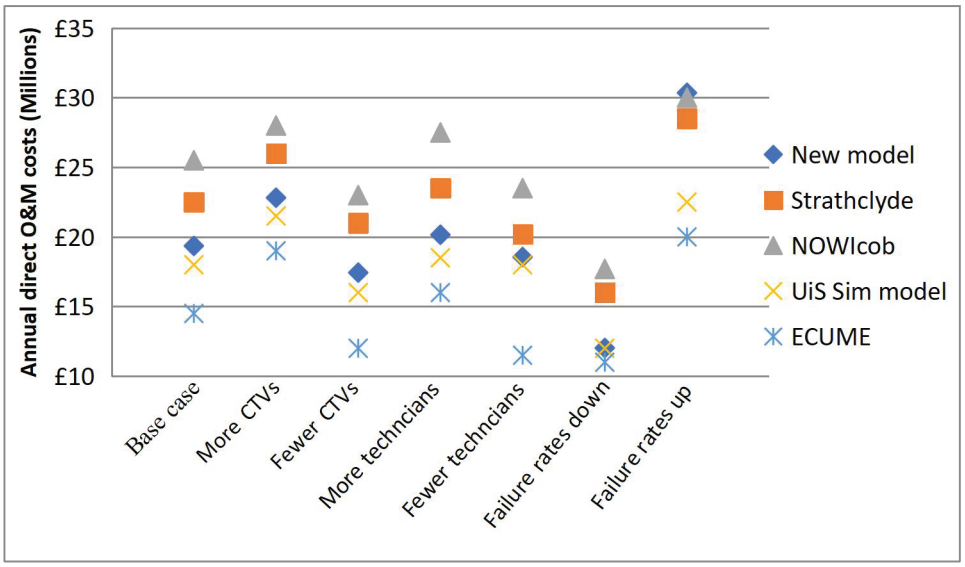


Figure 4. Annual direct O&M cost of the models in the reference cases.

5.3. A Sample Case Study

In order to evaluate the proficiency of the optimisation model, the data of the Rampion offshore wind farm is used as an example. The Rampion offshore wind farm is off the South Coast of the UK, and it is one of the new 'round 3' offshore wind sites designated by the UK government. As the data shown in Table 5, 116 wind turbines have been installed in the farm, which are specified technically by the rated capacity of 3.45 MW and the rated wind speed of 12.5 m/s. The average distance from the maintenance base to the OWF is 16.9 km. The mean wind speed over the last 10 years is 9.81 m/s.

Table 5. Offshore wind farm & turbine inputs.

Parameter	Value	Unit
Number of turbines	116	turbine
Generation capacity	3.45	MW
Cut-in wind speed	3	m/s
Rated wind speed	12.5	m/s
Cut-out wind speed	25	m/s
Distance to port	16.9	km
Water depth	19–39	m
Mean wind speed	9.81	m/s

In this study, the model was coded in the programming of A Mathematical Programming Language (AMPL) and then solved by the solver Gurobi on a laptop with CPU Core i5 2.4 GHz and 8 GB RAM. The optimal solutions with respect to different input parameter data were acquired by the solver of Gurobi within a reasonable computation time. With regard to the expected maintenance workload based on the wind-speed-dependent failure rates, the model estimates the number of hours for each transportation type and technicians in different maintenance categories on the range of wind turbine components. The total cost is minimised by the developed model with the maintenance demand.

All the involved costs, including transportation, labour, fuel, repair material and downtime costs, are taken into account in the experimentation. Fixed and charter costs occur in hiring a required transport. Fuel cost covers the expenditures of fuel consumption per m^3. Labour cost is assumed to be the annual salary of full-time or part-time technicians. The downtime cost is calculated by the energy potentially generated during the breakdown and the wholesale electricity price. The unit cost per MWh of both preventative and corrective maintenance is estimated, as the major outcome, with all the above types of costs and the amount of energy produced. The model is also able to determine the minimum O&M cost with the best transport selection; it assists decision-makers in making a decision on the most suitable maintenance plan. A sensitivity analysis considering different scenarios, namely 50% higher mean failure rates and 50% lower mean failure rates, is given in Table 6.

As the solution of the sample case shows in Table 6, three CTVs should be scheduled to meet the annual maintenance demand; eight chartering periods of FSV and five chartering periods of JUV are required. No helicopter is scheduled to execute maintenance service, although it was an optional maintenance transport. This could result from the relatively higher costs and restricted compatibility to maintenance categories on this transportation mode. From the row "Number of technicians", it can be seen that under the normal failure rate, four full-time and two part-time technicians are hired on CTVs to meet the maintenance labour demand. The full-time technicians on FSV or JUV can be deployed to work temporarily on CTVs. More part-time technicians will be used than full-time on FSVs and JUVs since major repair and replacement do not occur throughout the whole year.

Table 6. Comparison of sample case results with higher and lower failure rates.

	Sample Case	**50% Higher Failure Rates**	**50% Lower Failure Rates**
Maintenance working hours required			
Crew transfer vessel	6479	8677	2892
Field support vessel	632	996	325
Jack-up vessel	793	1440	418
Helicopter	0	0	0
Number of transport charter periods required			
Crew transfer vessel (charter period: 1 year)	3	5	2
Field support vessel (charter period: 1 week)	8	12	4
Jack-up vessel (charter period: 2 weeks)	5	9	3
Helicopter (charter period: 3 weeks)	0	0	0

Table 6. *Cont.*

	Sample Case	50% Higher Failure Rates	50% Lower Failure Rates
Number of technicians required			
F/T Crew transfer vessel	4	7	3
P/T Crew transfer vessel	2	3	1
F/T Field support vessel	1	2	0
P/T Field support vessel	4	3	5
F/T Jack-up vessel	8	15	1
P/T Jack-up vessel	13	7	20
F/T Helicopter	0	0	0
P/T Helicopter	0	0	0
Estimated costs			
Preventative maintenance cost per MWh (£)	6.70	6.70	6.70
Corrective maintenance cost per MWh (£)	19.42	31.95	9.23
Total maintenance cost per MWh (£)	26.12	38.65	15.93

In the 50% higher failure rates scenario, as shown in Table 6, the number of CTVs demonstrates an increase by two; and longer charter leases of FSV and JUV are also requested to satisfy the increased maintenance demands. Three additional full-time technicians and one additional part-time technician on CTVs are needed to match the maintenance workload with a 50% increase in failure rates. On the field support vessel, one more full-time technician replaces one part-time technician in order to provide more service time. A significant change is shown in the number of technicians on the Jack-up vessel; seven extra full-time persons are employed with a reduction of six part-time technicians. By considering 50% lower failure rates, more part-time technicians are hired on all types of transport because part-time employees are more cost-effective to satisfy the decreased maintenance demand.

The cost distribution is investigated for the three scenarios. According to the average results of the scenarios, transport charter cost contributes 51% of the total maintenance cost (shown in Figure 5). Loss of energy production during downtime gives 23%, and transport fixed cost occupies 15%. If the chartering period of FSV and JUV is extended to two months or longer per lease, then the total transport fixed cost could be reduced. Labour cost presents a small percentage (3%), and transportation fuel is less than 1% of the total cost.

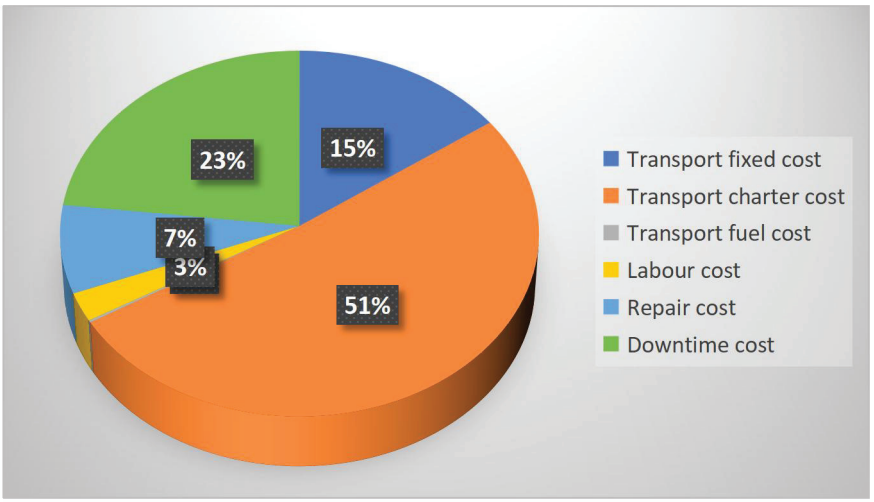


Figure 5. Distribution of each cost element in offshore wind maintenance.

6. Conclusions

Although offshore wind technology has been developed rapidly during the last decades, there are a limited number of optimisation models available to support O&M planning activities. This paper proposes a decision-making model to assist offshore wind project developers in planning cost-effective O&M decisions. The optimisation model aims to minimise the total cost of O&M activities, including transport fixed cost, transport charter cost, transport fuel cost, labour cost, repair cost and downtime cost, in offshore wind maintenance during a given period of time.

Five categories of maintenance and key components of wind turbines are considered in the developed model in order to produce realistic results. A transport type may be used to undertake maintenance activities on multiple maintenance categories. For example, jack-up vessels are compatible with carrying out major repairs and major replacements. Technicians may also be used flexibly between different transport types to execute different maintenance categories. For instance, the technicians on jack-up vessels or field support vessels are entitled to work on crew transfer vessels for smaller repairs. Wind-speed-dependent failure rates on different turbine components were considered in the optimisation model. The model takes into account the wind speed probabilities in a particular OWF, so it is able to supply a practical solution. By using the model solution, the effect of seasonal changes on wind turbine operation can be accounted for.

The results obtained from the optimisation model are able to contribute effectively to the planning of O&M resources and activities in advance to meet the necessary maintenance demand. Both the required transportation and labour will be used effectively to improve the performance cost. The revenue loss during downtime is regarded as another key element in O&M cost. According to the sensitivity analysis, the experimental results considering the sample OWF imply that the reliability of wind turbine components has an immediate effect on maintenance costs. Therefore, this proposed model can support offshore wind stakeholders in understanding the strategic resource requirement associated with the maintenance of an OWF.

Both transport and labour are utilised effectively in the optimal solutions. However, the utilisation could potentially be included as further objectives in the optimisation model rather than considering cost-related objectives only. A balance between service efficiency and cost-effectiveness could be achieved with a multi-objective optimisation model. The effect of sea state, such as wind speed, is one of the most significant factors causing uncertainty in the maintenance planning of OWFs. Weather forecast on a short timescale could be accurate, but it is not sufficient to support the strategic plan of offshore wind maintenance. Any sea state changes may result in a significantly different solution from that predicted by the mean value.

The Weibull distribution was applied for the weather forecast based on historical weather data in the OWF location. It is suggested to fit multiple weather scenarios into this optimisation model and different occurrence probabilities in each scenario. Alternatively, other stochastic modelling techniques might be considered for weather simulation, such as a Markov chain, and integrated into the decision-making model. Different stochastic models require different data inputs. Hence, this model can be deployed in a wider range of realistic cases with various data availability. Finally, the correlation of preventive maintenance with component failures can be investigated and included in the model as an extra parameter in future research.

Author Contributions: Conceptualization, X.L. and D.O.; Methodology, X.L. and X.S.; Validation, X.S.; Resources, X.S.; Writing—original draft, X.L.; Writing—review & editing, D.O.; Visualization, X.L.; Supervision, D.O.; Funding acquisition, X.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to the data format.

Acknowledgments: The paper presents the work as an extension of the project outcome we have developed for the 2OM (Offshore Operations & Maintenance Mutualisation) project, which is funded by the EU Interreg IVA France (Channel)—England Programme. This paper is an outcome of the research work supported by the EURO General Support Fund from the Association of European Operational Research Societies.

Conflicts of Interest: The authors declare no conflict of interest.

Nomenclature or Abbreviations

Nomenclatures:	
O&M	Operation and maintenance
OWF	Offshore wind farm
CTV	Crew transfer vessel
FSV	Field support vessel
JUV	Jack-up vessel
HEL	Helicopter
Sets:	
$i \in I$	Set of maintenance categories
$j \in J$	Set of wind turbine components
$k \in K$	Set of transport types
$w \in W$	Set of wind speeds
Parameters:	
C_k^{fixed}	Fixed cost of transport type k in a charter period
$C_k^{charter}$	Charter cost of transport type k per day
C_k^{fuel}	Fuel cost per m^3 of transport type k
$C_k^{labourFT}$	Labour cost of a full-time technician working on transport type k
$C_k^{labourPT}$	Labour cost of a part-time technician working on transport type k
C_{ij}^{repair}	Repair material cost of maintenance category i on component j
C^{down}	Expected downtime cost per MWh
D	Distance to maintenance base port
N	Number of wind turbines
Q_{ij}	Number of technicians required for maintenance category i on component j
F_{ij}	Mean failure rate of category i on component j
T_k^{travel}	Travel time of transport k from the base port to OWF
T_{ij}^{repair}	Expected repair time of category i on component j
T_{kw}^{wait}	Expected waiting time of transport k under wind speed w
S_k	Speed for transport k
O_k	Fuel consumption (m^3 per hour) of transport k
$H_k^{transport}$	Total hours per chartering period of transport k
$H_k^{labourFT}$	Total hours of a full-time technician on transport k
$H_k^{labourPT}$	Total hours of a part-time technician on transport k
L_k^{shift}	Working hours of a shift on transport k
$B_{ik}^{compatible}$	=1, if transport k is compatible with category i =0, otherwise
$B_{kw}^{deployable}$	=1, if transport k is deployable under wind speed w =0, otherwise
$Prob_w^{wind}$	Probability of wind speed w
$Prob_{w,ij}^{wind\|fail}$	Probability of wind speed w, given a failure occurred to category i on component j
$\lambda_{ij,w}^{fail\|wind}$	Probability of failure of category i on component j under wind speed w
G_w^{avail}	Power generated per hour under wind speed w

Cap^{rated}	Rated capacity of a wind turbine
A	Swept area of a wind turbine
ρ	Air density
Coe^{p}	Coefficient of the power of a wind turbine
WS_{w}	Value of wind speed w
WS^{in}	Cut-in wind speed
WS^{out}	Cut-out wind speed
WS^{rated}	Rated wind speed
Decision variables:	
x_{k}^{α}	Number of charter periods of transport k required
x_{kw}^{β}	Number of charter periods of transport k required under wind speed w
x_{ikw}^{γ}	Number of charter periods of transport k required for category i under wind speed w
x_{ijkw}^{δ}	Number of charter periods of transport k required for category i on component j under wind speed w
$y_{k}^{FT,\alpha}$	Number of full-time technicians on transport k
$y_{kw}^{FT,\beta}$	Number of full-time technicians on transport k under wind speed w
$y_{ikw}^{FT,\gamma}$	Number of full-time technicians on transport k for category i under wind speed w
$y_{ijkw}^{FT,\delta}$	Number of full-time technicians on transport k for category i on component j under wind speed w
$y_{k}^{PT,\alpha}$	Number of part-time technicians on transport k
$y_{kw}^{PT,\beta}$	Number of part-time technicians on transport k under wind speed w
$y_{ikw}^{PT,\gamma}$	Number of part-time technicians on transport k for category i under wind speed w
$y_{ijkw}^{PT,\delta}$	Number of part-time technicians on transport k for category i on component j under wind speed w
z_{ijkw}	= 1, if transport k is selected for category i on component j under wind speed w = 0, otherwise

References

1. Nielsen, J.J.; Sørensen, J.D. On risk-based operation and maintenance of offshore wind turbine components. *Reliab. Eng. Syst. Saf.* **2011**, *96*, 218–229. [CrossRef]
2. Gintautas, T.; Sørensen, J.D.; Vatne, S.R. Towards a Risk-based Decision Support for Offshore Wind Turbine Installation and Operation & Maintenance. *Energy Procedia* **2016**, *94*, 207–217. [CrossRef]
3. Li, X.; Ouelhadj, D.; Song, X.; Jones, D.; Wall, G.; Howell, K.E.; Igwe, P.; Martin, S.; Song, D.; Pertin, E. A decision support system for strategic maintenance planning in offshore wind farms. *Renew. Energy* **2016**, *99*, 784–799. [CrossRef]
4. Halvorsen-Weare, E.E.; Gundegjerde, C.; Halvorsen, I.B.; Hvattum, L.M.; Nonås, L.M. Vessel Fleet Analysis for Maintenance Operations at Offshore Wind Farms. *Energy Procedia* **2013**, *35*, 167–176. [CrossRef]
5. Scheu, M.; Matha, D.; Hofmann, M.; Muskulus, M. Maintenance Strategies for Large Offshore Wind Farms. *Energy Procedia* **2012**, *24*, 281–288. [CrossRef]
6. Ding, F.; Tian, Z. Opportunistic maintenance for wind farms considering multi-level imperfect maintenance thresholds. *Renew. Energy* **2012**, *45*, 175–182. [CrossRef]
7. Pahlke, T. Software & Decision Support Systems for Offshore Wind Energy Exploitation in North Sea Region. A report by Overspeed GmbH & Co. KG to the POWER project 2007. Available online: http://pcoe.nl/@api/deki/files/1900/=12wp1_executivesummary_sdss-studie_2007-06-05.pdf (accessed on 19 June 2023).
8. Hofmann, M.; Sperstad, I.B. NOWIcob—A Tool for Reducing the Maintenance Costs of Offshore Wind Farms. *Energy Procedia* **2013**, *35*, 177–186. [CrossRef]
9. Rademakers, L.W.M.M.; Braam, H.; Obdam, T.S.; Pieterman, R.P.V.D. Operation and maintenance cost estimator (OMCE) to estimate the future O&M costs of offshore wind farms. In Proceedings of the European Offshore Wind 2009 Conference, Stockholm, Sweden, 14-16 September 2009. Available online: http://www.ecn.nl/docs/library/report/2009/m09126.pdf (accessed on 19 June 2023).
10. Hofmann, M. A Review of Decision Support Models for Offshore Wind Farms with an Emphasis on Operation and Maintenance Strategies. *Wind Eng.* **2011**, *35*, 1–16. [CrossRef]
11. Zhang, Z.; Kusiak, A.; Song, Z. Scheduling electric power production at a wind farm. *Eur. J. Oper. Res.* **2013**, *224*, 227–238. [CrossRef]

12. Gundegjerde, C.; Halvorsen, I.B.; Halvorsen-Weare, E.E.; Hvattum, L.M.; Nonås, L.M. A stochastic fleet size and mix model for maintenance operations at offshore wind farms. *Transp. Res. Part C: Emerg. Technol.* **2015**, *52*, 74–92. [CrossRef]
13. Stålhane, M.; Halvorsen-Weare, E.; Nonås, L.M. A Decision Support System for Vessel Fleet Analysis for Maintenance. Unpublished Working Paper; 2016. Available online: http://www.sintef.no/globalassets/project/maritime-logistics-optimization-for-offshore-wind/staalhaneetal_wp_2016.pdf (accessed on 19 June 2023).
14. Irawan, C.A.; Ouelhadj, D.; Jones, D.; Stålhane, M.; Sperstad, I.B. Optimisation of maintenance routing and scheduling for offshore wind farms. *Eur. J. Oper. Res.* **2017**, *256*, 76–89. [CrossRef]
15. Fan, D.; Ren, Y.; Feng, Q.; Zhu, B.; Liu, Y.; Wang, Z. A hybrid heuristic optimization of maintenance routing and scheduling for offshore wind farms. *J. Loss Prev. Process. Ind.* **2019**, *62*, 103949. [CrossRef]
16. Stålhane, M.; Halvorsen-Weare, E.E.; Nonås, L.M.; Pantuso, G. Optimizing vessel fleet size and mix to support maintenance operations at offshore wind farms. *Eur. J. Oper. Res.* **2019**, *276*, 495–509. [CrossRef]
17. Gutierrez-Alcoba, A.; Hendrix, E.; Ortega, G.; Halvorsen-Weare, E.; Haugland, D. On offshore wind farm maintenance scheduling for decision support on vessel fleet composition. *Eur. J. Oper. Res.* **2019**, *279*, 124–131. [CrossRef]
18. Turan, H.H.; Elsawah, S.; Ryan, M.J. A long-term fleet renewal problem under uncertainty: A simulation-based optimization approach. *Expert Syst. Appl.* **2020**, *145*, 113158. [CrossRef]
19. Irawan, C.A.; Eskandarpour, M.; Ouelhadj, D.; Jones, D. Simulation-based optimisation for stochastic maintenance routing in an offshore wind farm. *Eur. J. Oper. Res.* **2021**, *289*, 912–926. [CrossRef]
20. Scheu, M.N.; Tremps, L.; Smolka, U.; Kolios, A.; Brennan, F. A systematic Failure Mode Effects and Criticality Analysis for offshore wind turbine systems towards integrated condition based maintenance strategies. *Ocean Eng.* **2019**, *176*, 118–133. [CrossRef]
21. McMorland, J.; Collu, M.; McMillan, D.; Carroll, J. Operation and maintenance for floating wind turbines: A review. *Renew. Sustain. Energy Rev.* **2022**, *163*, 112499. [CrossRef]
22. Karyotakis, A.; Bucknall, R. Planned intervention as a maintenance and repair strategy for offshore wind turbines. *J. Mar. Eng. Technol.* **2010**, *9*, 27–35. [CrossRef]
23. Kooijman, H.J.T.; Noord, M.D.; Uyterlinde, M.A.; Wals, A.F.; Herman, S.A.; Beurskens, H.J.M. Large Scale Offshore Wind Energy in the North Sea—A Technology and Policy Perspective. *Wind. Eng.* **2004**, *28*, 143–156. [CrossRef]
24. Carroll, J.; McDonald, A.; McMillan, D. Failure rate, repair time and unscheduled O&M cost analysis of offshore wind turbines. *Wind Energy* **2016**, *19*, 1107–1119. [CrossRef]
25. Wilson, G.; McMillan, D. Modelling the impact of the environment on offshore wind turbine failure rates. In Proceedings of the EWEA Offshore Conference, Frankfurt, Germany, 19–21 November 2013.
26. Wilson, G.; McMillan, D. Assessing wind farm reliability using weather dependent failure rates. *J. Phys. Conf. Ser.* **2014**, *524*, 012181. [CrossRef]
27. Garcia, A.; Torres, J.; Prieto, E.; de Francisco, A. Fitting wind speed distributions: A case study. *Sol. Energy* **1998**, *62*, 139–144. [CrossRef]
28. Krokoszinski, H.-J. Efficiency and effectiveness of wind farms—Keys to cost optimized operation and maintenance. *Renew. Energy* **2003**, *28*, 2165–2178. [CrossRef]
29. van Bussel, G.J.W.; Bierbooms, W.A.A.M. The DOWEC Offshore Reference Windfarm: Analysis of transportation for operation and maintenance. *Wind. Eng.* **2003**, *27*, 381–391. [CrossRef]
30. Obdam, T.; Rademakers, L.; Bramm, H.; Eecen, P. Estimating costs of operation & maintenance for offshore wind farms. In Proceedings of the Offshore Wind Energy Conference, Berlin, Germany, 7–10 May 2007.
31. Tavner, P.J.; Xiang, J.; Spinato, F. Reliability analysis for wind turbines. *Wind Energy Int. J. Progress Appl. Wind Power Convers. Technol.* **2007**, *10*, 1–18. [CrossRef]
32. Beerbühl, S.S.; Fröhling, M.; Schultmann, F. Combined scheduling and capacity planning of electricity-based ammonia production to integrate renewable energies. *Eur. J. Oper. Res.* **2015**, *241*, 851–862. [CrossRef]
33. Berbeglia, G.; Cordeau, J.-F.; Gribkovskaia, I.; Laporte, G. Rejoinder on: Static pickup and delivery problems: A classification scheme and survey. *TOP* **2007**, *15*, 45–47. [CrossRef]
34. Stålhane, M.; Hvattum, L.M.; Skaar, V. Optimization of Routing and Scheduling of Vessels to Perform Maintenance at Offshore Wind Farms. *Energy Procedia* **2015**, *80*, 92–99. [CrossRef]
35. Li, M.; Jiang, X.; Carroll, J.; Negenborn, R.R. A multi-objective maintenance strategy optimization framework for offshore wind farms considering uncertainty. *Appl. Energy* **2022**, *321*, 119284. [CrossRef]
36. Indhumathy, D.; Seshaiah, C.V.; Sukkiramathi, K. Estimation of Weibull parameters for wind speed calculation at Kanya-kumari in India. *Int. J. Innovative Res. Sci. Eng. Technol.* **2014**, *3*, 8340–8345.
37. Aksoy, H.; Toprak, Z.F.; Aytek, A.; Ünal, N.E. Stochastic generation of hourly mean wind speed data. *Renew. Energy* **2004**, *29*, 2111–2131. [CrossRef]
38. Jiang, H.; Li, Y.; Cheng, Z. Performances of ideal wind turbine. *Renew. Energy* **2015**, *83*, 658–662. [CrossRef]

39. Dinwoodie, I.; Endrerud, O.-E.V.; Hofmann, M.; Martin, R.; Sperstad, I.B. Reference Cases for Verification of Operation and Maintenance Simulation Models for Offshore Wind Farms. *Wind. Eng.* **2015**, *39*, 1–14. [CrossRef]
40. Dalgic, Y.; Lazakis, I.; Dinwoodie, I.; McMillan, D.; Revie, M. Advanced logistics planning for offshore wind farm operation and maintenance activities. *Ocean Eng.* **2015**, *101*, 211–226. [CrossRef]

mathematics

MDPI

Article

GIS-Integrated Data Analytics for Optimal Location-and-Routing Problems: The GD-ARISE Pipeline

Jun-Jae Won [1], Jong-Seung Lee [2] and Hyung-Tae Ha [1,2,*]

[1] Department of Applied Statistics, Gachon University, 1342 Seongnam-daero, Sujeong-gu, Seongnam-si 13120, Gyeonggi-do, Republic of Korea; willy49@gachon.ac.kr
[2] Department of Next Generation Smart Energy System Convergence, Gachon University, 1342 Seongnam-daero, Sujeong-gu, Seongnam-si 13120, Gyeonggi-do, Republic of Korea; new9279@gachon.ac.kr
* Correspondence: htha@gachon.ac.kr

Abstract

Optimizing the siting and servicing of urban facilities is a core operations research problem that must reconcile heterogeneous demand, spatial constraints, and network-realistic travel. We present GD-ARISE, a GIS-integrated and data analytics pipeline that maintains a pedestrian–road network metric from demand inference through siting to routing. The workflow has three modules: (i) GIS integration that unifies spatial layers on one network and distance metric; (ii) data analytics that builds multi-criteria suitability via the Analytic Hierarchy Process (AHP) and maps scores to adaptive service radii; (iii) optimal location-and-routing that selects nonoverlapping sites with a transparent greedy rule (SCASS) and computes depot-to-depot routes via simulated annealing on the same metric. A case study in Seoul's Gangnam District yields a high-coverage portfolio and feasible collection routes. We add a theoretical framework that casts SCASS as a conflict-graph problem, document the AHP elicitation with consistency checks, and report robustness analyses including sensitivity to AHP weights and to radius bounds. Results indicate that core hotspots remain stable to weighting, whereas mid-range corridors shift as criteria priorities or spatial parameters change.

Keywords: optimal location-and-routing problem; urban waste management; GIS integration; data analytics; analytic hierarchy process; maximal covering location problem; adaptive coverage; simulated annealing

MSC: 90B80

1. Introduction

1.1. Motivation

Urban waste disposal and storage can be framed as a rigorously defined operations research location-and-routing problem that is driven by demand data analytics and is GIS-native [1]. In megacities, rapid urbanization and rising population density intensify pressure on municipal services; public waste management often struggles to keep pace. A core cause is geometric mismatch: bins are not where people generate waste, and service routes ignore how people and vehicles actually move [1,2]. The result is a network-based operations research problem in which demand must be inferred, sites must be chosen, and service must be routed using a coherent metric and data stack. This is now feasible at the city scale because municipalities publish high-frequency geospatial datasets (e.g.,

floating populations, transit nodes, and points of interest), and open platforms such as OpenStreetMap, with OSMnx, enable the computation of network-consistent distances [3]. Planners increasingly require transparent analytics that map directly to policy levers—how many bins to deploy, where to place them, and how to service them given fleet and budget constraints—so there is a timely opportunity to connect demand estimation, siting, and routing under one geometry and one metric [2,4]. Despite practical need, algorithms for urban amenity siting that integrate heterogeneous demand data with geometry-consistent spatial and operational constraints have received limited attention. Few studies, to our knowledge, claim an end-to-end algorithm that explicitly integrates GIS-based geometry, demand-driven multi-criteria scoring that feeds back into spatial parameters, nonoverlap enforced with adaptive radii on a network metric, and routing on the same metric. Most prior workflows stop at ranking without altering geometry, mix Euclidean screening with network routing, or impose fixed buffers that ignore local demand. As a result, geometry-consistent siting-and-routing pipelines remain rare in the literature, leaving a gap for methods that are both operationally credible and analytically transparent.

1.2. Related Works

The literature on multi-criteria decision-making (MCDM), the maximal covering location problem (MCLP), and routing has been extensively developed for facility location in operations research. GIS strengthens these models by linking geospatial data and network analysis [3,5], enabling city-scale location allocation, coverage, and access studies [6]. For multi-criteria synthesis, the Analytic Hierarchy Process (AHP) [7] and related MCDM methods are widely used to build suitability surfaces for siting and infrastructure planning [6]. This approach typically involves weighting heterogeneous GIS layers to derive a composite suitability index. It has been widely applied across various domains, including landfill siting [8,9], healthcare facility planning [10], and renewable energy projects [11,12]. These studies formalize how multi-criteria weighting translates spatially into suitability gradients, underscoring MCDM's capacity to operationalize complex environmental and social criteria.

Church and ReVelle introduced MCLP, which chooses sites to cover as much demand as possible within a set distance [13]. Many studies have extended this idea to include budgets, fairness goals, and changes over time or uncertainty [14–17]. Ref. [18] investigated model placement as a single goal of maximizing coverage. In solid-waste systems, recent reviews summarize how covering models are used and point out that results can change a lot depending on the distance measure and the service radius assumed [19,20]. A primary critique of the classical MCLP concerns its rigid, binary definition of coverage. Berman et al. (2003) [21] addressed the model's unrealistic "all-or-nothing" assumption by introducing a gradual decay function, where service quality diminishes with distance rather than ceasing abruptly. Complementing this, Karasakal and Karasakal (2004) [22] highlighted that strict binary formulations can yield unjustified solutions where partial service is plausible, thereby motivating variants that permit partial coverage. More recent extensions have broadened the MCLP's focus from pure efficiency to include equity. Blanco and Gázquez (2023) [23] formalized the integration of fairness, employing concepts such as α-fairness and ordered-weighted objectives. Their work underscores that distributional equity is not inherent in the basic model and must be incorporated as an explicit objective. Despite these theoretical advancements, practical siting decisions increasingly rely on spatial data and multi-criteria integration. However, using such derived scores to dynamically adjust geometric or operational parameters, such as service radii or coverage decay, remains limited [24–26]. Spatial feasibility often requires nonoverlap [27,28], which connects to disk packing [29] and independent set formulations [30]; in practice, spacing is frequently

enforced with ad hoc buffers rather than with network metric optimization. In algorithms, greedy heuristics are common for NP-hard location models because they are fast and transparent [31], though not globally optimal [14,15,32].

For routing, simulated annealing [33,34] and other local search methods perform well on the traveling salesman problem (TSP) [35] and on modern VRP variants [36]. Recent applied work also moves toward joint decisions on where to locate facilities and how to route the service.

Rahmanifar et al. (2024) [37] present a non-linear multiobjective model that integrates warehouse location with vehicle routing in cold-chain logistics, and Hashemi-Amiri et al. (2023) [38] propose a tri-objective mixed-integer linear program (MILP) that unifies facility location, crew scheduling, and routing for municipal solid waste. These studies advance integrated planning, but they do not maintain a single, GIS-consistent distance metric from multi-criteria demand aggregation through siting constraints to downstream routing.

1.3. Contributions

This paper introduces GD-ARISE (GIS-integrated and Data analytic Adaptive Radius Integrated Siting and rEservicing), an end-to-end, GIS-integrated pipeline that maintains geometric consistency from demand analytics to operations. First, all spatial layers—administrative boundaries, pedestrian–road networks, floating populations, transit nodes, and waste-related points of interest—are reconciled onto a single pedestrian–road network and one distance metric, ensuring that measurements and decisions are expressed in the same geometry. Second, multi-criteria demand is constructed via AHP into a composite suitability score, and then mapped to geometry as an adaptive service radius, so that high-suitability areas receive smaller radii while lower-suitability areas receive larger radii to preserve access. Third, spatially constrained selection is formulated with site-specific radii on the network metric and implemented via a transparent greedy rule (SCASS) to produce a maximal nonoverlapping portfolio; the associated conflict structure admits an interpretation as a graph independent set problem. Fourth, depot-to-depot service routes are generated using simulated annealing on the identical network metric used upstream, closing the loop without changing geometry. Finally, the pipeline is demonstrated on waste bin siting and servicing in Seoul's Gangnam District, where the walkable network is sampled at fine resolution, composite scores and adaptive radii are computed, a nonoverlapping portfolio of sites is selected, and operational routes are produced, including a single depot case in Samsung 1–dong. By maintaining one geometry and one metric across all stages, GD-ARISE turns GIS analytics into operations research decisions about counts, placement, spacing, and service effort under realistic constraints.

2. An Integrated Planning Algorithm: GD-ARISE

We present GD-ARISE, a unified, GIS-integrated workflow for optimal location-and-routing. Figure 1 shows the workflow of the GD-ARISE algorithm, which has three modules: GIS integration, data analytics, and optimal location-and-routing. All steps use one network distance metric on the pedestrian–road network. This keeps geometry and units consistent. In GIS integration, load the administrative boundaries, the pedestrian–road network, floating population layers, transit nodes, and waste-related points of interest. Harmonize coordinate reference systems, clean fields, and clip layers to the study region. In data analytics, generate dense candidate sites along walkable streets. For each candidate site, compute criterion scores from the GIS layers such as population, transit access, and proximity to points of interest. Normalize all scores to $[0, 1]$. Use AHP to set weights and combine the criteria into a composite suitability. In optimal location-and-routing, first select sites with SCASS. Sort candidates by suitability and add a site if its radius does not overlap

any already selected radius. Then route service to build depot-to-depot tours. The workflow also reports the coverage attained, the number of selected sites, the route length, and the route time. Because all steps share one metric and one GIS substrate, results are easy to audit, map, and reproduce.

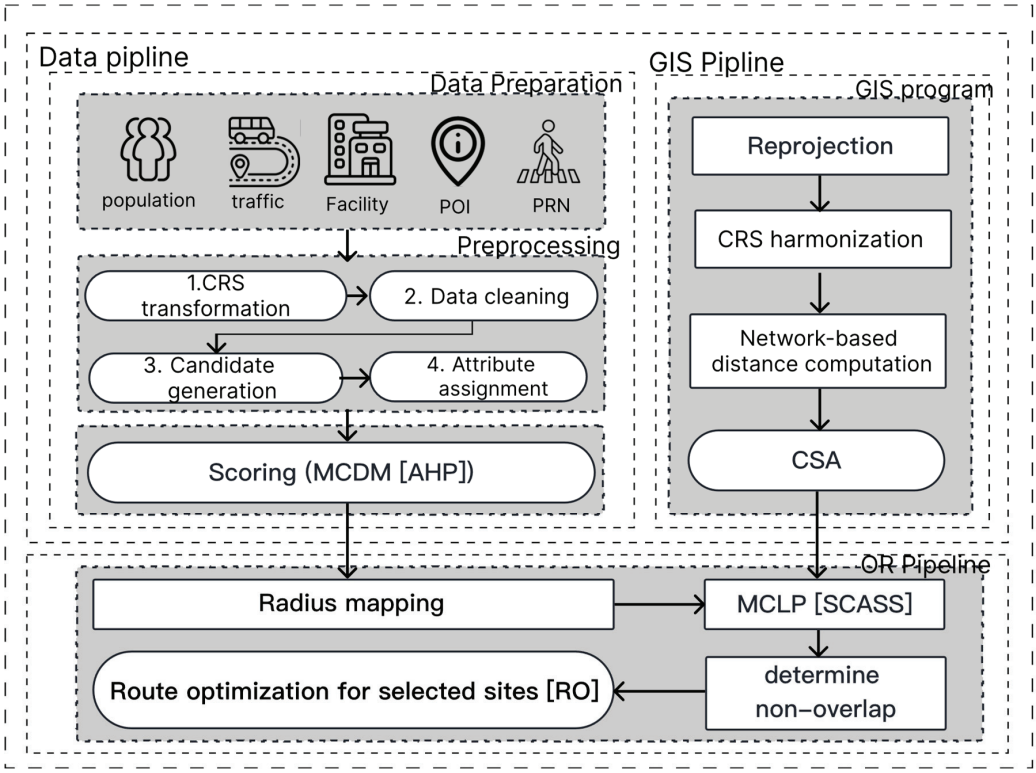


Figure 1. Workflow of the GD-ARISE.

We work on a single geographic region $L \subset \mathbb{R}^2$ with a pedestrian–road network-induced distance $d : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}_{\geq 0}$ (shortest path on the network). The region is discretized into candidate sites $J = \{j_1, \ldots, j_N\}$ sampled from the network at resolution $\delta > 0$, so adjacent samples are at most δ apart under d. A set of criteria $\mathcal{C} = \{C_1, \ldots, C_M\}$ evaluates each site. We assign to every $j \in J$ a composite suitability $S_j^* \in [0, 1]$ and an adaptive service radius $r_j \in [R_{\min}, R_{\max}]$ with $0 < R_{\min} < R_{\max} < \infty$. Then, we select a subset $J_{\text{sel}} \subseteq J$ of size $P_{\text{target}} \in \mathbb{N}$ while enforcing nonoverlap based on d and $\{r_j\}$. Finally, we route the service from a designated depot $D_0 \in L$. When routing is required, each served site carries demand $q_j \geq 0$, the fleet has $m \in \mathbb{N}$ vehicles, and each vehicle has capacity $Q > 0$. All distances and constraints use the same metric d.

Multi-criteria demand assessment will be performed by mapping heterogeneous raw measurements into a composite demand suitability score S_j^* and a site-specific spatial footprint r_j. The inputs are the raw criterion values or proximity-based syntheses for each $C_k \in \mathcal{C}$, together with a pairwise comparison matrix that encodes decision priorities. Each criterion is normalized to $[0, 1]$ via a monotone transformation so that larger values consistently mean greater desirability, and AHP weights are extracted from the principal eigenvector of the comparison matrix, subject to a consistency check. Aggregation by a convex combination yields $S_j^* = \sum_k \alpha_k S_{j,k}$. This first stage resolves two design tensions in a data-driven yet interpretable way: it fuses many incommensurate predictors of demand into a single, dimensionless score, and it ties spatial influence to local desirability so that high-value sites are modeled with tighter catchments while low-value sites expand to preserve access.

For every criterion $C_k \in \mathcal{C}$ and site $j \in J$, the non-negative value $V_{j,k} \in \mathbb{R}_{\geq 0}$ quantifies the magnitude of C_k at j. Two constructions cover the settings of interest. When C_k is directly observed at the site—such as floating population, footfall, or local demand intensity—we set $V_{j,k} = P_j$, where P_j denotes the measured quantity at j. When C_k reflects the influence of external features distributed in space, we let $F_k = \{f_s\}_{s=1}^{S_k} \subset L$ denote the relevant feature set and assign non-negative weights $\{\omega_{k,s}\}_{s=1}^{S_k}$ that sum to one and encode the relative importance of individual features or subtypes. A influence kernel $K : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ then maps distances to contributions so that

$$V_{j,k} = \sum_{s=1}^{S_k} \omega_{k,s}\, K\big(d(j, f_s)\big). \tag{1}$$

In the case study, we adopt $K(\rho) = \rho$ with $d(j, f_s)$, measured as Euclidean distance.

Because the criteria are measured in heterogeneous units, each raw value is transformed to a standardized unit-interval score via a criterion-specific monotone standardization map $f_k : \mathbb{R}_{\geq 0} \to [0, 1]$ defined by $S_{j,k} = f_k(V_{j,k})$. When C_k is a benefit-type attribute for which larger values are more desirable, a min–max mapping places all sites on a common scale,

$$S_{j,k} = \begin{cases} \dfrac{V_{j,k} - \min_{\ell \in J} V_{\ell,k}}{\max_{\ell \in J} V_{\ell,k} - \min_{\ell \in J} V_{\ell,k}}, & \text{if } \max_{\ell \in J} V_{\ell,k} > \min_{\ell \in J} V_{\ell,k}, \\ 0, & \text{otherwise,} \end{cases} \tag{2}$$

sending the empirical minimum to 0 and the empirical maximum to 1. To translate composite suitability into a spatial service footprint, two design constants $0 < R_{\min} < R_{\max}$ specify the admissible range of coverage radii. The adaptive coverage radius at site j is then defined by the affine mapping:

$$r_j = R_{\max} - (R_{\max} - R_{\min})\, S_j^*. \tag{3}$$

Because $S_j^* \in [0, 1]$, this definition guarantees $r_j \in [R_{\min}, R_{\max}]$. Differentiating shows that $\partial r_j / \partial S_j^* = -(R_{\max} - R_{\min}) < 0$, so sites with higher suitability are assigned proportionally smaller catchments by concentrating service in areas of strong demand while allowing sites in weaker areas to expand their reach to preserve access. The use of a network-based metric d is essential in urban contexts when it captures the true impedance of travel.

Now, we design an algorithm for selecting locations via adaptive coverage and greedy nonoverlap, so-called *SCASS*, in the spirit of MCLP. The aim is to choose the optimal number of target sites P_{target} that maximizes the amount of demand covered within the heterogeneous radii $\{r_j\}_{j \in J}$.

Proposition 1 (SCASS). *Let $L \subset \mathbb{R}^2$ be a geographic region endowed with a network metric d, let $J = \{j_1, \ldots, j_N\} \subset L$ be candidate sites with adaptive radii $\{r_j\}_{j \in J}$ determined from composite scores $\{S_j^*\}_{j \in J}$, and let $U \subseteq L$ be finite demand nodes with weights $\{w_u\}_{u \in U}$. For each $j \in J$ define the coverage set:*

$$C(j) = \{\, u \in U : \; d(u, j) \leq r_j \,\}, \quad \textit{and the service disk} \quad \mathbb{D}_j = \{x \in L : \; d(x, j) \leq r_j\}.$$

Construct the conflict graph $G_c = (J, \mathcal{E})$ with an edge $\{i, j\} \in \mathcal{E}$ iff $d(i, j) < r_i + r_j$.

(i) Nonoverlap $\Longleftrightarrow$ Independence. A subset $J_{\text{sel}} \subseteq J$ satisfies the SCASS nonoverlap requirement $\mathbb{D}_i \cap \mathbb{D}_j = \varnothing$ for all distinct $i, j \in J_{\text{sel}}$, if and only if J_{sel} is an independent set in G_c.

(ii) Adaptive radius MCLP with exclusions. For a given cardinality $P_{\text{target}} \in \mathbb{N}$, the problem of choosing $J_{\text{sel}} \subseteq J$ with $|J_{\text{sel}}| = P_{\text{target}}$ to maximize covered demand,

$$\sum_{u \in U} w_u \, \mathbf{1}\left\{ u \in \bigcup_{j \in J_{\text{sel}}} C(j) \right\}$$

subject to SCASS nonoverlap is equivalent to the maximal covering location problem on U with site-specific radii $\{r_j\}$ and the additional constraint that J_{sel} is an independent set of G_c.

(iii) Maximum-weight independent set (MWIS). If the siting objective is to maximize $\sum_{j \in J_{\text{sel}}} S_j^*$ subject to nonoverlap and $|J_{\text{sel}}| = P_{\text{target}}$, then the problem is a cardinality-constrained MWIS on G_c:

$$\max_{J_{\text{sel}} \subseteq J} \sum_{j \in J_{\text{sel}}} S_j^* \quad s.t. \quad |J_{\text{sel}}| = P_{\text{target}}, \;\; J_{\text{sel}} \text{ independent in } G_c.$$

Equivalently, with binaries $x_j \in \{0, 1\}$,

$$\begin{aligned} \max_{x \in \{0,1\}^N} \quad & \sum_{j=1}^{N} S_j^* \, x_j \\ s.t. \quad & x_i + x_j \leq 1 \quad \forall \, \{i, j\} \in \mathcal{E}, \\ & \sum_{j=1}^{N} x_j = P_{\text{target}}. \end{aligned}$$

Proof. For (i), if $\mathbb{D}_i \cap \mathbb{D}_j \neq \varnothing$, then there exists x with $d(i, j) \leq d(i, x) + d(x, j) \leq r_i + r_j$ by the triangle inequality; hence $\{i, j\} \in \mathcal{E}$ and the pair cannot be jointly selected. Conversely, if $d(i, j) \geq r_i + r_j$, then no x can lie in both disks, so $\mathbb{D}_i \cap \mathbb{D}_j = \varnothing$. This establishes the equivalence between nonoverlap and independence in G_c. Statement (ii) inserts this feasibility into the adaptive radius MCLP coverage objective, so a feasible solution is exactly an independent set of prescribed size. Statement (iii) is a direct translation of the selection objective into a cardinality-constrained MWIS, yielding the stated 0–1 formulation. □

Finally, we optimize the service route problem over the selected sites and depot using a capacitated vehicle-routing model. The node set is $V = \{0, 1, \ldots, K\}$ with $K = |J_{\text{sel}}|$. The node 0 means the depot D_0 of an operations center or garage, and node i corresponds to site $j_i \in J_{\text{sel}}$. Edge costs are the network distances $d_{ij} = d(j_i, j_j)$ with $j_0 \equiv D_0$. Demands $\{q_{j_i}\}_{i=1}^{K}$ and a common capacity Q define feasibility, and the goal is to partition J_{sel} into at most m depot-to-depot tours minimizing total travel cost while serving each site exactly once and respecting capacity. Because exact mixed-integer formulations are NP-hard at realistic scales, we adopt a simulated annealing search over permutations augmented with depot separators. A candidate solution is encoded as a sequence that starts and ends at 0 and contains each customer once, together with $m - 1$ additional depot symbols; cutting at the depot symbols yields the m routes. The total cost is the sum of d_{ij} along the sequence, and capacity feasibility amounts to verifying that the sum of demands on each between-depot segment does not exceed Q. This stage converts a strategic siting outcome into operationally feasible tours that coincide with the same network metric d used in the first two stages.

Routing operates on these selected sites and the designated depot $D_0 \in L$. To harmonize indexing, fix an arbitrary bijection between $\{1, \ldots, K\}$ and J_{sel} and write j_i for the site associated with index i. Define the node set $V = \{0, 1, \ldots, K\}$, where node 0 corresponds to the depot D_0 and node $i \in \{1, \ldots, K\}$ corresponds to site j_i. For any pair $(i, j) \in V \times V$, define $d_{ij} = d(j_i, j_j)$ with the convention $j_0 \equiv D_0$ and exclude self-loops via $x_{ii} = 0$. Each customer node $i \in \{1, \ldots, K\}$ is assigned a non-negative service demand $q_i \geq 0$ (e.g., expected daily pickups or deliveries), and vehicles have a common capacity $Q > 0$ in the

same units as the q_i. Let $m \in \mathbb{N}$ be the available fleet size. A vehicle route is a directed cycle that starts and ends at the depot, visits a subset of customers exactly once, and respects capacity. We encode routing decisions with binary arc variables $x_{ij} \in \{0, 1\}$ that indicate whether a vehicle travels directly from node i to node j. To enforce capacity and eliminate subtours, we introduce continuous load–flow variables using the classical single-commodity formulation: let y_i denote the cumulative load delivered up to and including customer i on the route that visits i, measured from zero at the depot. The capacitated vehicle routing problem on (V, d_{ij}) is then

$$\min_{x,y} \quad \sum_{i=0}^{K} \sum_{j=0}^{K} d_{ij}\, x_{ij} \tag{4}$$

subject to the depot degree constraints $\sum_{j=1}^{K} x_{0j} = m$ and $\sum_{i=1}^{K} x_{i0} = m$, the customer in- and out-degree constraints $\sum_{i=0}^{K} x_{ih} = 1$ and $\sum_{j=0}^{K} x_{hj} = 1$ for every $h \in \{1, \ldots, K\}$, and the load–flow constraints

$$y_j \;\geq\; y_i + q_j - Q\,(1 - x_{ij}) \qquad \text{for all } i \in V,\ j \in \{1, \ldots, K\}, \tag{5}$$

together with the bounds $y_0 = 0$ and $q_i \leq y_i \leq Q$ for all $i \in \{1, \ldots, K\}$, and the binary and no-nnegativity restrictions $x_{ij} \in \{0, 1\}$ and $y_i \geq 0$. The degree constraints ensure that every customer is entered and left exactly once and that exactly m tours depart from and return to the depot. The load–flow inequalities propagate cumulative load along used arcs: if $x_{ij} = 1$, then $y_j \geq y_i + q_j$. So, when a vehicle traverses (i, j), it must have delivered an additional q_j units by the time it leaves j; if $x_{ij} = 0$, the constraint is slack by at most Q. The bounds $y_0 = 0$ and $y_i \leq Q$ enforce vehicle capacity and, together with flow propagation, preclude subtours disconnected from the depot, since any positive delivery in a closed customer-only cycle would force y to grow without the possibility of resetting to 0. This mixed-integer program is NP-hard; indeed, when $m = 1$ and $Q \geq \sum_{i=1}^{K} q_i$, the problem reduces to the classical traveling salesman problem on $\{0, 1, \ldots, K\}$.

3. Application

Typical applications of the proposed algorithm include siting clinics, fire stations, micro-mobility docks, or public waste bins, each requiring a coherent pipeline. We illustrate the approach on public waste bin siting and servicing in the Gangnam District of Seoul, a dense mixed-use environment in which pedestrian flows, transit access, and commercial intensity co-produce spatially and temporally concentrated litter generation. Figure 2 presents the detailed computational workflow to apply the proposed GD-ARISE with the case study on public waste-bin siting and servicing in Seoul's Gangnam District.

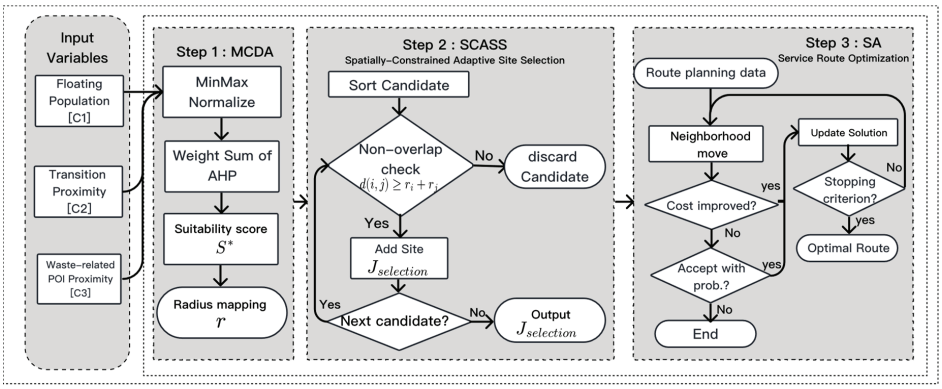


Figure 2. Computational workflow of the GD-ARISE for Gangnam District case.

3.1. GIS Analytics

The application domain is the Gangnam District of Seoul. We set the study region of the framework to $L = G \subset \mathbb{R}^2$, and we use a single distance metric $d : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}_{\geq 0}$ throughout all stages, instantiated in practice as the shortest-path metric induced by the pedestrian–road network $\mathcal{G}$ so that distances reflect walk times and barriers rather than straight lines. Gangnam's land incorporates corridors along Teheran–ro, high-street retail near Gangnam Boulevard, entertainment clusters in Apgujeong and Cheongdam, and multiple subway interchanges induce marked spatiotemporal variation in footfall. We represent this with a non-negative pedestrian density field $P : G \times [0, T] \to \mathbb{R}_{\geq 0}$ over a representative horizon $[0, T]$ (e.g., one day). Administrative boundaries for G were obtained as polygonal census layers and ingested into a GeoPandas workflow. To ensure geometric consistency across heterogeneous sources, all layers were reprojected to the common geographic CRS EPSG:4326 (WGS84) for integration with web data, and then to EPSG:5186 (Korea 2000/Central Belt) for all computations that require metric accuracy. All geoprocessing and network analyses were performed in Python (v3.11.11) using GeoPandas (v1.0.1) for spatial data handling, OSMnx (v2.0.2) for extracting the pedestrian network from OpenStreetMap, Shapely (v2.1.0) for geometric operations (e.g., buffering and distance), and Folium (v0.19.5) for interactive map visualization. Table 1 summarizes the GIS integration for the Gangnam case. It shows each spatial layer, how we prepare it, and how it is used in the model.

Table 1. GIS integration workflow for Gangnam case.

Variable Category	Source and Preprocessing	Role in Model
Administrative boundaries	Census polygons; reproject WGS84 $\to$ EPSG:5186; spatial joins to points	Study region G, attribution of J and U
Pedestrian network $\mathcal{G}$	OSM walkable edges via `OSMnx`; simplification; snapping tolerance	Distance metric d; candidate sampling; routing graph
Candidate locations J	Points every $\delta = 10$ m along $\mathcal{G}$; $N = 32{,}890$	Feasible siting set
Demand nodes U	Network vertices or grid centroids within G	Coverage evaluation, calibration, validation

Candidate facility sites are drawn from the walkable subgraph of OpenStreetMap within G. We extracted footways, pedestrian paths, sidewalks, and low-speed residential links using OSMnx, simplified the network to retain unique traversable edges, and then sampled points along these edges at a fixed network spacing $\delta = 10$ m. The result is a finite candidate set $J = \{j_1, \ldots, j_N\} \subset G$ with $N = 32{,}890$ points for feasible bin locations. Each $j \in J$ inherits attributes from intersecting administrative polygons via spatial join (e.g., sub-district codes) and is snapped to the nearest network node to avoid topological artifacts when computing $d(\cdot, \cdot)$. As shown in Figure 3, a random sample of candidate waste bin sites illustrates the spatial spread of the full feasible set across the walkable network. Blue points depict a random subset of 1000 candidates drawn from the full set of 32,890 network-anchored sites (black lines: pedestrian–road network; grey polygons: output-area boundaries). The sample visualizes coverage of feasible placements prior to scoring and selection.

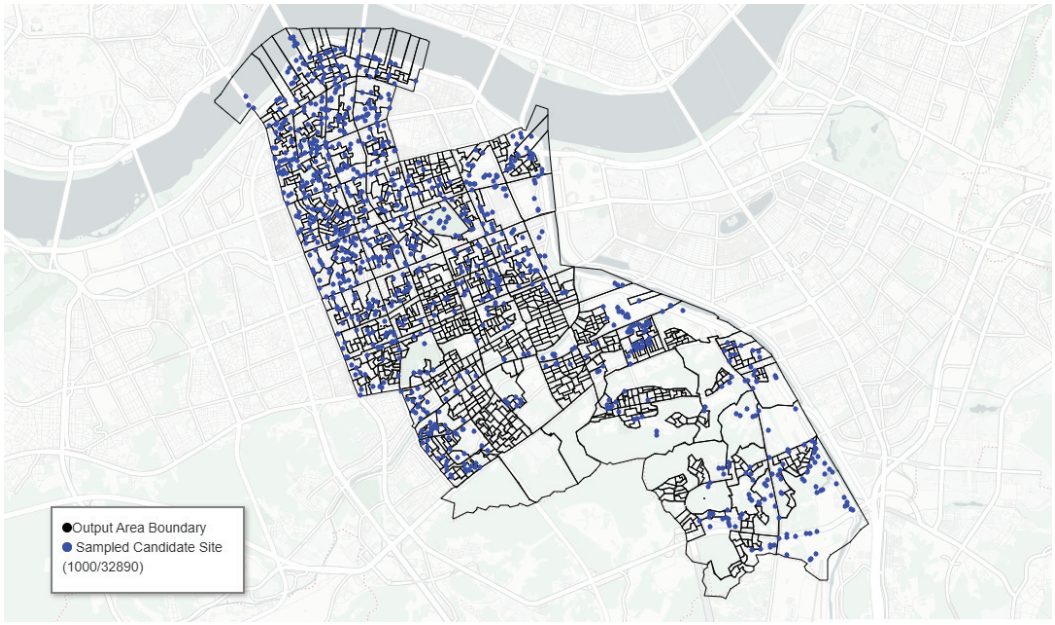


Figure 3. 1000 sample candidates of potential bin locations across Gangnam district.

3.2. Data Analytics: Demand Criteria and Feature Construction

Waste bin demand is driven by cumulative exposure to pedestrians and by proximity to attractors such as transit nodes or retail frontages. To support coverage modeling and validation, we also assemble a demand representation that is consistent with the framework's notation. Let $U \subset G$ denote a set of demand nodes at which pedestrian exposure and ancillary variables are tabulated; in practice, U may consist of network vertices in $\mathcal{G}$ inside G or centroids of census micro-polygons. All exogenous point datasets (e.g., transit stops and points of interest) are cleaned, deduplicated within a tolerance in EPSG:5186, and converted to GeoDataFrame for nearest-neighbor and kernel computations. Where a dataset is temporally indexed, we aggregate to representative daily means so that demand scores represent a typical day on the planning horizon. The dense sampling of the pedestrian network at $\delta = 10$ m yields $|J| = 32{,}890$ feasible bin locations, allowing the adaptive radii $\{r_j\}$ to respond to fine-scale variation in the built environment. Distances used in all subsequent computations—spatial-exclusion checks $d(i, j) \geq r_i + r_j$ and routing costs $d_{ij} = d(j_i, j_j)$—are evaluated with the same network metric d on $\mathcal{G}$. Table 2 summarizes the data analytics used in the Gangnam district case, which lists each variable category, how the data are sourced and preprocessed, and how they enter the model as criteria scores $S_{j,k}$ or routing inputs.

Table 2. Data analytics workflow for Gangnam District case.

Variable Category	Source and Preprocessing	Role in Model
Floating population	Daily counts by block; temporal mean; areal join to J and U	Criterion C_1: raw $V_{j,1} \to S_{j,1}$
Transit nodes	Bus stops and subway exits; nearest-neighbor on $\mathcal{G}$	Proximity criterion C_2: raw via kernel $\to S_{j,2}$
Waste-related POIs	Convenience stores, cafés, food trucks, parks; subtype weights ω	Proximity criterion C_3: raw via kernel $\to S_{j,3}$
Depot D_0, fleet and service params	Operations center location; $(m, Q, v_{\text{speed}}, \tau)$	Routing inputs and constraints

In connection with open data acquisition provided by the Seoul Open Data platform, first, Table 3 summarizes the demographic datasets integrated into the GD-ARISE pipeline for the Gangnam district case study. The datasets include Seoul's living-population esti-

mates for domestic and foreign residents, and each living-population dataset represents the estimated number of people present in a specific location and time, derived by combining administrative records (resident registry, transport, business, and building databases) with KT (Korea Telecom) big data. The census-block boundary layer (EPSG:5186) provides the spatial framework for aggregating and visualizing these population estimates. Together, these layers constitute the demographic foundation for quantifying spatial patterns of pedestrian exposure and population density within the GD-ARISE framework.

Table 3. Population and boundary datasets used for demographic analysis (C_1).

Dataset	Format/Unit	Description	Source
Seoul Living Population (Domestic Residents)	CSV (by census block)	• Daily population counts (15–21 May 2025) • Variables: DateID, TimeType, DistrictCode, CensusBlockCode, TotalPopulation • Estimated using public data (resident registry, transport, business, and building DBs) combined with KT telecom big data	https://data.seoul.go.kr/dataList/OA-14979/F/1/datasetView.do (accessed on 2 July 2025) (Seoul Open Data)
Seoul Living Population (Short-term Foreign Residents)	CSV (by census block)	• Estimated short-term foreign resident population by census block • Method identical to domestic dataset; includes mobility-based estimation using transport and telecom data	https://data.seoul.go.kr/dataList/OA-14980/F/1/datasetView.do (accessed on 2 July 2025) (Seoul Open Data)
Seoul Living Population (Long-term Foreign Residents)	CSV (by census block)	• Long-term foreign residents measured via public and telecom data integration • Same variable structure as domestic dataset	https://data.seoul.go.kr/dataList/OA-14978/F/1/datasetView.do (accessed on 2 July 2025) (Seoul Open Data)
Census Block Boundary	SHP/SBN (EPSG:5186)	• Geospatial boundaries of census blocks where living population is estimated • Spatial unit for integrating demographic data	https://data.seoul.go.kr/dataVisual/seoul/seoulLivingPopulation.do (accessed on 2 July 2025) (Seoul Open Data)

Table 4 summarizes the transit-related GIS layers integrated into the GD-ARISE pipeline to represent multimodal accessibility across Gangnam District. The datasets include bus stop locations and subway entrance coordinates, each obtained from verified public sources and re-projected to a unified coordinate reference system (EPSG:5186) for geometric consistency. The bus stop layer provides detailed node attributes, such as stop IDs, names, and coordinates, while the subway entrance layer contains manually extracted latitude–longitude pairs for all access points within the study area. Together, these layers constitute the transit node component of criterion C_2, serving as the spatial foundation for the proximity-based accessibility analysis in subsequent stages of the GD-ARISE framework. In addition, Table 5 details the source datasets and preprocessing for the waste-related POIs criterion C_3.

Now, we discretize G by sampling $\mathcal{G}$ at the spatial resolution $\delta > 0$ to obtain a candidate set of feasible bin sites $J = \{j_1, \ldots, j_N\} \subset G$, and, independently, a demand lattice $U = \{u_1, \ldots, u_M\} \subset G$. The criterion collection $\mathcal{C} = \{C_1, \ldots, C_M\}$ is expressed for waste bin siting, with observable correlates of litter pressure and disposal opportunity, in a manner that matches the demand formulation. A floating population is treated as a direct, site-specific magnitude: block-level counts from the observation period are averaged to obtain a mean daily exposure and then spatially joined to both J and U. When a candidate point or demand node falls within overlapping administrative polygons, the exposure is taken as the mean of the overlapping values, which yields well-defined raw measurements $V_{j,k}$ for the population criterion and preserves mass under areal interpolation. Public transit proximity is encoded via proximity to bus stops and subway exits. After loading both datasets and projecting to EPSG:5186, we compute for each $j \in J$ the shortest path distance along $\mathcal{G}$ to the nearest stop and to the nearest exit. Local waste generation potential is represented by proximity to POIs that tend to generate street litter, such as convenience

stores, cafés, food trucks, and public parks. Each POI set is harmonized into a single layer with subtype weights $\omega_{k,s} \geq 0$ that sum to one within the criterion.

Table 4. Transit-related datasets used for transit proximate disposal C_2.

Dataset	Format/Unit	Description	Source
Seoul Bus Stop Locations	CSV (EPSG:5186)	• NodeID, StopID • StopName • Coordinates (X, Y) • StopType	https://data.seoul.go.kr/dataList/OA-15067/S/1/datasetView.do (accessed on 2 July 2025) (Seoul Open Data)
Subway Entrance Coordinates	CSV (manual extraction)	• Latitude–Longitude pairs (manual extraction) • Subway entrance locations for accessibility criterion C_2	http://map.esran.com/ (accessed on 2 July 2025) (Seoul Open Data)

Table 5. POI datasets used for waste source facilities C_3.

Dataset	Format / Unit	Description	Source
Gangnam Food-Service Licensing (cafe_conv_stfood)	CSV (EPSG:5186)	• Fields: business name, type, address, coordinates (X, Y) • Filtered to cafés, convenience stores, food trucks; active only • Updated: 26 May 2025	https://data.seoul.go.kr/dataList/OA-18699/S/1/datasetView.do (accessed on 2 July 2025) (Seoul Open Data)
Gangnam Urban Parks	CSV (EPSG:5186)	• Fields: park name, type, latitude–longitude • Subtype of C_3 (parks; weight = 0.10) • Updated: 1 July 2024	https://data.seoul.go.kr/dataList/OA-15004/F/1/datasetView.do (accessed on 2 July 2025) (Seoul Open Data)

As shown in Figure 4, population scores are highly right-skewed, indicating many low-exposure segments and a small fraction of hotspots that dominate the upper tail. Most candidates have low values, with a long tail and few very high-exposure locations created by concentrated corridors. A floating population C_1 is a direct site-specific magnitude derived from block-level daily counts aggregated over the observation period. Let P_j denote the mean daily floating population assigned to site j via areal interpolation from its containing (or overlapping) census blocks. To place P_j on a unit scale while preserving ranks, we apply the benefit of min–max normalization over all candidates,

$$S_{j,\text{pop}} = \frac{P_j - \min_{i \in J} P_i}{\max_{i \in J} P_i - \min_{i \in J} P_i}, \tag{6}$$

which yields $S_{j,\text{pop}} \in [0, 1]$ and encodes relative pedestrian exposure.

