Spare parts recommendation for corrective maintenance of capital goods considering demand dependency

Dursun, I.; Grishina, Anastasiia; Akcay, Alp; van Houtum, Geert Jan

DOI
10.1016/j.ejor.2024.04.024

Publication date
2024

Document Version
Final published version

Published in
European Journal of Operational Research

Citation (APA)
Dursun, İ., Grishina, A., Akcay, A., & van Houtum, G. J. (2024). Spare parts recommendation for corrective maintenance of capital goods considering demand dependency. *European Journal of Operational Research*, *318*(1), 71-86. https://doi.org/10.1016/j.ejor.2024.04.024

European Journal of Operational Research 318 (2024) 71–86

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/eor

Production, Manufacturing, Transportation and Logistics

Spare parts recommendation for corrective maintenance of capital goods considering demand dependency

İpek Dursun [a,b], Anastasiia Grishina [c,1], Alp Akcay [a,*], Geert-Jan van Houtum [a]

[a] Eindhoven University of Technology, Department of Industrial Engineering and Innovation Sciences, P.O. Box 513, 5600 MB, Eindhoven, The Netherlands
[b] Delft University of Technology, Faculty of Technology, Policy and Management, Jaffalaan 5, 2628 BX, Delft, The Netherlands
[c] Simula Research Laboratory, Department of Data-Driven Software Engineering, Kristian Augusts gate 23, 0164, Oslo, Norway

ARTICLE INFO

Keywords:
Maintenance
Repair kit problem
Service control tower
Spare part recommendation

ABSTRACT

We consider a maintenance service provider that services geographically dispersed customers with its local service engineers. Traditionally, when a system failure is reported, a service engineer makes a diagnostic visit to the customer's location to perform corrective maintenance. If spare parts are required, they are ordered and a second visit is scheduled at a later date to complete the corrective maintenance. In this paper, the service provider can proactively send spare parts to the customer to avoid the costly second visit. Motivated by a real-world problem in the high-tech industry, our model considers the cost of a second visit, fixed shipment costs, retrieval costs for the parts that are sent to the customer, and send-back costs for the parts that are sent but not used for corrective maintenance. The uncertainty in the set of parts required for corrective maintenance is modeled with a general distribution that can capture the dependencies between demands for different spare parts. We formulate an integer linear program to find the optimal set of spare parts that minimizes the expected total cost. We derive analytical results for the structure of the optimal policy and compare the optimal policy with two benchmark policies from practice. We observe that the policies from practice often find the optimal policy, and a new heuristic policy that exploits the structure of the optimal policy, on average, performs better than the benchmark policies.

1. Introduction

High system availability is required for many complex capital goods. In order to keep the systems up and running, it is key to have a timely access to maintenance services. For complex high-tech systems (e.g., wafer steppers, MRI scanners, industrial printers), it is common that the manufacturers who design and build these systems also provide maintenance service and spare part support. We refer to them as *service providers* in the rest of this paper. A service provider may have many different customers at geographically dispersed locations. In order to manage such a network of after-sales operations smoothly, the concept of service control towers (SCT) has been developed (Song, van Houtum, & Van Mieghem, 2020; Topan, Eruguz, Ma, van der Heijden, & Dekker, 2020). The data collected by the SCT from different systems can be centrally processed in real-time for managing the resources (e.g., spare parts, engineers) needed for on-site maintenance.

Our work has been motivated by a collaboration with a global manufacturer of high-tech equipment, who is also responsible for providing maintenance services to a large customer base as the service provider. Traditionally, upon a system failure, the customer reports it to the service provider and a corrective maintenance case is created. A local service engineer then makes a so-called diagnostic visit to the customer's location to resolve the maintenance case. If a spare part turns out to be necessary but it is not readily available on site during the diagnostic visit, the service engineer must make a second visit at a later date. The second visit can be very costly due to the additional downtime caused by not resolving the maintenance case right away. For expensive capital equipment operated at high capacity (e.g., lithography machines in semiconductor fabs, imaging systems in hospitals), even a few hours of downtime can cause serious problems. Therefore, service providers invest in technologies that enable them, when there is a failure at a customer, to match the relevant data (e.g., sensor data related to the critical components remotely collected from various modules in the system) with the data from the past corrective maintenance cases to predict which spare parts may be needed. In our work, we consider a service provider equipped with such technologies. Sending spare parts proactively is especially common for complex high-tech equipment

* Corresponding author.
E-mail addresses: anastasiia@simula.no (A. Grishina), a.e.akcay@tue.nl (A. Akcay), g.j.v.houtum@tue.nl (G. van Houtum).
[1] Research conducted at Eindhoven University of Technology.

https://doi.org/10.1016/j.ejor.2024.04.024
Received 31 December 2022; Accepted 25 April 2024
Available online 30 April 2024

due to high downtime costs for them enforced by strict service level agreements.

We refer to distinct spare parts as Stock Keeping Units (SKUs). We assume that all SKUs are kept in a central warehouse. The service provider does not exactly know which specific SKUs (if any) will be needed until the diagnostic visit is performed, but can proactively send spare parts to the failed system to avoid the costly second visit. The service provider can choose to send too many SKUs from the central warehouse to the failed system, but this may lead to unnecessary shipment costs as well as SKU-specific retrieval costs and send-back costs for unused SKUs. On the other hand, sending too few SKUs may delay resolving the corrective maintenance case and requires a costly second visit. The objective is to determine the set of SKUs that will be sent to the failed system before the diagnostic visit that minimizes the expected total cost.

In the remainder of the paper, we refer to the probabilities over the possible *sets of SKUs* needed for the corrective maintenance as the demand distribution. We emphasize that the demand distribution is specific to a corrective maintenance case, and in practice, it can be built by using the historical data on related cases. For example, Grishina, Stolikj, Gao, and Petkovic (2020) already built a data-driven prediction model that generates a list of SKU sets that may be required to be able to close the maintenance case during the diagnostic visit along with the likelihood of each SKU set being the true set of SKUs required. To generate these likelihoods, the prediction model takes a text query (e.g., failure code, description of the failure by the customer), retrieves the related maintenance cases from the database of resolved cases based on a similarity metric, and calculates the percentage of the maintenance cases for each set of SKUs required in the past. The output of such a prediction model (i.e., the relative frequencies of the sets of SKUs required) can be used as the demand distribution in our model. The way we model the demand distribution allows us to capture the dependencies between the demands for different SKUs, which is one of the main novel features of our model (see Section 2 for further details).

The contributions of our paper can be summarized as follows:

- We formulate an Integer Linear Programming (ILP) model to optimize the selection of spare parts that will be sent to a failed system before the diagnostic visit of a field service engineer.
- We derive the optimal policy structure for problem instances with one or two SKUs, and we obtain analytical properties on the structure of the optimal policy for the problem instances with a general number of SKUs.
- We compare the optimal policy against two practically motivated benchmark policies: ‘send nothing’ (Policy 1) and ‘send a fixed number of SKUs with the highest demand probabilities’ (Policy 2). Furthermore, we propose a new policy (Policy 3) that is capable of exploiting the analytical properties on the structure of the optimal policy.
- Our numerical analysis on realistic problem instances shows that Policy 1, which reflects the traditional industry practice, is generally optimal when the marginal demand probabilities are similar for different SKUs and the cost of a second visit is relatively low. However, Policy 1 can be far from the optimal policy in other situations. We observe that Policy 2 is, on average, 12.2% and 17.8% costlier than the optimal policy in two sets of experiments, while these values are only 6.2% and 10.0% for Policy 3 in the same experiments.

The remainder of the paper is organized as follows. We discuss the relevant literature in Section 2, and provide a detailed problem description in Section 3. Section 4 presents the ILP formulation, Section 5 presents the analytical results for our structural analysis of the optimal policy, and Section 6 describes the details of the benchmark policies and their practical motivation. Section 7 presents our numerical experiments and insights, and Section 8 concludes the paper.

2. Literature review

The problem of which spare parts to choose to send in advance of an on-site visit is related to the so-called Repair Kit Problem (RKP) in the literature (Teunter, 2006), where a repair kit refers to the set of spare parts taken by service engineers with themselves to perform maintenance. In both problems, before the service engineer visits the failed equipment, it is not precisely known which spare parts are needed to be able to repair the equipment.

In the RKP, if all required parts are present in the repair kit, the system can be repaired immediately, otherwise, a follow-up action is needed. The RKP was first introduced by Smith, Chambers, and Shlifer (1980) to optimize multi-item inventories necessary for repairing field equipment by considering holding costs and the probability of job completion (i.e., repairing a failed equipment). The main trade-off is between the cost of holding parts in the repair kit and the cost of not meeting the required service levels. There are two kinds of models in the literature to model this trade-off: cost models (introduced by Smith et al. 1980) and service models (introduced by Graves 1982). In cost models, the holding cost is minimized, while not completing a job during the first time is penalized with a cost (see Bijvank, Koole, and Vis 2010, Mamer and Shogan 1987, Mamer and Smith 1982, Neves-Moreira, Veldman, and Teunter 2021, Saccani, Visintin, Mansini, and Colombi 2017, Smith et al. 1980, Teunter 2006). In service models, the holding cost is minimized subject to a service level constraint (see Bijvank et al. 2010, Graves 1982, Heeremans and Gelders 1995, Mamer and Shogan 1987, Prak, Saccani, Syntetos, Teunter, and Visintin 2017, Rippe and Kiesmüller 2023a, Teunter 2006). The first papers in the literature assume that a repair kit is used for a single job (Graves, 1982; Mamer & Shogan, 1987; Mamer & Smith, 1982; Smith et al., 1980). Heeremans and Gelders (1995) are the first to relax the single job assumption by introducing a multi-job model. In a multi-job model, multiple on-site visits can be done with the same repair kit (see Bijvank et al. 2010, Heeremans and Gelders 1995, Neves-Moreira et al. 2021, Prak et al. 2017, Rippe and Kiesmüller 2023a, Saccani et al. 2017, Teunter 2006).

The early papers assume that at most one unit from each stock-keeping unit (SKU) is needed during a single job (Graves, 1982; Heeremans & Gelders, 1995; Smith et al., 1980). Mamer and Smith (1982) relax this assumption by introducing a multi-unit model. (Mamer & Smith, 1982) are also the first to relax the assumption of independency between the failure behavior of different SKUs by defining representative job types. When the failure behavior of SKUs depends on each other, demand for these SKUs during a corrective maintenance visit is also dependent.

Similar to Mamer and Shogan (1987) and Teunter (2006), we also assume demand dependency between SKUs. Different from (Teunter, 2006), Mamer and Shogan (1987), Mamer and Smith (1982), we do not specifically define representative job types, but explicitly model the probability for each possible part combination that can appear for a maintenance case. This approach allows us to capture the likelihood of all possible relationships between the demands for each distinct part, which we refer to as the full dependency.

Our problem can be considered as an extension of the RKP. In our problem, field service engineers travel to a failed system without any spare parts, but the parts can be proactively sent from a central warehouse to make them ready at the moment of the diagnostic visit. At a central level, the maintenance service provider (e.g., an SCT that collects data from all the equipment in the field) decides which parts will be sent to the customer by using Advanced Demand Information (ADI), for which we assume a *general* probability distribution that can be obtained from the historical data on similar cases. Similar to our problem, Rippe and Kiesmüller (2023a, 2023b) study the RKP with ADI. The source of ADI in Rippe and Kiesmüller (2023a) is the sensors that monitor the condition of a subset of parts. On the other hand, Rippe and Kiesmüller (2023b) use the error codes provided by customers

Table 1
An overview of the related RKP literature.

	Smith et al. (1980)	Graves (1982)	Mamer and Smith (1982)	Mamer and Shogan (1987)	Heeremans and Gelders (1995)	Teunter (2006)	Bijvank et al. (2010)	Saccani et al. (2017)	Prak et al. (2017)	Neves-Moreira et al. (2021)	Rippe and Kiesmüller (2023a)	Rippe and Kiesmüller (2023b)	Current work
Number of jobs per tour													
Single	X	X	X	X								X	X
Multiple					X	X	X	X	X	X	X	X	
Number of units needed per SKU per job													
Single	X	X			X							X	X
Multiple			X	X		X	X	X	X	X	X		
Demand dependency between different SKUs													
No dependency	X	X			X		X	X	X	X	X	X	
Dependency via job types			X	X		X							
Full dependency													X
Model characteristic													
Service model		X		X	X	X	X		X		X		X
Cost model	X		X	X		X	X	X		X		X	X
Additional features													
Retrieval cost per item									X				X
Fixed transportation cost								X					X
Sending back cost per item												X	X
Advanced demand information per corrective maintenance													
Probability per SKU											X	X	
Probability per any set of SKUs													X

to describe the failures as the source of ADI. Similar to our work, both (Rippe & Kiesmüller, 2023a) and Rippe and Kiesmüller (2023b) consider the probability of having a demand for an SKU as the ADI in their model. However, they do not consider dependency between SKU demands.

A key characteristic of our model is to consider the retrieval cost, sending back cost of SKUs, and fixed transportation cost. However, previous work which explicitly models these costs is limited. The ordering cost per item (referred to as retrieval cost in our model) is considered by Prak et al. (2017), fixed transportation cost is considered by Saccani et al. (2017), and sending back cost is considered by Rippe and Kiesmüller (2023b). Table 1 provides an overview of the related RKP literature and the positioning of our paper within this literature.

Finally, we note that our work is related to the literature on the multi-product newsvendor problem in inventory control, where most of the early work is on developing exact or heuristic solution approaches under budget or space constraints, e.g., Abdel-Malek and Areeratchakul (2007), Abdel-Malek, Montanari, and Morales (2004), Moon and Silver (2000), Nahmias and Schmidt (1984). We refer the reader to Turken, Tan, Vakharia, Wang, Wang et al. (2012) for a review of these approaches. The literature on multi-product newsvendor problems has expanded in many directions over the years. Some examples include modeling the risk-sensitiveness of the decision makers (e.g., Choi and Ruszczyński 2011, Özler, Tan, and Karaesmen 2009), substitution between different products (e.g., Dutta and Chakraborty 2010, Lei, Ru, Shi, and Zhang 2022), advanced demand information (e.g., Bernstein and DeCroix 2015, Kuthambalayan, Mehta, and Shanker 2015), and uncertainty in the multivariate demand distribution (e.g., Olivares-Nadal 2024, Wang, Xiao, and Luo 2023). The spare-part selection problem we study in this paper is related to the multi-product newsvendor problem because in both problems there is a trade-off between the costs of ordering more than needed (i.e., overage) and ordering less than needed (i.e., underage). However, our problem is different from several angles. First of all, in case of an underage (i.e., when an SKU is not proactively sent to the failed system but that SKU turns out to be necessary for corrective maintenance), the underage cost has the cost of a second visit as a fixed setup cost in addition to typical part-specific shortage costs. Also, we assume the most general demand structure with demand as a binary vector and our demand model is capable of capturing any feasible covariance structure between different SKU demands. In the multi-product newsvendor literature, it is common to adopt continuous demand assumptions or special forms of dependence structures to generate analytical results.

