http://pubsonline.informs.org/journal/msom

MANUFACTURING & SERVICE OPERATIONS MANAGEMENT
Vol. 24, No. 4, July–August 2022, pp. 2221–2239
ISSN 1523-4614 (print), ISSN 1526-5498 (online)

Warranty Reserve Management: Demand Learning and Funds Pooling

Xiao-Lin Wang,[a] Yuanguang Zhong,[b] Lishuai Li,[c] Wei Xie,[b,]* Zhi-Sheng Ye[d]

[a] Business School, Sichuan University, Chengdu 610065, China; [b] School of Business Administration, South China University of Technology, Guangzhou 510641, China; [c] Faculty of Aerospace Engineering, Delft University of Technology, 2628 CD Delft, Netherlands; [d] Department of Industrial Systems Engineering and Management, National University of Singapore, Singapore 117576, Singapore
*Corresponding author

Contact: xlwang28-c@my.cityu.edu.hk, https://orcid.org/0000-0003-0100-8154 (X-LW); bmygzhong@scut.edu.cn, https://orcid.org/0000-0003-2805-8522 (YZ); lishuai.li@tudelft.nl, https://orcid.org/0000-0002-0990-5119 (LL); bmwxie@scut.edu.cn, https://orcid.org/0000-0002-5079-3315 (WX); yez@nus.edu.sg, https://orcid.org/0000-0001-5731-3911 (Z-SY)

Received: October 25, 2019
Revised: January 18, 2021; December 14, 2021
Accepted: December 24, 2021
Published Online in Articles in Advance: March 31, 2022

https://doi.org/10.1287/msom.2022.1086

Abstract. ***Problem definition*:** Warranty reserves are funds used to fulfill future warranty obligations for a product. In this paper, we investigate the warranty reserve planning problem faced by a manufacturing firm who manages warranties for multiple products. ***Academic/practical relevance*:** It is nontrivial to determine a proper amount of reserves to hold, because warranty expenditures are random in nature and reserving either excess or insufficient cash would incur losses. How can warranty reserve levels be optimized and promptly adjusted is a focal issue, especially for firms selling multiple products. ***Methodology*:** Inspired by the general pattern of empirical warranty claims data, we first develop an aggregate warranty cost (AWC) forecasting model for a single product by coupling stochastic product sales and failure processes, which can be used to plan for warranty reserves periodically. The reserve levels are then optimized via a distributionally robust approach, because the exact distribution of AWC is generally unknown. To reduce the losses generated from managing the funds, we further investigate two potential loss-reduction approaches: demand learning and funds pooling. ***Results*:** For the demand learning algorithm, we prove that, as the sales period grows, the optimal learning parameter asymptotically converges to a constant in a fairly fast rate; our simulation experiments show that the performance of demand learning is promising and robust under general warranty claim patterns. Moreover, we find that the benefits of funds pooling change over different stages of the warranty life cycle; in particular, the relative pooling benefit in terms of reserve losses is nonincreasing over time. ***Managerial implications*:** This study offers guidelines on how manufacturers should adaptively forecast and dynamically plan warranty reserves over the warranty life cycle.

Funding: This work was supported by the National Natural Science Foundation of China [Grants 71971085, 71871097, and 72071138] and the Natural Science Foundation of Guangdong Province [Grant 2020A1515011270].
Supplemental Material: The online supplement is available at https://doi.org/10.1287/msom.2022.1086.

Keywords: adaptive learning • distribution free • newsvendor • reserve management • risk pooling

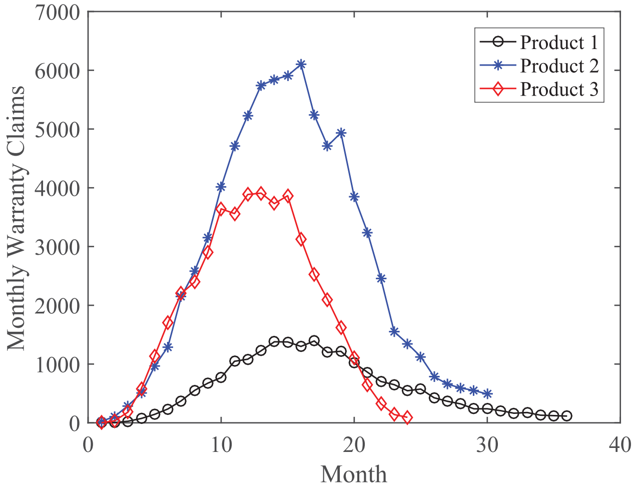
1. Introduction

As one of the most important after-sales services, product warranty, which provides protection against premature failures for customers, has been adopted by most manufacturers to signal the quality and reliability of their products. In general, a warranty contract defines a specific protection period, in which free rectifications of or refund compensations for failed units will be offered by the manufacturer (Gallego et al. 2015, Pinçe et al. 2016). Although warranty services can address customers' concerns, managing a warranty program is costly from the manufacturer's perspective, which can account for as much as 15% of net sales. Essentially, warranty expenses are random in nature, because they are closely related to two coupled stochastic processes—product sales and failure processes. For the manufacturer, servicing future warranty claims can incur liquidity risk, such as a shortage of cash, when unanticipated claims occur. A common solution to these issues adopted by manufacturers is to create an independent warranty reserve fund to cover contingent liabilities that arise from warranty obligations. As an industry practice, the main functions of a warranty reserve fund include: (1) signaling the firm's future performance to the stock market; (2) fulfilling contingent liabilities for future warranty service demands; and (3) playing as a tool of earnings management (Cohen et al. 2011).

However, an ever-increasing warranty reserve level is a major challenge for many manufacturers, especially the big ones. For example, Apple had a $3.834 billion balance in its warranty reserve fund at the end of September 2017, and the figures for Ford and General Motors were $5.031 billion and $8.479 billion, respectively.[1] Many manufacturers, especially automakers, tend to keep much more money in their reserve funds. According to a report by *Warranty Week*,[2] U.S.-based manufacturers put aside 1.8% of product sales (on average) for future warranty claims, which is 17.2 times the amount they actually pay out in warranty claims per month. One of the main reasons is that many manufacturers are unable to accurately predict future warranty expenses, causing an over stock in warranty reserves. Conversely, it presents a bad image to customers if a manufacturer's actual warranty expense is beyond its expectation, which reduces its earnings and upsets its shareholders.[3] Therefore, warranty reserve forecasting and management is an urgent yet challenging problem for manufacturers.

Typically, the *warranty life cycle*[4] of a product follows the product life cycle to support the associated service claims (Murthy et al. 2004, Khawam et al. 2007). Servicing field warranty claims accounts for the biggest share of reserve consumptions, because the manufacturer is fully or partially responsible for the associated servicing expenses. Figure 1 shows the monthly warranty claims for three products of a leading electronics manufacturer over their warranty life cycles. One can observe that the warranty life cycle can be roughly partitioned into three stages: ramp-up, steady, and ramp-down. In the first stage, a new product is introduced to the market and its sales start growing. As a result, more and more potential failures will occur, resulting in an increasing number of warranty claims. After that, the warranty life cycle enters a mature stage in which both the number of units under warranty and

Figure 1. (Color online) Monthly Warranty Claims for Three Electronic Products



the number of claims remain roughly stable but with considerable variability. The reason is that new sales consecutively occur and warranties on earlier-sold units expire. Finally, the cycle reaches a ramp-down stage when the manufacturer discontinues the product, that is, no new sales are observed and the warranties on remaining units in service gradually expire. This pattern of warranty claims is quite general and has been reported by some existing studies (Khawam et al. 2007, Calmon and Graves 2017, Calmon et al. 2021). Despite the generality of this time-varying pattern, how to adaptively forecast future warranty claims and then dynamically schedule necessary warranty reserves, at the operational level, is underexplored.

This research is motivated by the warranty reserving issues faced by the aforementioned electronics manufacturer. The firm would like to determine optimal reserve levels throughout the warranty life cycles of its products to reduce the corresponding reserve losses. One should distinguish *warranty costs* from *reserve losses*: The former stands for the expenses resulting from the servicing of warranty claims, whereas the latter represents the costs or losses resulting from reserving excess and/or insufficient funds. In this work, we first develop an adaptive aggregate warranty cost (AWC) forecasting model to explicitly capture the underlying mechanism of warranty claim generation. In particular, to synthesize the uncertainties induced by product sales and failure processes, we derive two important statistics, the mean and (approximate) variance of AWC increment within an arbitrary time interval (Theorem 1), which are two key metrics for warranty reserve planning. The proposed model is a discrete variation of the AWC forecasting model in Xie and Ye (2016) and raises an obvious advantage compared with theirs (only focused on the AWC within $[0, t]$); that is, one can dynamically forecast the AWC increment and then plan the associated reserves within an arbitrary time interval (e.g., a fiscal quarter) to facilitate short-term and faster liquidity. Then, given the mean and variance of AWC increments, we derive optimal reserve levels periodically for a product throughout its warranty life cycle via a distributionally robust approach (Proposition 2). In addition, we compare the total reserve losses generated by the worst-case scenario with those by a benchmark—the simulated "true-distribution" scenario. The result shows that the worst-case reserving policy performs fairly well.

We further investigate two efficient approaches—demand learning and funds pooling—to explore potential reserve-loss-reduction opportunities. The demand learning model dynamically updates the reserving plan with field warranty data through an exponential smoothing-type mechanism. We prove that, as the sales period grows, the optimal learning parameter asymptotically converges to a constant in probability in a fairly fast rate (Propositions 3 and 4).

The performance of demand learning for general claim patterns is demonstrated via simulation experiments. Moreover, we analyze the benefits of funds pooling for a manufacturer who manages warranties for multiple products (e.g., Apple handles warranty claims for Mac, iPhone, and iPad). Under some mild conditions, we obtain an important finding, that is, the relative pooling benefit in terms of reserve losses decreases as the relative range of standard deviations increases (Proposition 6). In particular, we examine how pooling benefits change over different stages of the warranty life cycle. An interesting result is that when there are only two products, the relative pooling benefit in terms of reserve losses shows a nonincreasing trend (Corollary 2). Finally, a case study using a field data set is presented to demonstrate the effectiveness of the proposed reserve planning methodologies.

The remainder of this paper is structured as follows. Section 2 briefly reviews the relevant literature. Section 3 presents the adaptive AWC forecasting technique and the distributionally robust reserve planning model. The potential loss reduction approaches, that is, demand learning and funds pooling, are investigated and discussed in Sections 4 and 5, respectively. Section 6 presents a real-world case study. Section 7 concludes the paper. Some extended results and all proofs are relegated to the online supplement.

2. Literature Review

Warranty reserve management has been a crucial topic in the literature since the 1960s. Early efforts were mostly devoted to the estimation/prediction of warranty reserve demand (Menke 1969, Patankar and Worm 1981, Thomas 1989, Eliashberg et al. 1997, Ja et al. 2002). These studies focus predominately on estimating/predicting total warranty expenses over the warranty period or the entire product life cycle, which overlook the time-varying characteristics of warranty expenses and can only be used to determine warranty reserves from a long-term perspective. Nevertheless, the research related to (dynamic) cash management of warranty reserves is still scarce. Tapiero and Posner (1988) study a warranty reserving problem to specify the fraction of revenues added to the reserve fund after each sale. They define a compound Poisson process for warranty claims, which is generally not the case in reality. Buczkowski and Kulkarni (2006) aim at optimizing the initial reserve level and the contribution amount from each sale to ensure that the fund covers warranty liabilities over a given time period with a prespecified probability. Gurgur (2011) tackles dynamic cash management of warranty reserves with the objective of minimizing total reserve losses. It is found that when there is no contribution after each sale, the optimal policy is a *base reserve* policy—resembling the base stock policy in the inventory theory—that achieves a base reserve level at the beginning of each period by cash deposits or withdrawals. Fundamentally, our work is closely related to the papers that analogously treat cash flows as "commodity inventories" (i.e., modeling the cash management problem in a mathematically equivalent form to an inventory control problem). As an early attempt to exploit this analogy, Baumol (1952) applies the classical lot size model to solve a deterministic cash transactions problem. Then, Miller and Orr (1966) propose a two-parameter control limit policy for a stochastic cash balance problem. Following the two seminal works, cash management problems have received much attention (Eppen and Fama 1969, Neave 1970, Gormley and Meade 2007, Chen and Simchi-Levi 2009). In essence, warranty reserve management falls into the broad scope of corporate cash management that focuses on the control and planning of cash balance for business firms.

Our research differs from the previously mentioned ones in the following aspects: First, their models cannot characterize the dynamics in the process of generating warranty claims over the entire warranty life cycle, whereas the proposed AWC forecasting model overcomes this issue by fully coupling stochastic sales and failure processes. Second, because it is difficult to derive the exact distribution of warranty costs, they usually use the normal approximation technique to address this problem. In contrast, our research adopts a *distributionally robust* approach for reserve planning purposes, which is more conservative yet robust. Third, although our research simplifies the decision-making process involved in the dynamic reserve management problem to a static approximation, the exponential smoothing–type learning mechanism incorporates the "dynamics" into our model and improves its performance significantly. Finally, we take the same inventory view (precisely, in a "newsvendor" framework) as in the aforementioned cash management literature but deal with a distinct type of stochastic demand source—warranty reserve demand—that exhibits a general increasing-stable-decreasing pattern.

Another research stream—demand learning in inventory management (see Mišić and Perakis (2020) for a recent overview)—is also closely related to our work. Existing studies on this topic focus on either single-period or multiperiod inventory control settings. In the former (i.e., newsvendor) setting, demand distribution is generally assumed to be unknown while historical data are available (Huber et al. 2019, Besbes and Mouchtaki 2021, Keskin et al. 2021). In the latter setting, demands across periods are usually assumed to be independent and identically distributed (Chen and Chao 2019, Zhang et al. 2020). In general, the warranty reserving problem is a multiperiod problem and, from Figure 1, the reserve demands across periods are essentially dependent and nonidentically distributed. This necessitates new data-driven methodologies that are capable to learn such demand from data. In this work, we propose an exponential smoothing–type learning mechanism, in which the optimal learning parameter is sequentially optimized by minimizing empirical total reserve losses directly. By leveraging the optimization problem structure (i.e., the objective function) in the learning process, we combine learning and optimization together to generate a better reserving plan. This also provides a response to Mišić and Perakis (2020) who call for new approaches to *predict-then-optimize* in operations management applications.

Finally, the paper contributes to the risk pooling literature in operations management as well. The idea of funds pooling is similar to that of inventory pooling, where demand variability is mitigated by aggregating demands across products or locations to reduce safety stocks and consequently inventory costs (Eppen 1979, Benjaafar et al. 2005, Berman et al. 2011). In recent years, many efforts have been devoted to studying the behaviors of risk pooling in various inventory settings, for example, heavy-tailed demands (Bimpikis and Markakis 2016), multilocation newsvendor (Govindarajan et al. 2021, Yang et al. 2021), and resource allocation (Karsten et al. 2015, Zhong et al. 2018). Nevertheless, most previous literature on inventory pooling deals only with the variability of product sales and considers relatively simple forms of demand distributions. The time-varying characteristics of warranty reserve demands—a new finite horizon demand source—over the warranty life cycle distinguish our problem from traditional inventory pooling issues and pose a challenge to modeling and analysis. In addition, the benefits of risk pooling in a multiperiod inventory setting have received limited attention (Tagaras and Cohen 1992, Bimpikis and Markakis 2016). In this research, we are particularly interested in an important perspective of risk pooling, *time-varying benefit behavior*, in multiperiod warranty reserve management.

3. Warranty Reserve Forecasting and Planning

In this section, we first develop a discrete-time reserve demand forecasting model that is adaptive to the uncertainties from product sales and failure processes. Then, we use a distributionally robust optimization model to periodically determine optimal reserve levels and evaluate the performance of the worst-case reserving model.

3.1. Warranty Reserve Demand Forecasting

Let L be the life cycle (sales period) of a product, that is, from its launch time to the date it is discontinued. Empirically, sales periods are finite and short for consumer electronics, due to fast technology obsolescence. For instance, Apple releases new generations of its products every year while previous generations are soon discontinued. Denote $N(t)$ as the cumulative number of units sold within $[0, t]$, $t \in [0, L]$.

Assumption 1. *The cumulative product sales $N(t)$ followsa homogeneous Poisson process (HPP) with rate λ (i.e., HPP(λ)).*

This assumption is relatively strong, especially when the magnitude of sales volume is large. However, it is most natural to assume that new-product demand arrivals follow an HPP, when sales curves do not exhibit obvious patterns. In addition, this assumption might also be acceptable for the sales of parts, which is relatively stationary due to their long life cycles.[5] Nevertheless, we acknowledge that this assumption is to make the problem analytically tractable. To relax the HPP assumption and generalize the adaptivity of our method to a larger scope of products with nonstationary sales, in Section 4.3, we will incorporate a nonhomogeneous Poisson process (NHPP)—which can capture a variety of time-varying patterns—to simulate real-world sales processes to further validate the robustness of the proposed methodologies.

Assumption 2. *The time to first failure of this product follows an exponential distribution $G(t;\theta)$ with a constant failure rate θ.*

The failure processes of most products exhibit three phases—infant mortality, useful life, and wear out, which correspond to a decreasing failure rate, a constant failure rate, and an increasing failure rate, respectively. In real applications, a constant failure rate has been demonstrated to be a good approximation for a unit that is under warranty (Blischke and Murthy 2000). This is because (i) the units must pass quality and reliability testing (e.g., burn-in and screening tests) before being released into the market, for which most of them have passed through the infant mortality period and entered the useful life period, and (ii) the warranty periods for most products are much shorter than their useful life periods. In particular, the failure rates for most electronics are, inherently, almost constant (Blischke and Murthy 2000).

Assumption 3. *The product is sold with a nonrenewing free replacement warranty policy of length W.*

Under this policy, if a unit is sold at time zero and fails at time $t \in [0, W]$, then the manufacturer will provide a free replacement that comes with a remaining warranty period of $W - t$. The nonrenewing free replacement warranty policy has been widely adopted for products that are nonrepairable or, economically, not worth repairing—for example, smartphones, battery packs, and some electric tools. Under this warranty policy, a failed unit or its critical component will be restored to a functioning state that is as good as new, through replacement by an identical new one.

Based on the aforementioned assumptions, we now introduce an adaptive model that can fully capture the interaction of product sales and under-warranty failures to forecast the AWC increments, which forms the foundation for short-term warranty reserve planning. Suppose that the warranty period is shorter than the sales period (i.e., $W \le L$), which is quite common in practice. The warranty life cycle $[0, W+L]$ is of interest to us. Let T_i ($i \in \{1, 2, \ldots, N(t)\}$) be the random purchase date of the ith sold unit. Let $C_i(\tau_i)$ ($\tau_i = \min\{t - T_i, W\}$) quantify the random warranty cost of the ith sold unit up to time $t \ge T_i$. Given T_i, the expected warranty cost of the ith sold unit up to time t can be calculated by the well-known renewal equation:

$$\mathbb{E}[C_i(\tau_i)] = C_W m(\tau_i) = C_W\left[G(\tau_i;\theta) + \int_0^{\tau_i} m(\tau_i - x)\mathrm{d}G(x;\theta)\right] = C_W\theta\tau_i, i \in \{1, 2, \ldots, N(t)\}, \quad (1)$$

where C_W is the average replacement cost per failure and $m(\cdot)$ is the renewal function associated with $G(t;\theta)$. Servicing a warranty claim incurs various costs: administration, transportation, labor, and materials, among others. For the sake of convenience, we combine these costs into a single term C_W.