As shown in Figure 5, disposal scores are mostly moderate with a broad mode around the mid-range, plus a spike at zero for candidates far from transit. The distribution is broadly spread with a mid-range mode and a mass at zero reflecting locations beyond the 300 m influence of both bus stops and subway exits. Transit proximate disposal opportunity C_2 is modeled as proximity to the nearest bus stop and the nearest subway exit, recognizing that on–off flows around stations correlate with both waste generation and appropriate placement of receptacles. Let $d_{j,\text{bus}}$ and $d_{j,\text{sub}}$ denote the network distances from j to the nearest bus stop and the nearest subway exit, respectively. A convex combination of linear distance decay kernels with a common maximum influence range $D_{\text{max,disp}} = 300$ m produces a unit-interval score

$$\begin{gathered} S_{j,\text{disp}} = \omega_{\text{bus}} \max\left\{0,\, 1 - \tfrac{d_{j,\text{bus}}}{300}\right\} + \omega_{\text{sub}} \max\left\{0,\, 1 - \tfrac{d_{j,\text{sub}}}{300}\right\}, \\ \text{where } \omega_{\text{bus}} = 0.75 \text{ and } \omega_{\text{sub}} = 0.25. \end{gathered} \tag{7}$$

which reflects the stronger baseline frequency of bus stops relative to subway exits, while allowing both to contribute when they are nearby.

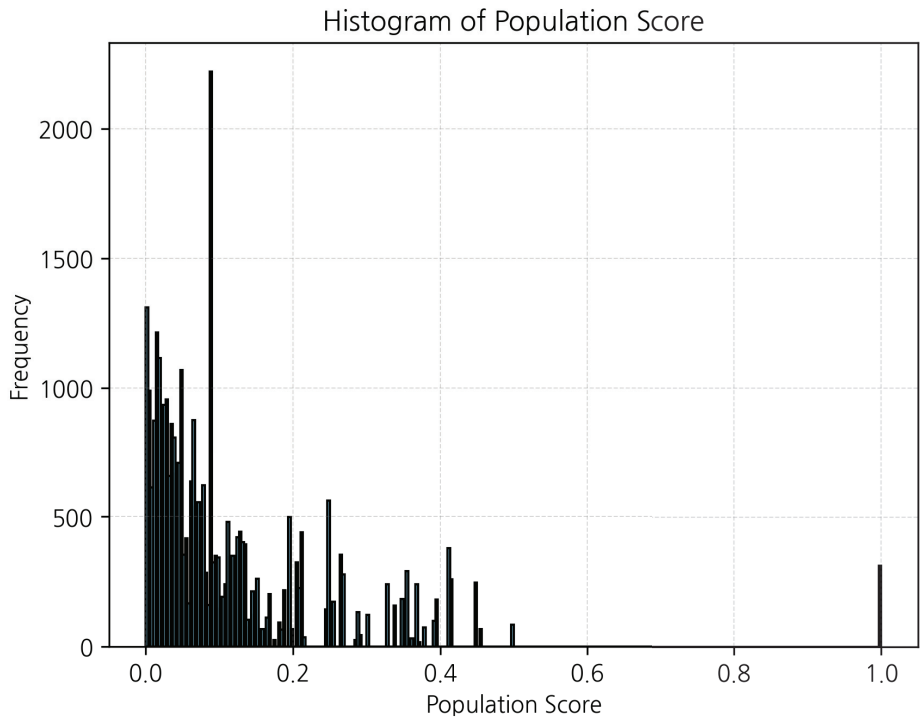


Figure 4. Histogram of the population score (floating population, min–max normalized).

As shown in Figure 6, shop proximity scores exhibit two masses: a spike at zero for sites with no nearby POIs and a broad peak around 0.5–0.6 where multiple POIs lie within walking range. The spike at zero reflects POI-sparse areas; the main peak indicates neighborhoods with several POIs inside the decay radius. Waste source proximity C_3 aggregates the influence of POIs associated with street litter. Let $\mathcal{K} = \{\text{conv}, \text{cafe}, \text{truck}, \text{park}\}$ index convenience stores, cafés, food trucks, and parks, and let $d_{j,k}$ denote the network distance from j to the nearest POI of subtype k. Subtype weights $\{w_k\}_{k\in\mathcal{K}}$ encode relative propensities for waste exposure and satisfy $\sum_k w_k = 1$. Using the same 300 m influence range, we set

$$S_{j,\text{shop}} = \sum_{k\in\mathcal{K}} w_k \max\left\{0,\, 1 - \tfrac{d_{j,k}}{300}\right\}, \quad \text{where } (w_{\text{conv}}, w_{\text{cafe}}, w_{\text{truck}}, w_{\text{park}}) = (0.35, 0.35, 0.20, 0.10). \tag{8}$$

so that co-location near multiple waste-related attractors increases the score while contributions taper linearly to zero at 300 m.

The empirical distributions of the component scores and the composite reveal substantial spatial heterogeneity across Gangnam's pedestrian network and inform the subsequent radius mapping. The floating population score $S_{j,\text{pop}}$ spans $[0, 1]$ and is right-skewed, with a mean of 0.1133, a standard deviation of 0.1380, and a 75th percentile of 0.1339, reflecting concentrated corridors of high exposure. The transit disposal score $S_{j,\text{disp}}$ ranges over $[0, 0.9780]$ and is more uniform, with a mean of 0.4170, a standard deviation of 0.2257, and a median of 0.4450, consistent with the dense but heterogeneous distribution of bus stops and subway exits. The POI proximity score $S_{j,\text{shop}}$ is generally small, with a maximum of 0.0987, a mean of 0.0275, a standard deviation of 0.0288, and a 75th percentile of 0.0515, indicating that only a minority of network points lie within short walks of multiple waste-generating attractors.

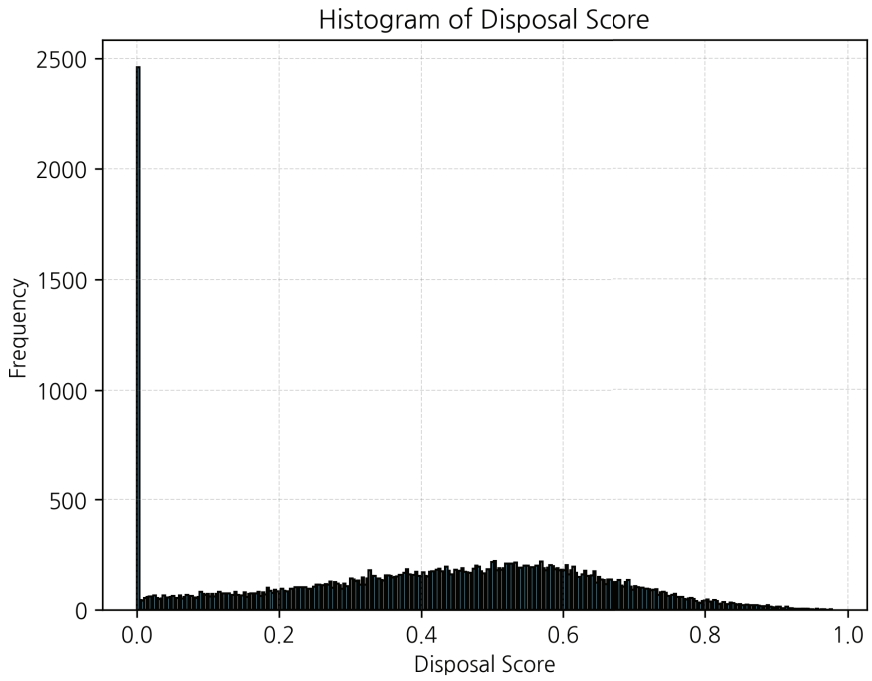


Figure 5. Histogram of the disposal score (transit proximity).

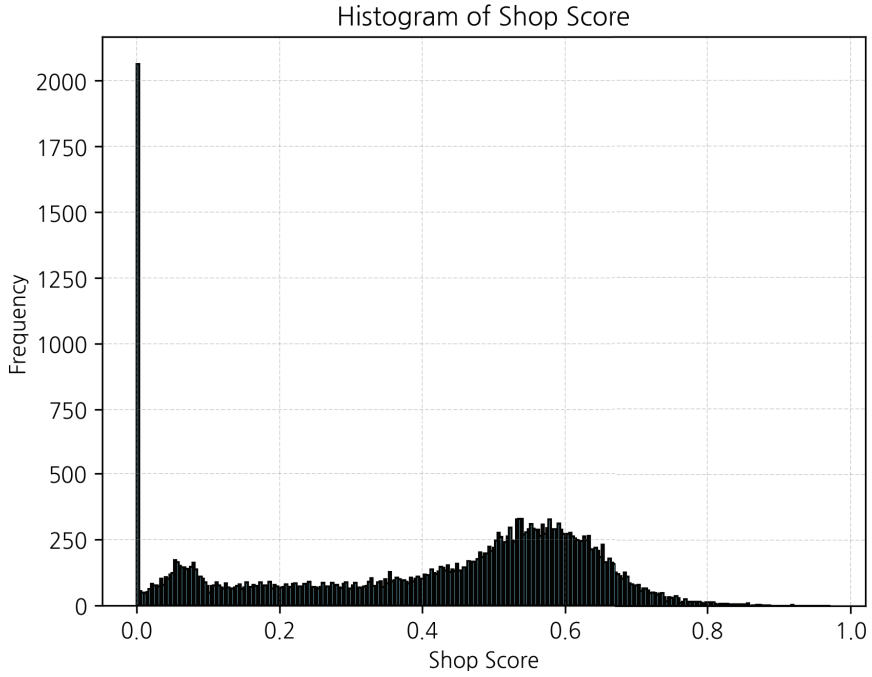


Figure 6. Histogram of the shop score (proximity to convenience stores, cafés, food trucks, and parks).

We specify three criteria $\mathcal{C} = \{C_1, C_2, C_3\}$ that capture complementary drivers of litter pressure and disposal opportunity: C_1 encodes floating population intensity, C_2 captures transit-proximate disposal likelihood, and C_3 reflects proximity to waste-generating points of interest (POIs). For each $j \in J$ and criterion C_k, a raw, non-negative measurement $V_{j,k}$ is constructed and mapped to a unit-interval score $S_{j,k} \in [0, 1]$ via a monotone transformation f_k so that larger values consistently indicate greater suitability for siting a bin at j. With the criterion scores in hand, the relative importance of population exposure, transit opportunity, and POI proximity is computed via the AHP. To clarify the AHP elicitation, we provide

the pairwise comparison scale and the exact questions used for expert judgments. AHP pairwise comparison scale used for expert elicitation.

Scale	Definition
1	Equally important
3	Slightly more important
5	Moderately more important
7	Strongly more important
9	Absolutely more important
2, 4, 6, 8	Intermediate between adjacent judgments
Reciprocals	Inverse when criterion B is preferred over A

Pairwise comparison questions used for expert elicitation were as follows:

Q1. How much more important is distance to waste-source facilities (cafés, convenience stores) than floating population? Answer: 1/5 (population is moderately more important than shop).

Q2. How much more important is disposal likelihood near transit stops (bus, subway) than floating population? Answer: 1/5 (population is moderately more important than disposal).

Q3. How much more important is disposal likelihood near transit stops than waste-source proximity? Answer: 3 (disposal is slightly more important than shop).

Two municipal officers completed the elicitation independently. Their responses were aggregated into the following pairwise comparison matrix:

$$A = \begin{bmatrix} 1 & 5 & 5 \\ \frac{1}{5} & 1 & \frac{1}{3} \\ \frac{1}{5} & 3 & 1 \end{bmatrix}.$$

This matrix indicates that population was judged substantially more important than both disposal likelihood and shop proximity, and that disposal likelihood was moderately more important than shop proximity. Using the normalized average method on A, we obtain the AHP weights,

$$\alpha_{\text{pop}} = 0.6864, \qquad \alpha_{\text{disp}} = 0.2114, \qquad \alpha_{\text{shop}} = 0.1022,$$

reported to four decimal places. The composite suitability at site j is the convex combination given by

$$\begin{aligned} S_j^* &= \alpha_{\text{pop}}\, S_{j,\text{pop}} + \alpha_{\text{disp}}\, S_{j,\text{disp}} + \alpha_{\text{shop}}\, S_{j,\text{shop}} \\ &= 0.6864\, S_{j,\text{pop}} + 0.2114\, S_{j,\text{disp}} + 0.1022\, S_{j,\text{shop}} \end{aligned} \tag{9}$$

which, by construction, lies in $[0, 1]$ and is strictly increasing in each constituent score. This S_j^* serves as the demand score output used downstream to assign adaptive radii and to prioritize candidates during siting. Aggregation by the AHP weights produces a composite S_j^* that remains right-skewed, with a mean of 0.2083, a standard deviation of 0.1199, a median of 0.1883, a maximum of 0.9301, and a 75th percentile of 0.2551. Under the adaptive radius map $r_j = R_{\max} - (R_{\max} - R_{\min})S_j^*$, these statistics translate into smaller service

radii in the highest-scoring corridors and larger radii in peripheral or residential areas, ensuring fine spatial granularity where litter pressure is greatest while preserving baseline access elsewhere. All distances entering the kernels are evaluated with the same network metric d as used in the siting and routing stages, maintaining geometric consistency across the full GD-ARISE pipeline.

Figure 7 shows a heatmap for the composite demand surface that is highly clustered with the highest percentiles concentrated along the northern–central corridors and decreasing toward the periphery. Colors show percentile ranks from low (light) to high (dark); road segments and administrative polygons are overlaid for reference. The surface reveals pronounced high-demand bands in the northern–central corridors, tapering toward the southeast and peripheral areas.

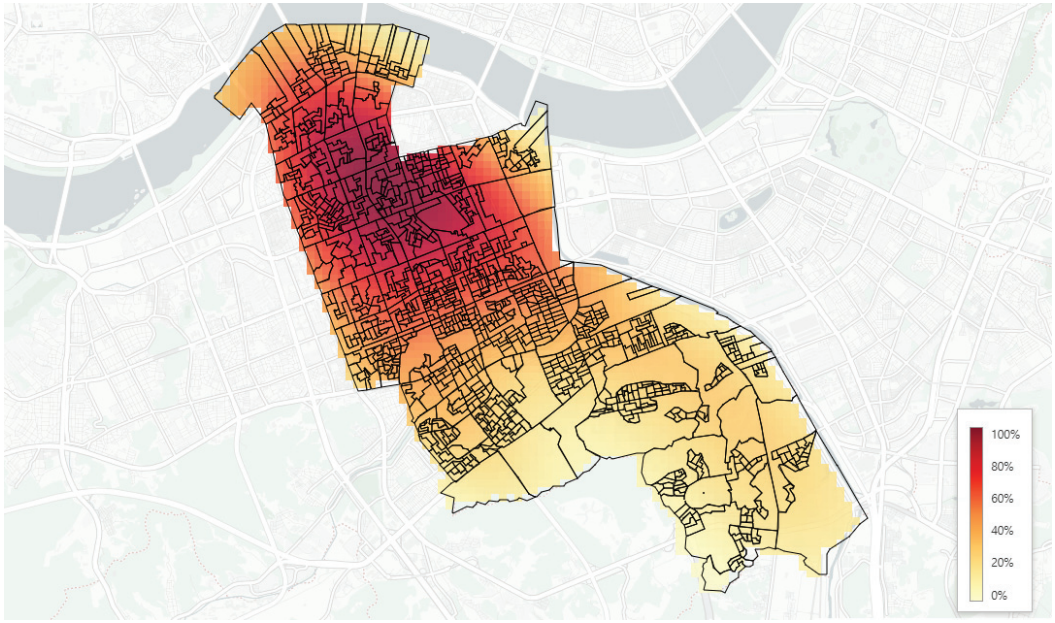


Figure 7. Composite demand heatmap across Gangnam district.

3.3. SCASS

This stage specializes the SCASS formulation to the Gangnam candidate set $J = \{j_1, \ldots, j_N\}$ with $N = 32{,}890$ points sampled every $\delta = 10\,\text{m}$ along the pedestrian–road network $\mathcal{G}$, using the same network-based distance metric d as in the scoring stage. The input at each site $j \in J$ is the composite suitability $S_j^* \in [0, 1]$ obtained from the AHP-based demand assessment. Two preprocessing choices are made to align data coverage with plausible service contexts while preserving the mathematical structure of SCASS. First, a proximity sanity check is applied to screen out isolated candidates: for each j we test whether there exists at least one demand facility (bus stop, subway exit, or waste-related POI) within 1 km under d. In the Gangnam dataset, every candidate satisfies this condition, so the working set remains J. Second, each candidate is assigned an adaptive coverage radius r_j via the affine map

$$r_j \;=\; R_{\max} - (R_{\max} - R_{\min})\, S_j^*, \tag{10}$$

so that higher suitability implies a smaller catchment, consistent with dense placement in demand hotspots, while lower suitability enlarges catchments to preserve baseline access elsewhere. Applied to Gangnam with $(R_{\min}, R_{\max}) = (30\,\text{m}, 150\,\text{m})$ and $P_{\text{target}} = 500$, the greedy procedure selects $|J_{\text{sel}}| = 347$ sites before all remaining candidates conflict with at least one already-selected site; this binding of nonoverlap rather than the numerical target determines the achieved count. The realized J_{sel} concentrates in high-scoring corridors

around major commercial and transit axes, while respecting minimum separations implied by the adaptive radii and thereby avoiding redundant service areas.

As shown in Figure 8, the final set of selected sites for this analysis is displayed. Network edges are shown in black; selections were obtained by the SCASS greedy nonoverlap procedure on the pedestrian–road network. Table 6 summarizes key parameters and resulting outcomes for the SCASS stage. It lists the candidate set size and sampling, the proximity filter, the adaptive radius bounds and mapping, the target portfolio size, the number of selected sites, and the shared network distance metric.

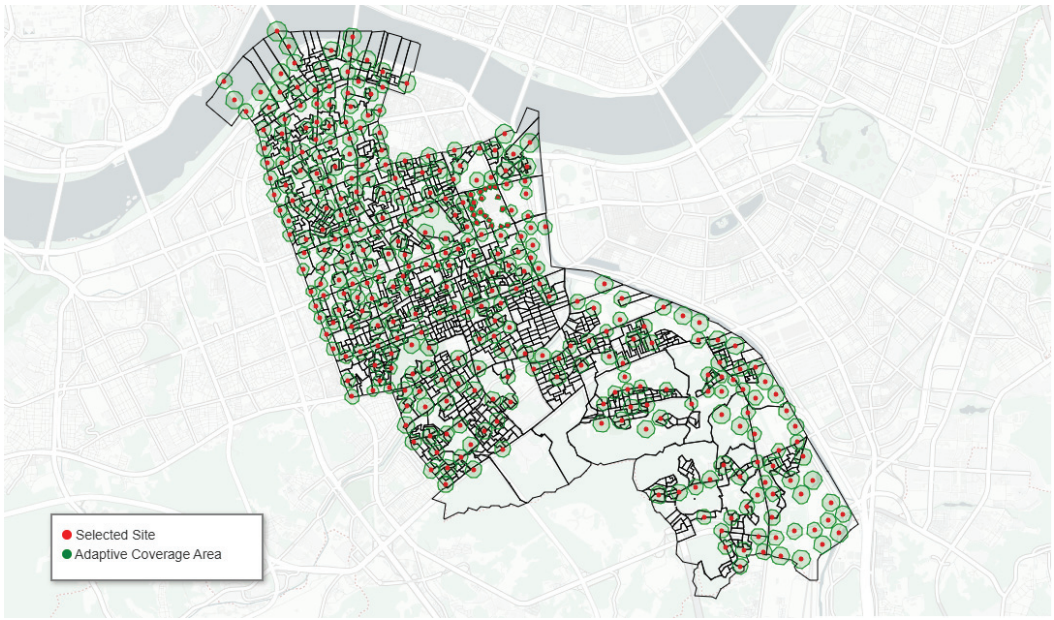


Figure 8. Final selected waste bin sites (red) and their adaptive coverage areas (green).

Table 6. Parameters and outcomes for adaptive site selection.

Quantity	Value	Definition/Role
Candidate set size N	32,890	Points sampled every $\delta = 10\,\mathrm{m}$ along $\mathcal{G}$ within G
Proximity sanity radius	1000 m	Filter to exclude candidates with no nearby facilities under d
Adaptive radius bounds	$R_{\min} = 30\,\mathrm{m}$, $R_{\max} = 150\,\mathrm{m}$	Affine map $r_j = R_{\max} - (R_{\max} - R_{\min})S_j^*$
Target portfolio P_{target}	500	Desired number of nonoverlapping bins
Selected sites $\lvert J_{\text{sel}}\rvert$	347	Maximal nonoverlapping set obtained by greedy scan under d
Distance metric d	Network shortest path	Used for radii, nonoverlap checks, coverage, and routing

3.4. Collection Route Optimization: Samsung 1–dong

To instantiate routing optimization on a concrete sub-area, we focus on Samsung 1–dong within Gangnam and route a single collection vehicle to visit the $K = 37$ waste–bin sites that were previously selected there. The depot is fixed at D_0 = (37.50867° N, 127.086466° E), and we retain the notation of the main framework by labeling the selected bins $\{j_1, \ldots, j_{37}\}$ and defining the node set $V = \{0, 1, \ldots, 37\}$ with $j_0 \equiv D_0$. Consistent with the end-to-end pipeline, inter-node travel costs are evaluated using the same distance metric d induced by the pedestrian–road network $\mathcal{G}$ restricted to the Samsung 1–dong

extent. Because all bins must be serviced exactly once and, for this small cluster, a single vehicle is sufficient, we solve the capacitated vehicle routing model in the special case $m = 1$ and $Q \geq \sum_{i=1}^{37} q_i$, which reduces to a traveling salesman problem on the node set V. A route is represented by a closed walk $(\pi_0, \pi_1, \ldots, \pi_{37}, \pi_{38})$ with $\pi_0 = \pi_{38} = 0$ and $\{\pi_1, \ldots, \pi_{37}\} = \{1, \ldots, 37\}$. Its network cost is

$$C(\pi) = \sum_{t=0}^{37} d_{\pi_t, \pi_{t+1}}, \tag{11}$$

and the optimization objective is to find π that minimizes $C(\pi)$.

As shown in Figure 9, simulated annealing produces a single depot-to-depot tour that visits all selected Samsung 1–dong sites in sequence on the same network used for siting. The green marker denotes the depot; blue points mark the 37 selected sites labeled in visit order; orange lines depict the annealed tour; the black polygon outlines the Samsung 1–dong boundary. The route is the outcome of simulated annealing, evaluated using the network metric consistent with location selection. The resulting annealed tour visits each of the 37 Samsung 1–dong bins exactly once and returns to the depot. Because the encoding, acceptance logic, and termination criteria conform exactly to the routing specification, this Samsung 1–dong case demonstrates that the GD-ARISE pipeline maintains geometric consistency from demand scoring through siting to routing, and that the final routing layer can be solved to high quality with modest computation even when distances respect real network impedances rather than idealized straight lines.

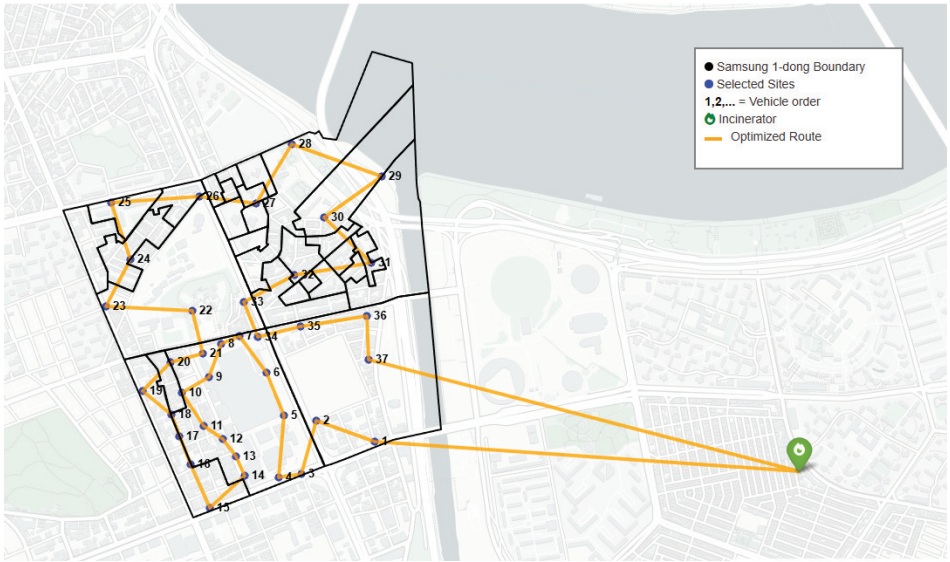
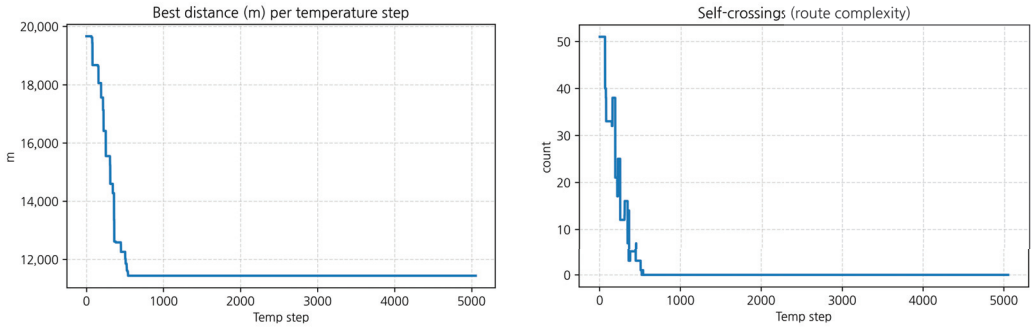


Figure 9. Optimized collection route in Samsung 1–dong.

For the computation of the simulated annealing (SA) algorithm, the parameters of the number of nodes in the route (n), the number of iterations per temperature stage (M), the initial temperature (T_0), the stopping temperature (T_{stop}), and the cooling rate ($\alpha, 0 < \alpha < 1$) are considered, so that the total number of temperature stages (L_a) and the time complexity (L_b) are respectively computed as $L_a = \left\lfloor \frac{\ln(T_{\text{stop}}/T_0)}{\ln(\alpha)} \right\rfloor$ and $L_b = O\big(L_a \cdot (M \cdot n + n^2)\big)$. For the single vehicle case, the parameters were set to $M = 1000$, $n = 39$, $T_0 = 1000$, $T_{\text{stop}} = 10^{-8}$, and $\alpha = 0.995$, yielding $L_a = 5054$ and $L_b = 204{,}793{,}134$. The optimized SA route resulted in a total travel distance of approximately 11.43 km, with a mean leg distance of about 301 m. The self-intersection count was none, and the tortuosity was 2.44 times the straight round-trip length from the depot. Figure 10 shows the total route length decreased sharply before about 1000 steps, from approximately 19,500 m to 11,430 m, and then stabilized.



(**a**) Total route length (**b**) Crossings

Figure 10. Variation of total route length and crossings with temperature (iteration) steps.

For the two-vehicle case, instead of one long tour (TSP), the route can first be divided into two subroutes, each optimized independently by SA. The sites were partitioned via the K-means clustering method to minimize overlap between vehicle coverages and to balance total travel distances, among many other alternative clustering methods. As shown in Figure 11, K-means clustering partitioned the candidate sites between the two vehicles, and inspection of the routes confirms that the two vehicles divided the four census blocks in Samsung 1–dong, with each vehicle traversing two blocks. This process results in two single depot-to-depot tours assigned to the two vehicles, each visiting all trash bins in Samsung 1–dong in sequence, as in the single vehicle case. The green marker denotes the depot; purple points mark the 37 selected sites; the blue lines depict the annealed tour of the 26 selected sites visited by Vehicle A, labeled in visit order; the orange lines depict the annealed tour of the 11 selected sites visited by Vehicle B, labeled in visit order; the black polygon outlines the Samsung 1–dong boundary.

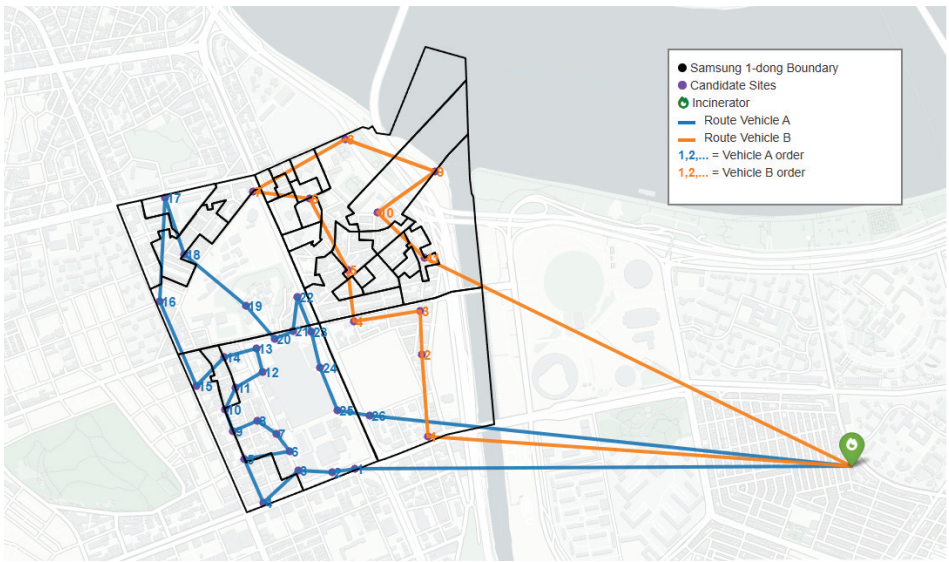
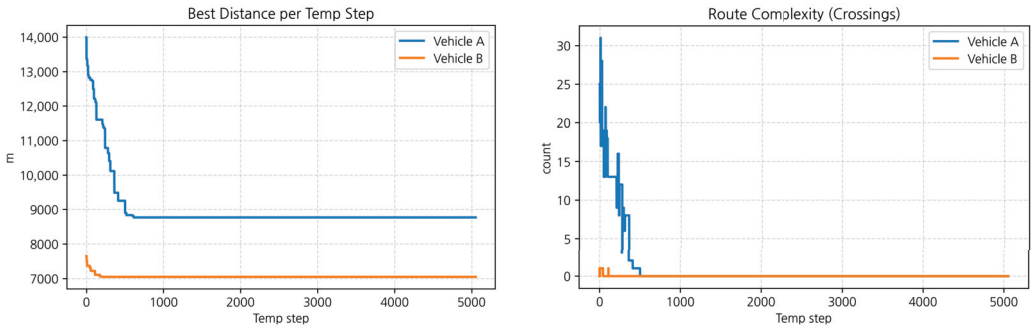


Figure 11. Optimized collection routes for two vehicles in Samsung 1–dong.

After clustering, Vehicle A was assigned $n_A = 28$ nodes and Vehicle B was assigned $n_B = 13$ nodes, excluding the depot. All other SA parameters were kept identical to the single-vehicle case, except for n. The estimated number of operations was 145,474,336 for Vehicle A, and 66,556,126 for Vehicle B. When the two vehicles divided the candidate sites for collection, the total travel distance of both vehicles combined was approximately 15.81 km, which is longer than the single-vehicle distance of about 11.43 km. However, the individual travel burdens were reduced, with Vehicle A and Vehicle B traveling approximately 8.77 km and

7.04 km, respectively. Each vehicle's route had a mean leg distance of approximately 324.8 m and 587.0 m, tortuosity values of 1.83 and 1.84, and a self-intersection count of zero for both.

Figure 12 shows the total route lengths of both vehicles decreased rapidly during the early temperature steps, with Vehicle A dropping from approximately 14,000 m to 8700 m and Vehicle B from about 7600 m to 7040 m within the first 1000 steps, after which both stabilized. The number of crossings followed a similar trend, falling from roughly 30 and 0 at the beginning to nearly zero within 1000 steps for Vehicle A and Vehicle B, respectively.



(**a**) Total route length (**b**) Crossings

Figure 12. Variation of total route length and crossings with temperature steps for the two-vehicle case.

3.5. Sensitivity Analysis

We analyze the impact of parameter settings on GD-ARISE's results. Two sensitivity experiments have been conducted by focusing on (i) the AHP weighting in the multi-criteria scoring stage and (ii) the adaptive coverage parameters in the SCASS site-selection stage. These experiments give ideas about how variations in subjective or geometric settings propagate through the pipeline and influence the resulting suitability distribution and selected facility portfolio.

3.5.1. AHP Sensitivity

To gauge the effect of AHP weighting, we repeated the analysis with near-uniform weights across the three criteria. Under uniform weights, the composite suitability distribution broadened and shifted upward (Figure 13), implying that more candidates attained moderate-to-high suitability. The feasible nonoverlapping selection also expanded and reconfigured, with changes concentrated in mid-range corridors while core hotspots remained stable (Figure 14). Overall, equal weighting increases admissible placements and shifts marginal sites rather than overturning the highest-demand areas.

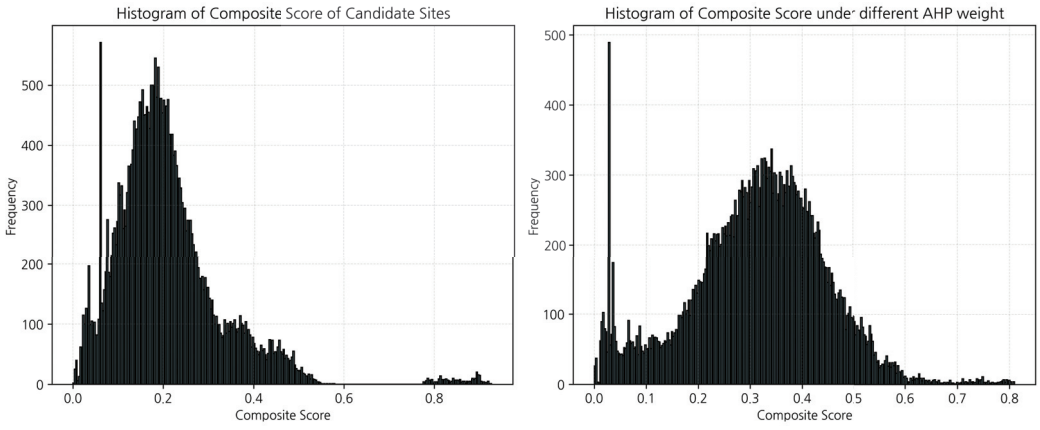
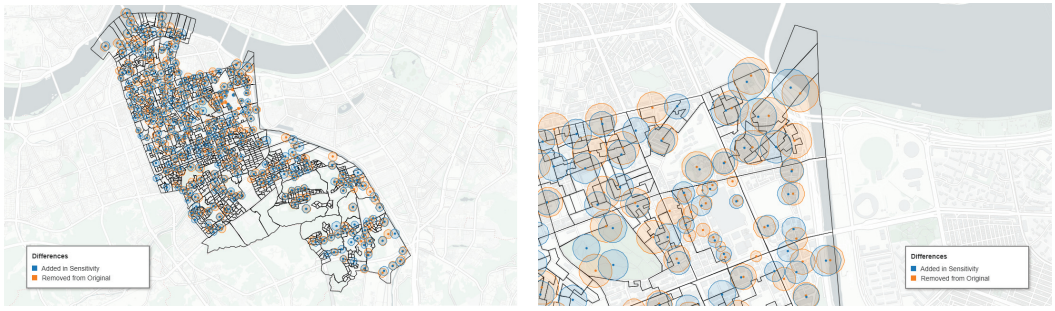


Figure 13. Histograms of composite suitability scores S_j^* of candidate sites under baseline AHP weights (**left**) and changed weights for sensitivity analysis (**right**).

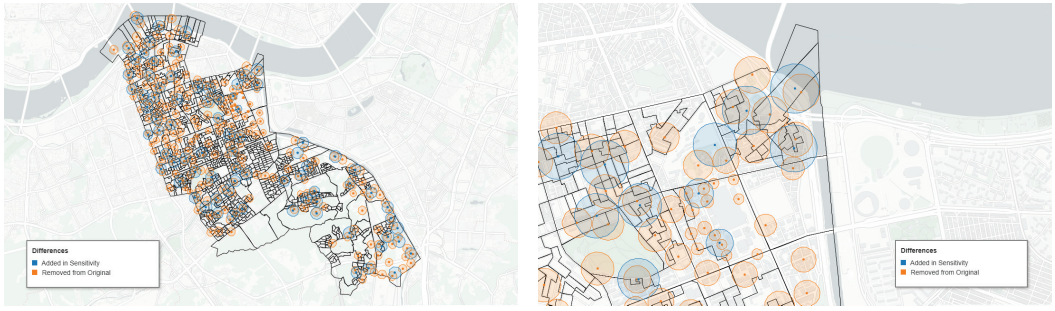


(**a**) Overall view (**b**) Enlarged view

Figure 14. Spatial comparison between the baseline and AHP-sensitivity selections.

3.5.2. SCASS Sensitivity

To assess sensitivity to spatial parameters in SCASS, we increased the adaptive-radius bounds from $(R_{\min}, R_{\max}) = (30\,\text{m}, 150\,\text{m})$ to $(100\,\text{m}, 200\,\text{m})$. Larger radii produced fewer but broader catchments: the feasible nonoverlapping set shrank (from 347 to 180 sites) while coverage areas grew across the entire distribution (range, mean, and quartiles all increased). Spatially, intensified exclusions reduced site density and shifted feasible locations, with removals and additions concentrated where spacing became binding (Figure 15). Overall, expanding $(R_{\min}, R_{\max})$ trades off count for reach, yielding sparser deployments with larger service footprints.



(**a**) Overall view (**b**) Enlarged view

Figure 15. Spatial comparison between the baseline and SCASS-sensitivity selections.

3.6. Managerial Implications

The Gangnam application converts analytics into direct choices about how many bins to deploy, where to place them, and how to service them under real network constraints. The realized portfolio J_{sel} provides a defensible deployment count given nonoverlap and adaptive radii; managers can tune P_{target} and $(R_{\min}, R_{\max})$ to meet coverage or budget targets. Adaptive radii yield a clear siting rule—smaller in high-pressure corridors, larger in quieter areas—so spacing reflects actual walkability rather than Euclidean distance, and nearby nonconflicting alternatives can be identified when frontage or regulations prevent a proposed point.

Two indicators support oversight: the share of demand nodes U covered by $\bigcup_{j \in J_{\text{sel}}} \mathbb{D}_j$, and the distribution of separations relative to $r_i + r_j$, revealing where spacing is tight or slack. The routing layer turns network distance into service hours with average speed and per-stop time, enabling staffing and shift design and making comparisons to incumbent routes straightforward. Sensitivity to AHP weights highlights "swing" sites and shows whether small preference shifts change coverage or route time; robust outcomes ease stakeholder agreement, while sensitive zones can be prioritized for pilots or field checks. Because all stages share one metric and dataset, the pipeline can be re-run seasonally as demand or networks change, with re-routing for minor updates and re-siting reserved for

larger deviations. The same geometry-consistent workflow transfers to curb-level assets, such as micro-mobility docks or hydration stations, by swapping criteria and data sources while retaining the sensitivity and monitoring protocol.

4. Conclusions

This paper framed public waste bin placement and servicing as a GIS-grounded operations research problem and developed a GD-ARISE data analytic pipeline that unifies multi-criteria demand assessment, spatially constrained site selection, and operational routing on a shared network metric. The Gangnam district case demonstrated that, with readily available administrative layers, pedestrian network data, floating population estimates, transit nodes, and points of interest, the pipeline can sample the walkable network at fine resolution, construct composite suitability scores, translate them into adaptive service radii, select a maximal nonoverlapping portfolio of facility sites, and generate depot-to-depot routes that are feasible on the actual network.

Methodologically, four elements define our contribution. First, GIS integration unifies all spatial layers on one pedestrian–road network and one distance metric. Second, demand data analytics converts heterogeneous inputs into transparent multi-criteria scores. Third, site selection is cast as coverage with nonoverlap on the same network. A fast greedy procedure enforces spacing with a degree-based performance guarantee and scales on GIS graphs. Fourth, routing closes the loop on the identical metric used upstream. Together, these steps form a reproducible, end-to-end pipeline for location-and-routing that is defensible, scalable, and operationally relevant for municipal planning. The Gangnam application highlights several substantive findings. Composite demand is highly skewed and spatially clustered along commercial corridors and transit interchanges; adaptive radii concentrate bins where pressure is greatest while preserving access in quieter areas; nonoverlap constraints prevent redundant service footprints and enforce minimum spacing that respects sidewalk conditions; and simulated annealing delivers shorter, operationally coherent tours once distances reflect real network impedances. These outcomes suggest that a GIS-integrated operations research approach can simultaneously improve perceived cleanliness, reduce reactive cleanup, and control service effort by aligning siting and routing with observed urban activity. While our case study centers on waste bins, the framework generalizes to other street-level amenities—micro-mobility docks, kiosks, sensors, hydration stations—where demand is heterogeneous, space is scarce, and operations matter.

The work also has limitations that point to avenues for extension. Criteria weights were elicited via AHP and calibrated to available data; while consistency checks were enforced, robustness to alternative weight sets and to additional criteria (e.g., complaints, street-furniture density, event schedules) merits further study. The adaptive radius mapping is intentionally simple and interpretable; in settings with strong regulatory or equity requirements, radius rules could be learned from outcomes, made time-dependent, or augmented with minimum-access guarantees. Finally, SCASS employs a greedy selection to ensure transparency and scalability, but we did not benchmark its potential suboptimality against exact or metaheuristic methods; future work should quantify optimality gaps (e.g., via MILP on subregions or lightweight local improvements) to bound performance under the stated constraints.

Author Contributions: Conceptualization, H.-T.H.; Methodology, J.-J.W., J.-S.L., and H.-T.H.; Software, J.-J.W. and J.-S.L.; Formal analysis, J.-J.W.; Data curation, J.-J.W. and J.-S.L.; Writing—original draft, J.-S.L. and H.-T.H.; Writing—review and editing, H.-T.H.; Visualization, J.-S.L.; Supervision, H.-T.H.; Funding acquisition, H.-T.H. All authors have read and agreed to the published version of the manuscript.

Funding: This Research was supported by the MSIT (Ministry of Science and ICT), Republic of Korea, under the ITRC (Information Technology Research Center) support program (RS-2025-00259004) supervised by the IITP (Institute for Information and Communications Technology Planning and Evaluation) and by the Korea Institute of Energy Technology Evaluation and Planning(KETEP) grant funded by the Korea government(MOTIE) (20214000000060, Department of Next Generation Energy System Convergence based-on Techno-Economics-STEP).

Data Availability Statement: The data presented in this study are openly available from the Seoul Open Data Portal (public repository) at the following URLs (accessed on 2 July 2025): https://data.seoul.go.kr/dataList/OA-14979/F/1/datasetView.do; https://data.seoul.go.kr/dataList/OA-14980/F/1/datasetView.do; https://data.seoul.go.kr/dataList/OA-14978/F/1/datasetView.do; https://data.seoul.go.kr/dataVisual/seoul/seoulLivingPopulation.do; https://data.seoul.go.kr/dataList/OA-15067/S/1/datasetView.do; https://data.seoul.go.kr/dataList/OA-18699/S/1/datasetView.do; https://data.seoul.go.kr/dataList/OA-15004/F/1/datasetView.do; In addition, one auxiliary dataset (subway entrance coordinates used for accessibility analysis) was derived from a public-domain web map and is available at: http://map.esran.com/ (accessed on 2 July 2025) [Seoul Open Data Portal].

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript (Figure 1):

CRS	Coordinate Reference System
MCDM	Multi-Criteria Decision-Making
AHP	Analytic Hierarchy Process
CSA	Coverage Suitability Analysis
MCLP	Maximal Covering Location Problem
SCASS	Spatially-Constrained Adaptive Site Selection
RO	Route Optimization
GD-ARISE	GIS-integrated and Data analytic Adaptive Radius Integrated Siting and rEservicing
OR	Operations Research
POI	Point of Interest
PRN	Pedestrian–Road Network

References

1. Hess, C.; Dragomir, A.G.; Doerner, K.F.; Vigo, D. Waste collection routing: A survey on problems and methods. *Cent. Eur. J. Oper. Res.* **2024**, *32*, 399–434. [CrossRef]
2. Han, J.; Zhang, J.; Guo, H.; Zhang, N. Optimizing location-routing and demand allocation in the household waste collection system using a branch-and-price algorithm. *Eur. J. Oper. Res.* **2024**, *316*, 958–975. [CrossRef]
3. Boeing, G. OSMnx: New methods for acquiring, modeling, analyzing, and visualizing complex street networks. *Comput. Environ. Urban Syst.* **2017**, *65*, 126–139. [CrossRef]
4. Mahéo, A.; Rossit, D.G.; Kilby, P. Solving the integrated bin allocation and collection routing problem for municipal solid waste: A Benders decomposition approach. *Ann. Oper. Res.* **2023**, *322*, 441–465. [CrossRef]
5. Haklay, M.; Weber, P. OpenStreetMap: User-generated street maps. *IEEE Pervasive Comput.* **2008**, *7*, 12–18. [CrossRef]
6. Nyimbili, P.H.; Erden, T. A hybrid approach integrating entropy–AHP and GIS for suitability assessment of urban emergency facilities. *ISPRS Int. J. -Geo-Inf.* **2020**, *9*, 419. [CrossRef]

7. Saaty, T.L. *The Analytic Hierarchy Process*; McGraw–Hill: New York, NY, USA, 1980.
8. Şener, B.; Süzen, M.L.; Doyuran, V. Landfill site selection by using geographic information systems. *Environ. Geol.* **2006**, *49*, 376–388. [CrossRef]
9. Rahmat, Z.G.; Niri, M.V.; Alavi, N.; Goudarzi, G.; Babaei, A.A.; Baboli, Z.; Hosseinzadeh, M. Landfill site selection using GIS and AHP: A case study: Behbahan, Iran. *KSCE J. Civ. Eng.* **2017**, *21*, 111–118. [CrossRef]
10. Ahmed, A.; Kheraj, T.; Mohammadi, A.; Bergquist, R. Hybrid GIS-MCDM approach for Hospital Site Selection Suitability Analysis in Poonch District, Jammu and Kashmir, India. *GeoJournal* **2024**, *89*, 186. [CrossRef]
11. Mostafaeipour, A.; Hosseini Dehshiri, S.S.; Hosseini Dehshiri, S.J.; Almutairi, K.; Taher, R.; Issakhov, A.; Techato, K. A thorough analysis of renewable hydrogen projects development in Uzbekistan using MCDM methods. *Int. J. Hydrogen Energy* **2021**, *46*, 31174–31190. [CrossRef]
12. Janmontree, J.; Zadek, H.; Ransikarbum, K. Analyzing solar location for green hydrogen using multi-criteria decision analysis. *Renew. Sustain. Energy Rev.* **2025**, *209*, 115102. [CrossRef]
13. Church, R.L.; ReVelle, C.S. The maximal covering location problem. *Pap. Reg. Sci.* **1974**, *32*, 101–118. [CrossRef]
14. Daskin, M.S. *Network and Discrete Location: Models, Algorithms, and Applications*; Wiley: Hoboken, NJ, USA, 1995.
15. ReVelle, C.; Eiselt, H.A. Location analysis: A synthesis and survey. *Eur. J. Oper. Res.* **2005**, *165*, 1–19. [CrossRef]
16. Marsh, M.T.; Schilling, D.A. Equity measurement in facility location analysis: A review and framework. *Eur. J. Oper. Res.* **1994**, *74*, 1–17. [CrossRef]
17. Daskin, M.S. A maximum expected covering location model: Formulation, properties and heuristic solution. *Transp. Sci.* **1983**, *17*, 48–70. [CrossRef]
18. Bonnet, B.; Dessavre, D.G.; Kraus, K.; Ramirez-Marquez, J.E. Optimal placement of public-access AEDs in urban environments. *Comput. Ind. Eng.* **2015**, *90*, 269–280. [CrossRef]
19. Farahani, R.Z.; Asgari, N.; Heidari, N.; Hosseininia, M.; Goh, M. Covering problems in facility location: A review. *Comput. Ind. Eng.* **2012**, *62*, 368–407. [CrossRef]
20. Adeleke, O.J.; Olukanni, D.O. Facility location problems: Models, techniques, and applications in waste management. *Recycling* **2020**, *5*, 10. [CrossRef]
21. Berman, O.; Krass, D.; Drezner, Z. The gradual covering decay location problem on a network. *Eur. J. Oper. Res.* **2003**, *151*, 474–480. [CrossRef]
22. Karasakal, O.; Karasakal, E.K. A maximal covering location model in the presence of partial coverage. *Comput. Oper. Res.* **2004**, *31*, 1515–1526. [CrossRef]
23. Blanco, V.; Gázquez, R. Fairness in maximal covering location problems. *Comput. Oper. Res.* **2023**, *157*, 106287. [CrossRef]
24. Berman, O.; Drezner, Z.; Krass, D.; Wesolowsky, G.O. The variable radius covering problem. *Eur. J. Oper. Res.* **2009**, *196*, 516–525. [CrossRef]
25. Özkan, B.; Özceylan, E.; Sarıçiçek, İ. GIS-based MCDM modeling for landfill site suitability analysis: A comprehensive review of the literature. *Environ. Sci. Pollut. Res.* **2019**, *26*, 30711–30730. [CrossRef] [PubMed]
26. Araújo, E.J.; Chaves, A.A.; Lorena, L.A.N. A mathematical model for the coverage location problem with overlap control. *Comput. Ind. Eng.* **2020**, *146*, 106548. [CrossRef]
27. Cherri, L.H.; Carravilla, M.A.; Ribeiro, C.; Toledo, F.M.B. Optimality in nesting problems: New constraint programming models and a new global constraint for non-overlap. *Oper. Res. Perspect.* **2019**, *6*, 100125. [CrossRef]
28. Gola, A.; Kłosowski, G.; Świć, A. Facility Layout Problem with Alternative Facility Variants. *Appl. Sci.* **2023**, *13*, 5032. [CrossRef]
29. Hifi, M.; M'Hallah, R. A literature review on circle and sphere packing problems: Models and methodologies. *Adv. Oper. Res.* **2009**, *2009*, 150624. [CrossRef]
30. Erlebach, T.; Jansen, K.; Seidel, E. Polynomial-time approximation schemes for geometric intersection graphs. *SIAM J. Comput.* **2005**, *34*, 1302–1323. [CrossRef]
31. Osman, I.H.; Laporte, G. Metaheuristics: A bibliography. *Ann. Oper. Res.* **1996**, *63*, 513–623. [CrossRef]
32. ReVelle, C.; Schlossberg, M.; Williams, J.C. Solving the maximal covering location problem with heuristic concentration. *Comput. Oper. Res.* **2008**, *35*, 427–435. [CrossRef]
33. Kirkpatrick, S.; Gelatt, C.D.; Vecchi, M.P. Optimization by simulated annealing. *Science* **1983**, *220*, 671–680. [CrossRef]
34. Aarts, E.; Korst, J. *Simulated Annealing and Boltzmann Machines: A Stochastic Approach to Combinatorial Optimization and Neural Computing*; Wiley: Hoboken, NJ, USA, 1989.
35. Lin, S.; Kernighan, B.W. An effective heuristic algorithm for the traveling-salesman problem. *Oper. Res.* **1973**, *21*, 498–516. [CrossRef]
36. Yu, V.F.; Lin, C.-H.; Maglasang, R.S.; Lin, S.-W.; Chen, K.-F. An efficient simulated annealing algorithm for the vehicle routing problem in omnichannel distribution. *Mathematics* **2024**, *12*, 3664. [CrossRef]

37. Rahmanifar, G.; Mohammadi, M.; Golabian, M.; Sherafat, A.; Hajiaghaei-Keshteli, M.; Fusco, G.; Colombaroni, C. Integrated location and routing for cold chain logistics networks with heterogeneous customer demand. *J. Ind. Inf. Integr.* **2024**, *38*, 100573. [CrossRef]
38. Hashemi-Amiri, O.; Ji, R.; Tian, K. An integrated location–scheduling–routing framework for a smart municipal solid waste system. *Sustainability* **2023**, *15*, 7774. [CrossRef]

mathematics

MDPI

Article

Addressing Daigou from the Perspective of Channel Competition: Strategy for Retail Management

Keqin Chang, Rachael Kwai Fun Ip and Pak Hou Che *

School of Business, Macau University of Science and Technology, Macao SAR, China;
1220008928@student.must.edu.mo (K.C.); kfip@must.edu.mo (R.K.F.I.)
* Correspondence: pahche@must.edu.mo

Abstract

In China's on-demand service platforms, daigou agents utilize locational differences through proxy purchasing. Daigou creates an informal supply chain that directly competes with official channels. This study incorporates daigou arbitrage into the channel competition framework via a multi-stage Stackelberg game-theoretic model. An analysis of the subgame perfect Nash equilibrium shows that daigou activity disrupts the manufacturer's profits. We have thus developed a strategy based on mathematical optimization and compared its effectiveness and side effects with those of existing methods. We came to identify purchase restrictions as one of the most powerful strategies. Equilibrium analysis and numerical experiments confirm that proper purchase restriction choices reduce daigou arbitrage and minimize negative impacts on legitimate demand. This work provides the first game-theoretic model that integrates informal proxy-purchase supply chains into dual-channel competitions.

Keywords: daigou; retail management; channel competition

MSC: 90B06

1. Introduction

Daigou is derived from the Chinese term for "purchasing on behalf". Daigou activities involve individuals or networks acting as intermediaries to acquire and resell goods. It often involves price markup through informal channels such as social media platforms or e-commerce apps. The daigou phenomenon originated in cross-border contexts. As Chinese consumers sought overseas luxury goods and pharmaceuticals due to tariffs, quality concerns, and price advantages, daigou has become a billion-dollar informal economy [1]. In recent years, resellers in China have benefited from localized supply constraints and viral demand for premium regional products. This shift is further accelerated by digital mobile applications. They enable real-time inventory sharing and livestreaming sales, increasing competition with formal retail channels by creating parallel distribution networks that affect pricing control and supply chain predictability [2,3].

The rise of daigou (also known as "proxy shopping") in China shows a transformation of retail and consumer behavior, driven by fragmented supply chains, domestic demand asymmetries, and the growth of digital platforms. Initially, daigou was primarily associated with cross-border e-commerce driven by factors such as exchange rate differences and tax advantages. However, it has since become a domestic phenomenon, particularly in urban retail, where consumers increasingly rely on intermediaries to access high-demand goods

from niche or regionally concentrated suppliers. This practice has created digital platform-based labor networks, redefining the boundaries between traditional retail, non-traditional work arrangements, and consumer-driven market dynamics. For example, Pang Donglai (www.azpdi.cn), a regional supermarket chain in China renowned for its premium service with a limited physical expansion, served as a notable example of how cross-border daigou activity has shifted to the domestic market. Although Pang Donglai currently operates in only two cities within Henan Province, China, its reputation for high product quality and customer experiences has gained significant demands from other places outside Henan Province. However, Pang Donglai provides no online shopping and delivery services to its customers. Thus, when customers want to buy Pang Donglai's high-quality products, they have to shop at their physical stores. Quite often, Pang Donglai customers have to queue outside these stores before they can shop. Customers from outside Henan Province must endure long travel times just to reach Pang Donglai stores, while even those from nearby regions often face hours of queuing before they can shop. To provide convenient and time-saving services to these Pang Donglai customers, domestic daigou agents emerged, who would collect orders from customers living in local regions or outside Henan Province, queue up outside Pang Donglai's stores, and purchase the items for their customers. The orders are completed by shipping these products to the customers via delivery networks. Many leading professional daigou operators would set up physical bases near Pang Donglai stores to complete their daigou activities in a more efficient way. However, the large volume of these daigou customers has overwhelmed Pang Donglai's operational capacity and compromised its service standards. To tackle these problems, Pang Donglai has implemented targeted interventions, including purchase restrictions on high-demand items to reduce the chances of product shortage. The conflicts between institutionalized retail models and platform-driven consumption practices reveal the need for a systematic analysis of daigou activities to understand the implications for supply chain resilience, labor organization, and evolving consumer behavior.

From a channel competition perspective, daigou represents a disruptive force in the retail ecosystem. Informal intermediaries compete with formal channels for consumer access and brand loyalty. In China, this is reflected in resellers benefiting from arbitrage opportunities such as regional exclusivity or promotional pricing. Daigou resells goods at a premium for convenience and accessibility, which results in the substitution effects that disrupt existing sales offers and reduces brand control. For instance, cross-border daigou for luxury brands often involves lower prices than domestic prices through tax avoidance, accounting for over 20% of China sales in some categories [2]. However, domestic iterations extend this competition to retail. This is reflected in large-scale retailers like Pang Donglai facing bulk-purchase resellers who charge markups to distant consumers. This increases the pressure on store operations and thus requires managerial intervention to safeguard service quality and supply chain resilience.

The case of Pang Donglai shows how domestic daigou can disrupt institutionalized retail models by creating unanticipated demand pressures and operational challenges. The existing literature has focused on cross-border daigou such as luxury goods and pharmaceuticals. Scalper-driven daigou operations often focus on limited-production products. However, the domestic daigou remains underexplored in standard retail settings. For example, studies often consider daigou as a cross-border phenomenon involving overseas procurement for mainland Chinese consumers, such as in Hong Kong or Western markets. The daigou serves as a "middleman" connecting customers to global brands through informal networks [3]. Prior research has extensively examined cross-border daigou in luxury goods and scalping in limited-edition products, but few studies have addressed its domestic iterations in standard retail settings. Therefore, the literature

limits the understanding of how daigou influences channel integration, supply chain agility, and competitive strategies in fragmented markets [4]. Thus, it raises questions about how domestic daigou operates within retail ecosystems, its implications for supply chain resilience and consumer behavior, and the policy responses needed to address its challenges. Recent research also point out that daigou is still largely "under-researched" as an unconventional entrepreneurial setup, especially in non-luxury retail markets [5]. Hence, these questions show a gap in the literature, which this study tries to address.

To address these gaps, this study examines domestic daigou from the perspective of channel competition through a comparison between a base model without daigou, a model incorporating daigou, and a model with daigou under constraints. This comparative analysis explores the following research questions:

- RQ1: How does the presence of daigou affect retail channel conflicts compared to scenarios without daigou?
- RQ2: What are the impacts of constraints like purchase limits on daigou dynamics and retailer strategies for management and resilience?
- RQ3: How do asymmetric substitution effects between daigou and official channels influence equilibrium outcomes and competition in the models?

From the analysis of theoretical models of multi-channel competition, this research contributes a framework for retail managers to navigate daigou dynamics, offering actionable strategies to reclaim channel control.

2. Literature Review

Daigou is defined as a practice that combines operations management, digital commerce, and economic theory by purchasing goods and reselling them at higher prices. The daigou agents act as speculators, taking advantage of price differences rather than making direct sales. On the other hand, daigou also shares some similarities with the on-demand service platforms. In this case, independent agents serve as intermediaries who respond to individual consumer requests, sourcing goods or services and exploiting price or availability differences. However, few analyses have applied quantitative frameworks to capture daigou's distinct setup. To fill this gap, our work proposes a Stackelberg leader–follower strategic model. In this model, the manufacturer sets rules to increase market efficiency and consumer benefit, while the daigou adjusts both price and volume decisions under these constraints.

2.1. Speculative Selling and Scalping

In terms of speculative selling, systematic research on this dates back to the 1970s, when academics began studying how inventories are transferred across periods from an arbitrage perspective. Rather than depending on general intuition, Kohn [6] employed a rigorous structure to illustrate how the cost of carrying and changing market prices drive stock adjustments in commodities, thereby focusing attention on real rather than rational behavior. Drawing on these studies, Tirole [7] questioned the results of the former study and instead explained persistent profits through a heterogeneity of beliefs across traders. While these models provide utility, they often struggle to accommodate barriers encountered in a real-life transaction environment. Ticket scalping is a good example of speculative behavior in the market. Scalpers purchase tickets at lower prices, then sell at higher prices as the ticket date approaches, taking advantage of on-demand fluctuations and temporal price variations [8–10].

More recent research expands the definition of speculative selling to include both speculators and forward-looking customers [11–14]. The study [11] proposes an equilibrium where the manufacturer limits initial availability. The manufacturer can increase

profit through willing-to-wait consumers while reducing the volume of speculative purchases at the same time. The study [14] shows that dynamic pricing mechanisms can also increase profit by adjusting prices in response to inventory risk and market demand fluctuations. On the other hand, ref. [15] further develops this direction and demonstrates how manufacturers can observe the uncertainties of market speculation while effectively managing consumer purchasing timing. These approaches, in the end, allow manufacturers to increase their profits even when facing limited resources and unpredictable consumer behavior. However, these models assume symmetric profit maximization and regulated markets. For the gray markets, they exploit the price differences, regulatory gaps, and cost asymmetries that disrupt the conventional supply chain. Therefore, this shows the importance of adjusting these models to include daigou contexts.

2.2. Cross-Border Daigou and Gray Markets

Cross-border daigou takes advantage of price differences between countries or regions. Daigou agents purchase from markets with lower prices, then sell in markets with higher prices. This approach adds a spatial factor to gray market corss-border arbitrage. There is extensive descriptive research on investigating cross-border daigou practices. Some studies show that cross-border daigou sellers choose product categories where pricing differences are large; this includes cosmetics, infant formula, and also luxury items [1,16,17]. The work [3] investigates the bodily motions that facilitate the flow of people between Mainland China and Hong Kong. It also shows how physical travel, combined with digital platforms, creates labor hierarchies and power asymmetries. Bai [18] suggests that luxury brands often struggle with controlling their widespread businesses, which can often bypass official supply chains. Recent work [5] frame daigou participants as entrepreneurial intermediaries, which allow them to fill market gaps that formal supply chains leave open.

Recent mathematical models consider daigou-driven arbitrage as a strategic game. Game-theoretic models suggest that manufacturers can oppose parallel imports, like daigou actors, by changing how they distribute products, lowering product quality, or timing market entrance based on regional differences in how many consumers are willing to purchase [19]. Another analytical model in [20] suggests that allowing restricted gray market activity could potentially increase total profits. This matches when luxury markets are large and the costs of international transactions are relatively low. A recent empirical study of informal secondary markets support this idea. Prasetyo [21] investigated how digital platforms operate in informal economies. The study showed how social networks help to promote gray market activities in reselling.

2.3. Domestic Daigou and Sharing Economy

Cross-border daigou focuses on price differences and global supply chain challenges. However, a unique practice has developed in domestic and regional markets. This practice is closely related to the sharing economy principles. Platforms like errand-running services (paotui, meaning “running legs” in Chinese) and on-demand delivery networks support that identity develops through access, not ownership [22]. They emphasize spatial proximity and real-time responsiveness instead of the price differences between countries or regions. So, the domestic models in fact differ from cross-border daigou, that is, the domestic models emphasize time-sensitive transactions. This includes acquiring limited-edition items, addressing domestic supply chain issues, and meeting hyperlocal consumer demands. This shows how the sharing economy aims to increase access to underused resources. As a result, digital platforms play a key role by improving trust and efficiency in these domestic operations [23,24].

The recent growth of these platforms is supported by mobile technologies, real-time data analytics, and matching systems. These tools provide effective coordination between consumers and micro-entrepreneurs or freelancers [25–27]. Research uses game-theoretic and queuing-theoretic frameworks to analyze the interplay between supply and demand. For example, Stackelberg hierarchies represent platforms as leaders that create surge pricing [28] and commissions [29]. On the other hand, Hotelling models analyze competition in multi-homing applications [30]. Queuing-theoretic approaches [27,31,32] analyze decisions on pricing, wages, and staffing under uncertainty. Reputation systems [33] help minimize matching issues to improve the system performance. A recent study [34] on informal markets in the sharing economy explored dual-channel supply chain competition during pandemic-induced demand disruptions. The authors discuss how informal reselling changes to platform-mediated constraints in e-commerce.

2.4. Dual-Channel Competition and Omni-Channel Retailing

Research on dual-channel competition and omni-channel retailing provides a deeper understanding of how daigou disrupts traditional supply chains. Recent studies on omni-channel retail demonstrated the importance of platform-mediated channels and informal markets in modern e-commerce. The work [35] examined omni-channel retailing technologies through a systematic review. It showed that integrating online and offline channels can reduce informal reselling by directing demand to formal platforms. The work [36] examined the challenges of retail in omni-channel. It also developed a scale to identify omni-channel obstacles, showing that digital tools can reduce the reliance on informal intermediaries such as daigou. Also, the work [37] proposed a systematic review and a future research for omni-channel management in the modern retailing setup. It emphasized the need to integrate informal channels into formal strategies to improve the resilience of the supply chain.

In this research, we focus on the manufacturer-centric control in the presence of domestic daigou. We develop a Stackelberg leader–follower game structure. The manufacturer acts as the leader, and the domestic daigou agent acts as the follower. This approach is different from the prior platform-centric models through pricing and commission strategies [28]. Instead, we focus on the manufacturer strategy that can influence the domestic daigou behavior. Then, we optimize the supply chain efficiency, pricing, and profits. Our analysis explores a novel manufacturer strategy with purchasing limitation. This strategy is closely related to the regulatory effects imposed by China's E-Commerce Law, which adds purchase restrictions via mandatory registration and taxation [38]. However, our approach achieves a similar outcome by limiting arbitrage, likely by redirecting the demand back to the official channels through supply chain mechanisms rather than policy interventions.

3. The Model

In this study, we analyze a supply chain with a single manufacturer and a single domestic daigou. We also focus on a single product to isolate the strategic interactions within this market structure. As empirical motivations often emphasize the networks of daigou agents and retailers, our choice of a single daigou agent simplifies the model to focus solely on the strategy. This is common in the supply chain literature for analytical tractability [39]. Hence, it allows us to derive closed-form expressions in pricing and demand, though future extensions could incorporate multiple agents to capture competition among daigou. On the other hand, we consider that the manufacturer also operates as a retailer, offering the product directly to consumers through traditional sales channels. Then, the domestic daigou enters the market by sourcing the same product from the manufacturer

and reselling it. The domestic daigou prioritizes the convenience of remote purchasing and "proxy shopping".

