Proceedings of the 8th European Conference on Industrial Engineering and Operations Management
Paris, France, July 2-4, 2025

Proceedings of the International Conference on Industrial Engineering and Operations Management

Publisher: IEOM Society International, USA DOI: 10.46254/EU08.20250279
Published: July 2, 2025

Smart Swaps: Optimizing Rotable Inventory for Exchange-Based Maintenance and Overhaul Operations

Murat Erkoc
Department of Industrial and Systems Engineering
University of Miami, Coral Gables, FL USA
merkoc@miami.edu

Salvador Romo-Fragoso
Stahl LLC
Miami, FL USA
salvador@sthal.net

Abstract

We investigate optimal rotable inventory control for maintenance, repair, and overhaul (MRO) service providers operating under stochastic demand and service rates within the context of an equipment exchange policy. In this setting, the provider fulfills service requests by exchanging the customer's equipment with a ready-to-use unit from its rotable (exchange) inventory. Orders are admitted only when inventory is available for immediate exchange, making inventory availability a key determinant of effective demand. We develop a queueing-based model to characterize the system dynamics and derive profit-maximizing inventory policies. Through analytical results and extensive numerical analysis, we explore how cost parameters, arrival rates, service rates, and capacity interact to influence optimal inventory positions. Our findings demonstrate that optimal inventory levels are highly sensitive to both economic and operational factors, such as the ratio of holding cost to exchange premium, traffic intensity, and service capacity. The results also highlight several non-monotonic relationships between operational parameters and inventory levels, offering nuanced insights into the strategic design of exchange-based MRO systems.

Keywords
Maintenance and overhaul; Rotable inventory; Inventory exchange; Queuing models; Markov processes.

1. Introduction
Maintenance, repair, and overhaul (MRO) services are essential for ensuring the continuous, safe, and efficient operation of service and production systems. This is particularly true in capital-intensive industries, where the uninterrupted performance of heavy equipment is vital for both profitability and long-term sustainability. MRO activities in such industries typically incur high costs—not only due to the sophisticated and specialized nature of the services required but also because of the disruptions they may cause to revenue-generating operations. For instance, in the commercial airline industry, equipment downtime can result in costs amounting to hundreds of thousands of dollars per day. Consequently, MRO service contracts are highly sensitive to both cost and time considerations for equipment owners as well as service providers (Vieira and Loures, 2016, Mofokeng et al., 2020, Chandola et al.. 2022).

In response to the need for reducing downtime-related disruptions, many MRO service providers have, in recent years, adopted programs that involve loaning or exchanging equipment to customers while their own equipment undergoes overhaul or repair. These initiatives, commonly referred to as exchange programs, aim to virtually eliminate service interruptions for equipment operators. While these programs offer competitive advantages and generate additional revenue through exchange premiums, they also require MRO providers to maintain a costly inventory of high-value equipment. In practice, this inventory is known as exchange inventory, which is conceptually similar to rotable inventory in literature. We note that while all exchange inventory is rotable, not all rotable inventory is managed under exchange agreements; exchange inventory is specifically maintained to support time-definite loan or swap arrangements with customers, whereas rotable inventory more broadly refers to any reusable components managed through repair-and-return cycles. Exchange inventory consists of equipment that cycles between being loaned to current customers and awaiting redeployment at the service provider's facility.

Under typical exchange programs, a service provider fulfills a customer request only if it has available inventory for immediate exchange—an approach referred to as the *rotable exchange policy*. Although this policy entails that not all incoming orders are admitted, its implementation enables higher turnover of expensive inventory items. As a result, the exchange policy has the potential to enhance profitability, particularly when the associated exchange premium is sufficiently high.

In this paper, we propose a queueing-based model to determine optimal inventory levels under the exchange policy. We derive optimality conditions and analyze the impact of key cost, demand, and service capacity parameters on inventory decisions. Our work is motivated by a real-world collaboration with a leading aviation MRO company specializing in landing gear services for a wide array of customers, including commercial airlines, freight operators, military organizations, and individual aircraft owners. Given the capital-intensive nature of exchange inventory, such companies face critical questions about whether offering an exchange program is economically viable for certain product lines—and, if so, how to determine optimal inventory levels.

Our analytical framework assumes service processes with parallel servers, where each server represents a work cell or a specialized team performing the entire MRO process for an order. This setup reflects the operational structure of regional shops, particularly those servicing landing gear for narrow-body aircraft, where individual teams handle repairs for specific customers. While many large-scale MRO operations follow job shop structures involving multiple teams and workstations, we argue that the managerial insights generated from our model are generalizable to a broader set of contexts to a certain extent.

This paper seeks to answer two strategic questions. First, under what conditions is the exchange program economically justifiable for an MRO service provider? Second, how do demand and capacity parameters influence the optimal sizing of exchange inventory? Our analysis shows that the exchange strategy is not universally optimal. Specifically, under conditions of high inventory costs and low system load, the program may not be economically viable regardless of the admission policy. However, when exchange premiums are sufficiently high or inventory holding costs are low, the policy becomes advantageous. We also find that the effects of arrival rates, service rates, and the number of parallel servers on optimal inventory levels are non-monotonic and context dependent. These dynamics are explored through extensive numerical analysis, from which we derive managerial insights.

The remainder of the paper is organized as follows. Section 2 summarizes the relevant literature. The basic setting and the notation are presented in Section 3. Sections 4 presents the optimality conditions and the analytical results for the exchange policy. Finally, Section 5 concludes the paper.

2. Literature review

Early work in literature pertaining to operations management in the MRO industry mostly focus on development and planning of service tasks (Dekker 1996). In the last two decades, with the growing role of MRO in both production and service industries, an increasing number of researchers turned their attentions to analytical modeling and optimization in this area. Review papers due to Nicolai and Dekker (2008), van Horenbeek et al. (2013), Levi et al. (2014), Hu et al. (2018), and Dinis (2025) provide a comprehensive view of the progression of the analytical models related to MRO services over the past twenty years. Majority of the reported work focus on forecasting, spare parts inventory management, scheduling, and contracting in MRO settings.