3. Problem description

We consider a maintenance service provider who is fully responsible for addressing the failures encountered in the systems operated by its customers. Suppose that a customer reports a system failure to the service provider on day t. At this moment, the service provider creates a corrective-maintenance case in its case-management system and determines the spare parts that have been used in similar maintenance cases in the past. Suppose that in total N different SKUs are demanded in matching cases. This means that the system failure can be caused by the (possibly joint) failure of N different SKUs, and the replacement of up to N SKUs may be needed to fix the system failure. We let $I = \{1, 2, \dots, N\}$ denote the set of SKUs. We assume that at most one part is needed for each SKU to successfully complete the corrective maintenance of the system. The main challenge in practice is not knowing which SKUs will require replacement during the corrective maintenance. The SKUs that require replacement are precisely known only after a physical examination of the system during the diagnostic visit. In the most general case, there are 2^N possible scenarios for the set of SKUs required for corrective maintenance. Let $M = \{0, 1, \dots, 2^N - 1\}$ denote the indices of these scenarios. Specifically, $\mathbf{s}_m = (s_{1m}, \dots, s_{im}, \dots, s_{Nm})$ denotes the binary vector indicating the SKUs required in scenario $m \in M$, where $s_{im} = 1$ denotes that SKU i is in the set of SKUs required in scenario m, and $s_{im} = 0$ denotes otherwise. The indices of the scenarios are ordered such that they represent a situation where the binary vectors $\mathbf{s}_m$ are ordered lexicographically. If $N = 2$, for example, then $\mathbf{s}_0 = (0, 0)$, $\mathbf{s}_1 = (0, 1)$, $\mathbf{s}_2 = (1, 0)$, and $\mathbf{s}_3 = (1, 1)$.

Remark 1. Our model is capable of capturing a situation where not just one but multiple units of a specific SKU are needed for corrective maintenance by incorporating the additional units as distinct SKUs. For example, suppose that there are 2 SKUs in total, and 2 units of SKU 1 and 1 unit of SKU 2 may be needed for corrective maintenance. We can model the second unit of SKU 1 as a third SKU: $(0, 0, 0)$ indicates no parts are needed, $(1, 0, 0)$ indicates 1 unit of SKU 1 is needed, $(0, 1, 0)$ indicates 1 unit of SKU 2 is needed, $(1, 1, 0)$ indicates 1 unit from each SKU is needed, $(1, 0, 1)$ indicates 2 units of SKU 1 are needed, and $(1, 1, 1)$ means 2 units of SKU 1 and 1 unit of SKU 2 are needed. In the remainder of the paper, we continue with the assumption that at most 1 unit of an SKU is needed for corrective maintenance.

We let $\hat{p}_m$ denote the probability that the set of SKUs that require replacement during corrective maintenance is represented by scenario m. By explicitly modeling the probabilities $\{\hat{p}_m\}_{m \in M}$, we can capture the

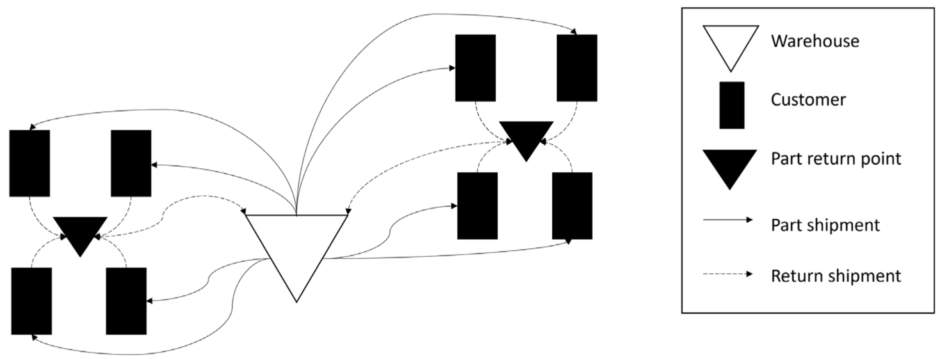


Fig. 1. Illustration of the spare parts shipment process for corrective maintenance.

most general distribution that represents the SKU demands, allowing us to consider any possible dependency between SKU demands. In practice, the service provider may have access to various degrees of information about the failures (via the analysis of data on previous maintenance cases as mentioned in Section 1) to determine a specific distribution for each particular corrective maintenance case.

All spare parts are stored at a central warehouse. The service provider can choose to send a set of SKUs from the central warehouse to the failed system in advance of the diagnostic visit (note that this can be an empty set, meaning that no spare part is sent to the customer). We assume that the required SKUs are always in stock and that no spare parts are kept on site. Retrieving one spare part for SKU i from the warehouse and including it in the shipment to the customer incurs a cost r_i (€/unit) (> 0). After the system failure occurs on day t, the spare parts are shipped to the customer overnight as one batch with a fixed transportation cost F (€) (> 0), and they become available on the site at the beginning of day $t+1$. On day $t+1$, the diagnostic visit of the field service engineer also takes place. At that moment, the problem is identified and the failed parts are replaced with the corresponding spare parts.

If all required spare parts are already sent to the site (or it turns out the failure is not because of a failed component and thus no spare part is needed), the corrective maintenance case is completed on day $t+1$. It is possible that not all of the parts sent to the site are needed on day $t+1$. The engineer collects and returns the unused parts to a local return point (see Fig. 1 for an illustration). These parts are later sent back to the central warehouse together with the old parts that come out of the other failed systems in the region. For unused parts, an additional cost arises related to checking and possibly re-packing them. Also, the shipment of a part from the warehouse decreases the availability of that SKU by one unit until it is returned back to the warehouse. We introduce the cost parameter b_i (€/unit) (> 0) to capture these aspects related to re-packing and temporary unavailability of unused parts for SKU i.

In case the system cannot be repaired during the diagnostic visit because not all the required spare parts are available, the missing parts are ordered on the same day (i.e., day $t+1$) based on the diagnosis of the service engineer. Similar to earlier, a variable cost of r_i for sending a spare part of SKU i from the warehouse to the customer, and the fixed transportation cost F must be charged. In addition, a second visit must be performed by the service engineer for resolving the case on the next day with the correct spare parts available on site. The cost of the second visit, which we denote by D (€) (> 0), includes the operational costs to arrange the visit and any penalty for the additional system downtime. On day $t+2$, the corrective maintenance is completed with the second visit. The problem is to determine the optimal set of SKUs to make available at the customer site during the diagnostic visit by minimizing the expected total cost. Fig. 2 summarizes the order of events in a corrective maintenance case.

4. ILP formulation

Decision variables

We let $\mathbf{x} = (x_1, \dots, x_i, \dots, x_N)$ denote the decision variables that indicate whether a spare part from a particular SKU is made available for the service engineer during the diagnostic visit. Specifically, $x_i = 1$ denotes that a spare part from SKU i is selected to be sent to the customer site so that it can be used during the diagnostic visit of the service engineer, and $x_i = 0$ otherwise. We introduce the binary variable z to indicate whether at least one part is sent, i.e., $z = 1$ means at least one SKU is chosen to make its spare part available during the diagnostic visit of the service engineer, and $z = 0$ means no SKU is chosen. We use $z \geq x_i, \forall i \in I$ as constraint (2) for the relation between decision variables z and x_i.

To describe the costs related to a possible second visit, we consider all possible sets $m \in M$ that form the set of parts that is needed to resolve the maintenance case. Here, we only consider sets of SKUs $m \in M$ with a strictly positive probability $\hat{p}_m$, and we denote these sets with $M' \subseteq M$. This reduces the number of variables and hence the computation time when solving the ILP. The vector $\mathbf{s}_m = (s_{1m}, \dots, s_{Nm})$ defines a set of spare parts that can be shipped during the diagnostic visit. If part $i \in I$ is needed for corrective maintenance (i.e. $s_{im} = 1$) but it is not sent for the diagnostic visit (i.e. $x_i = 0$), then that part i will be needed for the second visit. We let u_{im} denote a binary variable that is equal to 1 when SKU i is sent for the second visit, and 0 otherwise. We require $u_{im} \geq s_{im} - x_i$ because a part should be sent either in the first or in the second visit if it is needed. To be specific, if $s_{im} = 1$ and $x_i = 0$, then u_{im} will be forced to be 1. For example, suppose that $N = 2$ and a spare part from SKU 2 is not made available for the diagnostic visit. Then, $u_{21} = 1$ because we know $\mathbf{s}_1 = (0, 1)$ (i.e., SKU 2 is needed for corrective maintenance in scenario $m = 1$) but SKU 2 is not available during diagnostic visit. Finally, the variable $\hat{u}_m$ is 0 if $u_{im} = 0$ for all $i \in I$ and 1 otherwise. This variable denotes whether a second visit is needed if set m is the true set of parts. For the same example where SKU 2 is not made available for the diagnostic visit, $u_1 = 1$ and $u_3 = 1$ because $s_1 = (0, 1)$ and $s_3 = (1, 1)$ (i.e. SKU 2 is needed for corrective maintenance in scenarios $m = 1, 3$).

Objective function

Now, we introduce the objective function of the ILP model. The objective function has three parts. The first part, $\left(Fz + \sum_{i \in I} r_i x_i\right)$, represents the total cost related to the diagnostic visit which consists of the fixed transportation cost and variable transportation cost of parts. Let p_i denote the so-called marginal demand probability for SKU i representing the likelihood that the SKU i is needed for corrective maintenance, and it is given by

$$p_i = \sum_{m \in M} s_{im} \hat{p}_m, \ \forall i \in I. \quad (1)$$

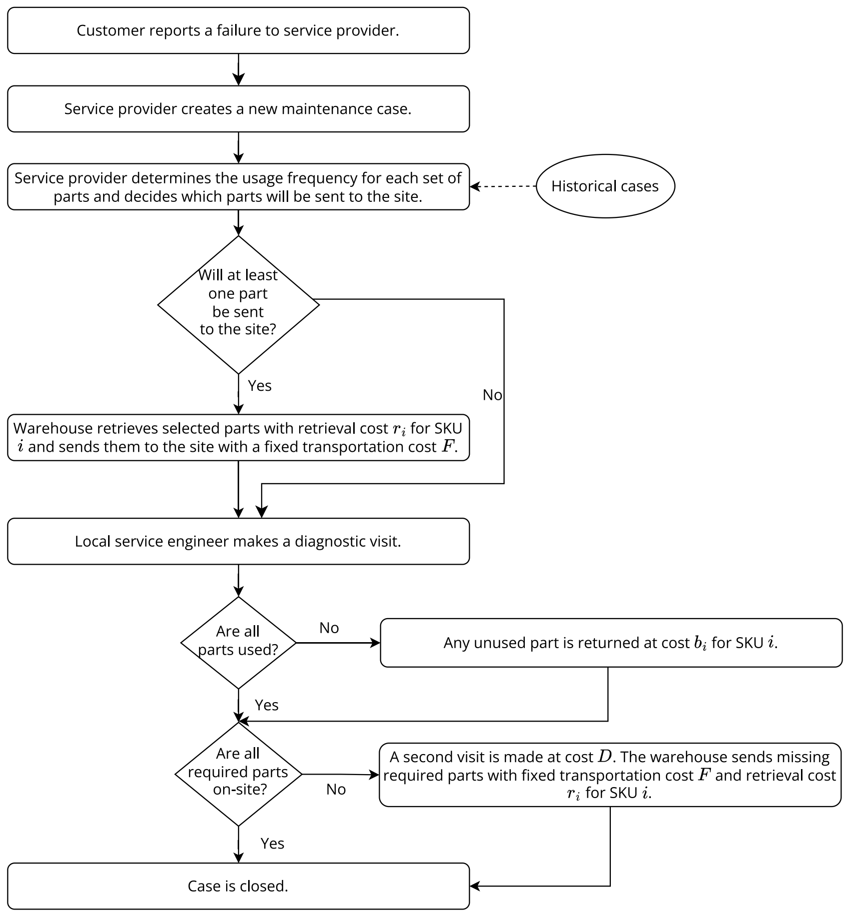


Fig. 2. Process diagram for corrective maintenance.

The second part in the objective, $\left(\sum_{i\in I} b_i x_i(1-p_i)\right)$, is the expected cost of returning the parts that are sent to the failed system but not needed. The third part, $\left((D+F)(\sum_{m\in M'} \hat{u}_m \hat{p}_m) + \sum_{i\in I} r_i(1-x_i)p_i\right)$, is the expected cost for a second engineer visit if any required parts for maintenance are not brought during the diagnostic visit. The overall objective function is

$$C(\mathbf{x}) = \left(Fz+\sum_{i\in I} r_i x_i\right)+\left(\sum_{i\in I} b_i x_i(1-p_i)\right)+\left((D+F)(\sum_{m\in M'} \hat{u}_m\hat{p}_m)+\sum_{i\in I} r_i(1-x_i)p_i\right)$$

and it can be simplified as

$$C(\mathbf{x}) = Fz + \sum_{i\in I}(r_i+b_i)(1-p_i)x_i + (D+F)(\sum_{m\in M'} \hat{u}_m\hat{p}_m) + \sum_{i\in I} r_i p_i.$$

For brevity in notation, we denote $r_i + b_i$ with c_i. We also remove the term $\sum_{i\in I} r_i p_i$ from the objective function because it is a constant. Then, the objective function is further simplified into

$$C(\mathbf{x}) = Fz + \sum_{i\in I} c_i(1-p_i)x_i + (D+F)(\sum_{m\in M'} \hat{u}_m\hat{p}_m)$$

Integer linear programming (ILP) model

The complete ILP model is summarized below:

$$\min_{\mathbf{x}} \; C(\mathbf{x}) = Fz + \sum_{i\in I} c_i(1-p_i)x_i + (D+F)(\sum_{m\in M'} \hat{u}_m\hat{p}_m)$$

s.t.

$$z \geq x_i, \forall i \in I \tag{2}$$

$$u_{im} \geq s_{im} - x_i, \forall i \in I, \forall m \in M' \tag{3}$$

$$\hat{u}_m \geq u_{im}, \forall i \in I, \forall m \in M' \tag{4}$$

$$x_i \in \{0,1\}, \forall i \in I \tag{5}$$

$$z \in \{0,1\} \tag{6}$$

$$u_{im} \in \{0,1\}, \forall i \in I, \forall m \in M' \tag{7}$$

$$\hat{u}_m \in \{0,1\}, \forall m \in M' \tag{8}$$

5. Structural analysis

In this section, we execute a structural analysis for the optimal policy. First, we present our analysis for the single SKU case ($N = 1$), then for the two-SKU case ($N = 2$), and finally for a general number of SKUs. All of the proofs can be found in Appendix A.

5.1. Optimal policy structure for $N = 1$

In the presence of a single SKU, the decision variables $\mathbf{x}$ reduce to x_1 and we denote the optimal policy with x_1^*. We derive the optimal policy structure in Proposition 1.