Define *warranted base* as the number of sold units that are still under warranty. Figure 2 illustrates the counting process of AWC during the warranty life cycle. It is important to point out that all units are eligible to generate warranty claims during their warranty periods, whereas the warranted base is changing over time. Because different units have different dates of purchase, the AWC up to time t can be formulated as

$$WC(t) = \begin{cases} \sum_{i=1}^{N(t)} C_i(\tau_i), & 0 \le t \le L, \\ \sum_{i=1}^{N(L)} C_i(\tau_i), & L < t \le L + W. \end{cases} \quad (2)$$

Industry practice shows that warranty reserves are planned periodically, for example, quarterly, semiyearly, or yearly, for accounting purposes (Cohen et al. 2011). In accordance with this practice, we need to evaluate the AWC increments within a sequence of equal-length time intervals over the warranty life cycle. Suppose that the length of reserve planning cycle is Δ, say, one quarter. We assume that the ratios of warranty period W and sales period L to Δ are integers, which is a common practice. Without loss of generality, we rewrite $W = m\Delta$ ($m \ge 1$), $L = n\Delta$ ($n \ge m$ as $L \ge W$), and $L + W = (n+m)\Delta$. Hence, the total number of planning periods is $n + m$ over the warranty life cycle. Let $WC_k = WC(k\Delta) - WC((k-1)\Delta)$ denote the AWC increment in the kth planning period ($k \in \{1, 2, \ldots, n+m\}$). In general, it is difficult, if not impossible, to obtain the exact distribution of WC_k because of the complexity of warranty costs (Ja et al. 2002, Buczkowski and Kulkarni 2006, Gurgur 2011). Fortunately, we can derive closed-form expressions of the mean and approximate variance of WC_k, which are two essential statistics for further analysis.

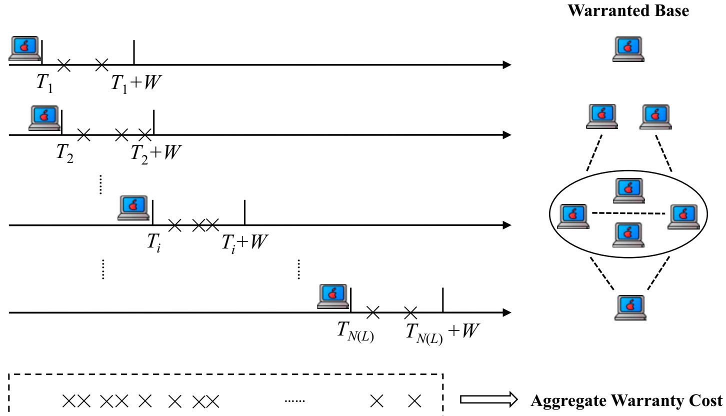
Theorem 1. *Let μ_k and σ_k^2 be the mean and variance of WC_k, namely, $\mu_k = \mathbb{E}[WC_k]$ and $\sigma_k^2 = \mathbb{V}\mathrm{ar}(WC_k)$. Then, one has*

$$\mu_k = \begin{cases} C_W\lambda\theta\Delta^2\left(k - \frac{1}{2}\right), & k = 1, \ldots, m, \\ C_W\lambda\theta\Delta^2 m, & k = m+1, \ldots, n, \\ C_W\lambda\theta\Delta^2\left(n + m - k + \frac{1}{2}\right), & k = n+1, \ldots, n+m, \end{cases} \quad (3)$$

and

$$\sigma_k^2 \simeq \begin{cases} C_W^2\lambda\theta\left[\Delta^2\left(k - \frac{1}{2}\right) + \theta\Delta^3\left(k^2 - k + \frac{1}{3}\right)\right], & k = 1, \ldots, m, \\ C_W^2\lambda\theta(1 + \theta m\Delta)\Delta^2 m, & k = m+1, \ldots, n, \\ C_W^2\lambda\theta\left[\begin{array}{l}\Delta^2\left(n + m - k + \frac{1}{2}\right) \\ +\theta\Delta^3\left(m^2 - n^2 - n + k + 2nk - k^2 - \frac{1}{3}\right)\end{array}\right], & k = n+1, \ldots, n+m. \end{cases} \quad (4)$$

Figure 2. (Color online) AWC Accumulation Process During the Warranty Life Cycle



It is worth pointing out that the asymptotic approximation in (4) is fairly accurate under practical settings.[6] Hereafter, we will use the terms *approximate variance* and *variance* interchangeably. The mean μ_k and variance σ_k^2 are inherently correlated, because both are dependent on k, C_W, λ, and θ. In this scenario, we cannot change one while keeping the other unchanged. This property renders the warranty reserving problem different from traditional inventory problems, in which the mean and variance of product demands are mostly assumed to be independent.

In practice, it is not uncommon that the means and variances of AWC increments for different planning periods or products might have distinct magnitudes. As a consequence, the coefficient of variation, $CV_k = \sigma_k/\mu_k$, is a useful alternative to measuring the risks and comparing the uncertainties involved in AWC increments for different planning periods or products. The following lemma shows the role of CV_k.

Lemma 1. *For any $\epsilon > 0$, we have*

$$\Pr\left\{\left|\frac{WC_k}{\mu_k} - 1\right| > \epsilon\right\} \le \frac{CV_k^2}{\epsilon^2}, k = 1, 2, \ldots, n+m.$$

Lemma 1 tells us that given an acceptable error ϵ, the variability of WC_k relative to estimated μ_k can be well controlled by the bound CV_k^2/ϵ^2. This means that the risks of relative forecast errors follow the same pattern as that of CV_k for different periods and/or products. According to Equations (3) and (4), CV_k can be approximated by

$$CV_k \simeq \begin{cases} \dfrac{\sqrt{(k-1/2)+\theta\Delta(k^2-k+1/3)}}{\sqrt{\lambda\theta}\Delta(k-1/2)}, & k = 1, \ldots, m, \\ \dfrac{\sqrt{1+\theta m\Delta}}{\sqrt{\lambda\theta m\Delta}}, & k = m+1, \ldots, n, \\ \dfrac{\sqrt{\begin{array}{l}(n+m-k+1/2) \\ +\theta\Delta(m^2-n^2-n+k+2nk-k^2-1/3)\end{array}}}{\sqrt{\lambda\theta}\Delta(n+m-k+1/2)}, & k = n+1, \ldots, n+m. \end{cases} \tag{5}$$

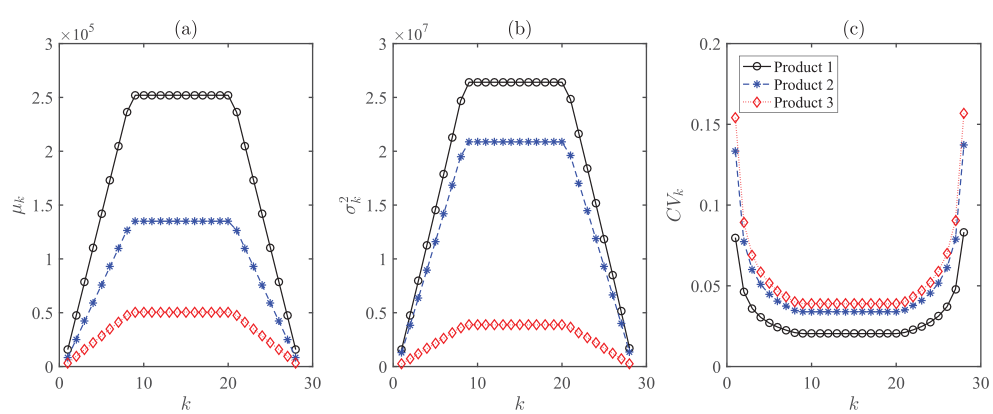
Proposition 1. *The terms μ_k and σ_k^2 are increasing in $k \in \{1, \ldots, m\}$, constant in $k \in \{m+1, \ldots, n\}$, and decreasing in $k \in \{n+1, \ldots, n+m\}$. In contrast, with respect to the same counting sets, CV_k shows an opposite decreasing-constant-increasing trend.*

The tendencies of mean μ_k, variance σ_k^2, and coefficient of variation CV_k over the warranty life cycle are illustrated in Figure 3. One can observe that the curves of both μ_k and σ_k^2 exhibit an upside-down bathtub shape, whereas that of CV_k is bathtub shaped. Basically, the trends of μ_k and σ_k^2 are analogous to the pattern in Figure 1, whereas the time-varying behavior of CV_k over the warranty life cycle is an interesting result. In the first stage of the warranty life cycle (i.e., $1 \le k \le m$), the variability indicated by CV_k shows a downward trend as time goes by. This is because the larger the warranted base, the lower the risk (Calmon and Graves 2017). Then, in the second stage (i.e., $m+1 \le k \le n$), demand variability becomes stable because the warranted base moves into a relatively steady state. Finally, the warranted base shrinks in the third stage (i.e., $n+1 \le k \le n+m$), which in turn increases the risk caused by uncertainties. That is to say, the AWC variability measured by the coefficient of variation decreases (respectively, increases) as the expected AWC increases (respectively, decreases). This result is consistent with our investigation of the behavior of warranted base over the warranty life cycle. It demonstrates the importance and necessity of adopting the coefficient of variation to capture demand variability in relation to mean μ_k.

3.2. Distributionally Robust Warranty Reserve Planning

Based on the previous reserve demand forecast, we develop an optimization model to periodically plan for warranty reserves over the warranty life cycle. Setting a proper amount of warranty reserves is nontrivial in

Figure 3. (Color online) Patterns of μ_k, σ_k^2, and CV_k for Three Products



Notes. The parameters are arbitrarily set as W = 2 years, L = 5 years, $\Delta = 1/4$ year, $\lambda_1 = 210{,}000$, $\theta_1 = 0.024$, $C_{W1} = \$100$, $\lambda_2 = 120{,}000$, $\theta_2 = 0.015$, $C_{W2} = \$150$, $\lambda_3 = 135{,}000$, $\theta_3 = 0.010$, and $C_{W3} = \$75$, for illustrative purposes. In this setting, the warranty life cycle is $L + W = 7$ years, $m = W/\Delta = 8$, and $n = L/\Delta = 20$.

general, because the warranty expenditures are random—induced by stochastic product sales and failure processes—and both over reserving and under reserving may cause additional losses. More specifically, over reserving results in an opportunity cost in the form of lost interest, whereas under reserving might necessitate calling for emergency funds from other sources at an above-average interest rate (Gurgur 2011). Therefore, the risk of holding excess or insufficient reserves should be taken into consideration when optimizing the warranty reserve levels.

In the previous section, we have derived the mean and variance of the AWC increment in any reserve planning period. Under the assumption that there is no cash contribution to the reserve fund from each sale, the AWC increment in each period is exactly the reserve demand. As mentioned earlier, however, its exact distribution is generally unknown. To overcome this difficulty, we adopt the well-known distributionally robust approach in Scarf (1958) to determine the optimal warranty reserve level for each period. Let C_h be the unit over reserving cost and C_b be the unit under reserving cost, which can be estimated by the time values of money.[7] Unlike the material property of commodity inventories, the monetary property of warranty reserves allows the manufacturer to calculate reserve levels for different periods independently (without leftover from previous periods). For a specific period $k \in \{1, 2, \ldots, n+m\}$, the manufacturer should determine an optimal amount of warranty reserves to minimize the associated expected reserve loss $\pi_k(WR_k)$, which is expressed as

$$\begin{aligned}\pi_k(WR_k) &= \mathbb{E}[C_h(WR_k - WC_k)^+ + C_b(WC_k - WR_k)^+] \\ &= C_h\mathbb{E}\big[(WR_k - WC_k) + (WC_k - WR_k)^+\big] \\ &\quad + C_b\mathbb{E}[WC_k - WR_k]^+ \\ &= C_hWR_k - C_h\mu_k + (C_h + C_b)\mathbb{E}[WC_k - WR_k]^+,\end{aligned} \tag{6}$$

where $x^+ = \max\{0, x\}$.

Because deriving the exact distribution of WC_k is highly intractable, we minimize Equation (6) against the worst case. According to Gallego and Moon (1993), the following lemma holds.

Lemma 2. *The following inequality holds.*

$$\mathbb{E}[WC_k - WR_k]^+ \le \frac{\sqrt{\sigma_k^2 + (WR_k - \mu_k)^2} - (WR_k - \mu_k)}{2}. \tag{7}$$

Moreover, for each WR_k, there exists a distribution of WC_k in which the bound (7) is tight.

Based on Lemma 2, the distributionally robust warranty reserving problem turns to minimize the following upper bound:

$$\widehat{\pi}_k(WR_k) = C_hWR_k - C_h\mu_k + \frac{1}{2}(C_h + C_b) \times\left[\sqrt{\sigma_k^2 + (WR_k - \mu_k)^2} - (WR_k - \mu_k)\right]. \tag{8}$$

In general, it is reasonable to assume $C_h < C_b$ in real applications. Then, the following result can be obtained.

Proposition 2. *The distributionally robust optimal warranty reserve level for the kth planning period is given by*

$$WR_k^* = \mu_k + \frac{1}{2}A\sigma_k, \tag{9}$$

where $A = \sqrt{C_b/C_h} - \sqrt{C_h/C_b} > 0$.

Substituting Equation (9) into Equation (8), we can obtain the minimized expected reserve loss in the kth period as

$$\widehat{\pi}_k^*(WR_k^*) = \frac{1}{2}B\sigma_k, \tag{10}$$

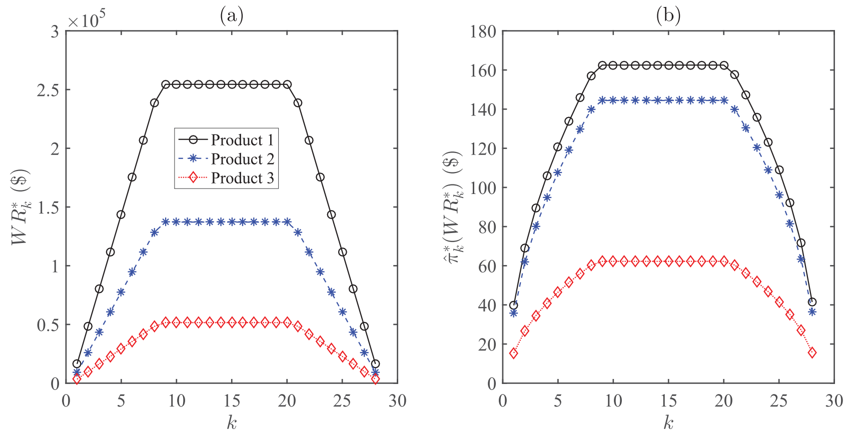
where $B = C_h(\sqrt{1 + A^2/4} + A/2) + C_b(\sqrt{1 + A^2/4} - A/2) > 0$.

Remark 1. The optimal warranty reserve level WR_k^* increases as C_b/C_h, μ_k, or σ_k increases, whereas the expected reserve loss $\widehat{\pi}_k^*(WR_k^*)$ is independent of μ_k but increasing in σ_k.

An intuitive result from Remark 1 is that in any planning period, if the ratio of C_b/C_h or the mean and/or standard deviation of AWC increment becomes larger, the manufacturer should specify a higher reserve level, whereas the expected reserve loss resulting from over-reserving and/or under-reserving depends only on the standard deviation of AWC increment. Figure 4 shows the warranty reserve levels and the associated losses for the three products (with the same parameter setting as in Figure 3). As observed, the warranty reserve levels exhibit similar trends over the warranty life cycle, which can be well explained by the time-varying warranted base. In addition, the tendency of expected reserve losses is related to, yet slightly different from, that of warranty reserves.

Although the distributionally robust optimization technique can generate a robust reserving plan against the worst case, the plan tends to be conservative in most cases, possibly causing additional losses. To assess the performance of the proposed worst-case reserving model, we compare it with a benchmark—the simulated nonparametric reserving model. Specifically, based on Assumptions 1–3, we implement Monte Carlo simulation to generate an empirical distribution of the AWC increment in each period, which can be regarded as the "true distribution" and serves as a perfect benchmark for the worst-case performance

Figure 4. (Color online) Warranty Reserve Levels and Expected Reserve Losses of the Three Products (C_h = \$0.02 and C_b = \$0.05)



assessment. Our simulation studies show that, compared with the benchmark, the distributionally robust reserving model performs fairly well under a wide range of practically reasonable parameter settings. For more details, please refer to Section EC.1 in the online supplement.

In what follows, based on the worst-case reserving model, we investigate two possible reserve-loss-reduction approaches: demand learning and funds pooling. The results will be discussed in two separate sections for the sake of convenience.

4. Reserve Loss Reduction via Demand Learning

In real applications, warranty claims data are usually subject to a high degree of variability due to complex sales and failure patterns. For instance, if the sales record of a product presents a nonstationary pattern, then the reserving model under the HPP sales assumption cannot guarantee a satisfactory performance. As a result, although the proposed simple model in Section 3 delivers a fairly good performance under some specific assumptions, there is still a potential to improve its adaptivity and performance for general claims data. In this section, we aim to augment the performance of our simple reserving model through demand learning.

4.1. Exponential Smoothing–Type Learning

As discussed before, most of the existing demand learning models in the inventory literature cannot be directly applied in our context, because warranty reserve demands across periods are dependent and nonidentically distributed. Another practical hurdle is that one cannot repeatedly collect warranty claims data for a specific product; actually, only one single data point can be collected in each planning period. Therefore, there is not enough data to learn the distributional properties of warranty reserve demands across periods.

Inspired by the demand-chasing heuristics in Bolton and Katok (2008) and Bostian et al. (2008), we implement a simple exponential smoothing–type mechanism to update the reserving plan derived from the distributionally robust model using observed warranty data, in which the smoothing factor is learnt dynamically to chase the time-varying reserve demands. Specifically, the adjusted reserve level $\widehat{WR}_k$ in the kth planning period is formulated as

$$\widehat{WR}_k = WR_k^* + \phi_k(\Delta Q_{k-1} - \mu_{k-1}), k = 2, 3, \ldots, \tag{11}$$

where $\phi_k \geq 0$ is the smoothing/learning factor (which is dynamically updated from period to period), WR_k^* is the "optimal" reserve level given by Equation (9), and ΔQ_k represents actual warranty expenses (i.e., the AWC increment) in period k. In essence, the learning model anchors on the reserve level WR_k^* derived from the distributionally robust model and tends to pull the reserve level down when the forecasted mean μ_{k-1} in the previous period is high and vice versa.

It is worth noting that, unlike the smoothing factor in a classic exponential smoothing model, the learning parameter ϕ_k could take a value that is greater than one. The optimal value of ϕ_k^* in period $k \in \{3, 4, \ldots\}$ is obtained by solving the following optimization problem:

$$\begin{aligned} &\min_\phi \sum_{i=1}^{k-1} \{C_h[WR_i^* + \phi(\Delta Q_{i-1} - \mu_{i-1}) - \Delta Q_i]^+ \\ &+ C_b[\Delta Q_i - WR_i^* - \phi(\Delta Q_{i-1} - \mu_{i-1})]^+\}. \end{aligned} \tag{12}$$

In the first period, we set $\widehat{WR}_i = WR_i^*$, whereas the learning parameter ϕ_2 in the second period could be arbitrarily chosen (say, $\phi_2 = 1$) for setup purposes. In this manner, we learn the parameter ϕ_k^* by minimizing empirical total reserve losses directly, using the collected data from the first $k-1$ periods. By leveraging the objective function in the learning process, we attempt to combine learning and optimization together to generate a better reserving plan. It is straightforward to verify that the optimization problem in (12) is convex; thus, an optimal ϕ_k^* exists.

4.2. Convergence of the Learning Algorithm

We are now in a position to explore some properties of the proposed learning mechanism. We first examine its asymptotic characteristic. To this end, define a convex function $\mathbb{H}(\phi) = \frac{1}{2}C_h A\sigma_{m+2} + (C_b + C_h)\mathbb{E}[\Delta Q_{m+2} - WR^*_{m+2} - \phi(\Delta Q_{m+1} - \mu_{m+1})]^+$, and let ϕ^* be its minimizer. Then, we have the following result.

Proposition 3. *Given the assumed model and planning cycle Δ, ϕ_k^* converges to ϕ^* in probability when the sales period $L = n\Delta \to \infty$.*

Proposition 3 shows that under the aforementioned assumptions, the optimal learning parameters obtained from successive periods asymptotically converge to a constant in probability if the sales period is long enough. This asymptotic characteristic guarantees a stable output of the learning mechanism in the long run. In practice, however, warranty reserve planning horizons (precisely, the warranty life cycles) are never "infinitely long," especially for consumer electronics that are depreciating quickly. Consequently, the rate of learning, that is, the convergence rate of ϕ_k in our case, provides a more important indication that whether the learning mechanism would perform well for products with short warranty life cycles.