We develop three models to capture different market dynamics, and the three models are shown below in Figure 1:

1. **Direct Channel Model (B)**: The manufacturer acts as the sole retailer, selling directly to consumers. This model is a benchmark for evaluating the impact of introducing a domestic daigou. It focuses on optimal pricing and demand in a simple direct channel setting.
2. **Daigou Entry Model (D)**: When a domestic daigou is introduced, the product is sourced from the manufacturer at the retail price and is resold to convenience-seeking consumers. This model focuses on the impact of domestic daigou entry on pricing and demand, as well as the profits.
3. **Constrained Daigou Model (C)**: Building on the Daigou Entry Model, we introduce a constraint on the purchase quantity for the daigou. This extends the Daigou Entry Model and analyzes how purchase constraint affects strategic decisions and channel efficiency.

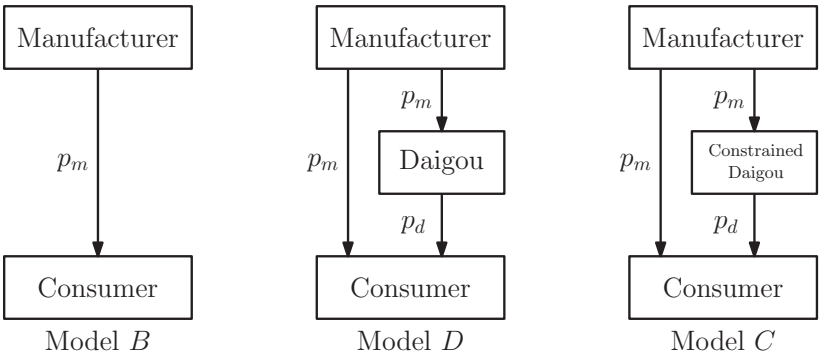


Figure 1. Direct Channel Model, Daigou Entry Model and Constrained Daigou Model.

The primary objective of this research is to analyze the pricing, demand, and overall profitability within a single-manufacturer single-domestic daigou supply chain. We developed and analyzed three sequential models for the Stackelberg game for analysis. The baseline is the Direct Channel Model. We then introduce the domestic daigou in the Daigou Entry Model, and find the impact of the presence of this intermediary. Based on the Daigou Entry Model, the Constrained Daigou Model introduces a purchase limit on the domestic daigou, allowing us to analyze how such purchase constraints affect the decisions and channel efficiency.

In this study, we use the following notation to describe the key variables and parameters. The manufacturer's sales price is denoted by p_m, while the daigou's sales price is denoted by p_d. In Models D and C, the parameters θ_d and θ_m are the asymmetric channel substitution parameters, where θ_d and θ_m represent the degree of substitution from the domestic daigou to the manufacturer and from the manufacturer to the domestic daigou, respectively. The superscripts $\{B, D, C\}$ are used to distinguish the models. In Models D and C, they incorporate the asymmetric channel substitution effects. Therefore, they capture the competitive dynamics between the formal manufacturer's direct channel and informal daigou channels. From the channel competition and differentiation [40–42], these parameters (θ_m^D and θ_d^D) represent the degrees of substitutability between the two channels, where $\theta_m^D, \theta_d^D \in (0, 1)$. Lower values of these parameters indicate weaker substitution; higher values represent near-perfect substitutability.

We assume $\theta_d^D > \theta_m^D$; this means the domestic daigou has a greater substitutive impact on the manufacturer's direct channel demand. Theoretically, this asymmetry arises from

the daigou's provision of differentiated value-added services that the manufacturer's direct channel often cannot copy. These services attract consumers who might initially consider the direct channel but switch to the daigou for convenience [43,44]. On the other hand, the manufacturer's direct channel typically does not match the daigou's service level. Therefore, it is less likely to draw away the daigou's consumers.

We assume the manufacturer's marginal production cost is constant at c, with $c \in (0, a)$, where a represents the market size. For the domestic daigou in Models D and C, the procurement cost equals the manufacturer's direct-channel retail price, $c_d^D = p_m^D$, as daigou operate as unauthorized, informal resellers who source products at standard retail prices from official websites or stores, without access to wholesale agreements, minimum-order quantity waivers, volume discounts, or tiered pricing [16,44,45]. This retail price procurement is a feature of retail-level gray markets and daigou channels, where intermediaries effectively act as high-volume consumers rather than contracted distributors. Manufacturers can employ discriminatory or wholesale pricing to deter arbitrage by authorized distributors, such strategies are ineffective against informal daigou, who source covertly or through consumer-facing retail channels. This assumption therefore reflects actual daigou practices in China-dominated markets and is standard in gray-market models of unauthorized parallel importation. Additional operational costs are omitted for analytical tractability, following the common practice detailed in the dual-channel literature [46,47].

We denote e as the level of value-adding effort deployed by the domestic daigou, which may include personalized marketing, faster local delivery, etc. Accordingly, $\gamma > 0$ represents the marginal effectiveness of this effort in expanding primary demand. The expanded market potential available to the daigou channel is thus modeled as $a + \gamma e$, where a represents the original market size reachable via the manufacturer's direct channel alone, and the incremental term $\gamma e > 0$ captures the daigou's role in stimulating new primary demand among previously time-sensitive or convenience-seeking consumers, following established models of channel expansion and effort-driven demand growth [48–51].

Building on these channel dynamics, in modeling the daigou's effort e in Model D and Model C, we assume it expands demand linearly via parameter γ, as this captures constant marginal returns to effort in stimulating consumer interest and simplifies the demand function for tractability [52] to represent realistic advertising for market share. The costs are quadratic, parameterized by η, to reflect increasing marginal costs, diminishing returns, and convexity, ensuring the profit function is concave for unique equilibria. This is a structure commonly used in marketing and supply chain models [48,53]. This relationship is empirically substantiated by studies on intermediary efforts in e-commerce and proxy markets, where increased personalization, marketing, and service efforts correlate with demand growth but require ongoing investments that can escalate with scale. For instance, qualitative interviews with 27 daigou practitioners highlighted how efforts like physical sourcing trips build trust and facilitate demand in cross-border commodity chains [3].

The game sequences in all three models follow a consistent decision sequence: the manufacturer first determines the price, and the domestic daigou then sets the selling prices for the product based on the observed strategy from the manufacturer.

- In Model B, the manufacturer sets the direct sales price p_m^B, and there is no domestic daigou in this model.
- In Model D, the manufacturer first sets the direct sales price p_m^D, after which the domestic daigou purchases at this price and sets the retail price p_d^D, where $p_d^D > p_m^D$.
- In Model C, the manufacturer first sets the direct sales price p_m^C and the purchase quantity limit Q^C for the consumers. Then, the domestic daigou purchases at the price p_m^C and sets the retail price p_d^C, where $p_d^C > p_m^C$.

The steps outlined above are also illustrated in the visual summary of the game sequence in Figure 2.

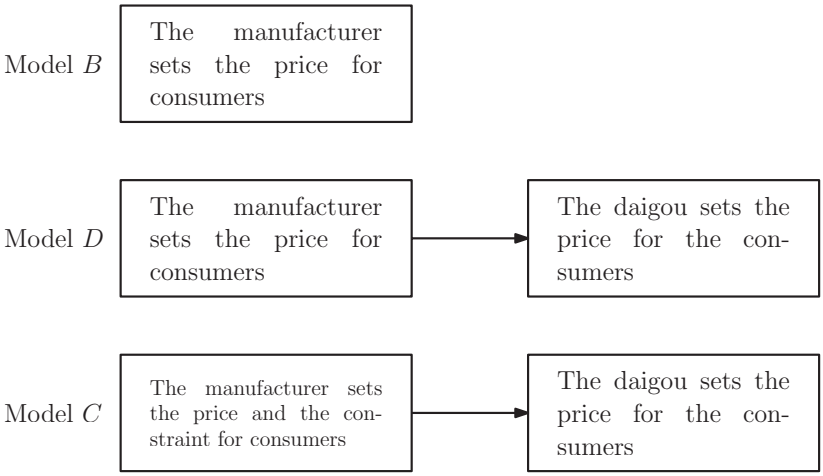


Figure 2. Game sequences of the models.

The notations used in this paper are summarized in Table 1.

Table 1. Summary of notations.

Notation	Explanation
x	$x \in \{m, d\}$, where m and d denote the manufacturer and the daigou, respectively
y	$y \in \{B, D, C\}$, where B, D, and C denote the base model, daigou model, and constrained daigou model, respectively
θ_x^y	The substitution effect of x in y
a	The market size
c	The unit cost of the product
p_x^y	The retail price of x in y
q_x^y	The quantity of x in y
e_x^y	The daigou's effort level of x in y
γ	The influence of efforts on demand
η	The parameter of the market cost
Π_x^y	The profit of x in Mode y
Q^C	The constraint of the daigou's purchase quantity in Model C

3.1. Direct Channel Model

In the Direct Channel Model, the manufacturer operates as the sole retailer in a single-channel market, serving consumers directly without intermediaries. As defined in the general setup, the baseline market size is $a^B = a > 0$, which represents the maximum consumer willingness-to-pay in the absence of daigou entry. The manufacturer incurs a constant marginal production cost c, where $0 < c < a$, to ensure positive profitability. Let $q_m^B \geq 0$ denote the manufacturer's equilibrium sales quantity. With standard linear inverse demand functions in supply chain models, the manufacturer's retail price is determined by

$$p_m^B = a^B - q_m^B \tag{1}$$

$$= a - q_m^B \tag{2}$$

So, we have the profit function

$$\Pi_m^B = (p_m^B - c)q_m^B. \tag{3}$$

Then, we have the following proposition.

Proposition 1. *In the Direct Channel Model, the maximum profit*

$$\Pi_m^{B*} = \frac{(a-c)^2}{4}, \tag{4}$$

where the optimal sales quantity $q_m^{B*} = \frac{a-c}{2}$ *and the optimal sales price* $p_m^{B*} = \frac{a+c}{2}$.

Proof. See Appendix A. □

3.2. Daigou Entry Model

We extend our baseline Direct Channel Model by incorporating the entry of a domestic daigou, establishing a dual-channel supply chain structure as detailed in the general setup. In this model, the manufacturer and daigou compete as independent retailers serving distinct consumer segments, introducing channel competition that expands market access for consumers facing barriers in the direct channel. This competition is modulated by the asymmetric substitution parameters $\theta_d^D > \theta_m^D$, where the daigou's stronger pull θ_d^D reflects its value-added services, potentially cannibalizing the manufacturer's demand more than vice versa.

Let q_m^D and q_d^D represent the equilibrium quantities supplied by the manufacturer's direct channel and the daigou, respectively. As per the effort-driven demand expansion in the general model, the daigou's presence increases the overall market size from the baseline $a^B = a$ to $a^D = a + \gamma e > a^B$, reflecting greater consumer reach through the informal channel [48–51]. The daigou incurs quadratic costs parameterized by η for these efforts, introducing a trade-off that influences its pricing and quantity decisions while ensuring concave profits for equilibrium analysis.

Hence, the inverse demand function is

$$p_m^D = a - \theta_d^D q_d^D - q_m^D \tag{5}$$

and

$$p_d^D = a + \gamma e^D - q_d^D - \theta_m^D q_m^D. \tag{6}$$

Then, the profit functions are obtained as

$$\Pi_m^D = (p_m^D - c)(q_m^D + q_d^D) \tag{7}$$

and

$$\Pi_d^D = (p_d^D - c_d^D)q_d^D - \frac{\eta(e^D)^2}{2} \tag{8}$$

$$= (p_d^D - p_m^D)q_d^D - \frac{\eta(e^D)^2}{2}, \tag{9}$$

where η is the effort cost coefficient.

Proposition 2. *In the Daigou Entry Model, given that* $\frac{\gamma^2}{\eta} < 2(1-\theta_d^D)$*, the maximum profit*

$$\Pi_m^{D*} = \frac{(\eta(3-2\theta_d^D-\theta_m^D)-\gamma^2)(a-c)^2}{4(\eta(2-\theta_d^D-\theta_d^D\theta_m^D)-\gamma^2)}, \tag{10}$$

and

$$\Pi_d^{D*} = \frac{\eta(2\eta(1-\theta_d^D)-\gamma^2)(1-\theta_m^D)^2(a-c)^2}{4(\eta(2-\theta_d^D-\theta_d^D\theta_m^D)-\gamma^2)}, \tag{11}$$

where the optimal effort is $e^{C*} = \frac{\gamma(1-\theta_m^D)(a-c)}{2(\eta(2-\theta_d^D-\theta_d^D\theta_m^D)-\gamma^2)}$*, the optimal sales quantities are* $q_m^{D*} = \frac{(2\eta(1-\theta_d^D)-\gamma^2)(a-c)}{2(\eta(2-\theta_d^D-\theta_d^D\theta_m^D)-\gamma^2)}$ *and* $q_d^{D*} = \frac{\eta(1-\theta_m^D)(a-c)}{2(\eta(2-\theta_d^D-\theta_d^D\theta_m^D)-\gamma^2)}$*, and the optimal sales prices are* $p_m^{D*} = \frac{a+c}{2}$ *and* $p_d^{D*} = \frac{a(\eta(3-2\theta_d^D-\theta_m^D)-\gamma^2)+c(\eta(2-\theta_m^D-\theta_d^D\theta_m^D)-\gamma^2)}{2(\eta(2-\theta_d^D-\theta_d^D\theta_m^D)-\gamma^2)}$.

Proof. See Appendix A. □

From Proposition 2, we see that the constraint $\frac{\gamma^2}{\eta} < 2(1-\theta_d^D)$ is required in the proof of Proposition 2 (see Appendix A). This constraint also guarantees the denominator of Π_m^{D*} and $\Pi_d^{D*} > 0$. To show this, let the denominator term in both Π_m^{D*} and Π_d^{D*} be $4(\eta(2-\theta_d^D-\theta_d^D\theta_m^D)-\gamma^2)$. So, we have

$$\eta(2-\theta_d^D-\theta_d^D\theta_m^D)-\gamma^2 > 0 \tag{12}$$

$$\Leftrightarrow \quad \theta_d^D < \frac{2}{1+\theta_m^D}\left(1-\frac{\gamma^2}{2\eta}\right). \tag{13}$$

Moving the terms from the constraint, we have $\theta_d^D < 1-\frac{\gamma^2}{2\eta}$. As $\theta_m^D < 1$, we have $\frac{2}{1+\theta_m^D} > 1$. Thus, we have

$$\theta_d^D < 1-\frac{\gamma^2}{2\eta} < \frac{2}{1+\theta_m^D}\left(1-\frac{\gamma^2}{2\eta}\right). \tag{14}$$

Hence, this constraint implies the profits $\Pi_m^{D*} > 0$ and $\Pi_d^{D*} > 0$.

3.3. Constrained Daigou Model

In this section, we examine a constrained variant of the Daigou Entry Model in which the domestic daigou faces a purchase limit Q^C per period. We keep the same parameterization as the Daigou Entry Model and replace the superscript D with C to denote the constrained setting. This limit represents a strategic mechanism that the manufacturer can impose to mitigate service degradation for primary customers by restricting daigou procurement. In practice, such constraints can be implemented through mechanisms like membership cards, which limit purchase quantities within a specified period. We characterize the equilibrium as a function of Q^C and conduct a comparative analysis relative to the unconstrained domestic daigou.

Let $\bar{q_d^C} = \min\{q_d^C, Q^C\}$. So, from the inverse functions

$$p_m^C = a-\theta_d^C\bar{q_d^C}-q_m^C, \tag{15}$$

$$p_d^C = a+\gamma e-\bar{q_d^C}-\theta_m^C q_m^C. \tag{16}$$

Then, we have the profit functions

$$\Pi_m^C = (p_m^C-c)(q_m^C+\bar{q_d^C}) \tag{17}$$

and

$$\Pi_d^C = (p_d^C - c_d^C)\bar{q_d^C} - \frac{\eta(e^C)^2}{2} \tag{18}$$

Proposition 3. *In the Constrained Daigou Model, if the purchase limit* $Q^C < \frac{\eta(1-\theta_m^C)(a-c)}{2(\eta(2-\theta_d^C-\theta_d^C\theta_m^C)-\gamma^2)}$, *then the maximum profit*

$$\Pi_m^{C*} = \frac{((a-c)+(1-\theta_d^C)Q^C)^2}{4}, \tag{19}$$

and

$$\Pi_d^{C*} = \frac{(1-\theta_m^C)(a-c)Q^C}{2} - \frac{(\eta(4-(1+\theta_d^C)(1+\theta_m^C))-\gamma^2)(Q^C)^2}{2\eta}, \tag{20}$$

where the optimal effort is $e^{C*} = \frac{\gamma Q^C}{\eta}$, *the optimal sales quantities are* $q_m^{C*} = \frac{a-c-(1-\theta_d^C)Q^C}{2}$ *and* $q_d^{C*} = Q$, *and the optimal sales prices are* $p_m^{C*} = \frac{a+c+(1-\theta_d^C)Q^C}{2}$ *and* $p_d^{C*} = \frac{\gamma^2 Q^C}{\eta} + \frac{2(a-Q^C)+\theta_m^C(1+\theta_d^C)Q^C-\theta_m^C(a-c)}{2}$.

For $Q^C \geq \frac{\eta(1-\theta_m^C)(a-c)}{2(2\eta-\eta\theta_d^C(1+\theta_m^C)-\gamma^2)}$, *the equilibrium coincides with that described in Proposition 2.*

Proof. See Appendix A. □

4. The Impacts of Domestic Daigou

In this section, we will show the impact of domestic daigou presence on the manufacturer's direct sales channel.

Proposition 4. *For Models B and D, the manufacturer's quantities and the prices are*

1. $q_m^{D*} < q_m^{B*} < q_m^{D*} + q_d^{D*}$;
2. $p_m^{B*} = p_m^{D*}$.

Proof. See Appendix A. □

The proposition highlights two key outcomes from the entry of a domestic daigou into the market: an unchanged manufacturer's retail price and a shift in sales quantities, where direct-channel sales decline while total output rises. Economically, this shows the dual role of the daigou as both a competitor and a market expander in an informal, unregulated resale channel.

First, we consider the demand dynamics. In Model B, the manufacturer enjoys a monopoly position with no daigou. It optimizes the direct sales q_m^{B*} to a baseline market of consumers who are willing to buy through the manufacturer's direct channels. In Model D, as the daigou enters, it introduces competition by offering value-added services. This attracts some consumers away from the manufacturer's direct channel, leading to a reduction in direct sales, that is, $q_m^{D*} < q_m^{B*}$. However, the daigou also invests effort into time-sensitive or remote buyers. It expands the overall demand beyond the original market size. This net growth provides the manufacturer's total sales increases, so we have $q_m^{D*} + q_d^{D*} > q_m^{B*}$. Intuitively, the daigou acts like an "uncontracted retailer" that boosts aggregate consumption without the manufacturer bearing the full cost of market expansion, turning potential lost sales into indirect gains.

As Stackelberg leader, the manufacturer sets $p_m^{B*} = p_m^{D*}$ to balance marginal revenue and costs. In Model B, the price is set at the monopoly level to balance marginal revenue

and cost in a single channel. In Model D, the manufacturer anticipates the daigou's entry and resale behavior but cannot directly control it, nor can it offer discounted prices, since daigou operators covertly source goods at retail prices. The daigou's presence introduces asymmetric substitution but the overall demand elasticity does not shift enough to warrant a price adjustment. Instead, the manufacturer maintains the same price to maximize profits from the combined direct and indirect channels, treating daigou purchases as additional demand. Thus, the equilibrium price remains the same, allowing the manufacturer to capture value from the daigou's market-broadening efforts without explicit coordination.

Proposition 5. *For Models B and D, the manufacturer's profit* $\Pi_m^{B*} < \Pi_m^{D*}$.

Proof. See Appendix A. □

Proposition 5 indicates that daigou entry boosts the manufacturer's profit $\Pi_m^{B*} < \Pi_m^{D*}$, despite partial cannibalization of direct sales. This arises because the daigou serves as an informal, cost-free extension of the manufacturer's distribution network. The daigou sources at retail prices, which increases the manufacturer's indirect revenue, and invests its own effort to expand demand into new segments, raising total output on the manufacturer.

5. The Impacts of Manufacturer's Strategy

Proposition 6. *For Models D and C, let* $\theta_d = \theta_d^D = \theta_d^C$ *and* $\theta_m = \theta_m^D = \theta_m^C$*; then*

1. $q_m^{D*} < q_m^{C*}$, *if* $Q^C < \frac{2\theta_d}{1+\theta_d} \cdot \frac{\eta(1-\theta_m)(a-c)}{2(\eta(2-\theta_d-\theta_d\theta_m)-\gamma^2)}$, *and* $q_m^{D*} \geq q_m^{C*}$ *otherwise.*
2. $p_m^{D*} < p_m^{C*}$.
3. $p_d^{D*} < p_d^{C*}$ *if* $\theta_d < \frac{2}{1+\theta_m}\left(1 - \frac{\gamma^2}{\eta}\right)$, *and* $p_d^{D*} \geq p_d^{C*}$ *otherwise.*

Proof. See Appendix A. □

Proposition 6 provides a comparative analysis between the Daigou Entry Model and the Constrained Daigou Model, assuming consistent substitution parameters θ_d and θ_m across both models. This establishes a purchase limit Q^C on the daigou as a strategic tool for the manufacturer to reclaim the market control. It also balances the trade-off between restricting informal channel growth. Tight constraints can shift the demand back to the direct channel, raising prices and potentially profits. On the other hand, loose constraints mimic the unconstrained case, reducing the manufacturer's dominance.

For the manufacturer's quantity, we have $q_m^{D*} < q_m^{C*}$ when Q^C is below the threshold $\frac{2\theta_d}{1+\theta_d} \cdot \frac{\eta(1-\theta_m)(a-c)}{2(\eta(2-\theta_d-\theta_d\theta_m)-\gamma^2)}$, but $q_m^{D*} \geq q_m^{C*}$ otherwise. This is because the constraint limits the daigou's volume in sales. It thus reduces the daigou's ability to influence the convenience-seeking consumers and forcing them to take the manufacturer's direct channel. However, if Q^C is loose, it shifts Model C back to the unconstrained Model D. Therefore, the unconstrained daigou effort leads to greater market expansion and reduces the manufacturer's control compared to the tight constrained case. This shows the constraint's role as a threshold mechanism: below the threshold, the limit binds to protect direct sales; otherwise, the system behaves as in Model D. The threshold itself depends on substitution asymmetry $\theta_d > \theta_m$ and effort cost-effectiveness γ^2/η, where a stronger daigou pulls or efficient efforts raise the bar for effective constraints, as the daigou can still thrive under moderate limits.

The manufacturer's optimal price is higher in Model C than in Model D, $p_m^{D*} < p_m^{C*}$. This shows reduced price competition under constraints. The manufacturer faces less downward pressure from the informal channel's lower-cost, service-enhanced offerings, enabling monopoly-like pricing in the direct channel given the Q^C. Without constraints, the daigou's presence intensifies rivalry, so the manufacturer needs to keep prices lower to retain shares amid substitution risks.

For the daigou, its optimal price is lower in Model D than in Model C, $p_d^{D*} < p_d^{C*}$ when $\theta_d < \frac{2}{1+\theta_m}\left(1 - \frac{\gamma^2}{\eta}\right)$, but higher or equal otherwise. When daigou substitution is weak, that is, θ_d is small, or efforts are inefficient, that is, γ^2/η is small, constraints force scarcity, allowing the daigou to charge a premium for limited stock. Conversely, strong substitution or efficient efforts empower the daigou in the unconstrained case to support market expansion and consumer loyalty for higher pricing, $p_d^{D*} \geq p_d^{C*}$, as it faces less threat from the manufacturer's channel and can pass on value-added costs. This condition shows how the daigou can "power" via asymmetry and their efforts change the pricing dynamic, turning constraints into a disadvantage for its margins.

Proposition 7. *For Models D and C, let* $\theta_d = \theta_d^D = \theta_d^C$ *and* $\theta_m = \theta_m^D = \theta_m^C$. *Let* $K = \eta(2 - \theta_d - \theta_d\theta_m) - \gamma^2$, $M = 2\eta(1 - \theta_d) - \gamma^2$, $N = \eta(4 - (1 + \theta_d)(1 + \theta_m)) - \gamma^2$, *and* $L = (1 - \theta_m)(a - c)$. *The profits satisfy*

1.
   - $\Pi_m^{D*} \geq \Pi_m^{C*}$ *if* $0 \leq Q^C \leq \frac{2\sqrt{K}}{\sqrt{K}+\sqrt{M+N-K}} \cdot \frac{\eta L}{2K}$;
   - $\Pi_m^{D*} < \Pi_m^{C*}$ *if* $\frac{2\sqrt{K}}{\sqrt{K}+\sqrt{M+N-K}} \cdot \frac{\eta L}{2K} < Q^C < \frac{\eta L}{2K}$;
   - $\Pi_m^{D*} = \Pi_m^{C*}$ *if* $Q^C \geq \frac{\eta L}{2K}$.
2. *Consider the quadratic equation* $f(t) = MK - Kt + \frac{N}{2}t^2 = 0$, *and let* $t_1 \leq t_2$ *be its roots (assuming they exist and are real).*
   - *(a) If* $2MN \geq K$, *then* $\Pi_d^{D*} \geq \Pi_d^{C*}$.
   - *(b) If* $2MN < K$, *then*
     - $\Pi_d^{D*} < \Pi_d^{C*}$ *when* $\max\{0, t_1\} \cdot \frac{\eta L}{2K} < Q^C < \min\{t_2, 1\} \cdot \frac{\eta L}{2K}$;
     - $\Pi_d^{D*} \geq \Pi_d^{C*}$ *otherwise.*

Proof. See Appendix A. □

Proposition 7 compares profits between the unconstrained Daigou Entry Model D and the Constrained Daigou Model C, with the same substitution parameters θ_d and θ_m. Economically, the purchase limit Q^C serves as a tunable instrument for the manufacturer to navigate the pressure between utilizing the daigou for market expansion and mitigating its cannibalizing effects on direct sales. The manufacturer can strategically reallocate demand by constraining the daigou's purchase quantities. It potentially enhances its own profits through reduced competition and higher pricing power, while the daigou's response depends on its ability to adapt via effort and substitution advantages.

For the manufacturer, profits in Model D are lower than in Model C when the purchase limit Q^C is greater than $\frac{2\sqrt{\eta(2-\theta_d-\theta_d\theta_m)-\gamma^2}}{\sqrt{\eta(2-\theta_d-\theta_d\theta_m)-\gamma^2}+\sqrt{\eta(3-2\theta_d-\theta_m)-\gamma^2}} \cdot \frac{\eta(1-\theta_m)(a-c)}{2(\eta(2-\theta_d-\theta_d\theta_m)-\gamma^2)}$ but less than $\frac{\eta(1-\theta_m)(a-c)}{2(\eta(2-\theta_d-\theta_d\theta_m)-\gamma^2)}$, and greater than or equal to those in Model C otherwise. This threshold behavior shows a "sweet spot" for the constraints: very low Q^C suppresses the daigou too much, losing the demand expansion benefits and yielding lower manufacturer profits than in the unconstrained case. A moderate Q^C optimally controls the cannibalization, redirecting consumers to the direct channel for higher margins without fully eliminating the daigou's role. Loose Q^C shows the constraint to be ineffective, turning profits to Model D levels, as the daigou operates freely, intensifying competition but also growing the overall profit at the manufacturer's expense.

On the other hand, the daigou's profit is greater than or equal to that in Model C if $2[2\eta(1 - \theta_d) - \gamma^2][\eta(4 - (1 + \theta_d)(1 + \theta_m)) - \gamma^2] \geq \eta(2 - \theta_d - \theta_d\theta_m) - \gamma^2$, and, otherwise, it is lower than in Model C for Q^C in a specific range defined by the roots of the associated quadratic equation and greater than or equal otherwise. Intuitively, when parameters favor the daigou, it can maintain or exceed unconstrained profits even under constraints

by creating scarcity premiums. However, if the condition fails, for example, a strong manufacturer pull via high θ_m, moderate constraints squeeze the daigou's margins by restricting Q^C without allowing full price adjustments, leading to lower profits than in the free-entry scenario.

This explains the observations in luxury goods supply chains, where firms usually implement targeted resale limits to optimize inventory allocation and profitability under gray market pressures, particularly in regions with significant price arbitrage opportunities. In practice, this aligns with the resilience of daigou networks in China's e-commerce ecosystem, where platforms like WeChat facilitate high-volume reselling in the absence of strict controls, yet adaptive strategies under constraints can yield comparable profits by creating scarcity premiums. Hence, these results inform strategic decisions on channel governance, emphasizing the need for controlled constraints to balance short-term profit maximization with long-term market sustainability in the presence of informal intermediaries.

Proposition 8. *For Models D and C,*

1. $\frac{\partial \Pi_m^{D*}}{\partial \theta_m^D} < 0$; $\frac{\partial \Pi_d^{D*}}{\partial \theta_m^D} < 0$.
2. $\frac{\partial \Pi_m^{C*}}{\partial \theta_m^C} = 0$; $\frac{\partial \Pi_d^{C*}}{\partial \theta_m^C} > 0$ *if* $Q^C > \frac{a-c}{1-\theta_d^C}$, *and* $\frac{\partial \Pi_d^{C*}}{\partial \theta_m^C} \leq 0$ *otherwise.*

Proof. See Appendix A. □

Proposition 8 examines how changes in θ_m affect profits in Models D and C, under asymmetric substitution with $\theta_d > \theta_m$. We see that a higher θ_m reduces the daigou's segmentation advantage by making the manufacturer's channel more attractive, therefore increasing cross-channel rivalry and potentially reducing margins for both parties unless constraints change the dynamics.

In the unconstrained daigou Model D, increases in θ_m^D reduce profits for both the manufacturer and the daigou, that is, $\frac{\partial \Pi_m^{D*}}{\partial \theta_m^D} < 0$ and $\frac{\partial \Pi_d^{D*}}{\partial \theta_m^D} < 0$, indicating that enhanced manufacturer pull increases overall channel competition, reducing the segmentation benefits and leading to lower margins with asymmetric substitution where the daigou retains a stronger influence. This shows heightened competition, where a stronger manufacturer pull draws convenience-seekers back from the daigou, reducing its resale volume and forcing price concessions to retain share. For the manufacturer, while this controls some cannibalization, it also reduces the daigou's incentive to invest effort in market expansion, shrinking overall demand and indirect sales, therefore leading to net profit losses in the fiercer rivalry.

In the constrained daigou Model C, the manufacturer's profit remains invariant to changes in θ_m^C, as pricing and quantity decisions are independent of this parameter due to the fixed daigou purchase limit. For the daigou, however, the derivative $\frac{\partial \Pi_d^{C*}}{\partial \theta_m^C} < 0$ when $Q^C < \frac{a-c}{1-\theta_d^C}$, where direct channel demand is positive and the substitution effect harms daigou pricing, and equals zero otherwise.

Proposition 9. *For Models D and C,*

1. $\frac{\partial \Pi_m^{D*}}{\partial \theta_d^D} < 0$ *if* $\theta_m^D < 1 - \frac{\gamma^2}{\eta}$, *and* $\frac{\partial \Pi_m^{D*}}{\partial \theta_d^D} \geq 0$ *otherwise;* $\frac{\partial \Pi_d^{D*}}{\partial \theta_d^D} > 0$ *if* $\theta_d^D < \frac{\theta_m^D}{1+\theta_m^D}\left(2 - \frac{\gamma^2}{\eta}\right)$, *and* $\frac{\partial \Pi_d^{D*}}{\partial \theta_d^D} \leq 0$ *otherwise.*
2. $\frac{\partial \Pi_m^{C*}}{\partial \theta_d^C} < 0$; $\frac{\partial \Pi_d^{C*}}{\partial \theta_d^C} > 0$.

Proof. See Appendix A. □

Proposition 9 examines the sensitivity of optimal profits to changes in the substitution parameter θ_d, which captures the daigou's ability to pull the demand from the manufacturer's direct channel, across the unconstrained Daigou Entry Model and the Constrained Daigou Model. So, the higher θ_d strengthens the daigou's competitiveness through value-added services, increasing the channel conflict and altering profit allocation in a business landscape where informal intermediaries like daigou disrupt traditional retail by arbitraging. This is often observed in luxury and cross-border supply chains, where gray market channels exploit regional price differentials and address consumer demands for personalization, lower costs, and time-convenience through informal resale mechanisms.

In Model D, manufacturer profits decline with increasing θ_d^D, as increased daigou substitution amplifies channel conflict and reduces direct sales margins under asymmetric competition. That is, the daigou obtains more consumers, which disrupts the manufacturer's monopoly-like position and forcing indirect reliance on daigou sales at lower effective margins. On the other hand, if the condition fails, that is, high θ_m^D or low-effort efficiency γ^2/η, profits rise or stabilize, which benefits total output in a cooperative-like equilibrium despite no formal contracts.

For the daigou, profits rise with $\theta_d^D < \frac{\theta_m^D}{1+\theta_m^D}\left(2 - \frac{\gamma^2}{\eta}\right)$, but reduce thereafter, illustrating a non-monotonic effect where substitution enhances market segmentation and daigou value capture, initially enhancing resale markups and returns on effort investments in segmented markets; yet, at higher levels, it increases inter-channel competition, leading to margin disruption as demand substitution worsens.

In Model C, manufacturer profits decrease with θ_d^C, showing persistent cannibalization even under purchase limits, as a stronger daigou pull diverts demand despite caps, reducing direct-channel dominance and highlighting the limits of constraints in fully insulating formal channels from informal competition. On the other hand, daigou profits increase with θ_d^C, as stronger substitution enhances pricing power on constrained volumes, creating scarcity-driven premiums, allowing the daigou to capitalize on its service advantages even under restrictions, therefore demonstrating resilience in adaptive business models where intermediaries turn regulatory problems into opportunities for higher per-unit margins.

Proposition 10. *For Model C,* $\frac{\partial \Pi_m^{C*}}{\partial Q^C} > 0$; $\frac{\partial \Pi_d^{C*}}{\partial Q^C} > 0$ *if* $Q^C < \frac{\eta(1-\theta_m^C)(a-c)}{2(2\eta(3-\theta_d^C-\theta_m^C-\theta_d^C\theta_m^C)-\gamma^2)}$ *and* $\frac{\partial \Pi_d^{C*}}{\partial Q^C} \leq 0$ *otherwise.*

Proof. See Appendix A. □

Proposition 10 analyzes the profits with respect to the purchase limit Q^C in the Constrained Daigou Model C. The manufacturer's profit increases with Q^C, $\frac{\partial \Pi_m^{C*}}{\partial Q^C} > 0$, implying that relaxing the constraint enhances profitability. So, this arises because higher Q^C enables the daigou to procure and resell at a larger quantity, effectively extending the manufacturer's distribution reach at no additional cost. Therefore, it increases overall demand through daigou effort in the daigou's channel while generating indirect revenue from retail-priced sales to the daigou. Although this may lead to partial cannibalization of direct-channel sales, the overall increase in total sales volume more than compensates, consistent with the gray market tolerance in luxury supply chains where manufacturers capitalize on informal intermediaries to enhance system-wide throughput without additional formal investments.

For the daigou, profits increase with Q^C below the threshold $\frac{\eta(1-\theta_m^C)(a-c)}{2(2\eta(3-\theta_d^C-\theta_m^C-\theta_d^C\theta_m^C)-\gamma^2)}$ ($\frac{\partial \Pi_d^{C*}}{\partial Q^C} > 0$), but remain unaffected above $\frac{\partial \Pi_d^{C*}}{\partial Q^C} = 0$. This threshold represents a saturation point of Q^C. At low Q^C, additional procurement capacity enables the daigou to spread

fixed effort costs across greater volumes, and improving margin capture via specialized resale in the daigou's channel. Beyond the threshold, the constraint becomes non-binding; therefore, the daigou achieves unconstrained operations in Model *D*. In business practice, this shows constraint calibration in omni-channel strategies, where moderate limits can incentivize intermediary performance without fully conceding control.

Proposition 11 (Limiting Case Analysis of θ_d^D)**.** *For Model D, let* $\theta_d^D \to 1 - \frac{\gamma^2}{2\eta}$,

1. $p_d^{D*} \to \frac{a+c-c\frac{\gamma^2}{\eta}}{2-\frac{\gamma^2}{\eta}}$;
2. $q_m^{D*} \to 0$ *and* $q_d^{D*} \to \frac{a-c}{2-\frac{\gamma^2}{\eta}}$;
3. $\Pi_m^{D*} \to \frac{(a-c)^2}{2\left(2-\frac{\gamma^2}{\eta}\right)}$ *and* $\Pi_d^{D*} \to 0$.

Proof. The proof is straightforward by substituting the value θ_d^D at the limit $1 - \frac{\gamma^2}{2\eta}$ back into the equation, so we omit the details here. □

From Proposition 11, we see that as the daigou's substitution θ_d^D approaches $1 - \frac{\gamma^2}{2\eta}$, this limit pushes the entire system toward extremes. The daigou's resale price is at

$$p_d^{D*} \to \frac{a + c - c\frac{\gamma^2}{\eta}}{2 - \frac{\gamma^2}{\eta}} > \frac{a+c}{2} = p_m^{D*},$$

which shows a markup that incorporates its effort. On the other hand, we see that direct-channel quantity $q_m^{D*} \to 0$, as consumers switch to the daigou. So, the daigou's quantity is at

$$q_d^{D*} \to \frac{a-c}{2-\frac{\gamma^2}{\eta}}.$$

From the daigou's profit, we see that $\Pi_d^{D*} \to 0$ due to the fact that quadratic effort costs are subtracted from the revenues. Therefore, we see that the manufacturer's profit Π_m^{D*} is all generated from the daigou's direct purchase.

6. Numerical Experiments

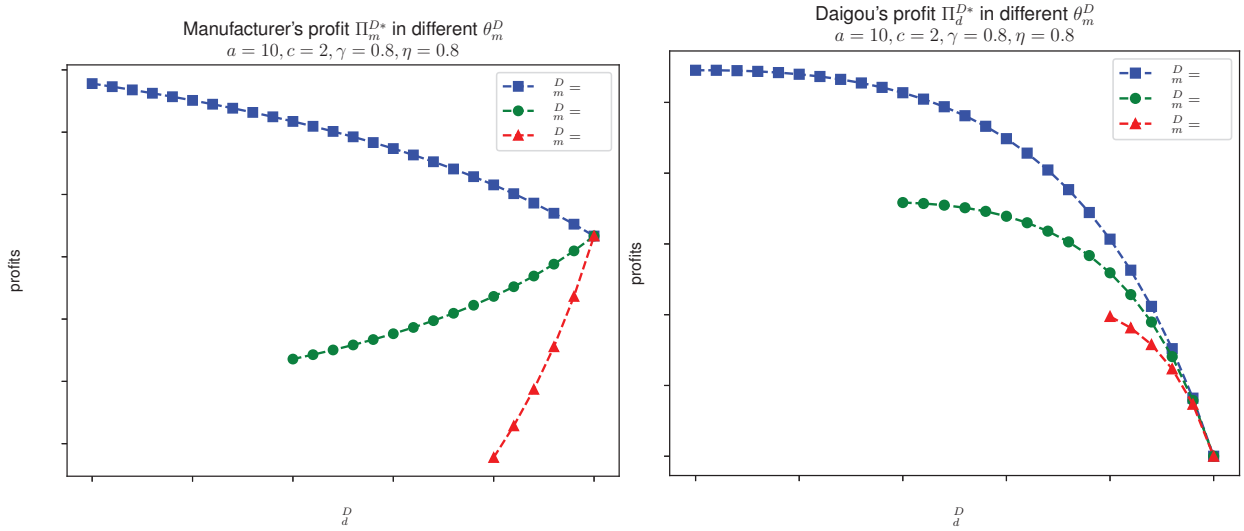
In this section, we perform the numerical experiments for Models *D* and *C*, the manufacturer's profit and daigou's profit, respectively. We first set the market size $a = 10$ and production cost $c = 2$. The optimal manufacturer price $p_m^{D*} = \frac{a+c}{2} = 6$ and a gross profit margin of 66.7%.

In Figure 3, we have the manufacturer's profit and daigou's profit in different values of θ_m^D.

In the following Figure 4, we have Model *C* manufacturer's profit and daigou's profit in different values of θ_m^C.

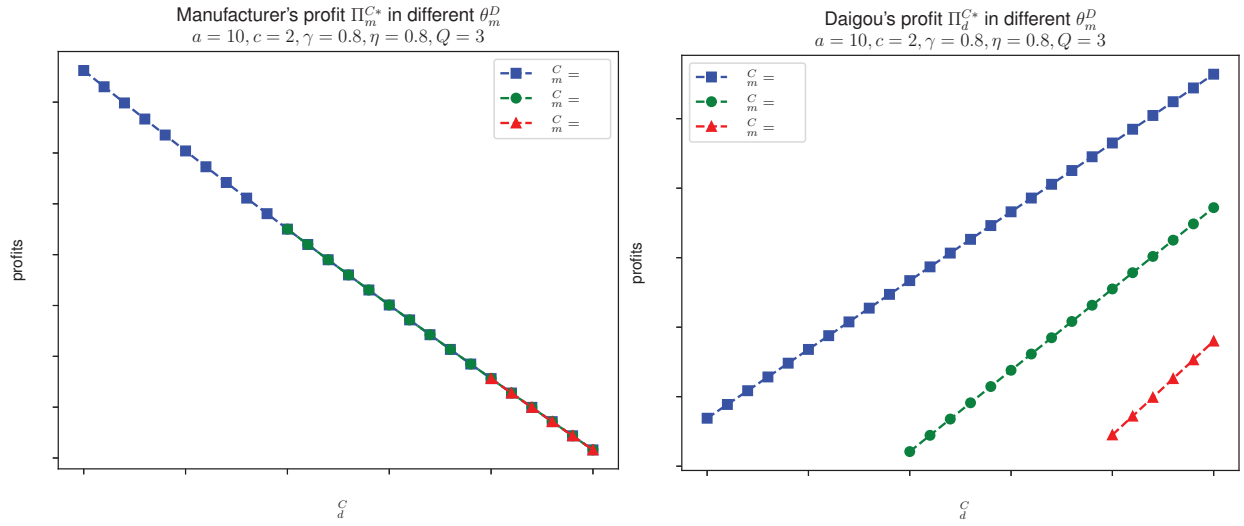
We see that Figures 3 and 4 align with the monotonicity in Proposition 9. In each figure, we vary θ_m across 0.1, 0.3, and 0.5 for comparison, selected with respect to the model's asymmetric assumption that $\theta_m < \theta_d$. So, θ_d is plotted starting from values slightly above each corresponding θ_m. Also, as the chosen effort parameters $\gamma = 0.8$ and $\eta = 0.8$, from Models *D* and *C*, we have $\theta_d < 1 - \frac{\gamma^2}{2\eta} = 0.6$ in this case. Therefore, the upper limit of θ_d in the figures is $\theta_d = 0.6$.

The following Figure 5 shows the daigou's profit Π_d^{C*} against the quantity restriction Q^C, with different values θ_m^C for $a = 10$, $c = 2$, $\gamma = 0.8$, and $\eta = 0.8$. The figure aligns with the monotonicity of Π_d^{C*} in Q^C as shown in Proposition 10.



(**a**) (**b**)

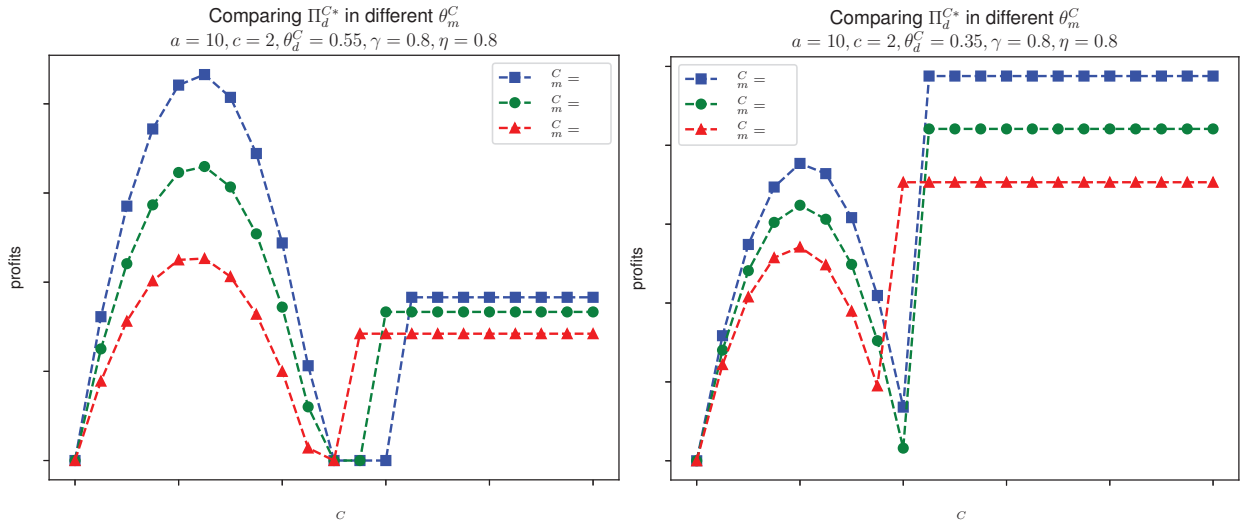
Figure 3. In Model D, manufacturer's profit Π_m^{D*} and daigou's profit Π_d^{D*} in different values of θ_m^D, given that $a = 10$, $b = 2$, $\gamma = 0.8$, and $\eta = 0.8$. (**a**) Π_m^{D*}. (**b**) Π_d^{D*}.



(**a**) (**b**)

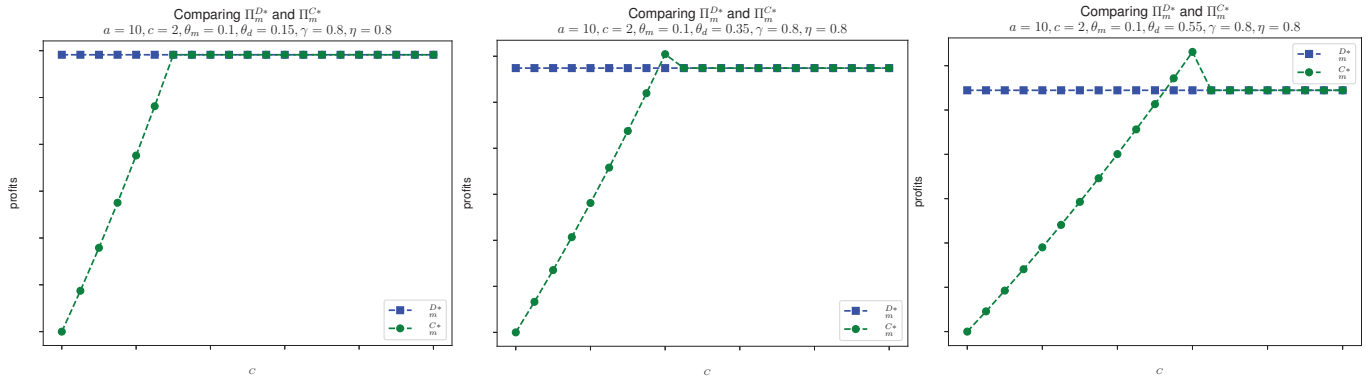
Figure 4. In Model C, manufacturer's profit Π_m^{C*} and daigou's profit Π_d^{C*} in different values of θ_m^C, given that $a = 10$, $b = 2$, $\gamma = 0.8$, $\eta = 0.8$ and $Q = 3$. (**a**) Π_m^{C*}. (**b**) Π_d^{C*}.

In Figure 6, we compare the manufacturer's profit in Models D and C, given that $a = 10$, $c = 2$, $\theta_m = 0.1$, $\gamma = 0.8$ and $\eta = 0.8$. We see that the manufacturer's profit in Model C increases below the threshold for Q^C. For some loose Q^C, the manufacturer's profit in Model C outperforms Model D. This illustration also confirms Proposition 7. A similar case is seen for daigou's profit in Models D and C as shown in Figure 7. For Model C, when Q^C is below the threshold, the daigou's profit is quadratic with respect to Q^C. When the substitution rate $\theta_d = 0.55$, we see that for tight values of Q^C, the daigou's profit is larger than in Model D. This is because for the high substitution, consumers tend to purchase through daigou. Therefore, due to the tight value of Q^C, it creates the scarcity and, therefore, the price at daigou is higher. Hence, daigou's profit in Model C outperforms that in Model D.



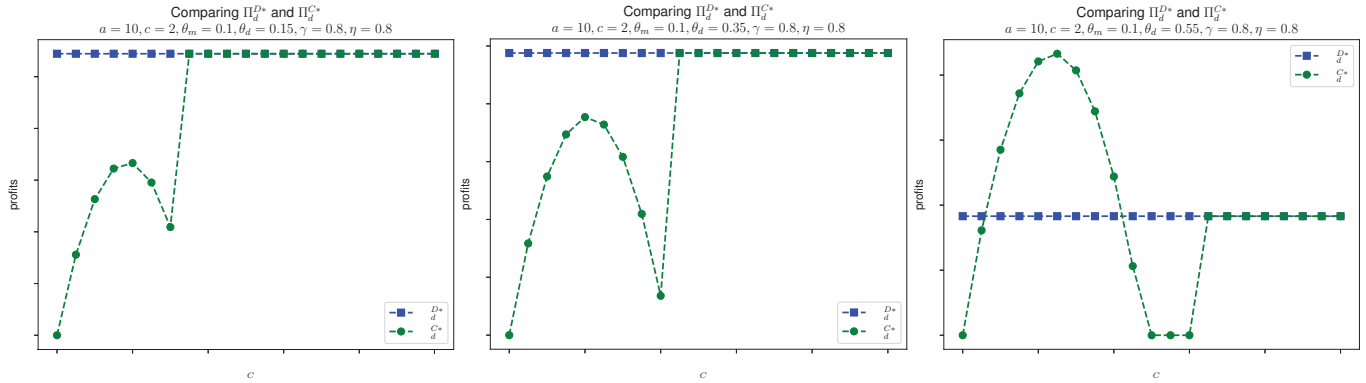
(**a**) (**b**)

Figure 5. In Model *C*, manufacturer's profit Π_m^{C*} and daigou's profit Π_d^{C*} in different values of θ_m^C, given that $a = 10$, $b = 2$, $\gamma = 0.8$, $\eta = 0.8$. (**a**) $\theta_d^C = 0.55$. (**b**) $\theta_d^C = 0.35$.



(**a**) (**b**) (**c**)

Figure 6. Comparing Π_m^{D*} and Π_m^{C*}, given that $a = 10$, $c = 2$, $\theta_m = 0.1$, $\gamma = 0.8$ and $\eta = 0.8$. (**a**) $\theta_d = 0.15$. (**b**) $\theta_d = 0.35$. (**c**) $\theta_d = 0.55$.



(**a**) (**b**) (**c**)

Figure 7. Comparing Π_d^{D*} and Π_d^{C*}, given that $a = 10$, $c = 2$, $\theta_m = 0.1$, $\gamma = 0.8$ and $\eta = 0.8$. (**a**) $\theta_d = 0.15$. (**b**) $\theta_d = 0.35$. (**c**) $\theta_d = 0.55$.

7. Conclusions and Future Research

7.1. Conclusions

The daigou activities at Pang Donglai illustrate how daigou agents can increase sales for Pang Donglai. It also disrupts the traditional supply chain and reduces local customers from the manufacturer's direct sales. The long queues outside Pang Donglai stores have made daigou one of the most effective and convenient channels for completing purchases. Such domestic daigou activities can generate additional sales opportunities for Pang Donglai, but at the same time, reduce local customers from the manufacturer's direct sales. Also, it could flood the manufacturer's direct sales capacity and therefore reduce its excellent customer service standards.

In conclusion, our study shows that daigou actually reduces channel conflicts. When comparing Model B and Model D, the manufacturer's profit is higher in Model D. This outcome arises because daigou informally integrates into the supply chain, expanding market reach and generating additional revenue. Therefore, it enhances overall profitability for the manufacturer as daigou sources products directly from the manufacturer without discounts.

Constraints like purchase limits impact daigou and retailer strategies by balancing profitability with resilience. Compared to Model D, these limits in Model C reduce the manufacturer's profit by reducing daigou's purchase volumes and market expansion. However, for daigou, carefully chosen limits under specific conditions can increase their profits through supply scarcity and higher resale prices. More importantly, such constraints restrict daigou's purchase quantity from the manufacturer. This ensures better service quality for regular consumers, improved inventory management, and stronger resilience against informal competition.

We also show that when daigou substitution dominates, $\theta_d > \theta_m$, it increases channel rivalry. It also disrupt the manufacturer profits in both unconstrained Model D and constrained Model C by shifting demand to informal channels. However, for daigou, higher θ_d initially enhances profits in Model D up to a threshold via better market segmentation. On the other hand, in Model C, it supports resilience through scarcity.

7.2. Limitations and Future Research

This study provides insights into strategic interactions in daigou activities; the findings are constrained by several simplifying assumptions. These include the following:

1. In the models, this work considers only a single daigou in Models D and C; the current analysis does not consider multiple-daigou competition as multiple-daigou competition could cause price wars and reduce individual daigou profits.
2. The models treat substitution effects as constant and asymmetric. It reduces the generalizability to scenarios where consumer preferences evolve dynamically, such as due to marketing or economic shifts. This may bias profit comparisons in the Propositions herein as it may downplay the impact of substitutions on channel conflicts and overall supply chain efficiency.
3. The models are static and do not consider dynamic interactions over time. This may result bias toward short-term equilibria and constrain the applicability to long-term scenarios where daigou adaptation or manufacturer responses could alter profit thresholds.

The assumptions in the analysis were to ensure analytical tractability and focus on the core mechanisms of manufacturer–daigou strategies. There are also several areas that could be improved to better align the model with practical business scenarios. Therefore, future research could address these gaps by conducting the following:

1. Exploring multi-agent dynamics and competition by incorporating multiple daigou agents to analyze intra-daigou rivalry, cooperation thresholds, and network effects on market equilibrium.
2. Incorporate discrete choice models to segment consumers into local versus remote groups, analyzing preference heterogeneity, loyalty dynamics, and the efficacy of manufacturer counter-strategies such as targeted loyalty programs or personalized pricing to mitigate daigou.
3. Develop time-dependent models that consider for inventory, seasonal demand variations, or supply chain uncertainties. It enables the optimization of adaptive purchase limits and resilience strategies in volatile global markets.
4. Investigate how digital platforms can implement governance policies, such as algorithmic monitoring or incentive structures. So, it could regulate daigou activities, balance informal innovation with formal channel protection and exploring impacts on transaction efficiency and trust.
5. Examine the role of government regulations, including tariffs and anti-gray market laws in shaping daigou ecosystems. Then, build the models where interventions could alter profit distributions and assess unintended consequences like supply chain disruptions.

7.3. Managerial Implications

7.3.1. Manufacturer

In the case when informal resale presents, manufacturers need to regard daigou not just as adversaries, it could also be potential business partners. The Daigou Entry Model illustrates how these agents can increase market access. We see that domestic daigou expands the market to remote or convenience-oriented consumers, which increases profit. Yet, this comes at a cost. Daigou activity risks the manufacturer's direct sales profit, which is shown in Proposition 7. Therefore, good purchase limits in the Constrained Daigou Model can outperform Model D's daigou profits when Q^C is chosen within a certain range. Such limits can also bring the demand back to the direct channels. In practice, these constraints can be implemented through mechanisms like membership cards. The membership card record system can be used to restrict purchase quantities within a specified period because typical consumers do not repurchase identical items multiple times in short periods. This method can target daigou without unnecessarily harming regular consumers. Also, managers must constantly track substitution parameters and adjust constraints dynamically. This can reduce the influence from the daigou. Proposition 9 shows the urgency of stricter controls at high daigou substitution θ_d. Early action can help preserve the consumer from direct channels. Managers learn to formalize daigou with partnership, therefore enhancing supply chain robustness.

7.3.2. Daigou

Daigou grows in the unrestricted Model D, whereby utilizing its effort and networks can increase the demand. In Proposition 7, a counterintuitive advantage is observed as a certain range of constraints in Model C yields higher profits via scarcity. Daigou should emphasize operational efficiency over raw volume, perhaps via niche customer ecosystems, to sustain earnings. Proposition 9 demonstrates that robust θ_d increases profits in Model C and even earlier in Model D. Furthermore, to enhance sustainability, daigou could explore formalization as authorized resellers. It could negotiate product discounts from manufacturers to legitimize arbitrage opportunities. Therefore, it can reduce risks associated with informal channels.

Author Contributions: Conceptualization, K.C. and P.H.C.; methodology, K.C. and P.H.C.; validation, R.K.F.I. and P.H.C.; formal analysis, K.C. and P.H.C.; investigation, K.C. and P.H.C.; writing—original draft preparation, K.C., R.K.F.I. and P.H.C.; writing—review and editing, K.C., R.K.F.I. and P.H.C.; visualization, P.H.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: No new data were created or analyzed in this study.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. Proofs

Proof of Proposition 1. As $p_m^B = a - q_m^B$, and

$$\Pi_m^B = (p_m^B - c)q_m^B \tag{A1}$$

$$= (a - q_m^B - c)q_m^B. \tag{A2}$$

From the first-order condition of differentiating Π_m^B with respect to q_m^B, and setting the derivative to be 0, we have $\frac{\mathrm{d}\Pi_m^B}{\mathrm{d}q_m^B} = (a-c) - 2q_m^B = 0$, yielding $q_m^{B*} = \frac{a-c}{2}$. Taking the second-order derivative, we have $\frac{\mathrm{d}^2\Pi_m^B}{\mathrm{d}(q_m^B)^2} < 0$; hence, $q_m^{B*} = \frac{a-c}{2}$ maximizes Π_m^B.

Substituting q_m^{B*} into the inverse demand function, the optimal price is $p_m^{B*} = \frac{a+c}{2}$.

Finally, substituting q_m^{B*} into a profit function, the maximum profit is $\Pi_m^{B*} = \frac{(a-c)^2}{4}$. □

Proof of Proposition 2. From the Stackelberg game, we first optimize the domestic daigou's profit.

Taking the first-order derivatives on the profit Π_d^D with respect to q_m^D and e^D, and setting the derivatives to be 0, we have

$$\begin{cases} \frac{\partial \Pi_d^D}{\partial q_m^D} = e\gamma - 2q_d^D(1-\theta_d^D) + q_m(1-\theta_m^D) & = 0 \\ \frac{\partial \Pi_d^D}{\partial e^D} = -e\eta + \gamma q_d^D & = 0. \end{cases} \tag{A3}$$

Then, we obtain

$$\begin{cases} q_d^{D*} = \frac{\eta q_m^D(1-\theta_m^D)}{2\eta(1-\theta_d^D)-\gamma^2} \\ e^{D*} = \frac{\gamma q_m^D(1-\theta_m^D)}{2\eta(1-\theta_d^D)-\gamma^2}. \end{cases} \tag{A4}$$

So, the Hessian is

$$H^D = \begin{bmatrix} \frac{\partial^2 \Pi_d^D}{\partial (q_d^D)^2} & \frac{\partial^2 \Pi_d^D}{\partial q_d^D \partial e^D} \\ \frac{\partial^2 \Pi_d^D}{\partial e^D \partial q_d^D} & \frac{\partial^2 \Pi_d^D}{\partial (e^D)^2} \end{bmatrix} = \begin{bmatrix} -2(1-\theta_d^D) & \gamma \\ \gamma & -\eta \end{bmatrix}. \tag{A5}$$

Then, the determinant $\det(H^D) = 2\eta(1-\theta_d^D) - \gamma^2$, so, given that $\frac{\gamma^2}{\eta} < 2(1-\theta_d^D)$, we thus have $\det(H^D) > 0$. So, H^D is negative definite, and q_d^{D*} and e^{D*} maximize the domestic daigou's profit:

$$\Pi_d^{D*} = (a + \gamma e^{D*} - q_d^{D*} - \theta_m^D q_m^D - (a - \theta_d^D q_d^{D*} - q_m^D))q_d^{D*} - \frac{\eta (e^{D*})^2}{2} \tag{A6}$$

$$= \frac{\eta (q_m^D)^2 (1-\theta_m^D)^2}{2(2\eta(1-\theta_d^D) - \gamma^2)}. \tag{A7}$$

And the domestic daigou's price is

$$p_d^{D*} = a - \frac{q_m^D(\eta(1+\theta_m^D - 2\theta_d^D\theta_m^D) - \gamma^2)}{2\eta(1-\theta_d^D) - \gamma^2}. \tag{A8}$$

Then, we optimize the manufacturer's profit by setting the first-order derivative $\frac{\partial \Pi_m^D}{\partial q_m^D} = 0$ and the second-order derivative $\frac{\partial^2 \Pi_m^D}{\partial (q_m^D)^2} = -2 < 0$. The profit Π_m^{D*} is thus maximized given that $q_m^{D*} = \frac{(a-c)(2\eta(1-\theta_d^D)-\gamma^2)}{2(\eta(2-\theta_d^D-\theta_d^D\theta_m^D)-\gamma^2)}$. We then substitute q_m^{D*} back to q_d^{D*} and e^{D*}, and obtain

$$q_d^{D*} = \frac{\eta(1-\theta_m^D)(a-c)}{2(\eta(2-\theta_d-\theta_d\theta_m)-\gamma^2)}. \tag{A9}$$

and

$$e^{D*} = \frac{\gamma(1-\theta_m^D)(a-c)}{2(\eta(2-\theta_d-\theta_d\theta_m)-\gamma^2)}. \tag{A10}$$

Similarly, we have the prices

$$p_m^{D*} = \frac{a+c}{2} \tag{A11}$$

and

$$p_d^{D*} = \frac{a(\eta(3-2\theta_d-\theta_m)-\gamma^2) + c(\eta(2-\theta_m-\theta_d\theta_m)-\gamma^2)}{2(\eta(2-\theta_d-\theta_d\theta_m)-\gamma^2)}. \tag{A12}$$

□

Proof of Proposition 3. Since the profit functions are quadratic with a negative leading coefficient, they attain a maximum. When the domestic daigou quantity constraint satisfies $Q^C \geq \frac{\eta(1-\theta_m^C)(a-c)}{2(\eta(2-\theta_d^C-\theta_d^C\theta_m^C)-\gamma^2)}$, the maximum profit is attained at $q_d^C = \frac{\eta(1-\theta_m^C)(a-c)}{2(\eta(2-\theta_d^C-\theta_d^C\theta_m^C)-\gamma^2)}$, as shown in Proposition 2.

Conversely, if the domestic daigou quantity constraint $Q^C < \frac{\eta(1-\theta_m^C)(a-c)}{2(\eta(2-\theta_d^C-\theta_d^C\theta_m^C)-\gamma^2)}$, the profits are maximized by taking $q_d^C = Q^C$. That is, $\Pi_d^C = (p_d^C - p_m^C)Q^C - \frac{\eta(e^C)^2}{2}$. So, taking the first-order derivatives on the profit Π_d^C with respect to e^C, we have

$$\frac{\partial \Pi_d^C}{\partial e^C} = \gamma Q^C - \eta e^C = 0, \tag{A13}$$

and $e^{C*} = \frac{\gamma Q^C}{\eta}$ maximizes the domestic daigou's profit as the second derivative is negative.

We then optimize the manufacturer's profit by setting the first-order derivative $\frac{\partial \Pi_m^C}{\partial q_m^C} = 0$. We have

$$\frac{\partial \Pi_m^C}{\partial q_m^C} = a - c - (1+\theta_d^C)Q^C - 2q_m^C = 0, \tag{A14}$$

and $q_m^{C*} = \frac{a-c-(1+\theta_d^C)Q^C}{2}$ maximizes the manufacturer's profit as the second derivative $\frac{\partial^2 \Pi_m^C}{\partial (q_m^C)^2} = -2 < 0$.

Substituting e^{C*} and q_m^{C*} back into the profits, we have

$$\Pi_m^{C*} = \frac{(a - c + (1 - \theta_d^C)Q^C)^2}{4} \tag{A15}$$

and

$$\Pi_d^{C*} = \frac{(1 - \theta_m^C)(a - c)Q^C}{2} + \frac{(\eta(4 - (1 + \theta_d^C)(1 + \theta_m^C)) - \gamma^2)(Q^C)^2}{2\eta}. \tag{A16}$$

The optimal prices are

$$p_m^{C*} = \frac{a + c + (1 - \theta_d^C)Q^C}{2}, \tag{A17}$$

and

$$p_d^{C*} = \frac{\gamma^2 Q^C}{\eta} + \frac{2(a - Q^C) + \theta_m^C(1 + \theta_d^C)Q^C - \theta_m^C(a - c)}{2}. \tag{A18}$$

□

Proof of Proposition 4. Note that q_m^B and q_m^{D*} are the quantities of the product sold directly by the manufacturer, and $q_d^{D*} + q_m^{D*}$ is the total quantity sold by the manufacturer, including direct sales and domestic daigou sales as domestic daigou purchases at full price from the manufacturer.