Our research diverges from the extant literature by focusing on exchange strategies which have become quite popular in the MRO industry in recent years. In that respect, our study falls in line with research that tackles optimization models that involve rotable inventories in the context of MRO operations. Most of these papers primarily focus on deterministic scheduling problems where the rotable inventory positions are treated as exogenous inputs (Luh et al.

2005; Joo 2009; Ertogral et al. 2015; Erkoc and Ertogral 2016). These papers tackle the scheduling problem either by exact solutions based on mathematical programming models or heuristic approaches. Arts and Flapper (2015) and Ertogral and Ozturk (2019) explicitly incorporate the rotable inventory decisions into the scheduling problem. They propose deterministic mathematical programming models to determine aggregate workforce levels, turn-around stock levels of equipment, and overhaul and replacement quantities over multiple periods for a given set of MRO orders with deadlines. They employ cost minimization objectives that comprise labor costs, material costs, and stocking costs. Our work differs from this group of models in that we consider uncertainty in demand and service rates. We incorporate uncertainty in our problem by employing queuing models.

In the maintenance and repair operations context, queuing models are mostly employed to tackle inventory management problems for repairable items (Muckstadt 2005). The typical goal is to determine optimal inventory levels for spare parts that help reduce the downtime costs incurred by machine failures in a closed loop system (Spanjers et al. 2005; Mirzahosseinian and Piplani 2011). Such studies usually involve queuing networks resulting in computationally complex problems. As such, they resort to approximation approaches (Perlman 2001; Zijm and Avsar 2003; Avsar and Zijm 2014). In this field, the most relevant and closest work to ours is due to Buyukkaramikli et al. (2015). The authors propose a queuing-based Markov decision model for the joint optimization of rotable inventory and capacity for a repair shop whose objective is to minimize the total cost of capacity, downtime and inventory. In this setting, failed components received at the shop are replaced by a spare part carried in the inventory of the shop. Once the failed component is repaired, it is added to the shop's spare part inventory. If a spare part is not available for immediate replacement, the arriving repair order has to wait until one is available. Our work differs in several aspects: *i*) while their model assumes a single server, we employ multi-server models; *ii*) instead of cost minimization, we focus on profit maximization; and *ii*) more importantly, we consider a pure exchange policy that limits the number of accepted orders based on inventory position.

Our model for the rotable exchange policy can be related to the service systems where order admissions are limited by the number of servers (Ormeci et al. 2002; Erkoc and Romo-Fragoso 2015; Dudin et al. 2015) or the number of inventory items (Savin et al. 2005; Gans and Savin 2007). The first group of papers investigates optimal admission control policies across multiple classes of customers that differ in their service requirements and revenue contributions. Our main difference from this group is that our model involves inventory management, and the order admissions are limited by inventory which may go above the number of servers. The second group also considers the admission control problem but in the context of rentals where the rental duration is akin to the service time in our model. Mathematically, these settings are reduced to a special case where the number of servers and number of inventory items are equal. The rotable exchange program, on the other hand, is more general in this respect.

3. Basic Setting, Assumptions and Notation

We consider an MRO service provider that receives job requests at random intervals. Each arriving job corresponds to a piece of equipment or a device requiring maintenance or repair. Under an exchange policy, the service provider fulfills the request by immediately replacing the customer's equipment with a ready-to-use unit from its in-stock inventory. This approach virtually eliminates turnaround time for the customer. In practice, this inventory is referred to as *exchange* inventory.

Because equipment downtime can be extremely costly for operators, many are willing to pay a premium for rapid turnaround, making the exchange model attractive for MRO providers. However, implementing an exchange policy requires the provider to maintain a substantial inventory of high-value equipment, which incurs significant holding costs. As a result, the viability of the exchange strategy hinges on the trade-off between the additional revenue generated through exchange premiums and the cost of maintaining exchange inventory.

Under the exchange policy, all incoming overhaul requests are fulfilled through immediate equipment replacement, contingent on the availability of exchange inventory. If no ready-to-use unit is available at the time of request, the job is not accepted. While this policy ensures minimal delay for admitted jobs, it also imposes strict inventory requirements that directly influence the service provider's operational and economic outcomes.

To model this setting, we adopt a Markov process framework with parallel service lines. MRO requests arrive according to a Poisson process with rate λ, and service times are exponentially distributed with rate μ. We let c denote the number of parallel service lines, each corresponding to a dedicated service team or work cell. The exchange inventory position is denoted by S, which represents the total number of high-value units maintained in the exchange pool. At any given time, this inventory is composed of two components S_f, and the number of exchange-ready units available in stock, and S_w, the number of units currently exchanged out to customers. We note that, while S_f and S_w vary depending on the state of the system, their sum, $S = S_f + S_w$ is constant at all times. As noted earlier, maintaining this inventory incurs a holding cost for the MRO provider, which we denote by h per unit per unit time.

4. Exchange Model

In this model, the system state is defined by means of the (S_f, S_w) pair. All arrivals are rejected when the system state is $(0, S)$. When $S > 0$, at each service request, S_f decreases and S_w increases by one unit each. The opposite occurs on the completion of an MRO service. The service provider charges a fee, denoted by *r*, for each accepted demand for service. To avoid pathological cases, we assume that $r * \min(\lambda, \mu) > h$.

The dynamics of the Markov process that represents the pure exchange setting depends on the comparison between the inventory position (*S*) and the number of service lines (*c*). Clearly, if $S \leq c$, the number of busy service lines cannot exceed *S*. Consequently, the service system effectively operates with *S* servers since demand beyond *S* will not be accepted by the service provider under this condition. This also implies that there will be no queue for service. On the other hand, if $S > c$, there will be a queue for the service when $S_w \leq c$. Based on this observation, we derive the steady state conditions for both the case of ample capacity (*i.e.*, $S \leq c$) and scarce capacity ($S > c$). We use the steady state parameters to derive the expected profit function for the service provider and investigate the optimality conditions for the selection of the exchange inventory position.