Proposition 4. *Given the assumed model and planning cycle Δ, suppose $\phi \longmapsto \mathbb{H}(\phi)$ is second order differentiable at $\phi = \phi^*$. Then, $\phi_k^* - \phi^* = O_P(1/\sqrt{k})$ and $\sqrt{k}(\phi_k^* - \phi^*)$ converges weakly to a mean-zero Gaussian distribution with a finite asymptotic variance.*

Proposition 4 reveals that the rate of convergence of the learning parameter is $1/\sqrt{k}$ and is asymptotically normal. This is to be expected since ϕ_k^* can be regarded as an *M*-estimator. Our finite sample simulation in Section 4.3 also reveals that when k is greater than 10, ϕ_k^* is close enough to ϕ^* in most cases. This demonstrates that the proposed demand learning mechanism is expected to deliever an acceptable performance for products with short warranty life cycles.

In addition, Propositions 3 and 4, combined together, provide a meaningful guidance to algorithmic implementation in real applications. That is, when a manufacturer collects historical warranty claims data of a product for a relatively long period of time, it is expected that the optimal learning parameter ϕ_k^* will converge to a constant ϕ^* in probability as time goes by. In this situation, the optimization step in Equation (12) can be skipped and the adjusted reserve level $\widehat{WR}_k$ can be updated by directly substituting ϕ^* into Equation (11), which simplifies the process.

4.3. Performance Evaluation with General Sales and Failure Patterns

The results in Propositions 3 and 4 hinge heavily on the underlying assumptions on product sales and failure processes (i.e., Assumptions 1 and 2). However, it is difficult, if not impossible, to explore the case when the assumptions are violated. Thus, we resort to simulation experiments to evaluate the performance of the demand learning mechanism. For this purpose, a general setting of sales and failure patterns is adopted to generate warranty claims.

In the simulation experiment, we first simulate sales data by synthesizing the NHPP and the well-known Bass model (Bass 1969). The Bass model is probably the most influential new-product diffusion model. It postulates that the trajectory of expected cumulative sales of a new product follows a deterministic function of time with three parameters: the potential market size κ, the innovator factor p, and the imitator factor q. For the NHPP-Bass model, the expected cumulative sales can be expressed as

$$\Lambda(t) = \kappa \frac{1 - e^{-(p+q)t}}{1 + \frac{q}{p}e^{-(p+q)t}}, \tag{13}$$

and the saturation point is $t^* = \ln(q/p)/(p+q)$, that is, the sales rate reaches its peak at t^*. Given the ith purchase date t_i $(i \in \{1, 2, \dots\})$, we can generate the ith interpurchase time x_i from the distribution function $Z_i(x) = 1 - \exp(-\Lambda(t_i + x) + \Lambda(t_i))$. Then, the $(i+1)$th purchase date is simply $t_{i+1} = t_i + x_i$.

Moreover, to relax the product failure assumption, we assume that the product's time to first failure follows a Weibull distribution with $G(t;\eta,\beta) = 1 - \exp\{-(t/\eta)^\beta\}$, where η and β are scale and shape parameters, respectively. Simulation is conducted in the following steps. For each simulation run, following the illustration in Figure 2, we first generate all the purchase dates within $[0, L]$ according to the NHPP-Bass model. Given the purchase date of each sold unit, we then generate the associated under-warranty failure instant(s) based on the Weibull distribution $G(t;\eta,\beta)$. Finally, counting the total number of warranty claims and the corresponding costs within each planning period yields the simulated AWC increments over the entire warranty life cycle.

Let S be the total number of simulation runs and $\Delta Q_{i,k}$ $(i \in \{1, 2, \dots, S\}, k \in \{1, 2, \dots, m+n\})$ be the simulated AWC increment for planning period k in simulation run i. For each simulation run i, we need to estimate the parameters λ_i and θ_i through least squares (see Section 6 for details). With estimated $\widehat{\lambda}_i$ and $\widehat{\theta}_i$, we can calculate the mean $\mu_{i,k}$ and variance $\sigma_{i,k}^2$ of the AWC increment in period k by Equations (3) and (4) and then determine the corresponding optimal reserve level $WR_{i,k}^*$ by Equation (9). Based on $\{\Delta Q_{i,k}\}_{k\in\{1,2,\ldots,m+n\}}$ and $\{WR_{i,k}^*\}_{k\in\{1,2,\ldots,m+n\}}$, the adjusted reserve level $\widehat{WR}_{i,k}$ can be obtained by Equation (11). Then, substituting $\widehat{WR}_{i,k}$ and $WR_{i,k}^*$ into $\tilde{\pi}_{i,k}(WR_{i,k}) = C_h(WR_{i,k} - \Delta Q_{i,k})^+ + C_b(\Delta Q_{i,k} - WR_{i,k})^+$ yields the simulated reserve losses, with and without learning, respectively. To evaluate the learning benefit, we define $\Omega_i = (\sum_{k=1}^{m+n}\tilde{\pi}_{i,k}(WR_{i,k}^*) - \sum_{k=1}^{m+n}\tilde{\pi}_{i,k}(\widehat{WR}_{i,k}))/\sum_{k=1}^{m+n}\tilde{\pi}_{i,k}(WR_{i,k}^*)$ to represent the percentage of total reserve loss reduction resulting from demand learning, for each run i.

In this simulation experiment, we propose to examine the algorithm performance with respect to two critical factors—the location of sales peak (t^*) and the shape of failure rate function ($\beta t^{\beta-1}/\eta^\beta$). We consider three types of sales peak, that is, early ($t^* < L/2$), middle ($t^* \approx L/2$), and late ($t^* > L/2$), as well as three shapes of failure rate function, that is, constant ($\beta = 1$), concave increasing ($1 < \beta < 2$), and convex increasing ($\beta > 2$). To this end, we fix the values of κ, q, and η at $\kappa = 200{,}000$, $q = 0.5$, and $\eta = 5$, respectively, and alter the values of p and β to satisfy the aforementioned settings. More specifically, we consider $p = 0.238$ ($t^* \approx 1$), $p = 0.109$ ($t^* \approx 2.5$), and $p = 0.054$ ($t^* \approx 4$) for the three types of sales peak, and $\beta = 1.0$, $\beta = 1.5$, and $\beta = 3.5$ for the three shapes of failure rate function, respectively (see Figure 5 for illustrations). Thus, we have nine ($= 3 \times 3$) different yet comprehensive scenarios in total. Other parameters are set as $W = 2$ years, $L = 5$ years, $\Delta = 1/4$ year, $C_h = \$0.02$, $C_b = \$0.05$, and $C_W = \$100$.

To illustrate the performance of demand learning, we arbitrarily simulate one sample for each of the nine scenarios. Figure 6 shows the corresponding simulated warranty costs and planned reserve levels (with and without learning) over the warranty life cycle. It is clear that the curves of simulated warranty costs are generally right-skewed (respectively, left-skewed) when the sales peak occurs early (respectively, late), whereas they are relatively symmetric when the sales peak locates at the middle of the sales period. More interestingly, we can see that the simulated cost curves in panels (b) and (c) are more symmetric than that in panel (a). This phenomenon is driven by the influence of failure rate function: When the shape parameter β is greater than one, the failure rate function increases faster over time (Figure 5), implying that more warranty claims tend to occur later. With a similar logic, although the simulated cost curves in the second row of Figure 6 are relatively symmetric, those in panels (e) and (f) tend to be left-skewed due to the impact of increasing failure rates. Furthermore, all the simulated cost curves in the third row are left-skewed, whereas the skewness of the curve in panel (g) is the smallest.

We note that the curves of mean AWC increments in Equation (3) are symmetric with respect to the midpoint of the warranty life cycle (Figures 3 and 6). As a result, when simulated cost curves are more symmetric, the proposed simple model (with simplified sales and failure processes) performs better for curve fitting. When the curves of simulated warranty costs are either left- or right-skewed, the distributionally robust optimal reserve levels deviate from simulated costs significantly. Nevertheless, the reserve levels adjusted by the learning algorithm are quite close to the simulated costs in all scenarios. In addition, Figure 7 illustrates the evolution of optimal learning parameters over the warranty life cycle for the nine scenarios in Figure 6. One can observe that in general sales and failure settings, the optimal learning parameters tend to converge as well. Moreover, the rate of convergence

Figure 5. (Color online) Sales Rates (a) and Failure Rates (b) Under Various Parameter Settings

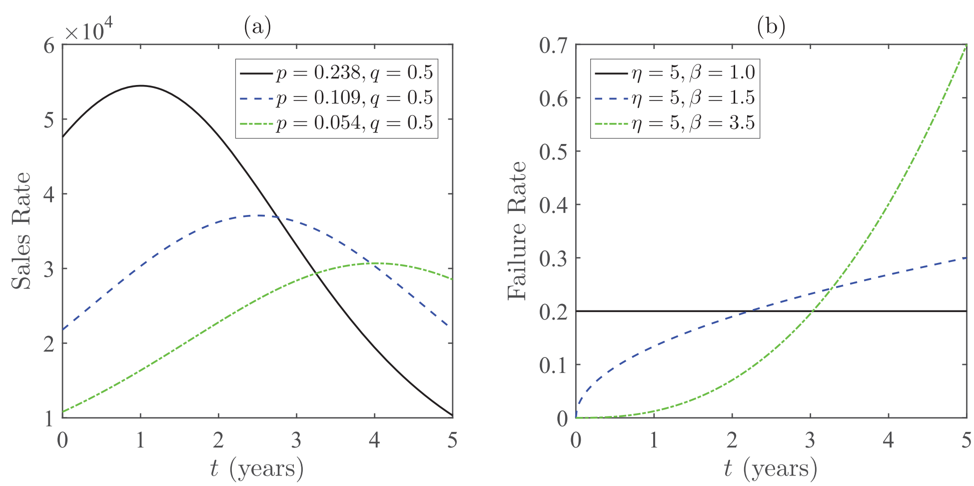
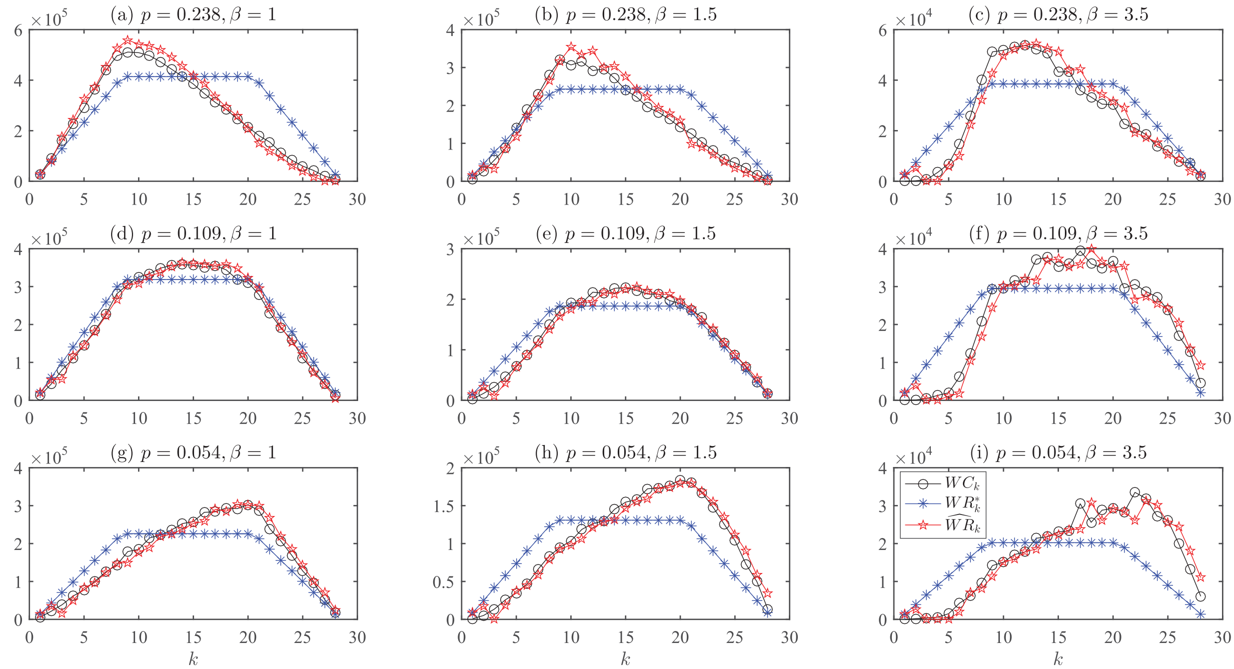


Figure 6. (Color online) Simulated Warranty Costs and Planned Reserve Levels (Without and with Learning) for Nine Scenarios



is fairly fast: when k is greater than 10, ϕ_k^* is close enough to ϕ^* in most scenarios. This implies that the exponential smoothing-type mechanism is capable to learn warranty reserve demands for products with both long and short life cycles.

We further examine the cost performance of demand learning through a large number of simulation runs. Figure 8 shows the box plots of Ω_i over 1,000 simulation runs for the nine scenarios. It is clear from this figure that demand learning can lead to a significant reduction in total reserve losses. The value of Ω_i can be as high as 83.4% (the maximum value in scenario 9). An interesting observation is that the smaller (respectively, larger) the discrepancy between analytical model and simulated curve, the smaller (respectively, larger) the learning benefit Ω_i. As discussed previously, the discrepancies between analytical model and simulated curve in scenarios 3, 4, and 7 (corresponding to panels (c), (d), and (g) in Figure 6) are comparatively small in their respective subgroups.

Figure 7. (Color online) Convergence of Optimal Learning Parameters ϕ_k^* in the Corresponding Nine Scenarios in Figure 6

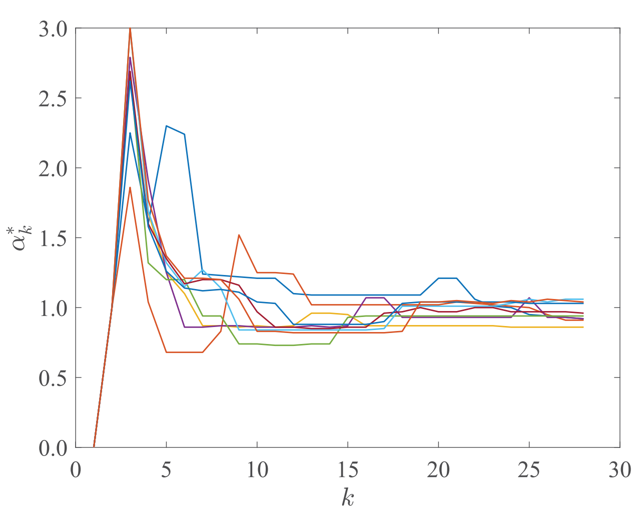
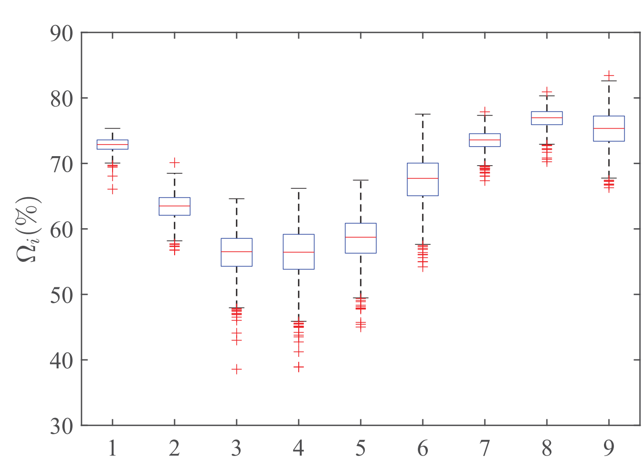


Figure 8. (Color online) Box Plots of Ω_i in the Nine Scenarios



As a result, the benefits of demand learning in these scenarios are relatively small. More specifically, the learning benefit in scenario 3 (respectively, 4 and 7) is the smallest among the first (respectively, middle and last) three scenarios. This observation indicates that the proposed demand learning algorithm performs better when the AWC forecasting model fails to fully capture the dynamics of actual warranty costs.

Another interesting phenomenon is regarding the dispersion—reflected by the interquartile range of each box—of learning benefits Ω_i over the 1,000 simulation runs. One can observe that scenarios 4, 5, and 6 (corresponding to middle sales peak) have a larger dispersion, relative to those for early and late sales peaks. The reason is that under this case, the analytical model matches the simulated curve better, for which the learning benefits are subject to a higher degree of randomness. In addition, when $\beta = 3.5$ (corresponding to scenarios 3, 6, and 9), the associated dispersion is relatively larger as well. This is because when $\beta = 3.5$, the failure rate function over the warranty period (i.e., $0 \le t \le 2$ years) is lower (Figure 5). Thus, the expected number of warranty claims per unit sold would be fewer. As a result, the AWC variability would enlarge, which in turn translates into a larger dispersion of Ω_i.

5. Reserve Loss Reduction via Funds Pooling

In practice, a manufacturing firm usually produces and sells multiple products at the same time, and the firm is thus responsible to maintain enough reserves to fulfill warranty obligations for these products. Because warranty expenditures are random and both over reserving and under reserving risks are involved in warranty reserve planning, it is reasonable to expect, just like with physical inventories, that pooling warranty reserves for multiple products could reduce reserve levels and reserve losses. In particular, given the nature of reserves—cash that can be freely used for any product sold by the firm and easily added to or withdrawn from a fund—pooling is especially simple and attractive. However, it is surprising that a considerable number of firms have not yet pooled their warranty reserves.[8] In this section, we aim to investigate the warranty reserve pooling scheme to quantify the corresponding (time-varying) pooling benefits, which is able to reveal the value of funds pooling.

5.1. Funds Pooling Scheme

We consider a manufacturing firm that manages the warranties for M ($M > 1$) different products. The following assumption is then introduced.

Assumption 4. *The M products are released onto the market at the same time and share the same warranty period W and sales period L.*

Although this assumption is adopted mainly for the sake of simplicity and tractability, it is not uncommon in reality. For example, Apple often releases products (e.g., iPhone, iPod, iPad, Mac) of different generations simultaneously. The warranty periods of these products are typically one year, and their sales periods are roughly the same.

Hereafter, we use the following notation to facilitate our presentation. Subscript j is attached to each parameter to indicate the jth product ($j \in \{1, 2, \ldots, M\}$). For instance, we use λ_j, θ_j, and C_{Wj} to denote the sales rate, failure rate, and replacement cost of the jth product, respectively. In this way, the optimal warranty reserve level and the corresponding expected reserve loss, for the jth product in the kth planning period, become $WR^*_{j,k} = \mu_{j,k} + \frac{1}{2}A\sigma_{j,k}$ and $\hat{\pi}^*_{j,k}(WR^*_{j,k}) = \frac{1}{2}B\sigma_{j,k}$, respectively (according to Equations (9) and (10)). Hence, without pooling, the total amount of warranty reserves planned for the M products in the kth period can be summed up as

$$TWR^*_k = \sum_{j=1}^{M} WR^*_{j,k} = \sum_{j=1}^{M} \mu_{j,k} + \frac{A}{2}\sum_{j=1}^{M} \sigma_{j,k}. \tag{14}$$

Conversely, the firm can run a combined fund for the M products to reduce reserve levels and pool risks. In the pooled case, the distributionally robust optimal warranty reserve level PWR^*_k in the kth period can be determined by minimizing

$$\begin{aligned}\pi^c_k(PWR_k) &= \mathbb{E}\left[C_h\left(PWR_k - \sum_{j=1}^{M} WC_{j,k}\right)^+ \right.\\ &\qquad \left. + C_b\left(\sum_{j=1}^{M} WC_{j,k} - PWR_k\right)^+\right]\\ &= C_h PWR_k - C_h\sum_{j=1}^{M}\mu_{j,k}\\ &\qquad + (C_h + C_b)\mathbb{E}\left[\sum_{j=1}^{M} WC_{j,k} - PWR_k\right]^+. \end{aligned} \tag{15}$$

Based on the result in Lemma 2, we can obtain the following proposition.