We see that $q_m^{B*} = \frac{a-c}{2}$ from Model B, and $q_m^{D*} = \frac{(2\eta(1-\theta_d^D)-\gamma^2)(a-c)}{2(\eta(2-\theta_d^D-\theta_d^D\theta_m^D)-\gamma^2)}$. To compare q_m^{B*} and q_m^{D*}, it is equivalent to check the value of $\frac{2\eta(1-\theta_d^D)-\gamma^2}{\eta(2-\theta_d^D-\theta_d^D\theta_m^D)-\gamma^2}$. So, we have

$$\frac{2\eta(1 - \theta_d^D) - \gamma^2}{\eta(2 - \theta_d^D - \theta_d^D\theta_m^D) - \gamma^2} - 1 = \frac{\eta\theta_d^D\theta_m^D - \eta\theta_d^D}{\eta(2 - \theta_d^D - \theta_d^D\theta_m^D) - \gamma^2} \tag{A19}$$

$$= -\frac{\eta\theta_d^D(1 - \theta_m^D)}{\eta(2 - \theta_d^D - \theta_d^D\theta_m^D) - \gamma^2} < 0, \tag{A20}$$

as θ_d^D and θ_m^D are valued within $[0, 1)$. This shows $q_m^{D*} < q_m^{B*}$.

On the other hand, we have

$$q_d^{D*} + q_m^{D*} = \frac{a - c}{2} \cdot \frac{\eta(1 - \theta_m^D) + (2\eta(1 - \theta_d^D) - \gamma^2)}{\eta(2 - \theta_d^D - \theta_d^D\theta_m^D) - \gamma^2}, \tag{A21}$$

and we have

$$\frac{\eta(1 - \theta_m^D) + (2\eta(1 - \theta_d^D) - \gamma^2)}{\eta(2 - \theta_d^D - \theta_d^D\theta_m^D) - \gamma^2} - 1 = \frac{\eta - \eta\theta_m^D - \eta\theta_d^D + \eta\theta_d^D\theta_m^D}{\eta(2 - \theta_d^D - \theta_d^D\theta_m^D) - \gamma^2} \tag{A22}$$

$$= \frac{\eta(1 - \theta_d^D)(1 - \theta_m^D)}{\eta(2 - \theta_d^D - \theta_d^D\theta_m^D) - \gamma^2} > 0. \tag{A23}$$

So, this shows $q_m^{B*} < q_d^{D*} + q_m^{D*}$.

And, for the prices, we have $p_m^{B*} = p_m^{D*} = \frac{a+c}{2}$ from Propositions 1 and 2. □

Proof of Proposition 5. From Propositions 1 and 2, we have

$$\Pi_m^{B*} = \frac{(a - c)^2}{4} \tag{A24}$$

and

$$\Pi_m^{D*} = \frac{(a-c)^2}{4} \cdot \frac{\eta(3-2\theta_d^D-\theta_m^D)-\gamma^2}{\eta(2-\theta_d^D-\theta_d^D\theta_m^D)-\gamma^2}. \tag{A25}$$

Then, taking

$$\frac{\eta(3-2\theta_d^D-\theta_m^D)-\gamma^2}{\eta(2-\theta_d^D-\theta_d^D\theta_m^D)-\gamma^2} - 1 = \frac{\eta(1-\theta_d^D)(1-\theta_m^D)}{\eta(2-\theta_d^D-\theta_d^D\theta_m^D)-\gamma^2} > 0, \tag{A26}$$

we thus have $\Pi_m^{B*} < \Pi_m^{D*}$. □

Proof of Proposition 6. Let the substitution parameter $\theta_m = \theta_m^D = \theta_m^C$ and $\theta_d = \theta_d^D = \theta_d^C$, and let $\theta_m = \theta_m^D = \theta_m^C$ and $\theta_d = \theta_d^D = \theta_d^C$, from Propositions 2 and 3; we see that $p_m^{D*} < p_m^{C*}$. Also, we have $q_m^{D*} = \frac{(2\eta(1-\theta_d)-\gamma^2)(a-c)}{2(\eta(2-\theta_d-\theta_d\theta_m)-\gamma^2)}$, $q_m^{C*} = \frac{-Q*\theta_d-Q+a-c}{2}$ and $Q_{\max}^C = \frac{\eta(1-\theta_m)(a-c)}{2(\eta(2-\theta_d-\theta_d\theta_m)-\gamma^2)}$.

We see that $q_m^{D*} = \frac{a-c}{2} - \theta_d \cdot Q_{\max}^C$ and $q_m^{C*} = \frac{a-c}{2} - \frac{1+\theta_d}{2}Q^C$. So, let $Q^{C*} = \frac{2\theta_d}{1+\theta_d}Q_{\max}^C < Q_{\max}^C$ as $\theta_d < 1$; then, if $Q^C < Q^{C*}$, we have $q_m^{D*} < q_m^{C*}$. Otherwise, $q_m^{D*} \geq q_m^{C*}$.

On the other hand, to compare p_d^{D*} and p_d^{C*}, we first let

$$M = \eta(2-\theta_d-\theta_d\theta_m)-\gamma^2 \tag{A27}$$

$$N = (a-c)\eta(1-\theta_d)(1-\theta_m), \tag{A28}$$

and $Q_{\max}^C = \frac{N}{2M(1-\theta_d)}$.

We then have

$$p_d^{D*} = \frac{a(\eta(3-2\theta_d-\theta_m)-\gamma^2)+c(\eta(2-\theta_m-\theta_d\theta_m)-\gamma^2)}{2(\eta(2-\theta_d-\theta_d\theta_m)-\gamma^2)} \tag{A29}$$

$$= \frac{a+c}{2} + \frac{N}{2M} \tag{A30}$$

and

$$p_d^{C*} = \frac{\gamma^2 Q^C}{\eta} + \frac{2(a-Q^C)+\theta_m^C(1+\theta_d)Q^C-\theta_m(a-c)}{2} \tag{A31}$$

$$= \frac{2a-\theta_m(a-c)}{2} - \frac{M-\gamma^2}{2\eta} \cdot Q^C. \tag{A32}$$

We see that $p_d^{C*}(Q^C)$ is decreasing in Q^C if $\eta > \frac{2\gamma^2}{2-\theta_d-\theta_d\theta_m}$; so taking $p_d^{C*}(Q_{\max}^C)$, we have

$$p_d^{C*}(Q_{\max}^C) = \frac{2a-\theta_m(a-c)}{2} - \frac{M-\gamma^2}{2\eta}Q_{\max}^C \tag{A33}$$

$$= \frac{a+c}{2} + \frac{a(1-\theta_m)}{2} - \frac{M-\gamma^2}{2\eta} \cdot \frac{N}{2M(1-\theta_d)} \tag{A34}$$

$$= \frac{a+c}{2} + \frac{a(1-\theta_m)}{2} - \frac{(a-c)(1-\theta_m)(M-\gamma^2)}{4M} \tag{A35}$$

$$= \frac{a+c}{2} + \frac{1-\theta_m}{4M}(2aM-(a-c)(M-\gamma^2)) \tag{A36}$$

$$= \frac{a+c}{2} + \frac{1-\theta_m}{4M}((a+c)M+(a-c)\gamma^2). \tag{A37}$$

So, taking $p_d^{C*}(Q_{\max}^C) - p_d^{D*}$, we have

$$p_d^{C*}(Q_{\max}^C) - p_d^{D*} \tag{A38}$$
$$= \frac{1-\theta_m}{4M}((a+c)M + (a-c)\gamma^2) - \frac{N}{2M} \tag{A39}$$
$$= \frac{1}{4M}((a+c)(1-\theta_m)M + (a-c)(1-\theta_m)\gamma^2 - 2N) \tag{A40}$$
$$= \frac{1}{4M}((a+c)(1-\theta_m)M + (a-c)(1-\theta_m)\gamma^2 - 2(a-c)\eta(1-\theta_d)(1-\theta_m)) \tag{A41}$$
$$= \frac{1-\theta_m}{4M}((a+c)(\eta(2-\theta_d-\theta_d\theta_m) - \gamma^2) + (a-c)\gamma^2 - 2(a-c)\eta(1-\theta_d)) \tag{A42}$$
$$= \frac{1-\theta_m}{4M}((a+c)\eta(2-\theta_d-\theta_d\theta_m) - 2c\gamma^2 - 2(a-c)\eta(1-\theta_d)) \tag{A43}$$
$$= \frac{1-\theta_m}{4M}(\eta(2(a+c) - (a+c)\theta_d - (a+c)\theta_d\theta_m - 2(a-c) + 2(a-c)\theta_d) - 2c\gamma^2) \tag{A44}$$
$$= \frac{1-\theta_m}{4M}(\eta(4c + a\theta_d(1-\theta_m) - c\theta_d(3+\theta_m)) - 2c\gamma^2) > 0 \tag{A45}$$

if $\eta > \frac{2\gamma^2}{4+\frac{a}{c}\theta_d(1-\theta_m)-\theta_d(3+\theta_m)}$. As we have $\eta > \frac{2\gamma^2}{2-\theta_d-\theta_d\theta_m}$, and we need

$$4 + \frac{a}{c}\theta_d(1-\theta_m) - \theta_d(3+\theta_m) > 2 - \theta_d - \theta_d\theta_m \tag{A46}$$
$$\Leftrightarrow \quad 2(1-\theta_d) + \frac{a}{c}\theta_d(1-\theta_m) > 0; \tag{A47}$$

hence, we have $p_d^{C*} > p_d^{C*}(Q_{\max}^C) > p_d^{D*}$ as $\eta > \frac{2\gamma^2}{2-\theta_d-\theta_d\theta_m}$. □

Proof of Proposition 7. Assuming we have the substitution parameter $\theta_m^D = \theta_m^C$ and $\theta_d^D = \theta_d^C$, and let $\theta_m = \theta_m^D = \theta_m^C$ and $\theta_d = \theta_d^D = \theta_d^C$. Let

$$K = \eta(2-\theta_d-\theta_d\theta_m) - \gamma^2, \tag{A48}$$
$$M = 2\eta(1-\theta_d) - \gamma^2, \tag{A49}$$
$$N = \eta(4-(1+\theta_d)(1+\theta_m)) - \gamma^2, \tag{A50}$$
$$L = (1-\theta_m)(a-c). \tag{A51}$$

We have

$$\Pi_m^{D*} = \frac{(3\eta - 2\eta\theta_d^D - \eta\theta_m^D - \gamma^2)(a-c)^2}{4(\eta(2-\theta_d^D-\theta_d^D\theta_m^D) - \gamma^2)} \tag{A52}$$
$$= \frac{(3\eta - 2\eta\theta_d - \eta\theta_m - \gamma^2)(a-c)^2}{4(\eta(2-\theta_d-\theta_d\theta_m) - \gamma^2)}, \tag{A53}$$
$$= \frac{(M+N-K)(a-c)^2}{4K}, \tag{A54}$$

and

$$\Pi_m^{C*} = \frac{(a-c+(1-\theta_d^C)Q^C)^2}{4} \tag{A55}$$
$$= \frac{(a-c+(1-\theta_d)Q^C)^2}{4}. \tag{A56}$$

Solving Q^C by setting $\Pi_m^{D*} = \Pi_m^{C*}$, we obtain

$$\frac{M+N-K}{K} \cdot \frac{(a-c)^2}{4} = \frac{((a-c)+(1-\theta_d)Q^C)^2}{4}. \tag{A57}$$

Multiplying 4 on both sides, and taking the square root, as $a - c > 0$ and $Q^C > 0$, we have

$$\sqrt{\frac{M+N-K}{K}}(a-c) = (a-c+(1-\theta_d)Q^C), \tag{A58}$$

$$Q^{C*} = \frac{(a-c)\left(\sqrt{\frac{M+N-K}{K}}-1\right)}{1-\theta_d} = \frac{\eta L}{K\left(\sqrt{\frac{M+N-K}{K}}+1\right)}. \tag{A59}$$

Then, we verify whether $Q^{C*} < Q^C_{\max} = \frac{\eta L}{2K}$, let $r = \frac{M+N-K}{K}$, and taking

$$\frac{Q^{C*}}{Q^C_{\max}} = \frac{\frac{\eta L}{K(\sqrt{r}+1)}}{\frac{\eta L}{2K}} \tag{A60}$$

$$= \frac{2}{\sqrt{r}+1} \tag{A61}$$

$$< 1 \tag{A62}$$

as $r > 1$ and $\sqrt{r}+1 > 2$. Therefore, we have shown that $Q^{C*} < Q^C_{\max}$. Taking $Q^C = 0$, we have $\Pi^{C*}_m(Q^C = 0) = \frac{(a-c)^2}{4} < \frac{M+N-K}{K} \cdot \frac{(a-c)^2}{4} = \Pi^{D*}_m$. As $\Pi^{C*}_m(Q^C)$ is monotonic increasing in Q^C, therefore, we can conclude that $\Pi^{C*}_m(Q^C = Q^{C*}) = \Pi^{D*}_m$ at $Q^C = Q^{C*}$ and $\Pi^{C*}_m(Q^C) > \Pi^{D*}_m$ for $Q^{C*} < Q^C < Q^C_{\max}$.

Next, we compare Π^{D*}_d and Π^{C*}_d. We first let

$$K = \eta(2-\theta_d-\theta_d\theta_m) - \gamma^2, \tag{A63}$$

$$M = 2\eta(1-\theta_d) - \gamma^2, \tag{A64}$$

$$N = \eta(4-(1+\theta_d)(1+\theta_m)) - \gamma^2, \tag{A65}$$

$$L = (1-\theta_m)(a-c). \tag{A66}$$

We then have

$$\Pi^{D*}_d = \frac{\eta M L^2}{4K} \tag{A67}$$

and

$$\Pi^{C*}_d = \frac{LQ^C}{2} - \frac{N(Q^C)^2}{2\eta}. \tag{A68}$$

Similarly, the upper bound of Q^C is denoted as $Q^C_{\max} = \frac{\eta L}{2K}$. Since $Q^C < Q^C_{\max}$, we let $t = \frac{Q^C}{Q^C_{\max}}$, where $t \in [0,1)$ and we have $Q^C = tQ^C_{\max}$. So,

$$\Pi^{C*}_d = \frac{LQ^C}{2} - \frac{N(Q^C)^2}{2\eta} \tag{A69}$$

$$= \frac{LtQ^C_{\max}}{2} - \frac{N(tQ^C_{\max})^2}{2\eta} \tag{A70}$$

$$= \frac{\eta LKtQ^C_{\max}}{2\eta K} - \frac{N(tQ^C_{\max})^2}{2\eta} \tag{A71}$$

$$= \frac{tK(Q^C_{\max})^2}{\eta} - \frac{N(tQ^C_{\max})^2}{2\eta} \tag{A72}$$

$$= \frac{(Q^C_{\max})^2}{\eta}\left(Kt - \frac{N}{2}t^2\right). \tag{A73}$$

Similarly, we have

$$\Pi_d^{D*} = \frac{MK}{\eta}(Q_{\max}^C)^2. \tag{A74}$$

Taking

$$\Pi_d^{D*} - \Pi_d^{C*} = \frac{(Q_{\max}^C)^2}{\eta}\left(MK - Kt + \frac{N}{2}t^2\right), \tag{A75}$$

we let $f(t) = MK - Kt + \frac{N}{2}t^2$; we see that the quadratic function $f(t)$ is convex, where the axis of symmetry is $t = \frac{N}{K} \in (0,1)$. Hence, the minimum value of $f_{\min}(t) = f\left(\frac{K}{N}\right)$. That is, we have

$$f\left(\frac{K}{N}\right) = MK - \frac{K^2}{N} + \frac{N}{2}\cdot\frac{K^2}{N^2} \tag{A76}$$

$$= MK - \frac{K^2}{2N} \tag{A77}$$

$$= K(M - \frac{K}{2N}), \tag{A78}$$

which is $f_{\min}(t) = f\left(\frac{K}{N}\right) = 0$ if $2MN = K$.

This implies that, if $2(2\eta(1-\theta_d) - \gamma^2)(\eta(4-(1+\theta_d)(1+\theta_m)) - \gamma^2) \geq \eta(2-\theta_d - \theta_d\theta_m) - \gamma^2$, we have $\Pi_d^{D*} \geq \Pi_d^{C*}$ for $Q^C < Q_{\max}^C$ and $\Pi_d^{D*} = \Pi_d^{C*}$ for $Q^C \geq Q_{\max}^C$. On the other hand, we have $\Pi_d^{D*} < \Pi_d^{C*}$ for $t_1 Q_{\max}^C < Q^C < t_2 Q_{\max}^C$; t_1 and t_2 are the two roots of $f(t) = 0$, where $t_1 < t_2$ and $\Pi_d^{D*} \geq \Pi_d^{C*}$ otherwise. □

Proof of Proposition 8. As the profit of Model D, we have

$$\Pi_m^{D*} = \frac{(3\eta - 2\eta\theta_d^D - \eta\theta_m^D - \gamma^2)(a-c)^2}{4(\eta(2-\theta_d-\theta_d\theta_m) - \gamma^2)}, \tag{A79}$$

and

$$\Pi_d^{D*} = \frac{\eta(2\eta(1-\theta_d^D) - \gamma^2)(1-\theta_m^D)^2(a-c)^2}{4(\eta(2-\theta_d-\theta_d\theta_m) - \gamma^2)}. \tag{A80}$$

So, taking the derivative of θ_m^D, we have

$$\frac{\partial \Pi_m^{D*}}{\partial \theta_m^D} = -\frac{\eta(a-c)^2(1-\theta_d^D)(2\eta(1-\theta_d^D) - \gamma^2)}{4(\eta(2-\theta_d^D - \theta_d^D\theta_m^D) - \gamma^2)^2} < 0 \tag{A81}$$

and

$$\frac{\partial \Pi_d^{D*}}{\partial \theta_m^D} = -\frac{\eta(a-c)^2(1-\theta_m^D)(2\eta(1-\theta_d^D) - \gamma^2)^2}{4(\eta(2-\theta_d^D - \theta_d^D\theta_m^D) - \gamma^2)^3} < 0. \tag{A82}$$

On the other hand, for Model C, we have

$$\frac{\partial \Pi_m^{C*}}{\partial \theta_m^C} = 0, \tag{A83}$$

since Π_m^{C*} does not depend on θ_m^C. For the daigou's profit $\Pi_d^{C*} = \frac{(1-\theta_m^C)(a-c)Q^C}{2} + \frac{(\eta(4-(1+\theta_d^C)(1+\theta_m^C)) - \gamma^2)(Q^C)^2}{2\eta}$, we have

$$\frac{\partial \Pi_d^{C*}}{\partial \theta_m^C} = \frac{(Q^C)^2(1+\theta_d^C) - Q^C(a-c)}{2}. \tag{A84}$$

Hence, $\frac{\partial \Pi_d^{C*}}{\partial \theta_m^C} > 0$ if $Q^C(1+\theta_d^C) > a-c$ and $\frac{\partial \Pi_d^{C*}}{\partial \theta_m^C} \leq 0$ otherwise. □

Proof of Proposition 9. As the profit of Model D, we have

$$\Pi_m^{D*} = \frac{(3\eta - 2\eta\theta_d^D - \eta\theta_m^D - \gamma^2)(a-c)^2}{4(\eta(2-\theta_d-\theta_d\theta_m)-\gamma^2)}, \tag{A85}$$

and

$$\Pi_d^{D*} = \frac{\eta(2\eta(1-\theta_d^D)-\gamma^2)(1-\theta_m^D)^2(a-c)^2}{4(\eta(2-\theta_d-\theta_d\theta_m)-\gamma^2)}. \tag{A86}$$

So, taking the derivative of θ_d^D, we have

$$\frac{\partial \Pi_m^{D*}}{\partial \theta_d^D} = -\frac{\eta(a-c)^2(1-\theta_m^D)(\eta(1-\theta_m^D)-\gamma^2)}{4(\eta(2-\theta_d^D-\theta_d^D\theta_m^D)-\gamma^2)^2} < 0. \tag{A87}$$

if $\theta_m^D < 1 - \frac{\gamma^2}{\eta}$. And for the daigou's profit,

$$\frac{\partial \Pi_d^{D*}}{\partial \theta_d^D} = \frac{\eta^2(a-c)^2(1-\theta_m^D)^2(\eta(2\theta_m^D-\theta_d^D-\theta_d^D\theta_m^D)-\gamma^2\theta_m^D)}{4(\eta(2-\theta_d^D-\theta_d^D\theta_m^D)-\gamma^2)^3} > 0, \tag{A88}$$

if $\theta_d^D < \frac{\theta_m^D}{1+\theta_m^D}\left(2-\frac{\gamma^2}{\eta}\right)$ and $\frac{\partial \Pi_d^{D*}}{\partial \theta_d^D} \leq 0$ otherwise.

Considering Model C, we have

$$\Pi_m^{C*} = \frac{((a-c)+(1-\theta_d^C)Q^C)^2}{4}, \tag{A89}$$

and

$$\Pi_d^{C*} = \frac{(1-\theta_m^C)(a-c)Q^C}{2} + \frac{(\eta(4-(1+\theta_d^C)(1+\theta_m^C))-\gamma^2)(Q^C)^2}{2\eta}. \tag{A90}$$

So, we have the derivative

$$\frac{\partial \Pi_m^{C*}}{\partial \theta_d^C} = -\frac{(a-c)Q^C + (1-\theta_d^C)(Q^C)^2}{2} < 0 \tag{A91}$$

and

$$\frac{\partial \Pi_d^{C*}}{\partial \theta_d^C} = \frac{(1+\theta_m^C)(Q^C)^2}{2} > 0. \tag{A92}$$

□

Proof of Proposition 10. We see that

$$\frac{\partial \Pi_m^{C*}}{\partial Q^C} = \frac{(1-\theta_d^C)(a-c+Q^C(1-\theta_d^C))}{2} > 0, \tag{A93}$$

and we have

$$\frac{\partial \Pi_d^{C*}}{\partial Q^C} = \frac{(1-\theta_m^C)(a-c)}{2} - Q^C(3-\theta_d^C-\theta_m^C-\theta_d^C\theta_m^C-\frac{\gamma^2}{\eta}), \tag{A94}$$

so $\frac{\partial \Pi_d^{C*}}{\partial Q^C} > 0$ if $Q^C < \frac{\eta(1-\theta_m^C)(a-c)}{2(4-(1+\theta_d^C)(1+\theta_m^C)-\gamma^2)}$, and $\frac{\partial \Pi_d^{C*}}{\partial Q^C} \leq 0$ otherwise. □

References

1. Zhang, X.; Tham, A.; Liu, Y.; Spinks, W.; Wang, L. 'To market, to market': Uncovering Daigou touristscapes within Chinese outbound tourism. *J. China Tour. Res.* **2021**, *17*, 549–569. [CrossRef]
2. Park, H.; Kim, S.; Jeong, Y.; Minshall, T. Customer entrepreneurship on digital platforms: Challenges and solutions for platform business models. *Creat. Innov. Manag.* **2021**, *30*, 96–115. [CrossRef]
3. Zhuoxiao, X. Im/materializing cross-border mobility: A study of mainland China–Hong Kong Daigou (Cross-border shopping services on global consumer goods). *Int. J. Commun.* **2018**, *12*, 4052–4065.
4. Zhu, Z.; Tang, X. Effect of integration capabilities with channel distributors on supply chain agility in emerging markets: An institution-based view perspective. *J. Enterp. Inf. Manag.* **2023**, *36*, 381–408. [CrossRef]
5. Glavas, C.; Mortimer, G.; Ding, H.; Grimmer, L.; Vorobjovas-Pinta, O.; Grimmer, M. How entrepreneurial behaviors manifest in non-traditional, heterodox contexts: Exploration of the Daigou phenomenon. *J. Bus. Ventur. Insights* **2023**, *19*, e00385. [CrossRef]
6. Kohn, M. Competitive speculation. *Econom. J. Econom. Soc.* **1978**, *46*, 1061–1076. [CrossRef]
7. Tirole, J. On the possibility of speculation under rational expectations. *Econom. J. Econom. Soc.* **1982**, *50*, 1163–1181. [CrossRef]
8. Courty, P. Some economics of ticket resale. *J. Econ. Perspect.* **2003**, *17*, 85–97. [CrossRef]
9. Courty, P. Ticket pricing under demand uncertainty. *J. Law Econ.* **2003**, *46*, 627–652. [CrossRef]
10. Geng, X.; Wu, R.; Whinston, A.B. Profiting from partial allowance of ticket resale. *J. Mark.* **2007**, *71*, 184–195. [CrossRef]
11. Su, X. Optimal pricing with speculators and strategic consumers. *Manag. Sci.* **2010**, *56*, 25–40. [CrossRef]
12. Lim, W.S.; Tang, C.S. Advance selling in the presence of speculators and forward-looking consumers. *Prod. Oper. Manag.* **2013**, *22*, 571–587. [CrossRef]
13. Feng, T.; Geunes, J. Speculation in a two-stage retail supply chain. *IIE Trans.* **2014**, *46*, 1315–1328. [CrossRef]
14. Huang, Y.S.; Gu, Y.H.; Fang, C.C. Pricing of perishable products with a speculator and strategic customers. *Int. J. Syst. Sci. Oper. Logist.* **2019**, *6*, 301–319. [CrossRef]
15. Kuksov, D.; Liao, C. Restricting Speculative Reselling: When "How Much" Is the Question. *Mark. Sci.* **2023**, *42*, 377–400. [CrossRef]
16. Zhao, K.; Zhao, X.; Deng, J. Online price dispersion revisited: How do transaction prices differ from listing prices? *J. Manag. Inf. Syst.* **2015**, *32*, 261–290. [CrossRef]
17. Zhao, K.; Zhao, X.; Deng, J. An empirical investigation of online gray markets. *J. Retail.* **2016**, *92*, 397–410. [CrossRef]
18. Bai, H.; McColl, J.; Moore, C. Luxury retailers' entry and expansion strategies in China. *Int. J. Retail Distrib. Manag.* **2017**, *45*, 1181–1199. [CrossRef]
19. Huang, H.; He, Y.; Chen, J. Competitive strategies and quality to counter parallel importation in global market. *Omega* **2019**, *86*, 173–197. [CrossRef]
20. Zhang, Z.; Feng, J. Price of identical product with gray market sales: An analytical model and empirical analysis. *Inf. Syst. Res.* **2017**, *28*, 397–412. [CrossRef]
21. Prasetyo, E.H. Digital platforms' strategies in Indonesia: Navigating between technology and informal economy. *Technol. Soc.* **2024**, *76*, 102414. [CrossRef]
22. Belk, R. You are what you can access: Sharing and collaborative consumption online. *J. Bus. Res.* **2014**, *67*, 1595–1600. [CrossRef]
23. Schlagwein, D.; Schoder, D.; Spindeldreher, K. Consolidated, systemic conceptualization, and definition of the "sharing economy". *J. Assoc. Inf. Sci. Technol.* **2020**, *71*, 817–838. [CrossRef]
24. Marth, S.; Hartl, B.; Penz, E. Sharing on platforms: Reducing perceived risk for peer-to-peer platform consumers through trust-building and regulation. *J. Consum. Behav.* **2022**, *21*, 1255–1267. [CrossRef] [PubMed]
25. Tang, W.; Xie, N.; Mo, D.; Cai, Z.; Lee, D.H.; Chen, X.M. Optimizing subsidy strategies of the ride-sourcing platform under government regulation. *Transp. Res. Part E Logist. Transp. Rev.* **2023**, *173*, 103112. [CrossRef]
26. Liu, Y.; Wang, X.; Gilbert, S.; Lai, G. On the participation, competition and welfare at customer-intensive discretionary service platforms. *Manuf. Serv. Oper. Manag.* **2023**, *25*, 218–234. [CrossRef]
27. Bai, J.; So, K.C.; Tang, C.S.; Chen, X.; Wang, H. Coordinating supply and demand on an on-demand service platform with impatient customers. *Manuf. Serv. Oper. Manag.* **2019**, *21*, 556–570. [CrossRef]

28. Chen, X.M.; Zheng, H.; Ke, J.; Yang, H. Dynamic optimization strategies for on-demand ride services platform: Surge pricing, commission rate, and incentives. *Transp. Res. Part B Methodol.* **2020**, *138*, 23–45. [CrossRef]
29. Benjaafar, S.; Kong, G.; Li, X.; Courcoubetis, C. Peer-to-peer product sharing: Implications for ownership, usage, and social welfare in the sharing economy. *Manag. Sci.* **2019**, *65*, 477–493. [CrossRef]
30. Belleflamme, P.; Peitz, M. Platform competition: Who benefits from multihoming? *Int. J. Ind. Organ.* **2019**, *64*, 1–26. [CrossRef]
31. Taylor, T.A. On-demand service platforms. *Manuf. Serv. Oper. Manag.* **2018**, *20*, 704–720. [CrossRef]
32. Dong, J.; Ibrahim, R. Managing supply in the on-demand economy: Flexible workers, full-time employees, or both? *Oper. Res.* **2020**, *68*, 1238–1264. [CrossRef]
33. Ert, E.; Fleischer, A.; Magen, N. Trust and reputation in the sharing economy: The role of personal photos in Airbnb. *Tour. Manag.* **2016**, *55*, 62–73. [CrossRef]
34. Raju, S.; Rofin, T.; Kumar, S.P. Pricing strategies for dual-channel supply chain members under pandemic demand disruptions. *Oper. Res.* **2025**, *25*, 58. [CrossRef]
35. Thaichon, P.; Quach, S.; Barari, M.; Nguyen, M. Exploring the role of omnichannel retailing technologies: Future research directions. *Australas. Mark. J.* **2024**, *32*, 162–177. [CrossRef]
36. Radomska, J.; Kawa, A.; Hajdas, M.; Klimas, P.; Silva, S.C. Unveiling retail omnichannel challenges: Developing an omnichannel obstacles scale. *Int. J. Retail Distrib. Manag.* **2024**, *53*, 1–20. [CrossRef]
37. Cai, Y.J.; Lo, C.K. Omni-channel management in the new retailing era: A systematic review and future research agenda. *Int. J. Prod. Econ.* **2020**, *229*, 107729. [CrossRef]
38. Tan, Y. The first-ever e-commerce law: How will the law impact individuals and businesses. *Tsinghua China L. Rev.* **2022**, *11*, 427.
39. Cachon, G.P.; Lariviere, M.A. Supply chain coordination with revenue-sharing contracts: Strengths and limitations. *Manag. Sci.* **2005**, *51*, 30–44. [CrossRef]
40. Yang, H.; Luo, J.; Zhang, Q. Supplier encroachment under nonlinear pricing with imperfect substitutes: Bargaining power versus revenue-sharing. *Eur. J. Oper. Res.* **2018**, *267*, 1089–1101. [CrossRef]
41. Matsui, K. Asymmetric product distribution between symmetric manufacturers using dual-channel supply chains. *Eur. J. Oper. Res.* **2016**, *248*, 646–657. [CrossRef]
42. Huang, S.; Guan, X.; Chen, Y.J. Retailer information sharing with supplier encroachment. *Prod. Oper. Manag.* **2018**, *27*, 1133–1147. [CrossRef]
43. Shao, J.; Krishnan, H.; McCormick, S.T. Gray markets and supply chain incentives. *Prod. Oper. Manag.* **2016**, *25*, 1807–1819. [CrossRef]
44. Ahmadi, R.; Iravani, F.; Mamani, H. Coping with gray markets: The impact of market conditions and product characteristics. *Prod. Oper. Manag.* **2015**, *24*, 762–777. [CrossRef]
45. Hong, D.; Sheng, J.; Wu, Z.; Fan, J. Authorized retailer's gray market activities and supply chain performance improvement. *J. Ind. Manag. Optim.* **2025**, *21*, 7161–7185. [CrossRef]
46. Chiang, W.y.K.; Chhajed, D.; Hess, J.D. Direct marketing, indirect profits: A strategic analysis of dual-channel supply-chain design. *Manag. Sci.* **2003**, *49*, 1–20. [CrossRef]
47. Tsay, A.A.; Agrawal, N. Channel conflict and coordination in the e-commerce age. *Prod. Oper. Manag.* **2004**, *13*, 93–110. [CrossRef]
48. Taylor, T.A. Supply chain coordination under channel rebates with sales effort effects. *Manag. Sci.* **2002**, *48*, 992–1007. [CrossRef]
49. Krishnan, H.; Kapuscinski, R.; Butz, D.A. Coordinating contracts for decentralized supply chains with retailer promotional effort. *Manag. Sci.* **2004**, *50*, 48–63. [CrossRef]
50. Gurnani, H.; Erkoc, M. Supply contracts in manufacturer-retailer interactions with manufacturer-quality and retailer effort-induced demand. *Nav. Res. Logist. (NRL)* **2008**, *55*, 200–217. [CrossRef]
51. Ma, P.; Wang, H.; Shang, J. Supply chain channel strategies with quality and marketing effort-dependent demand. *Int. J. Prod. Econ.* **2013**, *144*, 572–581. . [CrossRef]
52. Tsay, A.A.; Agrawal, N. Channel dynamics under price and service competition. *Manuf. Serv. Oper. Manag.* **2000**, *2*, 372–391. [CrossRef]
53. Han, F.; Wang, M.; Wu, Z. Information sharing and channel structure in e-commerce supply chain considering data-driven marketing. *PLoS ONE* **2025**, *20*, e0328040. [CrossRef] [PubMed]

mathematics

MDPI

Article

Coordination Mechanism and Profit Distribution of Traceability Information Sharing in the Prefabricated Food Supply Chain

Jiayi Zhang [1], Xinyi Sang [2,*] and Huini Zhou [2]

[1] School of Management, Universiti Sains Malaysia, Penang 11800, Malaysia
[2] School of Business Administration, Yangzhou University, Yangzhou 225009, China
* Correspondence: mz120242189@stu.yzu.edu.cn

Abstract

Against the backdrop of the rapid growth in the scale of the prepared food market, safety issues have gradually become prominent. Establishing a traceability system has become crucial to safeguarding consumer rights and promoting the sustainable development of the industry, with traceability information sharing serving as the core link. However, affected by differences in interest demands and information asymmetry between manufacturers and retailers in the prepared food supply chain, there are obstacles to traceability information sharing. To explore the coordination mechanism of traceability information-sharing behavior in the prepared food supply chain under different decision-making models and its impact on profit distribution, this paper constructs a two-level supply chain model including manufacturers and retailers, comprehensively considers the online–offline dual-channel sales model, and distinguishes four scenarios: centralized decision-making, decentralized decision-making, retailer-led cost-sharing contract decision-making, and manufacturer-led cost-sharing contract decision-making. Using a differential game model, the equilibrium results under different decision-making models are discussed. The validity of the model is verified through fitting with empirical analysis and numerical example analysis. The research results show the following: (1) The centralized decision-making model has the best effect on increasing the market share of the prepared food supply chain, and although the cost-sharing contract model can improve it, there is still a gap. (2) The centralized decision-making model is not the one with the maximum profit, and manufacturer-led cost-sharing decision-making basically achieves Pareto optimality. The main reasons are the insufficient incentive mechanism, high coordination costs, and uneven profit distribution in centralized decision-making. (3) The impact of manufacturers' offline channel traceability information-sharing behavior on profits is more significant than that of online channels. (4) In a market environment with information asymmetry, the impact of goodwill on the profits of prepared foods is more prominent. This research provides a theoretical basis for the management of the prepared food supply chain, helps optimize the traceability information-sharing mechanism and profit distribution plan, and promotes the healthy development of the industry. (5) When the coefficient measuring the intensity of traceability information sharing's impact on product quality across manufacturers' online and offline channels increases, only under the retailer-led model does product quality and goodwill exhibit a fluctuating trend of "rising from the bottom to the second place and then falling back to the bottom," while the profits of all subjects increase simultaneously. (6) As the system attenuation coefficient increases, the evolution of product quality and goodwill under different cooperation models shows significant differences; in terms of profits, the profits of manufacturers' online channels increase over time, while those of other subjects decrease. (7) When the discount rate rises, the manufacturer-led model presents distinct characteristics: both the ranking and absolute value of product quality

decline synchronously, the ranking of goodwill falls, but its absolute value rises against the trend, the evolution of product quality and goodwill shows obvious model heterogeneity, and the profits of all subjects generally decrease.

Keywords: prefabricated food supply chain; traceability information sharing; coordination mechanism; profit distribution; differential game

MSC: 91A23; 91A25; 90B50

1. Introduction

In the rapid development process of the modern food industry, prefabricated food, with its obvious advantages of convenience, standardization, and diversification, meets the demand of consumers for high-efficiency and high-quality food under the current fast-paced life model and quickly wins the favor of the market and realizes the leapfrog expansion of scale. According to industry statistics, the market size of China's prefabricated food has climbed to 680 billion yuan in 2024, a record high. Based on the current market growth trend and consumption demand trend forecast, this scale will exceed the trillion-yuan mark by 2026, reaching 1072 billion yuan, fully demonstrating the strong development vitality and broad market prospect of the prefabricated food industry.

However, behind the rapid expansion of the industry scale, the short board of development brought by rapid expansion has also gradually surfaced, among which food safety problems, such as the imperfect quality control system of raw materials and frequent hygiene hidden dangers in the production and processing links, are particularly prominent [1–4]. At the upstream raw material supply end, there are problems such as unclear origin source, non-uniform testing standards for pesticide and veterinary drug residues, and chaotic batch management, while at the middle production and processing link, some enterprises are facing loopholes such as simplified process flow for cost reduction, substandard workshop hygiene conditions, and inadequate temperature control of cold chain storage and logistics. These problems not only pose direct potential risks to the health of consumers but also trigger a crisis of public trust in prefabricated food, and this becomes a key obstacle that restricts the industry from "scale expansion" to "quality improvement" and the realization of long-term healthy development. Therefore, the construction of a chain-wide food safety management system and the improvement of quality control standards covering raw material procurement, production and processing, warehousing and logistics, terminal sales, and other links have become important tasks to be solved urgently in the prefabricated food industry.

With its convenience and standardization advantages in catering scenes, the popularity of prefabricated dishes is increasing [5–7]. At present, many scholars have carried out research on the quality and safety of prefabricated food. Shen et al. revealed through a questionnaire survey system that nutrition balance, technology safety, and governance trust have adverse effects on consumers' perceived risk of safety of prepared food [4]. Based on the MOA theory and SEM model, Hou et al. explored the influencing factors of consumers' willingness to consume prefabricated food in China, and found that motivation, opportunity, and ability positively affected consumption willingness, and the latter two mediated motivation through convenience and health factors [8]. Guo et al. proposed to apply the 3D-printing technology to food for the production of prefabricated food to solve the problems of microbial contamination, poor nutrition quality, and product standardization [9]. He et al., based on two (risk media: new media/traditional media) $\times$ two (consumption

promotion situation: strong/weak) inter-subject investigation experiments, explored the shaping mechanism of risk media on consumers' prefabricated food safety risk perception and tested the mediating effect of food safety awareness and the moderating effect of promotion intensity. The results showed that media environment and promotion intensity jointly affect this risk perception [2].

Industrial processing has exacerbated the complexity of the food supply chain, and consumer concerns about food-related risks have significantly increased their focus on food traceability [10]. As an effective tool to ease information asymmetry and ensure food quality and safety, traceability information sharing, the core link of the traceability system, has been a research focus in the field of food supply chain management for a long time, and it has been a broad concern in academia. Obonyo et al. discussed the traceability information sharing problem under the ternary relationship of the food supply chain with the background of Kenya's dairy supply chain and found that traceability information sharing has a differentiated impact on the development of social capital [11]. Ersoy et al. believe that traceability information-sharing technology can help more efficient knowledge sharing practice and thus promote the improvement of an efficient circular food supply chain [12]. Christensen et al. discussed the multidimensional impact of product attributes, demand conditions, supply patterns, and planned environmental characteristics of food manufacturers on information-sharing technology in food supply chains [13]. Yu et al. constructed a food supply chain game model with suppliers and retailers to study the profit trade-off decision between suppliers' quality disclosure and non-disclosure, considering the effect of asymmetric demand information in the vertical direction of the food supply chain on suppliers' quality disclosure [14]. The implementation of information sharing and related application technologies is regarded as the biggest challenge for food supply chain safety incident prevention and control. Van Beusekom-Thoolen et al. proposed the implementation of information sharing and related application technologies on the ground, which is regarded as the biggest challenge for food supply chain safety incident prevention and control [15]. Liu et al. studied the optimal decision-making and coordination of the two-channel food supply chain under information symmetric and asymmetric scenarios by using Stackelberg game theory and backward induction. After designing the coordination contract, they found that the information asymmetry will aggravate the deviation of the optimal decision-making of the supply chain members, the information sharing is only beneficial to the manufacturer, and the improved revenue-sharing/cost-sharing contract is an effective coordination mechanism [16]. Aiming at the multiple challenges faced by the food supply chain, Gruzauskas et al. proposed an information-sharing strategy based on interconnected vehicle technology and simulated the scene through the distribution model to improve food quality and enhance the sustainability and resilience of the supply chain [17]. León-Bravo and others focus on the short food supply chain, explore the operation practice of geographical, relational, and information proximity and its sustainability relevance, and analyze the realization form of information sharing, its synergy with relational proximity, and the internal mechanism of the three to jointly promote the sustainable development of the supply chain after establishing the proximity upstream and downstream of the supply chain [18]. Sharma et al. believe that traceability of the food supply chain and information sharing for customers have a positive impact on visibility; visibility not only promotes the adoption and response speed of sustainable practices, but also positively affects supply chain performance. However, information sharing with customers had no significant impact on performance, and information sharing with suppliers had no significant correlation with visibility [19].

Profit distribution, as the core element of supply chain management research, is crucial to the stable development of the prefabricated food supply chain. When Huang et al.

studied information sharing in supply chain coordination, they quantitatively evaluated the impact of reliability and availability of information transmission on supply chain profits, which provided a theoretical basis for the correlation research between traceability information sharing and profit distribution in the prefabricated food supply chain [20]. Hong et al. analyzed the impact of information sharing on supply chain competition and supplier profit from the perspective of information-sharing types, which is helpful to understand the mechanism of different types of traceability information sharing on profit distribution in the prefabricated food supply chain [21]. Diao et al. paid attention to the impact of competition between the sharing platform of the agricultural supply chain and retailers on profit, which provided a reference for similar research on channel relationship and profit distribution in the prefabricated food supply chain [22]. There are also a few scholars who apply the Shapley value method to food supply chains; for example, Gopalakrishnan and Sankaranarayanan apply the Shapley value method to investigate the practical feasibility of identifying bilaterally implementable security cost-sharing arrangements in relevant alliances for pollution problems in food supply chains or data leakage problems in technology networks [23]. However, due to the fact that food sales involve both online and offline channels, scholars have made more use of the differential game method to analyze this. Li et al. focused on the fresh food online and offline dual-channel supply chain under the disturbance of consumers' quality preferences, constructed centralized and decentralized decision-making models, explored the flexible decision-making mechanism of price, quality, and output, and achieved supply chain coordination through revenue-sharing contracts. It was proven that considering the disturbance could improve the profit of members, centralized decision-making was more advantageous, and that there was a flexible decision-making interval [24]. Li et al. considered the online and offline demand function of the relationship between promotion and substitution, analyzed the optimal decision-making and parameter effects of promotion intensity and online discount under three decision-making models of platform operation, and found that the centralized model had the highest supply chain profit, and the relationship between optimal promotion intensity and discount was regulated by a substitution coefficient and an online promotion influence coefficient [25]. Guo and others focused on the challenges of dual-channel operation of food retailing, based on the utility model, analyzed the optimal order volume and price discount of retailers under different transportation policies, revealed the impact of factors such as freight on pricing, and found that low freight rates can achieve a "win–win" situation and high freight rates can lead to target conflicts, while reasonable allocation of safety stocks and adjustment of transportation policies can improve efficiency, reduce waste, and enhance supply chain resilience [26]. Lin and Januardi combine multiple logistic regression, the logit model, and the Stackelberg competition bi-level programming model to analyze the customer channel preference, willingness to pay, and the pricing mechanism under competition in the dual sales channel system. It is found that channel price has a non-linear impact on both, and traditional retailers as leaders have higher profit [27]. A systematic comparison of the above-mentioned studies is summarized in Table 1.

Table 1. Summary of the related research.

	Food Quality and Safety	Traceability Information Sharing	Distribution of Profits	Differential Game
Shen et al. [4]	√			
Hou et al. [8]	√			
Guo et al. [9]	√			
He et al. [2]	√			
Obonyo et al. [11]		√		
Ersoy et al. [12]		√		
Christensen et al. [13]		√		
Yu et al. [14]		√		√

Table 1. *Cont.*

	Food Quality and Safety	Traceability Information Sharing	Distribution of Profits	Differential Game
Van Beusekom-Thoolen et al. [15]	√	√		√
Liu et al. [16]		√		√
Gruzauskas et al. [17]	√	√		
León-Bravo et al. [18]		√		
Sharma et al. [19]		√		
Huang et al. [20]			√	
Diao et al. [22]			√	
Gopalakrishnan and Sankaranarayanan [23]	√		√	
Li et al. [24]				√
Li et al. [25]				√
Guo et al. [26]				√
Lin and Januardi [27]				√
Thin paper	√	√	√	√

To sum it up, the attention paid by the existing research to the specific field of prefabricated food is still relatively insufficient, and the related discussions are mostly focused on basic aspects, such as technical improvement, method innovation, and identification of influencing factors. At the same time, although many scholars have carried out cross-method research under multiple scenarios for the influencing factors, action mechanism, practical challenges, and optimization path of information sharing in food supply chains, few have systematically and deeply analyzed the internal coordination mechanism and practical implementation path of traceability information-sharing behavior in supply chain coordination from the dual perspectives of supply chain participants, such as suppliers, manufacturers, retailers, and platform parties, and the overall performance optimization of the supply chain.

In view of this, this research takes the prefabricated food supply chain as the research object, based on the core orientation of "constructing a traceability system to crack the pain points of quality and safety"; the traceability system can realize the whole process of the traceability of product information and provide key support to protecting the rights and interests of consumers, enhancing the enterprise's crisis-response ability and the industry's overall reputation—focusing on the core of the traceability information sharing system, aiming at the problems of sharing obstacles and system efficiency limitations caused by the difference in interest demands and information asymmetry between manufacturers and retailers, and aiming at optimizing the sharing coordination mechanism and profit distribution scheme. Research and build a two-level supply chain model consisting of manufacturers (responsible for production, core traceability information sharing, and dual-channel operation functions) and retailers (responsible for sales and terminal traceability information sharing functions), simultaneously incorporate online and offline dual-channel sales models, set both parties as rational economic entities and form four decision scenarios around traceability information sharing, and clarify that information sharing affects market demand and revenue distribution patterns through its effect on product quality and brand goodwill. In terms of research methodology, a dynamic differential game analysis model is constructed, and the Hamilton–Jacobi–Bellman (HJB) equation, which can effectively depict the dynamic optimization problem under the continuous time dimension and accurately capture the interactive impact of goodwill accumulation and product quality, the coordination mechanism, and profit distribution logic of traceability information-sharing behavior, are systematically analyzed, and at the same time, the Shapley value method is innovatively integrated to scientifically and reasonably distribute the income of supply chain participants after forming a strategic alliance, which provides a theoretical support for the income distribution of food supply chain. On this basis, the study further focuses on the online and offline dual-channel scenarios, exploring the impact of the four decision-making modes on traceability sharing degree, market share, and profit distribution results,

as well as the difference in profit effect between manufacturers' channel sharing behavior and goodwill accumulation.

The contribution of this research is mainly reflected in three aspects:

1. Integrate the "multi-agent game" with the "online and offline dual channels" scenario of the pre-cooked food, anchor its "short-term protection and long-term chain" characteristics and traceability pain points, embed specific parameters such as freshness traceability weights, and construct a customized game model to solve the problem of insufficient pertinence of pan-food research to the pre-cooked food category.
2. Establish a coupling model of "dual-channel development and evolution-dynamic fluctuation of information asymmetry-real-time adaptation of coordination mechanism", quantify the transmission path of channel structure adjustment to traceability-sharing intention, design an incentive coefficient mechanism dynamically optimized with the proportion of channels, and solve the core problem of a lack of flexibility in the traditional mechanism.
3. Build a four-order conduction model of "traceability behavior-quality signal-trust accumulation-goodwill premium", prove that traceability sharing can be converted into a quality signal to form goodwill premium, and clarify the value-added path from traceability to profit.

2. Basic Assumptions

This study focuses on the coordination mechanism of information-sharing behavior within the traceability system of the prefabricated food supply chain, specifically between manufacturers and retailers. In this system, the manufacturer (denoted as m) is responsible for sharing traceability information, such as enterprise certification, production process details, logistics and warehousing data, and quality inspection reports. The manufacturer operates through traditional offline channels—supplying products to retailers—and online direct-to-consumer channels. The retailer (denoted as s), in turn, shares traceability information, such as product acceptance data and consumer feedback. The logical framework of this information-sharing coordination mechanism is depicted in Figure 1.

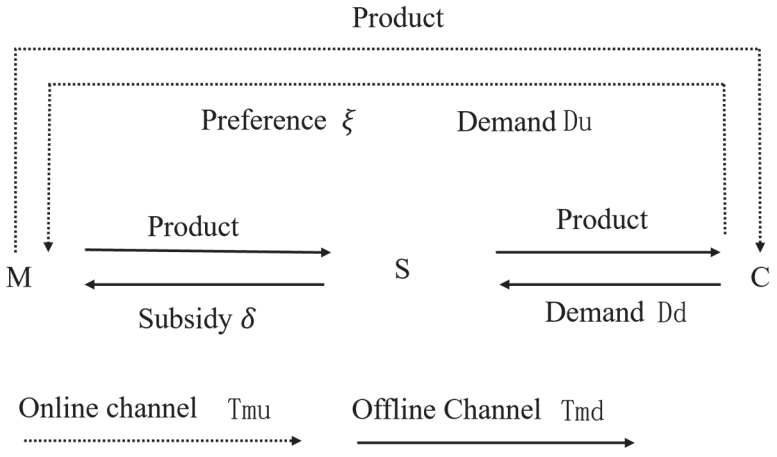


Figure 1. Logical framework of the coordination mechanism of traceability information-sharing behavior in the pre-prepared food supply chain.

The notation used throughout the paper is defined in Table 2. This study establishes a two-level supply chain model composed of manufacturers and retailers, in which the manufacturer adopts online and offline sales channels. At any time t, the manufacturer's traceability information-sharing behavior is divided into online and offline components. All members of the supply chain are regarded as rational economic entities, each having

full knowledge of the other's information-sharing costs and profit-related parameters. According to the literature [28–30], the cost functions of the traceable food manufacturer through online and offline channels, as well as those of the retailer, can be represented as follows: $C(Tm_u) = \frac{\gamma_1}{2} {Tm_u}^2$, $C(Tm_d) = \frac{\gamma_2}{2} {Tm_d}^2$, and $C(Ts) = \frac{\gamma_3}{2} Ts^2$, where $\gamma_i (i = 1, 2, 3)$ is a positive constant and serves as the key coefficient influencing the cost of information sharing among supply chain members.

Table 2. Definition of symbols.

γ_i $(i = 1, 2, 3)$	Coefficient of the cost impact of traceability information sharing.
Tm_u	Traceability information-sharing behavior of the manufacturer in the online channel.
Tm_d	Traceability information-sharing behavior of the manufacturer in the offline channel.
Ts	Traceability information-sharing behavior of the traceable food retailer.
η	Impact coefficient of the manufacturer's online channel traceability information sharing on traceable food quality.
β	Impact coefficient of the manufacturer's offline channel traceability information sharing on traceable food quality.
ε	Quality decay coefficient of traceable food.
$Q(t)$	Quality level of pre-made food at time t.
$G(t)$	Reputation level of pre-made food at time t.
ϕ	Impact coefficient of traceable food quality on reputation.
θ	Impact coefficient of retailer's traceability information-sharing behavior on the reputation of traceable food.
φ	Reputation decay coefficient of traceable food.
$D(t)$	Demand for pre-made food at time t.
α	Initial market demand.
ξ	Consumer preference coefficient for the online channel of pre-made food.
μ	Impact coefficient of food quality on demand for traceable food.
κ	Impact coefficient of reputation on demand for traceable food.
π_u^m, π_d^m, π_s	Unit marginal profits of the manufacturer's online channel, offline channel, and the retailer.
ρ	Discount rate.
δ	Subsidy provided by the retailer to the manufacturer.
ω	Subsidy provided by the manufacturer to the retailer.

On the one hand, the manufacturer's traceability information-sharing behavior directly affects product quality. The implementation of the traceability system ensures that prefabricated food complies with national standards during the transportation process [31–33]. As a dynamic variable $\dot{Q}(t)$, the rate of change in product quality over time can be expressed as follows:

$$\dot{Q}(t) = \eta Tm_u(t) + \beta Tm_d(t) - \varepsilon Q(t) \quad (1)$$

In Equation (1), Tmu and Tmd represent the manufacturer's traceability information-sharing behaviors in the online and offline channels, respectively. η denotes the impact coefficient of the manufacturer's online information-sharing behavior on product quality, whereas α_2 reflects the impact coefficient of the offline information-sharing behavior on product quality. ε represents the decay coefficient of product quality when the manufacturer does not engage in traceability information sharing, due to factors such as cost reduction and insufficient quality control. $Q(0) = Q_0$ defines the initial level of product quality.

On the other hand, the retailer's traceability information-sharing behavior and product quality influence the goodwill of prefabricated food. Based on a modification of the classic Nerlove–Arrow goodwill model [34], the dynamic evolution of goodwill can be expressed as follows:

$$\dot{G}(t) = \phi Q(t) + \theta Ts(t) - \varphi G(t) \quad (2)$$

In Equation (2), $G(t)$ denotes the goodwill of prefabricated food at time t; ϕ represents the coefficient reflecting the effect of product quality on goodwill; *Ts* denotes the retailer's traceability information-sharing behavior; θ indicates the coefficient representing the effect of the retailer's traceability information-sharing behavior on goodwill; φ represents the decay coefficient of goodwill or the effect of consumer forgetfulness on goodwill [35]; and $G(0) = G_0$ denotes the initial level of goodwill.

Given that the demand for prefabricated food is linearly related to its quality and goodwill, total market demand consists of three components: the baseline market demand, product quality, and goodwill. According to the literature [36], product quality and goodwill influence consumers' purchasing intentions. When the manufacturer and retailer engage in traceability information sharing, the demand functions for the online and offline channels can be expressed as follows:

$$D_u(t) = \xi\alpha + \mu Q(t) + \kappa G(t) \tag{3}$$

$$D_d(t) = (1-\xi)\alpha + \mu Q(t) + \kappa G(t) \tag{4}$$

In Equations (3) and (4), α represents the initial market demand; ξ denotes the consumer preference coefficient for online channels; μ reflects the degree to which product quality influences the demand for traceable food; κ represents the effect of goodwill on the demand for traceable food; and π_u^m, π_d^m, and π^s denote the unit marginal revenues of the manufacturer's online and offline channels and of the retailer, respectively. In addition, the manufacturer and the retailer adopt the same discount rate ρ for evaluation.

3. Model Construction and Solution

3.1. Centralized Decision-Making Model

Under the centralized decision-making model (denoted by subscript 1), the manufacturer and retailer of prefabricated food collaborate with the common objective of maximizing the overall profit of the entire supply chain. Both parties jointly determine their respective traceability information-sharing strategies. When evaluating the total profit over an infinite time horizon, a uniform discount rate ρ is applied for discounting. On this basis, the total profit of the food supply chain can be expressed as follows

$$R = \int_0^\infty e^{-\rho t}\left\{\begin{array}{l}\pi_u^m[\xi\alpha + \mu Q(t) + \kappa G(t)] + \pi_d^m[(1-\xi)\alpha + \mu Q(t) + \kappa G(t)] \\ +\pi^s[(1-\xi)\alpha + \mu Q(t) + \kappa G(t)] - \frac{\gamma_1}{2}Tm_{u1}{}^2 - \frac{\gamma_2}{2}Tm_{d1}{}^2 - \frac{\gamma_3}{2}Ts_1{}^2\end{array}\right\}dt \tag{5}$$

According to the Hamilton–Jacobi–Bellman (HJB) equation and the backward induction method, the equilibrium decisions of supply chain members are presented in Proposition 1.

Proposition 1. *The equilibrium results under the centralized decision-making scenario are as follows:*

(1) The optimal trajectories of product quality $Q(t)$ and goodwill $G(t)$ are given by the following:

$$Q_1(t) = \left\{\begin{array}{l}\frac{(\pi_u^m+\pi_d^m+\pi^s)(\eta^2\gamma_2+\beta^2\gamma_1)\mu}{\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)} + \frac{(\pi_u^m+\pi_d^m+\pi^s)(\eta^2\gamma_2+\beta^2\gamma_1)\phi\kappa}{\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)(\rho+\varphi)} \\ +\left\{Q_0 - \frac{(\pi_u^m+\pi_d^m+\pi^s)(\eta^2\gamma_2+\beta^2\gamma_1)\mu}{\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)} - \frac{(\pi_u^m+\pi_d^m+\pi^s)(\eta^2\gamma_2+\beta^2\gamma_1)\phi\kappa}{\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)(\rho+\varphi)}\right\}e^{-\varepsilon t}\end{array}\right. \tag{6}$$

$$G_1(t) = \left\{\begin{array}{l}\frac{(\pi_u^m+\pi_d^m+\pi^s)(\eta^2\gamma_2+\beta^2\gamma_1)\mu\phi}{\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)} + \frac{(\pi_u^m+\pi_d^m+\pi^s)(\eta^2\gamma_2+\beta^2\gamma_1)\phi^2\kappa}{\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)(\rho+\varphi)} \\ +\frac{(\pi_u^m+\pi_d^m+\pi^s)\theta^2\kappa}{\varphi\gamma_3(\rho+\varphi)} + \left\{G_0 - \frac{(\pi_u^m+\pi_d^m+\pi^s)(\eta^2\gamma_2+\beta^2\gamma_1)\mu\phi}{\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)}\right. \\ \left. -\frac{(\pi_u^m+\pi_d^m+\pi^s)(\eta^2\gamma_2+\beta^2\gamma_1)\phi^2\kappa}{\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)(\rho+\varphi)} - \frac{(\pi_u^m+\pi_d^m+\pi^s)\theta^2\kappa}{\varphi\gamma_3(\rho+\varphi)}\right\}e^{-\varphi t}\end{array}\right. \tag{7}$$

(2) The optimal value function of the food supply chain is given as follows:

$$V_1 = \begin{cases} [\frac{(\pi_u^m+\pi_d^m+\pi^s)\mu}{\rho+\varepsilon} + \frac{(\pi_u^m+\pi_d^m+\pi^s)\phi\kappa}{(\rho+\varepsilon)(\rho+\varphi)}]Q_1(t) + [\frac{(\pi_u^m+\pi_d^m+\pi^s)\kappa}{\rho+\varphi}]G_1(t) \\ +\frac{[\pi_u^m\xi+\pi_d^m-\pi_d^m\xi+\pi^s-\pi^s\xi]\alpha}{\rho} + \frac{\eta^2+\beta^2}{2\rho(\gamma_1+\gamma_2)}[\frac{(\pi_u^m+\pi_d^m+\pi^s)\mu}{\rho+\varepsilon} + \frac{(\pi_u^m+\pi_d^m+\pi^s)\phi\kappa}{(\rho+\varepsilon)(\rho+\varphi)}]^2 \\ +\frac{\theta^2}{2\rho\gamma_3}[\frac{(\pi_u^m+\pi_d^m+\pi^s)\kappa}{\rho+\varphi}]^2 \end{cases} \tag{8}$$

The detailed proof is provided in Appendix A.

Under the centralized decision-making framework, the manufacturer and the retailer of prefabricated food pursue the maximization of total supply chain profit as a common objective, jointly formulating their respective traceability information-sharing strategies. This approach emphasizes cooperation to optimize the overall economic efficiency of the supply chain, ensuring alignment of actions among all parties to achieve optimal market performance. The centralized decision-making model effectively coordinates the interests of all supply chain participants, promotes optimal resource allocation, and enhances operational stability and synergy within the supply chain, thereby reducing the risks associated with information asymmetry.

3.2. Decentralized Decision-Making Model

Under the decentralized decision-making model (denoted by subscript 2), the manufacturer and retailer of prefabricated food operate on an equal footing in the market, each independently bearing the costs incurred in the traceability information-sharing process. Both parties aim to maximize their own profits and make independent decisions regarding their respective information-sharing behaviors. On the basis of this assumption, the objective functions of the two parties can be defined as follows:

$$R(Tm_{u2}) = \int_0^\infty e^{-\rho t}\{\pi_u^m(\xi\alpha + \mu Q(t) + \kappa G(t)) - \frac{\gamma_1}{2}{Tm_{u2}}^2(t)\Big\}dt \tag{9}$$

$$R(Tm_{d2}) = \int_0^\infty e^{-pt}\Big\{\pi_d^m[(1-\xi)\alpha + \mu Q(t) + \kappa G(t)] - \frac{\gamma_2}{2}{Tm_{d2}}^2(t)\Big\}dt \tag{10}$$

$$R(Ts_2) = \int_0^\infty e^{-\rho t}\Big\{\pi^s[(1-\xi)\alpha + \mu Q(t) + \kappa G(t)] - \frac{\gamma_3}{2}{Ts_2}^2(t)\Big\}dt \tag{11}$$

Proposition 2. *The equilibrium results under the decentralized decision-making scenario are as follows:*
(1) The optimal trajectories of product quality $Q(t)$ and goodwill $G(t)$ are given by the following:

$$Q_2(t) = \begin{cases} [\frac{(\pi_u^m\eta^2\gamma_2+\pi_d^m\beta^2\gamma_1)\mu}{\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)} + \frac{(\pi_u^m\eta^2\gamma_2+\pi_d^m\beta^2\gamma_1)\kappa\phi}{\varepsilon\gamma_1\gamma(\rho+\varepsilon)(\rho+\varphi)}] \\ +\Big\{Q_0 - \frac{(\pi_u^m\eta^2\gamma_2+\pi_d^m\beta^2\gamma_1)\mu}{\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)} - \frac{(\pi_u^m\eta^2\gamma_2+\pi_d^m\beta^2\gamma_1)\kappa\phi}{\varepsilon\gamma_1\gamma(\rho+\varepsilon)(\rho+\varphi)}\Big\}e^{-\varepsilon t} \end{cases} \tag{12}$$

$$G_2(t) = \begin{cases} \frac{(\pi_u^m\eta^2\gamma_2+\pi_d^m\beta^2\gamma_1)\phi\mu}{\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)} + \frac{(\pi_u^m\eta^2\gamma_2+\pi_d^m\beta^2\gamma_1)\kappa\phi^2}{\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)(\rho+\varphi)} + \frac{\theta^2\pi^s\kappa}{\varphi\gamma_3(\rho+\varphi)} \\ +\Big\{G_0 - \frac{(\pi_u^m\eta^2\gamma_2+\pi_d^m\beta^2\gamma_1)\phi\mu}{\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)} - \frac{(\pi_u^m\eta^2\gamma_2+\pi_d^m\beta^2\gamma_1)\kappa\phi^2}{\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)(\rho+\varphi)} - \frac{\theta^2\pi^s\kappa}{\varphi\gamma_3(\rho+\varphi)}\Big\}e^{-\varphi t} \end{cases} \tag{13}$$

(2) The optimal value functions of the manufacturer's online and offline channels, of the retailer, and of the overall food supply chain are given as follows:

$$Vm_{u2} = \begin{cases} [\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa \phi}{(\rho+\varepsilon)(\rho+\varphi)}]Q_2(t) + \frac{\pi_u^m \kappa}{\rho+\varphi} G_2(t) + \frac{\eta^2}{2\gamma_1}[\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa \phi}{(\rho+\varepsilon)(\rho+\varphi)}]^2 \\ +\frac{\beta^2}{\gamma_2}[\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa \phi}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{\pi_d^m \mu}{\rho+\varepsilon} + \frac{\pi_d^m \kappa \phi}{(\rho+\varepsilon)(\rho+\varphi)}] + \frac{\theta^2 \pi_u^m \pi^s \kappa^2}{\gamma_3(\rho+\varphi)^2} + \pi_u^m \xi \alpha \end{cases} \tag{14}$$

$$Vm_{d2} = \begin{cases} [\frac{\pi_d^m \mu}{\rho+\varepsilon} + \frac{\pi_d^m \kappa \phi}{(\rho+\varepsilon)(\rho+\varphi)}]Q_2(t) + \frac{\pi_d^m \kappa}{\rho+\varphi} G_2(t) + \frac{\theta^2 \pi_d^m \pi^s \kappa^2}{\gamma_3(\rho+\varphi)^2} + \pi_d^m (1-\xi)\alpha \\ +\frac{\eta^2}{\gamma_1}[\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa \phi}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{\pi_d^m \mu}{\rho+\varepsilon} + \frac{\pi_d^m \kappa \phi}{(\rho+\varepsilon)(\rho+\varphi)}] + \frac{\beta^2}{2\gamma_2}[\frac{\pi_d^m \mu}{\rho+\varepsilon} + \frac{\pi_d^m \kappa \phi}{(\rho+\varepsilon)(\rho+\varphi)}]^2 \end{cases} \tag{15}$$

$$Vs_2 = \begin{cases} [\frac{\pi^s \mu}{\rho+\varepsilon} + \frac{\pi^s \kappa \phi}{(\rho+\varepsilon)(\rho+\varphi)}]Q_2(t) + \frac{\pi^s \kappa}{\rho+\varphi} G(t) + \frac{1}{2}(\frac{\pi^s \kappa}{\rho+\varphi})^2 + \pi^s (1-\xi)\alpha \\ +\frac{\eta^2}{\gamma_1}[\frac{\pi^s \mu}{\rho+\varepsilon} + \frac{\pi^s \kappa \phi}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa \phi}{(\rho+\varepsilon)(\rho+\varphi)}] \\ +\frac{\beta^2}{\gamma_2}[\frac{\pi_d^m \mu}{\rho+\varepsilon} + \frac{\pi_d^m \kappa \phi}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{\pi^s \mu}{\rho+\varepsilon} + \frac{\pi^s \kappa \phi}{(\rho+\varepsilon)(\rho+\varphi)}] \end{cases} \tag{16}$$

$$V_2 = \begin{cases} [\frac{(\pi_u^m + \pi_d^m + \pi^s)\mu}{\rho+\varepsilon} + \frac{(\pi_u^m + \pi_d^m + \pi^s)\phi\kappa}{(\rho+\varepsilon)(\rho+\varphi)}]Q_2(t) + \frac{(\pi_u^m + \pi_d^m + \pi^s)\kappa}{\rho+\varphi} G_2(t) \\ +\frac{\eta^2}{2\gamma_1}[\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa \phi}{(\rho+\varepsilon)(\rho+\varphi)}]^2 + (\frac{\beta^2}{\gamma_2} + \frac{\eta^2}{\gamma_1})[\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa \phi}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{\pi_d^m \mu}{\rho+\varepsilon} + \frac{\pi_d^m \kappa \phi}{(\rho+\varepsilon)(\rho+\varphi)}] \\ +\frac{\beta^2}{2\gamma_2}[\frac{\pi_d^m \mu}{\rho+\varepsilon} + \frac{\pi_d^m \kappa \phi}{(\rho+\varepsilon)(\rho+\varphi)}]^2 + \frac{\eta^2}{\gamma_1}[\frac{\pi^s \mu}{\rho+\varepsilon} + \frac{\pi^s \kappa \phi}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa \phi}{(\rho+\varepsilon)(\rho+\varphi)}] \\ +\frac{\beta^2}{\gamma_2}[\frac{\pi_d^m \mu}{\rho+\varepsilon} + \frac{\pi_d^m \kappa \phi}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{\pi^s \mu}{\rho+\varepsilon} + \frac{\pi^s \kappa \phi}{(\rho+\varepsilon)(\rho+\varphi)}] + \frac{1}{2}(\frac{\pi^s \kappa}{\rho+\varphi})^2 + \frac{(\pi_u^m + \pi_d^m)\pi^s \theta^2 \kappa^2}{\gamma_3(\rho+\varphi)^2} \\ +\pi_u^m \xi \alpha + \pi_d^m (1-\xi)\alpha + \pi^s (1-\xi)\alpha \end{cases} \tag{17}$$

The detailed proof is provided in Appendix A.

