Received 18 February 2023
Revised 4 April 2023
Accepted 20 April 2023

Inventory control strategy: based on demand forecast error

Yue Zhou, Xiaobei Shen and Yugang Yu
School of Management, University of Science and Technology of China, Hefei, China

Abstract

Purpose – This study examines the relationship between demand forecasting error and retail inventory management in an uncertain supplier yield context. Replenishment is segmented into off-season and peak-season, with the former characterized by longer lead times and higher supply uncertainty. In contrast, the latter incurs higher acquisition costs but ensures certain supply, with the retailer's purchase volume aligning with the acquired volume. Retailers can replenish in both phases, receiving goods before the sales season. This paper focuses on the impact of the retailer's demand forecasting bias on their sales period profits for both phases.

Design/methodology/approach – This study adopts a data-driven research approach by drawing inspiration from real data provided by a cooperating enterprise to address research problems. Mathematical modeling is employed to solve the problems, and the resulting optimal strategies are tested and validated in real-world scenarios. Furthermore, the applicability of the optimal strategies is enhanced by incorporating numerical simulations under other general distributions.

Findings – The study's findings reveal that a greater disparity between predicted and actual demand distributions can significantly reduce the profits that a retailer-supplier system can earn, with the optimal purchase volume also being affected. Moreover, the paper shows that the mean of the forecasting error has a more substantial impact on system revenue than the variance of the forecasting error. Specifically, the larger the absolute difference between the predicted and actual means, the lower the system revenue. As a result, managers should focus on improving the quality of demand forecasting, especially the accuracy of mean forecasting, when making replenishment decisions.

Practical implications – This study established a two-stage inventory optimization model that simultaneously considers random yield and demand forecast quality, and provides explicit expressions for optimal strategies under two specific demand distributions. Furthermore, the authors focused on how forecast error affects the optimal inventory strategy and obtained interesting properties of the optimal solution. In particular, the property that the optimal procurement quantity no longer changes with increasing forecast error under certain conditions is noteworthy, and has not been previously noted by scholars. Therefore, the study fills a gap in the literature.

Originality/value – This study established a two-stage inventory optimization model that simultaneously considers random yield and demand forecast quality, and provides explicit expressions for optimal strategies under two specific demand distributions. Furthermore, the authors focused on how forecast error affects the optimal inventory strategy and obtained interesting properties of the optimal solution. In particular, the property that the optimal procurement quantity no longer changes with increasing forecast error under certain conditions is noteworthy, and has not been previously noted by scholars. Therefore, the study fills a gap in the literature.

Keywords Forecast error, Supply uncertainty, Inventory control, Seasonal goods

Paper type Research paper

1. Introduction

The growth of e-commerce has led to a more diverse set of consumer demands for products. Accurate forecasting market demand and managing inventory is now crucial for ensuring the

Xiaobei Shen acknowledges financial support from the National Natural Science Foundation of China [Grants72222015 and 72171215] and theYouth Innovation Key ProjectFund of the University of Science and Technology of China [GrantYD2040002018]. Yugang Yu acknowledges financial support from the National Natural Science Foundation of China [Grants72091215/72091210 and 71921001].

Modern Supply Chain Research and Applications
Vol. 5 No. 2, 2023
pp. 74-101
Emerald Publishing Limited
2631-3871
DOI 10.1108/MSCRA-02-2023-0009

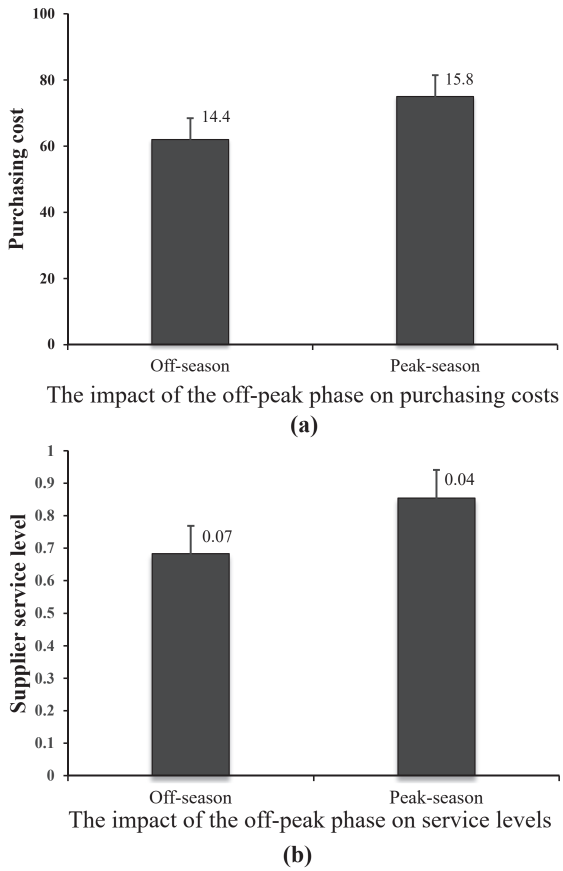
smooth operation and improved profits of businesses. While scholars have proposed various methods utilizing machine learning and deep learning to enhance prediction accuracy (Spiliotis *et al.*, 2022), uncertainty in demand predictions and errors remain prevalent, causing substantial losses in profits. Studies have demonstrated the negative impact of inaccurate predictions on transportation facilities (Cruz and Sarmento, 2020) and airlines' revenue (Fiig *et al.*, 2019). To improve competitiveness and sales, manufacturing enterprises may increase the number of product categories, but this increase in stock-keeping units (SKUs) can also lead to prediction errors and diminished decision quality (Wan and Sanders, 2017), further reducing profits. Consequently, reducing prediction error has become an important issue that must not be overlooked.

Furthermore, suppliers can be unreliable (Dada *et al.*, 2007). The manufacturing process can be prolonged due to unexpected failures of production infrastructure, unscheduled maintenance, shortages of raw materials and disorganized schedules. Delays in transportation logistics can also occur due to road congestion or restrictions. In addition, low-skilled workers with high turnover rates can result in quality and output problems (Ivanov *et al.*, 2017). The adverse impact of supply uncertainty on business performance has garnered significant attention from scholars in the field of supply chain management (Tomlin, 2006; Hu and Feng, 2017; Salem and Haouari, 2017). This study therefore focuses on the effect of demand forecast error on inventory decisions under the backdrop of supply uncertainty, with the aim of enhancing the profits of enterprises and providing valuable insights for future research in this field.

This research aims to tackle the issue of inventory optimization in the presence of uncertainty in both supply and demand. The study draws inspiration from a cooperative enterprise that sells primarily down jackets through online channels and utilizes real-world operational data of the enterprise. The primary feature of down jackets is their strong seasonality, with high sales during winter and low sales in summer. Additionally, the production cycle for down jackets is quite long, taking at least one month from the procurement of raw materials to the product being put up for sale. Therefore, enterprises need to adopt a stock-up approach to cope with this situation. Before the peak-season, enterprises need to purchase sufficient quantities of raw materials, purchase and store enough down jackets to ensure there is enough inventory to meet consumer demand when the peak-season arrives. Moreover, due to the long production cycle for down jackets, if the enterprise misses the production cycle, they may not be able to put enough products on the shelves before the peak-season. Therefore, stockpiling in advance is a necessary measure to ensure normal operation. In fact, the enterprise can use observed market signals to improve its demand forecasting accuracy. When purchasing during the off-season, the enterprise has less market information; when purchasing during the peak-season, the enterprise has more market information as it is closer to the sales season, allowing for more accurate demand forecasting and adjustments to order quantities.

Therefore, the study takes into account the existence of demand forecast errors, with the enterprise using observed market signals to improve its demand forecasting accuracy. Furthermore, the supply uncertainty faced by the enterprise also varies between peak and off- seasons, with lower procurement costs and lower supplier-service level during the off-season and higher procurement costs and higher supplier-service level during the peak-season. Utilizing actual sales, inventory and purchasing data from the cooperating enterprise, this study examines and tests the aforementioned observations. The results of independent sample T-tests demonstrate that purchasing costs in the low-season $(mean = 62, SD = 14.4)$ are significantly lower than those in the peak-season $(mean = 75, SD = 15.8)$ $(t(97) = -2.28, p < 0.05)$. Additionally, a covariance analysis indicates that the mean differences persist even when controlling for the number of purchased items. Furthermore, the results of the independent sample T-tests show that the supplier-service level in the low-season ($mean = 0.683, SD = 0.07$) is significantly lower than that in the peak-season ($mean = 0.854, SD = 0.04$) ($t(97) = -3.863, p < 0.01$). The covariance analysis also confirms that the mean differences persist when controlling for the number of purchased items. Hence, the study reveals, with a high level of certainty that the purchasing costs are significantly higher and supplier-service level is significantly lower in the peak-season than that in the low-season, as indicated in the following Figure 1.

The supplier-service level in this study refers to the ratio of the number of qualified products received by retailers within the contract period to the total procurement quantity. This result is consistent with intuition: during the off-season, there are fewer production lines and workers in the factory, resulting in lower production efficiency and quality. The procurement lead time in the purchasing contract between retailers and suppliers is also longer, leading to lower procurement costs. Apart from lower production efficiency in the factory, due to the longer lead time, there is more time for repairing defective products, which results in retailers having higher requirements for product quality during acceptance inspection. This also leads to lower supplier service levels during the off-season, while the opposite is true during the peak-season.



Source(s): Author's own work

Figure 1.
Supply uncertainty under different stages

Therefore, our research is built on such a practical background and seeks to investigate the following three questions:

RQ1. How does the demand forecasting error during the off-season impact the two-stage inventory decision and the optimal system profit?

RQ2. Are there any differences in the impact of different types of forecasting error on the optimal solution?

RQ3. How does the proposed inventory strategy perform in the real data of a cooperative enterprise?

To address these questions, the study establishes a two-stage inventory system composed of a single supplier and a single retailer, with the objective of maximizing system profit. The replenishment process is comprised of two stages, and the study considers the variation in supplier determinacy and the retailer's knowledge of demand information in the market. Firstly, the optimal inventory decision is obtained through optimization under two common demand distribution assumptions and the properties of the optimal solution are analyzed. Secondly, the demand forecasting error is divided into two types, and through sensitivity analysis, the study examines the impact of each error on the optimal inventory decision and profit, as well as the effect of other model parameters. The study finds that the mean error in demand forecasting results in a faster decline in the optimal profit compared to the variance error, thereby suggesting that enterprises may choose demand forecasting with a smaller mean deviation to secure benefits. Finally, the study applies the proposed optimal inventory control decision to the historical sales and procurement data of a cooperative enterprise, and finds that it can bring about a more than 7% increase in profit compared to the enterprise's current strategy.

Through theoretical analysis and empirical investigation, we arrive at the following conclusions: (1) The optimal replenishment strategy, under the assumption of either a two-point demand distribution or a uniform distribution, is an order-up-to strategy with an order-up-to level that is contingent upon the demand distribution. (2) Both the mean and variance of demand forecasting error have a detrimental impact on system profits, with mean error having a more pronounced effect than variance error. As a result, companies should focus on enhancing the accuracy of their mean demand forecasts. (3) In most scenarios, the optimal quantities for the two stages are inversely related to changes in forecasting error. However, when the forecast mean is lower than the true mean and the difference is substantial, the purchase quantity becomes independent of forecasting error.