Proposition 5. *The distributionally robust optimal level of combined warranty reserves for the kth planning period is given by*

$$PWR^*_k = \sum_{j=1}^{M} \mu_{j,k} + \frac{A}{2}\sqrt{\sum_{j=1}^{M} \sigma^2_{j,k}}. \tag{16}$$

The corresponding expected total reserve losses for the pooled case can be obtained as

$$\hat{\pi}^{c*}_k(PWR^*_k) = \frac{B}{2}\sqrt{\sum_{j=1}^{M} \sigma^2_{j,k}}. \tag{17}$$

5.2. Benefits of Funds Pooling

We investigate the benefits of funds pooling via four efficiency measures. In particular, we are interested in the time-varying properties of pooling benefits over the entire warranty life cycle. The first two measures are related to the levels of warranty reserves, that is, absolute benefit Υ_k^* and relative benefit Ψ_k^* ($k \in \{1, 2, \ldots, n+m\}$), which can be expressed as

$$\Upsilon_k^* = TWR_k^* - PWR_k^* = \frac{A}{2}\left(\sum_{j=1}^{M}\sigma_{j,k} - \sqrt{\sum_{j=1}^{M}\sigma_{j,k}^2}\right), \tag{18}$$

and

$$\Psi_k^* = \frac{TWR_k^* - PWR_k^*}{TWR_k^*} = \frac{\frac{A}{2}\left(\sum_{j=1}^{M}\sigma_{j,k} - \sqrt{\sum_{j=1}^{M}\sigma_{j,k}^2}\right)}{\sum_{j=1}^{M}\mu_{j,k} + \frac{A}{2}\sum_{j=1}^{M}\sigma_{j,k}}. \tag{19}$$

We also examine the absolute benefit Υ_k^o and relative benefit Ψ_k^o in terms of expected total reserve losses:

$$\begin{aligned}\Upsilon_k^o &= \sum_{j=1}^{M}\hat{\pi}_{j,k}^*(WR_{j,k}^*) - \hat{\pi}_k^{c*}(PWR_k^*)\\ &= \frac{B}{2}\left(\sum_{j=1}^{M}\sigma_{j,k} - \sqrt{\sum_{j=1}^{M}\sigma_{j,k}^2}\right),\end{aligned} \tag{20}$$

and

$$\begin{aligned}\Psi_k^o &= \frac{\sum_{j=1}^{M}\hat{\pi}_{j,k}^*(WR_{j,k}^*) - \hat{\pi}_k^{c*}(PWR_k^*)}{\sum_{j=1}^{M}\hat{\pi}_{j,k}^*(WR_{j,k}^*)}\\ &= 1 - \frac{\sqrt{\sum_{j=1}^{M}\sigma_{j,k}^2}}{\sum_{j=1}^{M}\sigma_{j,k}}.\end{aligned} \tag{21}$$

It is obvious that all the four efficiency measures are positive, even though the products are released at different dates (i.e., Assumption 4 does not hold). This implies that funds pooling is indeed beneficial. In terms of time-varying behaviors, it is straightforward to verify that absolute benefits Υ_k^* and Υ_k^o exhibit an upside-down bathtub shape (refer to Section EC.2.1 in the online supplement for details). In practice, the relative benefits are of more interest to manufacturing firms. In what follows, we focus only on the time-varying properties of relative benefits Ψ_k^* and Ψ_k^o.

Define $\sigma_{\max,k} = \max_{j=1,\ldots,M}\{\sigma_{j,k}\}$ and $\sigma_{\min,k} = \min_{j=1,\ldots,M}\{\sigma_{j,k}\}$. The following lemma examines the monotonicity of relative benefit Ψ_k^* with respect to the mean and standard deviation of the AWC increment.

Lemma 3. *The relative benefit Ψ_k^* decreases as $\mu_{j,k}$ increases and as $\sigma_{\min,k}$ decreases.*

Lemma 3 presents an interesting finding that somewhat explains the relationship between the pooling benefit Ψ_k^* and the range of standard deviations of different products. If we decrease the value of $\sigma_{\min,k}$ while keeping other standard deviations unchanged—that is, the left end-point of interval $[\sigma_{\min,k}, \sigma_{\max,k}]$ becomes smaller (the range is partially enlarged)—then the relative benefit Ψ_k^* becomes smaller. It is difficult to obtain the general monotonicity of Ψ_k^* with respect to k, because of its complex formulation. Nevertheless, when the AWC increments of all products share the same variance, we have the following result.

Corollary 1. *If the AWC increments of the M products share the same variance, then Ψ_k^* would follow the same time-varying monotonicity of $CV_{j,k}$ ($j \in \{1, 2, \ldots, M\}$) over the warranty life cycle.*

Corollary 1 implies that Ψ_k^* would present a bathtub shape that is consistent with the pattern of coefficients of variation (Figures 9(a) and 10(a)). This supports the use of coefficient of variation to describe AWC variability, as discussed earlier. It is worth pointing out that when the variances of AWC increments for different products are identical, their means are not necessarily the same.[9]

Next, we investigate the time-varying property of relative benefit Ψ_k^o in terms of reserve losses. Define $\delta_{\max,k} := \sigma_{\max,k}/\sigma_{\min,k}$ as the "relative range" of the standard deviations. Then, we have the following key result.

Proposition 6. *The relative benefit Ψ_k^o decreases as the ratio $\delta_{\max,k}$ increases.*

Proposition 6 shows that when all other standard deviations (i.e., $\sigma_{j,k} \neq \sigma_{\min,k}, \sigma_{\max,k}$) are fixed, enlarging the relative range of AWC standard deviations will weaken the pooling benefit. This phenomenon can be explained by the so-called *resonance* effect. That is, when the magnitudes of AWC variabilities of different products become larger, the manufacturer must allocate more reserves to those with higher variabilities, which clearly reduces the relative benefit of funds pooling. This result shows that the relative range $\delta_{\max,k}$ plays a critical role in comparing relative benefits Ψ_k^o in different planning periods. By defining $\Gamma_k^{(j,l)} := \sigma_{j,k}/\sigma_{l,k}$, $\forall j \neq l$, the following lemma is useful for investigating the time-varying behavior of the pooling benefit.

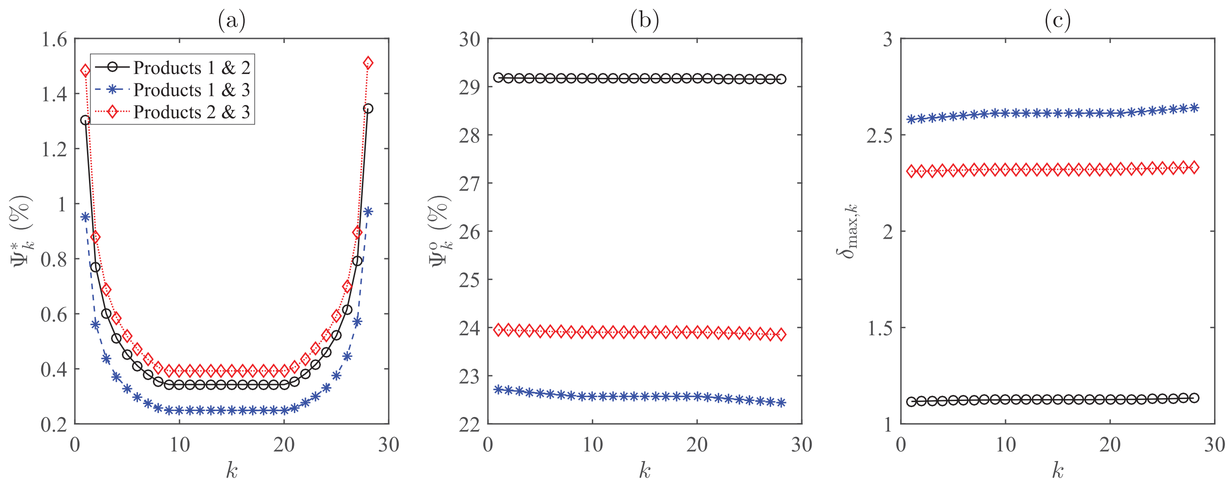
Lemma 4. *If $\Gamma_1^{(j,l)} > 1$ and $\theta_j > \theta_l$ hold in the first planning period, then (1) $\Gamma_k^{(j,l)} > 1$ holds for $k \in \{2, 3, \ldots, n+m\}$, and (2) $\Gamma_k^{(j,l)}$ is a nondecreasing function of k.*

Combining Proposition 6 and Lemma 4 (notice that $\delta_{\max,k} = \Gamma_k^{(\max,\min)}$), the time-varying characteristic of Ψ_k^o can be captured by the following corollary.

Corollary 2. *Define $\theta_{\max}$ and $\theta_{\min}$ as the failure rates associated with $\sigma_{\max,1}$ and $\sigma_{\min,1}$, respectively. For $M = 2$, if $\theta_{\max} > \theta_{\min}$, then Ψ_k^o is a nonincreasing function of $k \in \{1, 2, \ldots, n+m\}$.*

This finding is quite insightful. Although the relative benefit Ψ_k^o is also varying over time, its tendency

Figure 9. (Color online) Relative Benefits Ψ_k^* and Ψ_k^o with Respect to k When Pooling Two of the Three Products

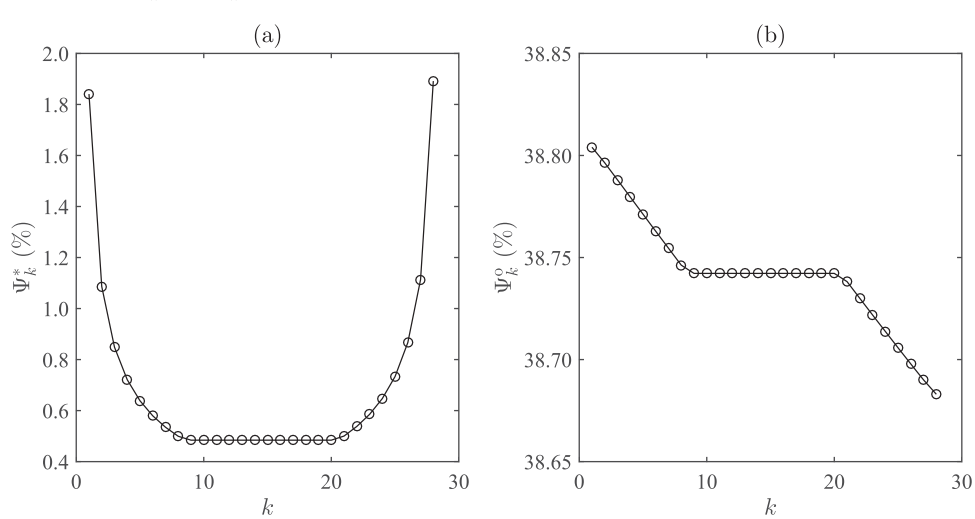


differs from that of Ψ_k^*. Figure 9 shows the relative benefits Ψ_k^o when pooling any two of the three products, with the parameter settings in Figures 3 and 4. It can be observed that Ψ_k^o shows an opposite tendency to $\delta_{\max,k}$. Moreover, in a specific period k, a larger (respectively, smaller) $\delta_{\max,k}$ corresponds to a smaller (respectively, larger) Ψ_k^o. For example, the relative benefit Ψ_k^* of pooling warranty reserves of products 1 and 2 (black circle in Figure 9) is the largest one among the three cases, because the associated ratio $\delta_{\max,k}$ is the smallest. These observations are consistent with the analytical results in Proposition 6 and Corollary 2.

It is important to note that Corollary 2 strictly holds only for $M = 2$; that is, there are only two products. This is because only when $M = 2$, Ψ_k^o decreases exclusively with $\delta_{\max,k}$, can the chain rule be applied to obtain Corollary 2. For $M > 2$, it is difficult, if not impossible, to obtain the analytical relationship between Ψ_k^o and k. Nevertheless, Figure 10 illustrates relative benefits Ψ_k^* and Ψ_k^o when the warranty reserves of all three products are pooled. It shows that Ψ_k^* has a bathtub shape and Ψ_k^o exhibits a nonincreasing pattern over the entire warranty life cycle. Moreover, it is noteworthy that when pooling three products, both Ψ_k^* and Ψ_k^o are higher than those in the case of pooling any two products. This demonstrates that pooling benefits would be more significant when the warranty reserves of more products are pooled.

Another interesting observation from Figures 9 and 10 is that the relative benefit Ψ_k^o varies within a small range. In particular, Ψ_k^o is almost constant in some cases. This is because when sales rate λ and

Figure 10. Relative Benefits Ψ_k^* and Ψ_k^o with Respect to k When Pooling the Three Products



replacement cost C_W are large and failure rate θ is small (which is quite common for electronics), we know from Equation (4) that the ratio $\Gamma_k^{(j,l)} = \sigma_{j,k}/\sigma_{l,k}$ between any two products j and l would be almost constant over time, which in turn leads to an almost constant Ψ_k^{o}.

The previous discussions of time-varying pooling benefits rely heavily on the simultaneous release assumption (i.e., Assumption 4). In practice, however, manufacturing firms often introduce their products in a one-by-one or batch-by-batch manner. We thus extend our discussion to the case of periodic release (which is common for consumer electronics); please refer to Section EC.2.2 in the online supplement. In addition, we provide a brief discussion in Section EC.2.3 of the online supplement on optimal release scheduling from a reserve management perspective.

6. Case Study with Field Data

In this section, we demonstrate the performance of the proposed reserve planning model and the reserve-loss-reduction strategies through a real warranty data set provided by the aforementioned electronics manufacturer. The firm manufactures various kinds of electronic products and components. We select the sales data and associated warranty claims data of three products as examples (data are scaled for confidentiality). The warranty periods of the three products are identical ($W = 12$ months), whereas their release dates and sales periods are different (Table 1). Monthly sales data are plotted in Figure 11(a), which shows that the sales volumes fluctuate over time but do not exhibit obvious patterns. We apply the Augmented Dickey-Fuller test[10] to examine the sales data and find that we are at least 70% confident that the data are stationary. Thus, it is "safe" to adopt the HPP sales assumption in our problem. In particular, the monthly sales rates are estimated as 61,316, 48,187, and 59,103, respectively. Moreover, the associated monthly warranty claims are shown in Figure 1. In the case study, warranty period W and sales period L are directly obtained from the firm and sales rate λ is fitted from the sales data, whereas failure rate θ is the only parameter that must be estimated from the claims data. For this purpose, we adopt the following least squares method:

$$\widehat{\theta} = \arg\min \sum_{k=1}^{n+m} (\mu_k - \Delta Q_k)^2,$$

where μ_k is given by Equation (3) and ΔQ_k is observed warranty expenses over month k. By solving this optimization problem, the estimated monthly failure rates are summarized in Table 1. Figure 11(b) further plots actual and estimated monthly warranty expenses. One can observe that, although the stylized AWC model can roughly capture real warranty claim patterns, the curve-fitting performance presents certain deficiency; that is, a systematic error exists. More specifically, the fitted curves tend to overestimate the actual data at early and late stages while underestimating the actual data at the middle stage.

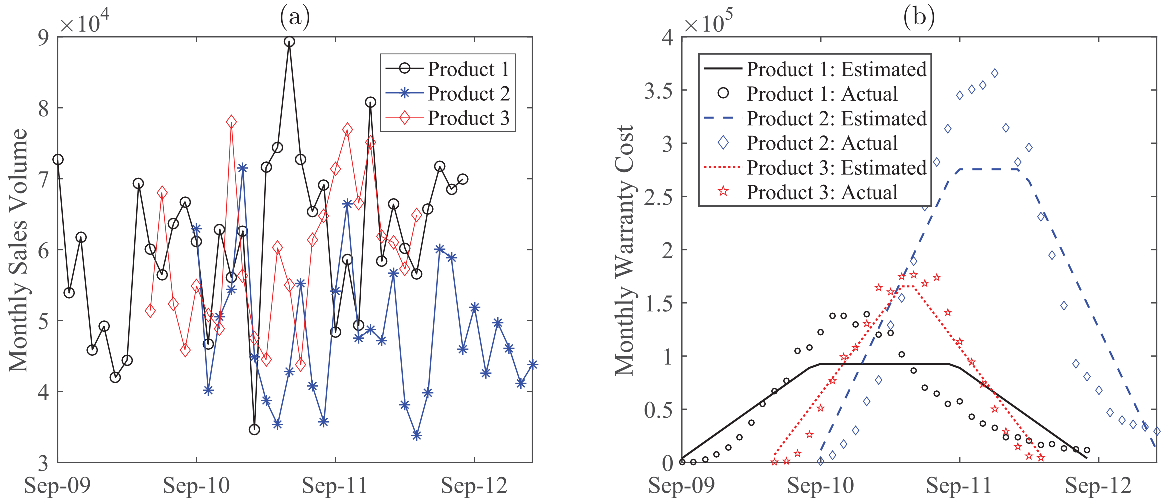
Based on the product sales, reliability, and cost information in Table 1, the mean $\mu_{j,k}$ and variance $\sigma_{j,k}^2$ of the AWC increment for each product j in each planning month k can be computed by Equations (3) and (4), respectively. Then, the distributionally robust optimal reserve level $WR_{j,k}^*$ is determined by Equation (9). Finally, the corresponding actual warranty reserve loss can be calculated through $\pi_{j,k}(WR_{j,k}^*) = C_h(WR_{j,k}^* - \Delta Q_{j,k})^+ + C_b(\Delta Q_{j,k} - WR_{j,k}^*)^+$. Figure 12 shows actual warranty costs and planned reserve levels for the three products. It can be observed that for each product, the planned reserves are excessive at first, insufficient in the middle, and over-stocked again at the last, which results from the aforementioned systematic error in curve fitting.

The unsatisfactory performance of the model-driven reserving policy necessitates demand learning. In this scenario, the adjusted warranty reserve levels $\widehat{WR}_{j,k}$ for the three products are dynamically updated by Equation (11), which are depicted in Figure 12. We can observe that the reserve levels with learning are much closer to actual warranty costs than those without learning. Moreover, funds pooling is also applied to improve warranty reserve planning performance. By pooling warranty reserves of the three products (with different release dates and sales periods), the combined reserve levels and corresponding reserve losses are evaluated by Equations (16) and (17), respectively (see Figure 12(d) for pooled reserve levels). Not surprisingly, the planned reserve level in the pooled case is closer to actual combined warranty expenses due to the risk pooling effect. One thing noteworthy is that it is feasible to apply demand learning to the pooled case, but the funds adjusted by learning cannot be used to evaluate pooling benefits.

Table 1. Product Information Used in the Case Study

Product	Sales period	W	L	$\widehat{\lambda}$	$\widehat{\theta}$	C_W	C_h	C_b
1	September 2009 to August 2011	12	24	61,316	0.00126	\$100	\$0.01	\$0.025
2	September 2010 to February 2012	12	18	48,187	0.00794	\$60	\$0.01	\$0.025
3	May 2010 to April 2011	12	12	59,103	0.00541	\$45	\$0.01	\$0.025

Figure 11. (Color online) Monthly Sales Volumes (a) and Aggregate Warranty Costs (b) of the Three Products



To compare the scenarios with and without learning, Table 2 presents total warranty reserve levels and total reserve losses for each product over its entire warranty life cycle. Compared with the no-learning scenario, the total reserve levels and total reserve losses in the learning scenario are indeed lower. In particular, the reductions in total reserve losses are significant in terms of percentage (27.64%–62.76%). This shows that adopting the demand learning mechanism in warranty reserve planning can save a large proportion of money for the firm, either in the pooled or nonpooled case. Moreover, when examining the pooling benefit in the learning scenario, we find that the pooled scheme results in even a slightly higher reserve level (increased by $48,979, 0.55%) than the nonpooled case. However, the corresponding total reserve losses are reduced significantly (decreased by $8496, 47.25%). Nevertheless, a higher total reserve level in the pooled case is not surprising. This is because, in the nonpooled cases, the reserve levels in some periods are insufficient, which results in large shortage expenses. In contrast, the pooled case puts more reserves in these periods to mitigate the stockout effect. This further demonstrates the effectiveness of demand learning and funds pooling. Furthermore, the magnitudes of reductions in total reserve losses due to learning and pooling are generally higher than those of total reserve levels, which implies that the two reserve-loss-reduction approaches have more significant impacts on reserve losses than reserve levels.

7. Conclusions and Implications

The goal of this paper is to investigate a warranty reserve planning problem during the entire warranty life cycle from a manufacturing firm's perspective. For this purpose, we first develop an adaptive AWC model by fully characterizing the mechanism that governs the generation of warranty claims, which allows the firm to plan for warranty reserves periodically.