Under the decentralized decision-making scenario, the manufacturer and the retailer of prefabricated food independently determine their traceability information-sharing decisions, with the goal of maximizing their own profits. This approach emphasizes that each entity operates based on its individual economic interests, ensuring autonomy and specificity in the decision-making process. However, in this model, information asymmetry and conflicts of interest between the manufacturer and the retailer become more pronounced, which may negatively affect the coordination and operational efficiency of the supply chain. Moreover, as both parties operate independently, their incentives to engage in traceability information sharing are relatively limited, resulting in overall supply chain profits that are generally lower than those achieved under centralized decision-making.

3.3. Retailer-Led Cost-Sharing Model

In the retailer-led cost-sharing model (denoted by subscript 3), as a core participant in the supply chain, the prefabricated food retailer provides a certain proportion of subsidy to the manufacturer to encourage traceability information sharing. In this case, both parties adopt a static feedback Stackelberg equilibrium strategy. According to the principle of profit maximization, the retailer initially determines its own traceability information-sharing strategy and undertakes part of the manufacturer's cost for offline traceability information sharing as a subsidy δ. Subsequently, the manufacturer decides its own information-sharing strategy based on the retailer's sharing behavior and the level of subsidy provided. Under this setting, the objective function of the manufacturer can be expressed as follows:

$$R(Tm_{u3}) = \int_0^\infty e^{-\rho t}\left\{\pi_u^m[\xi\alpha + \mu Q(t) + \kappa G(t)] - \frac{\gamma_1}{2} {Tm_{u3}}^2\right\} dt \tag{18}$$

$$R(Tm_{d3}) = \int_0^\infty e^{-pt}\left\{\pi_d^m[(1-\xi)\alpha + \mu Q(t) + \kappa G(t)] - \frac{\gamma_2}{2}(1-\delta){Tm_{d3}}^2\right\} dt \tag{19}$$

The objective function of the pre-prepared food retailer is given as follows:

$$R(Ts_3) = \int_0^{\infty} e^{-\rho t}\left\{\pi^s[(1-\xi)\alpha + \mu Q(t) + \kappa G(t)] - \frac{\gamma_3}{2}Ts_3{}^2(t) - \frac{\delta\gamma_2}{2}Tm_{d3}{}^2(t)\right\}dt \quad (20)$$

Proposition 3. *The equilibrium results under the retailer-led cost-sharing contract scenario are as follows:*

(1) The optimal trajectories of product quality $Q(t)$ and goodwill $G(t)$ are given by the following:

$$Q_3(t) = \left\{\begin{array}{l} \frac{[2\pi_u^m\eta^2\gamma_2+(2\pi^s+\pi_d^m)\beta^2\gamma_1]\mu}{2\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)} + \frac{[2\pi_u^m\eta^2\gamma_2+(2\pi^s+\pi_d^m)\beta^2\gamma_1]\kappa}{2\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)(\rho+\varphi)} \\ +\left\{Q_0 - \frac{[2\pi_u^m\eta^2\gamma_2+(2\pi^s+\pi_d^m)\beta^2\gamma_1]\mu}{2\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)} + \frac{[2\pi_u^m\eta^2\gamma_2+(2\pi^s+\pi_d^m)\beta^2\gamma_1]\kappa}{2\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)(\rho+\varphi)}\right\}e^{-\varepsilon t} \end{array} \quad (21)$$

$$G_3(t) = \left\{\begin{array}{l} \frac{[2\pi_u^m\eta^2\gamma_2+(2\pi^s+\pi_d^m)\beta^2\gamma_1]\phi\mu}{2\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)} + \frac{[2\pi_u^m\eta^2\gamma_2+(2\pi^s+\pi_d^m)\beta^2\gamma_1]\phi\kappa}{2\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)(\rho+\varphi)} + \frac{\theta^2\pi^s\kappa}{\varphi\gamma_3(\rho+\varphi)} \\ +\left\{G_0 - \frac{[2\pi_u^m\eta^2\gamma_2+(2\pi^s+\pi_d^m)\beta^2\gamma_1]\phi\mu}{2\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)} - \frac{\theta^2\pi^s\kappa}{\varphi\gamma_3(\rho+\varphi)}\right. \\ \left. -\frac{[2\pi_u^m\eta^2\gamma_2+(2\pi^s+\pi_d^m)\beta^2\gamma_1]\phi\kappa}{2\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)(\rho+\varphi)}\right\}e^{-\varphi t} \end{array} \quad (22)$$

(2) The optimal value functions of the manufacturer's online channel, offline channel, the retailer, and the overall food supply chain are given as follows:

$$Vm_{u3} = \left\{\begin{array}{l} [\frac{\pi_u^m\mu}{\rho+\varepsilon} + \frac{\pi_u^m\kappa}{(\rho+\varepsilon)(\rho+\varphi)}]Q(t) + \frac{\pi_u^m\kappa}{\rho+\varphi}G(t) \\ +\frac{\pi_u^m\xi\alpha}{\rho} + \frac{\eta^2}{2\rho\gamma_1}[\frac{\pi_u^m\mu}{\rho+\varepsilon} + \frac{\pi_u^m\kappa}{(\rho+\varepsilon)(\rho+\varphi)}] \\ +\frac{\theta^2\pi_u^m\pi^s\kappa^2}{\rho\gamma_3(\rho+\varphi)^2} + \frac{\beta^2}{2\rho\gamma_2}[\frac{\pi_u^m\mu}{\rho+\varepsilon} + \frac{\pi_u^m\kappa}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{(2\pi^s+\pi_d^m)\mu}{\rho+\varepsilon} + \frac{(2\pi^s+\pi_d^m)\kappa}{(\rho+\varepsilon)(\rho+\varphi)}] \end{array} \quad (23)$$

$$Vm_{d3} = \left\{\begin{array}{l} [\frac{\pi_d^m\mu}{\rho+\varepsilon} + \frac{\pi_d^m\kappa}{(\rho+\varepsilon)(\rho+\varphi)}]Q(t) + \frac{\pi_d^m\kappa}{\rho+\varphi}G(t) + \frac{\pi_d^m(1-\xi)\alpha}{\rho} \\ +\frac{\eta^2}{\rho\gamma_1}[\frac{\pi_u^m\mu}{\rho+\varepsilon} + \frac{\pi_u^m\kappa}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{\pi_d^m\mu}{\rho+\varepsilon} + \frac{\pi_d^m\kappa}{(\rho+\varepsilon)(\rho+\varphi)}] \\ +\frac{\beta^2}{4\rho\gamma_2}[\frac{\pi_d^m\mu}{\rho+\varepsilon} + \frac{\pi_d^m\kappa}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{(2\pi^s+\pi_d^m)\mu}{\rho+\varepsilon} + \frac{(2\pi^s+\pi_d^m)\kappa}{(\rho+\varepsilon)(\rho+\varphi)}] + \frac{\theta^2\pi_d^m\pi^s\kappa^2}{\rho\gamma_3(\rho+\varphi)^2} \end{array} \quad (24)$$

$$Vs_3 = \left\{\begin{array}{l} [\frac{\pi^s\mu}{\rho+\varepsilon} + \frac{\pi^s\kappa}{(\rho+\varepsilon)(\rho+\varphi)}]Q(t) + \frac{\pi^s\kappa}{\rho+\varphi}G(t) + \frac{\pi^s(1-\xi)\alpha}{\rho} \\ +\frac{\eta^2}{\rho\gamma_1}[\frac{\pi_u^m\mu}{\rho+\varepsilon} + \frac{\pi_u^m\kappa}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{\pi^s\mu}{\rho+\varepsilon} + \frac{\pi^s\kappa}{(\rho+\varepsilon)(\rho+\varphi)}] \\ +\frac{\beta^2}{8\rho\gamma_2}[\frac{(2\pi^s-\pi_d^m)\mu}{\rho+\varepsilon} + \frac{(2\pi^s-\pi_d^m)\kappa}{(\rho+\varepsilon)(\rho+\varphi)}] + \frac{\theta^2\pi^s{}^2\kappa^2}{2\rho\gamma_3(\rho+\varphi)^2} \end{array} \quad (25)$$

$$V_3 = \left\{\begin{array}{l} [\frac{(\pi_u^m+\pi_d^m+\pi^s)\mu}{\rho+\varepsilon} + \frac{(\pi_u^m+\pi_d^m+\pi^s)\kappa}{(\rho+\varepsilon)(\rho+\varphi)}]Q_3(t) + \frac{(\pi_u^m+\pi_d^m+\pi^s)\kappa}{\rho+\varphi}G_3(t) \\ +\frac{\alpha}{\rho}[\pi_u^m\xi + \pi_d^m(1-\xi) + \pi^s(1-\xi)] + \frac{\beta^2}{8\rho\gamma_2}[\frac{(2\pi^s-\pi_d^m)\mu}{\rho+\varepsilon} + \frac{(2\pi^s-\pi_d^m)\kappa}{(\rho+\varepsilon)(\rho+\varphi)}] \\ +\frac{\eta^2}{\rho\gamma_1}[\frac{\pi_u^m\mu}{\rho+\varepsilon} + \frac{\pi_u^m\kappa}{(\rho+\varepsilon)(\rho+\varphi)}][1 + \frac{(2\pi_d^m+2\pi^s)\mu}{\rho+\varepsilon} + \frac{(2\pi_d^m+2\pi^s)\kappa}{(\rho+\varepsilon)(\rho+\varphi)}] \\ +\frac{\theta^2}{\rho\gamma_3(\rho+\varphi)^2}(\pi_u^m\pi^s\kappa^2 + \pi_d^m\pi^s\kappa^2 + \frac{\pi^{s2}\kappa^2}{2}) \\ +\frac{3\beta^2}{4\rho\gamma_2}[\frac{\pi_u^m\mu}{\rho+\varepsilon} + \frac{\pi_u^m\kappa}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{(2\pi^s+\pi_d^m)\mu}{\rho+\varepsilon} + \frac{(2\pi^s+\pi_d^m)\kappa}{(\rho+\varepsilon)(\rho+\varphi)}] \end{array} \quad (26)$$

The detailed proof is provided in Appendix A.

Under the retailer-led cost-sharing contract scenario, the retailer provides subsidies to the manufacturer to incentivize traceability information sharing, thereby altering the manufacturer's cost–benefit structure. Compared with the decentralized model, the manufacturer's enthusiasm for traceability information sharing through online channels is significantly improved, motivating greater investment in information sharing. Consequently, this model partially alleviates the shortcomings of decentralized decision-making and promotes the manufacturer's participation in information sharing. However, for the

prefabricated food retailer, its traceability information-sharing behavior does not show improvement. This indicates that although the model positively affects the manufacturer, it remains insufficient for enhancing the retailer's motivation to share information. Therefore, the overall synergy of the supply chain is not yet fully realized, and further optimization of strategies is required to improve overall supply chain performance.

3.4. Manufacturer-Led Cost-Sharing Model

In the manufacturer-led cost-sharing model (denoted by subscript 4), as the core enterprise in the supply chain, the prefabricated food manufacturer provides a certain proportion of subsidy to the retailer during food safety incidents to encourage traceability information sharing. In this case, both parties adopt a static feedback Stackelberg equilibrium strategy. According to the principle of profit maximization, the manufacturer initially determines its own traceability information-sharing strategy and undertakes part of the retailer's information-sharing cost as a subsidy ω. Subsequently, the retailer determines its own traceability information-sharing behavior based on the manufacturer's sharing strategy and the level of subsidy provided. Under this setting, the objective function of the retailer can be expressed as follows:

$$R(Ts_4) = \int_0^\infty e^{-\rho t}\left\{\pi^s[(1-\xi)\alpha + \mu Q(t) + \kappa G(t)] - \frac{\gamma_3}{2}(1-\omega){Ts_4}^2(t)\right\}dt \tag{27}$$

The objective function of the prefabricated food manufacturer is given as follows:

$$R(Tm_{u4}) = \int_0^\infty e^{-\rho t}\left\{\pi_u^m[\xi\alpha + \mu Q(t) + \kappa G(t)] - \frac{\gamma_1}{2}{Tm_{u4}}^2\right\}dt \tag{28}$$

$$R(Tm_{d4}) = \int_0^\infty e^{-pt}\left\{\pi_d^m[(1-\xi)\alpha + \mu Q(t) + \kappa G(t)] - \frac{\gamma_2}{2}{Tm_{d4}}^2(t) - \frac{\omega\gamma_3}{2}{Ts_4}^2(t)\right\}dt \tag{29}$$

Proposition 4. *The equilibrium results under the manufacturer-led cost-sharing contract scenario are as follows:*

(1) The optimal trajectories of product quality $Q(t)$ and goodwill $G(t)$ are given by the following:

$$Q_4(t) = \begin{cases} \frac{(\pi_u^m\eta^2\gamma_2+\pi_d^m\beta^2\gamma_1)\mu}{\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)} + \frac{(\pi_u^m\eta^2\gamma_2+\pi_d^m\beta^2\gamma_1)\kappa}{\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)(\rho+\varphi)} \\ +\left\{Q_0 - \frac{(\pi_u^m\eta^2\gamma_2+\pi_d^m\beta^2\gamma_1)\mu}{\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)} - \frac{(\pi_u^m\eta^2\gamma_2+\pi_d^m\beta^2\gamma_1)\kappa}{\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)(\rho+\varphi)}\right\}e^{-\varepsilon t} \end{cases} \tag{30}$$

$$G_4(t) = \begin{cases} \frac{(\pi_u^m\eta^2\gamma_2+\pi_d^m\beta^2\gamma_1)\mu\phi}{\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)} + \frac{(\pi_u^m\eta^2\gamma_2+\pi_d^m\beta^2\gamma_1)\kappa\phi}{\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)(\rho+\varphi)} + \frac{\theta^2\kappa(2\pi_d^m+\pi^s)}{2\varphi\gamma_3(\rho+\varphi)} \\ +\left\{G_0 - \frac{(\pi_u^m\eta^2\gamma_2+\pi_d^m\beta^2\gamma_1)\mu\phi}{\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)} - \frac{(\pi_u^m\eta^2\gamma_2+\pi_d^m\beta^2\gamma_1)\kappa\phi}{\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)(\rho+\varphi)} - \frac{\theta^2\kappa(2\pi_d^m+\pi^s)}{2\varphi\gamma_3(\rho+\varphi)}\right\}e^{-\varphi t} \end{cases} \tag{31}$$

(2) The optimal value functions of the manufacturer's online and offline channels, of the retailer, and of the overall food supply chain are given as follows:

$$Vm_{u4} = \begin{cases} [\frac{\pi_u^m\mu}{\rho+\varepsilon} + \frac{\pi_u^m\kappa}{(\rho+\varepsilon)(\rho+\varphi)}]Q(t) + \frac{\pi_u^m\kappa}{\rho+\varphi}G(t) + \frac{\eta^2}{2\rho\gamma_1}[\frac{\pi_u^m\mu}{\rho+\varepsilon} + \frac{\pi_u^m\kappa}{(\rho+\varepsilon)(\rho+\varphi)}]^2 \\ +\frac{\beta^2}{\rho\gamma_2}[\frac{\pi_u^m\mu}{\rho+\varepsilon} + \frac{\pi_u^m\kappa}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{\pi_d^m\mu}{\rho+\varepsilon} + \frac{\pi_d^m\kappa}{(\rho+\varepsilon)(\rho+\varphi)}] + \frac{\theta^2\pi_u^m\pi^s\kappa^2}{\rho\gamma_3(\rho+\varphi)^2} + \frac{\pi_u^m\xi\alpha}{\rho} \end{cases} \tag{32}$$

$$Vm_{d4} = \begin{cases} [\frac{\pi_d^m\mu}{\rho+\varepsilon} + \frac{\pi_d^m\kappa}{(\rho+\varepsilon)(\rho+\varphi)}]Q(t) + \frac{\pi_d^m\kappa}{\rho+\varphi}G(t) + \frac{\pi_d^m(1-\xi)\alpha}{\rho} + \frac{\theta^2}{8\rho\gamma_3}[\frac{(2\pi_d^m+\pi^s)\kappa}{\rho+\varphi}]^2 \\ +\frac{\eta^2}{\rho\gamma_1}[\frac{\pi_u^m\mu}{\rho+\varepsilon} + \frac{\pi_u^m\kappa}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{\pi_d^m\mu}{\rho+\varepsilon} + \frac{\pi_d^m\kappa}{(\rho+\varepsilon)(\rho+\varphi)}] + \frac{\beta^2}{2\rho\gamma_2}[\frac{\pi_d^m\mu}{\rho+\varepsilon} + \frac{\pi_d^m\kappa}{(\rho+\varepsilon)(\rho+\varphi)}]^2 \end{cases} \tag{33}$$

$$Vs_4 = \begin{cases} [\frac{\pi^s\mu}{\rho+\varepsilon} + \frac{\pi^s\kappa}{(\rho+\varepsilon)(\rho+\varphi)}]Q(t) + \frac{\pi^s\kappa}{\rho+\varphi}G(t) + \frac{\pi^s(1-\xi)\alpha}{\rho} \\ + \frac{\eta^2}{\rho\gamma_1}[\frac{\pi_u^m\mu}{\rho+\varepsilon} + \frac{\pi_u^m\kappa}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{\pi^s\mu}{\rho+\varepsilon} + \frac{\pi^s\kappa}{(\rho+\varepsilon)(\rho+\varphi)}] \\ + \frac{\beta^2}{\rho\gamma_2}[\frac{\pi_d^m\mu}{\rho+\varepsilon} + \frac{\pi_d^m\kappa}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{\pi^s\mu}{\rho+\varepsilon} + \frac{\pi^s\kappa}{(\rho+\varepsilon)(\rho+\varphi)}] + \frac{\theta^2\pi^s\kappa^2(2\pi_d^m+\pi^s)}{4\rho\gamma_3(\rho+\varphi)^2} \end{cases} \quad (34)$$

$$V_4 = \begin{cases} [\frac{(\pi_u^m+\pi_d^m+\pi^s)\mu}{\rho+\varepsilon} + \frac{(\pi_u^m+\pi_d^m+\pi^s)\kappa}{(\rho+\varepsilon)(\rho+\varphi)}]Q(t) + \frac{(\pi_u^m+\pi_d^m+\pi^s)\kappa}{\rho+\varphi}G(t) \\ + \frac{\eta^2}{\rho\gamma_1}[\frac{\pi_u^m\mu}{\rho+\varepsilon} + \frac{\pi_u^m\kappa}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{(\pi_u^m+2\pi_d^m+2\pi^s)\mu}{\rho+\varepsilon} + \frac{(\pi_u^m+2\pi_d^m+2\pi^s)\kappa}{(\rho+\varepsilon)(\rho+\varphi)}] \\ + \frac{\beta^2}{\rho\gamma_2}[\frac{\pi_d^m\mu}{\rho+\varepsilon} + \frac{\pi_d^m\kappa}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{(\pi_u^m+\pi_d^m+2\pi^s)\mu}{\rho+\varepsilon} + \frac{(\pi_u^m+\pi_d^m+2\pi^s)\kappa}{(\rho+\varepsilon)(\rho+\varphi)}] \\ + \frac{\theta^2}{\rho\gamma_3(\rho+\varphi)^2}[\pi_u^m\pi^s\kappa^2 + \frac{\pi^s\kappa^2(2\pi_d^m+\pi^s)}{4}] + \frac{\alpha}{\rho}[\pi_u^m\xi + \pi_d^m(1-\xi) + \pi^s(1-\xi)] \end{cases} \quad (35)$$

The detailed proof is provided in Appendix A.

Under the manufacturer-led cost-sharing contract model, the manufacturer provides subsidies to incentivize the retailer to engage in traceability information sharing. In this scenario, the retailer's enthusiasm for participating in information sharing is enhanced; however, the level of engagement may still be lower than that observed under the centralized decision-making model. This phenomenon indicates that the manufacturer-led model possesses certain potential for further development in optimizing the traceability information-sharing behaviors of manufacturers and retailers.

4. Comparative Analysis

Based on the results of the previous models' construction and solution, as summarized in Table 3, the following conclusions can be drawn:

Table 3. Optimal solutions under different scenarios.

Scenario	Optimal Solution	
Centralized Decision-Making	$Tm_{u1} = \frac{(\pi_u^m+\pi_d^m+\pi^s)\eta\mu}{\gamma_1(\rho+\varepsilon)} + \frac{(\pi_u^m+\pi_d^m+\pi^s)\eta\phi\kappa}{\gamma_1(\rho+\varepsilon)(\rho+\varphi)}$; $Tm_{d1} = \frac{(\pi_u^m+\pi_d^m+\pi^s)\beta\mu}{\gamma_2(\rho+\varepsilon)} + \frac{(\pi_u^m+\pi_d^m+\pi^s)\beta\phi\kappa}{\gamma_2(\rho+\varepsilon)(\rho+\varphi)}$; $Ts_1 = \frac{(\pi_u^m+\pi_d^m+\pi^s)\theta\kappa}{\gamma_3(\rho+\varphi)}$	$Q_\infty^1 = \frac{(\pi_u^m+\pi_d^m+\pi^s)(\eta^2\gamma_2+\beta^2\gamma_1)\mu}{\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)} + \frac{(\pi_u^m+\pi_d^m+\pi^s)(\eta^2\gamma_2+\beta^2\gamma_1)\phi\kappa}{\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)(\rho+\varphi)}$; $G_\infty^1 = \frac{(\pi_u^m+\pi_d^m+\pi^s)(\eta^2\gamma_2+\beta^2\gamma_1)\mu\phi}{\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)} + \frac{(\pi_u^m+\pi_d^m+\pi^s)(\eta^2\gamma_2+\beta^2\gamma_1)\phi^2\kappa}{\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)(\rho+\varphi)} + \frac{(\pi_u^m+\pi_d^m+\pi^s)\theta^2\kappa}{\varphi\gamma_3(\rho+\varphi)}$
Decentralized Decision-Making	$Tm_{u2} = \frac{\eta}{\gamma_1}[\frac{\pi_u^m\mu}{\rho+\varepsilon} + \frac{\pi_u^m\kappa\phi}{(\rho+\varepsilon)(\rho+\varphi)}]$; $Tm_{d2} = \frac{\beta}{\gamma_2}[\frac{\pi_d^m\mu}{\rho+\varepsilon} + \frac{\pi_d^m\kappa\phi}{(\rho+\varepsilon)(\rho+\varphi)}]$; $Ts_2 = \frac{\theta\pi^s\kappa}{\gamma_3(\rho+\varphi)}$	$Q_\infty^2 = \frac{(\pi_u^m\eta^2\gamma_2+\pi_d^m\beta^2\gamma_1)\mu}{\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)} + \frac{(\pi_u^m\eta^2\gamma_2+\pi_d^m\beta^2\gamma_1)\kappa\phi}{\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)(\rho+\varphi)}$; $G_\infty^2 = \frac{(\pi_u^m\eta^2\gamma_2+\pi_d^m\beta^2\gamma_1)\phi\mu}{\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)} + \frac{(\pi_u^m\eta^2\gamma_2+\pi_d^m\beta^2\gamma_1)\kappa\phi^2}{\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)(\rho+\varphi)} + \frac{\theta^2\pi^s\kappa}{\varphi\gamma_3(\rho+\varphi)}$
Retailer-Led Model	$Tm_{u3} = \frac{\eta}{\gamma_1}[\frac{\pi_u^m\mu}{\rho+\varepsilon} + \frac{\pi_u^m\kappa}{(\rho+\varepsilon)(\rho+\varphi)}]$; $Tm_{d3} = \frac{\beta(2\pi^s+\pi_d^m)}{2\gamma_2}[\frac{\mu}{\rho+\varepsilon} + \frac{\kappa}{(\rho+\varepsilon)(\rho+\varphi)}]$; $Ts_3 = \frac{\theta\pi^s\kappa}{\gamma_3(\rho+\varphi)}$	$Q_\infty^3 = \frac{[2\pi_u^m\eta^2\gamma_2+(2\pi^s+\pi_d^m)\beta^2\gamma_1]\mu}{2\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)} + \frac{[2\pi_u^m\eta^2\gamma_2+(2\pi^s+\pi_d^m)\beta^2\gamma_1]\kappa}{2\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)(\rho+\varphi)}$; $G_\infty^3 = \frac{[2\pi_u^m\eta^2\gamma_2+(2\pi^s+\pi_d^m)\beta^2\gamma_1]\phi\mu}{2\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)} + \frac{[2\pi_u^m\eta^2\gamma_2+(2\pi^s+\pi_d^m)\beta^2\gamma_1]\phi\kappa}{2\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)(\rho+\varphi)} + \frac{\theta^2\pi^s\kappa}{\varphi\gamma_3(\rho+\varphi)}$
Manufacturer-Led Model	$Tm_{u4} = \frac{\eta}{\gamma_1}[\frac{\pi_u^m\mu}{\rho+\varepsilon} + \frac{\pi_u^m\kappa}{(\rho+\varepsilon)(\rho+\varphi)}]$; $Tm_{d4} = \frac{\beta}{\gamma_2}[\frac{\pi_d^m\mu}{\rho+\varepsilon} + \frac{\pi_d^m\kappa}{(\rho+\varepsilon)(\rho+\varphi)}]$; $Ts_4 = \frac{\theta\kappa(2\pi_d^m+\pi^s)}{2\gamma_3(\rho+\varphi)}$;	$Q_\infty^4 = \frac{(\pi_u^m\eta^2\gamma_2+\pi_d^m\beta^2\gamma_1)\mu}{\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)} + \frac{(\pi_u^m\eta^2\gamma_2+\pi_d^m\beta^2\gamma_1)\kappa}{\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)(\rho+\varphi)}$; $G_\infty^4 = \frac{(\pi_u^m\eta^2\gamma_2+\pi_d^m\beta^2\gamma_1)\mu\phi}{\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)} + \frac{(\pi_u^m\eta^2\gamma_2+\pi_d^m\beta^2\gamma_1)\kappa\phi}{\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)(\rho+\varphi)} + \frac{\theta^2\kappa(2\pi_d^m+\pi^s)}{2\varphi\gamma_3(\rho+\varphi)}$

Lemma 1.

1. *For product quality Q*:
   when $0 \le Q_0 < \frac{(\pi_u^m+\pi_d^m+\pi^s)(\eta^2\gamma_2+\beta^2\gamma_1)\mu}{\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)} + \frac{(\pi_u^m+\pi_d^m+\pi^s)(\eta^2\gamma_2+\beta^2\gamma_1)\phi\kappa}{\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)(\rho+\varphi)}$, *Q1 shows an increasing trend.*

when $0 \le Q_0 < \frac{(\pi_u^m \eta^2 \gamma_2 + \pi_d^m \beta^2 \gamma_1)\mu}{\varepsilon \gamma_1 \gamma_2 (\rho+\varepsilon)} + \frac{(\pi_u^m \eta^2 \gamma_2 + \pi_d^m \beta^2 \gamma_1)\kappa\phi}{\varepsilon \gamma_1 \gamma_2 (\rho+\varepsilon)(\rho+\varphi)}$, *Q2 shows an increasing trend.*

when $0 \le Q_0 < \frac{[2\pi_u^m \eta^2 \gamma_2 + (2\pi^s + \pi_d^m)\beta^2 \gamma_1]\mu}{2\varepsilon \gamma_1 \gamma_2 (\rho+\varepsilon)} + \frac{[2\pi_u^m \eta^2 \gamma_2 + (2\pi^s + \pi_d^m)\beta^2 \gamma_1]\kappa}{2\varepsilon \gamma_1 \gamma_2 (\rho+\varepsilon)(\rho+\varphi)}$, *Q3 shows an increasing trend.*

when $0 \le Q_0 < \frac{(\pi_u^m \eta^2 \gamma_2 + \pi_d^m \beta^2 \gamma_1)\mu}{\varepsilon \gamma_1 \gamma_2 (\rho+\varepsilon)} + \frac{(\pi_u^m \eta^2 \gamma_2 + \pi_d^m \beta^2 \gamma_1)\kappa}{\varepsilon \gamma_1 \gamma_2 (\rho+\varepsilon)(\rho+\varphi)}$, *Q4 shows an increasing trend.*

2. *For goodwill G*:

when $0 \le G_0 < \frac{(\pi_u^m + \pi_d^m + \pi^s)(\eta^2 \gamma_2 + \beta^2 \gamma_1)\mu\phi}{\varphi \varepsilon \gamma_1 \gamma_2 (\rho+\varepsilon)} + \frac{(\pi_u^m + \pi_d^m + \pi^s)(\eta^2 \gamma_2 + \beta^2 \gamma_1)\phi^2 \kappa}{\varphi \varepsilon \gamma_1 \gamma_2 (\rho+\varepsilon)(\rho+\varphi)} + \frac{(\pi_u^m + \pi_d^m + \pi^s)\theta^2 \kappa}{\varphi \gamma_3 (\rho+\varphi)}$, *G1 shows an increasing trend.*

when $0 \le G_0 < \frac{(\pi_u^m \eta^2 \gamma_2 + \pi_d^m \beta^2 \gamma_1)\phi\mu}{\varphi \varepsilon \gamma_1 \gamma_2 (\rho+\varepsilon)} + \frac{(\pi_u^m \eta^2 \gamma_2 + \pi_d^m \beta^2 \gamma_1)\kappa\phi^2}{\varphi \varepsilon \gamma_1 \gamma_2 (\rho+\varepsilon)(\rho+\varphi)} + \frac{\theta^2 \pi^s \kappa}{\varphi \gamma_3 (\rho+\varphi)}$, *G2 shows an increasing trend.*

when $0 \le G_0 < \frac{[2\pi_u^m \eta^2 \gamma_2 + (2\pi^s + \pi_d^m)\beta^2 \gamma_1]\phi\mu}{2\varphi \varepsilon \gamma_1 \gamma_2 (\rho+\varepsilon)} + \frac{\theta^2 \pi^s \kappa}{\varphi \gamma_3 (\rho+\varphi)} + \frac{[2\pi_u^m \eta^2 \gamma_2 + (2\pi^s + \pi_d^m)\beta^2 \gamma_1]\phi\kappa}{2\varphi \varepsilon \gamma_1 \gamma_2 (\rho+\varepsilon)(\rho+\varphi)}$, *G3 shows an increasing trend.*

when $0 \le G_0 < \frac{(\pi_u^m \eta^2 \gamma_2 + \pi_d^m \beta^2 \gamma_1)\phi\mu}{\varphi \varepsilon \gamma_1 \gamma_2 (\rho+\varepsilon)} + \frac{(\pi_u^m \eta^2 \gamma_2 + \pi_d^m \beta^2 \gamma_1)\kappa\phi}{\varphi \varepsilon \gamma_1 \gamma_2 (\rho+\varepsilon)(\rho+\varphi)} + \frac{\theta^2 \kappa (2\pi_d^m + \pi^s)}{2\varphi \gamma_3 (\rho+\varphi)}$, *G4 shows an increasing trend.*

These trends are consistent with the optimal solutions presented in Table 3.

Corollary 1. *As time progresses—thus,* $t \to \infty$ *—product quality and goodwill eventually reach a steady state. According to the steady-state values* Q_∞ *and* G_∞*, consumers' preference for prefabricated food with high quality and good reputation increases over time, leading to a continuous rise in product sales in the market and, consequently, an increase in overall profit.*

Lemma 2. *Online traceability information-sharing behavior of prefabricated food manufacturers:* $Tm_{u1} > Tm_{u2}$; $Tm_{u3} = Tm_{u4} > Tm_{u2}$.

Proof. $Tm_{u1} - Tm_{u2} = \frac{(\pi_d^m + \pi^s)\eta\mu}{\gamma_1(\rho+\varepsilon)} + \frac{(\pi_d^m + \pi^s)\eta\phi\kappa}{\gamma_1(\rho+\varepsilon)(\rho+\varphi)} > 0$, therefore $Tm_{u1} > Tm_{u2}$;

$$Tm_{u3} = Tm_{u4} = \frac{\eta}{\gamma_1}\left[\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}\right],\ Tm_{u3} - Tm_{u2} = \frac{\pi_u^m \eta\kappa(1-\phi)}{\gamma_1(\rho+\varepsilon)(\rho+\varphi)}$$

Since $\phi \in (0,1]$, therefore $1-\phi \in [0,1)$, hence $Tm_{u3} - Tm_{u2} > 0$, that is, $Tm_{u3} = Tm_{u4} > Tm_{u2}$. □

Corollary 2. *In terms of enthusiasm for traceability information sharing, retailers perform significantly better under the centralized decision-making model than under the decentralized one. For instance, the prefabricated food brand under Haidilao adopts a centralized decision-making model, in which retail stores actively share traceability information, such as consumer feedback, whereas production plants comprehensively share product traceability information. By analyzing consumer preference data, the production plants promptly adjust product flavors and packaging, launching prefabricated food products that better meet market demand. As a result, sales have increased substantially and brand goodwill has improved, providing strong evidence that the centralized decision-making model enhances retailers' enthusiasm for traceability information sharing.*

Compared with the decentralized model, the cost-sharing contract injects stronger motivation into traceability information sharing. This is because the subsidy reduces the sharing cost of the subsidized party, thereby increasing its marginal benefits from information sharing. From the perspective of profit maximization, the subsidized party becomes more willing to increase its investment in information sharing. For example, in the cooperation between Yonghui Superstores and Synear Food, Yonghui provided subsidies for the cost of online traceability information sharing. In response, Synear Food increased its online traceability efforts by providing detailed information on raw material

origins, processing procedures, and quality inspection reports. Consumers could obtain comprehensive product details by scanning QR codes, which enhanced their confidence in product quality. Consequently, online sales increased significantly, driving higher brand recognition and offline sales, thereby fully validating the positive effect of this model on the manufacturer's enthusiasm for online traceability information sharing.

Lemma 3. *Offline traceability information-sharing behavior of prefabricated food manufacturers:*

$$\begin{cases} Tm_{d1} > Tm_{d2}; \\ Tm_{d3} > Tm_{d4} > Tm_{d2} \end{cases}$$

Proof. $Tm_{d1} - Tm_{d2} = \frac{(\pi_u^m + \pi^s)\beta\mu}{\gamma_2(\rho+\varepsilon)} + \frac{(\pi_u^m + \pi^s)\beta\phi\kappa}{\gamma_2(\rho+\varepsilon)(\rho+\varphi)} > 0$, hence $Tm_{d1} > Tm_{d2}$;

$$Tm_{d4} - Tm_{d2} = \frac{\pi_d^m \eta\kappa(1-\phi)}{\gamma_2(\rho+\varepsilon)(\rho+\varphi)},$$

Since $\phi \in (0,1]$, therefore $1-\phi \in [0,1)$, hence $Tm_{d4} - Tm_{d2} > 0$, that is, $Tm_{d4} > Tm_{d2}$.

$$Tm_{d3} - Tm_{d4} = \frac{\beta(2\pi^s - \pi_d^m)}{2\gamma_2}[\frac{\mu}{\rho+\varepsilon} + \frac{\kappa}{(\rho+\varepsilon)(\rho+\varphi)}],$$

Therefore, if $2\pi^s > \pi_d^m$, then $Tm_{d3} > Tm_{d4}$; since the subsidy implicitly contains $2\pi^s > \pi_d^m$ as discussed earlier, thus $Tm_{d3} > Tm_{d4}$. □

Corollary 3. *According to Proposition 2, the intensity of the manufacturer's offline traceability information-sharing behavior under the centralized decision-making scenario is greater than that under the decentralized decision-making scenario. For example, in large-scale chain prefabricated food enterprises, the headquarters centrally allocates resources, encouraging each production base to actively collect and share traceability information across all offline channel stages—from production to sales terminals. Detailed records of product environmental parameters during warehousing and logistics processes are maintained to ensure product quality and efficient supply chain operation.*

Under the cost-sharing model, when the subsidy ratio is set at an appropriate level, when $2\pi^s > \pi_d^m$, the manufacturer-led cost-sharing model can more effectively stimulate all parties' enthusiasm for traceability information sharing. Its level of enthusiasm exceeds that of the retailer-led model and of the decentralized decision-making model.

For instance, Guolian Aquatic Products, positioned as a high-end brand in the market, faces consumers who are particularly concerned about product traceability information. Under the manufacturer-led cost-sharing model, the manufacturer provides the retailer with a higher proportion of subsidies for traceability information-sharing costs while also increasing its own investment in offline traceability information sharing. From the raw material procurement stage, detailed records are kept regarding the fishing area and the time of seafood capture. During the processing stage, hygiene standards and processing parameters are meticulously documented, and in the warehousing and logistics stages, temperature and humidity data are precisely recorded. Through comprehensive and detailed offline traceability information sharing, consumers gain a clear understanding of the entire process from ocean to table, leading to high recognition of product quality. Consequently, product sales and profits have increased substantially.

By contrast, another prefabricated food brand in the same region that adopted a retailer-led cost-sharing model offered limited subsidies to manufacturers, resulting in lower investment in offline traceability information sharing. Consequently, consumers'

trust and purchase intentions toward their products were relatively weak, and their market performance was markedly inferior. This fully verifies that, when the subsidy ratio is appropriately designed, the manufacturer-led cost-sharing model outperforms the retailer-led model in stimulating manufacturers' enthusiasm for offline traceability information sharing, and both are superior to the decentralized decision-making model.

Lemma 4. *Traceability information-sharing behavior of prefabricated food retailers:*

$$Ts_1 > Ts_4 > Ts_2 = Ts_3$$

Proof. $Ts_1 - Ts_2 = \frac{(\pi_u^m + \pi_d^m)\theta\kappa}{\gamma_3(\rho+\varphi)} > 0$, $Ts_2 = Ts_3 = \frac{\theta\pi^s\kappa}{\gamma_3(\rho+\varphi)}$, therefore $Ts_1 > Ts_2 = Ts_3$;

$$Ts_1 - Ts_4 = \frac{(2\pi_u^m + \pi^s)\theta\kappa}{2\gamma_3(\rho+\varphi)} > 0, \text{ thus } Ts_1 > Ts_4; Ts_4 - Ts_2 = \frac{\theta\kappa(2\pi_d^m - \pi^s)}{2\gamma_3(\rho+\varphi)},$$

As implied in the preceding section $2\pi_d^m > \pi^s$, thus $Ts_1 > Ts_4 > Ts_2 = Ts_3$.

Since it is implicitly included in the previous analysis that $2\pi_d^m > \pi^s$, it follows that $Ts_1 > Ts_4 > Ts_2 = Ts_3$. □

Corollary 4. *For retailers, the retailer-led cost-sharing strategy and the decentralized decision-making model tend to reduce their enthusiasm for traceability information sharing. This is because information sharing requires additional investment but yields limited direct benefits, resulting in low motivation. For example, in a regional prefabricated food market, local retailers provided subsidies to manufacturers, leading manufacturers to increase their traceability information-sharing efforts. However, the retailers remained inactive in collecting and sharing consumer feedback and product acceptance information, which prevented manufacturers from optimizing their products.*

Under the centralized decision-making model, retailers' enthusiasm for traceability information sharing increases significantly and reaches the highest level. Conversely, under the manufacturer-led cost-sharing contract, retailers' enthusiasm can improve under certain conditions but still remains lower than that under the centralized model. For instance, Anjoy Foods provides subsidies and a management platform for its partner retailers. Some retailers, incentivized by this mechanism, actively share feedback information—such as reporting packaging issues—which helps manufacturers improve products and also increases their own profits. However, due to the limitations of interest alignment and coordination mechanisms in this model, the extent of improvement in retailers' enthusiasm remains limited and fails to reach the level observed under centralized decision-making.

5. Profit Distribution Mechanism of the Prefabricated Food Supply Chain Based on the Shapley Value Method

In a strategic alliance, member enterprises achieve synergistic effects through resource integration and complementary advantages. The fairness of profit distribution directly affects the stability and sustainability of the alliance. Since the Shapley value profit distribution model is an effective approach to addressing profit allocation in alliance games [37], this paper adopts this theory, with the specific allocation results presented in Table 4.

Table 4. Profit distribution of the pre-prepared food supply chain.

S	M_1	C	S_1	C
$V(S)$	Vm_1	V_c	Vs_1	V_c
$V(S\backslash H)$	0	Vs_1	0	Vm_1
$V(S)-V(S\backslash \mathrm{H})$	Vm_1	$V_c - Vs_1$	Vs_1	$V_c - Vs_1$
$\mid S \mid$	1	2	1	2
$W(S)$	1/2	1/2	1/2	1/2
$\Phi_i(v)$	$1/2(V_c - Vs_1 + Vm_1)$		$1/2(V_c - Vm_1 + Vs_1)$	

Among them, V represents all possible combinations of the manufacturer and the retailer. $V(S)$ represents the profit value under the corresponding coalition state. $V(S\backslash H)$ denotes the profit obtained by the subcoalition after excluding the strategic alliance H. $\mid S \mid$ indicates the number of members included in the coalition. $w(|S|) = \frac{(|S|-1)!(n-|S|)}{n!}$ denotes the weight coefficient, and n represents the total number of participants involved in the game. $\Phi_i(v) = \sum_{i=1}^{2} w(|S|)[V(S) - V(S\backslash H)]$ is used to define the profit distribution mechanism between the manufacturer and the retailer within the strategic alliance.

6. Case Analysis

Based on the aforementioned model, MATLAB 2016 is used to simulate the change processes of the interests of individual members and the overall supply chain in the food supply chain under different scenarios. In line with the findings of References [38,39] and combined with the actual situation of Anjoy Foods, some parameters are set as shown in Table 5.

Table 5. Parameter assignments.

Variable	ε	μ	π^s	π_d^m	π_u^m	β	γ_1	γ_2	γ_3	κ	ϕ	η	φ	θ
Value	1	0.7	1.1	1	2	0.3	0.7	0.8	0.6	0.3	0.5	0.3	1	0.4

According to the studies by Hong & Huang (2016) and Xu et al. (2016), the following is assumed: $\rho = 0.1, G_0 = 0, Q_0 = 0$ [38,39].

(1) Optimal Trajectories of Quality and Goodwill under Different Initial Values of G_0

Figures 2–4 illustrate the optimal trajectories of product quality when the initial quality levels are $Q_0 = 0$, $Q_0 = 5$, and $Q_0 = 10$, respectively. Under all scenarios, the quality values eventually converge to a steady state. Similarly, Figures 5–7 show the optimal trajectories of goodwill when the initial goodwill levels are $G_0 = 0$, $G_0 = 5$, and $G_0 = 10$, respectively. Likewise, in all cases, the goodwill values ultimately approach a stable equilibrium state.

The following conditions are satisfied:

① Centralized decision-making: $0 \leq Q_0 < 0.7487, 0 \leq G_0 < 0.675$;
② Decentralized decision-making: $0 \leq Q_0 < 0.281, 0 \leq G_0 < 0.22$;
③ Retailer-led cost-sharing model: $0 \leq Q_0 < 0.3866, 0 \leq G_0 < 0.22$;
④ Manufacturer-led cost-sharing model: $0 \leq Q_0 < 0.3268, 0 \leq G_0 < 0.276$

The optimal trajectories of quality and goodwill in the prefabricated food supply chain exhibit an upward trend. By contrast, when these conditions are not met, both variables show a declining trend until they reach a steady state. This result is consistent with Lemma 1 in the comparative analysis section above, further verifying the validity of the previous conclusions.

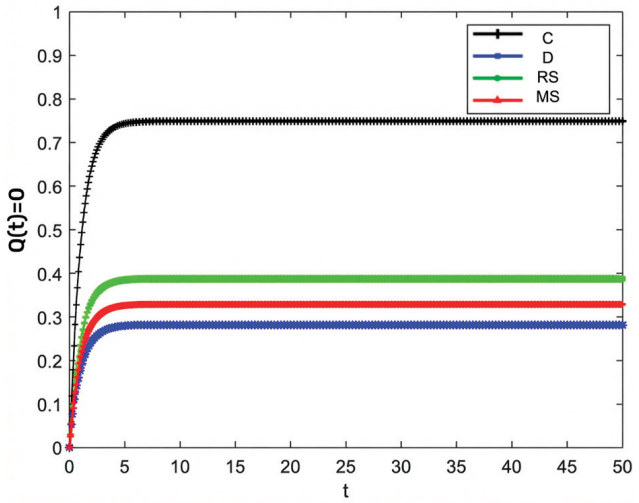


Figure 2. Optimal trajectory of product quality when $Q_0 = 0$.

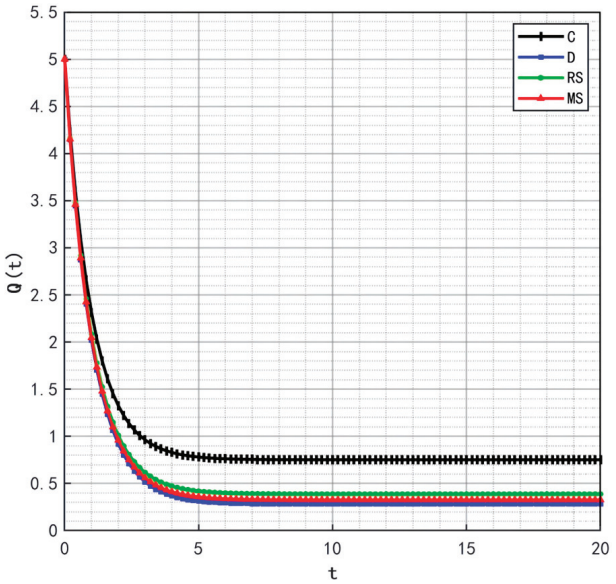


Figure 3. Optimal trajectory of product quality when $Q_0 = 5$.

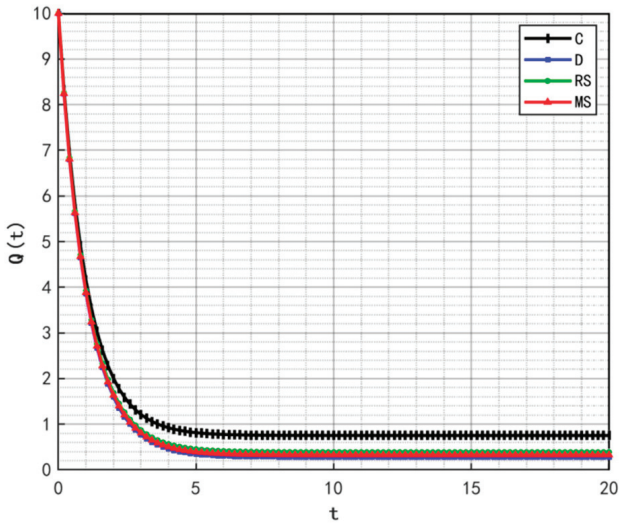


Figure 4. Optimal trajectory of product quality when $Q_0 = 10$.

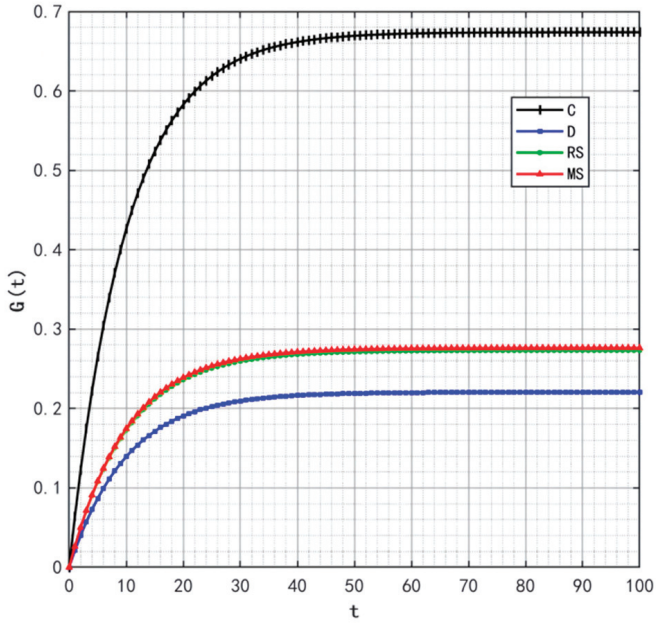


Figure 5. Optimal trajectory of goodwill when $G_0 = 0$.

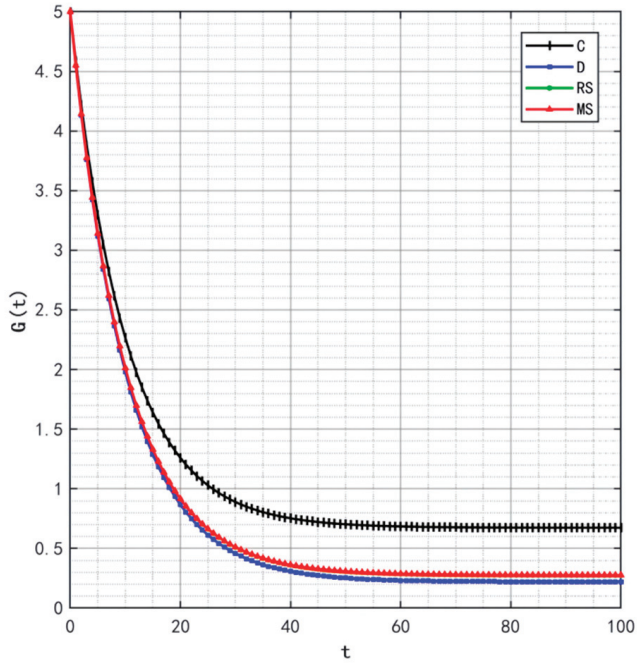


Figure 6. Optimal trajectory of goodwill when $G_0 = 5$.

Figures 8–10 show that the trajectories of quality and goodwill growth under the centralized decision-making scenario consistently outperform those in other scenarios, regardless of the initial levels of quality or goodwill. When the initial quality or goodwill is relatively low, the centralized decision-making model enables rapid improvement through coordination and collaboration among various stages of the supply chain, demonstrating its strong advantages in information sharing and resource integration.

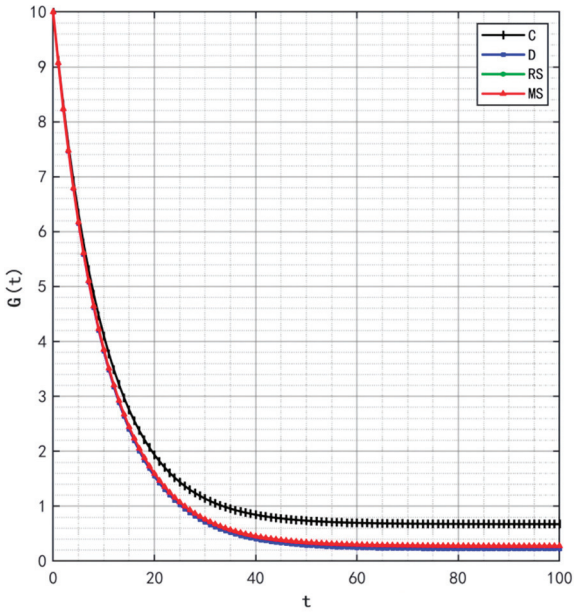


Figure 7. Optimal trajectory of goodwill when $G_0 = 10$.

Conversely, under the decentralized decision-making and cost-sharing contract scenarios (retailer-led and manufacturer-led), the growth of quality and goodwill is noticeably slower and tends to reach a steady state more quickly. This reflects the hindering effect of information asymmetry and conflicts of interest among supply chain members on goodwill accumulation in these settings. Although the cost-sharing contracts attempt to encourage information sharing through subsidies, the subsidizing party's own information-sharing behavior does not improve significantly. As a result, the growth patterns of quality and goodwill are similar to those under decentralized decision-making. These findings indicate that a subsidy-based mechanism alone has inherent limitations in enhancing the overall quality and goodwill performance of the supply chain.

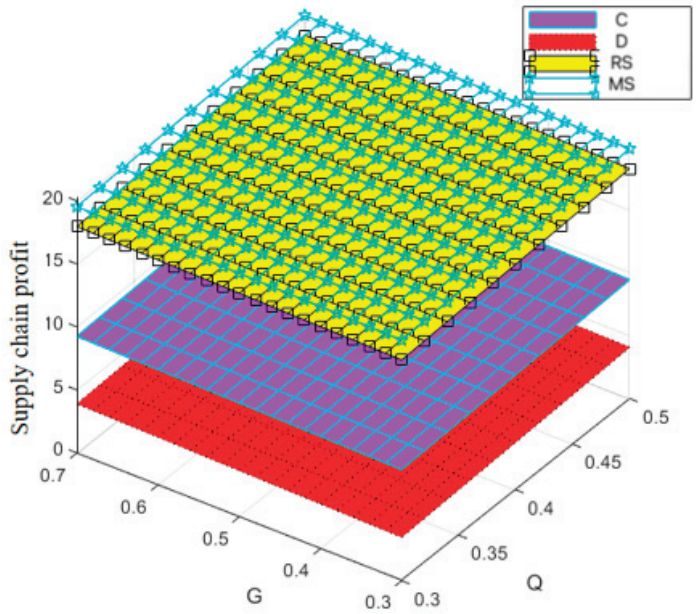


Figure 8. Variation trend of supply chain profit with changes in Q and G.

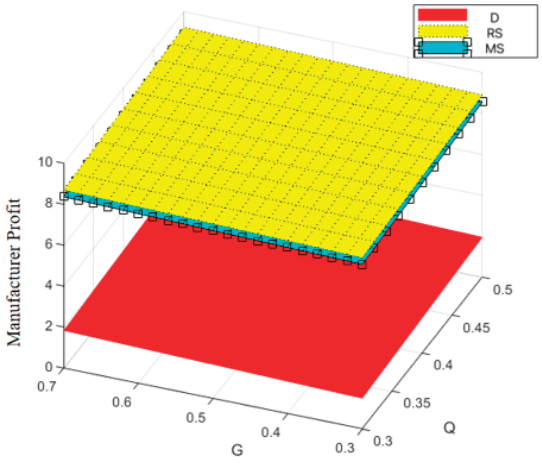


Figure 9. Variation trend of manufacturer profit with changes in Q and G.

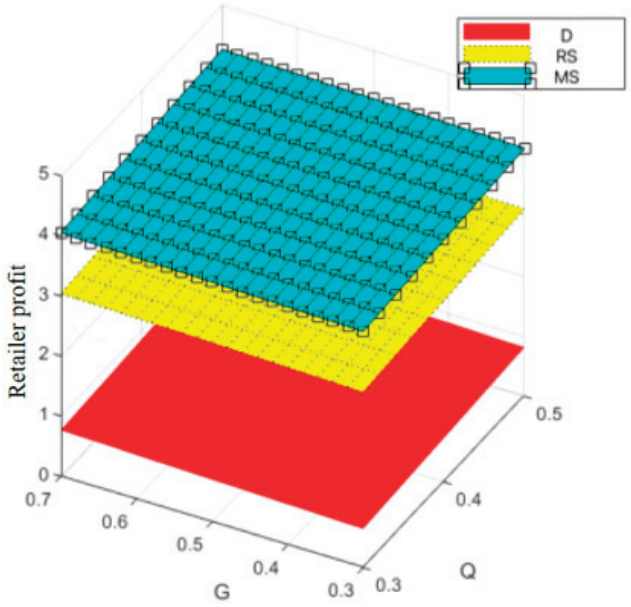


Figure 10. Variation trend of retailer profit with changes in Q and G.

(2) Variation Trends of quality Q and goodwill G with profit.

Figures 8–10 show that goodwill G significantly affects the profit of prefabricated food compared with product quality Q. The main reason behind this phenomenon may lie in the high information barriers associated with quality evaluation for prefabricated food. On the one hand, consumers find it difficult to directly verify product quality; on the other hand, the increasing prevalence of falsified ingredient labels in the prefabricated food market further complicates consumers' ability to distinguish between high- and low-quality products.

In such a market environment characterized by information asymmetry, a brand's strong goodwill has become an important proxy indicator of product quality for most consumers. Consumers tend to equate prefabricated foods with good reputations and with high-quality products that are positively endorsed by word of mouth, thereby showing a stronger preference for purchasing products with higher goodwill.

(3) Variation in market share over time under different decision-making scenarios.

Figure 11 illustrates the changes in market share of the prefabricated food supply chain over time under four scenarios: centralized decision-making, decentralized decision-making, retailer-led cost-sharing contract, and manufacturer-led cost-sharing contract.

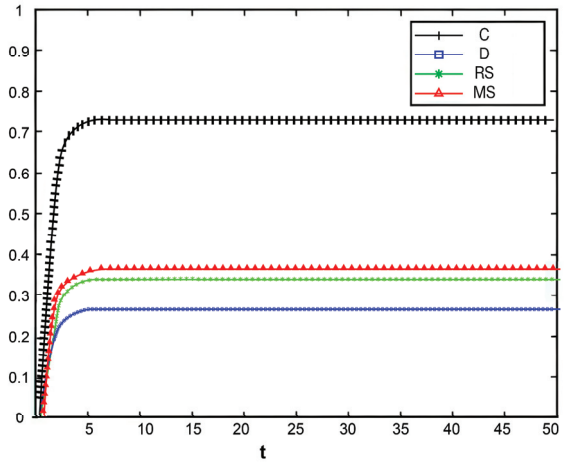


Figure 11. Variation of market share over time under different decision-making scenarios.

Under the centralized decision-making scenario, the supply chain exhibits significant synergy, with high enthusiasm for information sharing. As a result, product quality and goodwill continuously improve, driving a steady increase in market share. Conversely, under the decentralized decision-making scenario, information asymmetry and conflicts of interest are more pronounced, leading to lower enthusiasm for traceability information sharing and limited coordination and efficiency within the supply chain. The Figure shows that compared with centralized decision-making, the growth of market share under decentralized decision-making is notably slower.

Further analysis of the retailer-led cost-sharing contract scenario in Figure 11 shows that the manufacturer's market share increases faster than in the decentralized model. This indicates that the subsidies provided by retailers effectively reduce manufacturers' costs, enhance their enthusiasm for traceability information sharing, and thus promote market share growth. In the manufacturer-led cost-sharing contract scenario, the retailer's market share also grows faster than when under decentralized decision-making, but still fails to reach the level achieved under centralized decision-making.

Overall, different decision-making scenarios significantly affect the market share of supply chain members. The centralized decision-making model achieves overall optimization of the supply chain through collaborative cooperation, making it the most effective approach for improving market share. Although the cost-sharing contract models partially mitigate the shortcomings of decentralized decision-making, further optimization of subsidy strategies and market mechanisms is needed to better promote information sharing and the overall development of the supply chain.

Variation in profit over time.

Figures 12–15, respectively, illustrate the changes in profits of the manufacturer's online and offline channels, the total manufacturer profit, and the retailer profit over time. Figures 12–15 reveal that different decision-making models lead to significant differences in the profits of supply chain participants. Overall, the manufacturer-led cost-sharing contract model achieves a nearly Pareto-optimal outcome. Conversely, under the decentralized decision-making model, the absence of coordination mechanisms results in slower profit growth and the lowest profit levels among all participants.

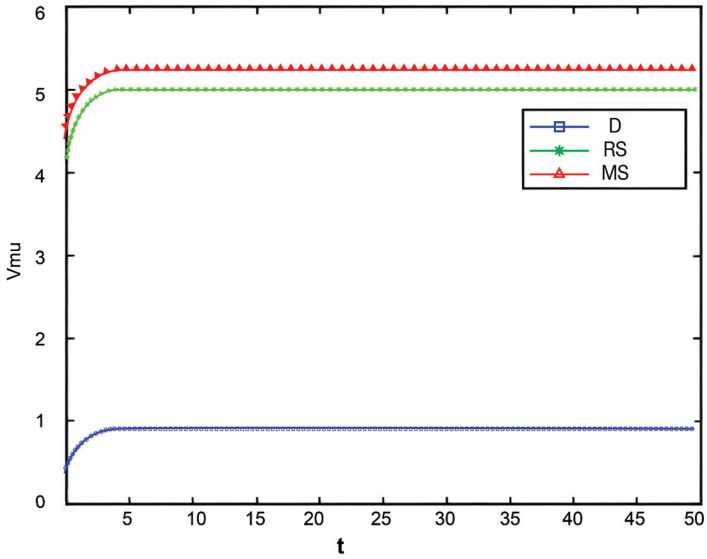


Figure 12. Variation in manufacturers' online channel profit over time.

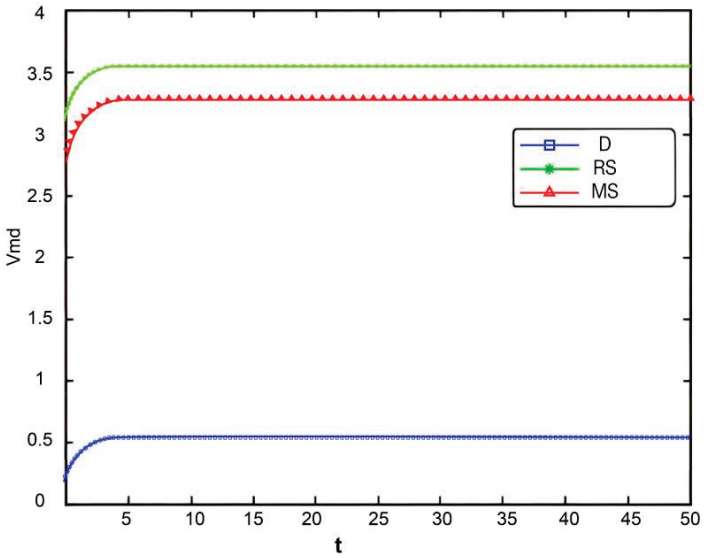


Figure 13. Variation in manufacturers' offline channel profit over time.

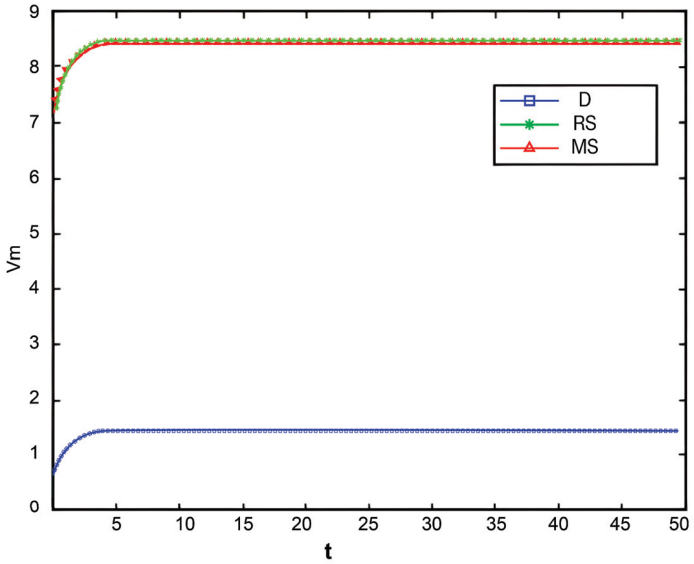


Figure 14. Variation in total manufacturer profit over time.