Theoretically, this study reveals the optimal strategies for a two-stage inventory model that takes into account stochastic yields and demand forecast quality. The study also provides an interesting finding that under certain conditions, the optimal purchase quantity does not depend on the forecast error. This means that the retailer does not adjust their order quantity as demand forecast accuracy further deteriorates. From a practical perspective, the results of this study can help managers choose better demand forecasting methods and optimize inventory levels to increase business efficiency.

The remainder of the paper is structured as follows: In Section 2, a comprehensive review of the relevant literature is provided. In Section 3, the problem of interest is outlined and a two-stage inventory model that accounts for demand forecasting updates is developed. Under the assumption of two-point and uniform demand distributions, the model is solved to obtain the optimal purchasing decision. Section 4 presents an analysis of the sensitivity of the two types of forecasting errors and other relevant parameters. The results of a numerical study are presented in Section 5. In Section 6, the model's performance is evaluated based on real-life data from a cooperative enterprise. Finally, the conclusion summarizes the key findings of the research and provides insights into potential avenues for future work.

2. Literature review

The following literature review is conducted from the perspectives of prediction error and supply uncertainty.

Accurate demand prediction is crucial for effective inventory control. In recent years, researchers have explored various demand prediction methods, with the aim of improving the performance of prediction models (Bunning *et al.*, 2020; Lakshmanan *et al.*, 2019; Petropoulos *et al.*, 2019). Effective methods for improving prediction accuracy have been identified, such as prediction combination (Barrow and Kourentzes, 2016) and demand source decomposition (Bruggen *et al.*, 2021). Despite these efforts, fitting real demand remains a challenging task (Spiliotis *et al.*, 2022; Prak *et al.*, 2021; Petropoulos *et al.*, 2019) and the prediction of mean and variance also significantly affects the prediction model's performance.

In response, some researchers have investigated the impact of prediction error on demand prediction. For example, Yuan *et al.* (2020) studied the impact of different prediction models in two parallel supply chains with a competitive relationship on the bullwhip effect and identified the conditions under which a retailer should choose a particular model. However, the prediction methods in this study were limited to moving average (MA), exponential smoothing (ES) and minimum mean square error (MMSE). Sanders and Graman (2016) studied the impact of amplified prediction error on supply chain costs and found that the amplification of prediction bias had a greater impact on supply chain costs than the amplification of prediction standard deviation, and that sharing demand prediction information could mitigate this impact.

Our research focuses on the impact of demand forecasting error on inventory control strategies in the field of management, an area that has received limited attention despite a wealth of research on optimizing prediction models. The calculation of replenishment decisions depends on the deviation between predicted and actual demand, making this an important gap in the field of research. Building on the findings of Sanders and Graman (2016), our study considers both forecasting bias and variance and directly examines their impact on inventory and firm profits. We compare the relative magnitude of these effects to aid decision-making.

In the realm of inventory control strategies, several factors contribute to supply uncertainty, including lead time, yield, cost and the risk of supply interruption (Svoboda *et al.*, 2021). A significant body of literature has explored optimal strategies in various scenarios. For example, Song (1994) and Kaplan (1970) proposed methods for setting safety stock in the presence of random lead time, while Rosling (2002) proposed an extension of the (s, S) structure for variable costs. Review articles by Paul *et al.* (2016) and Ivanov *et al.* (2017) examine the impact of supply chain interruption on inventory control.

This study focuses on the random yield of suppliers and the interplay between random yield and variable capacity. Previous works, such as Wang and Gerchak (1996), have shown that the order-up-to policy is optimal when only considering variable capacity; while the reorder point policy is optimal when both variable capacity and random yield are taken into account. Berling and Sonntag (2022) analyzed inventory systems with random yields and introduced the concept of rework, where products that do not meet standards are delivered at a later time. The authors used decomposition algorithms to determine the optimal base inventory level; however, their study is limited to a single stage and does not consider updates from the production stage.

Wu *et al.* (2022) studied the production and inventory planning problem for multiple products and stages, where the output of each stage is random, and proposed a heuristic algorithm. Xu *et al.* (2020) determined the optimal order quantity for each supply source under a multisource selection model with random yields and examined the impact on the retailer's service level. While the study extends to information updates, it focuses on the arrival information of previous orders rather than demand information. Liu *et al.* (2020) and Deng and Zheng (2020) considered the retailer's loss aversion and risk aversion in inventory control problems with random yields and random demands, respectively. Liu *et al.* (2020) employed a newsvendor model, while Deng and Zheng (2020) added chance constraints.

This study extends the previous studies of the retailer's replenishment problem with random yields in two stages by considering the updating of the retailer's demand knowledge. Our study simultaneously investigates the relationship between the optimal purchase quantity and different prediction errors, offering a novel perspective on the study of inventory models with random demand and yield.

Overall, in light of prior research that has considered the random yield and demand in inventory control, this study aims to expand upon these studies by incorporating the retailer's updating of market demand knowledge. Our work provides a novel perspective on the relationship between optimal purchase quantity and prediction error, making a significant contribution to the literature in this area and offering a fertile ground for future research to build upon.

3. Two-stage inventory model considering predictive update

3.1 Problem description and model establishment

In this study, we model a supply chain system comprising a supplier and a main retailer. The supplier produces and supplies the product to the retailer, who sells the product in the selling season. The system is characterized by three time periods, with the first two being production periods and the third being the selling period. The first production period is considered as the off-season, which is farther from the selling season, and the second production period is referred to as the peak-season, which is closer to the selling season. The supplier produces the products based on the retailer's orders during both production periods, and delivers them to the retailer's warehouse via a centralized shipment. Hence, the retailer has the ability to decide on its order quantity by forecasting market demand in both production periods. The demand forecast in the first period, represented by $\phi_0(x_0)$, is uncertain and has a wider range of possible demand. However, in the second period, the retailer can gather more updated market and external information through research and presales, which improves the demand forecast, represented by $\phi(x|x_0)$.

The retailer in the model engages in purchases in both the first and second stages of production. Prior to making these purchases, the retailer has a certain level of understanding regarding the supplier, namely, the modes of production in the two stages differ. Specifically, the production mode during the first stage is characterized by uncertainty, with the supplier-service level being a random variable (ε) that ranges between 0 and 1. In contrast, the production mode during the second stage is defined, with the supplier-service level guaranteed at 100% and higher purchasing costs incurred by the retailer if they opt to purchase during this stage.

These assumptions are reasonable. Firstly, the off-season is longer than the peak-season, as is the case for down jacket retailers who only operate during the winter season, which only accounts for a quarter of the year. During the off-season, retailers have longer lead times on purchase contracts with suppliers, meaning they have more time to send defective products back to the supplier's factory for repairs. Furthermore, the off-season purchase orders come with lower prices, which means retailers can tolerate lower supplier service levels. During the peak-season, retailers can obtain more external market-related information through research and presales to improve demand forecasting. They can adjust their purchasing plans accordingly, such as adding orders to suppliers. Because the sales season is about to begin, the lead time on purchase contracts will be shortened to meet consumer demand in a timely manner. In addition, to meet the delivery schedule, suppliers will open more production lines and recruit more workers to increase production, which leads to an increase in procurement costs in the second stage.

Many products with long production cycles face similar situations, such as traditional book publishing, which, like the down jacket industry, also has two production modes (Donohue, 2000). The fixed costs of installing printers and related software are high, but the demand for a particular book is often uncertain during printing, especially when the book's topic is trendy. When updated demand forecasts exceed initial expectations over time, a second production run is needed to print another batch of books to meet consumer demand.

Thus, in the context of retail operations, decision-makers are faced with a trade-off between achieving higher levels of accuracy and stability in their forecasts and supplier service levels, versus incurring higher unit purchase costs. In the first stage of the purchasing process, if the retailer orders too many products, they risk incurring holding costs due to stockpiling. Conversely, if the retailer orders too few products, they may face higher unit purchase costs in the second stage in order to make up for any shortages.

The retailer's purchase quantities in both stages are dependent on the results of their market forecast, which can be generated through a variety of methods, including regression analysis and random forests (Spiliotis *et al.*, 2022; Petropoulos *et al.*, 2019). However, no matter what forecasting method is used, it is impossible to accurately predict future demand. If the retailer's forecasted demand exceeds actual demand, this can lead to excessive stockpiling and increased costs, while the opposite scenario results in stockouts and decreased profits. Given the significance of forecast error to the retailer's optimal decision and resulting profit, this study aims to investigate the impact of market demand forecast error on purchasing decisions and total profits in the presence of supply uncertainties.

In this study, a two-stage optimization problem is formulated with the aim of maximizing the expected profit generated by the system. The model is structured as follows:

$$\max_{q_1 \geq 0} \pi(q_1) = \int_0^1 \left\{ -c_1 q_1 \varepsilon + \int_{x_0=0}^{\infty} H(x_0, q_1)\phi_0(x_0)dx_0 \right\} f(\varepsilon)d\varepsilon. \tag{3.1}$$

Where, $H(x_0, q_1) = \max_{q_2 \geq 0} h(x_e, q_1, q_2)$,

$$h(x_e, q_1, q_2) = -c_2 q_2 + \int_{x=0}^{\infty}\int_0^1 \{r\min[q_1\varepsilon + q_2, x] + v(q_1\varepsilon + q_2 - x)^+$$
$$-s(x - q_1\varepsilon - q_2)^+\}\phi(x|x_0)d\varepsilon dx.$$

Where q_1 and q_2 represent the respective quantities of purchases made by the retailer in the first and second production stages, which serve as the decision variables. The first stage is characterized by an uncertain supplier-service level, as the retailer's order of q_1 may be returned in a size of $q_1 * \varepsilon\,(with\, 0 \leq \varepsilon \leq 1)$. On the other hand, the second stage is equipped with a definite production mode. The unit production costs, including holding and delivery costs, are c_1 and c_2 for the first and second stages, respectively. Moreover, the retailer's prediction of the future market demand distributions in both stages are respectively denoted by $\phi_0(x_0)$ and $\phi(x|x_0)$, with $\phi(x|x_0)$ being closer to the true demand distribution. To emphasize the impact of prediction error on the optimal decision-making process, this paper assumes that the random demand distributions in both stages are known, and the retailer is aware of both $\phi_0(x_0)$ and $\phi(x|x_0)$ when making the decision of q_1. A benchmark scenario is also considered, where the retailer's prediction of future demand in both stages is the same and equal to $\phi_0(x_0)$. This setup serves as a means to more effectively illustrate the influence of prediction bias on the retailer's decisions.

The theoretical framework presented in Equation (3.1) draws inspiration from the work(Donohue, 2000) on supply contracts designed to coordinate the actions of manufacturers and distributors operating under two different production modes. In contrast to the previous study, our research incorporates the randomness of supplier service levels in the first stage of the model and emphasizes the analysis of the effects of forecast errors on the optimization of purchasing strategies and profit maximization.

The symbols used for other parameters in the model are as follows:

r: the retail price per unit of product.

s: the price that the retailer has to pay to the consumer for each unit of unmet demand.

v: the value of each unit of residual product at the end of the sales period.

As assumed in this paper, they satisfy the relationship: $r > c_2 > c_1 > v$.

3.2 Simplifying and solving the model

In order to facilitate a clearer comprehension of the results, this study simplifies the model from Section 3.1 and employs two prevalent assumptions regarding the demand distribution to derive an analytical representation for the optimal purchase quantity for the retailer in both production stages.