Proposition 1. *For $N = 1$, $x_1^* = 0$ for $p_1 \in [0, \check{p}_1]$ and $x_1^* = 1$ for $p_1 \in [\check{p}_1, 1]$, where $\check{p}_1 = \frac{F+c_1}{D+F+c_1}$.*

Proposition 1 states that the part should be sent to the system in advance of the diagnostic visit if and only if p_1, the probability that the part is needed for the corrective maintenance, exceeds a certain threshold, denoted by $\check{p}_1$ (sending the part in advance or not sending it has the same cost when $p_1 = \check{p}_1$). The optimal policy structure is shown in Fig. 3. As the cost of a second visit D increases, we see that $\check{p}_1$ approaches 0, meaning that the optimality region for sending the part in advance becomes larger. On the other hand, as the sum of

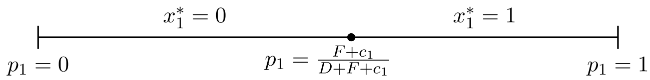


Fig. 3. Optimal policy structure for $N = 1$.

transportation cost F and cost c_1 (i.e., the cost of unnecessarily sending the part) increases, we see that the threshold value $\check{p}_1$ approaches 1, meaning that the optimality region for sending the part in advance becomes smaller.

5.2. Optimal policy structure for $N = 2$

In the presence of two SKUs, recall that there are 4 possible demand realizations: $\mathbf{s}_0 = (0, 0)$ (i.e. no parts are required for the corrective maintenance), $\mathbf{s}_1 = (0, 1)$ (i.e. only SKU 2 is required for the corrective maintenance), $\mathbf{s}_2 = (1, 0)$ (i.e. only SKU 1 is required for the corrective maintenance) and $\mathbf{s}_3 = (1, 1)$ (i.e. both SKU 1 and SKU 2 are required for the corrective maintenance), where $P(S = \mathbf{s}_m) = \hat{p}_m$ for $m \in \{0, 1, 2, 3\}$. Let $X \in \{0, 1\}$ denote the demand for SKU 1, and $Y \in \{0, 1\}$ denote the demand for SKU 2. We start our analysis by reformulating the bivariate demand distribution (which was presented in its most general form $\hat{p}_0, \dots, \hat{p}_3$ in Section 3) as an equivalent distribution that explicitly specifies the dependency between X and Y. To be specific, let $\sigma_{1,2}$ denote the covariance of the random variables X and Y:

$$\sigma_{1,2} = E[(X - E[X])(Y - E[Y])] \tag{9}$$

$$= E[XY - Xp_2 - Yp_1 + p_1p_2] \tag{10}$$

$$= E[XY] - p_1p_2 = \hat{p}_3 - p_1p_2. \tag{11}$$

where (9) is the definition of covariance, (10) follows from knowing that $E[X]$ is equal to $P(X = 1) = p_1$ and $E[Y]$ is equal to $P(Y = 1) = p_2$, and (11) follows from noting that $E[XY]$ is equal to $P(X = 1, Y = 1) = \hat{p}_3$. We observe from (11) that $\hat{p}_3$ is equal to $p_1p_2 + \sigma_{1,2}$. By using this observation, the relationships between the joint and marginal distributions of X and Y, and the fact that $\hat{p}_0 + \hat{p}_1 + \hat{p}_2 + \hat{p}_3 = 1$, it can further be verified that

$$P(S = \mathbf{s}_0) = \hat{p}_0 = (1 - p_1)(1 - p_2) + \sigma_{1,2}$$

$$P(S = \mathbf{s}_1) = \hat{p}_1 = (1 - p_1)p_2 - \sigma_{1,2}$$

$$P(S = \mathbf{s}_2) = \hat{p}_2 = p_1(1 - p_2) - \sigma_{1,2}$$

$$P(S = \mathbf{s}_3) = \hat{p}_3 = p_1p_2 + \sigma_{1,2}.$$

Given the relationship between the parameters $\hat{p}_0, \hat{p}_1, \hat{p}_2, \hat{p}_3$ and the parameters $p_1, p_2, \sigma_{1,2}$ characterized above, the requirement that the probabilities $\hat{p}_i$ must be in $[0, 1]$ for all i leads to the set of possible values for the covariance $\sigma_{1,2}$ as a function of p_1 and p_2:

$$\max\{p_1 + p_2 - p_1p_2 - 1, -p_1p_2\} \le \sigma_{1,2} \le \min\{p_2 - p_1p_2, p_1 - p_1p_2\}. \tag{12}$$

To put it in another way, for a given value of covariance $\sigma_{1,2}$, there is a feasible set of values that the probabilities p_1 and p_2 can take. Our goal is to characterize the optimal policy in the (p_1, p_2) space for a given $\sigma_{1,2}$. Definition 1 introduces a number of reference points and functions that will be used in the characterization of the optimal policy.

Definition 1.

(i) We define four points such that

- $(\underline{p}_1, \bar{p}_2) = \left(\frac{c_1}{D+F+c_1}, \frac{(F+c_2+(D+F)\sigma_{1,2})(D+F+c_1)}{(D+F+c_1)(D+F+c_2)-(D+F)c_1}\right)$,
- $(\bar{p}_1, \underline{p}_2) = \left(\frac{(F+c_1+(D+F)\sigma_{1,2})(D+F+c_2)}{(D+F+c_1)(D+F+c_2)-(D+F)c_2}, \frac{c_2}{D+F+c_2}\right)$,
- $(\tilde{p}_1, 0) = \left(\frac{F+c_1+(D+F)\sigma_{1,2}}{D+F+c_1}, 0\right)$, and
- $(0, \tilde{p}_2) = \left(0, \frac{F+c_2+(D+F)\sigma_{1,2}}{D+F+c_2}\right)$.

(ii) We define three functions of p_1 such that

- $f(p_1) = -\frac{F+c_2+(D+F)\sigma_{1,2}}{c_2+(1-p_1)(D+F)}$,
- $g(p_1) = \frac{(F+c_1+c_2+(D+F)\sigma_{1,2})-(D+F+c_1)p_1}{(D+F+c_2)-(D+F)p_1}$, and
- $h(p_1) = \frac{F+c_1+(D+F)\sigma_{1,2}-(D+F+c_1)p_1}{D+F}$.

In the remainder of this section, we will first analyze the special case where the SKUs have independent demands, i.e., $\sigma_{1,2} = 0$. Afterward, we will investigate the optimal policy structure in the general case with demand dependency.

5.2.1. Two SKUs with independent demands

Our analysis starts with Lemma 1 to show the relationship between the coordinates of the points specified in Definition 1(i).

Lemma 1. *For $\sigma_{1,2} = 0$, it holds that $0 < \underline{p}_1 < \tilde{p}_1 < \bar{p}_1$ and $0 < \underline{p}_1 < \tilde{p}_1 < 1$. Similarly, it holds that $0 < \underline{p}_2 < \tilde{p}_2 < \bar{p}_2$, and $0 < \underline{p}_2 < \tilde{p}_2 < 1$.*

Fig. 4 illustrates the points specified in Definition 1(i) on a unit square which represents the possible values of demand probabilities (p_1, p_2). These points and the functions specified in Definition 1(ii) split the unit square into four regions. Theorem 1 characterizes the optimal policy in each one of these four regions. Note that the decision variables $\mathbf{x}$ reduce to (x_1, x_2) when there are two SKUs. We denote the optimal policy with (x_1^*, x_2^*).

Theorem 1. *For two SKUs with independent demands (i.e., $\sigma_{1,2} = 0$),*

(i) $(x_1^*, x_2^*) = (0, 0)$ *for all values of* $(p_1, p_2) \in R_1$, *where the region* $R_1 = \{(p_1, p_2) \in [0, 1]^2 | p_1 \in [0, \underline{p}_1], p_2 \le f(p_1)\} \cup \{(p_1, p_2) \in [0, 1]^2 | p_1 \in [\underline{p}_1, \bar{p}_1], h(p_1) \le p_2 \le g(p_1)\}$.
(ii) $(x_1^*, x_2^*) = (0, 1)$ *for all values of* $(p_1, p_2) \in R_2$, *where the region* $R_2 = \{(p_1, p_2) \in [0, 1]^2 | p_1 \in [0, \underline{p}_1], p_2 \ge f(p_1)\}$.
(iii) $(x_1^*, x_2^*) = (1, 1)$ *for all values of* $(p_1, p_2) \in R_3$, *where the region* $R_3 = \{(p_1, p_2) \in [0, 1]^2 | p_1 \in [\underline{p}_1, \bar{p}_1], g(p_1) \le p_2\} \cup \{(p_1, p_2) \in [0, 1]^2 | p_1 \in [\bar{p}_1, 1], p_2 \ge \underline{p}_2\}$.
(iv) $(x_1^*, x_2^*) = (1, 0)$ *for all values of* $(p_1, p_2) \in R_4$, *where the region* $R_4 = \{(p_1, p_2) \in [0, 1]^2 | p_1 \in [\tilde{p}_1, \bar{p}_1], p_2 \le h(p_1)\} \cup \{(p_1, p_2) \in [0, 1]^2 | p_1 \in [\bar{p}_1, 1], p_2 \le \underline{p}_2\}$.

The four regions R_1, R_2, R_3, and R_4 that characterize the optimal policy in Theorem 1 are illustrated in Fig. 4. The shapes of these regions provide insights on how the optimal actions change as the (p_1, p_2) pair varies. For example, the optimal policy is symmetric with respect to the diagonal line $p_2 = p_1$ since the part costs are equal to each other

Fig. 5 provides further insights on how the cost parameters affect the optimal policy regions. We observe in Fig. 5(a) that an increased cost of the second visit, D, increases the size of region R_3 where $(x_1, x_2) = (1, 1)$ is optimal. In Fig. 5(b), it is shown that increasing the value of the fixed transportation cost, F, increases the size of region R_1 where $(x_1, x_2) = (0, 0)$ is optimal. It is intuitive that the optimal policy region for bringing SKU i becomes larger as the cost c_j of the SKU j increases $(j \ne i)$. The comparison of Fig. 5(c) and Fig. 5(d) shows how the optimal policy regions R_3 and R_4 (the ones where sending SKU 1 is optimal) become larger as the cost c_1 decreases from 100 to 20.

5.2.2. Two SKUs with dependent demands

Recall that Theorem 1 provided an explicit characterization of the optimal policy regions (i.e., the sets of (p_1, p_2) values where each action is optimal) under the assumption that $\sigma_{1,2} = 0$. However, an explicit characterization of the optimal policy regions is not straightforward when $\sigma_{1,2} \ne 0$ because the set of feasible (p_1, p_2) values cannot be derived in closed form for a given nonzero value of $\sigma_{1,2}$. The set of feasible (p_1, p_2) values can easily be obtained numerically, though. Fig. 6 shows the feasible set of (p_1, p_2) values for $\sigma_{1,2} = 0.05$ (left) and for $\sigma_{1,2} = 0.1$ (right). The shaded (purple color) areas in Fig. 6 denotes

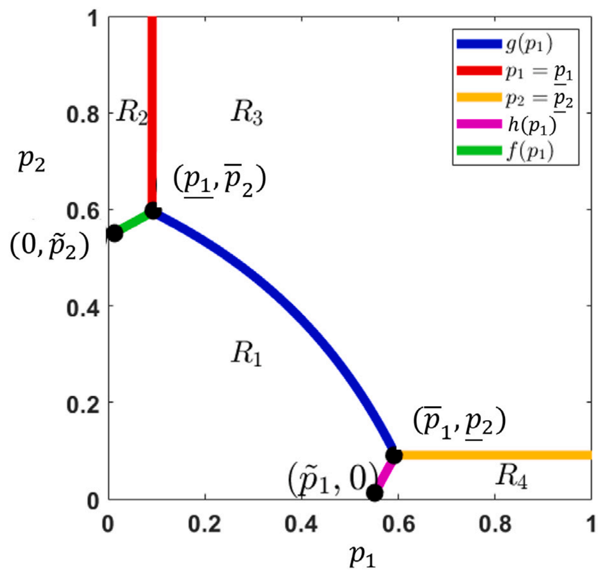


Fig. 4. Optimal policy regions for $\sigma_{1,2} = 0,\ F = 100,\ D = 100,\ c_1 = 20,\ c_2 = 20$.

the set of infeasible (p_1, p_2) values for the given $\sigma_{1,2}$ value. Note that the expressions for $\underline{p}_1$, $\underline{p}_2$, and the functions $f(\cdot)$, $g(\cdot)$, and $h(\cdot)$ from Definition 1 (which were introduced for a general $\sigma_{1,2}$ value) are used in Fig. 6 to determine the optimal actions. The roles of these expressions in determining the optimal actions follow from the arguments used in the proof of Theorem 1, so we omit the details to avoid repetition.

In Fig. 6, we observe that there are only two feasible solutions (i.e., $(x_1, x_2) = (0, 0)$ and $(x_1, x_2) = (1, 1)$). When there is a sufficiently large positive dependency between the demands of the two SKUs, it is intuitive that either both or neither of these SKUs will be needed. Furthermore, Fig. 6 shows that the switching curve between solutions $(x_1, x_2) = (1, 1)$ and $(x_1, x_2) = (0, 0)$ (i.e., the blue solid line) shifts upwards as the covariance $\sigma_{1,2}$ increases from 0.05 to 0.1. This means that the region of the (p_1, p_2) pairs where the solution $(x_1, x_2) = (1, 1)$ is optimal shrinks as the degree of positive dependency increases. That is, in this particular example, a larger covariance makes having a diagnostic visit without part shipment more preferable. Finally, by comparing the dashed blue switching curve (which was obtained under the assumption of no dependency) with the solid blue switching curve, we can understand the effect of ignoring demand dependency in decision making. For example, for all the (p_1, p_2) pairs below the dashed blue switching curve (i.e., the low values of the marginal probabilities), the optimal actions under a false assumption of independent demands are always the same as the optimal actions that would be obtained under demand dependency.

5.3. Characterization of the dominating and dominated solutions for a general N value

Suppose that we decide to send an SKU set $I' \subseteq I$ to the maintenance site for the diagnostic visit. The solution that corresponds to sending the parts in set I' is denoted with $\mathbf{x}_{I'} = (x'_1, \dots, x'_N)$, where

$$x'_i = \begin{cases} 1 & \text{if } i \in I' \\ 0 & \text{if } i \in I \backslash I'. \end{cases}$$

Any set I'' that is equal to $I' \setminus \{l\}$ with $l \in I'$ or equal to $I' \cup \{l\} \subseteq I$ with $l \in I \setminus I'$ is referred to as a neighboring set of I'.

Definition 2. A solution $\mathbf{x}_{I'}$ is dominated by another solution $\mathbf{x}_{I''}$ if $C(\mathbf{x}_{I'}) > C(\mathbf{x}_{I''})$. We refer to $\mathbf{x}_{I'}$ as the dominated solution and $\mathbf{x}_{I''}$ as the dominating solution.

In the remainder of this section, we will provide some analytical properties that can be used to determine the dominance between two neighboring solutions. Note that the probability that a second visit will be required under the solution $x_{I'}$ is $h_{I'} = \sum_{m \in M'} \hat{u}_m(I') \hat{p}_m$, where

$$\hat{u}_m(I') = \begin{cases} 0 & \text{if } I' \text{ contains all SKUs that are in } \mathbf{s}_m \\ 1 & \text{otherwise.} \end{cases}$$

Our analysis starts with establishing the monotonicity properties related to the probability of the second visit in Lemma 2.