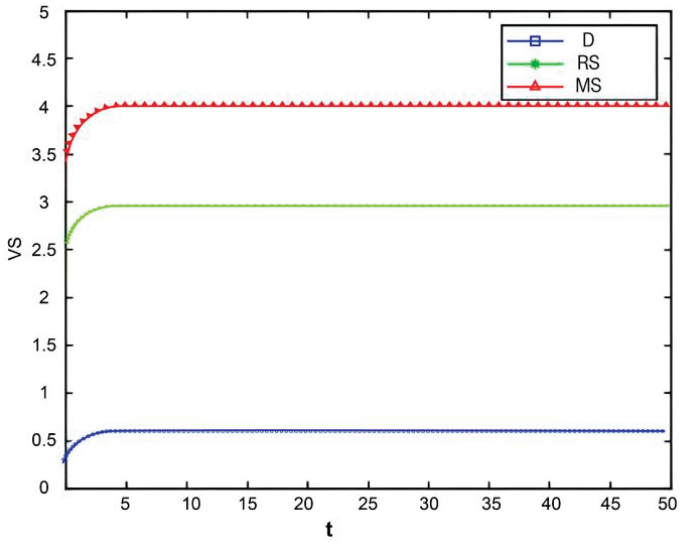


Figure 15. Variation in retailer profit over time.

Figure 16 illustrates the variation in the total supply chain profit over time. It can be seen from the Figure that the centralized decision-making model does not yield the maximum profit, a result that deviates from the conventional perception that centralized decision-making typically enables overall supply chain optimization. However, this finding is supported by practical cases—take Weizhixiang, a prefabricated food enterprise, as an example. In its early stage, the enterprise faced profit pressure after adopting a fully centralized decision-making model. The core cause lies in the disconnection between the unified management model and the regional attributes of prefabricated food consumption: the headquarters implemented unified production and distribution based on the taste preferences of the East China market, which failed to adapt to the differentiated demands of South China, North China, and other regions, resulting in supply–demand mismatch and inventory backlogs. Meanwhile, the unified channel policies and cold chain distribution standards struggled to meet the actual needs of regional terminals, and the inverted logistics costs in remote areas further eroded profits, ultimately undermining the enterprise's overall profitability.

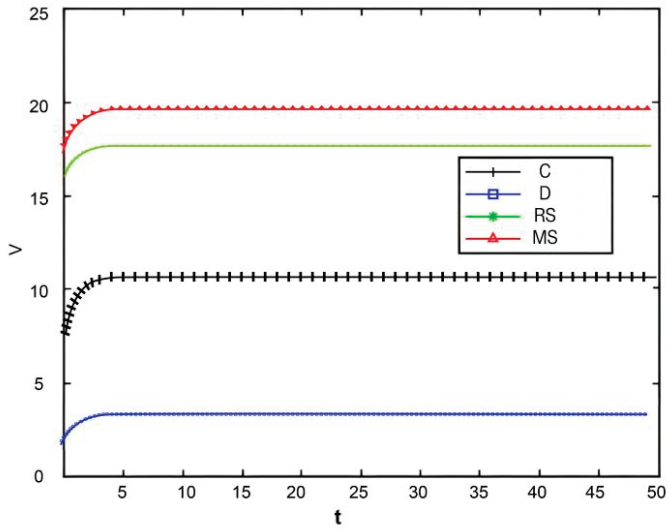


Figure 16. Variation in total supply chain profit over time.

The centralized decision-making model is generally believed to achieve optimal resource allocation through collaborative cooperation in supply chain management, thereby

maximizing the total supply chain profit. However, in this study, the centralized decision-making model did not achieve the highest profit, and the main reasons may be as follows:

1. Insufficient incentive mechanisms: Although centralized decision-making pursues overall profit maximization, it lacks explicit incentives for individual participants to share information, leading to low enthusiasm among manufacturers and retailers and reduced efficiency in information sharing.
2. High coordination costs: Centralized decision-making requires a high degree of coordination, which increases coordination costs and complicates decision-making processes. This may delay market responses and reduce the flexibility of the supply chain.
3. Unequal profit distribution: Profit distribution under centralized decision-making may be unbalanced. Manufacturers and retailers might experience reduced motivation if they bear excessive costs or receive insufficient returns.

(4) Relationship between manufacturers' traceability information-sharing behavior and profit.

Figures 17 and 18 show that, compared with online traceability information-sharing behavior, the manufacturers' offline traceability information-sharing behavior has a more pronounced effect on the profit of prefabricated food manufacturers. This is primarily because the offline channel directly interacts with end consumers. Once offline traceability information sharing enhances consumers' trust in the product, it can quickly translate into purchasing behavior, directly influencing sales and profits. Conversely, although online channels allow for broader information dissemination, consumers may find it difficult to assess the authenticity of online information, and their purchasing decisions are more easily influenced by price and promotional factors. Therefore, offline traceability information sharing exerts a more direct and effective influence on consumers' purchasing decisions.

Across different decision-making models, the effect of traceability information-sharing behavior on profits varies significantly. Under the centralized decision-making model, coordination among all parties allows the manufacturer's information-sharing efforts to receive sufficient support. The manufacturer's online information-sharing activities align effectively with retailer feedback, facilitating product optimization and profit growth. Under the decentralized decision-making model, however, manufacturers prioritize their own profit and may reduce investments in information sharing after weighing costs and benefits, thereby weakening product competitiveness and limiting profit growth.

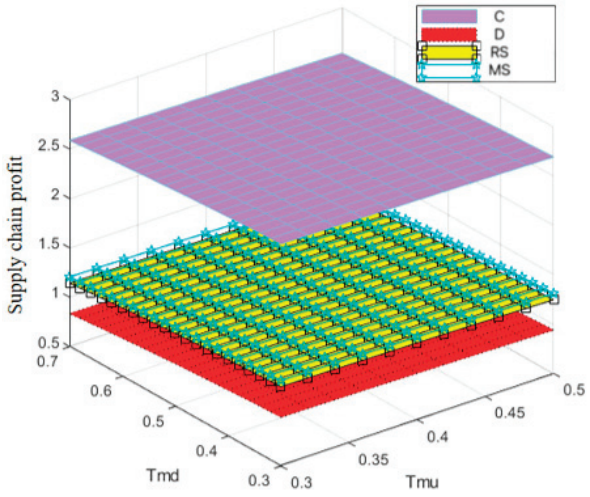


Figure 17. Variation trend of supply chain profit with changes in Tm_u and Tm_d.

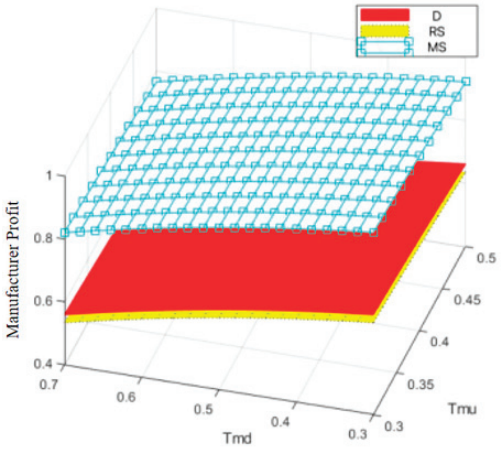
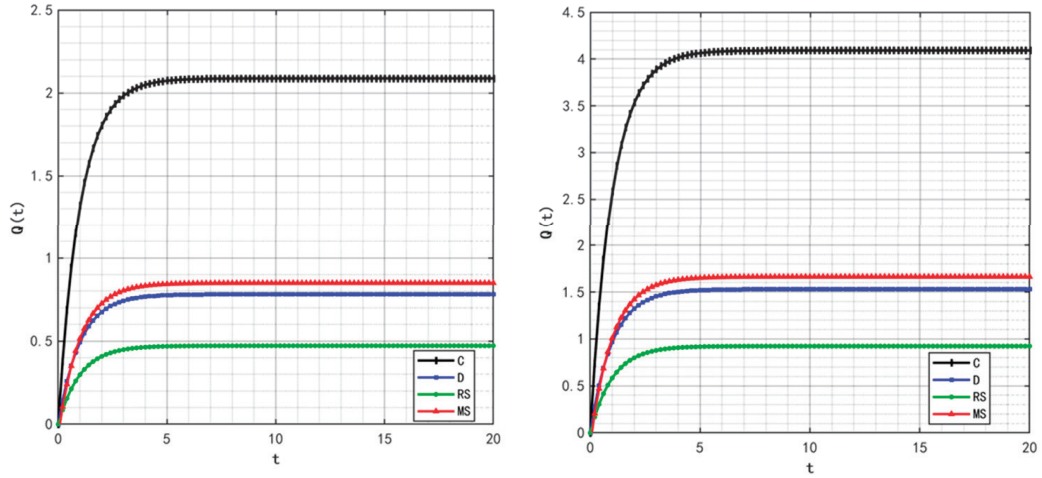


Figure 18. Variation trend of manufacturer profit with changes in Tm_u and Tm_d.

(5) Optimal trajectories of product quality, goodwill, and profit under different parameter combinations.

The quality and safety of prepared foods are crucial to public health. In the event of concentrated exposure to negative public opinion, the government will introduce stricter production supervision measures. At this point, the traceability information-sharing behavior between food manufacturers' online and offline channels becomes critical, which will promote the development of traceability information-sharing technology but also increase technical costs. Although some governments provide technical subsidies, continuous technological upgrading and the increase in implicit participants in the supply chain will still push up manufacturers' technology application costs. Based on this, we set two groups of values for the coefficients (η, β) that measure the impact of traceability information sharing on product quality in manufacturers' online and offline channels: (0.5, 0.5) and (0.7, 0.7). The optimal evolutionary trajectories of corresponding product quality and goodwill are shown in Figures 19 and 20.



(**a**) when $\beta = 0.5$ $\eta = 0.5$ (**b**) when $\beta = 0.7$ $\eta = 0.7$

Figure 19. Variation trends of product quality when (η, β) = (0.5, 0.5) (0.7, 0.7).

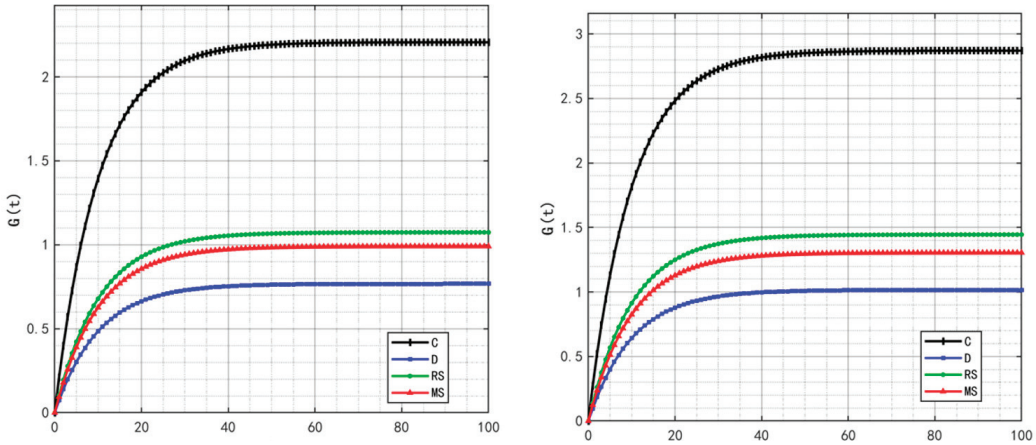


Figure 20. Variation trends of product goodwill when (η, β) = (0.5, 0.5) (0.7, 0.7).

By comparing the simulation results in Figures 2, 5, 19 and 20, it is clear that when the coefficients η and β (which measure the intensity of traceability information sharing's impact on product quality) increase, only under the retailer-led cooperation model does product quality show a fluctuating trend of "rising from the bottom to the second place and then falling back to the bottom." This variation rule also applies to the evolutionary process of product goodwill. Under the retailer-led model, the quality empowerment effect of channel information sharing exhibits "phased differences": in the initial stage, the increase in η and β directly enhances manufacturers' quality control capabilities, driving rapid improvement in product quality; however, as the degree of information sharing deepens, retailers' channel dominance will gradually squeeze manufacturers' quality investment space, while the marginal benefits of information sharing gradually diminish, ultimately leading to a decline in product quality and goodwill from high levels.

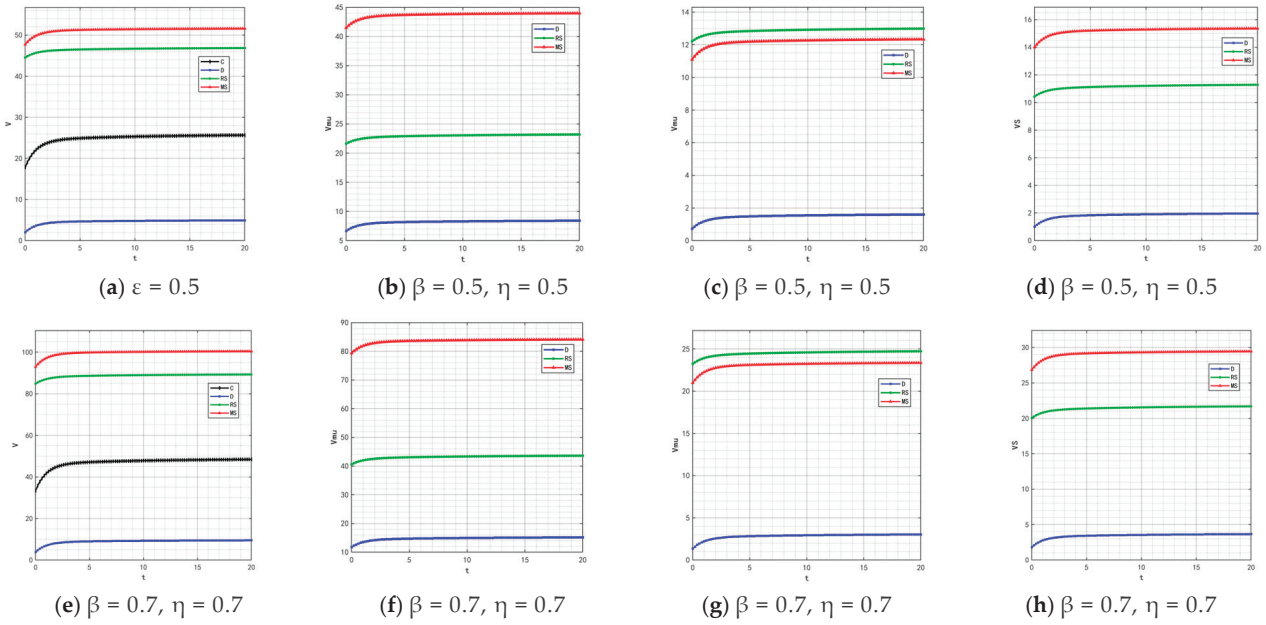
By comparing Figures 12, 14–16 and 21, it is found that as β and η (the coefficients of traceability information sharing's impact on product quality) increase, the profit ranking order of each subject under different models remains unchanged, with only numerical increases. This is because the increase in β and η only enhances the positive impact of traceability information sharing on product quality, thereby generally amplifying the profit level of each subject, but does not change the interest distribution logic under different models.

(6) Impact of changes in traceable food quality attenuation coefficient on product quality, goodwill, and profit.

With the rise in market demand, some unscrupulous merchants will seek profits by tampering with traceability information—such as relabeling expired food as freshly produced or delaying the update of key data. In addition, the frequent appearance of professional terms in traceability information makes it difficult for ordinary consumers to understand. These factors will significantly affect the value of the traceable food quality attenuation coefficient ε. Therefore, ε is set to 0.5 and 1.5, and the corresponding changes in product quality and goodwill are shown in Figures 22 and 23.

By comparing the simulation results in Figures 2, 5, 22 and 23, it is clear that as the system's attenuation coefficient ε gradually increases, the evolutionary laws of product quality and goodwill under different cooperation models show significant differentiated characteristics. Among them, only under the retailer-led cooperation model does product quality exhibit a unique fluctuating trajectory—first rapidly rising from the initial bottom to the second place, then gradually falling back to the bottom. The dynamic change in product goodwill is more complex: the goodwill under the centralized model always

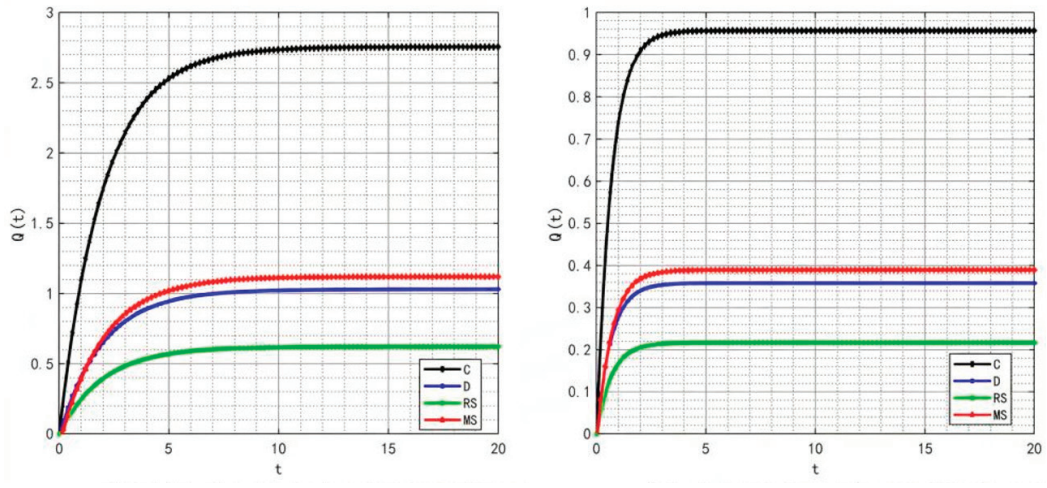
maintains the optimal level, while the goodwill ranking of the manufacturer-led model experiences continuous fluctuations of "falling from the second place to the bottom, then rising back to the third place." The root cause of this phenomenon lies in the differences in the interaction mechanism between the "attenuation effect" and "subject decision-making logic" under different models. For the retailer-led model, the initial increase in ε will force retailers to strengthen channel-side quality collaboration; however, as the attenuation effect intensifies, retailers' short-term profit orientation will lead them to shift resources to channel traffic, weakening the empowerment of manufacturers' quality investment, and ultimately resulting in a quality decline. The centralized model, relying on unified decision-making and scheduling capabilities, can offset the negative impact of the attenuation coefficient, thereby stably maintaining optimal goodwill. As for the manufacturer-led model, the initial increase in ε will weaken the long-term effectiveness of its quality investment, but in the later stage, manufacturers will passively increase goodwill repair investment to maintain market share, thereby promoting a slight recovery in their ranking.



(**a**) ε = 0.5 (**b**) β = 0.5, η = 0.5 (**c**) β = 0.5, η = 0.5 (**d**) β = 0.5, η = 0.5

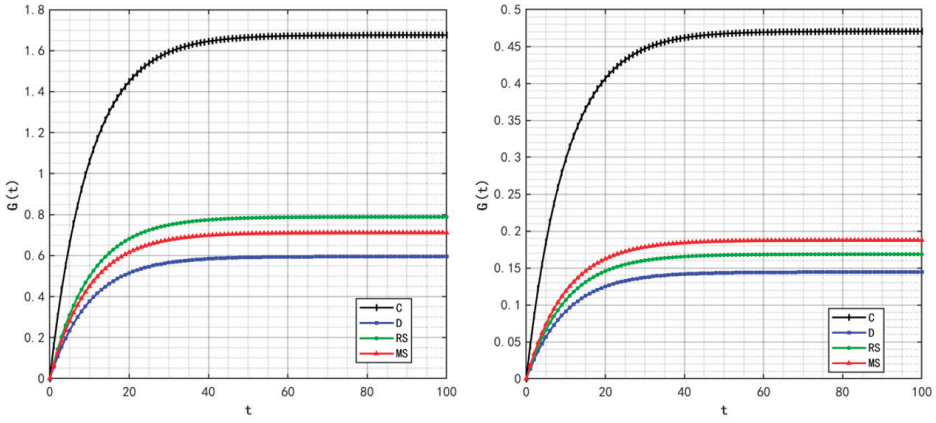
(**e**) β = 0.7, η = 0.7 (**f**) β = 0.7, η = 0.7 (**g**) β = 0.7, η = 0.7 (**h**) β = 0.7, η = 0.7

Figure 21. Variation trends of each subject's profit when (η, β) = (0.5, 0.5) (0.7, 0.7). The figures show the profit changes over time for the supply chain, manufacturer's online channel, manufacturer's offline channel and retailer changes respectively.



(**a**) when $\mathcal{E}$ =0.5 (**b**) when $\mathcal{E}$ =1.5

Figure 22. Variation trends of product quality when ε = 0.5 and 1.5.



(**a**) when ε =0.5 (**b**) when ε =1.5

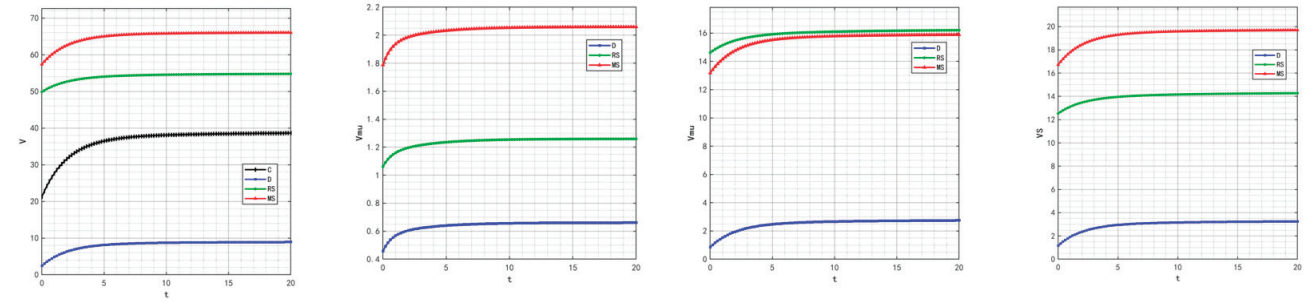
Figure 23. Variation trends of product goodwill when ε = 0.5 and 1.5.

By comparing Figures 12, 14–16 and 24, it is found that when the system's attenuation coefficient ε increases, the profit ranking order of each subject under different models remains unchanged, but the absolute profit values show differentiated fluctuations—the profit of manufacturers' online channels shows an upward trend over time, while the profits of other subjects decrease. This is because the core function of the system attenuation coefficient ε is to measure the time-sensitive loss intensity of factors such as traceability information and quality spillover. An increase in ε will accelerate the value attenuation of cross-channel factors relied on by most subjects. However, the interest distribution mechanism of different models is determined by the cooperation rules of the models and is not directly affected by factor attenuation, so the relative profit ranking under each model remains stable. The profit logic of manufacturers' online channels relies more on real-time information flow; although the increase in ε accelerates the attenuation of traditional factors, it also forces the channel to strengthen real-time information-sharing efficiency, enabling it to more accurately capture short-term market demand and reduce information asymmetry costs, ultimately offsetting the negative impact of attenuation and achieving profit growth. In contrast, other subjects such as offline channels and retailers, rely more on long-term factor accumulation and are unable to quickly adapt to the value loss caused by attenuation, leading to a natural decrease in profits.

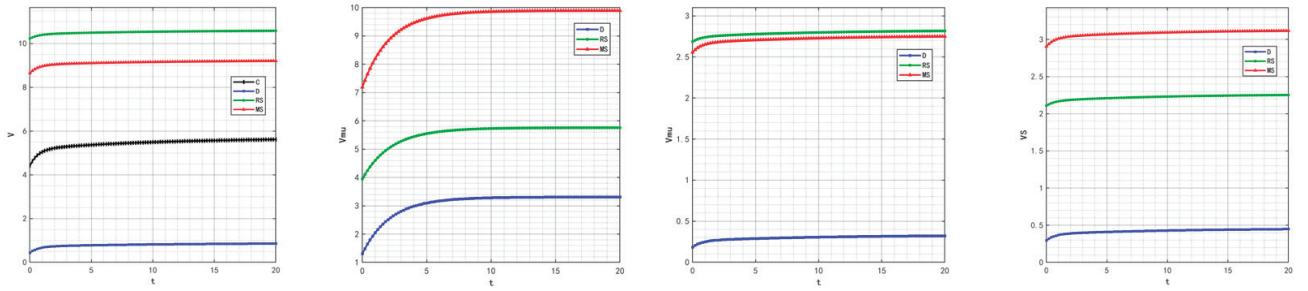
(7) Impact of changes in discount rate on product quality, goodwill, and profit.

Market risk is a key variable affecting the discount rate ρ, and this correlation is particularly significant in the food industry. When the food industry faces risks such as supply chain instability (e.g., sharp fluctuations in raw material prices, logistics disruptions, and insufficient production capacity in key links), policy and regulatory adjustments (e.g., upgrades to food safety standards, stricter environmental protection requirements, and changes in import and export restrictions), or sudden changes in consumer demand, the discount rate ρ tends to show an upward trend. Therefore, ρ is set to 0.15 and 0.3, and the results are shown in Figures 25 and 26.

By comparing the simulation results in Figures 2, 5, 25 and 26, it can be clearly observed that as the discount rate ρ gradually rises, the evolution of product quality shows obvious model heterogeneity—under the manufacturer-led cooperation model, its quality ranking continuously drops from the initial second place to the bottom, and the absolute quality value also decreases simultaneously. This change logic also holds in the product goodwill dimension, but there are significant differences in numerical trends: the goodwill ranking of the manufacturer-led model also falls from the second place to the bottom, but its absolute goodwill value shows a gradual upward trend.

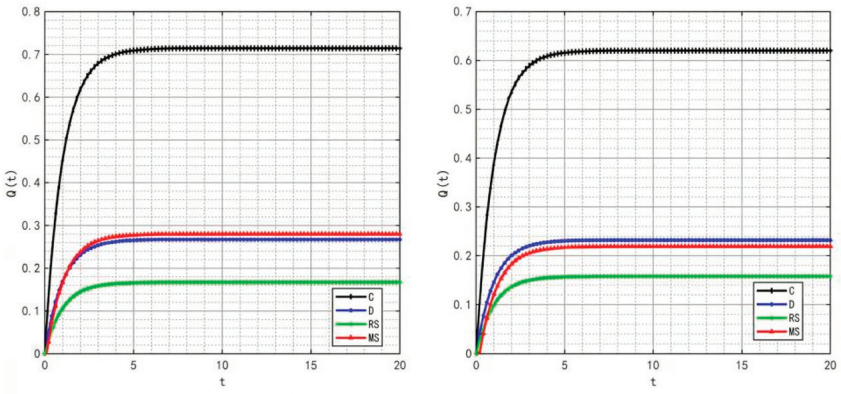


(**a**) ε = 0.5 (**b**) ε = 0.5 (**c**) ε = 0.5 (**d**) ε = 0.5



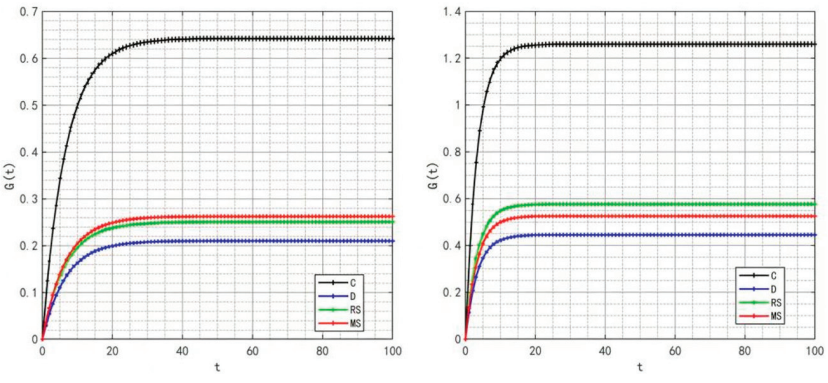
(**e**) ε = 1.5 (**f**) ε = 1.5 (**g**) ε = 1.5 (**h**) ε = 1.5

Figure 24. Variation trends of each subject's profit when ε = 0.5 and 1.5. The figures show the profit changes over time for the supply chain, manufacturer's online channel, manufacturer's offline channel and retailer changes respectively.



(**a**) ρ = 0.15 (**b**) ρ = 0.3

Figure 25. Variation trends of product quality when ρ = 0.15 and 0.3.



(**a**) ρ = 0.15 (**b**) ρ = 0.3

Figure 26. Variation trends of product goodwill when ρ = 0.15 and 0.3.

The root cause of this phenomenon lies in the "two-way mechanism" of the discount rate increase on the decision-making of food industry subjects: from the product quality

dimension, a higher discount rate means manufacturers assign lower weights to long-term returns. Improvements in food quality rely on long-cycle investments, such as raw material control and process upgrades, and the discounted present value of future returns from such investments will shrink as the discount rate rises. Therefore, manufacturers will take the initiative to reduce quality investment, directly leading to a simultaneous decline in quality ranking and numerical value. From the product goodwill dimension, although the increase in the discount rate also weakens manufacturers' emphasis on long-term goodwill accumulation, goodwill in the food industry is directly linked to market trust. To avoid consumer loss caused by quality decline, manufacturers will switch to short-term, effective goodwill maintenance methods. The costs of such investments can be covered by current returns, and the effects can be quickly reflected in goodwill statistics, thus forming a differentiated performance of falling rankings but rising numerical values.

By comparing Figures 12, 14–16 and 27, it is found that when the discount rate ρ increases, the profit ranking order of each subject under different models remains unchanged, and only the absolute profit values show a downward trend. This is because the discount rate reflects the time value discount of future returns by subjects—the higher ρ is, the lower the weight of future returns in current profit accounting. The relative profit ranking of supply chain models is determined by their cooperation mechanisms and is not directly affected by the logic of time value discount, so the relative strength of profits under different models remains stable. At the same time, each subject's profit includes a certain proportion of expected future returns (e.g., quality premiums from long-term cooperation and subsequent returns from channel collaboration). The increase in ρ will amplify the discount rate of these future returns, leading to a general decrease in the absolute current profit values of all subjects. However, this discount is a "same-logic weakening" of profits under all models and does not change the inherent interest distribution characteristics of different models, so the relative profit ranking remains unchanged.

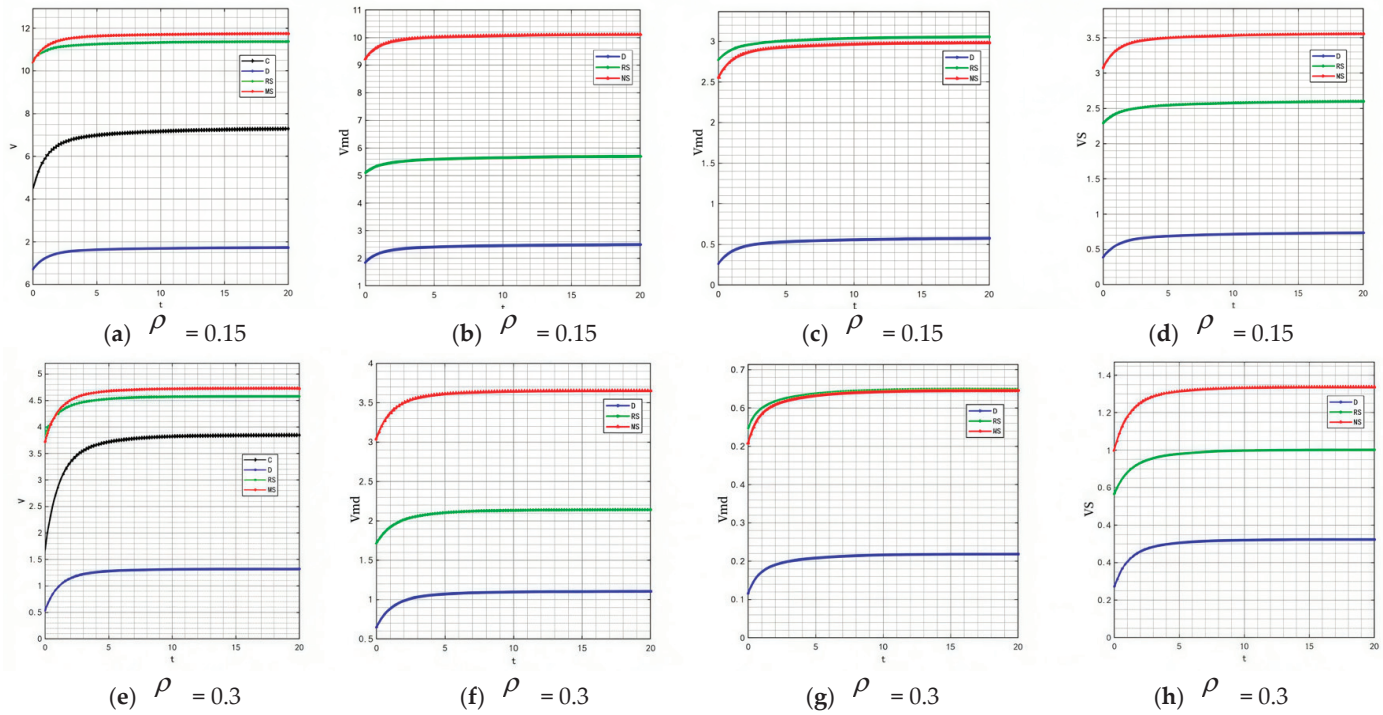


Figure 27. Variation trends of each subject's profit when ρ = 0.15 and 0.3. The figures show the profit changes over time for the supply chain, manufacturer's online channel, manufacturer's offline channel and retailer changes respectively.

7. Conclusions and Limitations

(1) Conclusions

This paper builds a two-level supply chain model of prefabricated food, which covers manufacturers and retailers, integrates an online and offline two-channel sales model, and systematically explores the coordination mechanism and profit distribution of traceability information-sharing behavior. By constructing a game model in the context of centralized decision-making, decentralized decision-making, and two kinds of cost-sharing contracts, the equilibrium results of different decision-making models are compared and analyzed, and the research conclusions are verified by an example analysis. The following research results are formed: (1) The centralized decision-making model is the best choice to improve the market share of the prefabricated food supply chain. Although the cost-sharing contract model can improve the growth rate of market share, there is still a gap between the centralized decision-making model and the cost-sharing contract model. (2) What is different from the traditional cognition is that the centralized decision-making model is not the model with the largest profit, but the cost-sharing decision led by the manufacturer has basically achieved Pareto optimality, among which the most important reasons may be the insufficient incentive mechanism of the centralized decision-making model, the high coordination cost, and the uneven distribution of benefits. (3) Manufacturer's offline channel traceability information-sharing behavior has a more severe impact on profit than online channels. (4) In the market environment of asymmetric information, consumers have difficulty directly judging the quality of pre-made food, so the impact of goodwill on profit is more prominent. (5) When the coefficient measuring the intensity of traceability information sharing's impact on product quality across manufacturers' online and offline channels increases, only under the retailer-led cooperation model do product quality and goodwill exhibit a fluctuating trend of "rising from the bottom to the second place and then falling back to the bottom," while the profits of all subjects increase simultaneously. (6) As the system attenuation coefficient increases, the evolution of product quality and goodwill under different cooperation models shows significant differences: under the retailer-led model, quality fluctuates in the trajectory of "bottom $\rightarrow$ second place $\rightarrow$ bottom"; under the manufacturer-led model, goodwill fluctuates as "second place $\rightarrow$ bottom $\rightarrow$ third place"; and the goodwill under the centralized model remains optimal at all times. In terms of profits, the profits of manufacturers' online channels increase over time, while those of other subjects decrease. (7) When the discount rate rises, the manufacturer-led model presents distinct characteristics: both the ranking and absolute value of product quality decline synchronously, the ranking of goodwill falls, but its absolute value rises against the trend, the evolution of product quality and goodwill shows obvious model heterogeneity, and the profits of all subjects generally decrease.

(2) Management Inspiration

To sum this up, the high-quality development of the prefabricated food supply chain needs to build a win–win ecosystem with traceability information sharing as the core link and fair distribution as the collaborative basis.

The sharing of traceability information is the key to strengthening the product quality defense line and enhancing the trust of consumers. The core enterprise needs to play a leading role. On the one hand, each entity should actively enhance the power of online and offline channel traceability information-sharing on product quality, opening up the whole channel traceability data link, and strengthening the guiding role of traceability information on production, processing, and quality control. This not only simultaneously enhances the product quality and brand reputation under each cooperation mode, but also realizes the general growth of profit of each entity in the supply chain. It is the core of leveraging the

positive cycle of "quality–goodwill–profit". On the other hand, through the cost-sharing contract, the initiative of upstream and downstream is mobilized, a differentiated allocation proportion is set according to contribution degree, and periodic incentives such as "initial advance + later profit return" are supplemented to break down the information barrier.

The government can also introduce targeted subsidy policies to lower the coordination threshold. Provide 20–30% subsidy for sourcing equipment procurement of small- and medium-sized suppliers and a one-time start-up capital subsidy for cross-link information platform construction led by core enterprises. At the same time, based on the Shapley value method, the profit distribution can be quantified according to the marginal contribution of each entity to the alliance, the quality risk cost can be bound with the incremental revenue, and transparent accounting can be realized by means of a third-party platform, so as to balance the multi-interests of core enterprises, suppliers, processors, distributors, and the like, and avoid weakening the cooperative power due to the imbalance of cost or revenue.

The enterprise also needs to establish a flexible cooperation model adaptation mechanism: considering that the comprehensive performance of "quality–goodwill–profit" of different cooperation models is significantly differentiated from environmental variables such as quality attenuation coefficient and industry discount rate change; flexible switching rules between models should be established (for example, the cooperation model should be dynamically adjusted based on the critical threshold of attenuation coefficient and discount rate), and meanwhile, information and resources interfaces between different models should be opened up, so as to eliminate the connection barriers of model switching and avoid the efficiency loss caused by operation faults.

In addition, enterprises need to combine the transparency of traceability information with the construction of brand goodwill in-depth, accumulate brand assets by virtue of the high-quality image of traceability, and enhance market competitiveness; at the industry level, government and enterprise should work together to speed up the establishment of a unified traceability platform, formulate standardized information standards, and address the root causes of information asymmetry; the regulatory authorities need to strengthen the whole chain of quality supervision to protect the rights and interests of consumers, and to ultimately promote the industry to move from decentralized competition to win–win cooperation and achieve standardized and high-quality development.

(3) Inadequacy

Although this research has formed corresponding research results in the aspects of model construction and analysis, there is still room for optimization. First, due to the convenience of solving the model, the research assumes that the supply chain members have complete information, and the case analysis relies on hypothetical parameters. This setting limits the practical application scope of the model to a certain extent. Secondly, the current research focuses on the category of secondary supply chains, and the follow-up can extend the exploration to more complex multi-level supply chain situations or include more participants, such as logistics providers and e-commerce platforms, so as to more comprehensively analyze the issue of traceability information sharing in prefabricated food supply chains. Thirdly, due to the complexity of the solving process of differential games, this study only constructs a two-level supply chain analysis framework composed of manufacturers and retailers, and does not include the key constraints commonly found in realistic scenarios, such as asymmetric information, limited rationality of decision makers, lack of trust among cooperating agents, etc. Subsequent research can further enrich the hypothetical dimensions of the model, and by introducing elements such as multi-level supply chain structure, dynamic information interaction mechanism, and limited rational decision-making criteria, an analysis paradigm that is closer to the real market environment can be constructed.

Author Contributions: Conceptualization, J.Z., X.S. and H.Z.; methodology, J.Z. and X.S.; software, J.Z.; validation, J.Z.; formal analysis, J.Z.; investigation, J.Z.; writing—original draft, J.Z.; writing—review and editing, J.Z., X.S. and H.Z.; visualization, X.S.; supervision, H.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Proof of Proposition 1. At this stage, the optimal control problem satisfies the HJB equation:

$$\rho V_1 = \max_{Tm,Ts}\left\{\begin{array}{l}\pi_u^m[\xi\alpha+\mu Q(t)+\kappa G(t)]+\pi_d^m[(1-\xi)\alpha+\mu Q(t)+\kappa G(t)]+\pi^s[(1-\xi)\alpha+\mu Q(t)\\ +\kappa G(t)]-\frac{\gamma_1}{2}Tm_{u1}{}^2-\frac{\gamma_2}{2}Tm_{d1}{}^2-\frac{\gamma_3}{2}Ts_1{}^2+V_1'(Q)[\eta Tm_{u1}(t)+\beta Tm_{d1}(t)-\varepsilon Q]\\ +V_1'(G)[\phi Q(t)+\theta Ts_1(t)-\varphi G(t)]\end{array}\right\} \quad (A1)$$

Taking the first-order partial derivatives of the right-hand side of Equation (A1) with respect to Tm_{u1}, Tm_{d1}, and Ts_1, setting them equal to zero yields the following:

$$Tm_{u1}=\frac{\eta V_1'(Q)}{\gamma_1};Tm_{d1}=\frac{\beta V_1'(Q)}{\gamma_2};Ts_1=\frac{\theta V_1'(G)}{\gamma_3} \quad (A2)$$

Substituting Equation (A2) into the HJB equation yields the following:

$$\rho V_1=\left\{\begin{array}{l}[(\pi_u^m+\pi_d^m+\pi^s)\mu-\varepsilon V_1'(Q)+\phi V_1'(G)]Q+[(\pi_u^m+\pi_d^m+\pi^s)\kappa-\varphi V_1'(G)]G\\ +[\pi_u^m\xi\alpha+\pi_d^m\alpha-\pi_d^m\xi\alpha+\pi^s\alpha-\pi^s\xi\alpha]+\frac{\eta^2V_1'^2(Q)}{2\gamma_1}+\frac{\beta^2V_1'^2(Q)}{2\gamma_2}+\frac{\theta^2V_1'^2(G)}{2\gamma_3}\end{array}\right. \quad (A3)$$

Let $V_1=e_1Q+e_2G+e_3$, where e_1, e_2, and e_3 are constant. Thus, $V_1'(Q)=e_1$, and $V_1'(G)=e_2$. Substituting into Equation (A3) yields the following:

$$\begin{array}{l}e_1=\frac{(\pi_u^m+\pi_d^m+\pi^s)\mu}{\rho+\varepsilon}+\frac{(\pi_u^m+\pi_d^m+\pi^s)\phi\kappa}{(\rho+\varepsilon)(\rho+\varphi)};e_2=\frac{(\pi_u^m+\pi_d^m+\pi^s)\kappa}{\rho+\varphi}\\ e_3=\left\{\begin{array}{l}\frac{[\pi_u^m\xi+\pi_d^m-\pi_d^m\xi+\pi^s-\pi^s\xi]\alpha}{\rho}+\frac{\eta^2+\beta^2}{2\rho(\gamma_1+\gamma_2)}\left[\frac{(\pi_u^m+\pi_d^m+\pi^s)\mu}{\rho+\varepsilon}+\frac{(\pi_u^m+\pi_d^m+\pi^s)\phi\kappa}{(\rho+\varepsilon)(\rho+\varphi)}\right]^2\\ +\frac{\theta^2}{2\rho\gamma_3}\left[\frac{(\pi_u^m+\pi_d^m+\pi^s)\kappa}{\rho+\varphi}\right]^2\end{array}\right.\end{array} \quad (A4)$$

Substituting Equation (A4) into Equation (A2) yields the static feedback Nash equilibrium under the centralized decision-making scenario:

$$\left\{\begin{array}{l}Tm_{u1}=\frac{(\pi_u^m+\pi_d^m+\pi^s)\eta\mu}{\gamma_1(\rho+\varepsilon)}+\frac{(\pi_u^m+\pi_d^m+\pi^s)\eta\phi\kappa}{\gamma_1(\rho+\varepsilon)(\rho+\varphi)};\\ Tm_{d1}=\frac{(\pi_u^m+\pi_d^m+\pi^s)\beta\mu}{\gamma_2(\rho+\varepsilon)}+\frac{(\pi_u^m+\pi_d^m+\pi^s)\beta\phi\kappa}{\gamma_2(\rho+\varepsilon)(\rho+\varphi)};\\ Ts_1=\frac{(\pi_u^m+\pi_d^m+\pi^s)\theta\kappa}{\gamma_3(\rho+\varphi)}\end{array}\right. \quad (A5)$$

At this point, the optimal trajectories of the prefabricated food quality $Q(t)$ and goodwill $G(t)$ are as follows:

$$Q_1(t)=\left\{\begin{array}{l}\frac{(\pi_u^m+\pi_d^m+\pi^s)(\eta^2\gamma_2+\beta^2\gamma_1)\mu}{\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)}+\frac{(\pi_u^m+\pi_d^m+\pi^s)(\eta^2\gamma_2+\beta^2\gamma_1)\phi\kappa}{\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)(\rho+\varphi)}\\ +\left\{Q_0-\frac{(\pi_u^m+\pi_d^m+\pi^s)(\eta^2\gamma_2+\beta^2\gamma_1)\mu}{\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)}-\frac{(\pi_u^m+\pi_d^m+\pi^s)(\eta^2\gamma_2+\beta^2\gamma_1)\phi\kappa}{\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)(\rho+\varphi)}\right\}e^{-\varepsilon t}\end{array}\right. \quad (A6)$$

$$G_1(t) = \begin{cases} \frac{(\pi_u^m+\pi_d^m+\pi^s)(\eta^2\gamma_2+\beta^2\gamma_1)\mu\phi}{\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)} + \frac{(\pi_u^m+\pi_d^m+\pi^s)(\eta^2\gamma_2+\beta^2\gamma_1)\phi^2\kappa}{\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)(\rho+\varphi)} \\ +\frac{(\pi_u^m+\pi_d^m+\pi^s)\theta^2\kappa}{\varphi\gamma_3(\rho+\varphi)} + \left\{ G_0 - \frac{(\pi_u^m+\pi_d^m+\pi^s)(\eta^2\gamma_2+\beta^2\gamma_1)\mu\phi}{\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)} \right. \\ \left. -\frac{(\pi_u^m+\pi_d^m+\pi^s)(\eta^2\gamma_2+\beta^2\gamma_1)\phi^2\kappa}{\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)(\rho+\varphi)} - \frac{(\pi_u^m+\pi_d^m+\pi^s)\theta^2\kappa}{\varphi\gamma_3(\rho+\varphi)} \right\} e^{-\varphi t} \end{cases} \tag{A7}$$

At this point, the optimal value function of the food supply chain is as follows:

$$V_1 = \begin{cases} [\frac{(\pi_u^m+\pi_d^m+\pi^s)\mu}{\rho+\varepsilon} + \frac{(\pi_u^m+\pi_d^m+\pi^s)\phi\kappa}{(\rho+\varepsilon)(\rho+\varphi)}]Q_1(t) + [\frac{(\pi_u^m+\pi_d^m+\pi^s)\kappa}{\rho+\varphi}]G_1(t) \\ +\frac{[\pi_u^m\xi+\pi_d^m-\pi_d^m\xi+\pi^s-\pi^s\xi]\alpha}{\rho} + \frac{\eta^2+\beta^2}{2\rho(\gamma_1+\gamma_2)}[\frac{(\pi_u^m+\pi_d^m+\pi^s)\mu}{\rho+\varepsilon} + \frac{(\pi_u^m+\pi_d^m+\pi^s)\phi\kappa}{(\rho+\varepsilon)(\rho+\varphi)}]^2 \\ +\frac{\theta^2}{2\rho\gamma_3}[\frac{(\pi_u^m+\pi_d^m+\pi^s)\kappa}{\rho+\varphi}]^2 \end{cases} \tag{A8}$$

□

Proof of Proposition 2. At this stage, the optimal control problem satisfies the HJB equation as follows:

$$\rho Vm_{u2} = \max_{Tm_{u2}} \left\{ \begin{array}{l} \pi_u^m[\xi\alpha + \mu Q(t) + \kappa G(t)] - \frac{\gamma_1}{2} Tm_{u2}{}^2(t) + V'm_{u2}(Q)[\eta Tm_{u2}(t) + \beta Tm_{d2}(t) \\ -\varepsilon Q(t)] + V'm_{u2}(G)[\phi Q(t) + \theta Ts_2(t) - \varphi G(t)] \end{array} \right\} \tag{A9}$$

$$\rho Vm_{d2} = \max_{Tm_{d2}} \left\{ \begin{array}{l} \pi_d^m[(1-\xi)\alpha + \mu Q(t) + \kappa G(t)] - \frac{\gamma_2}{2} Tm_{d2}{}^2(t) + V'm_{d2}(Q)[\eta Tm_{u2}(t) + \beta Tm_{d2}(t) \\ -\varepsilon Q(t)] + V'm_{d2}(G)[\phi Q(t) + \theta Ts_2(t) - \varphi G(t)] \end{array} \right\} \tag{A10}$$

$$\rho Vs_2 = \max_{Tm_{s2}} \left\{ \begin{array}{l} \pi^s[(1-\xi)\alpha + \mu Q(t) + \kappa G(t)] - \frac{\gamma_3}{2} Ts_2{}^2(t) + V's_2(Q)[\eta Tm_{u2}(t) + \beta Tm_{d2}(t) \\ -\varepsilon Q(t)] + V's_2(G)[\phi Q(t) + \theta Ts_2(t) - \varphi G(t)] \end{array} \right\} \tag{A11}$$

Taking the first-order partial derivatives of Equations (A9)–(A11) with respect to Tm_{u2}, Tm_{d2}, and Ts_2, and setting them equal to zero yields the following:

$$Tm_{u2} = \frac{\eta V'm_{u2}(Q)}{\gamma_1}; Tm_{d2} = \frac{\beta V'm_{d2}(Q)}{\gamma_2}; Ts_2 = \frac{\theta V's_2(G)}{\gamma_3} \tag{A12}$$

Substituting Equation (A12) into the HJB equation and simplifying yields the following:

$$\rho Vm_{u2} = \begin{cases} [\pi_u^m\mu - \varepsilon V'm_{u2}(Q) + \phi V'm_{u2}(G)]Q + [\pi_u^m\kappa - \varphi V'm_{u2}(G)] \\ +\pi_u^m\xi\alpha + \frac{\eta^2 V'm_{u2}{}^2(Q)}{2\gamma_1} + \frac{\beta^2 V'm_{d2}(Q)V'm_{u2}(Q)}{\gamma_2} + \frac{\theta^2 V'm_{u2}(G)V's_2(G)}{\gamma_3} \end{cases} \tag{A13}$$

$$\rho Vm_{d2} = \begin{cases} [\pi_d^m\mu - \varepsilon V'm_{d2}(Q) + \phi V'm_{d2}(G)]Q + [\pi_d^m\kappa - \varphi V'm_{d2}(G)]G \\ +\pi_d^m(1-\xi)\alpha + \frac{\eta^2 V'm_{u2}(Q)V'm_{d2}(Q)}{\gamma_1} + \frac{\beta^2 V'm_{d2}{}^2(Q)}{2\gamma_2} + \frac{\theta^2 V'm_{d2}(G)V's_2(G)}{\gamma_3} \end{cases} \tag{A14}$$

$$\rho Vs_2 = \begin{cases} [\pi^s\mu - \varepsilon V's_2(Q) + \phi V's_2(G)]Q + [\pi^s\kappa - \varphi V's_2(G)]G \\ +\pi^s(1-\xi)\alpha + \frac{\eta^2 V's_2(Q)V'm_{u2}(Q)}{\gamma_1} + \frac{\beta^2 V'm_{d2}(Q)V's_2(Q)}{\gamma_2} + \frac{\theta^2 V's_2{}^2(G)}{2\gamma_3} \end{cases} \tag{A15}$$

Let $Vm_{u2} = b_1Q + b_2G + b_3$, $Vm_{d2} = b_4Q + b_5G + b_6$, $Vs_2 = b_7Q + b_8G + b_9$, where b_1, b_2, b_3, b_4, b_5, b_6, b_7, b_8, and b_9 are constants. Thus, substituting $V'm_{u2}(Q) = b_1$,

$V'm_{u2}(G) = b_2$, $V'm_{d2}(Q) = b_4$, $V'm_{u2}(G) = b_5$, $V's_2(Q) = b_7$, and $V's_2(G) = b_8$ into Equations (A13)–(A15) yields the following:

$$\begin{cases} b_1 = \frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa\phi}{(\rho+\varepsilon)(\rho+\varphi)}; b_2 = \frac{\pi_u^m \kappa}{\rho+\varphi}; \\ b_3 = \begin{cases} \frac{\eta^2}{2\rho\gamma_1}[\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa\phi}{(\rho+\varepsilon)(\rho+\varphi)}]^2 + \frac{\beta^2}{\rho\gamma_2}[\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa\phi}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{\pi_d^m \mu}{\rho+\varepsilon} + \frac{\pi_d^m \kappa\phi}{(\rho+\varepsilon)(\rho+\varphi)}] \\ + \frac{\theta^2 \pi_u^m \pi^s \kappa^2}{\rho\gamma_3(\rho+\varphi)^2} + \frac{\pi_u^m \xi\alpha}{\rho} \end{cases}; \\ b_4 = \frac{\pi_d^m \mu}{\rho+\varepsilon} + \frac{\pi_d^m \kappa\phi}{(\rho+\varepsilon)(\rho+\varphi)}; b_5 = \frac{\pi_d^m \kappa}{\rho+\varphi}; \\ b_6 = \begin{cases} \frac{\eta^2}{\rho\gamma_1}[\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa\phi}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{\pi_d^m \mu}{\rho+\varepsilon} + \frac{\pi_d^m \kappa\phi}{(\rho+\varepsilon)(\rho+\varphi)}] + \frac{\beta^2}{2\rho\gamma_2}[\frac{\pi_d^m \mu}{\rho+\varepsilon} + \frac{\pi_d^m \kappa\phi}{(\rho+\varepsilon)(\rho+\varphi)}]^2 \\ + \frac{\theta^2 \pi_d^m \pi^s \kappa^2}{\rho\gamma_3(\rho+\varphi)^2} + \frac{\pi_d^m (1-\xi)\alpha}{\rho} \end{cases}; \\ b_7 = \frac{\pi^s \mu}{\rho+\varepsilon} + \frac{\pi^s \kappa\phi}{(\rho+\varepsilon)(\rho+\varphi)}; b_8 = \frac{\pi^s \kappa}{\rho+\varphi}; \\ b_9 = \begin{cases} \frac{\eta^2}{\rho\gamma_1}[\frac{\pi^s \mu}{\rho+\varepsilon} + \frac{\pi^s \kappa\phi}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa\phi}{(\rho+\varepsilon)(\rho+\varphi)}] \\ + \frac{\beta^2}{\rho\gamma_2}[\frac{\pi_d^m \mu}{\rho+\varepsilon} + \frac{\pi_d^m \kappa\phi}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{\pi^s \mu}{\rho+\varepsilon} + \frac{\pi^s \kappa\phi}{(\rho+\varepsilon)(\rho+\varphi)}] + \frac{1}{2\rho}(\frac{\pi^s \kappa}{\rho+\varphi})^2 + \frac{\pi^s (1-\xi)\alpha}{\rho} \end{cases}; \end{cases} \tag{A16}$$

Substituting Equation (A16) into Equation (A12) yields the static feedback Nash equilibrium for the online channel, offline channel, and retailer of prefabricated foods under the decentralized decision-making scenario:

$$Tm_{u2} = \frac{\eta}{\gamma_1}[\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa\phi}{(\rho+\varepsilon)(\rho+\varphi)}], Tm_{d2} = \frac{\beta}{\gamma_2}[\frac{\pi_d^m \mu}{\rho+\varepsilon} + \frac{\pi_d^m \kappa\phi}{(\rho+\varepsilon)(\rho+\varphi)}], Ts_2 = \frac{\theta\pi^s \kappa}{\gamma_3(\rho+\varphi)} \tag{A17}$$

At this point, the optimal trajectories of prefabricated food quality $Q(t)$ and goodwill $G(t)$ are as follows:

$$Q_2(t) = \begin{cases} [\frac{(\pi_u^m \eta^2\gamma_2 + \pi_d^m \beta^2\gamma_1)\mu}{\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)} + \frac{(\pi_u^m \eta^2\gamma_2 + \pi_d^m \beta^2\gamma_1)\kappa\phi}{\varepsilon\gamma_1\gamma(\rho+\varepsilon)(\rho+\varphi)}] \\ + \left\{ Q_0 - \frac{(\pi_u^m \eta^2\gamma_2 + \pi_d^m \beta^2\gamma_1)\mu}{\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)} - \frac{(\pi_u^m \eta^2\gamma_2 + \pi_d^m \beta^2\gamma_1)\kappa\phi}{\varepsilon\gamma_1\gamma(\rho+\varepsilon)(\rho+\varphi)} \right\} e^{-\varepsilon t} \end{cases} \tag{A18}$$

$$G_2(t) = \begin{cases} \frac{(\pi_u^m \eta^2\gamma_2 + \pi_d^m \beta^2\gamma_1)\phi\mu}{\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)} + \frac{(\pi_u^m \eta^2\gamma_2 + \pi_d^m \beta^2\gamma_1)\kappa\phi^2}{\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)(\rho+\varphi)} + \frac{\theta^2\pi^s\kappa}{\varphi\gamma_3(\rho+\varphi)} \\ + \left\{ G_0 - \frac{(\pi_u^m \eta^2\gamma_2 + \pi_d^m \beta^2\gamma_1)\phi\mu}{\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)} - \frac{(\pi_u^m \eta^2\gamma_2 + \pi_d^m \beta^2\gamma_1)\kappa\phi^2}{\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)(\rho+\varphi)} - \frac{\theta^2\pi^s\kappa}{\varphi\gamma_3(\rho+\varphi)} \right\} e^{-\varphi t} \end{cases} \tag{A19}$$

Under this Nash equilibrium scenario, the optimal value functions of the prefabricated food manufacturer's online channel, offline channel, retailer, and the overall food supply chain are as follows:

$$Vm_{u2} = \begin{cases} [\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa\phi}{(\rho+\varepsilon)(\rho+\varphi)}]Q_2(t) + \frac{\pi_u^m \kappa}{\rho+\varphi}G_2(t) + \frac{\eta^2}{2\gamma_1}[\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa\phi}{(\rho+\varepsilon)(\rho+\varphi)}]^2 \\ + \frac{\beta^2}{\gamma_2}[\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa\phi}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{\pi_d^m \mu}{\rho+\varepsilon} + \frac{\pi_d^m \kappa\phi}{(\rho+\varepsilon)(\rho+\varphi)}] + \frac{\theta^2\pi_u^m\pi^s\kappa^2}{\gamma_3(\rho+\varphi)^2} + \pi_u^m \xi\alpha \end{cases} \tag{A20}$$

$$Vm_{d2} = \begin{cases} [\frac{\pi_d^m \mu}{\rho+\varepsilon} + \frac{\pi_d^m \kappa\phi}{(\rho+\varepsilon)(\rho+\varphi)}]Q_2(t) + \frac{\pi_d^m \kappa}{\rho+\varphi}G_2(t) + \frac{\theta^2\pi_d^m\pi^s\kappa^2}{\gamma_3(\rho+\varphi)^2} + \pi_d^m(1-\xi)\alpha \\ + \frac{\eta^2}{\gamma_1}[\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa\phi}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{\pi_d^m \mu}{\rho+\varepsilon} + \frac{\pi_d^m \kappa\phi}{(\rho+\varepsilon)(\rho+\varphi)}] + \frac{\beta^2}{2\gamma_2}[\frac{\pi_d^m \mu}{\rho+\varepsilon} + \frac{\pi_d^m \kappa\phi}{(\rho+\varepsilon)(\rho+\varphi)}]^2 \end{cases} \tag{A21}$$

$$Vs_2 = \begin{cases} [\frac{\pi^s \mu}{\rho+\varepsilon} + \frac{\pi^s \kappa\phi}{(\rho+\varepsilon)(\rho+\varphi)}]Q_2(t) + \frac{\pi^s \kappa}{\rho+\varphi}G(t) + \frac{1}{2}(\frac{\pi^s \kappa}{\rho+\varphi})^2 + \pi^s(1-\xi)\alpha \\ + \frac{\eta^2}{\gamma_1}[\frac{\pi^s \mu}{\rho+\varepsilon} + \frac{\pi^s \kappa\phi}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa\phi}{(\rho+\varepsilon)(\rho+\varphi)}] \\ + \frac{\beta^2}{\gamma_2}[\frac{\pi_d^m \mu}{\rho+\varepsilon} + \frac{\pi_d^m \kappa\phi}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{\pi^s \mu}{\rho+\varepsilon} + \frac{\pi^s \kappa\phi}{(\rho+\varepsilon)(\rho+\varphi)}] \end{cases} \tag{A22}$$

$$V_2 = \begin{cases} [\frac{(\pi_u^m+\pi_d^m+\pi^s)\mu}{\rho+\varepsilon} + \frac{(\pi_u^m+\pi_d^m+\pi^s)\phi\kappa}{(\rho+\varepsilon)(\rho+\varphi)}]Q_2(t) + \frac{(\pi_u^m+\pi_d^m+\pi^s)\kappa}{\rho+\varphi}G_2(t) \\ +\frac{\eta^2}{2\gamma_1}[\frac{\pi_u^m\mu}{\rho+\varepsilon} + \frac{\pi_u^m\kappa\phi}{(\rho+\varepsilon)(\rho+\varphi)}]^2 + (\frac{\beta^2}{\gamma_2} + \frac{\eta^2}{\gamma_1})[\frac{\pi_u^m\mu}{\rho+\varepsilon} + \frac{\pi_u^m\kappa\phi}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{\pi_d^m\mu}{\rho+\varepsilon} + \frac{\pi_d^m\kappa\phi}{(\rho+\varepsilon)(\rho+\varphi)}] \\ +\frac{\beta^2}{2\gamma_2}[\frac{\pi_d^m\mu}{\rho+\varepsilon} + \frac{\pi_d^m\kappa\phi}{(\rho+\varepsilon)(\rho+\varphi)}]^2 + \frac{\eta^2}{\gamma_1}[\frac{\pi^s\mu}{\rho+\varepsilon} + \frac{\pi^s\kappa\phi}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{\pi_u^m\mu}{\rho+\varepsilon} + \frac{\pi_u^m\kappa\phi}{(\rho+\varepsilon)(\rho+\varphi)}] \\ +\frac{\beta^2}{\gamma_2}[\frac{\pi_d^m\mu}{\rho+\varepsilon} + \frac{\pi_d^m\kappa\phi}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{\pi^s\mu}{\rho+\varepsilon} + \frac{\pi^s\kappa\phi}{(\rho+\varepsilon)(\rho+\varphi)}] + \frac{1}{2}(\frac{\pi^s\kappa}{\rho+\varphi})^2 + \frac{(\pi_u^m+\pi_d^m)\pi^s\theta^2\kappa^2}{\gamma_3(\rho+\varphi)^2} \\ +\pi_u^m\xi\alpha + \pi_d^m(1-\xi)\alpha + \pi^s(1-\xi)\alpha \end{cases} \quad \text{(A23)}$$

□

Proof of Proposition 3. The objective function of the traceable food manufacturer is as follows:

$$R(Tm_{u3}) = \int_0^\infty e^{-\rho t}\Big\{\pi_u^m[\xi\alpha + \mu Q(t) + \kappa G(t)] - \frac{\gamma_1}{2}{Tm_{u3}}^2\Big\}\,dt \quad \text{(A24)}$$

$$R(Tm_{d3}) = \int_0^\infty e^{-pt}\Big\{\pi_d^m[(1-\xi)\alpha + \mu Q(t) + \kappa G(t)] - \frac{\gamma_2}{2}(1-\delta){Tm_{d3}}^2\Big\}dt \quad \text{(A25)}$$

At this point, the optimal control problems of the prefabricated food manufacturer's online and offline channels satisfy the HJB equations:

$$\rho Vm_{u3} = \max_{Tm_{u3}}\begin{Bmatrix} \pi_u^m[\xi\alpha + \mu Q(t) + \kappa G(t)] - \frac{\gamma_1}{2}{Tm_{u3}}^2 + V'm_{u3}(Q)[\eta Tm_{u3}(t) + \beta Tm_{d3}(t) - \varepsilon Q(t)] \\ +V'm_{u3}(G)[\phi Q(t) + \theta Ts_3(t) - \varphi G(t)] \end{Bmatrix} \quad \text{(A26)}$$

$$\rho Vm_{d3} = \max_{Tm_{d3}}\begin{Bmatrix} \pi_d^m[(1-\xi)\alpha + \mu Q(t) + \kappa G(t)] - \frac{\gamma_2}{2}(1-\delta){Tm_{d3}}^2 + V'm_{d3}(Q)[\eta Tm_{u3}(t) + \beta Tm_{d3}(t) \\ -\varepsilon Q(t) + V'm_{d3}(G)[\phi Q(t) + \theta Ts_3(t) - \varphi G(t)] \end{Bmatrix} \quad \text{(A27)}$$

Evidently, Equations (A26) and (A27) are convex functions with respect to Tm_{u3} and Tm_{d3}. Maximizing them under the first-order conditions yields the following:

$$Tm_{u3} = \frac{\eta V'm_{u3}(Q)}{\gamma_1}; Tm_{d3} = \frac{\beta V'm_{d3}(Q)}{\gamma_2(1-\delta)} \quad \text{(A28)}$$

The objective function of the prefabricated food retailer is as follows:

$$R(Ts_3) = \int_0^\infty e^{-\rho t}\Big\{\pi^s[(1-\xi)\alpha + \mu Q(t) + \kappa G(t)] - \frac{\gamma_3}{2}{Ts_3}^2(t) - \frac{\delta\gamma_2}{2}{Tm_{d3}}^2(t)\Big\}dt \quad \text{(A29)}$$