(1) *Demand follows the two-point distribution*

In this section, we consider a scenario where the demand for a product in the market follows a two-point distribution. During the first stage, the supplier, leveraging the available market information, predicts the demand distribution and obtains the parameters (l, h, and p) that characterize the two-point distribution, where the future market demand size will take the value of l with probability p and h with probability $1-p$, given that $0 < l < h$. At this point, the retailer places an order for q_1 units of the product to the supplier, and the product is delivered to the retailer at the conclusion of the production phase. In the second stage, the retailer gains additional information, through market research, etc., which enables them to determine the actual demand distribution. The retailer now understands that the size of the future market demand will be determined to be m with probability p, and n with probability $1-p$, given that $0 < m < n$.

Solving the model, the optimal purchase quantities for the first and second stages are obtained as follows:

$$q_1^* = \sqrt{\frac{q' - A\kappa^2}{B}},$$

$$q_2^* = (\kappa - q_1\varepsilon')^+.$$

Where,

$$A = \frac{r+s-c_2}{r+s-v}, B = \frac{c_1 - v}{r+s-v},$$

$$\kappa = \max\{m\mathbb{I}(p - A \geq 0), n\mathbb{I}(p - A < 0)\},$$

$$q' = \begin{cases} pl^2 + (1-p)h^2, \kappa \in (0, l] \\ p\kappa^2 + (1-p)h^2, \kappa \in (l, h] \\ \kappa^2, \kappa \in (h, \infty). \end{cases}$$

In the first stage, the retailer replenishes its inventory level from 0 to the optimal level, denoted by q_1^*. In the second stage, the retailer further replenishes its inventory level to κ, based on the order placed in the first stage. The value of ε' reflects the supplier-service level observed by the retailer after the end of the first stage. The value of κ is influenced by the retailer's forecast of the future demand distribution in the second stage, while the value of q' is a function of the relative positioning of the demand distributions predicted in both stages. As such, the optimal ordering level is contingent upon both these factors; when the parameter p is relatively small, the optimal ordering quantity tends toward n, whereas when p is relatively large, the optimal ordering quantity tends toward m."

(2) *Demand follows the uniform distribution*

In this section, we assume that the demand in the market follows a uniform distribution. The retailer in the first stage assumes that the future market demand will take on values in the interval $[l, h]$ with equal likelihood $(0 < l < h)$, whereas the retailer in the second stage is aware that the future market demand will be uniformly distributed over the interval $[m, n](0 < m < n)$. The assumption of a uniform distribution is considered to be more representative of real-world scenarios compared to a discrete two-point distribution.

Solving the model, the optimal purchase quantities for the first and second stages are obtained as follows:

$$q_1^* = \sqrt{\frac{q'' - A\tau^2}{B}},$$

$$q_2^* = (\tau - q_1\varepsilon')^+.$$

Where,

$$\tau = An + (1 - A)m,$$

$$q'' = \begin{cases} \frac{1}{3}\left(h^2 + l^2 + hl\right), \tau \in (0, l] \\ \frac{h^3 + \tau^2(2\tau - 3l)}{3(h - l)}, \tau \in (l, h] \\ \tau^2, \tau \in (h, \infty). \end{cases}.$$

The structure of the optimal solution mirrors that obtained through a two-point distribution. The replenishment level in the second stage, denoted by τ, takes on values in the interval between n and m. The value of q'' is also influenced by the relative positioning of the predicted demand distributions in both stages. When τ is relatively low, the optimal order quantity aligns more closely with the first-stage retailer's forecast of future demand, with q_1^* taking values in the interval between l and h. On the other hand, when τ is relatively high, the optimal order quantity is more consistent with the second-stage retailer's estimate of future demand, resulting in q_1^* being slightly greater than m.

As a conclusion, this study provides an inventory control strategy that maximizes the system's profit under two distinct demand distribution assumptions. The order-up-to policy is the chosen replenishment strategy, with the order-up-to level determined as κ when demand follows a two-point distribution and τ when demand follows a uniform distribution. Prior to the start of the sales season, the retailer will replenish inventory up to the order-up-to level based on the current inventory levels, refraining from placing further purchase orders with the supplier if the inventory exceeds this level. Due to the differences in the production methods in each stage and the varying accuracy of demand information available to the retailer, the policy is not fully established until the second stage.

Moreover, the optimal profit that the system can attain can be represented as:

$$\pi\left(q_1^*\right) = (r - v)\mu_1 - (c_1 - v)q_1^*.$$

Where, μ_1 represents the mean of the retailer's predicted demand distribution for the first stage. If the demand prediction follows a two-point distribution, then $\mu_1 = pl + (1-p)h$, and if the demand prediction follows a uniform distribution, $\mu_1 = \frac{1}{2}(l + h)$.

4. Sensitivity analysis

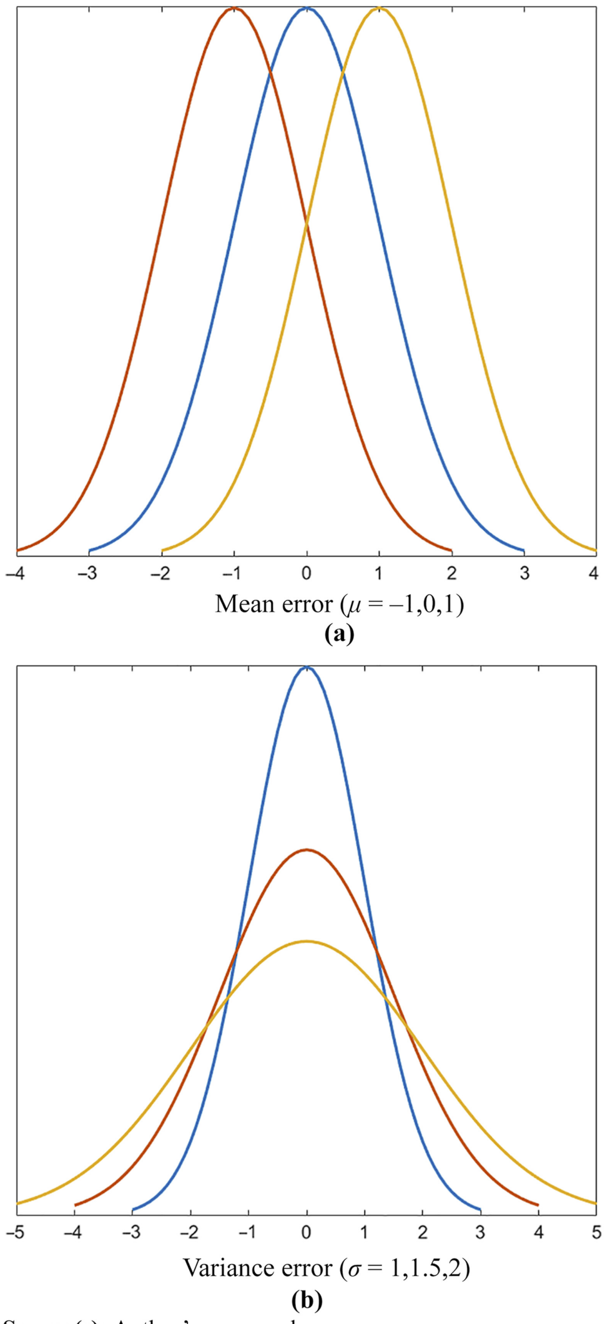
This study examines the sensitivity of relevant parameters in the optimal solution, with respect to different types of estimation and observation errors in the context of inventory control decision-making under uncertainty in supply. The focus of the analysis is on the mean error and variance error of demand forecasting, which are commonly employed measures in the literature to assess the quality of prediction methods from different perspectives.

As an illustration, taking the normal distribution as a reference, the distinction between mean error and variance error is depicted in Figure 2. The mean error of demand forecasting represents the difference between the mean of the predicted demand distribution and the actual demand distribution, assuming equal variances. This implies that the retailer is able to accurately predict the fluctuations in future market demand, but lacks a comprehensive understanding of the true demand range. For instance, when the mean error is substantial and the predicted demand distribution fails to overlap with the actual demand distribution, with the mean of the predicted demand exceeding that of the actual demand, the purchasing decisions of the retailer will be significantly impacted, leading to over purchasing and inventory buildup, thereby decreasing profits. On the other hand, the variance error of demand forecasting represents a higher variance of the predicted demand distribution than the actual demand distribution, with equal means. This implies that the retailer's estimate of the average value of future demand is accurate, but its assessment of the potential changes in demand is uncertain. When the variance error is excessive, it creates the illusion of a large range of future demand values, making it challenging for the retailer to make a sound judgment. A risk-averse retailer will opt for a smaller order, resulting in stockouts, while a risk-taking retailer will order excessively, leading to inventory buildup. Both the mean error and variance error of demand forecasting can induce the retailer to make overly optimistic or pessimistic predictions about future demand conditions, resulting in suboptimal decisions.

In this section, the notation μ_1, σ_1^2 is utilized to characterize the mean and variance of the demand distribution forecasted during the first stage, while μ_0, σ_0^2 symbolize the mean and variance in the subsequent stage. In instances where the error in prediction is expressed through variance, it holds that $\mu_1 = \mu_0$ and $\sigma_1 > \sigma_0$. Conversely, when the prediction error manifests in the mean, it can be noted that $\sigma_1 = \sigma_0$ and $\mu_1 \neq \mu_0$.

4.1 Variance error

In this section, we consider the difference between the variances of the demand distribution predicted in the two stages, represented by $\delta = \sigma_1^2 - \sigma_0^2$, with$\delta \geq 0$. Our assumption that $\delta \geq 0$ is motivated by the consideration that retailers may choose to estimate future demand within a broader range in order to mitigate potential risks, leading to a predicted variance that is greater than the actual variance. This assumption is deemed reasonable, as opposed to the alternative scenario where the retailer is overly confident in their prediction and the limited range of expected demand results in substantial losses.



Source(s): Author's own work

Figure 2.
Difference between mean error and variance error

Through our analysis, we have derived the optimal solution and first-order partial derivative of the optimal profit under different demand distributions.

For the case where the demand follows a two-point distribution, it is represented by

$$\frac{\partial q_1^*}{\partial \delta} = \frac{\sqrt{\frac{p}{1-p}}\left(\frac{1-p}{\sqrt{\delta+\sigma_0^2}} - I\frac{p-A}{\sqrt{\sigma_1^2-\delta}}\right)\mu_0 + \frac{p}{1-p}(1-IA)}{2\sqrt{\mathrm{B}\left[(1-p)\left(\mu_0+\sqrt{\frac{p}{1-p}(\delta+\sigma_0^2)}\right)^2 + (p-\mathrm{A})\left(\mu_0 + I\sqrt{\frac{p}{1-p}(\sigma_1^2-\delta)}\right)^2\right]}}. \tag{3.2}$$

In the case where the demand follows a uniform distribution, it is represented by

$$\frac{\partial q_1^*}{\partial \delta} = \frac{-h^3 - 2\tau^3 + 3\tau^2 Z + 3\sqrt{3}(\delta+\sigma_0^2)\left(\frac{h^2}{\sqrt{\delta+\sigma_0^2}} + \frac{2(2A-1)Z\tau}{\sqrt{-\delta+\sigma_1^2}}\right)}{4\sqrt{2}3^{\frac{3}{4}}B^{\frac{1}{2}}(\delta+\sigma_0^2)^{\frac{5}{4}}\sqrt{(h^3-3Z)(\tau+2\tau^3)}}. \tag{3.3}$$

Where,

$$I = \begin{cases} 1, p-A<0 \\ -1, p-A \geq 0 \end{cases}$$

$$Z = Al + (2A-1)h.$$

Additionally, the first-order partial derivative of the optimal profit with respect to the variance difference is given by

$$\frac{\partial \pi(q_1^*)}{\partial \delta} = -(c_1 - v)\frac{\partial q_1^*}{\partial \delta}.$$

Our analysis of the relationship between the symbols in equations (3.2) and (3.3) has revealed that the impact of prediction error, as reflected in the variance, on the inventory control decision is significant.