Lemma 2. *Let $l \in I$, $I' \subset I \backslash \{l\}$. It holds that $h_{I' \cup \{l\}} \leq h_{I'}$. It also holds that the difference $h_{I'} - h_{I' \cup \{l\}}$ is a non-decreasing function of I' on $I \backslash \{l\}$, i.e., $h_{I'_1} - h_{I'_1 \cup \{l\}} \leq h_{I'_2} - h_{I'_2 \cup \{l\}}$ where $I'_1 \subset I'_2 \subseteq I \backslash \{l\}$.*

It is intuitive that the probability of a second visit decreases as the set of parts sent to the site becomes larger. Lemma 2 further shows that the difference in the probabilities of a second visit for a set I' and $I' \cup \{l\}$ is greater than or equal to the difference in the probability of a second engineer visit for a subset of I' and the union of that subset with SKU $\{l\}$. Lemma 2 is helpful to provide results for the structure of the optimal policy for a general N number of SKUs.

Remark 2. Note that $h_I = 0$ and $h_{I'} - h_I = p_i$ for any $I' = I \backslash \{i\}$. Additionally, the following closed-form representations hold when the demands of the SKUs are independent:

- $h_\emptyset - h_I = 1 - \prod_{i \in I}(1 - p_i)$.
- $h_{I''} - h_{I'} = \left(\prod_{j \in I \backslash I''}(1 - p_j)\right)$ for any set $I' \subseteq I$ and $I'' = I' \backslash \{i\}$, where $I', I'' \neq \emptyset$.
- $h_\emptyset - h_{I'} = \left(\prod_{j \in I \backslash \{i\}}(1 - p_j)\right)$ for any set $I' = \{i\}$ and $I'' = \emptyset$.

We derive some conditions in Lemma 3 that can be used to determine the dominating solution and the dominated solution in a given pair of solutions. In other words, Lemma 3 characterizes when one solution is better than the other one.

Lemma 3.

(i) Solution $\mathbf{x}_\emptyset$ dominates the solution $\mathbf{x}_I$ if and only if

$$F + \sum_{i \in I} c_i(1 - p_i) - (D + F)h_\emptyset > 0. \tag{13}$$

(ii) For any set $I' \subseteq I$ with $\{i\} \in I'$ and $I'' = I' \backslash \{i\}$ where $I', I'' \neq \emptyset$, the solution $\mathbf{x}_{I''}$ dominates the solution $\mathbf{x}_{I'}$ if

$$c_i(1 - p_i) - (D + F)(h_{I''} - h_{I'}) > 0. \tag{14}$$

In a special case with $I' = I \backslash \{i\}$, the solution $\mathbf{x}_{I'}$ dominates the solution $\mathbf{x}_I$ if

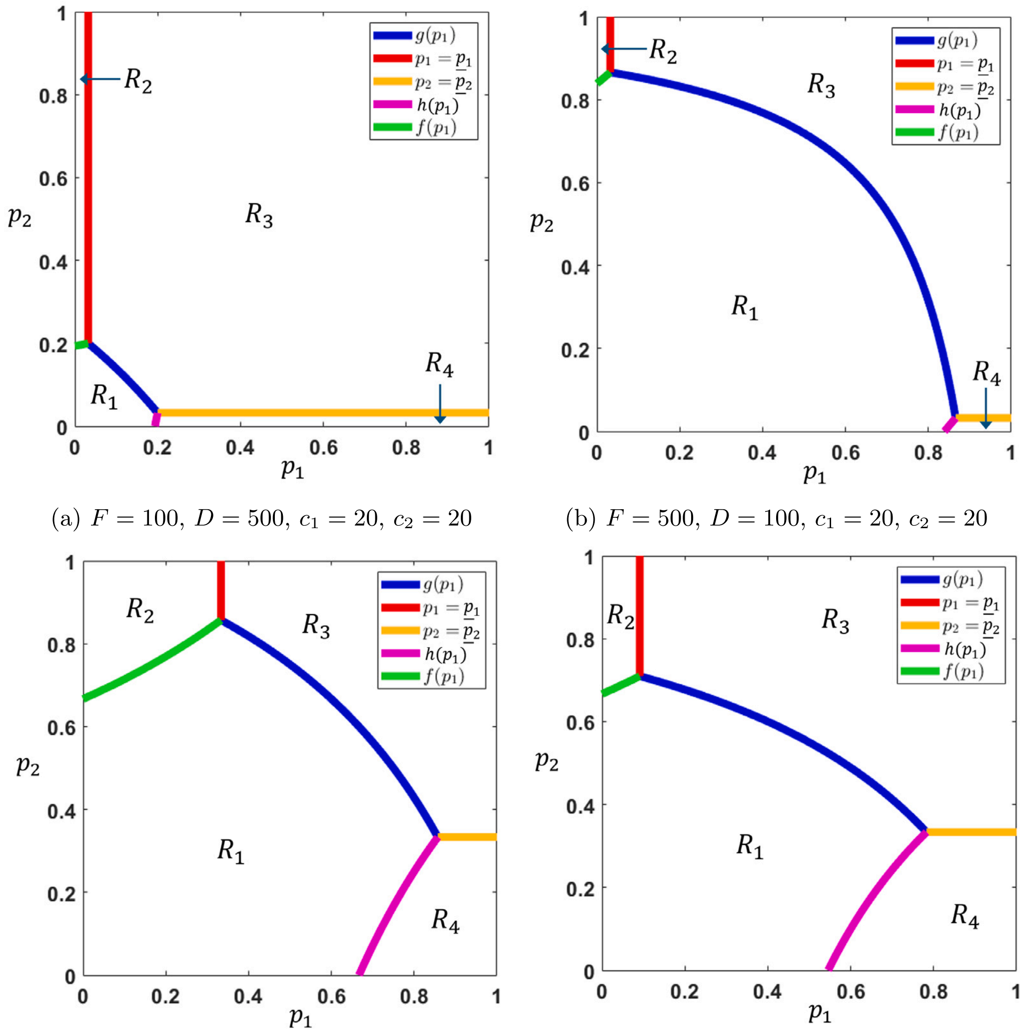
$$\frac{c_i}{D + F + c_i} > p_i. \tag{15}$$

(iii) For $I' = \{i\}$ and $I'' = \emptyset$, the solution $\mathbf{x}_{I''}$ dominates $\mathbf{x}_{I'}$ if

$$F + c_i(1 - p_i) - (D + F)(h_\emptyset - h_{I'}) > 0. \tag{16}$$

Lemma 3(i) shows that sending no parts is better than sending all parts when condition (13) is satisfied. Condition (14) in Lemma 3(ii) can be used to check whether removing a specific SKU from the solution improves the solution or not. This condition reduces to an easy-to-interpret formula (15) to check whether removing a certain SKU improves the "send-all-parts" policy. Finally, Lemma 3(iii) can be used to check whether adding a certain SKU improves the "send-no-parts" policy. The results in Lemma 3, together with the monotonicity result in Lemma 2, lead to Theorem 2.

Theorem 2. *If the solution $x_{I' \backslash \{i\}}$ is a better than the solution $x_{I'}$, then it also holds that the solution $x_{I'' \backslash \{i\}}$ is better than the solution $x_{I''}$, where $I'' \subset I'$ with $i \in I''$.*



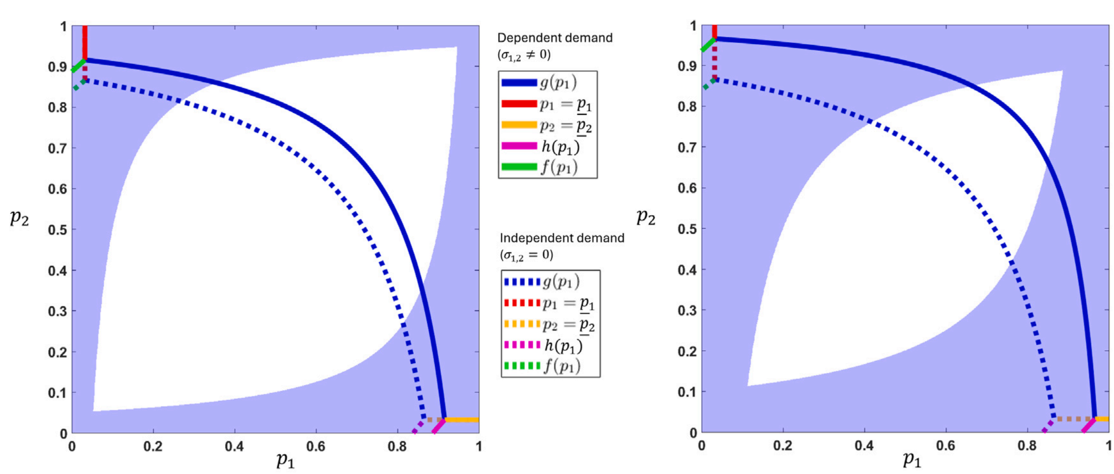
(a) $F = 100,\ D = 500,\ c_1 = 20,\ c_2 = 20$

(b) $F = 500,\ D = 100,\ c_1 = 20,\ c_2 = 20$

(c) $F = 100,\ D = 100,\ c_1 = 100,\ c_2 = 100$

(d) $F = 100,\ D = 100,\ c_1 = 20,\ c_2 = 100$

Fig. 5. Change of optimal policy regions for 2 SKUs with independent demand ($\sigma_{1,2} = 0$).



(a) $\sigma_{1,2} = 0.05$

(b) $\sigma_{1,2} = 0.1$

Fig. 6. Change of optimal policy regions for $F = 100$, $D = 100$, $c_1 = 20$, $c_2 = 20$ with dependent demand. The shaded area represents the infeasible combinations of p_1 and p_2 for the given $\sigma_{1,2}$ value. In the presence of demand dependency, $(x_1^*, x_2^*) = (0, 0)$ below the solid blue line, and $(x_1^*, x_2^*) = (1, 1)$ above the solid blue line.

Table 2
Characteristics of the demand distributions in Scenarios A, B, C and D.

	$(p_1, p_2, \dots, p_{10})$	Demand dependency	$\hat{p}_m$
A	$0.135, \forall i \in I$	Yes	0.09 for $m \in \{6, 24, 96, 384, 513\}$, 0.045 for $m \in \{1, 2, 4, 8, 16, 32, 64, 128, 256, 512\}$, 0.1 for $m = 0$, 0 otherwise
B	$0.135, \forall i \in I$	No	$\prod_{i \in I}\left(p_i s_{im} + (1-p_i)(1-s_{im})\right)$ for $m \in \{0, \dots, 2^{10}-1\}$
C	(0.5, 0.45, 0.4, 0.35, 0.3, 0.25, 0.2, 0.15, 0.1, 0.05)	Yes	0.05 for $m = 293$, 0.1 for $m \in \{390, 448\}$, 0.2 for $m = 392$, 0.25 for $m \in \{544, 592\}$, 0 otherwise
D	(0.5, 0.45, 0.4, 0.35, 0.3, 0.25, 0.2, 0.15, 0.1, 0.05)	No	$\prod_{i \in I}\left(p_i s_{im} + (1-p_i)(1-s_{im})\right)$ for $m \in \{0, \dots, 2^{10}-1\}$

According to Theorem 2, if removing the SKU i from a set I' is better than not removing it, this observation also holds for any subset of I' that contain SKU i. This result has key implications in reducing the solution space. For example, once it is confirmed that the condition (15) holds for a specific SKU i, it is for sure known that the optimal solution does not include SKU i, as stated in Corollary 1.

Corollary 1. *For an optimal solution* $\mathbf{x}^*$*, it will hold that* $x_i^* = 0$ *if* (15) *holds.*

The results in Lemma 3 and Theorem 2 can guide the decision makers to quickly determine whether a specific SKU can be eliminated from the list of candidate SKUs to be sent in advance of the diagnostic visit. Example 1 illustrates the practical use of our technical results in reducing the solution space.

Example 1. Let $N = 3$ with $I = \{1, 2, 3\}$, and suppose that $\frac{c_1}{D+F+c_1} > p_1$. Then, we know that the solution $(1, 1, 1)$ (i.e. sending the parts from all SKUs) is costlier than the solution $(0, 1, 1)$ (i.e. sending the parts only for SKU 2 and SKU 3). Since we already know that sending the set of parts $\{2, 3\}$ is better than sending $\{1, 2, 3\}$, Theorem 2 implies that sending $\{2\}$ is better than sending $\{1, 2\}$, sending $\{3\}$ is better than sending $\{1, 3\}$, and sending nothing is better than sending $\{1\}$. Thus, we can remove the solutions with $x_1 = 1$ from the list of possible optimal solutions. Suppose that the condition $c_2(1-p_2)-(D+F)(h_{I'}-h_{I''}) > 0$ also holds, where $I' = \{2, 3\}$ and $I'' = \{3\}$. That is, sending the part $\{3\}$ is better than sending $\{2, 3\}$. Theorem 2 then implies that sending nothing is better than sending $\{2\}$. Consequently, it can be concluded that the optimal solution must be one of the solutions $(0, 0, 1)$ and $(0, 0, 0)$.

In many real-world cases, the number of SKUs that represent the critical components related to a specific failure of a physical asset hardly exceeds $N = 10$, for which the ILP model presented in Section 4 can be instantly solved in an average personal computer. However, a solution approach that reduces the search space by exploiting the analytical properties presented in this section can still be useful as an alternative to the ILP model. Section 6 will introduce such a solution approach along with two other benchmark policies that are commonly used in practice.

6. Benchmark policies

In this section, we will introduce three benchmark policies as benchmark to the optimal policy. We choose these benchmark policies because they are easy to implement and can be attractive to the service providers for practical reasons as described below.

- **Policy 1**: No spare parts are sent to the failed system in advance of the on-site diagnostic visit by a service engineer. The maintenance process starts with the on-site diagnostic visit of the service engineer to determine which parts are needed for corrective maintenance. Then, the set of required spare parts (which are certainly known at that moment) is brought on-site after which the failed system is repaired in a second visit. Policy 1 is a commonly used policy in practice by service providers who do not have any spare part recommendation systems in place. The comparison of Policy 1 with the optimal policy provides insights on the potential business value of collecting historical data to build a demand distribution and proactively sending spare parts based on that distribution.
- **Policy 2**: The parts are ordered with respect to their frequency of usage and the top k parts with the highest frequency are sent to the failed system in advance of the on-site diagnostic visit. This policy is equivalent to ranking the parts with respect to their probabilities p_i specified in Eq. (1) and choosing the first k parts with the highest probability. We include this policy as a benchmark to the optimal policy as it resembles the current practice of an industry partner. The fixed number k that determines how many parts are sent for the diagnostic visit can be optimized. Note that if k is equal to zero, then Policy 2 is the same as Policy 1.
- **Policy 3**: While Policy 2 only considers the probabilities p_i, Policy 3 also takes the cost parameters into account to decide which parts to send to the failed system. It is essentially a greedy heuristic that is capable of exploiting the structural properties derived in Section 5.3. To be specific, Policy 3 decides which parts to send in three steps. In Step 1, the parts which satisfy condition (15) are eliminated. In Step 2, an ordered list of remaining parts is formed with an increasing ratio of p_i/c_i. Note that the parts early in this list with a lower p_i/c_i ratio are the ones that are less likely to be in the optimal solution, so they are better candidates to eliminate. In Step 3, the first part in the list is temporarily removed and the cost of the solution that includes the parts in this temporary list is compared with the cost of the solution that includes the parts in the original list by checking the condition (14) (or by checking the condition (16) if the temporary list does not include any part). If the condition is satisfied, the list is updated by permanently removing the part from the list, and Step 3 is repeated. If the condition is not satisfied (i.e., removing the part does not improve the solution), then the current list is returned as the solution of Policy 3.