Figure 12. (Color online) Actual Monthly Warranty Costs and Planned Reserve Levels (Without and with Learning) for the Three Products and the Pooled Case

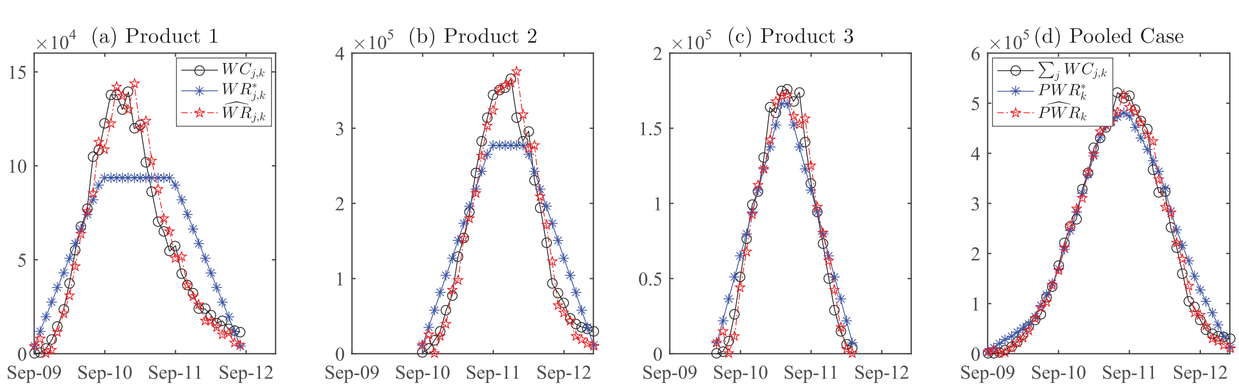


Table 2. Reserve Levels and Reserve Losses over the Warranty Life Cycle in Learning and No-Learning Scenarios

	Total reserve levels ($)			Pooled vs. nonpooled		
	Product 1	Product 2	Product 3	Pooled case	Reduction	Percentage of reduction
No-learning case	2,265,826	5,003,016	2,093,834	9,327,879	34,779	0.37%
Learning case	2,097,786	4,787,992	2,029,187	8,963,944	−48,979	−0.55%
Reduction	168,039	215,024	64,647	363,952		
Percentage of reduction	7.42%	4.30%	3.09%	3.90%		
	Total reserve losses ($)			Pooled vs. nonpooled		
	Product 1	Product 2	Product 3	Pooled case	Reduction	Percentage of reduction
No-learning case	12,506	19,569	5,258	14,629	22,704	60.82%
Learning case	4,657	9,520	3,805	9,485	8,496	47.25%
Reduction	7,849	10,049	1,453	5,144		
Percentage of reduction	62.76%	51.35%	27.64%	35.16%		

Because the exact distribution of AWC increments is generally unknown, the optimal warranty reserve level in each planning period is determined via a distributionally robust approach. On this basis, two reserve-loss-reduction strategies—demand learning and funds pooling—are further discussed. The demand learning algorithm can dynamically update the reserving plan by combining real warranty data, whereas the funds pooling scheme combines reserve funds of multiple products into a single fund to mitigate risks. Analytical results and simulation experiments show that demand learning and funds pooling are indeed beneficial for the firm in terms of reducing total reserve losses.

The main conclusions and implications of this work are summarized as follows:

(i). For the exponential-smoothing-type demand learning algorithm, an important finding is that when the sales period is long enough, the learning parameter will converge to a constant in probability, with a convergence rate of $1/\sqrt{k}$. This not only guarantees a stable output of the algorithm, but also simplifies the algorithm implementation. In addition, simulation experiments show that by incorporating the objective function in the learning process, the algorithm can significantly reduce total reserve losses under general sales and failure patterns, for which the proposed AWC model fails to capture the dynamics of warranty costs. We believe that this simple yet effective "predict-then-optimize" approach would be also useful for other operations management applications.

(ii). For the funds pooling scheme, we find that the pooling benefits change over different stages of the warranty life cycle, due to the dynamics of warranty expenses. More importantly, because reserve loss reduction is the primary objective of funds pooling, we prove an important result that the relative pooling benefit in terms of reserve losses exhibits a nonincreasing pattern (note that this result analytically holds only when there are two products). Another compelling insight is that the relative pooling benefit in terms of reserve losses decreases as the relative range of demand standard deviations enlarges, meaning that it is more attractive to pool warranty reserves for products whose AWC uncertainties are of similar magnitudes. This insight is quite general in the sense that it is applicable to traditional risk pooling problems in operations management.

(iii). Finally, our simulation experiments provide guidelines on the implementation of the proposed methodologies. Specifically, when the product sales and failure processes governing warranty claims generation obey those described in Assumptions 1 and 2, then implementing the distributionally robust reserving model, combined with funds pooling, can deliver a fairly good reserving plan; if this is not the case, then the demand learning mechanism can be further used to compensate the model's inefficiency by leveraging real claims data.

However, this paper presents several limitations that deserve future research efforts. First, because of the simplified assumptions on product sales and failure processes, the proposed AWC model is somewhat stylized, although it is still capable to roughly capture any inverted bathtub–like warranty claim patterns. Because deriving the mean and variance of warranty expenses in general settings is highly intractable, a promising future research topic is to develop an approximate, closed-form expression of the mean and variance to facilitate warranty reserve planning, without inducing much computational burden. Moreover, an implicit assumption adopted here is that the sales of different products are not correlated. In reality, however, products could be substitutes or complements. It is interesting to investigate the impact of such correlations on warranty reserve planning. One possible way is to apply diffusion-choice demand models (Li 2020) that integrate the Bass diffusion model and discrete choice models. Furthermore, only a specific form of warranty policy, that is, free replacement warranty, is considered in this work. Future research could incorporate more warranty policies, for example, pro rata and combination warranties, into the reserve planning framework.

Acknowledgments

The authors thank the department editor, associate editor, and two anonymous referees for constructive comments and suggestions that improved the paper significantly. The authors also thank Chung-Piaw Teo and Hanlin Liu for valuable discussions on an original version of this paper.

Endnotes

[1] Apple's warranties and service contracts: http://www.warrantyweek.com/archive/ww20171109.html.

[2] Warranty reserves: http://www.warrantyweek.com/archive/ww20060509.html.

[3] Warranty adjustments: http://www.warrantyweek.com/archive/ww20151105.html.

[4] One should distinguish *product life cycle* from *warranty life cycle*. The former starts from the time a product is first introduced to the market and extends the instant it is discontinued from the market, while the latter is the period from the beginning of a warranty program to the end of the program (when the warranties for final sales expire). In general, let L denote product life cycle and W denote warranty period; then the warranty life cycle should be $L+W$.

[5] Once all units of a product are sold to end consumers, the *installed base* (i.e., the total number of sold units that are currently in use) will remain almost constant for a relatively long period of time. As a result, the demand of parts (resulting from repair/replacement) gradually turns to be stationary, as the units have statistically identical quality and reliability.

[6] Our simulation studies reveal that the approximation accuracy is relatively insensitive to λ but exhibits an obvious increasing trend as θ decreases. When failure rate θ is small enough (say, $\theta < 0.1$), the approximation error vanishes.

[7] Suppose that the warranty reserve fund accrues interest at a constant rate ζ_0, the opportunistic investment return rate is ζ_1, and the interest rate for obtaining emergency funds is ζ_2 (in general, $\zeta_0 < \zeta_1 < \zeta_2$). Then, C_h and C_b can be roughly estimated by $C_h = \zeta_1 - \zeta_0$ and $C_b = \zeta_2 - \zeta_0$, respectively. Nevertheless, what matters to warranty reserve planning, as can be seen from Equation (9), is not the absolute values of C_h and C_b, but their relative value C_b/C_h or C_h/C_b.

[8] Based on our consultation with some firms, including a leading engine manufacturer, an electronics manufacturer, and a credit card company, all of them do not implement warranty reserve pooling for their products.

[9] For example, based on Equations (3) and (4), given two products j and l, combining the conditions $\theta_j = \theta_l$, $C_{Wj}^2\lambda_j = C_{Wl}^2\lambda_l$, and $C_{Wj}\lambda_j \neq C_{Wl}\lambda_l$ can guarantee that $\sigma_{j,k} = \sigma_{l,k}$ and $\mu_{j,k} \neq \mu_{l,k}$ for any $k \in \{1, 2, \ldots, n+m\}$. The following parameter setting satisfies the previous three conditions: $\theta_j = \theta_l = 0.05$, $C_{Wj} = 30$, $C_{Wl} = \sqrt{150}$, $\lambda_j = 500$, and $\lambda_l = 3{,}000$.

[10] The augmented Dickey-Fuller test is a popular method that assesses the presence of a unit root in an autoregressive model $y_t = \rho y_{t-1} + \sum_{j=1}^{p} \upsilon_j \Delta y_{t-j} + \epsilon_t$, where $\Delta y_t = y_t - y_{t-1}$, p is the lag order, and ϵ_t is a white noise. In this model, testing for a unit root is equivalent to testing $\rho = 1$, which implies that the time series y_t is nonstationary. The test statistic is $(\widehat{\rho} - 1)/\widehat{Std}(\widehat{\rho})$, where $\widehat{\rho}$ is the ordinary least squares estimator of ρ and $\widehat{Std}(\widehat{\rho})$ represents an estimator of the standard deviation of $\widehat{\rho}$. If the null hypothesis $H_0: \rho = 1$ is rejected at the significance level α, then we are $100(1-\alpha)\%$ confident that the time series sample is stationary. For more details, please refer to Fuller (1996, chapter 10).

References

Bass FM (1969) A new product growth for model consumer durables. *Management Sci.* 15(5):215–227.

Baumol WJ (1952) The transactions demand for cash: An inventory theoretic approach. *Quart. J. Econom.* 66(4):545–556.

Benjaafar S, Cooper WL, Kim JS (2005) On the benefits of pooling in production-inventory systems. *Management Sci.* 51(4):548–565.

Berman O, Krass D, Mahdi Tajbakhsh M (2011) On the benefits of risk pooling in inventory management. *Production Oper. Management* 20(1):57–71.

Besbes O, Mouchtaki O (2021) How big should your data really be? Data-driven newsvendor and the transient of learning. Working paper, Columbia Business School, New York.

Bimpikis K, Markakis MG (2016) Inventory pooling under heavy-tailed demand. *Management Sci.* 62(6):1800–1813.

Blischke WR, Murthy DNP (2000) *Reliability: Modeling, Prediction, and Optimization* (John Wiley & Sons, New York).

Bolton GE, Katok E (2008) Learning by doing in the newsvendor problem: A laboratory investigation of the role of experience and feedback. *Manufacturing Service Oper. Management* 10(3):519–538.

Bostian AJA, Holt CA, Smith AM (2008) Newsvendor "pull-to-center" effect: Adaptive learning in a laboratory experiment. *Manufacturing Service Oper. Management* 10(4):590–608.

Buczkowski PS, Kulkarni VG (2006) Funding a warranty reserve with contributions after each sale. *Probab. Engrg. Inform. Sci.* 20(3):497–515.

Calmon AP, Graves SC (2017) Inventory management in a consumer electronics closed-loop supply chain. *Manufacturing Service Oper. Management* 19(4):568–585.

Calmon AP, Graves SC, Lemmens S (2021) Warranty matching in a consumer electronics closed-loop supply chain. *Manufacturing Service Oper. Management* 23(5):1314–1331.

Chen B, Chao X (2019) Parametric demand learning with limited price explorations in a backlog stochastic inventory system. *IISE Trans.* 51(6):605–613.

Chen X, Simchi-Levi D (2009) A new approach for the stochastic cash balance problem with fixed costs. *Probab. Engrg. Inform. Sci.* 23(4):545–562.

Cohen D, Darrough MN, Huang R, Zach T (2011) Warranty reserve: Contingent liability, information signal, or earnings management tool? *Accounting Rev.* 86(2):569–604.

Eliashberg J, Singpurwalla ND, Wilson SP (1997) Calculating the reserve for a time and usage indexed warranty. *Management Sci.* 43(7):966–975.

Eppen GD (1979) Note–Effects of centralization on expected costs in a multi-location newsboy problem. *Management Sci.* 25(5):498–501.

Eppen GD, Fama EF (1969) Cash balance and simple dynamic portfolio problems with proportional costs. *Internat. Econom. Rev.* 10(2):119–133.

Fuller WA (1996) *Introduction to Statistical Time Series*, 2nd ed. (John Wiley & Sons, New York).

Gallego G, Moon I (1993) The distribution free newsboy problem: Review and extensions. *J. Oper. Res. Soc.* 44(8):825–834.

Gallego G, Wang R, Hu M, Ward J, Beltran JL (2015) No claim? Your gain: Design of residual value extended warranties under risk aversion and strategic claim behavior. *Manufacturing Service Oper. Management* 17(1):87–100.

Gormley FM, Meade N (2007) The utility of cash flow forecasts in the management of corporate cash balances. *Eur. J. Oper. Res.* 182(2):923–935.

Govindarajan A, Sinha A, Uichanco J (2021) Distribution-free inventory risk pooling in a multilocation newsvendor. *Management Sci.* 67(4):2272–2291.

Gurgur CZ (2011) Dynamic cash management of warranty reserves. *Engrg. Econom.* 56(1):1–27.

Huber J, Müller S, Fleischmann M, Stuckenschmidt H (2019) A data-driven newsvendor problem: From data to decision. *Eur. J. Oper. Res.* 278(3):904–915.

Ja SS, Kulkarni VG, Mitra A, Patankar JG (2002) Warranty reserves for nonstationary sales processes. *Naval Res. Logist.* 49(5):499–513.

Karsten F, Slikker M, Van Houtum GJ (2015) Resource pooling and cost allocation among independent service providers. *Oper. Res.* 63(2):476–488.

Keskin NB, Min X, Song JSJ (2021) The nonstationary newsvendor: Data-driven nonparametric learning. Working paper, Fuqua School of Business, Duke University, Durham, NC.

Khawam J, Hausman WH, Cheng DW (2007) Warranty inventory optimization for Hitachi Global Storage Technologies, Inc. *Interfaces* 37(5):455–471.

Li H (2020) Optimal pricing under diffusion-choice models. *Oper. Res.* 68(1):115–133.

Menke WW (1969) Determination of warranty reserves. *Management Sci.* 15(10):542.

Miller MH, Orr D (1966) A model of the demand for money by firms. *Quart. J. Econom.* 80(3):413–435.

Mišić VV, Perakis G (2020) Data analytics in operations management: A review. *Manufacturing Service Oper. Management* 22(1): 158–169.

Murthy DNP, Solem O, Roren T (2004) Product warranty logistics: Issues and challenges. *Eur. J. Oper. Res.* 156(1):110–126.

Neave EH (1970) The stochastic cash balance problem with fixed costs for increases and decreases. *Management Sci.* 16(7):472–490.

Patankar JG, Worm GH (1981) Note: Prediction intervals for warranty reserves and cash flows. *Management Sci.* 27(2):237–241.

Pinçe Ç, Ferguson M, Toktay B (2016) Extracting maximum value from consumer returns: Allocating between remarketing and refurbishing for warranty claims. *Manufacturing Service Oper. Management* 18(4):475–492.

Scarf H (1958) A min max solution of an inventory problem. Arrow K, Karlin S, Scarf H, eds. *Studies in the Mathematical Theory of Inventory and Production* (Stanford University Press, Stanford, CA), 201–209.

Tagaras G, Cohen MA (1992) Pooling in two-location inventory systems with non-negligible replenishment lead times. *Management Sci.* 38(8):1067–1083.

Tapiero CS, Posner MJ (1988) Warranty reserving. *Naval Res. Logist.* 35(4):473–479.

Thomas MU (1989) Note: A prediction model for manufacturer warranty reserves. *Management Sci.* 35(12):1515–1519.

Xie W, Ye ZS (2016) Aggregate discounted warranty cost forecast for a new product considering stochastic sales. *IEEE Trans. Reliability* 65(1):486–497.

Yang C, Hu Z, Zhou SX (2021) Multilocation newsvendor problem: Centralization and inventory pooling. *Management Sci.* 67(1): 185–200.

Zhang H, Chao X, Shi C (2020) Closing the gap: A learning algorithm for lost-sales inventory systems with lead times. *Management Sci.* 66(5):1962–1980.

Zhong Y, Zheng Z, Chou MC, Teo CP (2018) Resource pooling and allocation policies to deliver differentiated service. *Management Sci.* 64(4):1555–1573.

Online Supplement to "Warranty Reserve Management: Demand Learning and Funds Pooling"

EC.1. Worst-Case Performance Assessment

We assess the performance of the distributionally robust reserving model in Section 3.2 by comparing it with a benchmark—the simulated nonparametric model. In particular, we implement Monte Carlo simulation to generate an empirical distribution of the AWC increment in each period (based on Assumptions 1-3), which can be regarded as the "true distribution" and serves as a perfect benchmark for worst-case performance assessment.

The inputs of simulation include sales rate λ, failure rate θ, warranty period W, sales period L, planning cycle Δ, and the associated cost parameters (C_W, C_b, C_h). For each simulation run, we first generate all the purchase dates within $[0, L]$ according to $HPP(\lambda)$. For each purchase date, we then generate the associated under-warranty failure instant(s) based on the exponential distribution $G(t;\theta)$. Finally, counting the total number of warranty claims and the associated costs within each planning period yields the simulated AWC increments over the warranty life cycle.

Let S be the total number of simulation runs and $\Delta Q_{i,k}$ ($i \in \{1, 2, \ldots, S\}$, $k \in \{1, 2, \ldots, m+n\}$) denote the simulated AWC increment in simulation run i for planning period k. It is clear that when S is large enough, the mean and variance of $\{\Delta Q_{i,k}\}_{i\in\{1,2,\ldots,S\}}$ will converge to μ_k and σ_k^2 in Theorem 1, respectively. With a large number of simulated samples, we are able to determine the optimal warranty reserve level through a nonparametric newsvendor model. Suppose the distribution of AWC increment WC_k in the kth period is $F_k(\cdot)$, it is well known that the optimal warranty reserve level can be obtained as

$$WR_k^{\dagger} = F_k^{-1}(C_b/(C_b + C_h)). \tag{EC.1}$$

Therefore, we can leverage the simulated samples to form an empirical distribution $F_k(\cdot)$ and calculate the nonparametric optimal warranty reserve level $WR_k^{\dagger}$ based on the *critical fractile* $C_b/(C_b + C_h)$, which provides a promising benchmark for assessing the worst-case performance.

To compare the distributionally robust and nonparametric reserving models, we employ the following empirical newsvendor model to compute the *average* reserve loss $\tilde{\pi}_k(\cdot)$ in each period k:

$$\tilde{\pi}_k(WR_k) = \frac{1}{S}\sum_{i=1}^{S}\Big[C_h(WR_k - \Delta Q_{i,k})^+ + C_b(\Delta Q_{i,k} - WR_k)^+\Big]. \tag{EC.2}$$

By substituting WR_k^* in (9) and $WR_k^{\dagger}$ in (EC.1) into (EC.2), respectively, we can assess the performance of the worst-case scenario through the simulated samples. There is no doubt that the nonparametric model can result in a better decision that leads to a lower reserve loss. We thus

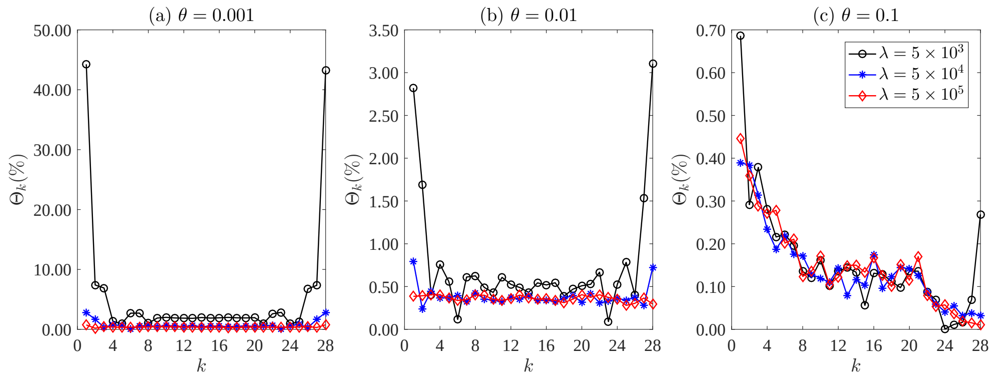


Figure EC.1 Percentage of Reserve Loss Growth from the Worst-Case Scenario.

define $\Theta_k = (\tilde{\pi}_k(WR_k^*) - \tilde{\pi}_k(WR_k^\dagger))/\tilde{\pi}_k(WR_k^\dagger)$ as the percentage of reserve loss growth from the worst-case scenario, in the kth planning period, to evaluate the associated performance. Accordingly, we further define $\Theta = (\sum_{k=1}^{m+n} \tilde{\pi}_k(WR_k^*) - \sum_{k=1}^{m+n} \tilde{\pi}_k(WR_k^\dagger))/\sum_{k=1}^{m+n} \tilde{\pi}_k(WR_k^\dagger)$ as the percentage of overall reserve loss growth from the worst-case scenario. When running the simulation experiment, we consider three sales-rate levels ($\lambda = 5,000$, $\lambda = 50,000$, $\lambda = 500,000$) and three failure-rate levels ($\theta = 0.001$, $\theta = 0.01$, $\theta = 0.1$). In total, we have nine sales-failure parameter combinations to cover a wide range of possible scenarios. Other parameters are set as $W = 2$ years, $L = 5$ years, $\Delta = 1$ quarter, $C_W = \$100$, $C_h = \$0.02$, and $C_b = \$0.05$. We generate 100,000 simulation runs for each parameter combination to guarantee a high simulation accuracy.