P1. Given that the mean of the demand distribution forecasted by the retailer in the first stage is equal to that in the second stage (i.e. $\mu_1 = \mu_0$), the greater the variance discrepancy between the forecasted and actual demand distributions, represented by δ ($\delta \geq 0$), the larger the optimal quantity of the first-stage procurement, q_1^*, the smaller the optimal quantity of the second-stage procurement, q_2^*, and the lower the optimal profit, $\pi(q_1^*)$.

The coincidence of means implies that the error in the prediction manifests itself in the form of variance error, with the variance of the demand prediction distribution in the first stage exceeding that of the actual demand distribution in the second stage. A larger variance discrepancy results in a more dispersed estimation of demand. In this scenario, the retailer perceives that the market demand has a certain probability of being relatively small (above zero) and a certain probability of being quite large. Due to the unreliability of the supplier in the first stage, only a fraction of the qualified products in the purchase order may arrive, prompting the retailer to place more orders in the first stage to accommodate potential high market demand. In the second stage, the retailer adjusts the order quantity by reducing q_2^* based on the actual demand distribution. Concurrently, an imprecise demand forecast exposes the retailer to the risks of stockouts and inventory accumulation, thus rendering a greater variance discrepancy δ, to result in lower optimal profit.

4.2 Mean error

In this section, we examine the scenario where the variances of the demand distributions predicted by the retailer in the two production stages are equivalent. The difference between the means is represented as $\omega = \mu_1 - \mu_0$. In the event that $\omega \geq 0$, it signifies that the retailer's prediction for future demand in the first stage is overestimated. Conversely, if $\omega < 0$, it implies that the retailer's prediction for future demand is underestimated. Our analysis encompasses the calculation of the first derivative of both the optimal solution and the optimal profit with respect to the aforementioned parameter under varying conditions.

For the case where the demand follows a two-point distribution, it is represented by

$$\frac{\partial q_1^*}{\partial \omega} = \begin{cases} \dfrac{\mu_0 + \omega + (1-2p)\sqrt{\dfrac{p}{1-p}}\sigma_0 + AX}{\sqrt{B\left[p\left(\mu_0 + \omega - \sqrt{\dfrac{p}{1-p}}\sigma_0\right)^2 + (1-p)\mathrm{T}^2 - AX^2\right]}}, \omega \in [u, +\infty) \\ \dfrac{(1-p)\left(\mu_0 + \omega + \sqrt{\dfrac{p}{1-p}}\sigma_0\right) - (p-A)X}{\sqrt{B\left[\mathrm{T}^2(1-p) + (p-A)X^2\right]}}, \omega \in [-v, u) \\ -\sqrt{\dfrac{1-A}{B}}, \omega \in (-\infty, -\nu). \end{cases} \quad (3.4)$$

Where, $X = \mu_1 - \omega + I\sqrt{\frac{p}{1-p}}\sigma_0, \mathrm{T} = \mu_0 + \omega + \sqrt{\frac{p}{1-p}}\sigma_0, u = \mathbb{I}(p < A)2\sqrt{\frac{p}{1-p}}\sigma_0, \nu = \mathbb{I}(p \geq A)2\sqrt{\frac{p}{1-p}}\sigma_0.$

For the case where the demand follows a uniform distribution, it is represented by

$$\frac{\partial q_1^*}{\partial \omega} = \begin{cases} \dfrac{\mu_0 + \omega + A\tau}{\sqrt{B[(\mu_0+\omega)^2 + \sigma_0^2 - A\tau^2]}}, \omega \in [u', +\infty) \\ \dfrac{3^{\frac{1}{4}}(-3\tau^2 + (\mu_0 + \omega + \sqrt{3}\sigma_0)^2 + 2\tau)}{2^{\frac{3}{2}}\sqrt{B\sigma_0(2\tau^3 + (\mu_0 + \omega + \sqrt{3}\sigma_0)^3 - 3\tau^2)}}, \omega \in [-v', u') \\ -\sqrt{\dfrac{1-A}{B}}, \omega \in (-\infty, -v') \end{cases} \quad (3.5)$$

Where, $O = \mu_0 + \omega + (2A-1)\sqrt{3}\sigma_0$, $u' = 2\sqrt{3}A\sigma_0$, $v' = 2\sqrt{3}(1-A)\sigma_0$.

And the first-order partial derivative of the optimal profit with respect to the mean difference is given by

$$\frac{\partial \pi(q_1^*)}{\partial \omega} = (r - v) - (c_1 - v)\frac{\partial q_1^*}{\partial \omega}.$$

The analysis of equations (3.4) and (3.5) leads to the following conclusions:

P2.1. when $\omega \geq 0$, as $|\omega|$ increases, the optimal value of q_1^* increases, the optimal value of q_2^* decreases and the resulting optimal profit decreases.

P2.2. When demand satisfies condition$\mathcal{M}$, as $|\omega|$ increases, q_1^* decreases and q_2^* decreases, the optimal profit decreases.

P2.3. When demand satisfies condition$\mathcal{N}$, as $|\omega|$ increases, the optimal profit decreases. However, $|\omega|$ has no effect on the values of optimal purchase quantity, q_1^* and q_2^*. Instead, they are only dependent on the parameters m, n, and p of the demand distribution in the second stage.

Condition $\mathcal{M}$: When the demand follows a two-point distribution, the condition is $p - A \geq 0$ and $-2\sqrt{\frac{p}{1-p}}\sigma_0 \leq \omega < 0$. When the demand follows a uniform distribution, the condition is $-2\sqrt{3}(1-A)\sigma_0 \leq \omega < 0$.

Condition $\mathcal{N}$: When the demand follows a two-point distribution, the condition is$p - A \geq 0$ and $\omega < -2\sqrt{\frac{p}{1-p}}\sigma_0$, or $p - A < 0$ and $\omega < 0$. When the demand follows a uniform distribution, the condition is $\omega < -2\sqrt{3}(1-A)\sigma_0$.

The occurrence of prediction errors as mean errors in the estimation of future demand by retailers implies a bias in the mean value, despite the accurate dispersion of the predicted distribution ($\sigma_1 = \sigma_0$). A positive mean difference $\omega > 0$ indicates that the predicted mean of the distribution is higher than the true mean, and the larger the mean difference, the higher the predicted mean. Consequently, the retailer in the first stage believes that the overall future market demand is high and will place more orders in the current stage with lower procurement costs. In contrast, the retailer in the second stage will adjust their approach based on the actual situation. Conversely, when the mean difference $\omega < 0$, the predicted mean is lower than the true mean, and the larger the mean difference, the lower the predicted mean. In this case, the retailer in the first stage believes that the future market is weak, with low demand, and will therefore order fewer products to avoid inventory buildup. The retailer in the second stage will replenish based on the new demand distribution. The conclusions for the two-point distribution and the uniform distribution are comparable, with the distinction that when $p - A \geq 0$, the retailer believes that demand is more likely to be l, whereas when $p - A < 0$, the retailer believes that demand is more likely to be h. Furthermore, much like variance error, an imprecise mean prediction, regardless of whether it overestimates or underestimates, will lead to a decline in optimal profit.

It is noteworthy that Proposition 2.3 is an intriguing result, given the prevailing belief that order quantity is dependent on bias. However, this study reveals that order quantity may also be independent of bias. In the case of condition$\mathcal{N}$, where the predicted mean is less than the true mean ($\omega < 0$) and the mean difference $|\omega|$ is substantial, the parameters l and h of the predicted distribution are significantly smaller than the parameters m and n of the true distribution. For retailers who only consider the first stage, choosing not to order may be a superior decision. However, given the higher demand estimates in the second stage and the lower procurement costs in the first stage, prestocking can save costs and better meet future demand. Thus, the optimal procurement quantity q_1^* in the first stage is not contingent upon the increase of the mean difference $|\omega|$, but rather on the retailer's demand estimation in the second stage.

Overall, both mean and variance prediction errors result in reduced total profits that the system can obtain, with greater errors leading to lower profits. This is consistent with expectations, as inaccurate demand forecasts diminish business performance. The impact of prediction errors on the optimal solution is situation-specific. In most cases (variance error or mean-high bias error), the optimal procurement quantity in the first stage is directly proportional to the prediction error, while the optimal procurement quantity in the second stage is inversely proportional to the prediction error. In a few cases (small mean bias errors), the opposite holds true. In special cases (low mean bias errors), the optimal procurement quantity is independent of the prediction error and solely dependent on the retailer's predicted demand distribution.

4.3 Sensitivity analysis of other parameters

In the model presented in this paper, changes in parameters other than the prediction error can also impact the optimal solution. This section categorizes the remaining parameters in the model into two groups and conducts sensitivity analysis under two demand distribution assumptions. One group of parameters pertains to costs, including selling price r, stockout cost s, salvage value v, purchasing cost c_1, and c_2. The other group of parameters relates to demand distribution, encompassing predicted distribution parameters l and h, as well as true distribution parameters m and n. The process of sensitivity analysis is delineated in Appendix 1, and the principal conclusions obtained are highlighted below.

- *P3.1.* As the selling price r, stockout cost s, and salvage value v increase, the optimal order quantity in the first stage q_1^* increases. Under the assumption of demand following a uniform distribution, the optimal order quantity in the second stage q_2^* first increases and then decreases, while under the assumption of demand following a two-point distribution, q_2^* first monotonically decreases, when r, s and v increase to $p \geq A$, the value of q_2^* jumps from $m - q_1\varepsilon'$ to $n - q_1\varepsilon'$ and continues to monotonically decrease ($q_2^* \neq 0$).
- *P3.2.* When the purchasing cost in the first stage c_1 increases, q_1^* decreases and q_2^* increases ($q_2^* \neq 0$). When the cost in the second stage c_2 increases, q_1^* increases and q_2^* decreases ($q_2^* \neq 0$).

Increasing the selling price, stockout cost and salvage value prompts the retailer to stockpile more goods as a buffer against potential stockouts. This is because the profit margin for each sold product increases, the penalty of unsatisfied demand rises and the holding cost of excess inventory decreases. Since the first-stage purchasing cost is relatively lower, the retailer would prefer to increase the order quantity in this stage. However, when the first-stage purchasing cost (c_1) increases, the retailer may switch to the second stage to minimize costs.

- *P4.* An increase in the parameters l and h of the retailer's forecast distribution of future demand in the first stage leads to an increase in the optimal order quantity q_1^* in the first stage and a decrease in the optimal order quantity q_2^* in the second stage ($q_2^* \neq 0$).