7. Numerical analysis

Section 7.1 introduces the main test bed for our numerical analysis. Section 7.2 presents our numerical insights on the effects of demand dependency and cost parameters on the optimal policy and optimal cost. Section 7.3 compares the optimal policy with the benchmark policies introduced in Section 6 and provides insights into how the performance of the benchmark policies changes with varying levels of demand dependency and cost parameters. Note that our model assumes the demand distribution $\hat{p}_i$, $i \in M$, is known. In the real world, it is possible that the number of corrective maintenance cases may be limited, making it difficult to obtain an accurate demand distribution that adequately reflects the true demand behavior. Section 7.4 will present a numerical study that investigates the sensitivity of the policies introduced in our work with respect to the uncertainty associated with not knowing the true demand distribution.

7.1. Description of the test bed

We let $F \in \{25, 50, 100\}$ and $D \in \{100, 200, 400\}$, and consider instances with $N = 10$ SKUs. The SKU-specific cost parameters $\mathbf{c} = (c_1, \ldots, c_{10})$ are sampled from a uniform distribution on the interval [10, 25] such that $\mathbf{c} = (20.13, 17.65, 10.51, 12.87, 10.49, 10.44, 14.38, 17.3, 14.5, 24.86)$. In our experiments, our first goal is to generate insights into the effects of marginal demand distributions and the level of demand dependency on the optimal policy. So, we defined two types of marginal demand distributions (i.e., two distinct $(p_1, \ldots, p_{10})$ values) and two levels of dependency (i.e., dependent demands and independent demands), leading to four distinct scenarios. Scenarios A and B assume each SKU has the same marginal demand probability, while Scenarios C and D have marginal demand probabilities varying from 0.05 to 0.5. The case of equal marginal demand probabilities represents a situation where the service provider does not have enough information to distinguish between SKUs. The demand distribution $\hat{p}_i$, $i \in M$, is specified in Scenarios A and C such that there are SKUs that are likely to be demanded together, while Scenarios B and D do not have such dependence. Table 2 provides the details on these demand scenarios. Our second goal is to investigate how the part-specific costs affect the optimal policy. For this purpose, we will vary the cost parameter c_i for one of the SKUs to observe how the optimal policy changes.

7.2. Results on the optimal policy

The optimal solution and optimal cost for each instance are given in Table 3. In 10 out of 36 instances, the optimal policy is the same as Policy 1. In 13 out of 36 instances, the optimal policy is equal to the ‘send all parts’ solution. For the remaining instances, the optimal number of parts that should be sent for the diagnostic visit varies between seven and nine. We observe only ‘send no parts’ or ‘send all parts’ solutions under Scenarios A and B, where the marginal demand probabilities of SKUs are equal. The optimal solution is ‘send no parts’ when the second engineer visit cost D is relatively low in comparison to fixed transportation cost F, and the optimal solution is ‘send all parts’ when D is relatively high in comparison to F. When we look at Scenarios C and D, we observe that optimal solutions other than ‘send no parts’ and ‘send all parts’ are also possible. That is, when there is more variety in the marginal demand probabilities of different SKUs (i.e., p_i values are less similar to each other), optimal solutions are more likely to be heterogeneous (i.e., a mix of zeros and ones).

The comparison of Scenarios A and B and the comparison of Scenarios C and D in Table 3 provide insights into the effect of demand dependency on the optimal policy. For example, we observe that the demand dependency does not influence the optimal solution when the cost of the second on-site visit cost is high, i.e., all instances from Scenarios A and B and all instances from Scenarios C and D have the same optimal solution when D is equal to 400. For lower values of D, we generally observe that the dependency leads to having less parts in the optimal solution (e.g., instances 22–23 compared to instances 31–32), but there are also exceptions to this observation (e.g., instance 6 compared to instance 15). Furthermore, while the dependency leads to an increase in optimal costs for Scenarios A and B with identical marginal demand probabilities, it leads to a decrease in optimal costs for Scenarios C and D with varying marginal demand probabilities.

Next, we investigate the effect of part-specific costs on the optimal policy. We do this by inflating the part-specific cost of one of the SKUs while keeping everything else the same. For brevity, we only focus on Scenarios C and D with various levels of marginal demand probabilities, and increase the cost of SKU 4 to make it significantly higher than the costs of other SKUs. Table 4 presents the results.

The comparison of Tables 3 and 4 leads to some interesting insights. For example, instance 22 in Table 3 has SKU 4 in its optimal solution. In Table 3, on the other hand, we see that SKU 4 is not in the optimal set of spare parts due to the high cost of c_4. We also see that SKU 6 is not in the optimal solution due to the high demand dependency with SKU 4 and low marginal demand probability (i.e., $p_6 = 0.25$). Recall that the demand distribution in instance 22 was specified such that the demand vector s_{544} (which involves SKUs 1, 4, and 6) has probability $\hat{p}_{544} = 0.25$, so there is dependency between SKUs 1, 4, and 6 by construction. Despite the dependency of SKU 4 with also SKU 1, we observe that SKU 1 is still in the optimal solution because its marginal demand probability is relatively higher (i.e., $p_1 = 0.5$). In the instances 19, 20, and 28 of Table 3, x_i^* was equal to 1 for some SKU i. However, in Table 4, the optimal solution is ‘send no parts’ in these instances. When the transportation cost is high or there is a high part-specific cost for an SKU that has a high marginal demand probability, we observe a more conservative optimal solution (i.e., the parts are sent only after the diagnostic visit confirms which specific parts are needed).

7.3. Comparison of the optimal policy with the benchmark policies

The goal of this section is to compare the performance of the optimal policy with the benchmark policies introduced in Section 6. We denote the cost of Policy 1 by C_1, the cost of Policy 2 for a given $k \in \{1, \ldots, N\}$ (referred to as Policy 2(k)) by $C_2(k)$, the cost of Policy 3 by C_3, and the cost of the optimal policy by C_{opt}. Table 5 reports the relative differences $\Delta_1 = \frac{C_1 - C_{\text{opt}}}{C_{\text{opt}}} 100\%$, $\Delta_2(k) = \frac{C_2(k) - C_{\text{opt}}}{C_{\text{opt}}} 100\%$, and $\Delta_3 = \frac{C_3 - C_{\text{opt}}}{C_{\text{opt}}} 100\%$.

In Table 5, we observe that Policy 1 is the best policy for Scenarios A and B when the cost parameter D is the lowest (i.e, instances 1–3 and 10–12). That is, it is optimal to send no parts before the diagnostic visit when marginal demand probabilities are similar to each other and the cost of second visit is relatively lower. On the other hand, when Policy 2 finds the optimal or a close-to-optimal solution, the best k value seems to be closer to N, i.e., for the instances where Policy 2 finds the optimal solution, the value of k is at least 7. That is, for problem instances where the optimal solution includes many of the candidate parts, Policy 2 is a logical approach. We further observe in Table 5 that either Policy 1, Policy 2, or Policy 3 finds the optimal policy in each instance. However, one particular heuristic policy cannot find the optimal solution in all problem instances. On average, Policy 3 is the best benchmark policy with only 6.2% deviation from the optimal policy, while this deviation is 57.4% for Policy 1 and 12.2% for Policy 2 with the best k value. In addition, Policy 3 is better than Policy 1 in 26 (out of 36) instances, and it is better than the Policy 2 with the best k value in 23 (out of 36) instances. The superior performance of Policy 3 can be attributed to its ability to exploit the structural properties from Section 5.3.

Table 6 compares the benchmark policies with the optimal policies of the instances in Table 4 to investigate how the increase of a part-specific cost affects our insights. In general, we continue to observe that Policy 3 is the best policy on average. Policy 3 is the best benchmark policy with only 10.0% average deviation from the optimal policy, while this deviation is 30.7% for Policy 1 and 17.8% for Policy 2 with the optimal k value. Furthermore, we observe in Table 6 that neither Policy 1 nor Policy 2 finds the optimal policy in instance 31, while Policy 3 can find the optimal solution in that instance.

Table 3
Results on the optimal policy for SKU-specific costs equal to **c**.

Instance	F	D	Scenario	Optimal solution	Optimal cost
1	25	100	A	(0, 0, 0, 0, 0, 0, 0, 0, 0, 0)	112.5
2	50	100	A	(0, 0, 0, 0, 0, 0, 0, 0, 0, 0)	135.0
3	100	100	A	(0, 0, 0, 0, 0, 0, 0, 0, 0, 0)	180.0
4	25	200	A	(1, 1, 1, 1, 1, 1, 1, 1, 1, 1)	157.5
5	50	200	A	(1, 1, 1, 1, 1, 1, 1, 1, 1, 1)	182.5
6	100	200	A	(1, 1, 1, 1, 1, 1, 1, 1, 1, 1)	232.5
7	25	400	A	(1, 1, 1, 1, 1, 1, 1, 1, 1, 1)	157.5
8	50	400	A	(1, 1, 1, 1, 1, 1, 1, 1, 1, 1)	182.5
9	100	400	A	(1, 1, 1, 1, 1, 1, 1, 1, 1, 1)	232.5
10	25	100	B	(0, 0, 0, 0, 0, 0, 0, 0, 0, 0)	95.7
11	50	100	B	(0, 0, 0, 0, 0, 0, 0, 0, 0, 0)	114.8
12	100	100	B	(0, 0, 0, 0, 0, 0, 0, 0, 0, 0)	153.1
13	25	200	B	(1, 1, 1, 1, 1, 1, 1, 1, 1, 1)	157.5
14	50	200	B	(1, 1, 1, 1, 1, 1, 1, 1, 1, 1)	182.5
15	100	200	B	(0, 0, 0, 0, 0, 0, 0, 0, 0, 0)	229.6
16	25	400	B	(1, 1, 1, 1, 1, 1, 1, 1, 1, 1)	157.5
17	50	400	B	(1, 1, 1, 1, 1, 1, 1, 1, 1, 1)	182.5
18	100	400	B	(1, 1, 1, 1, 1, 1, 1, 1, 1, 1)	232.5
19	25	100	C	(1, 1, 1, 1, 1, 1, 1, 0, 0, 0)	104.9
20	50	100	C	(1, 1, 1, 1, 1, 1, 1, 0, 0, 0)	133.6
21	100	100	C	(0, 0, 0, 0, 0, 0, 0, 0, 0, 0)	190.0
22	25	200	C	(1, 1, 1, 1, 1, 1, 1, 0, 0, 0)	119.9
23	50	200	C	(1, 1, 1, 1, 1, 1, 1, 0, 0, 0)	148.6
24	100	200	C	(1, 1, 1, 1, 1, 1, 1, 1, 1, 0)	203.9
25	25	400	C	(1, 1, 1, 1, 1, 1, 1, 1, 1, 0)	135.1
26	50	400	C	(1, 1, 1, 1, 1, 1, 1, 1, 1, 0)	161.4
27	100	400	C	(1, 1, 1, 1, 1, 1, 1, 1, 1, 1)	212.5
28	25	100	D	(1, 1, 1, 1, 1, 1, 1, 1, 0, 0)	118.9
29	50	100	D	(0, 0, 0, 0, 0, 0, 0, 0, 0, 0)	145.1
30	100	100	D	(0, 0, 0, 0, 0, 0, 0, 0, 0, 0)	193.4
31	25	200	D	(1, 1, 1, 1, 1, 1, 1, 1, 1, 0)	125.1
32	50	200	D	(1, 1, 1, 1, 1, 1, 1, 1, 1, 0)	151.4
33	100	200	D	(1, 1, 1, 1, 1, 1, 1, 1, 1, 0)	203.9
34	25	400	D	(1, 1, 1, 1, 1, 1, 1, 1, 1, 0)	135.1
35	50	400	D	(1, 1, 1, 1, 1, 1, 1, 1, 1, 0)	161.4
36	100	400	D	(1, 1, 1, 1, 1, 1, 1, 1, 1, 1)	212.5

Table 4
Optimal policy when the SKU-specific costs are equal to **c** except SKU 4 with $c_4 = 121.87$ instead of 12.87.

Instance	Optimal solution	Optimal cost
19	(0, 0, 0, 0, 0, 0, 0, 0, 0, 0)	118.8
20	(0, 0, 0, 0, 0, 0, 0, 0, 0, 0)	142.5
21	(0, 0, 0, 0, 0, 0, 0, 0, 0, 0)	190.0
22	(1, 1, 1, 0, 1, 0, 1, 0, 0, 0)	183.7
23	(1, 1, 1, 1, 1, 1, 1, 0, 0, 0)	220.8
24	(1, 1, 1, 1, 1, 1, 1, 1, 1, 0)	276.0
25	(1, 1, 1, 1, 1, 1, 1, 1, 1, 0)	207.3
26	(1, 1, 1, 1, 1, 1, 1, 1, 1, 0)	233.5
27	(1, 1, 1, 1, 1, 1, 1, 1, 1, 1)	284.6
28	(0, 0, 0, 0, 0, 0, 0, 0, 0, 0)	120.9
29	(0, 0, 0, 0, 0, 0, 0, 0, 0, 0)	145.1
30	(0, 0, 0, 0, 0, 0, 0, 0, 0, 0)	193.5
31	(1, 1, 1, 0, 1, 1, 1, 1, 1, 0)	192.9
32	(1, 1, 1, 1, 1, 1, 1, 1, 1, 0)	223.5
33	(1, 1, 1, 1, 1, 1, 1, 1, 1, 0)	276.0
34	(1, 1, 1, 1, 1, 1, 1, 1, 1, 0)	207.3
35	(1, 1, 1, 1, 1, 1, 1, 1, 1, 0)	233.5
36	(1, 1, 1, 1, 1, 1, 1, 1, 1, 1)	284.6

It is a natural question to ask which policy should be adopted in practice. As mentioned at the end of Section 5.3, we choose $N = 10$ because, in our experience, it is a realistic value to consider as the number of SKUs. For highly complex capital equipment such as industrial printers and lithography systems, there are definitely more than 10 critical components, but recall that we assume the service provider is equipped with the technology that can match the current maintenance case to the previous maintenance cases. This initial analysis typically reduces the set of SKUs relevant for the current maintenance case to a subset of all critical components. Since the ILP can easily be solved in a standard personal computer for $N = 10$, the optimal policy is a natural choice to follow in practice. If an ILP solver is not available, the three benchmark policies can be implemented and the one with the best performance can be chosen for a specific instance.