When $S \leq c$, the pure exchange policy turns into a truncated queue where no line forms. Using Kendall notation, this system can be denoted by *M/M/S/S*. In this case, the steady state probability distribution is known to be captured by Erlang's first formula (Gross and Harris 1998), namely,

$$\pi_n = \frac{\rho^n}{n! \sum_{i=0}^{S} \frac{\rho^i}{i!}} \qquad (0 \leq n \leq S), \tag{1}$$

where $\rho = \lambda/\mu$, is the traffic intensity and π_n denotes the probability that there are *n* units of equipment in service (*i.e.*, $S_w = n$). The expected profit function depends on the effective arrival rate. As pointed out earlier, incoming orders are accepted as long as there is available inventory for exchange. When $S_w = S$, or equivalently $S_f = 0$, all inventory is exchanged out implying that additional orders cannot be expected until the next MRO service is completed. From (1), this probability is captured by π_S and known as the blocking probability. Consequently, the affective arrival rate is $\lambda(1 - \pi_S)$ and hence, the associated revenue can be captured by $r\lambda(1 - \pi_S)$. Using (1), we can compute the expected profit as follows:

$$G_{S \leq c} = r\lambda \frac{A_{S-1}}{A_S} - hS, \tag{2}$$

where

$$A_x = \sum_{i=0}^{x} \frac{\rho^i}{i!} \tag{3}$$

.

In (2), the first term is the expected revenue whereas the second term is the cost of the inventory position. To derive the optimality conditions, we first establish that $G_{S \leq c}$ is concave in *S*.

Lemma 1. *The steady state profit function $G_{S \leq c}$ is concave in S for $S \geq 1$.*

Proof of Lemma 1. To complete the proof, it is sufficient to show that $\frac{A_{S-1}}{(\rho^S/S!) + A_{S-1}}$ is concave in *S*. First we note that this term is a discrete function of *S*. A discrete function, *f*(*n*) is concave in *n* a given region if for any *n* in the said region the following holds:

$$f(n+1) + f(n-1) - 2f(n) \leq 0. \tag{4}$$

Therefore, for concavity we need

$$\frac{A_S}{A_{S+1}}+\frac{A_{S-2}}{A_{S-1}}-2\frac{A_{S-1}}{A_S}\leq 0. \tag{5}$$

We can easily observe that the left hand side of the above inequality returns a negative value for any $S > 1$ and $\rho > 0$ using a three dimensional graph. For the case $S = 1$, the above inequality reduces to

$$-\frac{1}{(1+\rho)\left(1+\rho+\left(\frac{1}{2}\right)\rho^2\right)} \tag{6}$$

which is clearly negative for $\rho > 0$. Consequently, we can conclude that $G_{S\leq c}$ is concave in S for $S \geq 1$.

◊

This result indicates that the profit function given in (2) has a unique maximizer and can be easily obtained using a line search. We note that although the steady state probability formula depends on the average service time, it is independent of how the service time is distributed. As such, the optimal inventory position obtained from (2) is valid for any service time distribution.

When the number of service lines is smaller than the inventory amount, (that is, $S > c$), we have a *M/M/c/S* queuing system. In this system, the blocking probability, where $S_f = 0$ and $S_w = S$, is

$$\pi_S=\frac{(\rho/c)^S}{c!c^{-c}}\pi_0, \tag{7}$$

where

$$\pi_0=1/\left(\sum_{i=0}^{c-1}\frac{\rho^i}{i!}+\frac{1}{c!\,c^{-c}}\sum_{i=c}^{S}\left(\frac{\rho}{c}\right)^i\right) \tag{8}$$

.

Letting $z = \rho/c$, we can re-write the blocking probability as follows:

$$\pi_S=\frac{c^c(1-z)z^S}{c!(1-z)A_c+zc^c(z^c-z^S)}. \tag{9}$$

Consequently, the expected profit function is

$$G_{S>c}=r\lambda\left(1-\frac{c^c(1-z)z^S}{c!(1-z)A_c+zc^c(z^c-z^S)}\right)-hS. \tag{10}$$

In what follows, we show that the above function is unimodular in *S*.

Lemma 2. *The steady state profit function* $G_{S>c}$ *is unimodular with a unique maximizer for S > 0.*

Proof of Lemma 2. To prove unimodularity, we show that at any stationary point for *S*>0 in (10), the second derivative returns a negative value implying that there must be a unique stationary point which maximizes the function. First let $X = c!\,(1-z)A_c + z^{c+1}c^c$. Then the first derivative of $G_{S>c}$ with respect to *S* is

$$G'_{S>c}=r\lambda\left(-\frac{c^c(1-z)ln(z)z^SX}{\left(X-c^cz^{(S+1)}\right)^2}\right)-h. \tag{11}$$

At any stationary point the above function should return zero. Now we look at the second derivative:

$$G''_{S>c}=r\lambda\left(-\frac{c^c(1-z)\,ln(z)^2z^SX(X+c^cz^{(S+1)})}{\left(X-c^cz^{(S+1)}\right)^3}\right). \tag{12}$$

Evaluated at the stationary point (say *S*=*S**), the second derivative is then

$$G''_{S>c}(S = S^*) = h\left(\frac{\left(X+c^c z^{(S^*+1)}\right)}{\left(X-c^c z^{(S^*+1)}\right)^3}\right) ln\,(z). \tag{13}$$

First suppose $c \geq \rho$. In this case, since $z \leq 1$, $ln\,(z) \leq 0$ and both the numerator and the denominator inside the above parenthesis are nonnegative. As such, the overall function returns a non-positive value. When $c < \rho$, $z > 1$ which implies that $ln(z) > 0$. In this case, it is straightforward to see that the denominator inside the parenthesis is strictly negative since $c \leq S^*$. Using the incomplete upper gamma function, we can rewrite the numerator as follows:

$$X + c^c z^{(S^*+1)} = (1 - z)e^{\rho}\Gamma(c + 1, \rho) + c^c\left(z^{c+1} + z^{(S^*+1)}\right). \tag{14}$$