At this point, the optimal control problem of the prefabricated food retailer satisfies the HJB equation:

$$\rho Vs_3 = \max_{Ts_3}\begin{Bmatrix} \pi^s[(1-\xi)\alpha + \mu Q(t) + \kappa G(t)] - \frac{\gamma_3}{2}{Ts_3}^2 - \frac{\delta\gamma_2}{2}{Tm_{d3}}^2 + V's_3(Q)[\eta Tm_{u3}(t) \\ +\beta Tm_{d3}(t) - \varepsilon Q) + V's_3(G)[\phi Q(t) + \theta Ts_3(t) - \varphi G(t)] \end{Bmatrix} \quad \text{(A30)}$$

Substituting Equation (A28) into Equation (A30), then taking the first-order partial derivatives of the right-hand side of Equation (A30) with respect to Ts_3 and δ, and setting them equal to zero, yields the following:

$$Ts_3 = \frac{\theta V's_3(G)}{\gamma_3}; \delta = \frac{2V's_3(Q) - V'm_{d3}(Q)}{2V's_3(Q) + V'm_{d3}(Q)} \quad \text{(A31)}$$

Substituting Equations (A28) and (A31) into the HJB equation and simplifying yields the following:

$$\rho V m_{u3} = \begin{cases} [\pi_u^m \mu - \varepsilon V' m_{u3}(Q) + \phi V' m_{u3}(G)]Q + [\pi_u^m \kappa - \varphi V' m_{u3}(G)]G + \pi_u^m \xi \alpha \\ + \frac{\eta^2 V\prime m_{u3}{}^2(Q)}{2\gamma_1} + \frac{\beta^2 V\prime m_{u3}(Q)[2V' s_3(Q) + V' m_{d3}(Q)]}{2\gamma_2} + \frac{\theta^2 V' s_3(G) V' m_{u3}(G)}{\gamma_3} \end{cases} \tag{A32}$$

$$\rho V m_{d3} = \begin{cases} [\pi_d^m \mu - \varepsilon V' m_{d3}(Q) + \phi V' m_{d3}(G)]Q + [\pi_d^m \kappa - \varphi V' m_{d3}(G)]G + \pi_d^m (1-\xi)\alpha \\ + \frac{\eta^2 V' m_{u3}(Q) V' m_{d3}(Q)}{\gamma_1} + \frac{\beta^2 V' m_{d3}(Q)[2V' s_3(Q) + V' m_{d3}(Q)]}{4\gamma_2} + \frac{\theta^2 V' s_3(G) V' m_{d3}(G)}{\gamma_3} \end{cases} \tag{A33}$$

$$\rho V s_3 = \begin{cases} [\pi^s \mu - \varepsilon V' s_3(Q) + \phi V' s_3(G)]Q + [\pi^s \kappa - \varphi V' s_3(G)]G + \pi^s (1-\xi)\alpha \\ + \frac{\eta^2 V' m_{u3}(Q) V' s_3(Q)}{\gamma_1} + \frac{\beta^2 [2V' s_3(Q) - V' m_{d3}(Q)]^2}{8\gamma_2} + \frac{\theta^2 V' s_3{}^2(G)}{2\gamma_3} \end{cases} \tag{A34}$$

Let $Vm_{u3} = d_1 Q + d_2 G + d_3$, $Vm_{d3} = d_4 Q + d_5 G + d_6$, and $Vs_3 = d_7 Q + d_8 G + d_9$, where d_1, d_2, d_3, d_4, d_5, d_6, d_7, d_8 and d_9 are constants. Thus, substituting $V' m_{u3}(Q) = d_1$, $V' m_{u3}(G) = d_2$, $V' m_{d3}(Q) = d_4$, $V' m_{d3}(G) = d_5$, $V' s_3(Q) = d_7$, and $V' s_3(G) = d_8$ into Equations (A32)–(A34) yields the following:

$$\begin{cases} d_1 = \frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}; d_2 = \frac{\pi_u^m \kappa}{\rho+\varphi} \\ d_3 = \begin{cases} \frac{\pi_u^m \xi \alpha}{\rho} + \frac{\eta^2}{2\rho\gamma_1}[\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}] + \frac{\theta^2 \pi_u^m \pi^s \kappa^2}{\rho\gamma_3(\rho+\varphi)^2} \\ + \frac{\beta^2}{2\rho\gamma_2}[\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{(2\pi^s + \pi_d^m)\mu}{\rho+\varepsilon} + \frac{(2\pi^s + \pi_d^m)\kappa}{(\rho+\varepsilon)(\rho+\varphi)}] \end{cases} \\ d_4 = \frac{\pi_d^m \mu}{\rho+\varepsilon} + \frac{\pi_d^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}; d_5 = \frac{\pi_d^m \kappa}{\rho+\varphi} \\ d_6 = \begin{cases} \frac{\pi_d^m (1-\xi)\alpha}{\rho} + \frac{\eta^2}{\rho\gamma_1}[\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{\pi_d^m \mu}{\rho+\varepsilon} + \frac{\pi_d^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}] \\ + \frac{\beta^2}{4\rho\gamma_2}[\frac{\pi_d^m \mu}{\rho+\varepsilon} + \frac{\pi_d^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{(2\pi^s + \pi_d^m)\mu}{\rho+\varepsilon} + \frac{(2\pi^s + \pi_d^m)\kappa}{(\rho+\varepsilon)(\rho+\varphi)}] + \frac{\theta^2 \pi_d^m \pi^s \kappa^2}{\rho\gamma_3(\rho+\varphi)^2} \end{cases} \\ d_7 = \frac{\pi^s \mu}{\rho+\varepsilon} + \frac{\pi^s \kappa}{(\rho+\varepsilon)(\rho+\varphi)}; d_8 = \frac{\pi^s \kappa}{\rho+\varphi} \\ d_9 = \begin{cases} \frac{\pi^s (1-\xi)\alpha}{\rho} + \frac{\eta^2}{\rho\gamma_1}[\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{\pi^s \mu}{\rho+\varepsilon} + \frac{\pi^s \kappa}{(\rho+\varepsilon)(\rho+\varphi)}] \\ + \frac{\beta^2}{8\rho\gamma_2}[\frac{(2\pi^s - \pi_d^m)\mu}{\rho+\varepsilon} + \frac{(2\pi^s - \pi_d^m)\kappa}{(\rho+\varepsilon)(\rho+\varphi)}] + \frac{\theta^2 \pi^s{}^2 \kappa^2}{2\rho\gamma_3(\rho+\varphi)^2} \end{cases} \end{cases} \tag{A35}$$

Substituting Equation (A35) into Equations (A28) and (A31) yields the static feedback Stackelberg equilibrium under the manufacturer-led scenario:

$$\begin{cases} Tm_{u3} = \frac{\eta}{\gamma_1}[\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}] \\ Tm_{d3} = \frac{\beta(2\pi^s + \pi_d^m)}{2\gamma_2}[\frac{\mu}{\rho+\varepsilon} + \frac{\kappa}{(\rho+\varepsilon)(\rho+\varphi)}] \\ Ts_3 = \frac{\theta \pi^s \kappa}{\gamma_3(\rho+\varphi)} \\ \delta = \frac{2\pi^s - \pi_d^m}{2\pi^s + \pi_d^m} \end{cases} \tag{A36}$$

At this point, the optimal trajectories of product quality $Q(t)$ and goodwill $G(t)$ are as follows:

$$Q_3(t) = \begin{cases} \frac{[2\pi_u^m \eta^2 \gamma_2 + (2\pi^s + \pi_d^m)\beta^2 \gamma_1]\mu}{2\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)} + \frac{[2\pi_u^m \eta^2 \gamma_2 + (2\pi^s + \pi_d^m)\beta^2 \gamma_1]\kappa}{2\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)(\rho+\varphi)} \\ + \left\{ Q_0 - \frac{[2\pi_u^m \eta^2 \gamma_2 + (2\pi^s + \pi_d^m)\beta^2 \gamma_1]\mu}{2\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)} + \frac{[2\pi_u^m \eta^2 \gamma_2 + (2\pi^s + \pi_d^m)\beta^2 \gamma_1]\kappa}{2\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)(\rho+\varphi)} \right\} e^{-\varepsilon t} \end{cases} \tag{A37}$$

$$G_3(t) = \begin{cases} \frac{[2\pi_u^m \eta^2 \gamma_2 + (2\pi^s + \pi_d^m)\beta^2 \gamma_1]\phi\mu}{2\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)} + \frac{[2\pi_u^m \eta^2 \gamma_2 + (2\pi^s + \pi_d^m)\beta^2 \gamma_1]\phi\kappa}{2\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)(\rho+\varphi)} + \frac{\theta^2 \pi^s \kappa}{\varphi\gamma_3(\rho+\varphi)} \\ + \left\{ G_0 - \frac{[2\pi_u^m \eta^2 \gamma_2 + (2\pi^s + \pi_d^m)\beta^2 \gamma_1]\phi\mu}{2\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)} - \frac{\theta^2 \pi^s \kappa}{\varphi\gamma_3(\rho+\varphi)} \right. \\ \left. - \frac{[2\pi_u^m \eta^2 \gamma_2 + (2\pi^s + \pi_d^m)\beta^2 \gamma_1]\phi\kappa}{2\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)(\rho+\varphi)} \right\} e^{-\varphi t} \end{cases} \tag{A38}$$

At this point, the optimal value functions of the prefabricated food manufacturer's online channel, offline channel, retailer, and the overall prefabricated food supply chain are as follows:

$$Vm_{u3} = \begin{cases} [\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}]Q(t) + \frac{\pi_u^m \kappa}{\rho+\varphi} G(t) \\ + \frac{\pi_u^m \xi \alpha}{\rho} + \frac{\eta^2}{2\rho\gamma_1}[\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}] \\ + \frac{\theta^2 \pi_u^m \pi^s \kappa^2}{\rho\gamma_3(\rho+\varphi)^2} + \frac{\beta^2}{2\rho\gamma_2}[\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{(2\pi^s+\pi_d^m)\mu}{\rho+\varepsilon} + \frac{(2\pi^s+\pi_d^m)\kappa}{(\rho+\varepsilon)(\rho+\varphi)}] \end{cases} \tag{A39}$$

$$Vm_{d3} = \begin{cases} [\frac{\pi_d^m \mu}{\rho+\varepsilon} + \frac{\pi_d^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}]Q(t) + \frac{\pi_d^m \kappa}{\rho+\varphi} G(t) + \frac{\pi_d^m (1-\xi)\alpha}{\rho} \\ + \frac{\eta^2}{\rho\gamma_1}[\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{\pi_d^m \mu}{\rho+\varepsilon} + \frac{\pi_d^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}] \\ + \frac{\beta^2}{4\rho\gamma_2}[\frac{\pi_d^m \mu}{\rho+\varepsilon} + \frac{\pi_d^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{(2\pi^s+\pi_d^m)\mu}{\rho+\varepsilon} + \frac{(2\pi^s+\pi_d^m)\kappa}{(\rho+\varepsilon)(\rho+\varphi)}] + \frac{\theta^2 \pi_d^m \pi^s \kappa^2}{\rho\gamma_3(\rho+\varphi)^2} \end{cases} \tag{A40}$$

$$Vs_3 = \begin{cases} [\frac{\pi^s \mu}{\rho+\varepsilon} + \frac{\pi^s \kappa}{(\rho+\varepsilon)(\rho+\varphi)}]Q(t) + \frac{\pi^s \kappa}{\rho+\varphi} G(t) + \frac{\pi^s (1-\xi)\alpha}{\rho} \\ + \frac{\eta^2}{\rho\gamma_1}[\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{\pi^s \mu}{\rho+\varepsilon} + \frac{\pi^s \kappa}{(\rho+\varepsilon)(\rho+\varphi)}] \\ + \frac{\beta^2}{8\rho\gamma_2}[\frac{(2\pi^s-\pi_d^m)\mu}{\rho+\varepsilon} + \frac{(2\pi^s-\pi_d^m)\kappa}{(\rho+\varepsilon)(\rho+\varphi)}] + \frac{\theta^2 {\pi^s}^2 \kappa^2}{2\rho\gamma_3(\rho+\varphi)^2} \end{cases} \tag{A41}$$

$$V_3 = \begin{cases} [\frac{(\pi_u^m+\pi_d^m+\pi^s)\mu}{\rho+\varepsilon} + \frac{(\pi_u^m+\pi_d^m+\pi^s)\kappa}{(\rho+\varepsilon)(\rho+\varphi)}]Q_3(t) + \frac{(\pi_u^m+\pi_d^m+\pi^s)\kappa}{\rho+\varphi} G_3(t) \\ + \frac{\alpha}{\rho}[\pi_u^m \xi + \pi_d^m(1-\xi) + \pi^s(1-\xi)] + \frac{\beta^2}{8\rho\gamma_2}[\frac{(2\pi^s-\pi_d^m)\mu}{\rho+\varepsilon} + \frac{(2\pi^s-\pi_d^m)\kappa}{(\rho+\varepsilon)(\rho+\varphi)}] \\ + \frac{\eta^2}{\rho\gamma_1}[\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}][1 + \frac{(2\pi_d^m+2\pi^s)\mu}{\rho+\varepsilon} + \frac{(2\pi_d^m+2\pi^s)\kappa}{(\rho+\varepsilon)(\rho+\varphi)}] \\ + \frac{\theta^2}{\rho\gamma_3(\rho+\varphi)^2}(\pi_u^m \pi^s \kappa^2 + \pi_d^m \pi^s \kappa^2 + \frac{{\pi^s}^2 \kappa^2}{2}) \\ + \frac{3\beta^2}{4\rho\gamma_2}[\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{(2\pi^s+\pi_d^m)\mu}{\rho+\varepsilon} + \frac{(2\pi^s+\pi_d^m)\kappa}{(\rho+\varepsilon)(\rho+\varphi)}] \end{cases} \tag{A42}$$

□

Proof of Proposition 4. The objective function of the traceable food retailer is as follows:

$$R(Ts_4) = \int_0^\infty e^{-\rho t}\Big\{\pi^s[(1-\xi)\alpha + \mu Q(t) + \kappa G(t)] - \frac{\gamma_3}{2}(1-\omega){Ts_4}^2(t)\Big\}dt \tag{A43}$$

At this point, the optimal control problem of the prefabricated food retailer satisfies the HJB equation:

$$\rho V s_4 = \max_{Ts_4}\left\{ \begin{array}{l} \pi^s[(1-\xi)\alpha + \mu Q(t) + \kappa G(t)] - \frac{\gamma_3}{2}(1-\omega){Ts_4}^2(t) + V's_4(Q)[\eta Tm_{u4}(t) \\ + \beta Tm_{d4}(t) - \varepsilon Q) + V's_4(G)[\phi Q(t) + \theta Ts_4(t) - \varphi G(t)] \end{array} \right\} \tag{A44}$$

Taking the first-order derivative of Equation (A44) with respect to Ts_4 and setting it equal to zero yields the following:

$$Ts_4 = \frac{\theta V's_4(G)}{\gamma_3(1-\omega)} \tag{A45}$$

The objective function of the prefabricated food manufacturer is as follows:

$$R(Tm_{u4}) = \int_0^\infty e^{-\rho t}\Big\{\pi_u^m[\xi\alpha + \mu Q(t) + \kappa G(t)] - \frac{\gamma_1}{2}{Tm_{u4}}^2\Big\}\,dt \tag{A46}$$

$$R(Tm_{d4}) = \int_0^\infty e^{-pt}\Big\{\pi_d^m[(1-\xi)\alpha + \mu Q(t) + \kappa G(t)] - \frac{\gamma_2}{2}{Tm_{d4}}^2(t) - \frac{\omega\gamma_3}{2}{Ts_4}^2(t)\Big\}dt \tag{A47}$$

At this point, the optimal control problem of the prefabricated food manufacturer satisfies the HJB Equation:

$$\rho V m_{u4} = \max_{Tm_{u4}} \left\{ \begin{array}{l} \pi_u^m[\xi\alpha + \mu Q(t) + \kappa G(t)] - \frac{\gamma_1}{2} Tm_{u4}{}^2 + V'm_{u4}(Q)[\eta Tm_{u4}(t) + \beta Tm_{d4}(t) - \varepsilon Q(t)] \\ + V'm_{u4}(G)[\phi Q(t) + \theta Ts_4(t) - \varphi G(t)] \end{array} \right\} \tag{A48}$$

$$\rho V m_{d4} = \max_{Tm_{d4}} \left\{ \begin{array}{l} \pi_d^m[(1-\xi)\alpha + \mu Q(t) + \kappa G(t)] - \frac{\gamma_2}{2} Tm_{d4}{}^2 - \frac{\omega\gamma_3}{2} Ts_4{}^2 + V'm_{d4}(Q)[\eta Tm_{u4}(t) \\ + \beta Tm_{d4}(t) - \varepsilon Q(t)] + V'm_{d4}(G)[\phi Q(t) + \theta Ts_4(t) - \varphi G(t)] \end{array} \right\} \tag{A49}$$

Substituting Equation (A45) into Equations (A48) and (A49), then taking the first-order partial derivatives of the right-hand sides of Equations (A48) and (A49) with respect to Tmu_4, Tmd_4, and ω, and setting them equal to zero, yields the following:

$$Tm_{u4} = \frac{\eta V'm_{u4}(Q)}{\gamma_1}; Tm_{d4} = \frac{\beta V'm_{d4}(Q)}{\gamma_2}; \omega = \frac{2V'm_{d4}(G) - V's_4(G)}{2V'm_{d4}(G) + V's_4(G)} \tag{A50}$$

Substituting Equations (A45) and (A50) into the HJB equation and simplifying yields the following:

$$\rho V m_{u4} = \left\{ \begin{array}{l} [\pi_u^m \mu - \varepsilon V'm_{u4}(Q) + \phi V'm_{u4}(G)]Q + [\pi_u^m \kappa - \varphi V'm_{u4}(G)]G + \pi_u^m \xi\alpha \\ + \frac{\eta^2 V'm_{u4}{}^2(Q)}{2\gamma_1} + \frac{\beta^2 V'm_{u4}(Q)V'm_{d4}(Q)}{\gamma_2} + \frac{\theta^2 V's_4(G)V'm_{u4}(G)}{\gamma_3} \end{array} \right. \tag{A51}$$

$$\rho V m_{d4} = \left\{ \begin{array}{l} [\pi_d^m \mu - \varepsilon V'm_{d4}(Q) + \phi V'm_{d4}(G)]Q + [\pi_d^m \kappa - \varphi V'm_{d4}(G)]G + \pi_d^m (1-\xi)\alpha \\ + \frac{\eta^2 V'm_{u4}(Q)V'm_{d4}(Q)}{\gamma_1} + \frac{\beta^2 V'm_{d4}{}^2(Q)}{2\gamma_2} + \frac{\theta^2[2V'm_{d4}(G)+V's_4(G)]^2}{8\gamma_3} \end{array} \right. \tag{A52}$$

$$\rho V s_4 = \left\{ \begin{array}{l} [\pi^s \mu - \varepsilon V's_4(Q) + \phi V's_4(G)]Q + [\pi^s \kappa - \varphi V's_4(G)]G + \pi^s (1-\xi)\alpha \\ + \frac{\eta^2 V'm_{u4}(Q)V's_4(Q)}{\gamma_1} + \frac{\beta^2 V'm_{d4}(Q)V's_4(Q)}{\gamma_2} + \frac{\theta^2 V's_4(G)[2V'm_{d4}(G)+V's_4(G)]}{4\gamma_2} \end{array} \right. \tag{A53}$$

Let $Vm_{u4} = c_1 Q + c_2 G + c_3$, $Vm_{d4} = c_4 Q + c_5 G + c_6$, and $Vs_4 = c_7 Q + c_8 G + c_9$, where c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_8, and c_9 are constants. Thus, substituting $V'm_{u4}(Q) = c_1$, $V'm_{u4}(G) = c_2$, $V'm_{d4}(Q) = c_4$, $V'm_{d4}(G) = c_5$, $V's_4(Q) = c_7$, and $V's_4(G) = c_8$ into Equations (A51)–(A53) yields the following:

$$\left\{ \begin{array}{l} c_1 = \frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}; c_2 = \frac{\pi_u^m \kappa}{\rho+\varphi} \\ c_3 = \left\{ \begin{array}{l} \frac{\pi_u^m \xi\alpha}{\rho} + \frac{\eta^2}{2\rho\gamma_1}\left[\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}\right]^2 \\ + \frac{\beta^2}{\rho\gamma_2}\left[\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}\right]\left[\frac{\pi_d^m \mu}{\rho+\varepsilon} + \frac{\pi_d^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}\right] + \frac{\theta^2 \pi_u^m \pi^s \kappa^2}{\rho\gamma_3(\rho+\varphi)^2} \end{array} \right. \\ c_4 = \frac{\pi_d^m \mu}{\rho+\varepsilon} + \frac{\pi_d^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}; c_5 = \frac{\pi_d^m \kappa}{\rho+\varphi} \\ c_6 = \left\{ \begin{array}{l} \frac{\pi_d^m (1-\xi)\alpha}{\rho} + \frac{\eta^2}{\rho\gamma_1}\left[\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}\right]\left[\frac{\pi_d^m \mu}{\rho+\varepsilon} + \frac{\pi_d^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}\right] \\ + \frac{\beta^2}{2\rho\gamma_2}\left[\frac{\pi_d^m \mu}{\rho+\varepsilon} + \frac{\pi_d^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}\right]^2 + \frac{\theta^2}{8\rho\gamma_3}\left[\frac{(2\pi_d^m + \pi^s)\kappa}{\rho+\varphi}\right]^2 \end{array} \right. \\ c_7 = \frac{\pi^s \mu}{\rho+\varepsilon} + \frac{\pi^s \kappa}{(\rho+\varepsilon)(\rho+\varphi)}; c_8 = \frac{\pi^s \kappa}{\rho+\varphi} \\ c_9 = \left\{ \begin{array}{l} \frac{\pi^s (1-\xi)\alpha}{\rho} + \frac{\eta^2}{\rho\gamma_1}\left[\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}\right]\left[\frac{\pi^s \mu}{\rho+\varepsilon} + \frac{\pi^s \kappa}{(\rho+\varepsilon)(\rho+\varphi)}\right] \\ + \frac{\beta^2}{\rho\gamma_2}\left[\frac{\pi_d^m \mu}{\rho+\varepsilon} + \frac{\pi_d^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}\right]\left[\frac{\pi^s \mu}{\rho+\varepsilon} + \frac{\pi^s \kappa}{(\rho+\varepsilon)(\rho+\varphi)}\right] + \frac{\theta^2 \pi^s \kappa^2 (2\pi_d^m + \pi^s)}{4\rho\gamma_3(\rho+\varphi)^2} \end{array} \right. \end{array} \right. \tag{A54}$$

Substituting Equation (A54) into Equations (A45) and (A50) yields the static feedback Stackelberg equilibrium under the manufacturer-led scenario:

$$\begin{cases} Tm_{u4} = \frac{\eta}{\gamma_1}[\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}]; \\ Tm_{d4} = \frac{\beta}{\gamma_2}[\frac{\pi_d^m \mu}{\rho+\varepsilon} + \frac{\pi_d^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}]; \\ Ts_4 = \frac{\theta\kappa(2\pi_d^m+\pi^s)}{2\gamma_3(\rho+\varphi)}; \\ \omega = \frac{2\pi_d^m-\pi^s}{2\pi_d^m+\pi^s} \end{cases} \tag{A55}$$

At this point, the optimal trajectories of prefabricated food quality $Q(t)$ and goodwill $G(t)$ are as follows:

$$Q_4(t) = \begin{cases} \frac{(\pi_u^m \eta^2\gamma_2+\pi_d^m \beta^2\gamma_1)\mu}{\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)} + \frac{(\pi_u^m \eta^2\gamma_2+\pi_d^m \beta^2\gamma_1)\kappa}{\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)(\rho+\varphi)} \\ +\left\{Q_0 - \frac{(\pi_u^m \eta^2\gamma_2+\pi_d^m \beta^2\gamma_1)\mu}{\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)} - \frac{(\pi_u^m \eta^2\gamma_2+\pi_d^m \beta^2\gamma_1)\kappa}{\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)(\rho+\varphi)}\right\}e^{-\varepsilon t} \end{cases} \tag{A56}$$

$$G_4(t) = \begin{cases} \frac{(\pi_u^m \eta^2\gamma_2+\pi_d^m \beta^2\gamma_1)\mu\phi}{\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)} + \frac{(\pi_u^m \eta^2\gamma_2+\pi_d^m \beta^2\gamma_1)\kappa\phi}{\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)(\rho+\varphi)} + \frac{\theta^2\kappa(2\pi_d^m+\pi^s)}{2\varphi\gamma_3(\rho+\varphi)} \\ +\left\{G_0 - \frac{(\pi_u^m \eta^2\gamma_2+\pi_d^m \beta^2\gamma_1)\mu\phi}{\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)} - \frac{(\pi_u^m \eta^2\gamma_2+\pi_d^m \beta^2\gamma_1)\kappa\phi}{\varphi\varepsilon\gamma_1\gamma_2(\rho+\varepsilon)(\rho+\varphi)} - \frac{\theta^2\kappa(2\pi_d^m+\pi^s)}{2\varphi\gamma_3(\rho+\varphi)}\right\}e^{-\varphi t} \end{cases} \tag{A57}$$

The optimal value functions of the prefabricated food manufacturer's online channel, offline channel, retailer, and the overall prefabricated food supply chain are given as follows:

$$Vm_{u4} = \begin{cases} [\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}]Q(t) + \frac{\pi_u^m \kappa}{\rho+\varphi}G(t) + \frac{\eta^2}{2\rho\gamma_1}[\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}]^2 \\ +\frac{\beta^2}{\rho\gamma_2}[\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{\pi_d^m \mu}{\rho+\varepsilon} + \frac{\pi_d^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}] + \frac{\theta^2\pi_u^m \pi^s\kappa^2}{\rho\gamma_3(\rho+\varphi)^2} + \frac{\pi_u^m \xi\alpha}{\rho} \end{cases} \tag{A58}$$

$$Vm_{d4} = \begin{cases} [\frac{\pi_d^m \mu}{\rho+\varepsilon} + \frac{\pi_d^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}]Q(t) + \frac{\pi_d^m \kappa}{\rho+\varphi}G(t) + \frac{\pi_d^m(1-\xi)\alpha}{\rho} + \frac{\theta^2}{8\rho\gamma_3}[\frac{(2\pi_d^m+\pi^s)\kappa}{\rho+\varphi}]^2 \\ +\frac{\eta^2}{\rho\gamma_1}[\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{\pi_d^m \mu}{\rho+\varepsilon} + \frac{\pi_d^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}] + \frac{\beta^2}{2\rho\gamma_2}[\frac{\pi_d^m \mu}{\rho+\varepsilon} + \frac{\pi_d^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}]^2 \end{cases} \tag{A59}$$

$$Vs_4 = \begin{cases} [\frac{\pi^s \mu}{\rho+\varepsilon} + \frac{\pi^s \kappa}{(\rho+\varepsilon)(\rho+\varphi)}]Q(t) + \frac{\pi^s \kappa}{\rho+\varphi}G(t) + \frac{\pi^s(1-\xi)\alpha}{\rho} \\ +\frac{\eta^2}{\rho\gamma_1}[\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{\pi^s \mu}{\rho+\varepsilon} + \frac{\pi^s \kappa}{(\rho+\varepsilon)(\rho+\varphi)}] \\ +\frac{\beta^2}{\rho\gamma_2}[\frac{\pi_d^m \mu}{\rho+\varepsilon} + \frac{\pi_d^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{\pi^s \mu}{\rho+\varepsilon} + \frac{\pi^s \kappa}{(\rho+\varepsilon)(\rho+\varphi)}] + \frac{\theta^2\pi^s\kappa^2(2\pi_d^m+\pi^s)}{4\rho\gamma_3(\rho+\varphi)^2} \end{cases} \tag{A60}$$

$$V_4 = \begin{cases} [\frac{(\pi_u^m+\pi_d^m+\pi^s)\mu}{\rho+\varepsilon} + \frac{(\pi_u^m+\pi_d^m+\pi^s)\kappa}{(\rho+\varepsilon)(\rho+\varphi)}]Q(t) + \frac{(\pi_u^m+\pi_d^m+\pi^s)\kappa}{\rho+\varphi}G(t) \\ +\frac{\eta^2}{\rho\gamma_1}[\frac{\pi_u^m \mu}{\rho+\varepsilon} + \frac{\pi_u^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{(\pi_u^m+2\pi_d^m+2\pi^s)\mu}{\rho+\varepsilon} + \frac{(\pi_u^m+2\pi_d^m+2\pi^s)\kappa}{(\rho+\varepsilon)(\rho+\varphi)}] \\ +\frac{\beta^2}{\rho\gamma_2}[\frac{\pi_d^m \mu}{\rho+\varepsilon} + \frac{\pi_d^m \kappa}{(\rho+\varepsilon)(\rho+\varphi)}][\frac{(\pi_u^m+\pi_d^m+2\pi^s)\mu}{\rho+\varepsilon} + \frac{(\pi_u^m+\pi_d^m+2\pi^s)\kappa}{(\rho+\varepsilon)(\rho+\varphi)}] \\ +\frac{\theta^2}{\rho\gamma_3(\rho+\varphi)^2}[\pi_u^m \pi^s\kappa^2 + \frac{\pi^s\kappa^2(2\pi_d^m+\pi^s)}{4}] + \frac{\alpha}{\rho}[\pi_u^m \xi + \pi_d^m(1-\xi) + \pi^s(1-\xi)] \end{cases} \tag{A61}$$

□

References

1. Gong, J.; Sun, Y.; Du, H.Y.; Jiang, X. Research on safety risk control of prepared foods from the perspective of supply chain. *Heliyon* **2024**, *10*, e25012. [CrossRef] [PubMed]
2. He, Q.Z.; Zhao, Y.N.; Sun, Y.; Ji, H. Consumer perceptions of pre-prepared food safety risks: Evidence from survey experiments. *J. Sci. Food Agric.* **2025**. [CrossRef]
3. Huang, J.J.; Zhang, M.; Fang, Z.X. Perspectives on Novel Technologies of Processing and Monitoring the Safety and Quality of Prepared Food Products. *Foods* **2023**, *12*, 3052. [CrossRef] [PubMed]
4. Shen, C.; Meng, W.Y.; Chen, X.; Liu, K.; Wu, X.; Yu, Q. Analysis of Factors Affecting Consumers' Perception of Food Safety Risks in the Prepared Food Market. *Foods* **2025**, *14*, 3463. [CrossRef] [PubMed]

5. Huang, J.J.; Zhang, M.; Mujumdar, A.S.; Li, C. AI-based processing of future prepared foods: Progress and prospects. *Food Res. Int.* **2025**, *201*, 115675. [CrossRef] [PubMed]
6. Petimar, J.; Grummon, A.H.; Simon, D.; Block, J.P. Nutritional Composition and Purchasing Patterns of Supermarket Prepared Foods Over Time. *Am. J. Prev. Med.* **2023**, *64*, 213–220. [CrossRef]
7. Wong, J.W.C.; Song, Y.X.; Xu, L.; Lu, A.Y.E. How do prepared dishes attract customers? Effects of prepared dish attributes on customers' trust, attitudes and adoption intentions. *Br. Food J.* **2025**, *127*, 4606–4625. [CrossRef]
8. Hou, B.; Chen, Y.; Wang, Z.W. Consumer-driven prepared food consumption policies: A sustainability perspective on reducing food waste. *Front. Sustain. Food Syst.* **2025**, *9*. [CrossRef]
9. Guo, J.; Zhang, M.; Law, C.L.; Luo, Z. 3D printing technology for prepared dishes: Printing characteristics, applications, challenges and prospects. *Crit. Rev. Food Sci. Nutr.* **2024**, *64*, 11437–11453. [CrossRef]
10. Lu, B.J.; Wen, D.C.; Li, H.; Chen, X. Food Traceability System Design Incorporating AI Chatbots: Promoting Consumer Engagement with Prepared Foods. *Foods* **2025**, *14*, 3731. [CrossRef]
11. Obonyo, E.; Formentini, M.; Ndiritu, S.W. Information sharing in agri-food supply chains: Insights from the Kenya dairy supply chain. *Supply Chain. Manag.-Int. J.* **2025**, *30*, 127–143. [CrossRef]
12. Ersoy, P.; Börühan, G.; Mangla, S.K.; Hormazabal, J.H.; Kazancoglu, Y.; Lafcı, C. Impact of information technology and knowledge sharing on circular food supply chains for green business growth. *Bus. Strategy Environ.* **2022**, *31*, 1875–1904. [CrossRef]
13. Christensen, F.M.M.; Jonsson, P.; Dukovska-Popovska, I.; Steger-Jensen, K. Information sharing for replenishment planning and control in fresh food supply chains: A planning environment perspective. *Prod. Plan. Control.* **2023**, *34*, 1371–1392. [CrossRef]
14. Yu, Y.N.; He, Y.; Guo, X.T.; Li, D.; Huang, H. Quality disclosure strategy with asymmetric demand information in food supply chains. *Transp. Res. Part E-Logist. Transp. Rev.* **2024**, *183*, 103427. [CrossRef]
15. Van Beusekom-Thoolen, P.; Holmes, P.; Jansen, W.; Vos, B.; de Boer, A. Interdisciplinary challenges associated with rapid response in the food supply chain. *Supply Chain. Manag. Int. J.* **2024**, *29*, 444–459. [CrossRef]
16. Liu, Y.P.; Yan, B.; Chen, X.X. Decisions of dual-channel fresh agricultural product supply chains based on information sharing. *Int. J. Retail. Distrib. Manag.* **2024**, *52*, 910–930. [CrossRef]
17. Gruzauskas, V.; Burinskiene, A.; Krisciunas, A. Application of Information-Sharing for Resilient and Sustainable Food Delivery in Last-Mile Logistics. *Mathematics* **2023**, *11*, 303. [CrossRef]
18. León-Bravo, V.; Borrello, B.; Ciccullo, F.; Caniato, F. Unpacking Proximity for Sustainability in Short Food Supply Chains. *Bus. Strategy Environ.* **2025**, *34*, 1792–1809. [CrossRef]
19. Sharma, M.; Alkatheeri, H.; Jabeen, F.; Sehrawat, R. Impact of COVID-19 pandemic on perishable food supply chain management: A contingent Resource-Based View (RBV) perspective. *Int. J. Logist. Manag.* **2022**, *33*, 796–817. [CrossRef]
20. Huang, Y.-S.; Ho, J.-W.; Kao, W.-Y. Availability and reliability of information transmission for supply chain coordination with demand information sharing. *Comput. Ind. Eng.* **2022**, *172*, 108642. [CrossRef]
21. Hong, X.; Gong, Y.; Rekik, Y.; Li, Q. Public versus private information: The impact of quality information sharing on under different channel structures. *Comput. Ind. Eng.* **2023**, *176*, 108937. [CrossRef]
22. Diao, W.; Tian, L.; Xu, Y. Effects of shared farms on the agricultural supply chain. *Int. J. Prod. Res.* **2024**, *62*, 7525–7539. [CrossRef]
23. Gopalakrishnan, S.; Sankaranarayanan, S. Cooperative security against interdependent risks. *Prod. Oper. Manag.* **2023**, *32*, 3504–3520. [CrossRef]
24. Li, Z.N.; Liu, Q.M.; Ye, C.M.; Dong, M.; Zheng, Y. Achieving Resilience: Resilient Price and Quality Strategies of Fresh Food Dual-Channel Supply Chain Considering the Disruption. *Sustainability* **2022**, *14*, 6645. [CrossRef]
25. Li, X.; Li, K.A.; Li, Y.J. Platform Operations under Dual-Channel Catering Supply Chain. *Mathematics* **2023**, *11*, 3610. [CrossRef]
26. Guo, X.T.; He, Y.; Yu, Y.N.; Li, D. Impact of shipping policies on ordering and pricing strategies in dual-channel perishable food retailing. *Int. Trans. Oper. Res.* **2025**. [CrossRef]
27. Lin, S.W.; Januardi, J. Willingness-to-pay experimental model for Stackelberg dual channel pricing decision. *Int. J. Retail. Distrib. Manag.* **2023**, *51*, 103–123. [CrossRef]
28. Yu, X.Q.; Ren, X.X. The Impact of Food Quality Information Services on Food Supply Chain Pricing Decisions and Coordination Mechanisms Based on the O2O E-Commerce Mode. *J. Food Qual.* **2018**, *2018*, 8956820. [CrossRef]
29. Zhang, Y.T.; Shang, J. Green technology investment and channel selection in sustainable food supply chain based on blockchain technology. *Front. Sustain. Food Syst.* **2025**, *9*, 1638313. (In Chinese) [CrossRef]
30. Musavi, M.; Taleizadeh, A.A.; Bozorgi-Amiri, A.; Moshtagh, M.S. Pricing decisions of organic and conventional products in a dual-channel competitive food supply chain. *Ann. Oper. Res.* **2022**, *354*, 789–825. [CrossRef]
31. Buechel, B.; Klössner, S.; Meng, F.Y.; Nassar, A. Misinformation due to asymmetric information sharing. *J. Econ. Dyn. Control.* **2023**, *150*, 104641. [CrossRef]
32. Gallo, A.; Accorsi, R.; Goh, A.; Hsiao, H.I.; Manzini, R. A traceability-support system to control safety and sustainability indicators in food distribution. *Food Control.* **2021**, *124*, 107866. [CrossRef]

33. Schudel, S.; Shoji, K.; Shrivastava, C.; Onwude, D.I.; Defraeye, T. Solution roadmap to reduce food loss along your postharvest supply chain from farm to retail. *Food Packag. Shelf Life* **2023**, *36*, 101057. [CrossRef]
34. Nerlove, M.; Arrow, K.J. Optimal Advertising Policy under Dynamic Conditions. *Economica* **1962**, *29*, 129–142. [CrossRef]
35. Starbird, S.A. Supply Chain Contracts and Food Safety. *Choices* **2005**, *20*, 123–127.
36. Ghosh, S.K.; Seikh, M.R.; Chakrabortty, M. Coordination and strategic decision making in a stochastic dual-channel supply chain based on customers' channel preferences. *Int. J. Syst. Assur. Eng. Manag.* **2024**, *15*, 2248–2270. [CrossRef]
37. Shapley, L.S. *A Value for n-Person Games*; Kuhn, H.W., Tucker, A.W., Eds.; Contributions to the Theory of Games II; Princeton University Press: Princeton, NJ, USA, 1953; pp. 307–317.
38. Hong, J.T.; Huang, P. Study on supply chain quality coordination based on differential game. *Chin. J. Manag. Sci.* **2016**, *24*, 100–107. (In Chinese)
39. Xu, C.Q.; Zhao, D.Z.; Yuan, B.Y.; He, L.F. Differential game model of upstream and downstream joint emission reduction and low-carbon promotion. *J. Manag. Sci. China* **2016**, *19*, 53–65. (In Chinese)

mathematics

MDPI

Article

Hybrid Graph Convolutional-Recurrent Framework with Community Detection for Spatiotemporal Demand Prediction in Micromobility Systems

Mayme Moon Zin [1], Karn Patanukhom [1,*], Merkebe Getachew Demissie [2] and Santi Phithakkitnukoon [1,*]

[1] Department of Computer Engineering, Faculty of Engineering, Chiang Mai University, Chiang Mai 50200, Thailand; maymemoonzin_1@cmu.ac.th
[2] Department of Civil Engineering, Schulich School of Engineering, University of Calgary, Calgary, AB T2N 1N4, Canada
* Correspondence: karn.patanukhom@cmu.ac.th (K.P.); santi@eng.cmu.ac.th (S.P.)

Abstract

The rapid growth of dockless electric scooter (e-scooter) sharing services has transformed short-distance urban mobility, offering convenience and sustainability benefits while amplifying challenges related to demand imbalance, fleet rebalancing, and spatial inequity. Accurate spatiotemporal demand prediction is therefore essential for optimizing resource allocation and supporting data-driven policy interventions. This study proposes a hybrid deep learning framework that integrates a Graph Convolutional Network (GCN) with a Gated Recurrent Unit (GRU) and community detection to enhance short-term prediction of e-scooter pick-up and drop-off demands. The Louvain algorithm is employed to partition urban areas into mobility-based communities, enabling the model to capture functional connectivity rather than relying solely on geographic proximity. Using real-world e-scooter trip data from Calgary, Canada, the model's performance is evaluated against established baselines, including a Masked Fully Convolutional Network (MFCN) and conventional GRU architectures. Results show that the proposed approach achieves up to 11.8% improvement in mean absolute error (MAE) compared with the MFCN baseline and more robust generalization across temporal horizons. The findings demonstrate that integrating community structures into graph-based learning effectively captures complex urban dynamics, providing practical insights for sustainable micromobility operation and service deployment.

Keywords: e-scooter demand prediction; spatiotemporal modeling; graph convolutional network (GCN); gated recurrent unit (GRU); community detection; micromobility; urban mobility management

MSC: 90B06

1. Introduction

Shared micromobility has rapidly transformed urban transportation over the past decade. Among various forms, dockless electric scooters (e-scooters) have expanded swiftly since their launch in Santa Monica, California, in 2017 by Bird. Within only a few years, major operators such as Bird and Lime scaled globally, embedding e-scooters as a convenient yet controversial feature of many city streets across North America, Europe, and Asia [1–3]. Their appeal lies in convenience, digital connectivity, and adaptability.

E-scooters can be unlocked through mobile applications, tracked via GPS, and paid for seamlessly using integrated digital platforms [4].

E-scooters are often described as a potential solution to the "first/last-mile" problem, connecting users between public transport nodes and final destinations [5,6]. Within the broader context of Mobility-as-a-Service (MaaS), they support multimodal travel, reduce car dependency, and promote sustainable and inclusive mobility [7]. However, their proliferation has also introduced new challenges concerning safety, environmental impact, and social equity. Accidents and pedestrian conflicts have raised public concerns [8], while life-cycle analyses reveal that much of their environmental footprint stems from manufacturing, maintenance, and fleet rebalancing rather than electricity consumption [9,10]. Moreover, accessibility remains uneven, as deployment tends to concentrate in central or affluent areas [11,12]. These issues highlight the need for operational efficiency and evidence-based policymaking driven by accurate demand prediction.

1.1. Importance of Demand Prediction in Micromobility

Accurate spatiotemporal demand prediction plays a key role in sustaining shared micromobility systems. Reliable prediction enables operators to optimize fleet redistribution, reduce idle time, and maintain service availability in high-demand areas. From a public-policy perspective, demand prediction can guide infrastructure planning, and equitable resource allocation. While statistical and classical machine learning models have provided foundational insights, they often fail to capture complex nonlinear interactions between spatial and temporal factors influencing micromobility demand [13,14].

With the rise of large-scale trip data, deep learning has become increasingly popular in mobility research. Early studies applied convolutional or recurrent neural networks to capture spatial or temporal correlations separately. Phithakkitnukoon et al. [15] proposed a Masked Fully Convolutional Network (MFCN) to predict e-scooter demand while handling data sparsity through a dual-branch classification and regression structure. Their approach achieved robust short- and long-term performance by weighting zero-demand regions and focusing on spatial areas of interest. Ham et al. [16] later introduced a convolutional autoencoder integrated with a recurrent decoder (ERD), demonstrating that latent feature extraction can improve short-term e-scooter demand prediction compared with long short-term memory (LSTM) models. Sahnoon et al. [17] evaluated three deep learning models, MFCN, UNet, and UNet Transformer (UNETR), for short-term prediction of e-scooter pick-ups and drop-offs. While Transformers are often favored for capturing long-range dependencies and achieving high accuracy, the study did not find definitive evidence of their overall superiority.

Kim et al. [18] incorporated community structure into a deep learning framework by clustering Seoul's e-scooter service areas using modularity optimization before predicting temporal demand with LSTM models. They found that activation functions significantly affected peak-period prediction performance, with the exponential linear unit and hyperbolic tangent (tanh) functions yielding the most accurate results. Liu et al. [19] analyzed three months of e-scooter trips in Indianapolis to uncover temporal variations, noting pronounced peaks during commuting hours and minimal activity at night. Bai and Jiao [20] compared usage patterns between Austin and Minneapolis, identifying the built environment, such as proximity to downtown areas and land-use diversity, as major drivers of demand. Similarly, Lee et al. [21] developed a multivariate log-linear regression model in Manhattan that incorporated demographic factors such as population density and median income, revealing that socioeconomic variables also influence ridership.

https://doi.org/10.3390/math14010116

1.2. Advances in Deep Learning for Spatiotemporal Demand Prediction

More sophisticated hybrid and graph-based models have recently emerged to address spatial dependencies more effectively. Yang et al. [22] proposed a Spatiotemporal Graph Convolutional Network (SGCN) combining graph convolution for spatial relations with LSTM for temporal dynamics, also including weather variables such as temperature, wind, and air quality. Using real-world data from Tianjin, their model outperformed baseline ARIMA and LSTM models. Li et al. [23] developed an irregular convolutional neural network (ICNN) to forecast bike-sharing demand in Singapore, Chicago, Washington D.C., New York, and London. The ICNN, which integrates convolutional and LSTM layers, achieved superior predictive accuracy across all sites.

Song et al. [24] tackled the issue of data imbalance using a Sparse Diffusion Convolutional Gated Recurrent Unit (SpDCGRU), which merges diffusion convolution with GRU architecture and introduces spatial data reclustering and fusion-loss strategies. Their approach achieved improved overall performance and interpretability. Xu et al. [25] advanced this direction by proposing a Spatiotemporal Multi-Graph Transformer (STMGT) that integrates four types of graphs, adjacency, functional similarity, demographic similarity, and transportation supply, to capture multiple forms of spatial dependency. Their model outperformed conventional deep learning frameworks and identified weather as the most influential predictor of demand. In parallel with these developments, attention-based architectures have gained prominence in spatiotemporal traffic and demand prediction. For example, Jiang et al. [26] proposed PDFormer, a transformer-based model designed to capture long-range propagation delays in complex traffic flows. Similarly, Lan et al. [27] introduced DSTAGNN, a dynamic spatiotemporal graph neural network that constructs adaptive spatial–temporal graphs from data to model evolving traffic patterns. More recently, Singh et al. [28] proposed an Integrated Spatiotemporal Graph Neural Network (ISTGCN) for traffic forecasting, demonstrating that jointly modeling spatial dependencies and temporal dynamics within a unified GNN architecture can significantly improve prediction accuracy. While these approaches highlight the effectiveness of advanced graph-based and attention-driven models, they typically rely on computationally intensive graph representations or fixed sensor networks, which limits direct functional interpretability in mobility systems.

These developments illustrate a clear progression toward deep, graph-based, and multi-feature architectures that jointly model spatial and temporal relationships. Nevertheless, most existing frameworks still rely on fixed geographic grids or administrative zones that may not represent functional travel communities. Few studies have explored spatial partitioning based on intrinsic travel-flow connectivity.

1.3. Graph-Based Learning and Community Detection

GCNs provide a powerful way to model complex dependencies between spatially linked nodes [13,14]. They have been successfully applied in transportation studies for traffic prediction [29], taxi-demand prediction [30], and ride-hailing demand analysis [31]. However, applications to micromobility remain relatively limited.

In parallel, community detection offers an alternative spatial structuring approach that reflects functional connectivity rather than fixed spatial boundaries. The Louvain algorithm [32] efficiently identifies cohesive communities within large networks by maximizing modularity [33]. Integrating this clustering approach into predictive models can capture localized collective behavior and improve interpretability in urban mobility studies. Dastjerdi and Morency [34] demonstrated that community-level modeling can enhance demand prediction performance in bike-sharing systems, particularly under changing conditions such as the COVID-19 pandemic.

https://doi.org/10.3390/math14010116

1.4. Research Gap and Contribution

Despite significant progress in deep learning-based micromobility short-term demand prediction, most existing studies still rely on fixed grid cells or administrative zones that may not represent functional mobility structures. Such spatial partitions often fail to capture the dynamic flow connectivity among urban areas, which limits interpretability and reduces prediction performance under sparse or unevenly distributed data. In addition, many state-of-the-art models such as SGCN, ICNN, SpDCGRU, and STMGT construct spatial graphs solely from geographic adjacency or predefined neighborhood structures. These approaches are not able to capture latent mobility communities that emerge from actual origin–destination travel behavior. Although these models advance graph-based and hybrid spatiotemporal learning, they do not incorporate functional communities as explicit and learnable features within the prediction framework. Furthermore, although GCNs effectively model spatial dependencies and GRUs capture temporal dynamics, few studies have integrated these approaches with community detection to represent mobility-driven spatial relationships. Existing hybrid models such as STMGT rely on multiple predefined spatial graphs or handcrafted proximity measures, but they do not include flow-derived communities or embed community identifiers as trainable inputs within the GCN. Likewise, prior spatiotemporal variants typically use a single regression output and do not separate the occurrence of demand from its magnitude. These methodological gaps highlight the need for a unified framework that combines mobility-informed spatial structures with multi-task prediction.

To address these gaps, this study proposes a hybrid deep learning framework that combines GCNs with community detection for short-term e-scooter demand prediction. The Louvain algorithm is employed to partition the city into demand-flow-based communities, producing a functional spatial representation that better reflects real-world travel behavior. Unlike prior studies that rely on fixed spatial grids or distance-based adjacency, our framework embeds community identifiers as a learnable feature, enabling the GCN to capture both local geographic correlations and higher-level functional similarity across urban regions. This community-aware design differs from existing models such as STMGT, which do not incorporate community detection outputs into the learning process, and it enables the model to jointly leverage geographic adjacency and mobility-based functional zones. The integration of a GRU-based temporal encoder and a dual-branch classification and regression architecture further allows the model to learn short-term temporal dynamics, spatial spillover effects, and demand sparsity within a single cohesive design. This multi-task structure, which distinguishes the probability of activity from the magnitude of demand, provides an additional level of modeling flexibility not present in existing spatiotemporal hybrids. These combined capabilities have not been addressed together in previous frameworks. The resulting community graph is integrated into a GCN-based predictive model, enabling the framework to jointly learn temporal patterns and inter-community dependencies while mitigating the effects of data sparsity and spatial heterogeneity. Using real-world e-scooter trip data from Calgary, Canada, the model is benchmarked against established baselines, including MFCN and LSTM architectures. Results demonstrate that incorporating community structures within graph learning substantially improves predictive accuracy and robustness.

Beyond methodological advancement, this research provides actionable insights for service providers and urban policymakers. More accurate demand prediction can support efficient fleet rebalancing, and service deployment, and sustainable integration of micromobility into urban transport systems. The findings also contribute to the broader field of urban informatics by showing how community-aware graph learning can uncover latent spatial organization within human mobility networks.

https://doi.org/10.3390/math14010116

2. Materials and Methods

2.1. Dataset

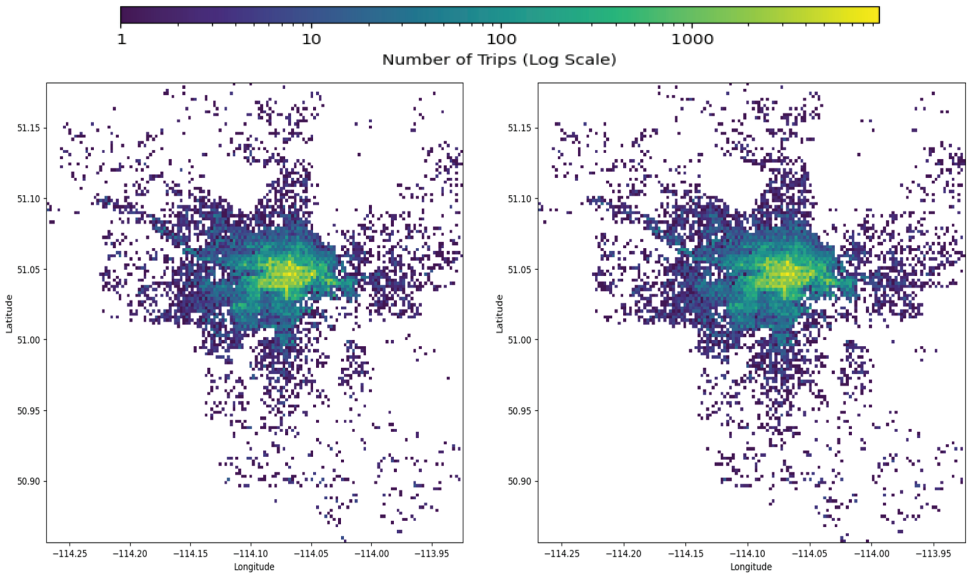
This research uses a real-world dataset from Calgary, Canada, where two companies, Lime and Bird, deployed 1000 and 500 e-scooters, respectively. Calgary is the largest city in the Canadian Province of Alberta. It covers an area of 820.62 km^2 with the population of 1,306,784. Dockless e-scooter sharing was introduced in Calgary in July of 2019. E-scooters are small, electric-powered vehicles, allowing users to rent and return scooters at any location via a mobile application. Each scooter trip record contains detailed usage information, including the date, hour of the day, day of the week, trip duration in minutes, trip length in kilometers, and starting/ending geolocations (latitude, longitude). The scope of the data was selected as a period of 75 days from 15 July 2019 to 27 September 2019.

In the original dataset, the City of Calgary aggregated trip geolocations into 300 m^2 hexagonal grids for privacy protection. In contrast, our approach adopts a 200 m square grid format to enable a detailed spatial analysis of scooter usage patterns. The total number of e-scooter rides in the dataset is 459,478 rides from 4080 different starting grid locations to 4367 ending grid locations. Although Calgary city is composed of approximately 21,702 square grids (200 m $\times$ 200 m), the analysis dataset is 19.45% of the city. Figure 1 illustrates this spatial coverage, highlighting the active grids used in the analysis and illustrating the spatial concentration of trips within the downtown core. This study aimed to predict the hourly and 24 h demand of pick up and drop off for each square grid. Analysis of the trip data indicates that e-scooter demand is higher on weekends than on weekdays. On average, there were 6570 trips per day on weekends and 5965 trips per day on weekdays, representing an increase of approximately 10% in weekend usage compared with weekday activity. Looking at the hourly demand patterns, shown in Figure 2, it provides a detailed look into the different hourly patterns of weekday and weekend e-scooter usage. During the weekdays, the demand clearly shows two peaks, which is a strong indicator of work-related commuting. A smaller rise occurs in the morning, around 8 a.m., as people travel to work, followed by a much larger peak in the late afternoon, around 4 p.m. (16:00), when people make their return trips. In contrast, weekend demand grows more gradually through the late morning and peaks in the mid-afternoon, suggesting a stronger focus on leisure and social trips rather than commuting. The demand reaches its highest point in the late afternoon, between 3 p.m. and 6 p.m. (15:00–18:00). This comparison suggests that e-scooters play a dual role in urban mobility, serving as a practical option for structured weekday commutes and as a flexible, leisure-based mode of transport on weekends. The analysis of trip duration and distance, shown in Figure 3, indicates that most e-scooter rides are short in both time and distance, a characteristic feature of micromobility usage. While the average trip duration was 12.88 min, the median was only 8.40 min. A similar pattern is seen for trip distance. The average trip covered 1.85 km, but the median distance was just 1.26 km. This data suggests that e-scooters in Calgary are most often used for solving the "first- and last-mile" problem, such as traveling between transit stations and workplaces, and to substitute short walking trips rather than longer car journeys.

The spatial distribution of e-scooter activity across Calgary, illustrated in Figure 4, highlights strong spatial clustering in both pick-up and drop-off demand. The highest concentration of trips occurs in and around the downtown core, particularly in areas with dense commercial activity and proximity to public transit stations. This pattern suggests that e-scooters are commonly used for short urban trips within the central business district and for connecting to public transit hubs. Beyond the downtown area, moderate levels of activity are also visible on main transit routes and near other popular areas for work and recreation, such as those adjacent to the Bow River pathway network. Overall comparison between weekdays and weekends shows that weekday trips are slightly more

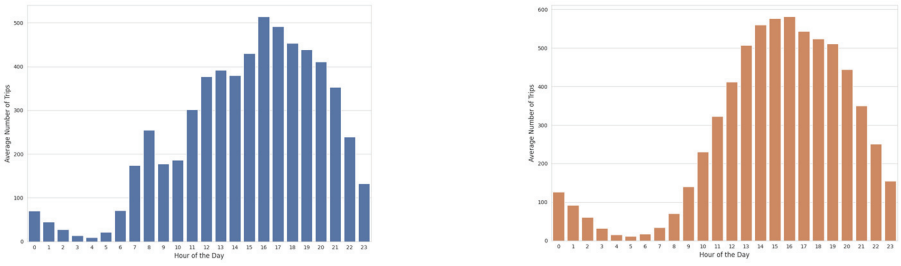
https://doi.org/10.3390/math14010116

centralized, reflecting commuter-oriented travel to and from the city center. In contrast, weekend trips are more spatially dispersed, with increased activity extending toward parks and leisure destinations. These spatiotemporal characteristics form the foundation for the predictive modeling described in the next section, where hourly and daily demand are estimated for each grid cell.



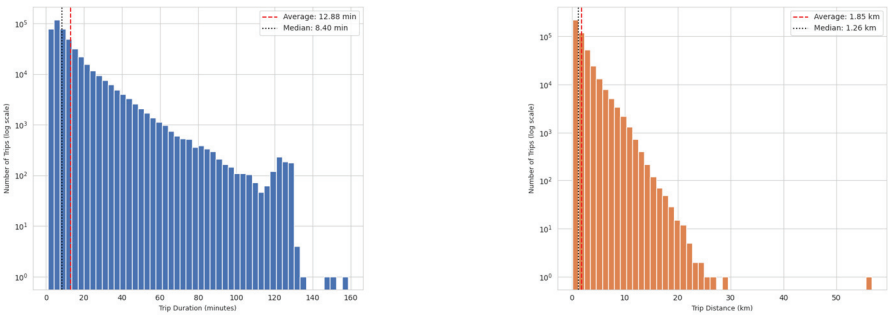
(**a**) (**b**)

Figure 1. Spatial distribution of e-scooter (**a**) pick-up and (**b**) drop-off activities in Calgary. The city contains 21,702 square grids (200 m $\times$ 200 m), but only about 19.45% recorded trips during the 75-day study period. Color intensity (log_{10} scale) indicates the number of trips, showing concentrated activity in the city center.



(**a**) (**b**)

Figure 2. Average hourly e-scooter demand: (**a**) weekdays and (**b**) weekends.



(**a**) (**b**)

Figure 3. Log_{10} scale distributions of e-scooter trip characteristics: (**a**) trip duration (minutes) and (**b**) trip distance (kilometers). The average (red dashed line) and median (black dotted line) values are indicated for both duration (12.88 and 8.40 min, respectively) and distance (1.85 and 1.26 km, respectively).

https://doi.org/10.3390/math14010116

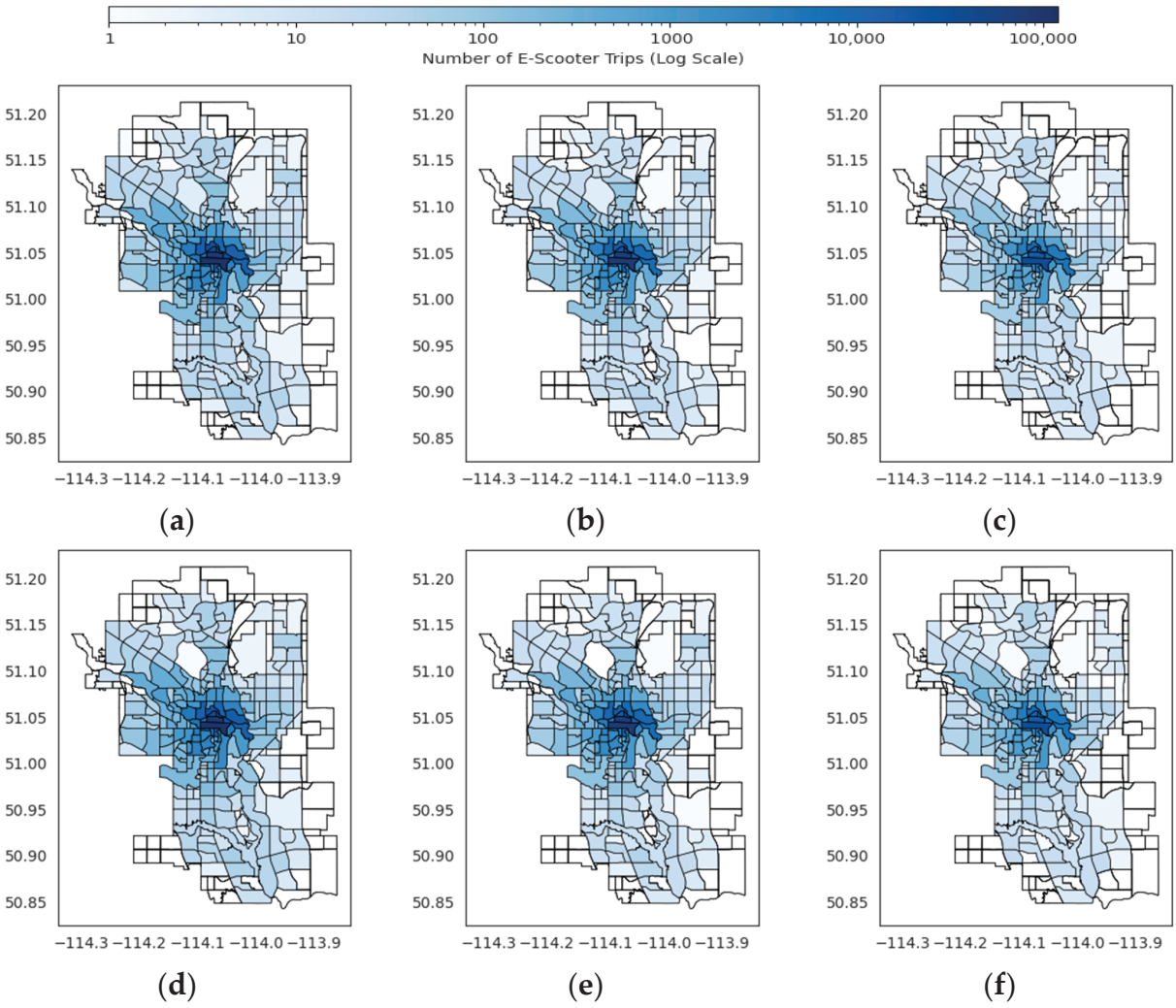


Figure 4. Spatial distribution of e-scooter demand in Calgary. (**a**–**c**) Pick-up demand for overall, weekday, and weekend periods. (**d**–**f**) Drop-off demand for overall, weekday, and weekend periods. Brighter colors on the log_{10} scale indicate higher trip volumes.

2.2. Proposed Model Architecture

To predict the spatiotemporal demand of e-scooters, we propose a deep learning model based on GCN and GRU, the workflow of the proposed model is summarized in detail in in Algorithm 1 and illustrated in Figure 5. A key aspect of our approach is the inclusion of a community feature, derived by applying the Louvain community detection algorithm to a travel network graph of the grids. This feature provides the model with a higher-level understanding of the urban structure. The model architecture consists of two main branches: a classification branch and a regression branch. The model input is a spatiotemporal tensor that represents 14 historical demand lag features and five contextual features (which include the sine–cosine components for the temporal dynamics and the weekend flag) and one categorical community identifier feature. In this implementation, only the lag features are processed by the GRU to learn short- and long-term temporal patterns. The resulting hidden states are then concatenated with the five contextual features, to which the embedded community feature is added, and then passed through two GCN layers, which capture spatial dependencies between adjacent grids and communities. This combined GRU-GCN architecture effectively integrates sequential and spatial dependencies within the same learning framework.

https://doi.org/10.3390/math14010116

Algorithm 1 Spatiotemporal E-Scooter Demand Prediction Using the Proposed GCN Model

Input: Feature tensor $X \in R^{B\times N\ \times H\times F}$, where B is the batch size, N is the number of active grid nodes, H is the historical time window, and F is the total number of input features.
Edge index matrix $\widehat{Adj} \in R^{2\times \mathrm{E}}$.
Output: Predicted e-scooter pick-up or drop-off demand matrix $\widehat{y} \in R^{B\times N\ \times 1}$.
Initialization: Define feature counts: $F_{lag} = 14$, $F_{context} = 5$, $F_{comm\_id} = 1$.
Models: GRU, Embedding, GCNConv$^{(1)}$, GCNConv$^{(2)}$
Split X into three parts based on feature indices:
$X_{\mathrm{lag}} \leftarrow X[:, :, :, 0{:}F_{lag}]$ // Contain the 14 lag-based demand features.
$X_{\mathrm{context}} \leftarrow X[:, :, :, F_{lag}{:}F_{lag} + F_{context}]$ // Contains the 5 contextual features: $\{h_{sin}, h_{cos}, d_{sin}, d_{cos}, w\}$, where h_{sin} and h_{cos} are sine–cosine encodings of the hour (24 h cycle), d_{sin} and d_{cos} are sine–cosine encodings of the day of the week (7-day cycle), and w is the weekend indicator.
$X_{\mathrm{comm\_id}} \leftarrow X[:, :, :, F_{lag} + F_{context}]$ // Contains the 1 categorical community ID feature $\{c_{id}\}$.
$X_{reshaped\_lag} \leftarrow \mathrm{reshape}\ (X_{\mathrm{lag}}, (B \times N, H, F_{lag}))$
$H_{out}, \_ \leftarrow \mathrm{GRU}\ (X_{reshaped\_lag})$
$H_{\mathrm{GRU}} \leftarrow H_{out}\ [:, -1, :]$
$X_{context\_last} \leftarrow X_{\mathrm{context}}\ [:, :, -1, :]$
$H_{\mathrm{context\_flat}} \leftarrow \mathrm{reshape}\ (X_{\mathrm{context\_last}}, (B \times N, F_{context}))$
$X_{\mathrm{comm\_id\_last}} \leftarrow X_{\mathrm{comm\_id}}\ [:, :, -1, :]$
$H_{comm\_id\_flat} \leftarrow \mathrm{reshape}\ (X_{\mathrm{comm\_id\_last}}, (B \times N))$
$H_{comm\_emb} \leftarrow \mathrm{Embedding}(H_{comm\_id\_flat})$
$H^{(0)} \leftarrow \mathrm{concat}[H_{GRU}, H_{\mathrm{context\_flat}}, H_{comm\_emb}]$
$H^{(1)} \leftarrow \mathrm{ReLU}\ (\mathrm{Dropout}(\mathrm{GCNConv}^{(1)}(H^{(0)}, \widehat{Adj})))$
$H^{(2)} \leftarrow \mathrm{ReLU}\ (\mathrm{Dropout}(\mathrm{GCNConv}^{(2)}(H^{(1)}, \widehat{Adj})))$
$P \leftarrow \sigma\ (W_{\mathrm{cls}} H^{(2)} + b_{\mathrm{cls}})$
$\mathrm{A} \leftarrow W_{\mathrm{reg}} H^{(2)} + b_{\mathrm{reg}}$
$\widehat{y} \leftarrow \max\ (0, \mathrm{P} \odot \mathrm{A})$
$\widehat{y} \leftarrow \mathrm{reshape}\ (\widehat{y}, (B,N,1))$
return $\hat{y}$

The classification branch addresses a binary classification task, estimating the probability (P) that any trip activity will occur in each community at the next time step. This branch is optimized using a binary cross-entropy loss function. In parallel, the regression branch predicts the expected magnitude of the demand (A), if activity does occur, and is optimized using a Weighted Mean Squared Error (WMSE) loss. The weighting scheme uses the square root of the historical demand as a sublinear scaling factor. This increases the importance of high-demand locations while preventing excessively large weights that could destabilize the training process. Using raw demand values as weights would cause the loss function to be dominated by a small number of extreme-demand samples, leading to overfitting and poor generalization. The square root operation provides a balanced compromise. It emphasizes accurate prediction in high-demand grids, where operational decisions such as rebalancing are most critical, while maintaining numerical stability and ensuring that moderate-demand regions also influence model learning in a meaningful way. Finally, the model combines the outputs of both branches by multiplying the probability by the predicted demand ($P \odot A$). Since the classification and regression branches are trained sequentially and independently rather than through a joint multi-task

https://doi.org/10.3390/math14010116

loss, a weighting coefficient λ is not required to balance the two objectives. This decoupled training approach avoids the sensitivity issues of multi-task methods and allows each branch to specialize effectively. This learning framework is particularly effective for sparse demand data, as it enables the model to first identify potential active locations before estimating demand intensity. As a result, the proposed model achieves improved accuracy and robustness in predicting both the spatial and temporal variations in e-scooter usage across urban environments.