7.4. Sensitivity with respect to the demand-distribution uncertainty

Recall that we argued the demand distribution $\hat{p}_m$, $m \in M$, is obtained by counting the frequencies of the part combinations that were needed to resolve the related corrective maintenance cases in the past. If there is a large number of such maintenance cases in the past, that leads to a data set that can be used to build an accurate empirical distribution of the true (but unknown) probability distribution of the demand scenarios. However, it may be possible that the number of corrective maintenance cases in the past is limited. The objective of this section is to investigate how the number of the past corrective maintenance cases, which we denote by T, affects the performance of the optimal policy.

In our experiments, we assume that the demand distributions specified in Table 2 are the true demand distributions and they are unknown by the service provider. The service provider uses the past maintenance cases to build an empirical demand distribution (i.e., the frequencies of each part combination in the historical maintenance cases), and

Table 5
Comparison of the optimal policy against Policies 1, 2, and 3 for SKU-specific costs equal to **c**.

Inst.	Δ_1	$\Delta_2(k)$										Δ_3
		$k=1$	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$	$k=9$	$k=10$	
1	**0.0%**	32.7%	41.2%	34.4%	39.3%	32.4%	35.4%	31.4%	39.7%	35.8%	40.0%	26.8%
2	**0.0%**	45.0%	51.3%	43.0%	46.2%	37.9%	39.6%	33.9%	39.9%	34.2%	35.2%	27.5%
3	**0.0%**	60.2%	63.7%	53.8%	54.9%	45.0%	45.0%	36.9%	40.2%	32.2%	29.2%	30.0%
4	28.6%	49.1%	52.4%	38.9%	39.5%	26.0%	25.3%	13.9%	17.0%	5.7%	**0.0%**	0.0%
5	23.3%	54.1%	56.3%	42.8%	42.7%	29.2%	28.0%	16.3%	18.4%	6.7%	**0.0%**	0.0%
6	16.1%	60.9%	61.6%	48.1%	47.1%	33.5%	31.7%	19.6%	20.2%	8.2%	**0.0%**	0.0%
7	142.9%	157.7%	155.2%	124.6%	119.5%	88.8%	82.4%	53.9%	51.2%	22.8%	**0.0%**	0.0%
8	121.9%	147.8%	145.0%	116.8%	111.8%	83.5%	77.3%	50.8%	47.9%	21.5%	**0.0%**	0.0%
9	93.5%	134.4%	131.3%	106.2%	101.3%	76.1%	70.4%	46.7%	43.4%	19.8%	**0.0%**	0.0%
10	**0.0%**	39.5%	49.9%	53.1%	57.4%	58.3%	57.8%	59.5%	61.9%	59.7%	64.6%	0.0%
11	**0.0%**	53.9%	61.7%	63.2%	65.6%	64.9%	62.9%	62.3%	62.2%	57.8%	58.9%	59.2%
12	**0.0%**	71.9%	76.4%	75.9%	75.8%	73.2%	69.2%	65.9%	62.5%	55.5%	51.9%	52.9%
13	9.4%	31.1%	34.7%	33.5%	32.5%	29.0%	23.9%	19.3%	14.3%	5.7%	**0.0%**	0.0%
14	4.9%	36.8%	39.4%	37.6%	36.0%	32.1%	26.6%	21.5%	15.8%	6.7%	**0.0%**	0.0%
15	**0.0%**	46.3%	47.5%	45.0%	42.5%	37.9%	32.0%	26.0%	19.3%	9.5%	1.2%	1.9%
16	106.6%	123.6%	121.9%	114.5%	106.3%	94.4%	79.7%	64.1%	46.3%	22.8%	**0.0%**	0.0%
17	88.8%	116.7%	114.6%	107.6%	99.7%	88.5%	74.8%	60.1%	43.5%	21.5%	**0.0%**	0.0%
18	64.6%	107.3%	104.7%	98.1%	90.8%	80.6%	68.2%	54.8%	39.5%	19.8%	**0.0%**	0.0%
19	13.3%	46.6%	56.0%	62.0%	58.0%	35.2%	12.9%	**0.0%**	14.0%	14.6%	31.1%	0.0%
20	6.7%	51.6%	58.9%	63.6%	58.6%	36.1%	13.8%	**0.0%**	11.0%	9.6%	21.6%	9.5%
21	**0.0%**	57.9%	63.1%	66.4%	60.2%	37.8%	15.6%	0.6%	8.3%	4.7%	11.8%	5.3%
22	78.3%	107.5%	115.7%	120.9%	109.1%	68.3%	27.9%	**0.0%**	12.3%	4.4%	14.7%	4.3%
23	59.8%	100.2%	106.8%	111.0%	99.8%	62.7%	25.9%	**0.0%**	9.9%	1.9%	9.4%	1.8%
24	39.8%	93.8%	98.5%	101.6%	91.0%	57.8%	24.9%	1.1%	8.3%	**0.0%**	4.2%	0.0%
25	198.8%	224.8%	232.0%	236.6%	211.4%	138.2%	65.3%	11.0%	21.8%	**0.0%**	1.8%	0.0%
26	164.9%	202.1%	208.1%	212.0%	189.3%	124.2%	59.3%	10.7%	19.8%	**0.0%**	0.7%	0.0%
27	123.5%	175.3%	179.9%	182.9%	163.3%	107.9%	52.8%	11.1%	18.0%	0.7%	**0.0%**	0.0%
28	1.7%	27.7%	30.2%	27.2%	23.0%	15.4%	6.7%	1.1%	**0.0%**	1.0%	15.6%	0.0%
29	**0.0%**	38.0%	39.1%	35.3%	30.0%	21.6%	11.9%	4.8%	1.7%	0.9%	12.0%	1.8%
30	**0.0%**	53.5%	53.0%	48.1%	41.3%	31.6%	20.6%	11.5%	5.9%	2.8%	9.8%	3.5%
31	73.9%	96.1%	94.2%	85.0%	72.4%	54.8%	34.9%	18.0%	6.6%	**0.0%**	9.9%	0.0%
32	59.7%	94.0%	91.5%	82.6%	70.5%	53.8%	34.9%	18.6%	7.1%	**0.0%**	7.3%	0.0%
33	42.3%	91.5%	88.4%	79.8%	68.2%	52.5%	35.0%	19.2%	7.6%	**0.0%**	4.2%	0.0%
34	204.3%	219.9%	210.3%	189.9%	162.5%	126.8%	86.9%	49.7%	20.2%	**0.0%**	1.8%	0.0%
35	169.7%	197.8%	188.8%	170.6%	146.0%	114.1%	78.4%	45.0%	18.4%	**0.0%**	0.7%	0.0%
36	127.6%	171.7%	163.7%	148.0%	126.7%	99.4%	68.9%	40.1%	16.8%	0.7%	**0.0%**	0.0%
Avg.	57.4%	95.0%	96.9%	90.6%	83.1%	62.5%	44.6%	27.2%	24.8%	13.5%	12.2%	6.2%

Table 6
Comparison of the optimal policy against Policies 1, 2, and 3 when SKU-specific costs are equal to **c** except SKU 4 with $c_4 = 121.87$ instead of 12.87.

Inst.	Δ_1	$\Delta_2(k)$										Δ_3
		$k=1$	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$	$k=9$	$k=10$	
19	**0.0%**	29.5%	37.7%	43.0%	99.2%	80.1%	60.4%	49.1%	61.4%	61.9%	76.5%	19.2%
20	**0.0%**	42.2%	49.0%	53.4%	98.5%	78.2%	57.4%	44.4%	54.7%	53.3%	64.6%	34.6%
21	**0.0%**	57.9%	63.1%	66.4%	97.5%	75.7%	53.6%	38.6%	46.3%	42.6%	49.8%	37.8%
22	16.3%	35.4%	40.7%	44.1%	75.0%	49.1%	22.8%	4.5%	12.5%	7.4%	14.1%	7.1%
23	7.6%	34.8%	39.2%	42.0%	66.6%	42.2%	17.4%	**0.0%**	6.7%	1.3%	6.3%	1.2%
24	3.3%	43.2%	46.7%	48.9%	66.8%	42.7%	18.4%	0.8%	6.2%	**0.0%**	3.1%	0.0%
25	94.8%	111.7%	116.4%	119.4%	137.1%	90.1%	42.5%	7.1%	14.2%	**0.0%**	1.2%	0.0%
26	83.1%	108.8%	113.0%	115.7%	130.3%	85.8%	41.0%	7.4%	13.7%	**0.0%**	0.5%	0.0%
27	66.9%	105.6%	109.0%	111.2%	121.5%	80.6%	39.4%	8.3%	13.5%	0.5%	**0.0%**	0.0%
28	**0.0%**	25.6%	28.1%	25.1%	79.6%	73.2%	64.7%	59.1%	58.1%	59.1%	73.4%	19.9%
29	**0.0%**	38.0%	39.1%	35.3%	78.8%	71.3%	61.7%	54.6%	51.4%	50.6%	61.7%	26.3%
30	**0.0%**	53.5%	53.0%	48.1%	77.9%	68.9%	57.9%	48.8%	43.2%	40.1%	47.1%	33.2%
31	12.9%	27.2%	26.0%	20.0%	48.6%	37.8%	24.9%	13.9%	6.6%	2.3%	8.7%	0.0%
32	8.2%	31.4%	29.8%	23.7%	47.2%	36.4%	23.7%	12.6%	4.8%	**0.0%**	5.0%	0.0%
33	5.1%	41.4%	39.2%	32.8%	49.9%	38.8%	25.8%	14.2%	5.6%	**0.0%**	3.1%	0.0%
34	98.3%	108.5%	102.2%	89.0%	105.3%	82.6%	56.6%	32.4%	13.2%	**0.0%**	1.2%	0.0%
35	86.4%	105.8%	99.7%	87.1%	100.4%	78.9%	54.2%	31.1%	12.7%	**0.0%**	0.5%	0.0%
36	69.9%	102.8%	96.8%	85.1%	94.2%	74.2%	51.4%	30.0%	12.6%	0.5%	**0.0%**	0.0%
Avg.	30.7%	61.3%	62.7%	60.6%	87.5%	65.9%	43.0%	25.4%	24.3%	17.8%	23.1%	10.0%

finds the ‘‘optimal’’ solution by solving the ILP that takes the empirical demand distribution as input. We evaluate the cost of the resulting policy under the *true* demand distribution, and report its percentage deviation from the true optimal policy (i.e., the policy that is obtained

Table 7
Percentage deviation from the true optimal cost when the optimal policy is calculated based on a demand distribution estimated from historical data of T maintenance cases.

Inst.	$T = 15$	$T = 30$	$T = 50$	$T = 100$
1	12.8 ±0.9%	3.8 ± 0.5%	0.6 ± 0.2%	0.0 ± 0.0%
2	9.0 ± 0.9%	1.0 ± 0.3%	0.1 ± 0.1%	0.0 ± 0.0%
3	2.9 ± 0.6%	0.0 ± 0.0%	0.0 ± 0.0%	0.0 ± 0.0%
4	15.0 ± 0.5%	11.1 ± 0.5%	6.5 ± 0.4%	2.0 ± 0.2%
5	15.1 ± 0.5%	9.9 ± 0.4%	5.5 ± 0.3%	2.0 ± 0.2%
6	16.4 ± 0.5%	9.6 ± 0.5%	3.7 ± 0.3%	1.2 ± 0.2%
7	30.4 ± 1.3%	9.5 ± 0.8%	2.9 ± 0.5%	0.1 ± 0.1%
8	23.9 ± 1.1%	9.7 ± 0.8%	2.8 ± 0.5%	0.1 ± 0.1%
9	21.9 ± 1.0%	9.1 ± 0.7%	2.3 ± 0.4%	0.1 ± 0.1%
10	4.4 ± 0.9%	0.3 ± 0.2%	0.0 ± 0.0%	0.0 ± 0.0%
11	1.1 ± 0.5%	0.2 ± 0.2%	0.0 ± 0.0%	0.0 ± 0.0%
12	0.3 ± 0.3%	0.0 ± 0.0%	0.0 ± 0.0%	0.0 ± 0.0%
13	18.8 ± 0.5%	12.3 ± 0.5%	7.7 ± 0.4%	3.0 ± 0.3%
14	16.3 ± 0.6%	8.9 ± 0.5%	5.4 ± 0.4%	2.1 ± 0.2%
15	12.9 ± 0.7%	5.7 ± 0.5%	1.9 ± 0.2%	0.9 ± 0.1%
16	34.7 ± 1.5%	10.4 ± 0.9%	2.9 ± 0.5%	0.2 ± 0.1%
17	28.9 ± 1.4%	7.8 ± 0.8%	2.2 ± 0.4%	0.4 ± 0.2%
18	25.7 ± 1.2%	8.7 ± 0.8%	3.3 ± 0.5%	0.0 ± 0.0%
19	9.8 ± 0.7%	5.4 ± 0.4%	2.4 ± 0.3%	1.0 ± 0.2%
20	7.0 ± 0.5%	2.8 ± 0.3%	2.1 ± 0.2%	0.8 ± 0.1%
21	3.0 ± 0.4%	1.9 ± 0.2%	0.8 ± 0.1%	0.6 ± 0.1%
22	7.1 ± 0.7%	2.9 ± 0.4%	2.4 ± 0.3%	1.3 ± 0.1%
23	5.6 ± 0.6%	1.8 ± 0.2%	1.3 ± 0.2%	0.7 ± 0.1%
24	3.9 ± 0.5%	1.5 ± 0.2%	1.1 ± 0.1%	0.8 ± 0.1%
25	8.7 ± 0.9%	3.8 ± 0.3%	2.8 ± 0.2%	1.8 ± 0.2%
26	8.8 ± 1.1%	2.3 ± 0.3%	1.8 ± 0.2%	1.1 ± 0.2%
27	5.1 ± 0.7%	1.6 ± 0.2%	1.4 ± 0.2%	0.6 ± 0.1%
28	6.7 ± 0.5%	2.1 ± 0.2%	1.3 ± 0.1%	0.9 ± 0.1%
29	5.9 ± 0.4%	2.9 ± 0.2%	2.0 ± 0.2%	1.2 ± 0.1%
30	5.2 ± 0.4%	2.8 ± 0.2%	1.6 ± 0.2%	0.7 ± 0.1%
31	9.8 ± 0.7%	5.7 ± 0.4%	4.0 ± 0.3%	1.3 ± 0.2%
32	6.5 ± 0.5%	3.6 ± 0.4%	2.6 ± 0.3%	0.9 ± 0.2%
33	5.6 ± 0.5%	3.1 ± 0.3%	2.6 ± 0.2%	0.7 ± 0.1%
34	13.2 ± 1.2%	3.6 ± 0.6%	1.8 ± 0.3%	1.0 ± 0.1%
35	11.0 ± 1.0%	2.5 ± 0.5%	1.4 ± 0.3%	0.3 ± 0.0%
36	10.5 ± 1.0%	2.3 ± 0.4%	1.4 ± 0.3%	0.4 ± 0.1%

by solving the ILP that takes the true demand distribution as input) in Table 7. Note that these percentage deviations are calculated by using simulation (i.e., by generating T samples of required parts from the true distribution to mimic the past corrective maintenance cases in each simulation). For each instance, we performed 300 simulations and reported the average percentage deviation with its standard error after the ± sign. As T increases, the empirical demand distribution is expected to converge to the true demand distribution. This explains why the percentage deviations shrink as T increases.