From the Gauss's continued fraction expansion of the incomplete gamma function, we can observe that $\Gamma(c + 1, \rho) < \rho^{c+1}e^{-\rho}/(\rho - c)$. Consequently, since $z > 1$,

$$X + c^c z^{(S^*+1)} > (1 - z)\frac{\rho^{c+1}}{\rho-c} + c^c\left(z^{c+1} + z^{(S^*+1)}\right), \tag{15}$$

where the right hand side reduces to ρ^{S+1}/c^{S-c+1}. Since this term is a lower bound for (14) and strictly positive, we can conclude that the equation in (14) always returns a positive value. Hence, the overall function in (13) must be strictly negative. As such, we observe that any stationary point must be maximizing for the profit function implying that it must be unique.

◊

From (2) and (10), now we can write the expected profit function for the MRO service provider in general form:

$$G(S) = \begin{cases} G_{S\leq c}, & \text{if } S \leq c \\ G_{S>c}, & \text{if } S > c \end{cases} \tag{16}$$

Given (16), we can now investigate the optimal inventory position for the service provider. We first need the following result:

Lemma 3. *For $S \to c$, the slope of $G_{S>c}$ is smaller than the slope of $G_{S\leq s}$.*

Proof of Lemma 3. To see this result holds, we first note that at $S = c$ both functions intersect. Second, beyond this point (*i.e.*, $S > c$) $G_{S\leq c}$ always returns higher values since this function is constructed based on the assumption that the maximum number of busy servers is S whereas in $G_{S>c}$ this number is c. As such the profit curve for the latter function should always be above the former one for any S such that $S > c$. This implies that at $S = c$, $G_{S>c}$ must have a lower slope than that of $G_{S\leq c}$.

◊

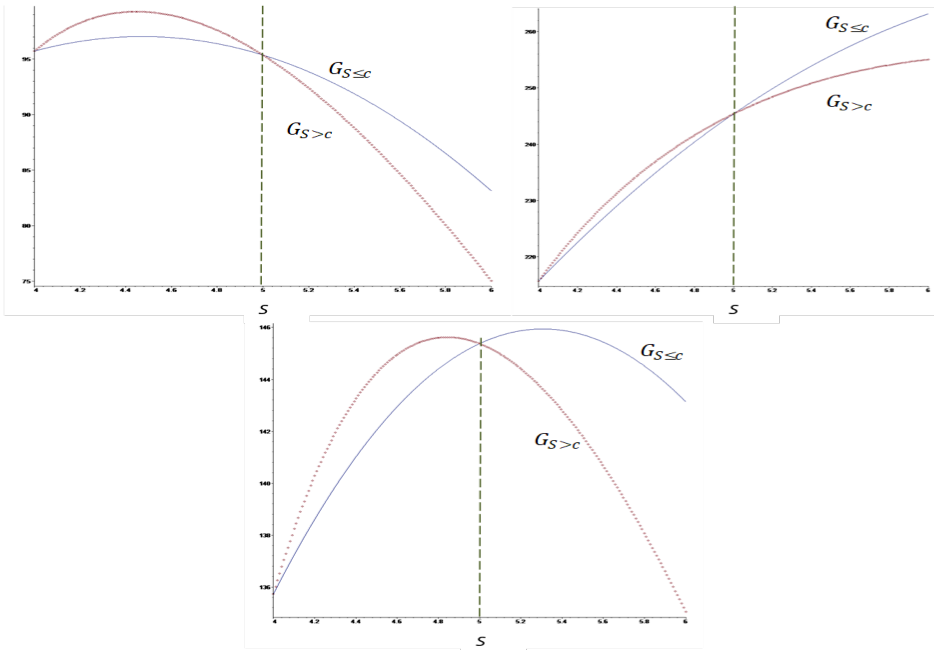

Figure 1. Expected profit curve scenarios.

The above result helps us deduce the possible scenarios for optimal inventory positioning. Basically, Lemma 3 implies that one of the three cases should occur at $S = c$: *i*) both curves are decreasing or *ii*) both curves are increasing or *iii*) while $G_{S\leq c}$ is increasing, $G_{S>c}$ is decreasing. All three cases are illustrated in Figure 1. In the first case, clearly the optimal inventory position is below c and computed based on $G_{S\leq c}$. In this case, the service provider's optimal exchange inventory is lower than the number of available servers and as such, there will be no queue for service. In the second case, the optimal inventory position must be above c implying that the optimal inventory level is above the number of available servers. The last case implies that at optimality, the exchange inventory level exactly matches the number of available service lines.

Based on this observation, we can derive the optimal inventory policies for the MRO service provider. We let S_L^* and S_H^* represent the integer inventory position values that maximize $G_{S\leq c}$ and $G_{S>c}$ respectively. Although we cannot derive closed form functions for these values, they can be easily computed using a simple line search.

Proposition 1. *For given arrival rates (λ), service rates (μ), and number of servers (c) the optimal inventory position, S^*, is such that*

i) $S^* = S_L^*$ *where* $S_L^* < c$, *if and only if*

$$h/r > \lambda\left(\frac{A_{c-1}}{A_c} - \frac{A_{c-2}}{A_{c-1}}\right). \tag{17}$$

ii) $S^* = c$ *if and only if*

$$\lambda\left(\frac{A_{c-1}}{A_c} - \frac{A_{c-2}}{A_{c-1}}\right) \geq h/r \geq \lambda c^c z^c \left(\frac{c^c z^{c+1} - c!(z-1)A_c}{c!A_c(c!A_c + c^c z^{c+1})}\right). \tag{18}$$

iii) $S^* = S_H^*$, *where* $S_H^* > c$, *if and only if*

$$h/r < \lambda c^c z^c \left(\frac{c^c z^{c+1} - c!(z-1)A_c}{c!A_c(c!A_c + c^c z^{c+1})}\right). \tag{19}$$

Proof of Proposition 1. The inequality (17) is derived from $G_{S\leq c}(c-1) - G_{S\leq c}(c) > 0$. That is, the inequality holds if and only if the profit with $S = c - 1$ is higher compared to $S = c$. This clearly indicates that in the latter case the profit function has a negative slope. Form Lemma 3, we can conclude that $G_{S>c}$ must have a negative slope as well. Since both functions are known to be unimodular (as shown by Lemmas 1 and 2), the optimal inventory position must be strictly below s and hence, the optimal solution should be determined based on (2). If the inequality does not hold, from concavity, it is straightforward to see that optimal inventory position must be equal to or greater than *s*. Hence, the first part (*i*) holds.

