# Reducing costs of repairable spare parts supply systems via dynamic scheduling 

## Citation for published version (APA):

Tiemessen, H. G. H., \& Houtum, van, G. J. J. A. N. (2010). Reducing costs of repairable spare parts supply systems via dynamic scheduling. (BETA publicatie : working papers; Vol. 329). Technische Universiteit Eindhoven.

## Document status and date:

Published: 01/01/2010

## Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

## Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.
Link to publication


## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25 fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

## Take down policy

If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.

# Reducing costs of repairable spare parts supply systems via dynamic scheduling 

H.G.H. Tiemessen, G.J. van Houtum

Beta Working Paper series 329

| BETA publicatie | WP 329 (working <br> paper) |
| ---: | :--- |
| ISBN | $978-90-386-2368-9$ |
| ISSN |  |
| NUR | 804 |
| Eindhoven | September 2010 |

# Reducing costs of repairable spare parts supply systems via dynamic scheduling 

H.G.H. Tiemessen ${ }^{1 *}$,<br>G.J. van Houtum ${ }^{2}$,<br>${ }^{1}$ Department of Mathematical and Computational Sciences, IBM Research Lab Zurich<br>Saümerstrasse 4, 8803 Rüschlikon, Switzerland<br>${ }^{2}$ School of Industrial Engineering, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, the Netherlands

September 15, 2010


#### Abstract

We study a system consisting of one repair shop and one stockpoint, where spare parts of repairables are kept on stock to serve an installed base of systems. Part requests are met from stock if possible, and backordered otherwise. Our objective is to determine initial stock levels and a policy for scheduling repair jobs such that holding and backorder cost are minimized. We propose two dynamic scheduling rules, compare their performance with the static priority rule, and show that even when stock levels and static priorities have been optimized simultaneously, dynamic scheduling rules often reduce total cost by more than $10 \%$.


Keywords: Inventory control, spare parts, dynamic scheduling, discrete event simulation.

[^0]
## 1 Introduction

In this paper, we study the provisioning of repairable spare parts for critical components of advanced technical equipment, such as airplanes, military systems, and large computer systems. Such equipment is often critical for the primary processes of its users, and thus high fractions of up-time are required. For high-tech systems, usually a large part of the total support cost consists of spare parts cost. This is for spare parts usage, for having spare parts on stock in locations at close distance of the installed systems, for repair of failed parts, and for transportation of the parts.

The inventory control for spare parts may have a large effect on the total spare parts cost. What has been studied extensively is the so-called system approach, in which the inventory control is directly focused on availability of systems instead of target service levels for individual Stock Keeping Units (SKU-s). This has been studied in single-location and multi-echelon settings, and it has been shown that, under given system availability constraints, the system approach may lead to large reductions in holding cost compared to the straightforward item approach; see Thonemann et al. (2001), and Rustenburg et al. (2002). In these comparisons, leadtimes for repair of failed spare parts are considered as given. Obviously, the optimal holding cost decreases if one would be able to reduce these repair leadtimes. Unfortunately, reducing repair lead times for all SKU-s usually requires additional investments in the logistics network and/or the capacity and productivity in the repair facility. An interesting research direction is however to investigate how one can schedule work at the repair facility such that the holding cost can be reduced without additional investments in repair capacity. A recent contribution of Adan et al. (2009) examined this challenging problem and showed that, in comparison to FCFS scheduling, large cost reductions (around 40\%) can be obtained by static priorities (and simultaneous optimization of the base stock levels). In this paper we study the same problem setting and build upon their results and observations.

The goal of this paper is to study the benefits of dynamic scheduling rules on the total system performance. We study a spare parts supply system consisting of one repair facility, one stockpoint, and multiple repairable SKU-s. Ready-for-use parts are kept on stock in a single stockpoint to serve an installed base of technical systems. When a part of one of the technical systems fails, the failed part is immediately sent to the repair facility and at the same time a ready-for-use part is requested at the stockpoint. Such a request is fulfilled immediately if there is a part of the requested SKU on stock and otherwise the request is backordered and fulfilled later. Each backordered request corresponds to a technical system that is down. The objective of our model is to minimize the sum of annual holding cost and expected annual down-time cost.

The motivation behind this research is that we expect to be able to further improve the system approach by using dynamic priorities instead of static priorities as proposed by Adan et al. (2009). We are encouraged by studies in Make-to-Stock (MTS) literature that show that dynamic scheduling rules often outperform static scheduling rules (see e.g. Perez and Zipkin, 1997).

To minimize the expected annual total cost, we must provide solutions for two types of problems: a tactical planning problem (calculating appropriate base stock levels), and an operational planning problem (calculating an appropriate dynamic scheduling policy). In this paper we focus on the operational planning problem and 'borrow' existing methods to obtain solutions for the tactical planning problem. We assume that there is only one opportunity to decide upon the base stock levels. This means that it is not allowed to acquire additional parts (or to sell off excess stock) later. We assume that the failure rates for all SKU-s are constant over time and known when the base stock levels must be calculated.

In this paper we introduce two dynamic scheduling rules that use online stock level information to schedule work at the repair facility. To examine the performance of our scheduling rules, we first investigate its optimality gap. We are able to do this for sufficiently small problem instances by formulating the problem as a Markov Decision Problem (MDP). Then, we examine the performance of our dynamic scheduling rules on problems with many SKU-s. For this, we use the same test bed as described in Adan et al. (2009). As we are not able to evaluate the performance our dynamic scheduling rules analytically, we will use discrete event simulation.