When the retailer is in the first stage, her decision-making process is prone to being influenced by current expectations. A higher l and h indicate that the retailer's estimation of future demand is more optimistic, based on past sales data and market research. Consequently, the retailer chooses to place more orders in the current stage, where the purchasing cost is relatively lower. However, when the retailer enters the second stage, the actual market demand remains unchanged, and she adjusts her response to reality by reducing the order quantity.

- *P5.* When $p - A \geq 0$ and $m \in (0, l]$, q_1^* decreases with an increase in m, while when $m \in (l, \infty)$, q_1^* increases with an increase in m. Conversely, when $p - A < 0$ and $n \in (h, \infty)$, q_1^* decreases with an increase in n, while when $n \in (0, h]$, q_1^* increases with an increase in n. For q_2^*, when $\sqrt{\frac{B}{1-A}} \leq \varepsilon' \leq 1$, the monotonicity of q_2^* is opposite to that of q_1^*; when $0 \leq \varepsilon' < \sqrt{\frac{B}{1-A}}$, q_2^* monotonically increases with an increase in m ($q_2^* \neq 0$).

Assuming a two-point distribution for demand, we find that the optimal order quantities, q_1^* and q_2^*, are influenced by the value of m when $p-A \geq 0$, but not by n. On the other hand, when $p-A < 0$, the changes in q_1^* and q_2^* depend on the value of n but not on m. The optimal order quantities are contingent upon the actual demand distribution predicted by the retailer in the second stage. Due to the unique properties of the two-point distribution, the actual demand is closer to m when $p-A \geq 0$ and closer to n when $p-A < 0$. When the actual demand is less than the demand forecasted by the retailer in the first stage, the retailer may realize that the first-stage forecast is too high and reduce the order quantity q_1^*. On the other hand, if the actual demand exceeds the first-stage forecast, an increase in m or n will lead to an increase in q_1^* to meet future market demands. When ε' is large, indicating that the orders placed in the first stage can be fulfilled at a high level, the change in the optimal order quantity q_2^* in the second stage is opposite to that of q_1^*, and the rate of change is slower. However, when ε' is small, implying that only a few of the orders placed in the first stage can be fulfilled, the retailer can only increase the order quantity in the second stage to replenish inventory as m or n increases. These results have significant implications for supply chain management and inventory control.

P6. When $\tau \in (0, Ah + (1-A)l]$, an increase in m/n results in a decrease in q_1^*; when $\tau \in (Ah + (1-A)l, \infty)$, an increase in m/n leads to an increase in q_1^*. For q_2^*, when $\sqrt{\frac{B}{1-A}} \leq \varepsilon' \leq 1$, the monotonicity of q_2^* is not completely opposite to that of q_1^*. Specifically, when $\tau \in (0, \tau']$, q_2^* monotonically increases with an increase in m/n;when $\tau \in (\tau', \infty)$, q_2^* monotonically decreases, where τ' is the value that satisfies equation $1 - \varepsilon'.\frac{\partial q_1^*}{\partial \tau} = 0$ andwhen $0 \leq \varepsilon' < \sqrt{\frac{B}{1-A}}$, q_2^* monotonically increases with an increase in m and $n(q_2^* \neq 0)$.

The result under the assumption of demand following a uniform distribution is similar to that under the assumption of a two-point distribution, the impact of the parameters m and n on the optimal order quantities q_1^* and q_2^* in the second stage depends on the value of τ, where τ represents a value between m and n. Notably, the impact of m and n on the optimal order quantity q_2^* in the second stage does not completely oppose that of q_1^* , and has a lag effect. Specifically, q_2^* only begins to decrease after q_1^* has increased to a certain extent, which is more representative of the general scenario and results in an increase in the total order quantity of the retailer with an increase in the parameters m and n of the true demand distribution.

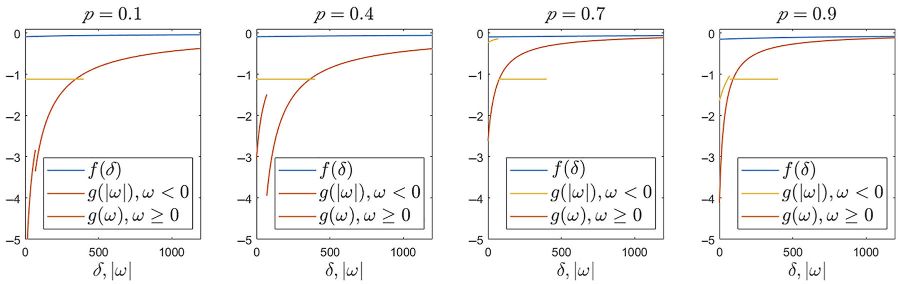
In this section, a sensitivity analysis is conducted on the parameters relevant to the optimal solution to address research question (1), regarding the impact of demand forecasting errors during the off-season on inventory decisions and the system's optimal profit in a two-stage procurement model. The findings reveal that both types of forecasting errors can result in a decline in the system's optimal profit and corresponding changes in the optimal order quantities. Additionally, the sensitivity analysis on other parameters contributes to the understanding of how cost and demand distribution parameters affect the optimal solution. Moving forward, the next section combines numerical experiments to provide a more in-depth analysis of the differences in the effect of the two types of forecasting errors on the optimal solution.

5. Numerical analysis of two types of prediction errors

Based on the discussion in Section 4, this paper has shown that both types of forecasting errors can cause a decline in the total profit of the system consisting of the retailer and the supplier under different scenarios, with corresponding changes in the optimal order quantities. However, in practice, businesses often face challenges in accurately matching the true demand distribution during demand forecasting. In such cases, the choice of a forecasting method with better reliability (lower mean error) or validity (lower variance error) becomes crucial. This section aims to investigate the differences in the rate of decline in the optimal profit under different mean and variance errors.

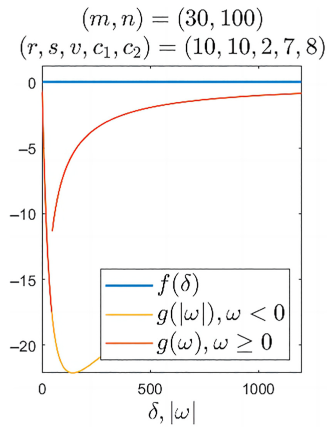
To achieve this, we define $f(\delta) = \frac{\partial \pi(q_1^*)}{\partial \delta}$ and $g(\omega) = \frac{\partial \pi(q_1^*)}{\partial |\omega|}$, where $\pi(q_1^*)$ represents the profit of the retailer. We fix the parameter sets (r, s, v, c_1, c_2) and (m, n) separately, and set various parameter combinations to observe how $f(\delta)$ and $g(\omega)$ change with forecast error under two different demand distributions. Specifically, we conduct experiments using multiple parameter combinations, following the parameter settings of Ban (2020): with (r, s, v, c_1, c_2) fixed, we vary n from 60 to 180; while with (m, n) fixed, we vary v from 0 to 10. The other parameters vary simultaneously under condition$r > c_2 > c_1 > v$.

Since the function graphs exhibit the same changing trend under different parameter combinations, we only present the graphs of one parameter combination($(r, s, v, c_1, c_2) = (10, 5, 1, 5, 6), (m, n) = (30, 100)$) to illustrate our key conclusion. The graphs obtained under other parameter combinations are included in Appendix 2 as supplementary material to demonstrate the robustness of the conclusion. Figure 3 shows the change of $f(\delta)$ and $g(\omega)$ with forecast error under four different levels of probability p, when the demand follows a two-point distribution. Figure 4 illustrates the change of $f(\delta)$ and $g(\omega)$ with forecast errors, when the demand follows a uniform distribution.



Source(s): Author's own work

Figure 3. Variation of $f(\delta)$ and $g(\omega)$ with prediction error under a two-point distribution, $p = 0.1, 0.4, 0.7, 0.9$ levels



Source(s): Author's own work

Figure 4. Variation of $f(\delta)$ and $g(\omega)$ with prediction error under uniform distribution

In the experiments, we observe that when $\omega \leq 0$, $\frac{\partial \pi(q_1^*)}{\partial \omega} \geq 0$. To aid comparison, we adopt an absolute value transformation of the horizontal axis, showing the corresponding graph in the fourth quadrant (indicated by the red curve in Figures 3 and 4).

From the figures, we can observe that we can observe that regardless of the demand distribution and parameter settings, the function values of $f(\delta)$ and $g(\omega)$ are always negative. Furthermore, the blue curve of $f(\delta)$ is always higher than the red and yellow curves of $g(\omega)$, which is a noteworthy conclusion.

This finding implies that both types of forecasting errors can lead to a decrease in the optimal profit that the system can achieve. ***However, mean errors have a greater impact on optimal profit than variance errors.*** Additionally, when $\delta\ and\ \omega \rightarrow \infty$, $f(\delta)\ and\ g(\omega) \rightarrow 0$, indicating that as the forecasting error approaches infinity, it no longer affects the optimal profit. This conclusion holds across multiple parameter combinations. Thus, managers should prioritize improving the demand forecasting quality, particularly the accuracy of mean forecasts, ahead of the sales season.

Drawing on the analysis presented above, this study provides a response to research question (2), which concerns the impact of different types of forecasting errors on the optimal solution. The findings indicate that, in comparison to variance errors, a greater absolute difference between predicted and true means results in diminished system performance. This outcome offers recommendations and support for practical enterprise management. For instance, when selecting a demand forecasting model, businesses can opt to trade reliability (variance) for validity (bias) and choose an unbiased estimator. Additionally, in keeping with the law of large numbers, enterprises can enhance the precision of mean forecasts by accumulating sufficient historical data to secure higher profits.

6. Case study of optimal inventory strategy

To validate the practicality of the inventory control strategy proposed in this study, we collected sales and inventory data for four representative products from our partner companies between June 2020 and June 2021, and estimated the model parameters based on the collected data. Using the "Medium-length Down Jacket without Collar" product as an example, we calculated the daily sales volume during the peak-season and the daily sales volume and mean of previous sales, which were distributed between 300 and 330 units, as shown in Figure 5. Since winter is typically considered to start in November and consumers tend to purchase down jackets during this period, suppliers receive numerous production orders at this time. Therefore, we assume that the peak-season is from November to January, while the remaining months constitute the off-season. The sales curve does not exhibit a unimodal distribution and cannot be accurately modeled using a normal distribution. Therefore, we conducted a case analysis using the analytical expression of the optimal procurement strategy obtained in Section 3.2.2. Although this assumption may not fully correspond to the actual data distribution of our partner companies, it is easy to understand and provides a simple expression that is conducive to practical application by business operators.