The proof of part (*iii*) is similar. The inequality in (19) is directly reduced from $G_{S>c}(c+1) - G_{S>c}(c) > 0$. That is, the inequality holds if and only if the profit defined in (10) with $S = c + 1$ is higher compared to $S = c$. This implies that at the latter point the profit function has a positive slope and as such, from Lemma 3, $G_{S\le c}$ must have a positive slope as well. Since both functions are unimodular, the optimal inventory position must be strictly above *s* and hence, the optimal solution should be determined based on (10). If the inequality does not hold, from concavity, it is straightforward to see that optimal inventory position must be equal to or lower than *s*.

The deductions obtained in the proofs of parts (*i*) and (*iii*) clearly implies that when (18) holds the optimal inventory position is equal to the number of service lines.

The result above characterizes the optimal inventory policy for the MRO service provider, which is primarily driven by the ratio of the holding cost to the per-unit revenue. As expected, when this ratio exceeds a certain threshold, the optimal inventory level falls below the number of available service lines, resulting in a system without queues and with idle servers. Conversely, when the ratio falls below a lower threshold, it becomes optimal to maintain an inventory level that exceeds the number of service lines. In the intermediate range between these two thresholds, the optimal policy is to align the inventory level exactly with the number of servers. More generally, the optimal inventory position is non-decreasing in the per unit revenue (*r*) and non-increasing in the holding cost (*h*), as expected.

The threshold curves are illustrated in Figure 2. As expected, for a large number of server lines, the optimal inventory level tends to fall below the number of service lines. When the holding cost-to-revenue ratio (h/r) is sufficiently high, it is not economically optimal for the service provider to maintain an inventory level exceeding the number of servers. Interestingly, in certain scenarios, the optimal inventory level initially matches the number of service lines, then exceeds it as the number of servers increases, before returning to match it again, and eventually falling below. This non-monotonic behavior is exemplified by the dashed curve in the middle in Figure 2. The implication is that the relationship between the number of servers and the optimal inventory level is not strictly monotonic. This outcome can be attributed to the diminishing marginal impact of each additional server on overall system performance. A detailed numerical analysis exploring the influence of system parameters on optimal inventory levels is presented in the next section.

4.1 Numerical Analysis for the Pure Exchange Policy

In this section, we investigate the impact of model parameters on optimal inventory policies using numerical analysis. Our analysis is based on 2,440 instances obtained by varying values of arrival rates, service rates, number of servers, and holding cost to revenue (HCR) ratio (i.e., h/r). Figure 3 summarizes the overall impact of the traffic density (*z*) and HCR ratio on the optimal inventory decisions and the associated profits in a scatter plot. We observe that the impact of the traffic density on the optimal inventory policy outcome is not always monotonic. Typically, the optimal inventory levels (and also the optimal expected profits) first increases and then decreases with traffic density. However, this pattern cannot be generalized as it depends on the specific composition of the arrival rates, service rates, and the number of service lines. To obtain more comprehensive results we investigate the influence of these parameters individually.

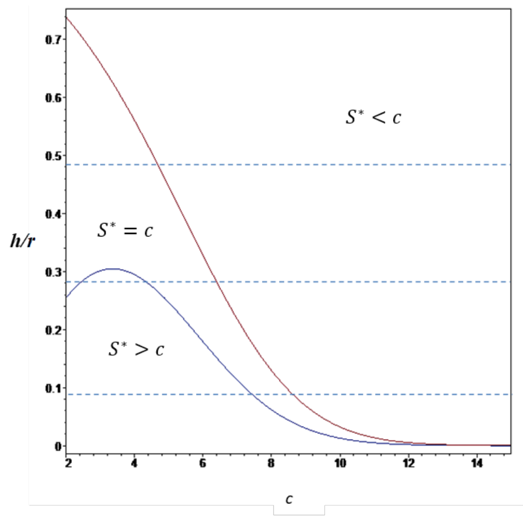

Figure 2. Threshold lines as functions of the number of servers

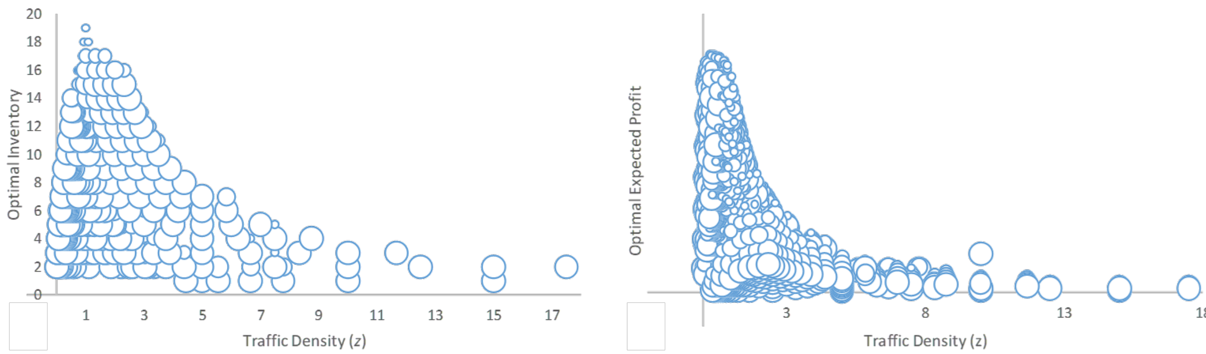

Figure 3. Traffic density and the optimal outcome (The bubble size represents $\boldsymbol{h/r}$ varying between 0.1 and 0.7)