For $T = 100$, Table 7 shows that the percentage deviation never exceeds about 3% and it is often much lower across all instances. This observation suggest that service providers can use our model with reasonable confidence after collecting data from nearly one hundred corrective maintenance cases performed in the past. Grishina et al. (2020) had already reported that their predictive model is the most useful when the number of past corrective maintenance cases is high enough (larger than 100 for their specific example). Our numerical results confirm this finding, and also show how the number of corrective maintenance cases in the historical data specifically affects the performance of the optimal policy.

8. Conclusion

In this paper, we consider a spare part recommendation problem from the perspective of a maintenance service provider. We develop a decision support model for the selection of spare parts that will be proactively sent to a failed system. In this way, a costly second visit caused by the unavailability of the right spare parts can be avoided and the corrective maintenance can be completed in the first visit of a local service engineer. We formulate an ILP model that minimizes the expected total cost that consists of the cost of a second visit in addition to the spare-part transportation costs, and SKU-specific retrieval and send-back costs. We provide analytical results on the structure of the optimal policy for the special cases of single and two SKUs. In addition, we derive analytical properties of the optimal policy for a general number of SKUs. We introduce a greedy heuristic policy that can exploit these structural properties and, on average, performs better than two other benchmark policies that are commonly used in practice. In this work, it is assumed that the service provider knows the probability distribution of the demand. But, in practice, this distribution is estimated from data related to past corrective maintenance cases. Our numerical analysis shows that the estimation errors becomes negligible once the historical data consists of about one hundred corrective maintenance cases. When the number of previous corrective maintenance cases is limited, an integrated estimation and decision-making approach to simultaneously learn the demand distribution and make spare-part recommendations can be an interesting research direction.

Funding

This research was partially supported by the EU project 'DayTiMe-Digital Lifecycle Twins for Predictive Maintenance' (grant agreement 'ITEA-2018-Daytime-17030').

CRediT authorship contribution statement

İpek Dursun: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. **Anastasiia Grishina:** Conceptualization, Data curation, Writing – original draft. **Alp Akcay:** Conceptualization, Funding acquisition, Supervision, Writing – original draft, Writing – review & editing, Investigation, Methodology. **Geert-Jan van Houtum:** Conceptualization, Funding acquisition, Supervision, Writing – original draft, Writing – review & editing, Investigation, Methodology.

Appendix A. Proofs

Proof of Proposition 1. For $N = 1$, there are two solutions for the ILP: $x_1 = 1$ and $x_1 = 0$. The costs under these solutions are

$$C(1) = F + c_1(1 - p_1)$$
$$C(0) = (D + F)p_1.$$

Note that $C(1)$ is strictly decreasing as a function of p_1 and $C(0)$ is strictly increasing. Let $\check{p}_1$ be the point where the cost under both solutions is equal. Then:

$$C(1) - C(0) = 0 \iff F + c_1(1 - \check{p}_1) - (D + F)\check{p}_1 = 0$$
$$\iff \check{p}_1 = \frac{F + c_1}{D + F + c_1}$$

When $C(1) \geq C(0)$ (i.e. $\frac{F+c_1}{D+F+c_1} \geq p_1$), solution $x_1 = 0$ is optimal. When $C(1) \leq C(0)$ (i.e. $\frac{F+c_1}{D+F+c_1} \leq p_1$), solution $x_1 = 1$ is optimal. □

Proof of Lemma 1. We assume that $\sigma_{1,2} = 0$. Then, $\tilde{p}_1 = \frac{F+c_1}{D+F_1+c_1}$, and

$$\bar{p}_1 = \frac{(F + c_1)(D + F + c_2)}{(D + F + c_2)(D + F + c_1) - (D + F)c_2} = \frac{F + c_1}{(D + F + c_1) - \frac{(D+F)c_2}{D+F+c_2}}.$$

Obviously, $0 < \frac{c_1}{D+F+c_1} < \frac{F+c_1}{D+F+c_1} < \frac{F+c_1}{(D+F+c_1) - \frac{(D+F)c_2}{D+F+c_2}}$. Hence, $0 < \underline{p}_1 < \tilde{p}_1 < \bar{p}_1$. Additionally, $0 < \underline{p}_1 < \tilde{p}_1 < 1$. It also holds that $0 < \underline{p}_2 < \tilde{p}_2 < \bar{p}_2$ and $0 < \underline{p}_2 < \tilde{p}_2 < 1$ from symmetry. □

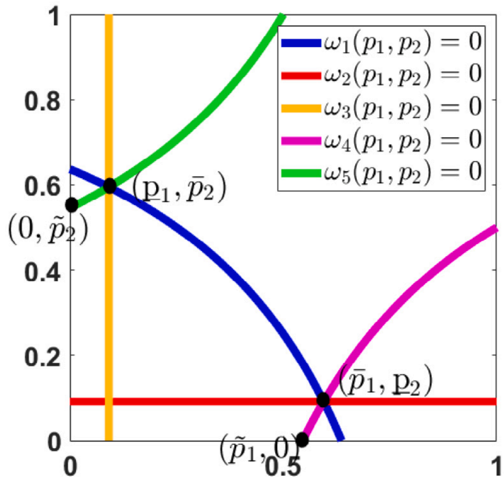


Fig. A.7. Points defined in Definition 1.

Proof of Theorem 1. Assume that $\sigma_{1,2}=0$ and $N=2$. There are four possible solutions: $\mathbf{x}_0=(x_1,x_2)=(0,0)$, $\mathbf{x}_1=(x_1,x_2)=(0,1)$, $\mathbf{x}_2=(x_1,x_2)=(1,0)$ and $\mathbf{x}_3=(x_1,x_2)=(1,1)$. Let $\tilde{h}_m$ be the probability of a second engineer visit if combination $\mathbf{x}_m$ is chosen for the repair kit. $\tilde{h}_0=p_1+p_2-p_1p_2$, $\tilde{h}_1=p_1$, $\tilde{h}_2=p_2$ and $\tilde{h}_3=0$. From these probabilities we calculate costs under each solution (x_1,x_2) as

$$C(0,0)=(D+F)(p_1+p_2-p_1p_2)$$
$$C(0,1)=F+c_2(1-p_2)+(D+F)p_1$$
$$C(1,0)=F+c_1(1-p_1)+(D+F)p_2$$
$$C(1,1)=F+c_1(1-p_1)+c_2(1-p_2).$$

We define $\omega_1(p_1,p_2)$ as the cost difference between the solutions $(0,0)$ and $(1,1)$.

$$\begin{aligned}\omega_1(p_1,p_2) &= C(0,0)-C(1,1)\\ &= (D+F)(p_1+p_2-p_1p_2)-(F+c_1(1-p_1)+c_2(1-p_2))\\ &= (D+F+c_1)p_1+(D+F+c_2)p_2-(p_1p_2)(D+F)-(F+c_1+c_2)\end{aligned}$$

We define $\omega_2(p_1,p_2)$ as the cost difference between the solutions $(1,1)$ and $(1,0)$.

$$\begin{aligned}C(1,1)-C(1,0) &= F+c_1(1-p_1)+c_2(1-p_2)-(F+c_1(1-p_1)+(D+F)p_2)\\ &= c_2-p_2(D+F+c_2)\end{aligned}$$

We define $\omega_3(p_1,p_2)$ as the cost difference between the solutions $(1,1)$ and $(0,1)$.

$$\begin{aligned}C(1,1)-C(0,1) &= F+c_1(1-p_1)+c_2(1-p_2)-(F+c_2(1-p_2)+(D+F)p_1)\\ &= c_1-p_1(D+F+c_1)\end{aligned}$$

We define $\omega_4(p_1,p_2)$ as the cost difference between the solutions $(1,0)$ and $(0,0)$.

$$\begin{aligned}C(1,0)-C(0,0) &= F+c_1(1-p_1)+(D+F)p_2-((D+F)(p_1+p_2-p_1p_2))\\ &= F+c_1+(D+F)(p_1p_2)-p_1(D+F+c_1)\end{aligned}$$

We define $\omega_5(p_1,p_2)$ as the cost difference between the solutions $(0,0)$ and $(0,1)$.

$$\begin{aligned}C(0,1)-C(0,0) &= (F+c_2(1-p_2)+(D+F)p_1)-(D+F)(p_1+p_2-p_1p_2)\\ &= F+c_2+(D+F)(p_1p_2)-p_2(D+F+c_2)\end{aligned}$$

We define $\omega_6(p_1,p_2)$ as the cost difference between the solutions $(1,0)$ and $(0,1)$.

$$\begin{aligned}C(0,1)-C(1,0) &= F+c_2(1-p_2)+(D+F)p_1-(F+c_1(1-p_1)+(D+F)p_2)\\ &= (c_2-c_1)-p_2(c_2+D+F)+p_1(c_1+D+F)\end{aligned}$$

i. Solution $\mathbf{x}=(0,0)$ is optimal if and only if $\omega_1(p_1,p_2)\leq 0$, $\omega_4(p_1,p_2)\geq 0$ and $\omega_5(p_1,p_2)\geq 0$.
ii. Solution $\mathbf{x}=(0,1)$ is optimal if and only if $\omega_3(p_1,p_2)\geq 0$, $\omega_5(p_1,p_2)\leq 0$, and $\omega_6(p_1,p_2)\leq 0$.
iii. Solution $\mathbf{x}=(1,1)$ is optimal if and only if $\omega_1(p_1,p_2)\geq 0$, $\omega_2(p_1,p_2)\leq 0$ and $\omega_3(p_1,p_2)\leq 0$.
iv. Solution $\mathbf{x}=(1,0)$ is optimal if and only if $\omega_2(p_1,p_2)\geq 0$, $\omega_4(p_1,p_2)\leq 0$, and $\omega_6(p_1,p_2)\geq 0$.

Please note that if $\omega_2(p_1,p_2)\geq 0$, then $\omega_5(p_1,p_2)\leq 0$. If $\omega_4(p_1,p_2)\leq 0$ and $\omega_5(p_1,p_2)\leq 0$, then $\omega_6(p_1,p_2)\geq 0$. So, solution $\mathbf{x}=(1,0)$ is optimal if and only if $\omega_2(p_1,p_2)\geq 0$ and $\omega_4(p_1,p_2)\leq 0$. Therefore, condition $\omega_6(p_1,p_2)\geq 0$ is redundant. Similarly, if $\omega_3(p_1,p_2)\geq 0$, then $\omega_4(p_1,p_2)\geq 0$. If $\omega_4(p_1,p_2)\geq 0$ and $\omega_5(p_1,p_2)\leq 0$, then $\omega_6(p_1,p_2)\leq 0$. Therefore, solution $\mathbf{x}=(0,1)$ is optimal if and only if $\omega_3(p_1,p_2)\geq 0$ and $\omega_5(p_1,p_2)\leq 0$. The condition $\omega_6(p_1,p_2)\leq 0$ is redundant.

In order to derive the structure of optimal policy, we determine the intersection points of $\omega_k(p_1,p_2)=0, k\in\{1,2,3,4,5\}$. We first calculate the intersection of $\omega_2(p_1,p_2)=0$ and $\omega_4(p_1,p_2)=0$. If $\omega_2(p_1,p_2)=0$, then $p_2=\frac{c_2}{D+F+c_2}$. We first plug p_2 into $\omega_4(p_1,p_2)=0$.

$$\begin{aligned}&F+c_1-p_1\left[(D+F+c_1)-\frac{c_2}{D+F+c_2}(D+F)\right]=0\\ \iff\ &(F+c_1)(D+F+c_2)\\ -\ &p_1\left[(D+F+c_2)(D+F+c_1)-c_2(D+F)\right]=0\\ \iff\ &p_1=\frac{(F+c_1)(D+F+c_2)}{(D+F+c_2)(D+F+c_1)-(D+F)c_2}=\frac{(F+c_1)}{(D+F+c_1)-\frac{(D+F)c_2}{D+F+c_2}}.\end{aligned}$$

We characterize the intersection point of $\omega_2(p_1,p_2)=0$ and $\omega_4(p_1,p_2)=0$ as $(\bar{p}_1,\underline{p}_2)$.

Next, we determine the intersection point of $\omega_1(p_1,p_2)=0$ and $\omega_2(p_1,p_2)=0$. We also plug the point $p_2=\frac{c_2}{D+F+c_2}$ in $\omega_1(p_1,p_2)=0$:

$$\begin{aligned}&(D+F+c_1)p_1+\frac{(D+F+c_2)c_2}{D+F+c_2}-\frac{(D+F)c_2p_1}{D+F+c_2}-(F+c_1+c_2)=0\\ \iff\ &(D+F+c_1)p_1+c_2-\frac{(D+F)c_2p_1}{D+F+c_2}-\sigma_{1,2}(D+F)-(F+c_1+c_2)=0\\ \iff\ &(D+F+c_1)p_1-\frac{(D+F)c_2p_1}{D+F+c_2}-(F+c_1)=0\\ \iff\ &p_1\left((D+F+c_1)-\frac{c_2}{D+F+c_2}(D+F)\right)=F+c_1\\ \iff\ &p_1=\frac{(F+c_1)(D+F+c_2)}{(D+F+c_2)(D+F+c_1)-(D+F)c_2}\end{aligned}$$

We characterize the intersection point of $\omega_1(p_1,p_2)=0$ and $\omega_2(p_1,p_2)=0$ as $(\bar{p}_1,\underline{p}_2)$. We see that $\omega_1(p_1,p_2)=0$, $\omega_2(p_1,p_2)=0$ and $\omega_4(p_1,p_2)=0$ intersects at the same point. From symmetry, the same holds for the intersection point of $\omega_1(p_1,p_2)=0$, $\omega_3(p_1,p_2)=0$ and $\omega_5(p_1,p_2)=0$. We characterize this point with $(\underline{p}_1,\bar{p}_2)$ where $(\underline{p}_1,\bar{p}_2)=\left(\frac{c_1}{D+F+c_1},\frac{(F+c_2)(D+F+c_1)}{(D+F+c_1)(D+F+c_2)-(D+F)c_1}\right)$.

At point $(\bar{p}_1,\underline{p}_2)$ solutions (1, 1), (1, 0) and (0, 0) are all optimal. At point $(\underline{p}_1,\bar{p}_2)$ solutions (1, 1), (0, 1) and (0, 0) are all optimal. $\omega_4(p_1,p_2)$ cuts the p_1-axis at the point $\tilde{p}_1=\frac{F+c_1}{D+F+c_1}$. $\omega_5(p_1,p_2)$ cuts the p_2-axis at $\tilde{p}_2=\frac{F+c_2}{D+F+c_2}$. We describe these points in Definition 1 and we demonstrate them in Fig. A.7.