Figure EC.1 illustrates the percentages of reserve loss growth Θ_k ($\forall k$) over the warranty life cycle. We can see that Θ_k's tend to be large when λ and θ are small. This can be explained by the fact that when sales and failure rates are small, the resultant warranted base is small and the units therein are less vulnerable to failure, which in turn yields a small amount of warranty claims, yet a high degree of uncertainty. In this situation, the reserving plan derived from the worst-case scenario would be too conservative and thus leads to a large percentage of reserve loss growth. Nevertheless, the percentages of overall reserve loss growth Θ for all the nine combinations are quite small (even in the extreme case of $\lambda = 5,000$ and $\theta = 0.001$, $\Theta \approx 3.05\%$ is still acceptably small), despite the variability induced by simulation inaccuracy. This is largely due to the newsvendor structure of the problem in the sense that the objective function is flat around the optimum. This demonstrates that the distributionally robust reserving model performs fairly well under the proposed assumptions.

EC.2. Additional Results on Funds Pooling

EC.2.1. Absolute Benefits Υ_k^* and Υ_k°

Equations (18) and (20) show that absolute benefits Υ_k^* and Υ_k° are independent of $\mu_{j,k}$ ($\forall j$). It is interesting to observe that Υ_k^* and Υ_k° have the same structure, except for coefficients $A/2$ and $B/2$. The following lemma shows the impact of $\sigma_{j,k}$ on Υ_k^* and Υ_k°.

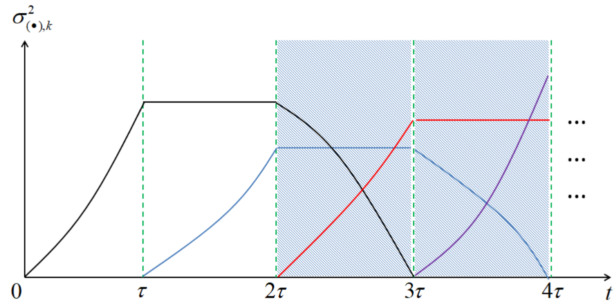


Figure EC.2 Variances of AWC Increments for Products with Periodic Release.

LEMMA EC.1. *Both Υ_k^* and Υ_k° increase as standard deviation $\sigma_{j,k}$ increases.*

This result is intuitive, because the absolute pooling benefits would become significant when demand uncertainty increases. Since the two absolute benefits rely only on the standard deviations of the demand, according to Proposition 1 and Lemma EC.1, one can apply the *chain rule* to obtain the following result.

COROLLARY EC.1. *Both Υ_k^* and Υ_k° increase in $k \in \{1,\dots,m\}$, remain constant in $k \in \{m+1,\dots,n\}$, and decrease in $k \in \{n+1,\dots,n+m\}$.*

Corollary EC.1 reveals the time-varying property of the absolute benefits over the warranty life cycle. Specifically, the curves of Υ_k^* and Υ_k° exhibit an upside-down bathtub shape. This time-variant characteristic stems from the monotonicity of standard deviation $\sigma_{j,k}$ (see Proposition 1).

EC.2.2. Extension: Funds Pooling for Products with Periodic Release

In Section 5, the discussions on funds pooling are limited to the case that all products are released simultaneously. In reality, however, this is not always the case. Here, we investigate a more realistic situation in which products are released periodically. Some consumer electronics, e.g., iPhone and ThinkPad X1 Carbon, are upgraded every year. In such cases, the inter-release time between two generations is typically one year, which is equal to the associated warranty period W. Let τ denote the inter-release time of two products. To simplify the discussion, we assume products are released year by year with the same life cycle $L = 2$ years, i.e., $\tau = W = \frac{1}{2}L = 1$ year. For instance, iPhone 6 and 6 Plus were launched on September 9, 2014 and discontinued on September 7, 2016.

In this situation, product release follows the pattern in Figure EC.2, which illustrates the trends of variances of warranty expenses for different products. It is easy to observe that starting from time point 2τ, funds pooling will always occur among three consecutively released products. Here, we only consider the relative pooling benefit Ψ_k° in terms of reserve losses, as it is the most important efficiency measure. If we order the products by their release sequences, i.e., $(1),(2),\dots,(M)$, then for the jth shadowed region $[j\tau,(j+1)\tau]$, $j \geq 2$, the relative pooling benefit Ψ_k° will only be affected by

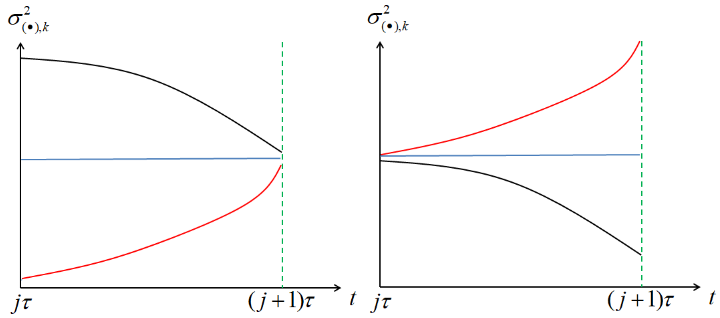


Figure EC.3 Variances of Three Products within an Overlapping Period.

the third variance segment $\sigma^2_{(j-1),2m+k}$ of product $(j-1)$ (a decreasing curve), the second variance segment $\sigma^2_{(j),m+k}$ of product (j) (a constant line), and the first variance segment $\sigma^2_{(j+1),k}$ of product $(j+1)$ (an increasing curve) for $1 \le k \le m$. Without loss of generality, we just need to focus on one of the time intervals $[j\tau,(j+1)\tau]$, $j \ge 2$, to analyze the relative pooling benefit Ψ°_k based on the following term

$$\Psi^{\circ}_k = 1 - \frac{\sqrt{\sigma^2_{(j-1),2m+k} + \sigma^2_{(j),m+k} + \sigma^2_{(j+1),k}}}{\sigma_{(j-1),2m+k} + \sigma_{(j),m+k} + \sigma_{(j+1),k}}, 1 \le k \le m. \tag{EC.3}$$

According to Equation (EC.3), the following proposition holds.

Proposition EC.1. *Let $a_{(j)} = C_{W(j)}\lambda_{(j)}\theta_{(j)}$. Under the periodic release assumption, one has (1) Ψ°_k monotonically increases in k when $\sigma_{(j-1),3m} \ge \sigma_{(j),2m} \ge \sigma_{(j+1),m}$ with parameter-wise sufficient conditions*

$$\begin{cases} \frac{a_{(j-1)}\theta_{(j-1)}(m-1/3)+\frac{1}{2}a_{(j-1)}}{a_{(j)}\theta_{(j)}m^2+a_{(j)}m} \ge 1, \\ \frac{a_{(j)}\theta_{(j)}m^2+a_{(j)}m}{a_{(j+1)}\theta_{(j+1)}(m^2-m+1/3)+a_{(j+1)}(m-1/2)} \ge 1; \end{cases} \tag{EC.4}$$

and (2) Ψ°_k monotonically decreases in k when $\sigma_{(j+1),1} \ge \sigma_{(j),m+1} \ge \sigma_{(j-1),2m+1}$ with parameter-wise sufficient conditions

$$\begin{cases} \frac{2a_{(j+1)}\theta_{(j+1)}+3a_{(j+1)}}{6(a_{(j)}\theta_{(j)}m^2+a_{(j)}m)} \ge 1, \\ \frac{a_{(j)}\theta_{(j)}m^2+a_{(j)}m}{a_{(j-1)}\theta_{(j-1)}(m^2-1/3)+a_{(j-1)}(m-1/2)} \ge 1. \end{cases} \tag{EC.5}$$

Figure EC.3 illustrates variance trends based on the results obtained in Proposition EC.1. One can observe that the range of variances decreases (resp. increases) as time goes by, when the corresponding sufficient condition in (EC.4) (resp. (EC.5)) holds. Hence, the relative pooling benefit Ψ°_k will be increasing (resp. decreasing) over time, which is consistent with the result in Proposition 6. According to Proposition EC.1, manufacturers can anticipate the trends of pooling benefits after estimating the parameters of three consecutively released products. In particular, when $\theta_{(j-1)} \gg \theta_{(j)} \gg \theta_{(j+1)}$—that is, product reliability is improved generation by generation—the relative pooling benefit Ψ°_k always shows an increasing trend within each interval $[j\tau,(j+1)\tau]$, $j \ge 2$. This result encourages manufacturers to put more investments in the research and development (R&D) process.

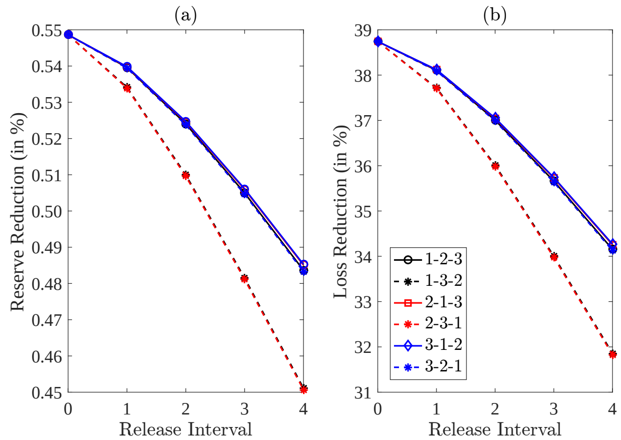


Figure EC.4 **Reductions of Reserve Levels and Reserve Losses versus Release Interval.**

EC.2.3. Product Release Scheduling from a Reserve Management Standpoint

Based on the above discussions on warranty reserve pooling, it is interesting to investigate the optimal product release strategy from a reserve management perspective. That is, how should a manufacturing firm determine its product release schedule to reduce total reserve losses? Though product release is a strategic decision that hinges largely on a firm's R&D progress, marketing strategy, and competition pressure, among others, our discussion offers a new perspective that can be taken into account additionally.

Without pooling, the total amount of reserves planned and the associated reserve losses over the warranty life cycle of each single product would be independent of release timing. The impact of release schedule is essentially reflected by the way of pooling reserves for multiple products. More specifically, if the products are released simultaneously, then their warranty reserves can always be pooled over their warranty life cycles. However, if they are released sequentially (see Figure EC.2 for an illustration), then there would be some periods in which the warranty reserves of only one or two products can be pooled.

Without loss of generality, we consider three new products that are to be introduced into the market. They can be released either simultaneously or sequentially. In the latter situation, there are six possible cases of release sequence: 1-2-3, 1-3-2, 2-1-3, 2-3-1, 3-1-2, and 3-2-1. For example, "1-2-3" implies that product 1 is released prior to product 2, which, in turn, is followed by product 3. In the sequential release scenario, we further assume that the three products are periodically released for the sake of simplicity; that is, the release interval between any two successively released products is a constant. This assumption is in line with most practical scenarios, especially for consumer electronics. Note that the case of release interval being zero represents simultaneous release, in which all release sequences degenerate.

Since it is difficult to examine this problem analytically, we resort to numerical studies. Based on the reserve planning and pooling methodologies in Sections 3 and 5, we calculate the reserve levels and losses, with pooling, over the whole planning horizon. The percentages of reserve reduction and loss reduction, relative to the non-pooled scenario, can then be evaluated for each release schedule. Figure EC.4 shows the percentages of reduction against various release intervals (in quarters), using the parameter settings in Figures 3 and 4. We can see that simultaneous release is optimal, as it can lead to the highest percentages of reduction in both reserve levels and losses. In terms of sequential release, “3-1-2” is the best option of release sequence. As can be observed in Figure 3, the means and variances of AWC increments for product 1 are larger than those of product 2, which, in turn, exceed those of product 3. Because only one or two products can be pooled at the beginning and end of the planning horizon (see Figure EC.2), it is beneficial to release products with smaller reserve demands either earlier or later than those with larger demands so as to fully leverage the advantage of funds pooling. Nevertheless, the benefits induced by simultaneous release are much larger than those by the best release sequence. Therefore, simultaneous release, if practically feasible, should be recommended from a reserve management perspective.

EC.3. Technical Proofs

Proof of Theorem 1. We start from the continuous-time AWC forecasting model for a single product. By coupling product sales and failure processes over the warranty life cycle, the mean and variance of AWC up to time $t \in [0, L+W]$ are given by

$$\mathbb{E}[WC(t)] = \begin{cases} C_W \lambda\theta \left(\dfrac{t^2}{2}\right), & 0 \le t \le W, \\ C_W \lambda\theta \left(Wt - \dfrac{W^2}{2}\right), & W < t \le L, \\ C_W \lambda\theta \left[(L+W)t - \dfrac{L^2+W^2}{2} - \dfrac{t^2}{2}\right], & L < t \le L+W, \end{cases} \tag{EC.6}$$

and

$$\mathbb{V}\mathrm{ar}(WC(t)) = \begin{cases} C_W^2 \lambda\theta \left(\dfrac{t^2}{2} + \dfrac{\theta t^3}{3}\right), & 0 \le t \le W, \\ C_W^2 \lambda\theta \left[(1+\theta W)Wt - W\left(\dfrac{W}{2} + \dfrac{2\theta W^2}{3}\right)\right], & W < t \le L, \\ C_W^2 \lambda\theta \left[\begin{array}{l} \left(\dfrac{\theta L^3}{3} - \dfrac{L^2}{2} - \dfrac{2\theta W^3}{3} - \dfrac{W^2}{2}\right) \\ +(L+W+\theta W^2 - \theta L^2)t + \left(\theta L - \dfrac{1}{2}\right)t^2 - \dfrac{\theta t^3}{3} \end{array}\right], & L < t \le L+W. \end{cases} \tag{EC.7}$$

Equations (EC.6) and (EC.7) are derived from Equation (2) via conditional expectation and variance formulas, respectively, by conditioning on the cumulative sales $N(t)$ and associated purchase date $\boldsymbol{T(t)} = (T_1, T_2, \ldots, T_{N(t)})$ up to time t. The detailed derivation is referred to Xie and Ye (2016) and thus omitted.

Now we turn to the discrete-time AWC forecasting model to facilitate periodic reserve planning. For this purpose, we need to evaluate the mean and variance of AWC increment in an arbitrary interval $[t_s, t_e]$, i.e., $WC(t_e) - WC(t_s)$. It is clear that $\mathbb{E}[WC(t_e) - WC(t_s)] = \mathbb{E}[WC(t_e)] - \mathbb{E}[WC(t_s)]$. For the variance part, following the basic rule of variance operation, we have

$$\mathbb{V}\mathrm{ar}(WC(t_e) - WC(t_s)) = \mathbb{V}\mathrm{ar}(WC(t_e)) + \mathbb{V}\mathrm{ar}(WC(t_s)) - 2\mathbb{C}\mathrm{ov}(WC(t_e), WC(t_s)).$$

It is known that

$$\begin{aligned}
\mathbb{C}\mathrm{ov}(WC(t_e), WC(t_s)) &= \mathbb{C}\mathrm{ov}(WC(t_e) + WC(t_s) - WC(t_s), WC(t_s)) \\
&= \mathbb{C}\mathrm{ov}(WC(t_s), WC(t_s)) + \mathbb{C}\mathrm{ov}(WC(t_e) - WC(t_s), WC(t_s)) \\
&\simeq \mathbb{C}\mathrm{ov}(WC(t_s), WC(t_s)) = \mathbb{V}\mathrm{ar}(WC(t_s)).
\end{aligned}$$

The asymptotic approximation above is constructed by investigating the effects of λ and θ on approximation errors via simulation. Thus, $\mathbb{C}\mathrm{ov}(WC(t_e) - WC(t_s), WC(t_s)) \simeq 0$. Therefore, we have

$$\mathbb{V}\mathrm{ar}(WC(t_e) - WC(t_s)) \simeq \mathbb{V}\mathrm{ar}(WC(t_e)) - \mathbb{V}\mathrm{ar}(WC(t_s)).$$

This implies that the mean (resp. variance) of AWC increment in any time interval can be computed (resp. approximated) by the difference of means (resp. variances) at the two end points. This is an essential finding that can facilitate warranty reserve planning in an arbitrary interval $[t_s, t_e]$. As a result, we have $\mu_k = \mathbb{E}[WC(k\Delta)] - \mathbb{E}[WC((k-1)\Delta)$ and $\sigma_k^2 \simeq \mathbb{V}\mathrm{ar}(WC(k\Delta)) - \mathbb{V}\mathrm{ar}(WC((k-1)\Delta))$, and thus Theorem 1 can be easily obtained from Equations (EC.6) and (EC.7). The proof is complete. □

Proof of Lemma 1. We start from the well-known Markov inequality:

$$\Pr\{X \geq \epsilon\} \leq \frac{\mathbb{E}[X]}{\epsilon}, \tag{EC.8}$$

where X is a nonnegative random variable.

By replacing X and ϵ in (EC.8) with $(WC_k - \mu_k)^2/\mu_k^2$ and ϵ^2, respectively, we have

$$\Pr\left\{\left(\frac{WC_k - \mu_k}{\mu_k}\right)^2 \geq \epsilon^2\right\} \leq \frac{\mathbb{E}\left[\left(\frac{WC_k - \mu_k}{\mu_k}\right)^2\right]}{\epsilon^2},$$

which is equivalent to

$$\Pr\left\{\left|\frac{WC_k - \mu_k}{\mu_k}\right| \geq \epsilon\right\} \leq \frac{\mathbb{E}\left[(WC_k - \mu_k)^2\right]}{\mu_k^2 \epsilon^2} = \frac{\sigma_k^2}{\mu_k^2 \epsilon^2} = \frac{CV_k^2}{\epsilon^2}.$$

The proof is complete. □

Proof of Proposition 1. For the mean, we need to examine $\mu_k - \mu_{k-1}$ for $k \in \{2, \dots, n+m\}$. We have

$$\mu_k - \mu_{k-1} = \begin{cases} C_W \lambda\theta\Delta^2 > 0, & k = 2, \dots, m, \\ 0, & k = m+2, \dots, n, \\ -C_W \lambda\theta\Delta^2 < 0, & k = n+2, \dots, n+m. \end{cases}$$

Further, by checking $\mu_{m+1} - \mu_m$ and $\mu_{n+1} - \mu_n$, we have

$$\mu_{m+1} - \mu_m = \frac{1}{2} C_W \lambda\theta\Delta^2 > 0,$$

$$\mu_{n+1} - \mu_n = -\frac{1}{2} C_W \lambda\theta\Delta^2 < 0.$$

Thus, μ_k is an increasing function in $k \in \{1, \dots, m\}$, a constant function in $k \in \{m+1, \dots, n\}$, and a decreasing function in $k \in \{n+1, \dots, n+m\}$.