As depicted in Figure 4, we first examine the impact of demand and number of service lines on the optimal inventory position. In most cases, our observations are consistent with our earlier conclusion regarding to the influence of traffic density. Specifically, the optimal inventory level initially increases and then decreases as the number of servers increases. When the number of servers is limited, service capacity rather than inventory becomes the bottleneck for the system. Increasing the number of servers enhances the return on inventory investment up to a certain point. Beyond that point, as the service rate becomes less of a bottleneck due to the increased number of servers, fewer inventories are needed because inventory can be replenished more quickly. This pattern typically emerges when the arrival rate is sufficiently high.

However, when the arrival rate is low, the effect of the number of servers is more ambiguous. In this case, when the holding cost is low, the optimal inventory position is generally non-increasing with the number of servers. The opposite holds when holding costs are high. In the former case, because the holding cost is low, it is optimal to maintain a relatively large inventory with few servers so as to compensate for the low effective service rate. As the effective service rate increases, so does the inventory turnaround, reducing the need for an exchange inventory buffer. In the latter case, low turnarounds resulting from few servers do not justify investment in inventory when holding cost is high. As the number of servers increases, improved turnaround performance supports maintaining a larger inventory position.

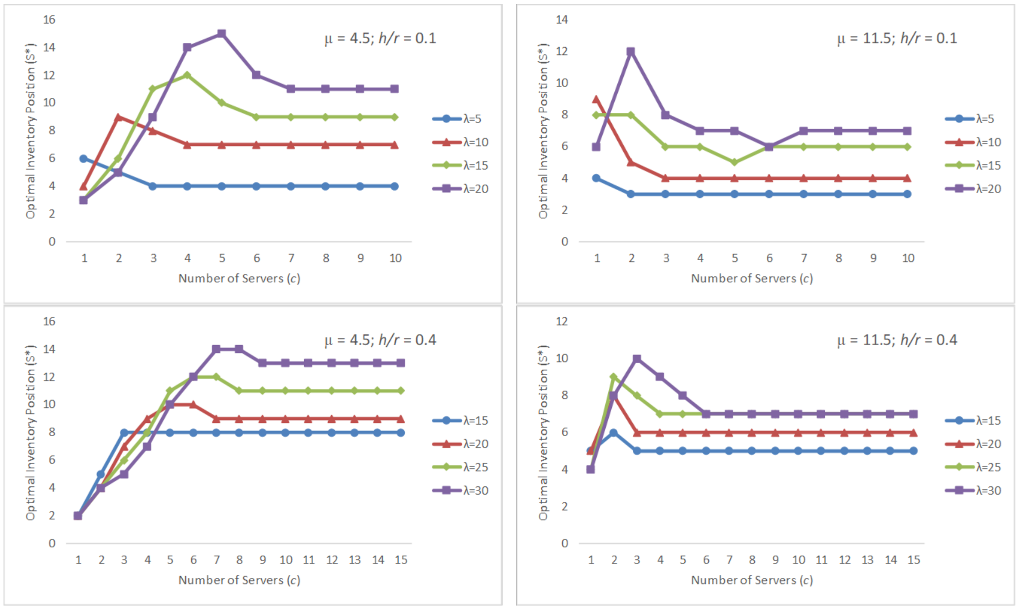

Figure 4. Impact of arrival rates and number of servers on optimal inventory position

Another interesting observation is that when there are few servers, higher arrival rates lead to lower inventory levels, whereas when the number of servers is sufficiently high, the opposite is true. As mentioned above, in the former case, the service rate becomes a bottleneck as arrival rates increase, which reduces inventory turnover and leads to lower inventory levels. In the latter case, with a high number of servers, higher arrival rates improve inventory turnover, resulting in increased inventory levels. When the number of service lines is moderate, neither low nor high—the relationship between arrival rates and optimal inventory becomes non-monotonic. In order to better understand this pattern, we first make the following observation regarding the trade-off between the arrival rates and the optimal inventory:

Lemma 4. *Assuming* $r\mu > h$, *as* $\lambda \to \infty$, $S^* \to c$.

Proof of Lemma 4. First, note that as the arrival rate grows too large there is always demand for any completed job. Therefore, the effective demand is as high as the throughput of the overall service system. As such the effective demand rate is $S\mu$ when $S<c$ and $c\mu$ when $S \geq c$. In the former case, the mean profit increases in S since $r\mu > h$ whereas in the latter case it decreases with S since the revenue is a function of S while the total holding cost is increasing with S. Consequently, optimal inventory position, S^*, converges to the number of service lines, c.

◊

The above result simply states that with sufficiently high demand rate, the optimal inventory position converges to the number of servers. This is intuitive in that at high arrival rate the system can admit arrivals as fast as its maximum throughput and as such, the service company cannot be better off by maintaining an inventory position either below or above c. In general, the impact of arrival rates on the optimal exchange inventory is as illustrated in Figure 5. Given that all other parameters are constant, the optimal inventory first increases with λ to a certain point and then decreases until it converges to c. When the demand rate is sufficiently low but increasing, the service company needs to increase its exchange inventory position in order to cope with increased demand. However, once the traffic density reaches a tipping point (usually around $\lambda/(c\mu) = 1$), the arrival rates become too dense that the work-in-process inventory (S_w) and thus, the service queue inflate. In the steady state, the system will rarely have ready-to-exchange finished inventory (*i.e.*, $S_f \sim 0$). Consequently, the overall service rate becomes the bottleneck for the system rather than the inventory. In this case, additional inventory will be rarely used and thus incur more cost than revenue. As the arrival rate further increases, the inventory level will converge to the number of servers.