Finally, we derive Theorem 1 from the following results. First, we reformulate $\omega_5(p_1,p_2)=0$ as $f(p_1)=p_2$, reformulate $\omega_1(p_1,p_2)=0$ as $g(p_1)=p_2$ and reformulate $\omega_4(p_1,p_2)=0$ as of $h(p_1)=p_2$ (please see Definition 1). Then we formulate Theorem 1.

(i) We let R_1 define a region where $\omega_1(p_1,p_2)\leq 0$, $\omega_4(p_1,p_2)\geq 0$ and $\omega_5(p_1,p_2)\geq 0$. Therefore, the solution $(0,0)$ is optimal in region R_1.
(ii) We let R_2 define a region where $\omega_5(p_1,p_2)\leq 0$ and $\omega_3(p_1,p_2)\geq 0$. In region R_2 the solution $(0,1)$ is optimal.

(iii) We let R_3 define a region, where $\omega_3(p_1,p_2) \le 0$, $\omega_1(p_1,p_2) \ge 0$, and $\omega_2(p_1,p_2) \le 0$. The solution $(1,1)$ is optimal in region R_3.

(iv) We let R_4 define the region, where $\omega_4(p_1,p_2) \le 0$, and $\omega_2(p_1,p_2) \ge 0$. The solution $(1,0)$ is optimal in the region R_4. □

Proof of Lemma 2. First, we show that $h_{I'} \ge h_{I'\cup\{l\}}$. Hence, we have to show

$$h_{I'} - h_{I'\cup\{l\}} = \sum_{m\in M'} (\hat{u}_m(I'_1) - \hat{u}_m(I'_1\backslash\{l\}))\hat{p}_m \ge 0. \tag{A.1}$$

It holds that

$$\hat{u}_m(I') - \hat{u}_m(I'\cup\{l\}) = \begin{cases} 1 & \text{if SKU } l \text{ in } s_m \text{ and } I' \text{ contains all other SKUs from } \mathbf{s}_m, \\ 0 & \text{otherwise.} \end{cases}$$

Hence, (A.1) holds. Next, we need to show that $h_{I'_1} - h_{I'_1\cup\{l\}} \le h_{I'_2} - h_{I'_2\cup\{l\}}$ where $I'_1 \subset I'_2 \subseteq I\backslash\{l\}$. This is equivalent to showing that

$$\sum_{m\in M'} (\hat{u}_m(I'_1) - \hat{u}_m(I'_1\cup\{l\}))\hat{p}_m \le \sum_{m\in M'} (\hat{u}_m(I'_2) - \hat{u}_m(I'_2\cup\{l\}))\hat{p}_m.$$

It holds that

$$\hat{u}_m(I'_1) - \hat{u}_m(I'_1\cup\{l\}) = \begin{cases} 1 & \text{if SKU } l \text{ in } s_m \text{ and } I'_1 \text{ contains all other SKUs from } \mathbf{s}_m, \\ 0 & \text{otherwise.} \end{cases}$$

$$\hat{u}_m(I'_2) - \hat{u}_m(I'_2\cup\{l\}) = \begin{cases} 1 & \text{if SKU } l \text{ in } s_m \text{ and } I'_2 \text{ contains all other SKUs from } \mathbf{s}_m, \\ 0 & \text{otherwise.} \end{cases}$$

Notice that $\hat{u}_m(I'_1) - \hat{u}_m(I'_1\cup\{l\}) = 1$ implies that $\hat{u}_m(I'_2) - \hat{u}_m(I'_2\cup\{l\}) = 1$ because $I'_1 \subset I'_2$. Hence, will hold. □

Proof of Lemma 3.

(i) Cost of bringing all parts is $C(\mathbf{x}_I) = F + \sum_{i\in I} c_i(1-p_i)$. The cost of bringing no parts is $C(\mathbf{x}_\emptyset) = (D+F)h_\emptyset$. Their difference is $C(\mathbf{x}_I) - C(\mathbf{x}_\emptyset) = F + \sum_{i\in I} c_i(1-p_i) - (D+F)h_\emptyset$. If this difference is greater than zero, (i.e. $F + \sum_{i\in I} c_i(1-p_i) - (D+F)h_\emptyset > 0$), then the solution $\mathbf{x}_\emptyset$ dominates the solution $\mathbf{x}_I$.

(ii) Now, we write the condition for removing a part i from any set I', where the new set is $I'' = I'\backslash\{i\}$. The cost of solution, $\mathbf{x}_{I'}$ is $C(\mathbf{x}_{I'}) = F + \sum_{j\in I'} c_j(1-p_j) + (D+F)h_{I'}$ and the cost of taking solution, $\mathbf{x}_{I''}$ is $C(\mathbf{x}_{I''}) = F + \sum_{j\in I''} c_j(1-p_j) + (D+F)h_{I''}$. The solution $\mathbf{x}_{I''}$ dominates the solution $\mathbf{x}_{I'}$ if the following holds.

$$\begin{aligned} C(\mathbf{x}_{I'}) - C(\mathbf{x}_{I''}) &= F + \sum_{j\in I'} c_j(1-p_j) + (D+F)h_{I'} \\ &\quad - F - \sum_{j\in I''} c_j(1-p_j) - (D+F)h_{I''} \\ &= c_i(1-p_i) - (D+F)(h_{I''} - h_{I'}) > 0. \end{aligned}$$

As a special case, the condition for removing part i from the solution I is

$$C(\mathbf{x}_I) - C(\mathbf{x}_{I\backslash\{i\}}) = c_i(1-p_i) - (D+F)p_i > 0,$$

which can be rewritten to condition (14).

(iii) Let assume $I' = \{i\}$ and $I'' = \emptyset$. Solution $\mathbf{x}_\emptyset$ dominates the solution $\mathbf{x}'_I$ if the following holds.

$$C(x_{I'}) - C(x_\emptyset) = F + c_i(1-p_i) - (D+F)(h_\emptyset - h_{I'}) \ge 0,$$

where $h_\emptyset - h_{I'} > 0$. □

Proof of Theorem 2. Let assume $I'\backslash\{i\}$ is a better solution than I'. $I'\backslash\{i\}$ is a better solution than I', if and only if $F\mathbb{1}_{\{I'\backslash\{i\}\neq\emptyset\}} + c_i(1-p_i) - (D+F)(h_{I'\backslash\{i\}} - h_{I'}) > 0$ holds. We need to show that $\mathbf{x}_{I''\backslash\{i\}}$ is a better solution than $\mathbf{x}_{I''}$, therefore, $C(\mathbf{x}_{I''}) - C(\mathbf{x}_{I''\backslash\{i\}}) > 0$, to complete the proof. Please note that $(h_{I'\backslash\{i\}} - h_{I'}) \ge (h_{I''\backslash\{i\}} - h_{I''})$ holds by Lemma 2.

$$\begin{aligned} C(\mathbf{x}_{I''}) - C(\mathbf{x}_{I''\backslash\{i\}}) &= F\mathbb{1}_{\{I''\backslash\{i\}\neq\emptyset\}} + c_i(1-p_i) - (D+F)(h_{I''\backslash\{i\}} - h_{I''}) \\ &\ge F\mathbb{1}_{\{I'\backslash\{i\}\neq\emptyset\}} + c_i(1-p_i) - (D+F)(h_{I'\backslash\{i\}} - h_{I'}) > 0. \end{aligned}$$

This shows that $\mathbf{x}_{I''\backslash\{i\}}$ is a better solution than $\mathbf{x}_{I''}$. □

Proof of Corollary 1. Let $\mathbf{x}^*$ be an optimal solution and suppose that (15) holds. Then, we need to show that $x^*_i = 0$. First consider a set I', where $i \in I'$. Then the cost of the solution $\mathbf{x}_{I'}$ is $C(\mathbf{x}_{I'}) = F + c_i(1-p_i) + \sum_{j\in I'\backslash\{i\}} c_j(1-p_j) + (D+F)h_{I'}$. In order to show that $x_i = 0$ is a better solution than $x_i = 1$, we need to show that $C(\mathbf{x}_{I'}) - C(\mathbf{x}_{I'\backslash\{i\}}) > 0$.

$$\begin{aligned} C(\mathbf{x}_{I'}) - C(\mathbf{x}_{I'\backslash\{i\}}) &= F\mathbb{1}_{\{I'\backslash\{i\}\neq\emptyset\}} + c_i(1-p_i) - (D+F)(h_{I'\backslash\{i\}} - h_{I'}) \\ &\ge F\mathbb{1}_{\{I'\backslash\{i\}\neq\emptyset\}} + c_i(1-p_i) - (D+F)p_i \\ &> F\mathbb{1}_{\{I'\backslash\{i\}\neq\emptyset\}} + c_i\left(1 - \frac{c_i}{c_i+D+F}\right) - (D+F)\frac{c_i}{c_i+D+F} \\ &= F\mathbb{1}_{\{I'\backslash\{i\}\neq\emptyset\}} + \frac{c_i(D+F)}{c_i+D+F} - \frac{c_i(D+F)}{c_i+D+F} = F\mathbb{1}_{\{I'\backslash\{i\}\neq\emptyset\}} > 0. \end{aligned}$$

Please note that from Remark 2 and Lemma 2 it holds that $(h_{I'\backslash\{i\}} - h_{I'}) \le p_i$. □

Appendix B. Service level constraint

Service level agreements with system users may enforce a minimum level on the level of service delivered by the service provider. Therefore, we introduce a service level constraint

$$1 - \sum_{m\in M'} \hat{u}_m\hat{p}_m \ge \alpha \tag{B.1}$$

to ensure the probability of resolving a maintenance case during the diagnostic visit is at least at the desired level α. Note that $\sum_{m\in M'} \hat{u}_m\hat{p}_m$ is the probability that a second visit is required. We let C^α_{opt} be the cost of the optimal solution with the service level constraint under a given α. In the presence of a service-level constraint, the cost parameter D for the second visit can be interpreted as the cost of travel (or additional operational costs incurred by the service provider due to the second visit) instead of the penalty cost for the additional downtime of the failed system.

References

Abdel-Malek, L. L., & Areeratchakul, N. (2007). A quadratic programming approach to the multi-product newsvendor problem with side constraints. *European Journal of Operational Research*, *176*(3), 1607–1619.

Abdel-Malek, L., Montanari, R., & Morales, L. C. (2004). Exact, approximate, and generic iterative models for the multi-product newsboy problem with budget constraint. *International Journal of Production Economics*, *91*(2), 189–198.

Bernstein, F., & DeCroix, G. A. (2015). Advance demand information in a multiproduct system. *Manufacturing & Service Operations Management*, *17*(1), 52–65.

Bijvank, M., Koole, G., & Vis, I. F. (2010). Optimising a general repair kit problem with a service constraint. *European Journal of Operational Research*, *204*(1), 76–85.

Choi, S., & Ruszczyński, A. (2011). A multi-product risk-averse newsvendor with exponential utility function. *European Journal of Operational Research*, *214*(1), 78–84.

Dutta, P., & Chakraborty, D. (2010). Incorporating one-way substitution policy into the newsboy problem with imprecise customer demand. *European Journal of Operational Research*, *200*(1), 99–110.

Graves, S. C. (1982). Note: A multiple-item inventory model with a job completion criterion. *Management Science*, *28*(11), 1334–1337.

Grishina, A., Stolikj, M., Gao, Q., & Petkovic, M. (2020). Improving spare part search for maintenance services using topic modelling. In *CEUR workshop proceedings: vol. 2699*, Galway, Ireland.

Heeremans, D., & Gelders, L. (1995). Multiple period repair kit problem with a job completion criterion: A case study. *European Journal of Operational Research*, *81*(2), 239–248.

Kuthambalayan, T. S., Mehta, P., & Shanker, K. (2015). Managing product variety with advance selling and capacity restrictions. *International Journal of Production Economics*, *170*, 287–296.

Lei, L., Ru, J., Shi, R., & Zhang, J. (2022). A two-product newsvendor problem with partial demand substitution. *Production and Operations Management*, *31*(3), 1157–1173.

Mamer, J. W., & Shogan, A. W. (1987). A constrained capital budgeting problem with applications to repair kit selection. *Management Science*, *33*(6), 800–806.

Mamer, J. W., & Smith, S. A. (1982). Optimizing field repair kits based on job completion rate. *Management Science*, *28*(11), 1328.

Moon, I., & Silver, E. A. (2000). The multi-item newsvendor problem with a budget constraint and fixed ordering costs. *Journal of the Operational Research Society*, *51*, 602–608.

Nahmias, S., & Schmidt, C. P. (1984). An efficient heuristic for the multi-item newsboy problem with a single constraint. *Naval Research Logistics Quarterly*, *31*(3), 463–474.

Neves-Moreira, F., Veldman, J., & Teunter, R. H. (2021). Service operation vessels for offshore wind farm maintenance: Optimal stock levels. *Renewable and Sustainable Energy Reviews*, *146*, Article 111158.

Olivares-Nadal, A. V. (2024). Constructing decision rules for multiproduct newsvendors: An integrated estimation-and-optimization framework. *European Journal of Operational Research*.

Özler, A., Tan, B., & Karaesmen, F. (2009). Multi-product newsvendor problem with value-at-risk considerations. *International Journal of Production Economics*, *117*(2), 244–255.

Prak, D., Saccani, N., Syntetos, A., Teunter, R., & Visintin, F. (2017). The repair kit problem with positive replenishment lead times and fixed ordering costs. *European Journal of Operational Research*, *261*(3), 893–902.

Rippe, C., & Kiesmüller, G. P. (2023a). The added value of advance demand information for the planning of a repair kit. *Central European Journal of Operations Research*, *31*, 311–335.

Rippe, C., & Kiesmüller, G. P. (2023b). The repair kit problem with imperfect advance demand information. *European Journal of Operational Research*, *304*(2), 558–576.

Saccani, N., Visintin, F., Mansini, R., & Colombi, M. (2017). Improving spare parts management for field services: a model and a case study for the repair kit problem. *IMA Journal of Management Mathematics*, *28*(2), 185–204.

Smith, S. A., Chambers, J. C., & Shlifer, E. (1980). Note—Optimal inventories based on job completion rate for repairs requiring multiple items. *Management Science*, *26*(8), 849–854.

Song, J. S., van Houtum, G. J., & Van Mieghem, J. A. (2020). Capacity and inventory management: Review, trends, and projections. *Manufacturing & Service Operations Management*, *22*(1), 36–46.

Teunter, R. H. (2006). The multiple-job repair kit problem. *European Journal of Operational Research*, *175*(2), 1103–1116.

Topan, E., Eruguz, A., Ma, W., van der Heijden, M., & Dekker, R. (2020). A review of operational spare parts service logistics in service control towers. *European Journal of Operational Research*, *282*(2), 401–414.

Turken, N., Tan, Y., Vakharia, A. J., Wang, L., Wang, R., & Yenipazarli, A. (2012). The multi-product newsvendor problem: Review, extensions, and directions for future research. *Handbook of Newsvendor Problems: Models, Extensions and Applications*, 3–39.

Wang, Y., Xiao, L., & Luo, X. G. (2023). On the optimal solution of a distributionally robust multi-product newsvendor problem. *Journal of the Operational Research Society*, *74*(12), 2578–2592.