Likewise, for the variance, we examine $\sigma_k^2 - \sigma_{k-1}^2$ for $k \in \{2, \dots, n+m\}$, as follows:

$$\sigma_k^2 - \sigma_{k-1}^2 = \begin{cases} C_W^2 \lambda\theta[\Delta^2 + 2\theta\Delta^3(k-1)] > 0, & k = 2, \dots, m, \\ 0, & k = m+2, \dots, n, \\ C_W^2 \lambda\theta[-\Delta^2 + 2\theta\Delta^3(1+n-k)] < 0, & k = n+2, \dots, n+m. \end{cases}$$

Further, by checking $\sigma_{m+1}^2 - \sigma_m^2$ and $\sigma_{n+1}^2 - \sigma_n^2$, we have

$$\sigma_{m+1}^2 - \sigma_m^2 = C_W^2 \lambda\theta \left[\frac{1}{2}\Delta^2 + \theta\Delta^3 \left(m - \frac{1}{3} \right) \right] > 0,$$

and

$$\sigma_{n+1}^2 - \sigma_n^2 = -C_W^2 \lambda\theta \left[\frac{1}{2}\Delta^2 + \frac{1}{3}\theta\Delta^3 \right] < 0.$$

Hence, σ_k^2 is an increasing function in $k \in \{1, \dots, m\}$, a constant function in $k \in \{m+1, \dots, n\}$, and a decreasing function in $k \in \{n+1, \dots, n+m\}$.

However, for the coefficient of variation, the result is not as straightforward as those of the mean and variance. We need to examine $CV_k - CV_{k-1}$ case by case, as follows.

Case 1: $k \in \{2, \dots, m\}$. Let $z = k - 1/2$, then $k^2 - k = z^2 - 1/4$. In this case, we have

$$\begin{aligned} CV_k - CV_{k-1} &= \frac{\sqrt{\frac{1}{z} + \theta\Delta + \frac{\theta\Delta}{12z^2}}}{\sqrt{\lambda\theta}\Delta} - \frac{\sqrt{\frac{1}{z-1} + \theta\Delta + \frac{\theta\Delta}{12(z-1)^2}}}{\sqrt{\lambda\theta}\Delta} \\ &= \frac{\left(\frac{1}{z} + \theta\Delta + \frac{\theta\Delta}{12z^2}\right) - \left(\frac{1}{z-1} + \theta\Delta + \frac{\theta\Delta}{12(z-1)^2}\right)}{\sqrt{\lambda\theta}\Delta \left(\sqrt{\frac{1}{z} + \theta\Delta + \frac{\theta\Delta}{12z^2}} + \sqrt{\frac{1}{z-1} + \theta\Delta + \frac{\theta\Delta}{12(z-1)^2}} \right)}. \end{aligned} \tag{EC.9}$$

The denominator of (EC.9) is strictly positive, so we only need to examine its numerator; that is,

$$\left(\frac{1}{z} + \theta\Delta + \frac{\theta\Delta}{12z^2}\right) - \left(\frac{1}{z-1} + \theta\Delta + \frac{\theta\Delta}{12(z-1)^2}\right) = \frac{1}{z} - \frac{1}{z-1} + \frac{\theta\Delta}{12}\left(\frac{1}{z^2} - \frac{1}{(z-1)^2}\right) < 0.$$

Thus, $CV_k - CV_{k-1} < 0$ when $k \in \{2, \dots, m\}$.

Case 2: $k\in\{m+2,\dots,n\}$. It is straightforward that CV_k is constant in this interval.

Case 3: $k\in\{n+2,\dots,n+m\}$. Let $z'=n+m-k+1/2$, then $m^2-n^2-n+k+2nk-k^2-1/3=-(z')^2+2mz'-1/12$. In this case, we have

$$CV_k-CV_{k-1}=\frac{\sqrt{\frac{1}{z'}-\theta\Delta+2m\theta\Delta\frac{1}{z'}-\frac{\theta\Delta}{12z'^2}}}{\sqrt{\lambda\theta}\Delta}-\frac{\sqrt{\frac{1}{z'+1}-\theta\Delta+2m\theta\Delta\frac{1}{z'+1}-\frac{\theta\Delta}{12(z'+1)^2}}}{\sqrt{\lambda\theta}\Delta}$$
$$=\frac{\left(\frac{1}{z'}-\theta\Delta+2m\theta\Delta\frac{1}{z'}-\frac{\theta\Delta}{12z'^2}\right)-\left(\frac{1}{z'+1}-\theta\Delta+2m\theta\Delta\frac{1}{z'+1}-\frac{\theta\Delta}{12(z'+1)^2}\right)}{\sqrt{\lambda\theta}\Delta\left(\sqrt{\frac{1}{z'}-\theta\Delta+2m\theta\Delta\frac{1}{z'}-\frac{\theta\Delta}{12z'^2}}+\sqrt{\frac{1}{z'+1}-\theta\Delta+2m\theta\Delta\frac{1}{z'+1}-\frac{\theta\Delta}{12(z'+1)^2}}\right)}. \tag{EC.10}$$

The denominator of (EC.10) is strictly positive, thus we only need to examine its numerator:

$$\left(\frac{1}{z'}-\theta\Delta+2m\theta\Delta\frac{1}{z'}-\frac{\theta\Delta}{12z'^2}\right)-\left(\frac{1}{z'+1}-\theta\Delta+2m\theta\Delta\frac{1}{z'+1}-\frac{\theta\Delta}{12(z'+1)^2}\right)$$
$$=\left(\frac{1}{z'}-\frac{1}{z'+1}\right)\left[1+2m\theta\Delta-\frac{\theta\Delta}{12}\left(\frac{1}{z'}+\frac{1}{z'+1}\right)\right].$$

Further let $\xi(z')=1+2m\theta\Delta-\frac{\theta\Delta}{12}\left(\frac{1}{z'}+\frac{1}{z'+1}\right)$. It is intuitive that $\xi(z')$ is an increasing function of z'. As $n+2\le k\le n+m$ and $z'=n+m-k+1/2$, we know that $1/2\le z'\le m-3/2$. Note that here we should have $m-3/2\ge 1/2$, i.e., $m\ge 2$. This mild condition is reasonable and practical, since it is meaningless to study the monotonicity of CV_k in the case of $m=1$ (the interval $[n+1, n+m]$ would degenerate). This implies that $\xi(z')$ increases from $\xi(1/2)$ to $\xi(m-3/2)$. Notice that $\xi(1/2)=1+2(m-1/9)\theta\Delta>0$. This means that $\xi(z')$ is always greater than zero when $1/2\le z'\le m-3/2$. Thus, $CV_k-CV_{k-1}>0$ when $k\in\{n+2,\dots,n+m\}$.

Further, by checking $CV_{m+1}-CV_m$ and $CV_{n+1}-CV_n$, we have

$$CV_{m+1}-CV_m=\frac{\sqrt{1+\theta m\Delta}}{\sqrt{\lambda\theta m}\Delta}-\frac{\sqrt{(m-1/2)+\theta\Delta(m^2-m+1/3)}}{\sqrt{\lambda\theta}\Delta(m-1/2)}$$
$$=\frac{\left(\frac{1}{m}-\frac{1}{m-1/2}\right)-\frac{\theta\Delta}{12(m-1/2)^2}}{\sqrt{\lambda\theta}\Delta\left(\sqrt{(1+\theta m\Delta)\frac{1}{m}}+\sqrt{\frac{1}{m-1/2}+\theta\Delta+\frac{\theta\Delta}{12(m-1/2)^2}}\right)}<0,$$

and

$$CV_{n+1}-CV_n=\frac{\sqrt{(m-1/2)+\theta\Delta(m^2-1/3)}}{\sqrt{\lambda\theta}\Delta(m-1/2)}-\frac{\sqrt{1+\theta m\Delta}}{\sqrt{\lambda\theta m}\Delta}$$
$$=\frac{\left(\frac{1}{m-1/2}-\frac{1}{m}\right)+\theta\Delta\frac{m^2-m+1/6}{m-1/2}}{\sqrt{\lambda\theta}\Delta\left(\sqrt{(1+\theta m\Delta)\frac{1}{m}}+\sqrt{\frac{1}{m-1/2}+\theta\Delta\frac{m^2-1/3}{m-1/2}}\right)}>0.$$

Therefore, CV_k is a decreasing function in $k\in\{1,\dots,m\}$, a constant function in $k\in\{m+1,\dots,n\}$, and an increasing function in $k\in\{n+1,\dots,n+m\}$. The proof is complete. □

Proof of Lemma 2. The proof of this lemma follows directly from Lemmas 1 and 2 in Gallego and Moon (1993). □

Proof of Proposition 2. To determine the optimal amount of warranty reserves in the kth period, we take the first derivative of $\hat{\pi}_k(WR_k)$ in Equation (8) with respect to WR_k:

$$\frac{\partial \hat{\pi}_k(WR_k)}{\partial WR_k} = \frac{C_h - C_b}{2} + \frac{C_h + C_b}{2} \frac{WR_k - \mu_k}{\sqrt{\sigma_k^2 + (WR_k - \mu_k)^2}}.$$

Further taking the second derivative of $\hat{\pi}_k(WR_k)$ with respect to WR_k, we have

$$\frac{\partial^2 \hat{\pi}_k(WR_k)}{\partial WR_k^2} = \frac{C_h + C_b}{2} \frac{\sigma_k^2}{[\sigma_k^2 + (WR_k - \mu_k)^2]^{3/2}} > 0.$$

Because $\partial^2 \hat{\pi}_k(WR_k)/\partial WR_k^2 > 0$, there must exist an optimal WR_k^* such that $\hat{\pi}_k(WR_k)$ is minimized. Setting $\partial \hat{\pi}_k(WR_k)/\partial WR_k$ to zero, we can easily obtain WR_k^* in Equation (9). $\square$

Proof of Proposition 3. Define $\widehat{WR}_i(\phi) = WR_i^* + \phi(\Delta Q_{i-1} - \mu_{i-1})$, $i = 2, 3, \ldots, k$. According to (12), the optimal learning parameter ϕ_k^* is obtained by minimizing the following criterion function

$$h_k(\phi) = \frac{1}{k-2} \sum_{i=2}^{k-1} \left[C_h \left(\widehat{WR}_i(\phi) - \Delta Q_i \right)^+ + C_b \left(\Delta Q_i - \widehat{WR}_i(\phi) \right)^+ \right]. \tag{EC.11}$$

To establish the convergence of the estimator ϕ_k^*, we need to show that $h_k(\phi)$ converges to a fixed function $\mathbb{H}(\phi)$ in probability. To this end, we rewrite $h_k(\phi) = h_{k1}(\phi) + h_{k2}(\phi)$, where

$$\begin{cases} h_{k1}(\phi) = \frac{1}{k-2} \sum_{i=2}^{k-1} C_h \Big[WR_i^* + \phi(\Delta Q_{i-1} - \mu_{i-1}) - \Delta Q_i \Big]^+, \\ h_{k2}(\phi) = \frac{1}{k-2} \sum_{i=2}^{k-1} C_b \Big[\Delta Q_i - WR_i^* - \phi(\Delta Q_{i-1} - \mu_{i-1}) \Big]^+. \end{cases}$$

The only source of randomness is $(\Delta Q_{i-1}, \Delta Q_i)$. When $i - 1 > m$ and $i \le n$, i.e., $m + 1 < i \le n$, $(\Delta Q_{i-1}, \Delta Q_i)$ has an identical distribution because of the HPP sales process and exponentially distributed failure times. In addition, WR_i^* given in Proposition 2 is a constant for $m + 1 < i \le n$. Hence, $(\Delta Q_{i-1}, \Delta Q_i)$ can be recognized as an $(m+1)$-dependent sequence. It is straightforward to show that the variances of both $h_{k1}(\phi)$ and $h_{k2}(\phi)$ converge to zero as $n \to \infty$. Therefore,

$$\begin{cases} h_{k1}(\phi) \to_{qm} \mathbb{H}_1(\phi) = C_h \mathbb{E} \Big[WR_{m+2}^* + \phi(\Delta Q_{m+1} - \mu_{m+1}) - \Delta Q_{m+2} \Big]^+, \\ h_{k2}(\phi) \to_{qm} \mathbb{H}_2(\phi) = C_b \mathbb{E} \Big[\Delta Q_{m+2} - WR_{m+2}^* - \phi(\Delta Q_{m+1} - \mu_{m+1}) \Big]^+, \end{cases} \tag{EC.12}$$

where $\to_{qm}$ denotes convergence in quadratic mean (which implies convergence in probability). Note that the expectations on the right-hand side (RHS) of Equation (EC.12) are finite according to Jensen's inequality, i.e., $\mathbb{E}[Y]^+ \le \mathbb{E}|Y| \le \sqrt{\mathbb{E}[Y^2]}$ for any random variable Y, and the random variables inside the square brackets on the RHS of the above display have finite variances.

Then, we can define $\mathbb{H}(\phi) = \mathbb{H}_1(\phi) + \mathbb{H}_2(\phi)$ as the desired fixed function. Because $h_k(\phi)$ is convex in ϕ for all $k > 1$, $\mathbb{H}(\phi)$ is also convex (Fleming and Harrington 1991, Lemma 8.3.1). The convexity of $\mathbb{H}(\phi)$ can also be directly verified by checking $\mathbb{H}_1(\phi)$ and $\mathbb{H}_2(\phi)$. Based on the same lemma,

ϕ_k^* (the minimizer of $h_k(\phi)$) converges to ϕ^* (the minimizer of $\mathbb{H}(\phi)$) in probability. In particular, according to Equations (6) and (9), the fixed formula of $\mathbb{H}(\phi)$ can be calculated as

$$\mathbb{H}(\phi) = \frac{1}{2}C_h A\sigma_{m+2} + (C_b + C_h)\mathbb{E}\left[\Delta Q_{m+2} - WR^*_{m+2} - \phi(\Delta Q_{m+1} - \mu_{m+1})\right]^+, \tag{EC.13}$$

which implies that ϕ_k^* converges to a constant as $n \to \infty$. $\square$

Proof of Proposition 4. To show the $\sqrt{k}$-convergence of ϕ_k^*, it is more convenient to center the random criterion function $h_k(\phi)$ in (EC.11) and the fixed criterion function $\mathbb{H}(\phi)$ in (EC.13) at ϕ^*. To this end, we first define $\tilde{\mathbb{H}}_k : \mathbb{R} \mapsto \mathbb{R}$ given by

$$\begin{aligned}\tilde{\mathbb{H}}_k(\chi) = (k-2)(C_b + C_h)\Bigg\{&\mathbb{E}\left[\Delta Q_{m+2} - WR^*_{m+2} - \left(\tfrac{\chi}{\sqrt{k-2}} + \phi^*\right)(\Delta Q_{m+1} - \mu_{m+1})\right]^+ \\ &-\mathbb{E}\left[\Delta Q_{m+2} - WR^*_{m+2} - \phi^*(\Delta Q_{m+1} - \mu_{m+1})\right]^+\Bigg\}.\end{aligned}$$

By definition, the minimum value of $\tilde{\mathbb{H}}_k(\chi)$ is reached at $\chi = 0$. In addition, based on the Lebesgue Dominated Convergence Theorem, we can verify that $\tilde{\mathbb{H}}_k(\chi)$ is differentiable at $\chi = 0$. Then, given the assumption that $\mathbb{H}(\phi)$ is twice differentiable at $\phi = \phi^*$ and the second derivative $\mathbb{H}''(\phi^*)$, we can apply Taylor's expansion to see that as $k \to \infty$,

$$\tilde{\mathbb{H}}_k(\chi) = \frac{1}{2}\mathbb{H}''(\phi^*)\chi^2 + o(1). \tag{EC.14}$$

Next, we turn to the random criterion function $h_k(\phi)$. Define $\varepsilon_i = WR_i^* + \phi^*(\Delta Q_{i-1} - \mu_{i-1}) - \Delta Q_i$. Thus, we can rescale the summand of the random criterion function as

$$\begin{aligned}m_{ki}(\chi) = C_h&\left\{\left[WR_i^* + \left(\tfrac{\chi}{\sqrt{k-2}} + \phi^*\right)(\Delta Q_{i-1} - \mu_{i-1}) - \Delta Q_i\right]^+ - \varepsilon_i^+\right\} \\ +C_b&\left\{\left[\Delta Q_i - WR_i^* - \left(\tfrac{\chi}{\sqrt{k-2}} + \phi^*\right)(\Delta Q_{i-1} - \mu_{i-1})\right]^+ - (-\varepsilon_i)^+\right\},\end{aligned}$$

and $h_k(\phi)$ can be correspondingly rescaled as $\mathbb{G}_k(\chi) = \sum_{i=2}^{k-1} m_{ki}(\chi)$. Observe that $\chi \mapsto \mathbb{G}_k(\chi)$ is a convex function. Let χ_k^* be the minimizer of $\mathbb{G}_k(\chi)$, we have $\chi_k^* = \sqrt{k-2}(\phi_k^* - \phi^*)$. In particular, the right derivative of $m_{ki}(\chi)$ at $\chi = 0$ can be expressed as

$$\mathcal{W}_{ki} = \frac{1}{\sqrt{k-2}}\Big(C_h\{\varepsilon_i \geq 0\} - C_b\{\varepsilon_i < 0\}\Big)(\Delta Q_{i-1} - \mu_{i-1}).$$

Then it is easy to verify that $\mathbb{E}\mathcal{W}_{ki} = (k-2)^{-1/2}\mathbb{H}'(\phi^*) = 0$ when $m+1 < i \leq k-1$. Let $\mathcal{R}_{ki}(\chi) = m_{ki}(\chi) - \mathcal{W}_{ki}\chi$. Then $\mathbb{G}_k(\chi)$ can be decomposed as

$$\mathbb{G}_k(\chi) = \tilde{\mathbb{H}}_k(\chi) + \sum_{i=2}^{k-1}\mathcal{W}_{ki}\chi + \sum_{i=2}^{k-1}\left[\mathcal{R}_{ki}(\chi) - \mathbb{E}\mathcal{R}_{ki}(\chi)\right] + \left[\sum_{i=2}^{k-1}\mathbb{E}\mathcal{R}_{ki}(\chi) - \tilde{\mathbb{H}}_k(\chi)\right].$$

Direct calculation yields $\sum_{i=2}^{k-1}\mathbb{E}\mathcal{R}_{ki}(\chi)-\tilde{\mathbb{H}}_k(\chi)=o(1)$ and $\mathbb{E}[\sum_{i=2}^{k-1}(\mathcal{R}_{ki}(\chi)-\mathbb{E}\mathcal{R}_{ki}(\chi))]^2=o(1)$, which implies that $\sum_{i=2}^{k-1}[\mathcal{R}_{ki}(\chi)-\mathbb{E}\mathcal{R}_{ki}(\chi)]=o_P(1)$. Hence, combined with Equation (EC.14), the above decomposition can be rewritten as

$$\mathbb{G}_k(\chi)=\frac{1}{2}\mathbb{H}''(\phi^*)\chi^2+\sum_{i=2}^{k-1}\mathcal{W}_{ki}\chi+o_P(1). \tag{EC.15}$$

Therefore, $\mathbb{G}_k(\chi)-\sum_{i=2}^{k-1}\mathcal{W}_{ki}\chi$ converges to $\frac{1}{2}\mathbb{H}''(\phi^*)\chi^2$ in probability for each fixed χ. Because $\chi\mapsto\mathbb{G}_k(\chi)-\sum_{i=2}^{k-1}\mathcal{W}_{ki}\chi$ is convex, we can strengthen the convergence to be uniform over any compact subset in $\mathbb{R}$ by the convexity lemma (Fleming and Harrington 1991, Lemma 8.3.1). Consequently, based on the Central Limit Theorem for m-dependent sequences, $\sum_{i=2}^{k-1}\mathcal{W}_{ki}$ converges to a mean-zero normal distribution (denoted as $\mathcal{W}$).