Secondly, this study employed a weighted average method with sales volume and procurement volume as weights to estimate the selling price r, production costs c_1 and c_2. The supplier-service level ε' during the off-season and the demand distribution parameters m and n during the peak-season was estimated using a maximum likelihood method. The estimation method for the parameters l and h, which predict the demand during the off-season for the peak-season, was consistent with the current forecasting strategy implemented by the enterprise. It was based on a sliding weighted average of historical demand data from the previous day and the demand from the previous 15 days, with weights of 2 for historical demand and the demand from the previous 1–5 days, 1.5 for demand from the previous

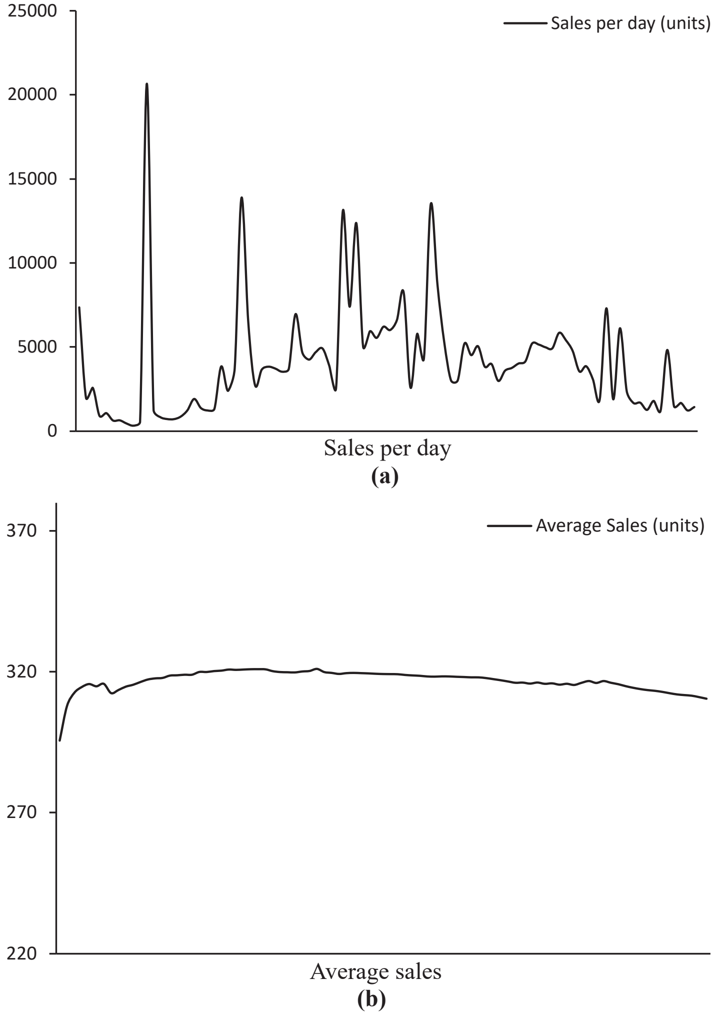


Figure 5.
Peak-season sales of the product

Source(s): Author's own work

6–10 days and 1 for demand from the previous 11–15 days. The out-of-stock cost s and the residual value v were determined through interviews with managers. The model parameters estimated based on real data from our partner companies are presented in the Table 1 below.

The results of parameter estimation indicate that $l < m < n < h$, which is a consequence of excluding outliers with excessively large or small values caused by brushing orders or returns during the computation of actual demand. When predicting demand, the lower sales volume during the off-season resulted in a smaller predicted value for l, while the higher sales volume during the previous peak-season led to a larger predicted value for h. The estimates for l and h represent the upper and lower limits of demand for the sales season as estimated by the managers.

Drawing on a real-world context, this study computed the optimal profit attainable by a company based on the expression for the optimal procurement strategy obtained in Section

	Product A	Product B	Product C	Product D
r	345	522	732	143
s	60	75	90	15
v	10	20	30	5
c_1	68	83	165	22
c_2	75	132	179	27
m	357,019	178,776	8,966	510,388
n	546,087	406,917	117,660	649,824
l	29,070	37,669	50,094	456,597
h	580,136	530,878	153,049	810,774
ε'	0.65	0.71	0.69	0.82

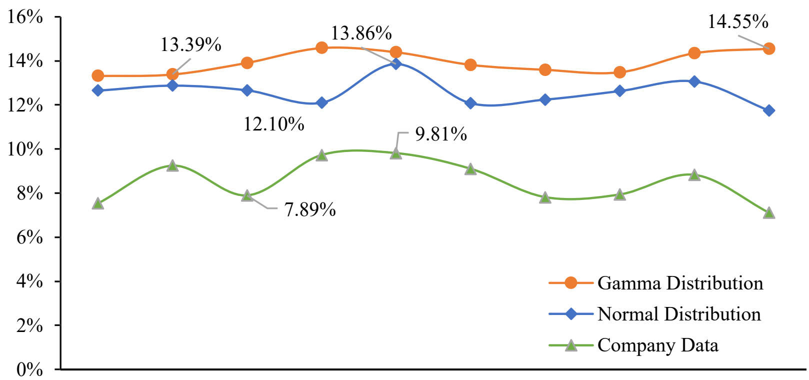
Source(s): Author's own work

Table 1.
Parameter estimation

3.2.2, and contrasted it with the company's current replenishment policy. The findings revealed an average profit increase of 7.02% across the four products. Furthermore, given that the out-of-stock cost s and residual value v were based on managerial expertise, this study depicted the average profit enhancement across different parameter combinations by varying s and v. The range for s was set from $0.9 * s$ to $1.1 * s$, with v adjusted similarly, as illustrated by the green curve (company data) in Figure 6.

In order to test the effectiveness of the proposed optimal procurement strategy under more general conditions, this study not only utilized actual demand data from a cooperative company, but also solved the problem using random demand data from gamma and normal distributions. Average profit improvement was calculated while varying values of s and v, as shown by the yellow and blue curves in Figure 6, respectively. Results indicate that the optimal strategy performed better under more general demand distributions, with an average profit improvement of 13%.

The analysis in this section addressed research question (3), which was focused on the performance of the inventory strategy using the actual demand data from the cooperative company. Despite being developed under certain assumptions, the excellent experimental results suggest that the proposed inventory control strategy is feasible in practice and can provide valuable guidance for retailers in making inventory decisions, thus improving their business operations.



Source(s): Author's own work

Figure 6.
Average profit improvement with multiple parameter combinations

These findings have important practical implications, highlighting the potential for the proposed model to enhance inventory control and decision-making in retail operations. Moreover, this research can be generalized to broader contexts beyond the specific case of the cooperative company.

7. Conclusion

When making business decisions, enterprise managers often face uncertainty in both demand and supply. While some demand forecasting methods have improved accuracy and bias performance thanks to technologies such as machine learning (Spiliotis *et al.*, 2022), demand forecast errors persist. Sanders and Graman (2016) have shown that amplifying forecasting bias and standard deviation can have a clear adverse impact on supply chain costs, with the effect of the forecasting standard deviation being greater. To address this, this paper develops an inventory model with stochastic production rates of suppliers and demand knowledge updates. The replenishment process undergoes two different production modes, with the aim of maximizing the total profit of the system. This paper studies the optimal inventory decision of retailers under supply uncertainty and the influence of two different types of forecasting errors on inventory decisions and optimal profits. Based on actual data from a cooperative company, the effectiveness of the proposed inventory strategy in practice is discussed. The main conclusions of this paper are as follows:

(1) The retailer's optimal replenishment policy follows an order-up-to policy, with the order-up-to level determined by the forecast distribution of future demand. Due to changes in the two-stage production mode and the retailer's market demand knowledge, the policy is only determined in the second stage.

(2) In many cases, as the forecasting error increases, the optimal order quantity q_1^* in the first stage increases, while the optimal order quantity q_2^* in the second stage decreases, resulting in a smaller optimal profit for the system. If the forecasting error is manifested in the mean (i.e. forecasting bias) and the retailer's forecast of future demand is lower than the actual demand, then in some cases, as the mean error increases, the optimal order quantity q_1^* in the first stage decreases, while the optimal order quantity q_2^* in the second stage increases, resulting in a smaller optimal profit. In other cases, the optimal order quantity is not related to the mean error, but only the optimal profit decreases as the mean error increases.

(3) Based on numerical simulations under various parameter combinations and demand distribution assumptions, it is found that the effect of mean error (i.e. forecasting bias) on optimal profit is generally greater than that of variance error (i.e. forecasting variance). This suggests that forecasting methods with larger bias and higher accuracy lead to faster optimal profit reduction under the same circumstances than those with smaller bias and lower accuracy.

Thus, theoretically, our study established a two-stage inventory optimization model that simultaneously considers random yield and demand forecast quality, and provides explicit expressions for optimal procurement strategies under two specific demand distributions that takes into account these features. Furthermore, we focused on how forecast error affects the optimal inventory strategy and obtained interesting properties of the optimal solution that are mainly reflected in conclusion (2). In particular, the property that the optimal procurement quantity no longer changes with increasing forecast error under certain conditions is noteworthy and has not been previously noted by scholars. Therefore, our study fills a gap in the literature.

From a practical standpoint, our findings can assist retailers in making inventory control decisions for products with long production cycles and significant seasonality while considering the impact of forecast errors. When forecasting demand, decision-makers can prioritize accuracy over confidence to improve profits, optimize inventory decisions and meet consumer demands. The inventory strategies outlined in this paper have been validated using actual demand data from collaborating companies as well as demand data generated by typical random distributions, yielding better profits than the current approach.

Future research directions: With the advancement of the big data era and the continuous progress of information technology, there are various directions for future extensions of inventory optimization research under demand uncertainty. These include exploring more general demand distribution assumptions, introducing other supply uncertainties (such as random lead times) and investigating the interactions among different forecasting errors. By continuously enriching research in inventory management, enterprises can gain greater profits, optimize inventory decisions, reduce inventory costs and accurately predict market demand.

References

Ban, G.Y. (2020), "Confidence intervals for data-driven inventory policies with demand censoring", *Operations Research*, Vol. 68, pp. 309-326.

Barrow, D.K. and Kourentzes, N. (2016), "Distributions of forecasting errors of forecast combinations: implications for inventory management", *International Journal of Production Economics*, Vol. 177, pp. 24-33.

Berling, P. and Sonntag, D.R. (2022), "Inventory control in production-inventory systems with random yield and rework: the unit-tracking approach", *Production and Operations Management*, Vol. 31, pp. 2628-2645.

Bruggen, A., Grabner, I. and Sedatole, K.L. (2021), "The folly of forecasting: the effects of a disaggregated demand forecasting system on forecast error, forecast positive bias and inventory levels", *Accounting Review*, Vol. 96, pp. 127-152.

Bunning, F., Heer, P., Smith, R.S. and Lygeros, J. (2020), "Improved day ahead heating demand forecasting by online correction methods", *Energy and Buildings*, Vol. 211, 109821.

Cruz, C.O. and Sarmento, J.M. (2020), "Traffic forecast inaccuracy in transportation: a literature review of roads and railways projects", *Transportation*, Vol. 47, pp. 1571-1606.

Dada, M., Petruzzi, N.C. and Schwarz, L.B. (2007), "A newsvendor's procurement problem when suppliers are unreliable", *M&Som-Manufacturing & Service Operations Management*, Vol. 9, pp. 9-32.

Deng, S.M. and Zheng, Z. (2020), "Optimal production decision for a risk-averse manufacturer faced with random yield and stochastic demand", *International Transactions in Operational Research*, Vol. 27, pp. 1622-1637.

Donohue, K.L. (2000), "Efficient supply contracts for fashion goods with forecast updating and two production modes", *Management Science*, Vol. 46, pp. 1397-1411.

Fiig, T., Weatherford, L.R. and Wittman, M.D. (2019), "Can demand forecast accuracy be linked to airline revenue?", *Journal of Revenue and Pricing Management*, Vol. 18, pp. 291-305.

Hu, B.Y. and Feng, Y. (2017), "Optimization and coordination of supply chain with revenue sharing contracts and service requirement under supply and demand uncertainty", *International Journal of Production Economics*, Vol. 183, pp. 185-193.