In the case of service rates, we observe that the inventory level converges to one as service rate grows too big.

Lemma 5. *Assuming,* $r\lambda > h$, *as* $\mu \to \infty$, $S^* \to 1$.

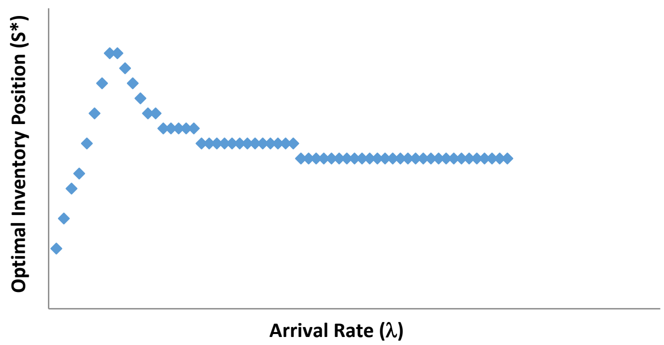

Figure 5. Arrival rate vs. optimal inventory

The proof follows directly from the observation that an infinite service rate implies instantaneous service completion, and therefore, immediate replenishment of the exchange inventory. Consequently, there will be no need to carry more than a single unit of inventory. Consistent with Lemma 5, the optimal inventory level generally decreases as the service rate increases. However, as illustrated in Figure 6, this relationship is not strictly monotonic. In certain cases—particularly when the service rate is low, and the arrival rate is high—the optimal inventory level initially increases with the service rate before eventually declining. This behavior arises because high arrival rates, when coupled with limited service capacity, lead to slower turnover and a reduced number of completed overhauls, thereby lowering the available exchange-ready inventory. As the service rate improves, it compensates for the high demand by accelerating turnover, which increases the revenue generated from exchanges. In this range, maintaining a higher inventory level becomes beneficial. However, as the service rate continues to increase, the system becomes more efficient, reducing the need for additional inventory. Thus, the optimal inventory level eventually declines with further increases in service rate.

In general, we conclude that the optimal inventory level tends to be lower under both low and high traffic density conditions. When the system load (i.e., traffic density) is low, the return on investment from maintaining a large exchange inventory does not justify the associated holding costs, as incoming demand is insufficient to utilize the inventory effectively. Conversely, when the system is heavily loaded, the service process becomes the primary bottleneck. Prolonged replenishment times hinder the turnover of inventory, limiting the benefit of maintaining a large exchange pool in steady state. As a result, in both extremes, a relatively smaller inventory position is optimal.

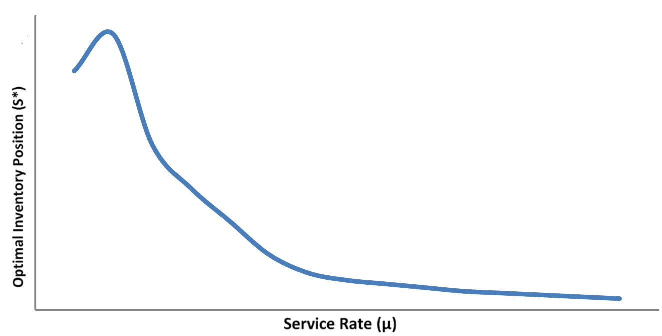

Figure 6. Service rate vs. optimal inventory ($c = 9$, $\boldsymbol{\lambda} = 25$, $h = 25$)

5. Conclusions

In this paper we investigate optimal exchange inventory policies for a maintenance and repair service provider operating under uncertainty in both order arrivals and service times. We model the system using Markov processes for both arrival and service rates, with parallel servers representing dedicated work cells or teams. Under the exchange policy considered, the service provider fulfills customer orders only if there is inventory available for immediate exchange. As such, the order admission process—and, by extension, the effective demand—is directly governed by the inventory position. Within this setting, we derive optimality conditions for maximizing the service provider's profit and analyze the influence of cost and operational parameters on the optimal inventory level.

Our analysis revealed that the ratio of holding cost to marginal revenue has a predictable and monotonic impact on inventory decisions: as this ratio increases, the optimal inventory level decreases (Figure 2). In contrast, the impact of operational parameters—such as arrival rates, service rates, and the number of servers—on the optimal inventory position is more nuanced and non-monotonic. Specifically, the optimal inventory level increases with the arrival rate up to a certain point, and then decreases, eventually converging to the number of servers (Figure 5). A similar trend is observed with respect to the service rate, where the optimal inventory converges to one as the service rate becomes very high (Figure 6). Regarding the number of servers, the optimal inventory is either non-increasing or follows a non-monotonic pattern, depending on the interplay between arrival and service rates (Figure 4). These findings indicate that diminishing returns from holding additional inventory can eventually erode the profitability advantage conferred by the exchange model.

An interesting extension of this work would be to study hybrid fulfillment strategies, where the service provider can flexibly choose between offering an exchange or scheduling a conventional repair based on system state and customer characteristics. Such a setting would introduce dynamic decision-making, where the provider balances quick turnaround (via exchange) and lower cost (via traditional repair) based on inventory levels, queue length, or customer priority. Incorporating this flexibility into the model could yield valuable insights into how service differentiation and dynamic order allocation influence overall profitability, resource utilization, and customer satisfaction. Additionally, coupling inventory decisions with dynamic pricing mechanisms could further enhance decision support for MRO firms, especially in markets with heterogeneous customer valuations and urgency levels.

References

Arts, J. and Flapper, S. D., Aggregate overhaul and supply chain planning for rotables, *Annals of Operations Research*, vol. 224, no. 1, pp. 77-100, 2015.

Avşar, Z. M. and Zijm, W. H., Approximate queueing models for capacitated multi-stage inventory systems under base-stock control, *European Journal of Operational Research*, vol. 236, no. 1, pp. 135–146, 2014.