Let $w_k^*=-[\mathbb{H}''(\phi^*)]^{-1}\sum_{i=2}^{k-1}\mathcal{W}_{ki}$ be the minimum of the two leading terms in Equation (EC.15), we next attempt to show that $\chi_k^*-w_k^*=o_P(1)$. For any $\delta,\epsilon>0$, we will show that $P(|\chi_k^*-w_k^*|>\delta)<\epsilon$ as $k\to\infty$. The weak convergence of $\sum_{i=2}^{k-1}\mathcal{W}_{ki}$ implies that w_k^* is uniformly tight. We find a number $\mathcal{M}$ such that $P(|w_k^*|>\mathcal{M}-\delta)<\epsilon$. For $\chi\in[-\mathcal{M},\mathcal{M}]$, the convexity lemma implies

$$\mathbb{G}_k(\chi)=\frac{1}{2}\mathbb{H}''(\phi^*)\left(\chi+[\mathbb{H}''(\phi^*)]^{-1}\sum_{i=2}^{k-1}\mathcal{W}_{ki}\right)^2-\frac{1}{2}[\mathbb{H}''(\phi^*)]^{-1}\left(\sum_{i=2}^{k-1}\mathcal{W}_{ki}\right)^2+o_P^{\mathrm{U}}(1),$$

where the superscript U is used to indicate that $o_P^{\mathrm{U}}(1)$ term is uniform over $\chi\in[-\mathcal{M},\mathcal{M}]$. Conditional on the event that $|w_k^*|<\mathcal{M}-\delta$, we can use the convexity of $\mathbb{G}_k(\chi)$ to show that

$$\frac{\delta}{\chi-w_k^*}\mathbb{G}_k(\chi)+\frac{\chi-w_k^*-\delta}{\chi-w_k^*}\mathbb{G}_k(w_k^*)\geq\mathbb{G}_k(w_k^*+\delta)\geq\mathbb{G}_k(w_k^*)+o_P^{\mathrm{U}}(1),$$

for $\chi-w_k^*>\mathcal{M}$ ($\forall\chi\in[-\mathcal{M},\mathcal{M}]$). The first inequality can be proved with the Jensen's inequality while the second inequality is due to the fact that w_k^* is the minimum of the two leading terms in Equation (EC.15). One can easily verify that $\mathbb{G}_k(\chi)>\mathbb{G}_k(w_k^*)+o_P^{\mathrm{U}}(1)$ when $\chi-w_k^*>\delta$ by rearranging the above inequalities. By repeating the above procedure and use the concavity of $-\mathbb{G}_k(\chi)$, we can show that $\mathbb{G}_k(\chi)>\mathbb{G}_k(w_k^*)+o_P^{\mathrm{U}}(1)$ when $\chi-w_k^*<-\delta$. Combining the above results, one can see that, conditional on $|w_k^*|<\mathcal{M}-\delta$, $|\chi_k^*-w_k^*|<\delta$ holds with probability tending to one. Thus, we can finally attain

$$\liminf P(|\chi_k^*-w_k^*|<\delta)\geq\liminf P(|w_k^*|\leq M-\delta)\geq 1-\epsilon.$$

Hence, we conclude that $\chi_k^*=w_k^*+o_P(1)$. Since $w_k^*\rightsquigarrow\mathcal{W}$, an appeal to Slutsky's lemma yields that $\chi_k^*=\sqrt{k-2}(\phi_k^*-\phi^*)\rightsquigarrow\mathcal{W}$. The proof is completed by noting that $\mathcal{W}$ is a mean-zero normal distribution. □

Proof of Proposition 5. The proof is similar to that of Proposition 2, we thus omit it. □

Proof of Lemma 3. We first take the first derivative of Ψ_k^* with respect to $\mu_{j,k}$:

$$\frac{\partial \Psi_k^*}{\partial \mu_{j,k}} = \frac{A}{2} \frac{\sqrt{\sum_{l=1}^M \sigma_{l,k}^2} - \sum_{l=1}^M \sigma_{l,k}}{\left(\sum_{l=1}^M \mu_{l,k} + \frac{A}{2}\sum_{l=1}^M \sigma_{l,k}\right)^2}.$$

Since $\sqrt{\sum_{l=1}^M \sigma_{l,k}^2} < \sum_{l=1}^M \sigma_{l,k}$ and $A>0$, we have $\partial \Psi_k^*/\partial \mu_{j,k} < 0$. Thus, Ψ_k^* is a decreasing function of $\mu_{j,k}$, $j=1,\dots,M$, $k=1,\dots,n+m$.

Then, we take the first derivative of Ψ_k^* with respect to $\sigma_{j,k}$:

$$\frac{\partial \Psi_k^*}{\partial \sigma_{j,k}} = \frac{A}{2} \frac{\left(\sqrt{\sum_{l=1}^M \sigma_{l,k}^2} - \sigma_{j,k}\right)\sum_{l=1}^M \mu_{l,k} + \frac{A}{2}\sum_{l=1, l\neq j}^M \sigma_{l,k}\left(\sigma_{l,k} - \sigma_{j,k}\right)}{\left(\sum_{l=1}^M \mu_{l,k} + \frac{A}{2}\sum_{l=1}^M \sigma_{l,k}\right)^2 \sqrt{\sum_{l=1}^M \sigma_{l,k}^2}}.$$

For $\sigma_{\min,k} = \min_{j=1,\dots,M}\{\sigma_{j,k}\}$, we have $\sqrt{\sum_{l=1}^M \sigma_{l,k}^2} - \sigma_{\min,k} > 0$ and $\sum_{l=1}^M \sigma_{l,k}\left(\sigma_{l,k} - \sigma_{\min,k}\right) > 0$. Thus, $\partial \Psi_k^*/\partial \sigma_{\min,k} > 0$. In this case, Ψ_k^* decreases as $\sigma_{\min,k}$ decreases. □

Proof of Corollary 1. By setting $\sigma_{1,k} = \sigma_{2,k} = \cdots = \sigma_{M,k} = \sigma_k$, we can further simplify Equation (19) as

$$\Psi_k^* = \frac{\frac{A}{2}(M\sigma_k - \sqrt{M}\sigma_k)}{\sum_{j=1}^M \mu_{j,k} + \frac{A}{2}M\sigma_k} = \frac{\frac{A}{2}(M - \sqrt{M})}{\frac{\sum_{j=1}^M \mu_{j,k}}{\sigma_k} + \frac{A}{2}M} = \frac{\frac{A}{2}(M - \sqrt{M})}{\sum_{j=1}^M \frac{1}{CV_{j,k}} + \frac{A}{2}M}.$$

Then, it is easy to verify that Corollary 1 holds. This completes the proof. □

Proof of Proposition 6. To prove this result, in the kth period, we need to sort out the standard deviations of the M products in an ascending order, i.e., $\sigma_{(1),k} < \sigma_{(2),k} < \cdots < \sigma_{(M),k}$. Let $\delta_{(1),k} = \sigma_{(1),k}/\sigma_{(1),k} = 1$, $\delta_{(2),k} = \sigma_{(2),k}/\sigma_{(1),k}$, $\cdots$, $\delta_{(M),k} = \sigma_{(M),k}/\sigma_{(1),k}$ (Notice that $\delta_{(M),k} = \delta_{\max,k} = \sigma_{\max,k}/\sigma_{\min,k}$). Then, we have $\delta_{(M),k} > \cdots > \delta_{(2),k} > \delta_{(1),k} = 1$. In this setting, Ψ_k° can be rewritten as $\Psi_k^{\circ} = 1 - \sqrt{\sum_{j=1}^M \delta_{(j),k}^2}/\sum_{j=1}^M \delta_{(j),k}$.

Taking the first derivative of Ψ_k° with respect to $\delta_{(M),k}$ gives

$$\frac{\partial \Psi_k^{\circ}}{\partial \delta_{(M),k}} = \frac{\sum_{j=1}^{M-1} \delta_{(j),k}\left(\delta_{(j),k} - \delta_{(M),k}\right)}{\left(\sum_{j=1}^M \delta_{(j),k}\right)^2 \sqrt{\sum_{j=1}^M \delta_{(j),k}^2}} < 0.$$

Thus, Ψ_k° decreases as $\delta_{(M),k}$ (or equivalently, $\delta_{\max,k}$) increases. □

Proof of Lemma 4. To prove this result, we turn to showing that $(\Gamma_k^{(j,l)})^2$ is a nondecreasing function of k. Then, once $(\Gamma_1^{(j,l)})^2 > 1$ (and $\theta_j > \theta_l$), we can get $(\Gamma_k^{(j,l)})^2 > 1$ for $k = 2,\dots,n+m$. However, since $\sigma_{j,k}^2$ is a piece-wise function, we need to examine the monotonicity of $(\Gamma_k^{(j,l)})^2$ during the three intervals, respectively.

Case 1: $k \in \{2, \dots, m\}$. Let $z = k - 1/2$ (similar to the proof of Proposition 1), then

$$(\Gamma_k^{(j,l)})^2 = \frac{\sigma_{j,k}^2}{\sigma_{l,k}^2} = \hat{C}\frac{z + \theta_j \Delta\left(z^2 + \frac{1}{12}\right)}{z + \theta_l \Delta\left(z^2 + \frac{1}{12}\right)},$$

where $\hat{C} = C_{Wj}^2 \lambda_j \theta_j / C_{Wl}^2 \lambda_l \theta_l$ is a constant. We have

$$\begin{aligned}(\Gamma_k^{(j,l)})^2 - (\Gamma_{k-1}^{(j,l)})^2 &= \hat{C}\frac{z + \theta_j \Delta\left(z^2 + \frac{1}{12}\right)}{z + \theta_l \Delta\left(z^2 + \frac{1}{12}\right)} - \hat{C}\frac{z - 1 + \theta_j \Delta\left((z-1)^2 + \frac{1}{12}\right)}{z - 1 + \theta_l \Delta\left((z-1)^2 + \frac{1}{12}\right)} \\ &= \hat{C}\frac{(\theta_j - \theta_l)\Delta\left(z^2 + z - \frac{1}{12}\right)}{[z + \theta_l \Delta\left(z^2 + \frac{1}{12}\right)][z - 1 + \theta_l \Delta\left((z-1)^2 + \frac{1}{12}\right)]} > 0.\end{aligned}$$

Thus, if $\theta_j > \theta_l$, then $(\Gamma_k^{(j,l)})^2$ is an increasing function of k for $k = 1, \dots, m$.

Case 2: $k \in \{m+2, \dots, n\}$. In this case, it is intuitive that $(\Gamma_k^{(j,l)})^2$ is a constant.

Case 3: $k \in \{n+2, \dots, n+m\}$. Let $z' = n + m - k + 1/2$, then we have

$$\begin{aligned}(\Gamma_k^{(j,l)})^2 - (\Gamma_{k-1}^{(j,l)})^2 &= \hat{C}\frac{z' + \theta_j \Delta\left(-z'^2 + 2mz' - \frac{1}{12}\right)}{z' + \theta_l \Delta\left(-z'^2 + 2mz' - \frac{1}{12}\right)} - \hat{C}\frac{z' + 1 + \theta_j \Delta\left(-(z'+1)^2 + 2m(z'+1) - \frac{1}{12}\right)}{z' + 1 + \theta_l \Delta\left(-(z'+1)^2 + 2m(z'+1) - \frac{1}{12}\right)} \\ &= \hat{C}\frac{(\theta_j - \theta_l)\Delta\left(z'^2 + z' - \frac{1}{12}\right)}{\left[z' + \theta_l \Delta\left(-z'^2 + 2mz' - \frac{1}{12}\right)\right]\left[z' + 1 + \theta_l \Delta\left(-(z'+1)^2 + 2m(z'+1) - \frac{1}{12}\right)\right]} > 0.\end{aligned}$$

Thus, $\Gamma_k^{(j,l)}$ is an increasing function of k for $k = n+1, \dots, n+m$.

Up to now, we have shown that $(\Gamma_k^{(j,l)})^2$ is increasing in k over $[1,\ m]$, is constant over $[m+1,\ n]$, and is increasing in k during $[n+1,\ n+m]$. To demonstrate its global monotonicity, we need to check $(\Gamma_{m+1}^{(j,l)})^2 - (\Gamma_m^{(j,l)})^2$ and $(\Gamma_{n+1}^{(j,l)})^2 - (\Gamma_n^{(j,l)})^2$, as follows:

$$\begin{aligned}(\Gamma_{m+1}^{(j,l)})^2 - (\Gamma_m^{(j,l)})^2 &= \hat{C}\frac{1 + \theta_j m\Delta}{1 + \theta_l m\Delta} - \hat{C}\frac{m - \frac{1}{2} + \theta_j \Delta\left(m^2 - m + \frac{1}{3}\right)}{m - \frac{1}{2} + \theta_l \Delta\left(m^2 - m + \frac{1}{3}\right)} \\ &= \hat{C}\frac{(\theta_j - \theta_l)\Delta\left(\frac{m}{2} - \frac{1}{3}\right)}{[1 + \theta_l m\Delta]\left[m - \frac{1}{2} + \theta_l \Delta\left(m^2 - m + \frac{1}{3}\right)\right]} > 0,\end{aligned}$$

and

$$\begin{aligned}(\Gamma_{n+1}^{(j,l)})^2 - (\Gamma_n^{(j,l)})^2 &= \hat{C}\frac{m - \frac{1}{2} + \theta_j \Delta\left(m^2 - \frac{1}{3}\right)}{m - \frac{1}{2} + \theta_l \Delta(m^2 - \frac{1}{3})} - \hat{C}\frac{1 + \theta_j m\Delta}{1 + \theta_l m\Delta} \\ &= \hat{C}\frac{(\theta_j - \theta_l)\Delta\left(\frac{m}{2} - \frac{1}{3}\right)}{[1 + \theta_l m\Delta]\left[m - 1/2 + \theta_l \Delta\left(m^2 - \frac{1}{3}\right)\right]} > 0.\end{aligned}$$

Therefore, both $(\Gamma_k^{(j,l)})^2$ and $\Gamma_k^{(j,l)}$ are nondecreasing functions of k over $[1,\ n+m]$. In this manner, once $\Gamma_1^{(j,l)} > 1$, we know that $\Gamma_k^{(j,l)} > 1$ for $k = 2, \dots, n+m$. $\square$

Proof of Proposition EC.1. Let $\psi_k = \frac{\sqrt{\sigma_{(j-1),2m+k}^2 + \sigma_{(j),m+k}^2 + \sigma_{(j+1),k}^2}}{\sigma_{(j-1),2m+k} + \sigma_{(j),m+k} + \sigma_{(j+1),k}}$. We are trying to examine the monotonicity of ψ_k in k. It is easy to know that the monotonicity of Ψ_k° holds as follows.

$$\begin{cases} \Psi_k^{\circ} \geq \Psi_{k+1}^{\circ}, \text{ if } \psi_k \leq \psi_{k+1}, \\ \Psi_k^{\circ} \leq \Psi_{k+1}^{\circ}, \text{ if } \psi_k \geq \psi_{k+1}, 1 \leq k \leq m-1. \end{cases}$$

Let $a_{(j)} = C_{W(j)}\lambda_{(j)}\theta_{(j)}$, $t = k - 1/2$, and $\Delta = 1$ (the value of Δ has no influence on the final results, so we can simply normalize it to 1, e.g., 1 month, for the sake of conciseness). Then, we can obtain a simplified expression for $\sigma^2_{(j-1),2m+k}$, $\sigma^2_{(j),m+k}$, and $\sigma^2_{(j+1),k}$ as

$$\begin{cases} \sigma^2_{(j-1),2m+k} = a_{(j-1)}\theta_{(j-1)}(m^2 - t^2 - 1/12) + a_{(j-1)}(m-t), \\ \sigma^2_{(j),m+k} = a_{(j)}\theta_{(j)}m^2 + a_{(j)}m, \\ \sigma^2_{(j+1),k} = a_{(j+1)}\theta_{(j+1)}t^2 + a_{(j+1)}t + a_{(j+1)}\theta_{(j+1)}/12. \end{cases}$$

It is easy to know that $\sigma_{(j-1),2m+k}$ decreases as k increases, $\sigma_{(j),m+k} = a_{(j)}\theta_{(j)}m^2 - a_{(j)}m$ is a constant, and $\sigma_{(j+1),k}$ increases as k increases. Then, we prove the monotonicity of ψ_k in $\sigma_{(j-1),2m+k}$ and $\sigma_{(j+1),k}$, respectively.

Case I: $\psi_k \le \psi_{k+1}$. We first investigate the condition under which ψ_k is an increasing function of k. Taking the first derivative of ψ_k with respect to $\sigma^2_{(j-1),2m+k}$ yields

$$\frac{\partial \psi_k}{\partial \sigma_{(j-1),2m+k}} = \frac{\sigma_{(j-1),2m+k}\left(\sigma_{(j),m+k} + \sigma_{(j+1),k}\right) - \left(\sigma^2_{(j),m+k} + \sigma^2_{(j+1),k}\right)}{\sqrt{\sigma^2_{(j-1),2m+k} + \sigma^2_{(j),m+k} + \sigma^2_{(j+1),k}}\left(\sigma_{(j-1),2m+k} + \sigma_{(j),m+k} + \sigma_{(j+1),k}\right)^2}.$$

To show $\frac{\partial \psi_k}{\partial \sigma_{(j-1),2m+k}} \le 0$, one has

$$\begin{aligned} &\sigma_{(j-1),2m+k}\left(\sigma_{(j),m+k} + \sigma_{(j+1),k}\right) - \left(\sigma^2_{(j),m+k} + \sigma^2_{(j+1),k}\right) \le 0 \\ \iff &\left(\frac{\sigma_{(j),m+k}}{\sigma_{(j-1),2m+k}} - 1\right)\sigma_{(j),m+k} + \left(\frac{\sigma_{(j+1),k}}{\sigma_{(j-1),2m+k}} - 1\right)\sigma_{(j+1),k} \ge 0. \end{aligned}$$

Notice that the standard deviations are always positive for $1 \le k \le m$. Based on the chain rule, if we want to show that ψ_k is increasing in k, then we need the following sufficient condition

$$\begin{cases} \frac{\sigma_{(j),m+k}}{\sigma_{(j-1),2m+k}} \ge 1, \\ \frac{\sigma_{(j+1),k}}{\sigma_{(j-1),2m+k}} \ge 1, 1 \le k \le m. \end{cases}$$

Similarly, for $\frac{\partial \psi_k}{\partial \sigma_{(j+1),k}}$, ψ_k is increasing in k if $\frac{\partial \psi_k}{\partial \sigma_{(j+1),k}} \ge 0$. Then, we need the following sufficient condition

$$\begin{cases} \frac{\sigma_{(j+1),k}}{\sigma_{(j-1),2m+k}} \ge 1, \\ \frac{\sigma_{(j+1),k}}{\sigma_{(j),m+k}} \ge 1, 1 \le k \le m. \end{cases}$$

Hence, if we have $\sigma_{(j+1),k} \ge \sigma_{(j),m+k} \ge \sigma_{(j-1),2m+k}$ $(1 \le k \le m)$, then ψ_k is an increasing function of k. Based on the monotonicity of $\sigma_{(j+1),k}$ and $\sigma_{(j-1),2m+k}$, we can obtain the following sufficient condition

$$\begin{cases} \frac{2a_{(j+1)}\theta_{(j+1)} + 3a_{(j+1)}}{6(a_{(j)}\theta_{(j)}m^2 + a_{(j)}m)} \ge 1, \\ \frac{a_{(j)}\theta_{(j)}m^2 + a_{(j)}m}{a_{(j-1)}\theta_{(j-1)}(m^2 - 1/3) + a_{(j-1)}(m - 1/2)} \ge 1. \end{cases}$$

In this case, the relative pooling benefit Ψ°_k is monotonically decreasing.

Case II: $\psi_k \geq \psi_{k+1}$. Similarly, we investigate the condition under which ψ_k is a decreasing function of k. Following the same procedure in Case I, we have the following sufficient condition

$$\begin{cases} \frac{\sigma_{(j),m+k}}{\sigma_{(j-1),2m+k}} \leq 1, \\ \frac{\sigma_{(j+1),k}}{\sigma_{(j-1),2m+k}} \leq 1, 1 \leq k \leq m, \end{cases} \quad \text{and} \quad \begin{cases} \frac{\sigma_{(j+1),k}}{\sigma_{(j-1),2m+k}} \leq 1, \\ \frac{\sigma_{(j+1),k}}{\sigma_{(j),m+k}} \leq 1, 1 \leq k \leq m. \end{cases}$$

Hence, if we have $\sigma_{(j-1),2m+k} \geq \sigma_{(j),m+k} \geq \sigma_{(j+1),k}$ $(1 \leq k \leq m)$, then ψ_k is a decreasing function of k. Based on the monotonicity of $\sigma_{(j+1),k}$ and $\sigma_{(j-1),2m+k}$, we can get the following sufficient condition

$$\begin{cases} \frac{a_{(j-1)}\theta_{(j-1)}(m-1/3)+\frac{1}{2}a_{(j-1)}}{a_{(j)}\theta_{(j)}m^2+a_{(j)}m} \geq 1, \\ \frac{a_{(j)}\theta_{(j)}m^2+a_{(j)}m}{a_{(j+1)}\theta_{(j+1)}(m^2-m+1/3)+a_{(j+1)}(m-1/2)} \geq 1. \end{cases}$$

In this case, the relative pooling benefit Ψ_k^o is monotonically increasing. $\square$