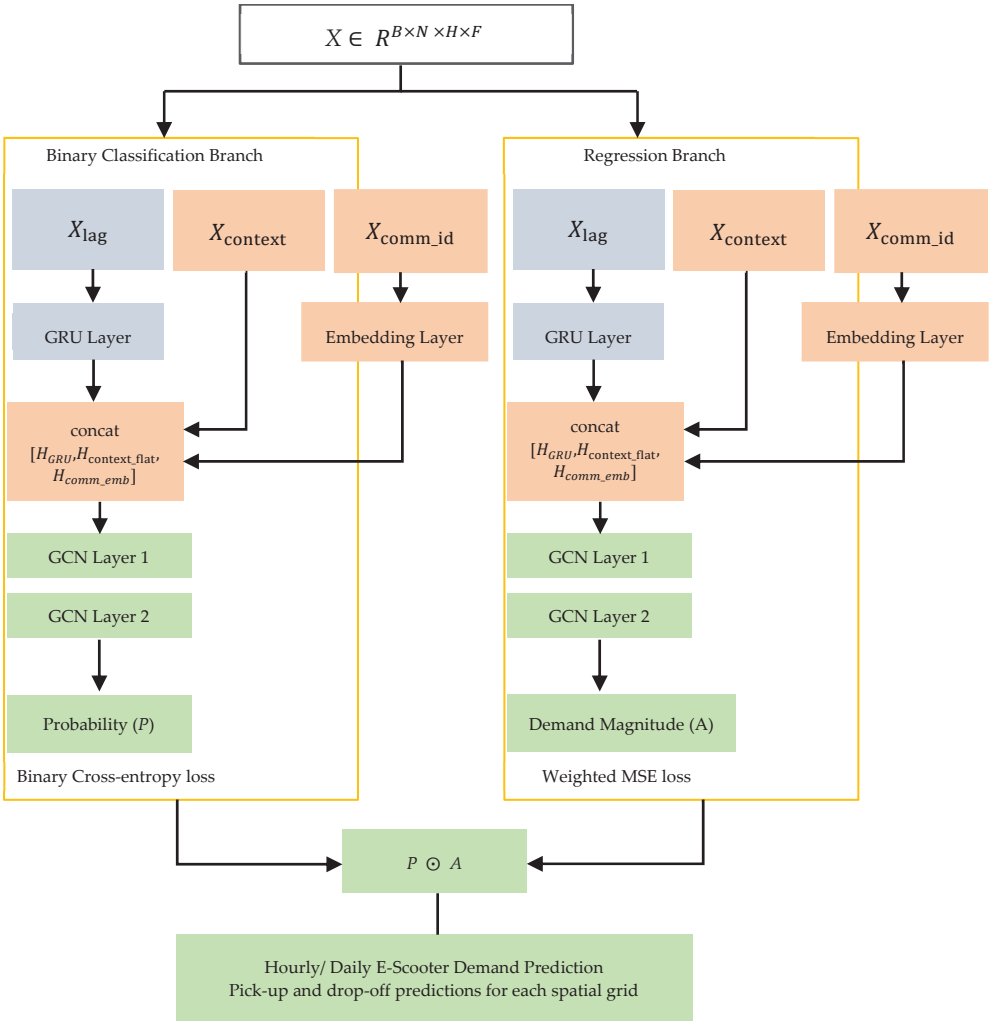


Figure 5. Proposed model architecture for scooter demand prediction. The input tensor $X \in \boldsymbol{R}^{\boldsymbol{B}\times\boldsymbol{N}\times\boldsymbol{H}\times\boldsymbol{F}}$ includes 14 lag features ($\boldsymbol{X_{lag}}$), five contextual variables: $\boldsymbol{h_{sin}}$, $\boldsymbol{h_{cos}}$, $\boldsymbol{d_{sin}}$, $\boldsymbol{d_{cos}}$, w with cyclic variables encoded as sine and cosine pairs and one categorical community feature ($\boldsymbol{c_{id}}$). The classification branch outputs the probability P, and the regression branch estimates the demand magnitude A.

2.2.1. Input Representation

The raw e-scooter trip data was transformed into a structured spatiotemporal tensor that serves as the model's input. This process began with spatial discretization, where the city was partitioned into the 200 m $\times$ 200 m grids defined in Section 2.1. Each trip was then assigned to a grid based on its pick-up and drop-off locations. Subsequently, the data was temporally aggregated by counting the number of pick-ups and drop-offs within each grid for every hour of the study period. To handle the inherent sparsity of the data and create a regular time series for each grid, all grid-time instances with no activity were explicitly assigned a demand value of zero.

https://doi.org/10.3390/math14010116

From this structured time series, a feature vector was engineered for each grid at each time step, consisting of 20 features. These features were organized into two distinct groups based on their function within the model architecture, as detailed in Table 1. The historical demand features include a set of fourteen lagged demand values from specific past time steps, serving as the primary predictive inputs. These lags are strategically selected to capture immediate trends (recent hours), daily seasonality (the same hour on the previous day), and weekly seasonality (the same hour in the previous week). The spatiotemporal context features include six variables that provide static contextual information for each prediction. These features were concatenated with the GRU output and passed directly into the GCN layers. Among these, the community feature enhanced the model's spatial understanding of urban structure. The Louvain algorithm was applied to detect clusters of functionally related grids, and each grid's community identifier was used as a categorical input feature. Each community identifier is encoded using a 10-dimensional embedding layer. This dimensionality was selected through empirical tuning (testing 5, 10, 16, and 32), with 10 yielding the best validation performance. This choice provides a compact yet expressive representation for the 257 identified communities, effectively balancing model complexity with the ability to capture latent functional similarities between urban zones. The resulting vectors are concatenated with the temporal and contextual features before being passed to the GCN layers. In addition, temporal dynamics were represented by sine–cosine encodings of the hour of the day and day of the week, which preserve the cyclic nature of these temporal variables and by a binary weekend indicator capturing behavioral differences between weekdays and weekends. Finally, this feature data was structured into sequences. Each training sample contains feature vectors for all active grids over a 24 h historical window ($H = 24$). This results in a final input tensor with the shape (B, N, H, F), where B is the batch size, N is the number of active grids, H is the historical time window, and F is the number of features.

Table 1. Features used for the next-hour ($t + 1$) prediction and the next 24 h ($t + 24$) prediction.

Feature Group	Prediction Task/Feature	Lag Values (Hours Before Prediction)	Description
Historical Demand Features (14)	Next-Hour Prediction ($t + 1$)	0, 1	Immediate trend (recent hours)
		22, 23, 24	Daily cycle (same hour on the previous day)
		143, 166, 167, 168, 191	Weekly cycle (same hour in the previous week)
		335, 336	Two-week cycle
		503, 504	Three-week cycle
	Next-24-h Prediction ($t + 24$)	0, 1	Immediate trend
		120, 121	Five days prior
		143, 144, 145	Six days prior
		168	One week prior (weekly cycle)
		312, 313	Thirteen days prior
		336	Two weeks prior
		480, 481	Twenty days prior
		648	Twenty-seven days prior
Spatiotemporal Context Features (6)	Temporal context : h_{sin}, h_{cos}	-	Represents 24 h cyclic variation (sine–cosine encoded)
	Temporal context : d_{sin}, d_{cos}	-	Represents seven-day cycle (sine–cosine encoded)
	Temporal context: w	-	Binary variable indicating weekends (1) vs. weekdays (0)
	Spatial context : c_{id}	-	Categorical feature from Louvain community detection

https://doi.org/10.3390/math14010116

2.2.2. Community Detection

To enhance the model's spatial understanding of Calgary's urban structure, we introduce a community feature derived from the origin–destination (O-D) flows of all e-scooter trips. A mobility network graph was constructed using the *networkx* library, where each active grid was treated as a node. For each individual trip record, an unweighted edge was created between its start grid and its end grid, representing the functional connectivity derived from user travel patterns. The Louvain community detection algorithm was then applied to the aggregated network to partition the city into communities of grids exhibiting strong internal trip connections [32]. In this study, community detection using the Louvain algorithm was performed once on the fully aggregated trip graph, which covers all days and hours in the dataset (75 days). This approach generates a single, robust set of mobility-based communities that remains stable across the entire study period, capturing long-term functional connectivity rather than transient daily fluctuations. The Louvain method is an efficient algorithm that identifies community structures by optimizing a metric known as modularity (Q) [33]. Modularity measures the quality of a partition by comparing the density of connections within communities to the density of connections in a random network. A high modularity score indicates a well-defined community structure. Modularity is formally defined as:

$$Q = \frac{1}{2m} \sum_{i,j} (A_{ij} - \frac{k_i k_j}{2m}) \delta(c_i, c_j), \quad (1)$$

where A_{ij} is an element of the adjacency matrix (1 if an edge exists between nodes i and j, and 0 otherwise), k_i is the degree of node i, $\delta(c_i, c_j)$ is 1 if nodes i and j are in the same community, and 0 otherwise, m is the total number of edges in the network [30].

The Louvain method operates iteratively through two repeating phases. In the first phase (modularity optimization), each node is initially assigned to its own community. For each node, the algorithm evaluates the modularity gain that would occur if the node were moved into each of its neighbors' communities, selecting the move that provides the largest positive gain. This process continues for all nodes until no individual move can further improve modularity. In the second phase (network aggregation), the communities identified in the first phase are aggregated into "super-nodes" to form a smaller network, where edges between super-nodes are weighted by the sum of the connections between their constituent nodes. These two phases are repeated until no additional improvement in modularity is achieved, yielding a stable community structure [32].

The detection was performed on the combined network of all active pick-up (4080) and drop-off (4367) grids, encompassing all unique locations appearing as either trip origins or destinations. The Louvain algorithm identified 257 communities ($\approx$17 grids per community) with a final modularity score of 0.368, indicating a meaningful community structure within Calgary's mobility network. A modularity value of this magnitude reflects a moderately strong clustering pattern, meaning that connections within communities are significantly denser than would be expected in a random graph. This ensures that the detected communities represent genuine functional regions shaped by actual mobility flows rather than artifacts of the algorithm. In practice, such a modularity score confirms that the resulting clusters, such as downtown commercial zones, university districts, and recreational corridors, are coherent and interpretable, making the community assignments informative as spatial features for demand prediction. Each grid cell was then assigned a unique community identifier (c_{id}), which was used as a categorical feature in the model's input to represent its belonging to a specific functional zone. Incorporating community membership provides the model with a clearer understanding of the city's spatial organization. It reduces data sparsity, groups grids with similar mobility patterns (such as downtown, university areas and recreational

https://doi.org/10.3390/math14010116

corridors) and enables the GCN to learn more meaningful relationships within and between these regions.

Community detection algorithms can be broadly categorized into several groups, including modularity-optimization approaches (such as Louvain and Leiden), statistical inference and generative models (for example, stochastic block models), spectral clustering methods, and evolutionary or heuristic algorithms. Recent studies have introduced hybrid intelligent approaches that combine fuzzy multi-criteria decision making with evolutionary search, such as the Fuzzy-AHP-based evolutionary algorithm proposed by Pourabbasi et al. [35]. These methods aim to improve accuracy in identifying communities within large and complex networks.

Although a variety of newer algorithms exist, the Louvain method remains well suited for mobility-based applications due to its strong scalability for large and sparse O-D networks, its ability to achieve high modularity without extensive parameter tuning, and its production of flat community labels that can be directly embedded as categorical features within a GCN. Since the objective of this study is to enhance spatiotemporal demand prediction rather than develop a new community detection algorithm, Louvain provides an efficient, interpretable, and practical solution for constructing mobility-aware spatial features.

2.2.3. Gated Recurrent Unit (GRU)

The GRU component is designed to capture the temporal dependencies of e-scooter demand over time. As discussed in Section 2.1, Calgary's e-scooter usage exhibits pronounced daily and hourly variations, making a sequential model like a GRU essential for accurate prediction. A GRU is an advanced type of Recurrent Neural Network (RNN) developed to process sequential data efficiently, making it well suited for time-series prediction [36]. Unlike a simple RNN [37,38], a GRU incorporates gating mechanisms—an update gate and a reset gate—that control the flow of information. This enables the model to learn long-term dependencies and decide which historical information is relevant for the current prediction, mitigating issues such as the vanishing-gradient problem [16].

In our model, at each hour t, the GRU receives a spatiotemporal input tensor $X_t \in R^{N\times F}$, where N denotes the number of active grid cells and F = 14 represents the number of lag-based features describing each grid. These lag features capture the historical demand patterns across the previous 24 h window (H = 24) and serve as the temporal sequence input to the GRU. For each grid cell, the GRU processes these sequences step by step, updating its hidden state at every time interval. The final hidden state thus encodes a compact representation of the temporal dynamics for that grid over the past 24 h. In the next stage, this learned temporal representation is concatenated with six contextual features (the four sine–cosine components derived from the hour of the day and day of the week, the weekend flag, and the community identifier derived from Louvain detection in Section 2.2.2). The combined vector is then passed to the GCN layers (Section 2.2.4), which integrate spatial correlations among adjacent grids and communities. The GRU component effectively models the sequential evolution of demand by using its update and reset gates to control the flow of information and mitigate the vanishing-gradient problem common in standard recurrent networks. Formally, for each time step t, given the input vector x_t and the previous hidden state h_{t-1}, the GRU updates its state using the following operations [33]:

$$\text{Reset Gate}: \quad r_t = \sigma\left(W_{ir}x_t + b_{ir} + W_{hr}h_{t-1} + b_{hr}\right), \tag{2}$$

$$\text{Update Gate}: \quad z_t = \sigma\left(W_{iz}x_t + b_{iz} + W_{hz}h_{t-1} + b_{hz}\right), \tag{3}$$

$$\text{New Gate}: \tilde{h}_t = \tanh\left(W_{in}x_t + b_{in} + r_t \odot \left(W_{hn}h_{t-1} + b_{hn}\right)\right), \tag{4}$$

https://doi.org/10.3390/math14010116

$$\text{Final Hidden State}: \; h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \widetilde{h}_t, \tag{5}$$

where x_t is the input, h_t is the hidden state, W and b are learnable parameters, σ is the sigmoid function, and $\odot$ denotes element-wise multiplication. The update gate z_t determines how much prior information is retained, while the reset gate r_t controls how much of the previous state to forget. After processing the full 24 h sequence, the GRU outputs the final hidden representation h_t, which summarizes the temporal context of demand for each grid.

In summary, the GRU functions as the temporal encoder within the proposed GRU-GCN framework, transforming sequences of historical demand into meaningful latent representations that provide a robust foundation for the spatial learning performed by the subsequent GCN component.

2.2.4. Graph Convolutional Network (GCN)

The GCN component builds upon the temporal representations produced by the GRU to model spatial dependencies in e-scooter demand across Calgary's grid structure. As discussed in Section 2.1, demand exhibits clear spatial clustering, particularly in downtown and transit-adjacent areas, with notable variation between weekdays and weekends. The GCN is well suited for this task because it operates on graph-structured data, propagating information between connected nodes to learn relational patterns such as demand spillover between adjacent grids or within functionally similar communities [39]. In this framework, the city is represented as a graph $G = (V, E, A)$. The components are explicitly defined as follows. Vertices (V) represent the set of N nodes corresponding to the active grid cells where e-scooter trips occurred. In our study, N represents the 4080 starting locations for the pick-up graph and 4367 ending locations for the drop-off graph. Edges (E) constitute the set of unweighted edges capturing spatial proximity, where an edge $(i, j) \in$ E is established if and only if grid cell i and grid cell j share a common geographic boundary. This structure models the direct diffusion of demand between immediately adjacent areas. The Adjacency Matrix (A) is an $N \times N$ symmetric binary matrix where $A_{ij} = 1$ if $(i, j) \in$ E and $A_{ij} = 0$ otherwise. Although the spatial relationship is inherently symmetric, self-loops ($A_{ii} = 1$) are implicitly handled during the GCN propagation step via the renormalization trick ($\hat{A} = A + I$) to ensure numerical stability and include node-specific features during aggregation.

The input to the GCN is the concatenated representation of the GRU output and the six contextual features (as described in Table 1). Specifically, the GRU provides a matrix of hidden states $H^{(0)} \in R^{N \times D}$, where D is the GRU hidden dimension. Each row $h_i^{(0)}$ represents the temporal embedding for grid i. These temporal embeddings are then augmented with the contextual features to form the complete node feature matrix used by the GCN. Each node also carries an additional categorical feature, c_{id}, derived from the Louvain community detection process (Section 2.2.2). This feature is numerically encoded and incorporated into the node feature matrix, enabling the GCN to jointly capture local spatial proximity through network edges and regional functional similarity through community membership. The GCN consists of two stacked layers that progressively aggregate spatial information. We adopt the spectral graph convolution formulation of [39], which approximates localized spectral filters using first-order Chebyshev polynomials [40]. For a given layer l, the propagation rule is defined as Equation (6).

$$H^{(l+1)} = \sigma(\hat{D}^{-\frac{1}{2}} \hat{A} \hat{D}^{-\frac{1}{2}} H^{(l)} W^{(l)}), \tag{6}$$

https://doi.org/10.3390/math14010116

where $H^{(l)}$ is the matrix of node features at layer l, $\hat{A} = A + I$ is the adjacency matrix of the graph with self-loops added, $\widehat{D}$ is the diagonal matrix of $\hat{A}$, $W^{(l)}$ is the layer-specific learnable weight matrix, and σ denotes the ReLU activation.

This operation allows each node to update its representation by aggregating features from its spatially connected neighbors. The first GCN layer captures local correlations, such as demand patterns propagating from busy transit hubs to adjacent grids, while the second layer refines higher-order relationships, producing the final spatiotemporal embeddings matrix H^2 for each grid. These embeddings are then fed into the parallel classification and regression branches (Section 2.2). The classification branch uses a fully connected layer followed by a sigmoid activation to output the activity probability $P \in [0, 1]$ per grid. The regression branch employs a similar structure but with a linear output to predict demand magnitude $A \geq 0$, incorporating weighed MSE loss considerations for count data. The branches share the initial GCN layers for efficiency but diverge after the GRU stage to specialize in their respective tasks.

To enhance generalization and ensure stable convergence, the model was trained using the Adam optimizer [41] with an initial learning rate of 0.0005 and a weight decay of 1×10^{-5} (L2 regularization). The training process utilized a batch size of 16 and ran for a maximum of 200 epochs. To prevent overfitting on the sparse spatial graph, dropout regularization with a rate of 0.2 was applied after each GCN layer [42], and an early stopping strategy was implemented with a patience of 10 epochs. Additionally, a learning rate scheduler reduced the learning rate by a factor of 0.5 if the validation loss stagnated for 5 epochs. Each branch is trained independently: the classifier minimizes Binary Cross-Entropy loss, while the regressor minimizes Weighted MSE loss. Instead of training separate models for different temporal periods, the model was trained on a unified dataset comprising both weekdays and weekends, using a single set of training hyperparameters. Since the framework does not use a joint multi-task loss, no weighting coefficient λ is required. Instead, temporal variations are captured directly through the sine–cosine encodings of the hour of the day and day of the week, together with a weekend indicator. These features allow the model to learn distinct demand patterns for weekdays and weekends within a single architecture. Final predictions combine both outputs via multiplication ($P \odot A$) at inference time. In summary, the GCN functions as the spatial encoder in the proposed GRU-GCN architecture, transforming grid-specific temporal embeddings into relational representations that reflect neighborhood dynamics and community structure. This integration enables precise prediction of demand hotspots while addressing the spatiotemporal sparsity inherent in urban e-scooter data.

2.3. Implementation Details

To ensure reproducibility, here we summarize the specific configuration used in our experiments. All models were implemented using PyTorch version 2.9.0 and PyTorch Geometric on an NVIDIA A100 GPU. The Adam optimizer was employed with an initial learning rate of 0.0005. To mitigate overfitting, a dropout rate of 0.2 was applied after each GCN layer, and early stopping was used with a patience of 10 epochs. The complete set of hyperparameters and model architecture settings is summarized in Table 2.

Table 2. Summary of model hyperparameters and training configurations used in the proposed GRU-GCN framework.

Hyperparameter	Value	Description
GRU Hidden Dimension	64	Hidden state size of the GRU layer
GRU Layers	1	Number of recurrent layers in the GRU block
GCN Hidden Dimension	64	Node embedding size in GCN layers

https://doi.org/10.3390/math14010116

Table 2. *Cont.*

Hyperparameter	Value	Description
GCN Layers	2	Number of Graph Convolutional layers
Community Embedding Dim	10	Dimension of learnable community embeddings
Training Settings		
Batch Size	16	Number of samples per batch
Training Epochs	200	Maximum iterations
Early Stopping Patience	10	Epochs to wait before stopping
Optimizer	Adam	Adaptive Moment Estimation
Learning Rate	0.0005	Initial learning rate
Weight Decay	1×10^{-5}	L2 regularization coefficient
Dropout Rate	0.2	Applied after GCN layers

3. Results

To evaluate the predictive performance of the proposed GRU-GCN framework, the model was trained and tested on a 75-day e-scooter dataset from the City of Calgary. The data was divided to preserve temporal consistency: the first 61 days (15 July–14 September 2019) were used for training, while the final 14 days (15–27 September 2019) were used for testing. The Masked Fully Convolutional Network (MFCN) developed by Al. [15] was used as the baseline for comparison. Building upon this benchmark, six variants of the proposed GCN framework were constructed to investigate how alternative spatial–temporal configurations influence predictive performance. The evaluation considers two prediction horizons: next-hour prediction ($t + 1$) and next 24 h prediction ($t + 24$). The performance of the model was quantified using Mean Absolute Error (*MAE*), defined as follows:

$$\boldsymbol{MAE} = \frac{\mathbf{1}}{\boldsymbol{N}} \sum_{\boldsymbol{i}=\mathbf{1}}^{\boldsymbol{N}} |\hat{\boldsymbol{y}}_{\boldsymbol{i}} - \boldsymbol{y}_{\boldsymbol{i}}|, \tag{7}$$

where $\boldsymbol{y}_{\boldsymbol{i}}$ and $\hat{\boldsymbol{y}}_{\boldsymbol{i}}$ denote the observed and predicted e-scooter demand for grid i, and N is the total number of grid cells evaluated. The MAE measures the average magnitude of prediction errors across all grids, providing an overall indicator of accuracy.

The following subsections summarize the performance of the proposed GCN model, which integrates a mobility-based community feature and consistently provides the most balanced prediction results across sparse, moderate, and high-demand regions.

3.1. Next-Hour Pick-Up and Drop-Off Prediction Performance

Six GCN model variants were developed to evaluate different spatiotemporal modeling strategies for short-term e-scooter demand prediction. The benchmark reference is the MFCN from the baseline study. The GCN variants gradually build from a baseline spatial GCN to more sophisticated architectures incorporating temporal encoding with GRU, dual-branch outputs (classification and regression), weighted loss adjustment, and community-based spatial features. To clarify the model configurations presented in Table 3, the variants are defined as follows. Proposed SB (Single Branch) utilizes a single regression output without a classification stage. Proposed WC1 (With Classification Version 1) model represents the initial two-stage architecture consisting of independent classification and regression steps using standard MSE, while Proposed WC2 (With Classification Version 2) model employs the enhanced two-stage framework that incorporates Weighted MSE loss to better handle high demand. Proposed GCN-MM (Multi-Model) involves training separate models for each community independently. Finally, the Proposed GCN represents the complete framework, integrating an optimized two-stage framework with community-based spatial features. Table 3 presents the performance of the six GCN variants for the one-hour-ahead pick-up

https://doi.org/10.3390/math14010116

demand prediction. The results highlight a trade-off between handling sparse regions and capturing high-demand peaks. The two-branch model, Proposed WC2, achieved the best accuracy for zero-demand grids (MAE = 0.0050), confirming its effectiveness for sparse data. Proposed SB, incorporating a GRU-based temporal encoder, proved highly effective in the medium-demand range ($y \in [6, 10]$, MAE = 2.4290), indicating that the inclusion of temporal dependencies effectively enhances short-term demand prediction. In contrast, Proposed GCN-MM, which trained individual models for each community, showed strong accuracy in low-demand active grids ($y > 0$, MAE = 1.1815 and $y \in [1, 5]$, MAE = 0.8979), but its high error in zero-demand cells ($y = 0$, MAE = 0.0978) reflects overfitting and poor generalization caused by limited community-level data.

Table 3. Performance comparison between the benchmark MFCN model and six variants of the proposed GCN architecture for the next-hour or $t + 1$ pick-up demand prediction, based on the mean absolute error (MAE) along with the standard deviation of absolute errors (SD of AE).

Model	Feature Size	No of Predictions	MAE (SD of AE)					
			y = 0	**y > 0**	**y ∈ [1, 5]**	**y ∈ [6, 10]**	**y ∈ [11, 15]**	**y ∈ [16, ∞]**
MFCN	14	2.4 k	0.0057 (0.1812)	1.5180 (3.1669)	1.2136 (2.773)	2.9911 (3.669)	4.5346 (5.396)	**8.4077** (9.330)
GCN	17	1.28 M	0.0734 (0.1084)	1.2374 (1.6946)	0.9675 (0.9469)	2.8630 (1.9336)	5.5368 (2.9251)	11.811 (10.6081)
Proposed SB	17	1.28 M	0.0123 (0.1010)	1.5091 (1.8267)	1.3275 (1.3030)	**2.4290** (1.7725)	4.4716 (2.6754)	10.6398 (10.3199)
Proposed WC1	17	1.28 M	0.0053 (0.0665)	1.4034 (1.7018)	1.1153 (1.0386)	3.2156 (1.9538)	5.7806 (2.9315)	12.2893 (11.1480)
Proposed WC2	17	1.28 M	**0.0050** (0.0748)	1.4197 (1.8226)	1.1993 (1.1119)	2.8480 (1.8205)	4.4732 (2.9049)	9.7067 (9.4221)
Proposed GCN-MM	17	1.28 M	0.0978 (0.2409)	**1.1815** (1.4523)	**0.8979** (0.9249)	3.0238 (1.9411)	5.1599 (2.9228)	11.5644 (10.5682)
Proposed GCN	18	1.28 M	0.0051 (0.0749)	1.4134 (1.8139)	1.1891 (1.1049)	2.8777 (1.8549)	**4.4186** (2.5946)	9.6895 (9.2734)

Our proposed GCN, which incorporates community detection as a spatial feature, achieved the most balanced and consistent results across all demand ranges. It closely matched the best sparse-demand model ($y = 0$, MAE = 0.0051) and significantly outperformed the MFCN baseline in both low-demand (1.1891 vs. 1.2136) and medium-demand (2.8777 vs. 2.9911) ranges. Additionally, it achieved the lowest error in the high–medium demand bracket ($y \in [11, 15]$, MAE = 4.4186), confirming its stability in dense urban areas where accurate demand prediction is most critical. Although the baseline MFCN exhibited slightly lower error in the rare, extreme-demand category ($y \geq 16$; 8.4077 vs. 9.6895 for the proposed GCN), these instances account for less than 0.01% of all observations. Overall, the proposed GCNs consistently superior or competitive performance across all typical demand levels, particularly its leading results in the $y \in [11, 15]$ range, confirms that the proposed GCN offers the most balanced and generalizable framework for short-term e-scooter demand prediction in complex urban environments.

Table 4 summarizes the one-hour-ahead drop-off demand prediction results. The performance trends are consistent with those of the pick-up task, revealing a similar trade-off between sparse and high-demand regions. Proposed WC1, using a dual-branch structure, achieved the best zero-demand accuracy (MAE = 0.0047) but showed weaker performance for high-demand cases (MAE = 13.9176). Adding a weighted loss in Proposed WC2 improved stability in high-demand intervals (MAE reduced to 9.2507) while maintaining low-demand accuracy (MAE = 1.2336 for $y \in [1, 5]$). Proposed SB, which performed well in the medium-demand range ($y \in [6, 10]$, MAE = 2.4657), confirming the value of temporal dependency modeling. Proposed GCN-MM, trained separately for each community, achieved the best results in low-demand active grids ($y > 0$, MAE = 1.1087, $y \in [1, 5]$, MAE = 0.8497) but

https://doi.org/10.3390/math14010116

had high error for zero-demand cells (MAE = 0.1011), indicating overfitting from limited per-community samples. Our Proposed GCN, which integrates community detection, provided the most balanced and generalizable results. It achieved MAEs of 0.0054 for zero-demand and 9.0302 for high-demand grids, with strong performance in the high-medium range (y ∈ [11, 15], MAE = 4.2972). Compared with the baseline MFCN (y = 0, MAE = 0.0063 and y ≥ 16, MAE = 9.8085), the proposed GCN consistently outperformed across most intervals.

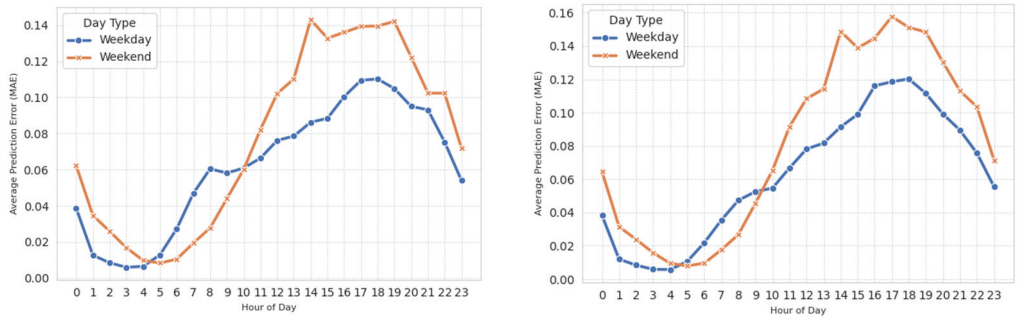
Table 4. Performance comparison between the benchmark MFCN model and six variants of the proposed GCN architecture for the next-hour or $t + 1$ drop-off demand prediction, based on the mean absolute error (MAE) along with the standard deviation of absolute errors (SD of AEs).

Model	Feature Size	No of Predictions	MAE (SD of AE)					
			y = 0	**y > 0**	**y ∈ [1, 5]**	**y ∈ [6, 10]**	**y ∈ [11, 15]**	**y ∈ [16, ∞]**
MFCN	14	2.4 k	0.0063 (0.1872)	1.5382 (2.8517)	1.2513 (2.4982)	3.1714 (3.6090)	4.5551 (4.0260)	9.8085 (9.6832)
GCN	17	1.37 M	0.0205 (0.1146)	1.4476 (1.7277)	1.2842 (1.2170)	2.4893 (1.9347)	4.5246 (2.8562)	11.6168 (12.2902)
Proposed SB	17	1.37 M	0.0235 (0.1220)	1.3856 (1.7646)	1.2123 (1.2445)	**2.4657** (1.8582)	4.7671 (2.8317)	12.2261 (12.2568)
Proposed WC1	17	1.37 M	**0.0047** (0.0625)	1.3420 (1.7380)	1.0798 (0.9123)	3.2769 (1.9168)	6.1085 (2.9796)	13.9176 (12.4829)
Proposed WC2	17	1.37 M	0.0056 (0.0752)	1.4184 (1.6779)	1.2336 (1.2186)	2.9032 (2.2759)	4.3552 (3.1804)	9.2507 (9.7960)
Proposed GCN-MM	17	1.37 M	0.1011 (0.2339)	**1.1087** (1.6985)	**0.8497** (0.9387)	3.0258 (1.8103)	5.7779 (2.8675)	13.0021 (11.5892)
Proposed GCN	18	1.37 M	0.0054 (0.0751)	1.4117 (1.6656)	1.2169 (1.1997)	3.0118 (2.1835)	**4.2972** (3.1262)	**9.0302** (9.9654)

Our proposed model outperforms the benchmark MFCN and the baseline GCN primarily because it better understands not only when e-scooter demand changes but also how different areas of the city influence each other. While the MFCN employs convolutional filters limited to fixed grid neighborhoods and the baseline GCN models only local geographic adjacency, the proposed GCN leverages graph-based connectivity derived from functional relationships between zones, enabling spatial information to propagate adaptively across the urban network. The integration of community detection as a node feature further enhances this process by clustering grids with similar mobility behavior, allowing the model to learn cross-zone dependencies that are not represented in either baseline. Moreover, the combination of GRU-based temporal encoding and GCN spatial aggregation allows the proposed framework to integrate short-term fluctuations with broader spatial trends, enhancing its generalization across varying demand levels. Consequently, the proposed GCN achieves a more coherent spatial understanding of urban mobility dynamics, resulting in lower overall prediction errors and stronger robustness than the baselines.

The overall temporal performance of the proposed GCN model for the next-hour prediction task is shown in Figure 6, which plots the average prediction error (MAE) for both (a) pick-up and (b) drop-off demand. It can be observed that for both tasks, the MAE is consistently low during off-peak hours (approximately 00:00 to 06:00). A clear divergence in performance between weekdays and weekends emerges as demand increases throughout the day. For both pick-up and drop-off predictions, the error is significantly higher on weekends, particularly during the afternoon peak from 14:00 to 20:00. This indicates that the model more accurately captures the structured, commuter-based demand on weekdays, while the less regular, leisure-driven activity on weekends results in a greater average prediction error.

https://doi.org/10.3390/math14010116



(**a**) Pick-up (**b**) Drop-off

Figure 6. Hourly average prediction error (MAE) for next-hour or (t + 1) demand: (**a**) pick-up and (**b**) drop-off, shown separately for weekday and weekend using the proposed GCN-4 model.

3.2. Next 24-h Pick-Up and Drop-Off Prediction Performance

Table 5 presents the results for the 24 h-ahead pick-up demand prediction. Overall, the performance trends are consistent with those observed in the one-hour prediction task. The two-branch models (Proposed WC1 and Proposed WC2) again demonstrated superior accuracy in sparse regions (y = 0) and moderate-demand zones, achieving MAE values of 0.0044 and 0.0050, respectively. The incorporation of a weighted loss (Proposed WC2) effectively reduced errors for higher demand levels, indicating improved robustness to data imbalance. Proposed SB performed strongly in the medium-demand range (y ∈ [6, 10], MAE = 2.3631), demonstrating that temporal dependency modeling remains beneficial even for long-term prediction. The proposed GCN, which integrates community detection as a spatial feature, achieved the most balanced performance across all demand categories. With an MAE of 0.0053 for zero-demand grids and 8.6355 for extreme high-demand cases, the proposed GCN consistently outperformed the benchmark MFCN model (MAE = 0.0072 and 11.2299, respectively). The inclusion of community structure enabled more effective spatial information propagation, allowing the model to generalize better across diverse urban regions. In contrast, proposed GCN-MM, which trained separate models for each community, achieved competitive accuracy in low-demand active grids (y > 0, MAE = 1.2128 and y ∈ [1, 5], MAE = 0.8514) but showed degraded performance in high-demand areas (MAE = 13.5130) due to limited training data and overfitting within smaller subgraphs. Therefore, the results confirm that incorporating community information enhances the stability and scalability of long-term e-scooter demand prediction.

Table 5. Performance comparison between the benchmark MFCN model and six variants of the proposed GCN architecture for the next 24 h or t + 24 pick-up demand prediction, based on the mean absolute error (MAE) along with the standard deviation of absolute errors (SD of AE).

Model	Feature Size	No of Predictions	MAE (SD of AE)					
			y = 0	**y > 0**	**y ∈ [1, 5]**	**y ∈ [6, 10]**	**y ∈ [11, 15]**	**y ∈ [16, ∞]**
MFCN	14	2.4 k	0.0072 (0.1870)	1.6850 (3.1231)	1.3307 (2.8816)	3.1609 (3.5391)	5.2500 (3.6433)	11.2299 (8.4332)
GCN	17	1.28 M	0.0139 (0.1204)	1.6534 (1.9371)	1.4911 (1.3790)	2.4091 (1.7625)	4.3537 (2.5127)	11.0606 (10.1379)
Proposed SB	17	1.28 M	0.0135 (0.1163)	1.5568 (1.9356)	1.3817 (1.3096)	**2.3631** (1.7463)	4.3800 (2.6205)	11.4834 (10.8860)
Proposed WC1	17	1.28 M	**0.0044** (0.0615)	1.4302 (1.8170)	1.1073 (0.8909)	3.3048 (2.1369)	5.9645 (2.7744)	12.2274 (11.9223)
Proposed WC2	17	1.28 M	0.0050 (0.0756)	1.4557 (1.8362)	1.2082 (1.0797)	2.8506 (1.8144)	4.8619 (2.7734)	10.3802 (9.8565)
Proposed GCN-MM	17	1.28 M	0.0924 (0.2247)	**1.2128** (1.7382)	**0.8514** (0.8686)	3.2655 (1.8970)	6.3713 (2.8842)	13.5130 (12.0103)
Proposed GCN	18	1.28 M	0.0053 (0.0799)	1.5293 (1.8994)	1.2911 (1.2978)	3.1125 (1.8276)	**4.3152** (2.4896)	**8.6355** (9.8171)

https://doi.org/10.3390/math14010116

The next 24 h drop-off prediction experiment was conducted in a similar way, and the results are presented in Table 6. The overall performance patterns were consistent with those observed in the short-term predictions. The two-branch models (Proposed WC1 and Proposed WC2) achieved strong accuracy in sparse grids, with MAE values of 0.0051 and 0.0054, respectively, and the weighted-loss variant Proposed WC2 improved robustness in higher-demand ranges (MAE = 11.4845). Proposed SB, with a GRU temporal encoder, performed the lowest error for active-demand cells (y > 0, MAE = 1.0935). In contrast, Proposed GCN-MM, trained independently for each community, showed lower overall accuracy, performing best in the low active bin (y ∈ [1, 5], MAE = 0.8154) but declining under high-demand conditions (MAE = 15.2888) due to limited training data. The baseline GCN achieved the best accuracy in the medium–high bracket (y ∈ [6, 10], MAE = 2.3938), indicating effective spatial modeling for those peaks. The proposed GCN, which incorporates community detection, provided the most balanced results overall: it achieved the best zero-demand accuracy (y = 0, MAE = 0.0050) and strong extreme-demand performance (y $\geq$ 16, MAE = 10.0982), outperforming the baseline MFCN (y = 0, MAE = 0.0078 and y $\geq$ 16, MAE = 12.4215). These results confirm that explicitly integrating community structure within a unified model enhances the capacity to capture spatial dependencies and leads to more reliable long-term e-scooter demand forecasts.

Table 6. Performance comparison between the benchmark MFCN model and six variants of the proposed GCN architecture for the next 24 h or *t* + 24 drop-off demand prediction, based on the mean absolute error (MAE) along with the standard deviation of absolute errors (SD of AEs).

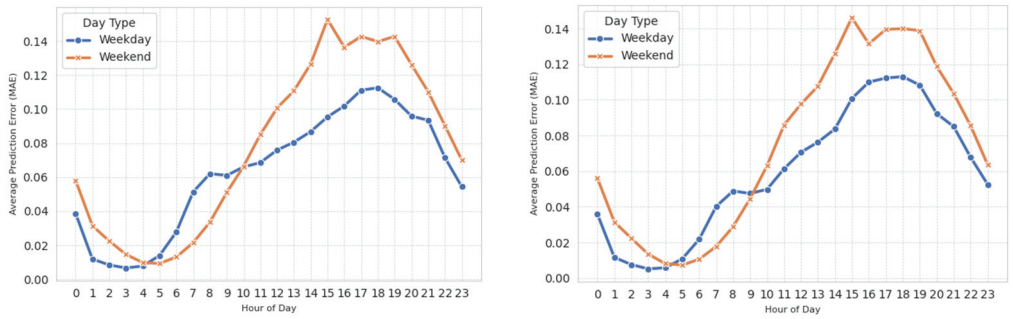
Model	Feature Size	No of Predictions	MAE (SD of AE)					
			y = 0	**y > 0**	**y ∈ [1, 5]**	**y ∈ [6, 10]**	**y ∈ [11, 15]**	**y ∈ [16, ∞]**
MFCN	14	2.4 k	0.0078 (0.1678)	1.6230 (2.567)	1.3130 (2.2940)	3.0823 (3.0603)	5.2701 (3.0756)	12.4215 (8.9490)
GCN	17	1.37 M	0.0147 (0.1264)	1.5506 (1.8655)	1.3953 (1.3368)	**2.3938** (1.7272)	4.6165 (2.7112)	13.1775 (11.4313)
Proposed SB	17	1.37 M	0.1820 (0.2270)	**1.0935** (1.6387)	0.8234 (0.8663)	2.9773 (1.8192)	6.1763 (2.8791)	15.3078 (12.4467)
Proposed WC1	17	1.37 M	0.0051 (0.0724)	1.3690 (1.7939)	1.0123 (0.8982)	3.2218 (1.9256)	6.4127 (3.0165)	15.4617 (12.5934)
Proposed WC2	17	1.37 M	0.0054 (0.0738)	1.3796 (1.8102)	1.1555 (0.9937)	2.8758 (1.7430)	5.0472 (2.7751)	11.4845 (10.0969)
Proposed GCN-MM	17	1.37 M	0.0977 (0.2202)	1.1478 (1.7006)	**0.8154** (0.8269)	3.3612 (2.0973)	6.8945 (3.0820)	15.2888 (12.4140)
Proposed GCN	18	1.37 M	**0.0050** (0.0682)	1.4455 (1.8547)	1.2290 (1.1590)	3.0437 (1.9116)	**4.6059** (2.4427)	**10.0982** (10.0180)

Overall performance for the proposed GCN model in the 24 h prediction scheme, illustrated in Figure 7, shows lower errors during off-peak hours and noticeably higher MAE on weekend afternoons for both pick-up and drop-off tasks. Within the 24 h forecasts for active regions (y > 0), the proposed GCN achieved a slightly lower MAE for drop-offs (1.4455) compared to pick-ups (1.5293), indicating stable long-range prediction capability. Figure 7 further illustrates the temporal distribution of errors, where both pick-up and drop-off predictions exhibit low MAE values during early morning hours (00:00–06:00) and increasing errors throughout the day as e-scooter activity intensifies. Higher error levels are observed during afternoon and evening peaks (14:00–20:00), particularly on weekends. This pattern suggests that weekday demand, driven primarily by structured commuting behavior, is more predictable than the irregular, leisure-oriented activity that characterizes weekends.

In summary, the performance comparison between one-hour and 24 h prediction time scales indicates that the proposed GCN performs more effectively in short-term prediction for both pick-up and drop-off tasks. For active demand regions (y > 0), the next-hour pick-up prediction achieved a Mean Absolute Error (MAE) of 1.4134, which

https://doi.org/10.3390/math14010116

slightly increased to 1.5293 for the 24 h horizon. A notable advantage of the proposed GCN over the MFCN benchmark and the baseline GCN is observed in high-demand scenarios. For instance, in predicting high-intensity drop-offs ($y \geq 16$), the proposed GCN reduced the MAE from 9.8085 (MFCN) and 11.6168 (GCN) to 9.0302 for the next-hour task and from 12.4215 (MFCN) and 13.1775 (GCN) to 10.0982 for the 24 h task. These results highlight that incorporating the community feature within the proposed GCN framework enhances spatial representation and model stability, enabling reliable performance across different temporal scales and particularly improving long-range e-scooter demand prediction.



(**a**) Pick-up (**b**) Drop-off

Figure 7. Hourly average prediction error (MAE) for next 24 hr or (t + 24) demand: (**a**) pick-up and (**b**) drop-off, shown separately for weekday and weekend using the proposed GCN-4 model.

4. Discussions

The findings show that embedding community detection within a hybrid GRU-GCN framework enhances both interpretability and predictive performance in modeling spatiotemporal micromobility demand. Compared with conventional deep learning baselines such as the MFCN and single-branch GCN models, the proposed framework provides a more robust representation of spatial dependencies and temporal patterns. By utilizing graph-theoretic representations of mobility flows, the model captures both structural and functional connectivity between urban areas, leading to more accurate and stable predictions.

The inclusion of the Louvain community feature enables the framework to identify functionally cohesive subregions such as downtown commercial zones, university areas, and recreational corridors that exhibit similar mobility behaviors. This transformation of raw origin–destination data into a modular network representation allows the model to learn the relationships between spatially connected communities more effectively. Embedding these community structures as node features allows the GCN to propagate information across the graph based on spectral graph theory, thereby modeling correlations that extend beyond adjacent geographic grids. Consequently, the model performs better in both densely and sparsely populated regions by leveraging the inherent regularities in urban mobility patterns.

Beyond predictive performance, the detected communities also enhance the interpretability and practical utility of the model. Several representative communities identified by the Louvain algorithm illustrate this point. For example, a large downtown-centered cluster captures the strong bidirectional flows between commercial blocks and transit-accessible grids, reflecting the core role of the central business district in shaping weekday mobility. Another community emerges around the University of Calgary, where trip patterns exhibit strong evening and weekend activity associated with student travel and campus-related trips. A third community aligns with the Bow River pathway network, showing higher leisure-oriented usage on weekends. These examples demonstrate that the community structure highlights meaningful functional regions and provides a compact representation of mobility behavior that is not available from geographic adjacency alone.

https://doi.org/10.3390/math14010116

The resulting interpretability helps urban planners identify mobility anchors, evaluate network connectivity, and design infrastructure or staging strategies tailored to corridor-level demand patterns.

The GRU component complements the graph structure by capturing recurring temporal dependencies, such as daily commuting and weekend activity patterns. This sequential learning process enhances the model's ability to handle both short-term fluctuations and longer-term seasonal trends. The higher prediction error observed during weekend afternoons reflects the inherent variability of leisure-related trips, which are often influenced by social or environmental factors not explicitly included in the current model. These findings are consistent with prior studies indicating that leisure and discretionary trips display greater unpredictability than routine commuting behavior. Although the dataset represents a single summer period, the temporal patterns suggest that extending the model to multi-seasonal or year-long datasets would likely improve robustness by capturing additional seasonal cycles. In such cases, incorporating exogenous variables such as weather or event indicators may be necessary to account for greater temporal variability.

From a methodological perspective, the adoption of a multi-branch structure that combines classification and regression tasks contributes to improved robustness under sparse data conditions. This design enables the model to first detect potential activity zones and then estimate the expected demand within them, reducing the sensitivity to data imbalance. Moreover, the use of graph convolution over irregular spatial structures allows information to diffuse efficiently across the network, accommodating complex urban topologies that cannot be represented well by fixed grid-based methods. Together, these components enable the proposed framework to capture nonlinear spatial and temporal interactions without relying on rigid spatial partitions.

The practical implications of these findings are significant. Improved demand prediction can support operational decisions such as fleet redistribution, charging logistics, and dynamic pricing, while also informing policymakers on issues related to service accessibility and urban equity. The enhanced interpretability offered by mobility-based communities further enables planners to identify high-demand corridors, evaluate functional connectivity between neighborhoods, and prioritize infrastructure upgrades in regions that serve as mobility anchors.

In addition to interpretability, computational efficiency is an important consideration for real-time deployment. The Louvain algorithm is executed only once during preprocessing, and its near-linear complexity on sparse graphs ensures scalability for large urban networks. During training, the GRU-GCN architecture requires moderately more computation than MFCN or single-branch GCN models, but the inference phase, critical for real-time forecasting, consists of a single forward pass with no need for repeated community detection or graph reconstruction. This keeps runtime efficiency suitable for near-real-time operational use.

Finally, although the study focuses on Calgary, the underlying methodology is expected to generalize well to other cities. The approach relies on three readily available data sources: grid-level demand counts, origin–destination flows, and simple temporal contextual features. Because both the spatial graph and the community structure are derived directly from local travel patterns, the model adapts naturally to different urban forms without the need for manual spatial tuning. In practical applications, transferring the framework to a new city would require either full retraining with local data or a domain-adaptation strategy in which GRU weights are partially retained while GCN layers and community assignments are recalibrated using the new origin–destination network. Even limited local data is typically sufficient to reconstruct the flow graph and detect mobility communities, enabling the framework to generalize across cities with varying

https://doi.org/10.3390/math14010116

scales, densities, and street layouts. Beyond its application to micromobility, the proposed approach illustrates how advanced mathematical modeling and network-based learning can be leveraged to address complex problems in urban analytics.

5. Conclusions

This study proposed a hybrid deep learning framework that combines Graph Convolutional Networks (GCNs) and Gated Recurrent Units (GRUs) with community detection for short- and long-term e-scooter demand prediction. Using trip data from Calgary, Canada, the proposed GCN model achieved an average reduction of 11.8% in mean absolute error (MAE) compared with the benchmark MFCN model. By incorporating community-based spatial features derived from the Louvain algorithm, the model successfully captured functional relationships among urban regions, resulting in improved predictive accuracy and interpretability. Unlike conventional grid-based or purely adjacency-driven approaches, the proposed framework embeds mobility-driven communities as learnable spatial features, enabling the model to jointly capture local spatial proximity and higher-level functional similarity across urban areas. The resulting community-aware representation also enhanced interpretability by highlighting mobility clusters that align with meaningful functional areas of the city.

Despite these promising results, several limitations remain. The model was developed and validated using data from a single city, and its applicability to other contexts requires further evaluation. Nevertheless, the framework is inherently transferable because it constructs both the spatial graph and community structure directly from observed origin–destination flows, which are commonly available in micromobility systems. This design allows the framework to move beyond fixed spatial partitions and adapt naturally to city-specific mobility patterns. Applying the model to a new city would involve reconstructing the flow network, redetecting communities, and retraining or fine-tuning the model with local demand data, making the adaptation process straightforward. Future research should evaluate the model's generalizability by testing it in cities with differing socioeconomic and spatial characteristics.

Additionally, the current framework does not explicitly account for exogenous factors such as weather, traffic, or public events, which are known to influence micromobility usage. Including these variables could enhance both prediction precision and adaptability to short-term fluctuations. The dataset used in this study covers a summer period, and incorporating multi-seasonal or year-long data may improve robustness by capturing seasonal variations and weather-driven behavioral shifts. Extending the temporal scope would allow the model to learn additional recurring patterns, though such expansion may require integration of contextual variables to account for increased variability. Another limitation lies in the static nature of the Louvain-based community detection approach, which does not capture temporal evolution in mobility patterns. Future extensions could explore dynamic or time-aware community detection to reflect evolving functional regions within cities. Finally, exploring attention mechanisms, graph transformers, or other advanced architectures could further improve learning efficiency and the interpretability of spatial dependencies.

In conclusion, the proposed GRU-GCN framework represents a promising step forward in the modeling of spatiotemporal demand for micromobility systems. The key contributions of this study lie in (i) integrating mobility-informed community structures into graph-based learning, (ii) jointly modeling temporal dynamics and functional spatial relationships within a unified framework, and (iii) addressing demand sparsity through a two-stage prediction design. By integrating community-aware graph structures with temporal sequence modeling, the framework provides both methodological rigor and practical relevance. The computational efficiency of the approach, particularly during inference where only a single forward pass is required, further supports its suitability

https://doi.org/10.3390/math14010116

for near-real-time deployment in operational settings. Overall, the proposed framework advances current spatiotemporal demand prediction practices by shifting from static spatial representations toward community-aware, functionally interpretable, and transferable modeling of urban mobility demand. This approach contributes to the advancement of data-driven urban analytics, supporting the development of intelligent, and sustainable urban mobility systems.

Author Contributions: Conceptualization, M.G.D. and S.P.; methodology, K.P. and M.M.Z.; software, M.M.Z.; validation, M.M.Z., K.P. and S.P.; formal analysis, K.P. and S.P.; investigation, M.M.Z., K.P. and S.P.; resources, M.M.Z.; data curation, M.M.Z.; writing—original draft preparation, M.M.Z.; writing—review and editing, S.P., K.P. and M.G.D.; visualization, M.M.Z.; supervision, K.P. and S.P.; project administration, S.P.; funding acquisition, S.P. and M.G.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially supported by the National Research Council of Thailand (NRCT) Thailand, the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant (RGPIN/03037-2022), and Chiang Mai University.

Data Availability Statement: Data available on request due to Non-Disclosure Agreement (NDA).

Acknowledgments: We thank the City of Calgary for providing the e-scooter data. The research was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant (RGPIN/03037-2022), the National Research Council of Thailand (NRCT) through the Hub of Talents in AI and Emerging Technology (AI-NEXT), and Chiang Mai University.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Gössling, S. Integrating e-Scooters in Urban Transportation: Problems, Policies, and the Prospect of System Change. *Transp. Res. Part D Transp. Environ.* **2020**, *79*, 102230. [CrossRef]
2. Chang, A.Y.; Miranda-Moreno, L.; Clewlow, R.; Sun, L. Trend or Fad? Deciphering the Enablers of Micromobility in the U.S. *Transp. Res. Rec.* **2019**, *2673*, 1–12.
3. McKenzie, G. Spatiotemporal Comparative Analysis of Scooter-Share and Bike-Share Usage Patterns in Washington. *D.C. J. Transp. Geogr.* **2019**, *78*, 19–28. [CrossRef]
4. Hosseinzadeh, A.; Algomaiah, M.; Kluger, R.; Li, Z. E-Scooters and Sustainability: Investigating the Relationship between the Density of E-Scooter Trips and Characteristics of Sustainable Urban Development. *Sustain. Cities Soc.* **2020**, *66*, 102624. [CrossRef]
5. Sedor, A.; Oriold, J. *Shared E-Bike and E-Scooter Final Pilot Report*; City of Calgary: Calgary, AB, Canada, 2020.
6. Abduljabbar, R.L.; Liyanage, S.; Dia, H. The Role of Micro-Mobility in Shaping Sustainable Cities: A Systematic Literature Review. *Transp. Res. Part D Transp. Environ.* **2021**, *92*, 102734. [CrossRef]
7. Palm, M.; Farber, S.; Shalaby, A.; Young, M. Equity Analysis and New Mobility Technologies: Toward Meaningful Interventions. *J. Plan. Lit.* **2021**, *36*, 16–29. [CrossRef]
8. Trivedi, T.K.; Liu, C.; Antonio, A.L.M.; Wheaton, N.; Kreger, V.; Yap, A.; Schriger, D.; Elmore, J.G. Injuries Associated with Standing Electric Scooter Use. *JAMA Netw. Open* **2019**, *2*, e187381. [CrossRef]
9. Hollingsworth, J.; Copeland, B.; Johnson, J.X. Are E-Scooters Polluters? The Environmental Impacts of Shared Dockless Electric Scooters. *Environ. Res. Lett.* **2019**, *14*, 084031. [CrossRef]
10. Moreau, H.; de Meux, L.J.; Zeller, V.; D'Ans, P.; Ruwet, C.; Achten, W.M.J. Dockless E-Scooter: A Green Solution for Mobility? Comparative Case Study between Dockless E-Scooters, Displaced Transport, and Personal E-Scooters. *Sustainability* **2020**, *12*, 1803. [CrossRef]
11. Abdelwahab, B.; Palm, M.; Shalaby, A.; Farber, S. Evaluating the Equity Implications of Ridehailing through a Multi-Modal Accessibility Framework. *J. Transp. Geogr.* **2021**, *95*, 103147. [CrossRef]
12. Liao, F.; Correia, G. Electric Carsharing and Micromobility: A Literature Review on Their Usage Pattern, Demand, and Potential Impacts. *Int. J. Sustain. Transp.* **2020**, *14*, 686–700. [CrossRef]
13. Zhang, S.; Tong, H.; Xu, J.; Maciejewski, R. Graph Convolutional Networks: A Comprehensive Review. *Comput. Soc. Netw.* **2019**, *6*, 11. [CrossRef] [PubMed]

14. Wu, Z.; Pan, S.; Chen, F.; Long, G.; Zhang, C.; Yu, P.S. A Comprehensive Survey on Graph Neural Networks. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *32*, 4–24. [CrossRef]
15. Phithakkitnukoon, S.; Patanukhom, K.; Demissie, M.G. Predicting Spatiotemporal Demand of Dockless E-Scooter Sharing Services with a Masked Fully Convolutional Network. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 773. [CrossRef]
16. Ham, S.W.; Cho, J.-H.; Park, S.; Kim, D.-K. Spatiotemporal Demand Prediction Model for E-Scooter Sharing Services with Latent Feature and Deep Learning. *Transp. Res. Rec.* **2021**, *2675*, 34–43. [CrossRef]
17. Sahnoon, M.; Manuel, A.; Demissie, M.G.; Souza, R. UNET and UNETR Based Frameworks for Predicting the Short-Term Spatiotemporal Demand of E-Scooter Sharing Services. In Proceedings of the 2024 IEEE 27th International Conference on Intelligent Transportation Systems (ITSC), Edmonton, AB, Canada, 24–27 September 2024; pp. 2804–2811. [CrossRef]
18. Kim, S.; Choo, S.; Lee, G.; Kim, S. Predicting Demand for Shared E-Scooter Using Community Structure and Deep Learning Method. *Sustainability* **2022**, *14*, 2564. [CrossRef]
19. Liu, M.; Mathew, J.K.; Seeder, S.; Li, H.; Bullock, D.M. Analysis of E-Scooter Trips and Their Temporal Usage Patterns. *ITE J.* **2019**, *89*, 44–51.
20. Bai, S.; Jiao, J. Dockless E-Scooter Usage Patterns and Urban Built Environments: A Comparison Study of Austin, TX, and Minneapolis, MN. *Travel Behav. Soc.* **2020**, *20*, 264–272. [CrossRef]
21. Lee, M.; Chow, J.Y.J.; He, B.Y.; Lee, H. Forecasting E-Scooter Competition with Direct and Access Trips by Mode and Distance in New York City. *Transp. Res. Part D Transp. Environ.* **2022**, *108*, 103300.
22. Yang, Y.; Shao, X.; Zhu, Y.; Yao, E.; Liu, D.; Zhao, F. Short-Term Forecasting of Dockless Bike-Sharing Demand with the Built Environment and Weather. *J. Adv. Transp.* **2023**, *2023*, 7407748. [CrossRef]
23. Li, X.; Xu, Y.; Zhang, X.; Shi, W.; Yue, Y.; Li, Q. Improving Short-Term Bike Sharing Demand Forecast through an Irregular Convolutional Neural Network. *Transp. Res. Part C Emerg. Technol.* **2023**, *147*, 103984. [CrossRef]
24. Song, J.-C.; Hsieh, I.Y.L.; Chen, C.S. Sparse Trip Demand Prediction for Shared E-Scooter Using Spatio-Temporal Graph Neural Networks. *Transp. Res. Part D Transp. Environ.* **2023**, *125*, 103962. [CrossRef]
25. Xu, Y.; Zhao, X.; Zhang, X.; Paliwal, M. Real-Time Forecasting of Dockless Scooter-Sharing Demand: A Spatio-Temporal Multi-Graph Transformer Approach. *IEEE Trans. Intell. Transp. Syst.* **2023**, *24*, 8507–8518. [CrossRef]
26. Jiang, J.; Han, C.; Zhao, W.X.; Wang, J. PDFormer: Propagation Delay-aware Dynamic Long-range Transformer for Traffic Flow Prediction. In Proceedings of the AAAI Conference on Artificial Intelligence, Washington, DC, USA, 7–14 February 2023; Volume 37, pp. 4365–4373.
27. Lan, S.; Ma, Y.; Huang, W.; Wang, W.; Yang, H.; Li, P. DSTAGNN: Dynamic Spatial-Temporal Aware Graph Neural Network for Traffic Flow Forecasting. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 20459–20469.
28. Singh, G.; Pal, S.; Kumar, Y. Integrated Spatio-Temporal Graph Neural Network for Traffic Forecasting. *Appl. Sci.* **2024**, *14*, 11477. [CrossRef]
29. Ke, J.; Zheng, H.; Yang, H.; Chen, X. Short-Term Forecasting of Passenger Demand under On-Demand Ride Services: A Spatio-Temporal Deep Learning Approach. *Proc. AAAI Conf. Artif. Intell.* **2019**, *33*, 1394–1401. [CrossRef]
30. Zhao, J.; Zhang, H.; Sun, J.; Chen, H. Prediction of Urban Taxi Travel Demand by Using Hybrid Dynamic Graph Convolutional Network. *Sensors* **2022**, *22*, 5982. [CrossRef]
31. Yang, Y.; Xu, W.; Wu, J.; Zhang, Y. Dynamic Graph Convolutional Network-Based Prediction of Urban Grid-Level Taxi Demand-Supply Imbalance. *ISPRS Int. J. Geo-Inf.* **2024**, *13*, 34. [CrossRef]
32. Blondel, V.D.; Guillaume, J.-L.; Lambiotte, R.; Lefebvre, E. Fast Unfolding of Communities in Large Networks. *J. Stat. Mech. Theory Exp.* **2008**, *2008*, P10008. [CrossRef]
33. Newman, M.E.J. Modularity and Community Structure in Networks. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 8577–8582. [CrossRef]
34. Dastjerdi, A.M.; Morency, C. Bike-Sharing Demand Prediction at Community Level under COVID-19 Using Deep Learning. *Sensors* **2022**, *22*, 1060. [CrossRef]
35. Pourabbasi, E.; Majidnezhad, V.; Farzi Veijouyeh, N.; Taghavi Afshord, S.; Jafari, Y. A Novel Intelligent Fuzzy-AHP Based Evolutionary Algorithm for Detecting Communities in Complex Networks. *Soft Comput.* **2024**, *28*, 7251–7269. [CrossRef]
36. Cho, K.; Van Merriënboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; Bengio, Y. Learning Phrase Representations Using RNN Encoder-Decoder for Statistical Machine Translation. In Proceedings of the EMNLP, Doha, Qatar, 25–29 October 2014; pp. 1724–1734.
37. Elman, J.L. Finding Structure in Time. *Cogn. Sci.* **1990**, *14*, 179–211. [CrossRef]
38. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef]
39. Kipf, T.N.; Welling, M. Semi-Supervised Classification with Graph Convolutional Networks. In Proceedings of the 5th International Conference on Learning Representations (ICLR 2017)-Conference Track Proceedings, Toulon, France, 24–26 April 2017.
40. Hammond, D.K.; Vandergheynst, P.; Gribonval, R. Wavelets on Graphs via Spectral Graph Theory. *Appl. Comput. Harmon. Anal.* **2011**, *30*, 129–150. [CrossRef]

41. Kingma, D.P.; Ba, J.L. Adam: A Method for Stochastic Optimization. In Proceedings of the 3rd International Conference on Learning Representations (ICLR 2015)-Conference Track Proceedings, San Diego, CA, USA, 7–9 May 2015.
42. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *J. Mach. Learn. Res.* **2014**, *15*, 1929–1958.

MDPI AG
Grosspeteranlage 5
4052 Basel
Switzerland
Tel.: +41 61 683 77 34

Mathematics Editorial Office
E-mail: mathematics@mdpi.com
www.mdpi.com/journal/mathematics

Academic Open
Access Publishing

mdpi.com

ISBN 978-3-7258-6833-9