Buyukkaramikli, N. C., van Ooijen, H. P., and Bertrand, J. W. M., Integrating inventory control and capacity management at a maintenance service provider, *Annals of Operations Research*, vol. 231, no. 1, pp. 185–206, 2015.

Chandola, D. C., Jaiswal, K., Verma, S., and Singh, B., Aviation MRO: A comprehensive review of factors affecting productivity of Aircraft Maintenance Organization, In *2022 Advances in Science and Engineering Technology International Conferences (ASET)*, pp. 1-7. IEEE, 2022.

Dekker, R., Applications of maintenance optimization models: a review and analysis, *Reliability Engineering and System Safety*, vol. 51, pp. 229–240, 1996.

Dinis, D., Maintenance management: A review on problems and solutions, *Procedia Computer Science*, vol. 253, pp. 3069–3077, 2025.

Dudin, A. N., Manzo, R., and Piscopo, R., Single server retrial queue with group admission of customers, *Computers & Operations Research*, vol. 61, pp. 89–99, 2015.

Erkoc, M. and Ertogral, K., Overhaul planning and exchange scheduling for maintenance services with rotable inventory and limited processing capacity, *Computers & Industrial Engineering*, vol. 98, pp. 30–39, 2016.

Erkoc, M. and Romo-Fragoso, S., Order admission and optimal pricing for regular jobs and big deals at a service company, *International Journal of Operations Research and Information Systems*, vol. 6, no. 1, pp. 1-18, 2015.

Ertogral, K., Erkoc, M., and Ulker, D. H., A production scheduling model and analysis for the maintenance repair and overhaul service providers, *Journal of the Faculty of Engineering and Architecture of Gazi University*, vol. 30, no. 3, pp. 513-522, 2015.

Ertogral, K. and Ozturk, F. S., An integrated production scheduling and workforce planning model for the maintenance and repair operations in airline industry, *Computers & Industrial Engineering*, vol. 127, pp. 832–840, 2019.

Gans, N. and Savin, S., Pricing and capacity rationing for rentals with uncertain durations, *Management Science*, vol. 53, no. 3, pp. 390–407, 2007.

Gross, D. and Harris, C. M., *Fundamentals of Queuing Theory*, John Wiley & Sons, New York, NY, 1998.

Hu, Q., Boylan, J. E., Chen, H., and Labib, A., OR in spare parts management: A review, *European Journal of Operational Research*, vol. 266, no. 2, pp. 395–414, 2018.

Joo, S. J., Scheduling preventive maintenance for modular designed components: A dynamic approach, *European Journal of Operational Research*, vol. 192, no. 2, pp. 512–520, 2009.
Levi, R., Magnanti, T., Muckstadt, J., Segev, D., and Zarybnisky, E., Maintenance scheduling for modular systems: modeling and algorithms, *Naval Research Logistics*, vol. 61, no. 6, pp. 472–488, 2014.
Luh, P. B., Yu, D., Soorapanth, S., Khibnik, A. I., and Rajamani, R., A Lagrangian relaxation based approach to schedule asset overhaul and repair services, *IEEE Transactions on Automation Science and Engineering*, vol. 2, no. 2, pp. 145–157, 2005.
Mirzahosseinian, H. and Piplani, R., A study of repairable parts inventory system operating under performance-based contract, *European Journal of Operational Research*, vol. 214, no. 2, pp. 256–261, 2011.
Mofokeng, T., Mativenga, P. T., and Marnewick, A., Analysis of aircraft maintenance processes and cost, *Procedia CIRP*, vol. 90, pp. 467–472, 2020.
Muckstadt, J. A., *Analysis and Algorithms for Service Parts Supply Chains*, Springer Series in Operations Research and Financial Engineering, Springer, Berlin, Germany, 2005.
Örmeci, E. L., Burnetas, A. N., and Emmons, H., Dynamic policies of admission to a two-class system based on customer offers, *IIE Transactions*, vol. 34, no. 9, pp. 813–822, 2002.
Perlman, Y., Mehrez, A., and Kaspi, M., Setting expediting repair policy in a multi-echelon repairable-item inventory system with limited repair capacity, *Journal of the Operational Research Society*, vol. 52, no. 2, pp. 198–209, 2001.
Savin, S. V., Cohen, M. A., Gans, N., and Katalan, Z., Capacity management in rental businesses with two customer bases, *Operations Research*, vol. 53, no. 4, pp. 617–631, 2005.
Spanjers, L., van Ommeren, J. C., and Zijm, W. H., Closed loop two-echelon repairable item systems, *OR Spectrum*, vol. 27, no. 2–3, pp. 369–398, 2005.
van Horenbeek, A., Bure, J., Cattrysse, D., Pintelon, L., and Vansteenwegen, P., Joint maintenance and inventory optimization: a review, *International Journal of Production Economics*, vol. 143, no. 2, pp. 499–508, 2013.
Vieira, D. R. and Loures, P. L., Maintenance, repair and overhaul (MRO) fundamentals and strategies: An aeronautical industry overview, *International Journal of Computer Applications* 135, no. 12 pp. 21-29, 2016.
Zijm, W. H. and Avşar, Z. M., Capacitated two-indenture models for repairable item systems, *International Journal of Production Economics*, vol. 81, pp. 573–588, 2003.

Biographies

Murat Erkoc is a professor at the Department of Industrial and Systems Engineering at the University of Miami, Florida. He holds a PhD degree in industrial engineering from Lehigh University. His research interests include supply chain management, game theoretical modeling, operations management, stochastic processes, and large-scale optimization. He is a member of the Institute of Industrial Engineers (IIE), Institute for Operations Research and Management Science (INFORMS), and Alpha Pi Mu.

Salvador Romo-Fragoso is the Chief Executive Officer & Owner at Sthall LLC. He holds a PhD degree in industrial engineering from the University of Miami, Florida. He was formally the Americas Service and Repair Customer Support Director at Hewlett-Packard. His research interests include operations research, strategic planning, operations digital mapping, and data analytics. He is a member of the Institute of Industrial Engineers (IIE).