Ivanov, D., Dolgui, A., Sokolov, B. and Ivanova, M. (2017), "Literature review on disruption recovery in the supply chain", *International Journal of Production Research*, Vol. 55, pp. 6158-6174.

Kaplan, R.S. (1970), "A dynamic inventory model with stochastic lead times", *Management Science*, Vol. 16, pp. 491-507.

Lakshmanan, B., Raja, P. and Kalathiappan, V. (2019), "Sales demand forecasting using LSTM network", *Advances in Intelligent Systems and Computing*, Vol. 1056, pp. 125-132.

Liu, W., Song, S.J., Qiao, Y., Zhao, H. and Wang, H. (2020), "The loss-averse newsvendor problem with random yield and reference dependence", *Mathematics*, Vol. 8, 1231.

Paul, S.K., Sarker, R. and Essam, D. (2016), "Managing risk and disruption in production-inventory and supply chain systems: a review", *Journal of Industrial and Management Optimization*, Vol. 12, pp. 1009-1029.

Petropoulos, F., Wang, X. and Disney, S.M. (2019), "The inventory performance of forecasting methods: evidence from the M3 competition data", *International Journal of Forecasting*, Vol. 35, pp. 251-265.

Prak, D., Teunter, R., Babaic, M.Z., Boylan, J.E. and Syntetos, A. (2021), "Robust compound Poisson parameter estimation for inventory control", *Omega-international Journal of Management Science*, Vol. 104, 102481.

Rosling, K. (2002), "Inventory cost rate functions with nonlinear shortage costs", *Operations Research*, Vol. 50, pp. 1007-1017.

Salem, R.W. and Haouari, M. (2017), "A simulation-optimisation approach for supply chain network design under supply and demand uncertainties", *International Journal of Production Research*, Vol. 55, pp. 1845-1861.

Sanders, N.R. and Graman, G.A. (2016), "Impact of bias magnification on supply chain costs: the mitigating role of forecast sharing", *Decision Sciences*, Vol. 47, pp. 881-906.

Song, J.-S. (1994), "The effect of leadtime uncertainty in a simple stochastic inventory model", *Management Science*, Vol. 40, pp. 603-613.

Spiliotis, E., Makridakis, S., Semenoglou, A.A. and Assimakopoulos, V. (2022), "Comparison of statistical and machine learning methods for daily SKU demand forecasting", *Operational Research*, Vol. 22, pp. 3037-3061.

Svoboda, J., Minner, S. and Yao, M. (2021), "Typology and literature review on multiple supplier inventory control models", *European Journal of Operational Research*, Vol. 293, pp. 1-23.

Tomlin, B. (2006), "On the value of mitigation and contingency strategies for managing supply chain disruption risks", *Management Science*, Vol. 52, pp. 639-657.

Wan, X. and Sanders, N.R. (2017), "The negative impact of product variety: forecast bias, inventory levels, and the role of vertical integration", *International Journal of Production Economics*, Vol. 186, pp. 123-131.

Wang, Y. and Gerchak, Y. (1996), "Periodic review production models with variable capacity, random yield and uncertain demand", *Management Science*, Vol. 42, pp. 130-137.

Wu, J., Zhang, D., Yang, Y., Wang, G. and Su, L. (2022), "Multi-stage multi-product production and inventory planning for cold rolling under random yield", *Mathematics*, Vol. 10 No. 4, 597.

Xu, X.L., Cai, X.Q. and Zhang, L.M. (2020), "Optimal purchasing policy for fresh products from multiple supply sources with considerations of random delivery times, risk and information", *Decision Sciences*, Vol. 51, pp. 1377-1410.

Yuan, X.G., Zhang, X.Q. and Zhang, D.L. (2020), "Analysis of the impact of different forecasting techniques on the inventory bullwhip effect in two parallel supply chains with a competition effect", *Journal of Engineering*, Vol. 2020, 2987218.

Corresponding author
Yugang Yu can be contacted at: ygyu@ustc.edu.cn

Appendix 1

The detailed formulas for Proposition 1 and Proposition 2 have been presented in the main text and are readily verifiable. Hence, this section is dedicated to the analysis and proof of Propositions 3 through 6.

(1) The sensitivity analysis of r

It is established as follows. Firstly, we have:

$$\frac{\partial q_1^*}{\partial r} = \frac{q - \mathrm{X}^2}{2(r+s-v)\sqrt{B(q - A\mathrm{X}^2)}} \geq 0,$$

$$\frac{\partial q_2^*}{\partial r} = \begin{cases} \frac{(n-m)(1-A)}{r+s-v} I - \varepsilon' . \frac{\partial q_1^*}{\partial r}, q_2^* \neq 0 \\ 0, q_2^* = 0 \end{cases}.$$

Additionally, when demand follows a two-point distribution $\mathrm{X} = \kappa, q = q', I = 0$; and when demand follows a uniform distribution, $\mathrm{X} = \kappa, q = q'', I = 1$.

We observe that $q'' - \tau^2 \in \begin{cases} \left(\frac{1}{3}(h^2 - 2l^2 + hl), \frac{1}{3}(h^2 + l^2 + hl)\right), \tau \in (0, l] \\ \left(0, \frac{1}{3}(h^2 - 2l^2 + hl)\right), \tau \in (l, h] \\ 0, \tau \in (h, \infty) \end{cases}$, since $l = \mu - \sigma > 0$, it follows that $\frac{1}{3}(h^2 - 2l^2 + hl) = 2\sigma(3\mu - \sigma) > 0$, and hence $q'' - \tau^2 \geq 0$. Similarly, we have $q' - \kappa^2 \geq 0$, where $q' - \kappa^2 \in \begin{cases} (pl^2 + (1-p)h^2, (1-p)(h^2 - l^2)), \kappa \in (0, l] \\ (0, (1-p)(h^2 - l^2)), \kappa \in (l, h] \\ 0, \kappa \in (h, \infty) \end{cases}$.

(2) The sensitivity analysis of s

It is analogous to that of r, and the partial derivative formula and proof method are the same as those described in 1. Hence, we avoid repeating them here.

(3) The sensitivity analysis of c_1

Taking the partial derivatives of q_1^* and q_2^* with respect to c_1, which result in the following expressions:

$$\frac{\partial q_1^*}{\partial c_1} = \frac{-\sqrt{q - A\mathrm{X}^2}}{2(r+s-v)\sqrt{B^3}} \leq 0.$$

$$\frac{\partial q_2^*}{\partial c_1} = \begin{cases} -\varepsilon' . \frac{\partial q_1^*}{\partial c_1}, q_2^* \neq 0 \\ 0, q_2^* = 0 \end{cases}.$$

(4) The sensitivity analysis of c_2

Taking the partial derivatives of q_1^* and q_2^* with respect to c_2, which yield:

$$\frac{\partial q_1^*}{\partial c_2} = \frac{\mathrm{X}^2}{2(r+s-v)\sqrt{B(q - A\mathrm{X}^2)}} \geq 0.$$

$$\frac{\partial q_2^*}{\partial c_2} = \begin{cases} \frac{m-n}{r+s-v} - \varepsilon' . \frac{\partial q_1^*}{\partial c_2}, q_2^* \neq 0 \\ 0, q_2^* = 0 \end{cases}.$$

(5) The sensitivity analysis of v

Firstly, we have:

$$\frac{\partial q_1^*}{\partial v} = \frac{(1-B)q - A\text{X}^2}{2(c_1 - v)\sqrt{B(q - A\text{X}^2)}} > 0.$$

$$\frac{\partial q_2^*}{\partial v} = \begin{cases} \frac{A(n-m)}{r+s-v} - \varepsilon' . \frac{\partial q_1^*}{\partial v}, q_2^* \neq 0 \\ 0, q_2^* = 0 \end{cases}.$$

Since $\frac{A}{1-B} = \frac{r+s-c_2}{r+s-c_1} < 1$, and according to the proof in 1, we have $q' - A\kappa^2 \geq 0$. Thus, we can conclude that $q'' - \frac{A}{1-B}\tau^2 > 0$, which completes the proof of Proposition 3.

(6) The sensitivity analysis of l

When demand follows a two-point distribution, we have $\frac{\partial q_1^*}{\partial l} = \begin{cases} \frac{pl}{\sqrt{B(q' - A\kappa^2)}}, \kappa \in (0, l) \\ 0, \kappa \in (l, \infty) \end{cases}$, which implies that $\frac{\partial q_1^*}{\partial l} \geq 0$. When demand follows a uniform distribution, we have $\frac{\partial q_1^*}{\partial l} = \frac{1}{2\sqrt{B(q'' - A\tau^2)}} . \frac{\partial q''}{\partial l}$, where $\frac{\partial q''}{\partial l} = \left\{ \frac{1}{3} \right. (2l + h), \tau \in (0, l) \frac{(\text{h} - \tau)^2 (h + 2\tau)}{3(h - l)^2}, \tau \in (l, h) 0, \tau \in (h, \infty)$. Since $\frac{\partial q''}{\partial l} \geq 0$, it follows that $\frac{\partial q_1^*}{\partial l} \geq 0$. In addition, we have $\frac{\partial q_2^*}{\partial l} = \begin{cases} -\varepsilon' . \frac{\partial q_1^*}{\partial l}, q_2^* \neq 0 \\ 0, q_2^* = 0 \end{cases}$.

(7) The sensitivity analysis of h

When demand follows a two-point distribution, we have $\frac{\partial q_1^*}{\partial h} = \begin{cases} \frac{(1-p)h}{\sqrt{B(q' - A\kappa^2)}}, \kappa \in (0, h] \\ 0, \kappa \in (h, \infty) \end{cases}$, which implies that $\frac{\partial q_1^*}{\partial h} \geq 0$. When demand follows a uniform distribution, we have $\frac{\partial q_1^*}{\partial h} = \frac{1}{2\sqrt{B(q'' - A\tau^2)}}$.

$\frac{\partial q''}{\partial h}$, where $\frac{\partial q''}{\partial h} = \begin{cases} \frac{1}{3}(2h + l), \tau \in (0, l] \\ \frac{h^2(2h - 3l) + \tau^2(3l - 2\tau)}{3(h - l)^2}, \tau \in (l, h] \\ 0, \tau \in (h, \infty) \end{cases}$. Since $\frac{\partial q''}{\partial h} \geq 0$, it follows that $\frac{\partial q_1^*}{\partial h} \geq 0$. In addition,

we have $\frac{\partial q_2^*}{\partial h} = \begin{cases} -\varepsilon' . \frac{\partial q_1^*}{\partial h}, q_2^* \neq 0 \\ 0, q_2^* = 0 \end{cases}$. Thus, we have completed the proof of Proposition 4.

(8) The sensitivity analysis of m

Taking the partial derivatives of q_1^* and q_2^* with respect to m, we can conclude the following conclusions:

When the demand follows a two-point distribution, $\frac{\partial q_1^*}{\partial m} = \mathbb{I}(p - A \geq 0)\frac{\partial q_1^*}{\partial \kappa}$, where $\frac{\partial q_1^*}{\partial \kappa} =$

$$\begin{cases} \dfrac{-A\kappa}{\sqrt{B(q' - A\kappa^2)}} < 0, \kappa \in (0, l] \\ \dfrac{(p - A)\kappa}{\sqrt{B(q' - A\kappa^2)}}, \kappa \in (l, h] \\ \dfrac{(1 - A)\kappa}{\sqrt{B(q' - A\kappa^2)}} > 0, \kappa \in (h, \infty) \end{cases}, \text{ and } \frac{\partial q_2^*}{\partial m} = \begin{cases} 1 - \varepsilon' . \dfrac{\partial q_1^*}{\partial m}, q_2^* \neq 0 \\ 0, q_2^* = 0 \end{cases}.$$

For the sign of $\frac{\partial q_1^*}{\partial m}$, it can be easily proved that when $p \in (0, A]$, we have $\frac{\partial q_1^*}{\partial m} \leq 0$; and when $p \in (A, \infty)$, we have $\frac{\partial q_1^*}{\partial m} > 0$. For the sign of $\frac{\partial q_2^*}{\partial m}$, when $\kappa \in (h, \infty)$, we have $\frac{\partial q_1^*}{\partial \kappa} = \sqrt{\frac{1-A}{B}}$. Then let $1 - \varepsilon' . \frac{\partial q_1^*}{\partial m}\kappa < 0$, it follows that $\varepsilon' > \sqrt{\frac{B}{1-A}}$, and when $\kappa \in (l, h]$ and $p - A \geq 0$, let $1 - \varepsilon' . \frac{\partial q_1^*}{\partial m} < 0$, it follows that $m > h . \sqrt{\frac{(1-p)B}{(p-A)[(p-A)\varepsilon'^2 - B]}}$. However, this leads to a contradiction because $m \leq h$ and $\sqrt{\frac{(1-p)B}{(p-A)[(p-A)\varepsilon'^2 - B]}}$, so the assumption is false.

When demand follows a uniform distribution, it is shown that $\frac{\partial q_1^*}{\partial m} = (1 - A)\frac{\partial q_1^*}{\partial \tau}$, where

$$\frac{\partial q_1^*}{\partial \tau} = \begin{cases} \dfrac{-A\tau}{\sqrt{B(q'' - A\tau^2)}}, \tau \in (0, l) \\ \dfrac{\tau(\frac{\tau - l}{h - l} - A)}{\sqrt{B(q'' - A\tau^2)}}, \tau \in (l, h) \\ \dfrac{(1 - A)\tau}{\sqrt{B(q'' - A\tau^2)}}, \tau \in (h, \infty) \end{cases}, \text{ and } \frac{\partial q_2^*}{\partial m} = \begin{cases} 1 - A - \varepsilon' . \dfrac{\partial q_1^*}{\partial m}, q_2^* \neq 0 \\ 0, q_2^* = 0 \end{cases}.$$

For the sign of $\frac{\partial q_1^*}{\partial m}$, it can be easily proved that $\frac{\partial q_1^*}{\partial m} \leq 0$ for $\tau \in (0, Ah + (1 - A)l]$, and $\frac{\partial q_1^*}{\partial m} > 0$ for $\tau \in (Ah + (1 - A)l, \infty)$. For the sign of $\frac{\partial q_2^*}{\partial m}$, we set $1 - A - \varepsilon' . \frac{\partial q_1^*}{\partial m} = 0$ and obtain $\tau^2(\frac{\tau - l}{h - l} - A)^2 \varepsilon'^2 = B(q'' - A\tau^2)$,where $f(\tau) = \tau^2(\frac{\tau - l}{h - l} - A)^2 \varepsilon'^2$ and $g(\tau) = B(q'' - A\tau^2)$. We have $f'(\tau) > 0$ and $g'(\tau) > 0$, and for $\tau \in (Ah + (1 - A)l, h]$, we have $0 < f(\tau) < (1 - A)^2 \varepsilon'^2 h^2$v, and because $\frac{h^3 - [Ah + (1 - A)l]^3}{3(h - l)} B < g(\tau) < (1 - A)Bh^3$, the intersection of $f(\tau)$ and $g(\tau)$ exists. Hence, Proposition 5 and 6 are proved.

(9) The sensitivity analysis of n

The analysis and method are the same as in 8, and therefore will not be repeated here.

Appendix 2

To verify the robustness of the conclusion in Section 5, we conducted numerical experiments with multiple parameter settings to compare the speed difference of the optimal profit changing with mean error and variance error. Where $f(\delta) = \frac{\partial \pi(q_1^*)}{\partial \delta}$ and $g(\omega) = \frac{\partial \pi(q_1^*)}{\partial |\omega|}$. Figures A1 and A2 represent the experimental results for demand following a two-point distribution and a uniform distribution, respectively.

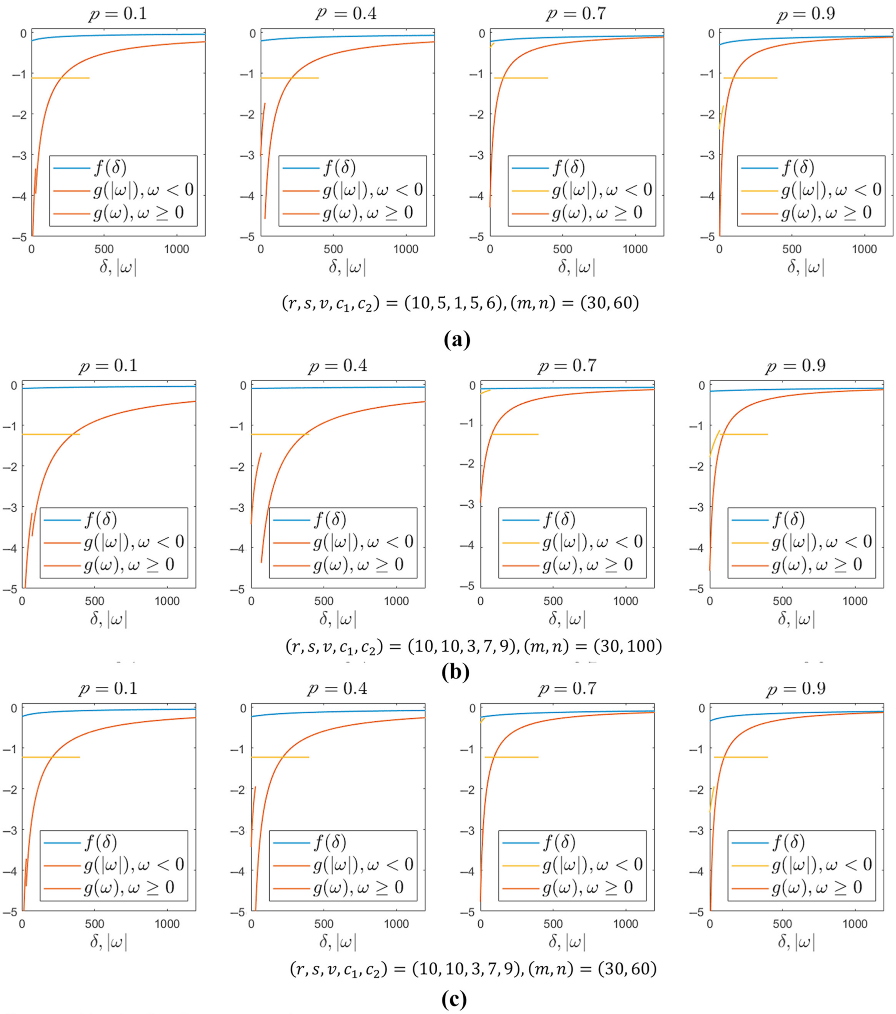
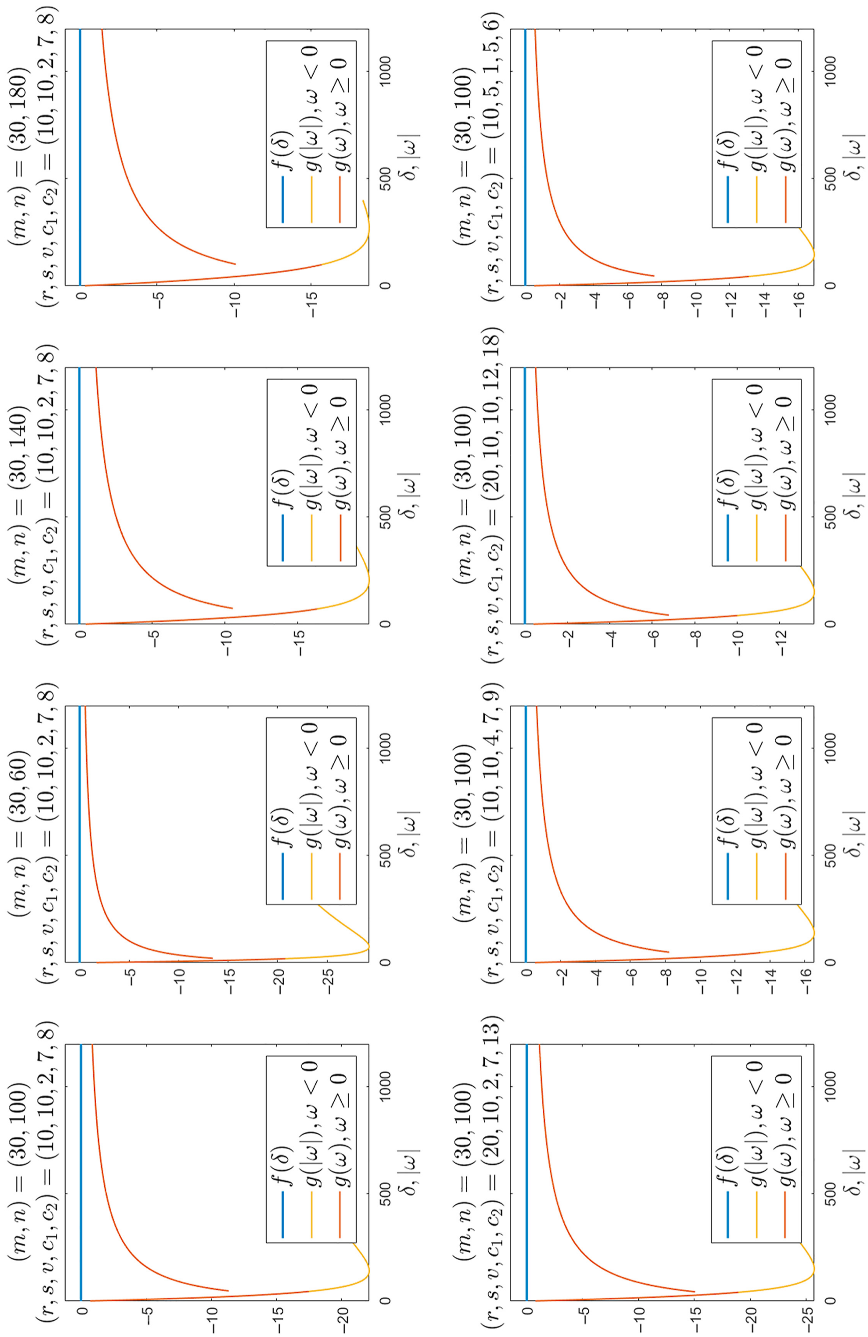


Figure A1. Variation of $\boldsymbol{f(\delta)}$ and $\boldsymbol{g(\omega)}$ under two-point distribution, $\boldsymbol{p = 0.1, 0.4, 0.7, 0.9}$ levels

Source(s): Author's own work



Source(s): Author's own work

Figure A2.
Variation
of$f(\delta)$and$g(\omega)$under
uniform distribution-
append