Our work contributes to a rich literature on spare parts inventory models. The literature that is most related to the work in this paper, comes from three streams: In the first stream of literature, ample repair capacities are assumed, the models are focused on optimal control for multiple items in multi-echelon systems, and targets are typically set in terms of system availability. This stream started with the seminal paper of Sherbrooke (1968) on the METRIC model. Since then, numerous extensions and variants have been developed such as MOD-METRIC (Muckstadt, 1973), and VARI-METRIC (Slay, 1984). For a comprehensive overview of this stream, see the references in Sherbrooke (2004) and Rustenburg et al. (2003). The assumption of ample repair capacities in this research stream facilitates the analysis and enables that systems with many SKU$s$ can be optimized. However, the assumption of ample repair capacity is not always justified. It can lead to a poor estimation of system performance and a poor allocation of stocks in systems with highly utilized repair facilities (cf. van Harten and Sleptchenko, 2003, and Sleptchenko et al., 2002).

Therefore, in a second stream of literature, various ways to model finite repair capacities have been studied. Most papers in this stream are based on queuing type models with exponential servers and First-Come First Served (FCFS) scheduling discipline; see Gross et al. (1983), Albright and Gupta (1993), Zijm and Avsar (2003). In much of this work, the focus is on the development of approximate evaluation algorithms, and if optimization is applied, this is generally limited to systems with limited numbers of SKU-s only.

In a third stream of literature, dynamic priority rules have been studied. These are rules that use the state of the system (e.g. online stock level information and work-in-progress information) when making scheduling decisions. Dynamic priority rules lie between static policies such as FCFS
and dynamic programming approaches that are interesting from a scientific point of view but impractical for real-life use because of their computational intractability. Several studies pointed out that scheduling rules that use online stock level information outperform scheduling rules that do not; see e.g. Hausman and Scudder (1982), and Pyke (1990).

A significant part of relevant papers in this stream comes from the multi-product, finite capacity, make-to-stock literature: Hax and Meal (1973) describe the development of a hierarchical planning and scheduling system for a multi-location multi-item seasonal demand environment. They consider (among many other things) the problem of calculating appropriate production quantities for all products for the next time period and propose to allocate production capacity among items in the same product family such that the expected run-out times are equalized. Perez and Zipkin (1997) develop myopic allocation rules for a stochastic production-inventory system where several products share a single processor of limited capacity. Johnson and Scudder (1999) study a scheduling problem in an MTS environment where several different products compete for production time on a single assembly line. They indicate that scheduling rules which consider the inventory position and demand forecast information outperform traditional fixed cycle rules. Their problem formulation is related to the well-known Stochastic Economic Lot sizing and Scheduling Problem (Stochastic ELSP). For a comprehensive survey on Stochastic ELSP we refer to Winands et al. (2011).

Two recent and very relevant contributions in this third stream have been made by Caggiano et al. (2006), and Kat and Avsar (2009). Caggiano et al. (2006), develop an integrated model for making real-time repair and inventory allocation decisions in a two-echelon repairable spare parts system with one central repair facility. Their model is a finite-horizon, periodic-review, mathematical programming model. The capacity of the repair facility is a decision variable and is expressed as a maximum on the number of parts that can be repaired during one time period. They propose to use their model in a rolling horizon setting, solving the mathematical model at the beginning of every new time period. Kat and Avsar (2009), study a problem similar to ours. An important difference is however that they incur a fixed backorder cost for each backordered request regardless of the time needed to satisfy the backordered demand. Based on numerical investigations, they show that the optimal policy to minimize the sum of holding cost and expected backorder cost is a stationary base stock policy with switching curves and fixed (not state-dependent) base-stock levels. Both Caggiano et al. (2006), and Kat and Avsar (2009) do not focus on optimizing base stock levels, and their analysis is limited to small numbers of SKU-s.

Our work fits in the third stream of research, and our main contribution consists of two parts:

- We present two new dynamic scheduling rules for a repairable spare parts system with one stocking location and an arbitrary number of SKU-s where the objective is to minimize the sum of holding cost and expected backorder cost: the Equalization of Backorder Times rule (EBT-rule), and the Myopic rule (M-rule). We compare these two scheduling rules with
the optimal scheduling rule and show that the optimality gap is small ( $<1 \%$ ) for most problem instances. For 'unbalanced' problem instances we show that the optimality gap of the EBT-rule can become big, whereas the M-rule still performs well. For technical reasons, this comparison is limited to relatively small problem instances, but there is no reason to believe that the conclusions would change for larger problem instances.
- We propose a new solution method for optimizing base stock levels and scheduling decisions. The solution consists of two parts: For calculating base stock levels, we use the method for the joint calculation of base stock levels and static priorities as proposed by Adan et al. (2009). For calculating scheduling decisions we do not use the optimized static priorities, but plug in one of our two dynamic scheduling rules.
We create a test bed of 540 problem instances of real-life size and show that the dynamic scheduling rules outperform the static priority rule. We find average reductions of the annual total cost of $8.9 \%$ for the EBT-rule and $9.6 \%$ for the M-rule. The cost reductions follow from reductions in the expected number of backorders of $31.2 \%$ and $34.9 \%$, respectively. Obviously, further improvements might be possible when we optimize the base stock levels for our proposed dynamic scheduling rules. We leave this for future research.

The organization of this paper is as follows. In Section 2, we formulate our model. Next, in Section 3, we describe all basic techniques in the development and evaluation of all algorithms and heuristics. In Section 4, we execute numerical experiments to test the effectiveness of various algorithms and heuristics. Finally, we conclude in Section 5.

## 2 Model

In this section we first describe the spare part supply system and formulate our model.

### 2.1 System Description

We consider a single location with one repair facility and one stockpoint, where spare parts of multiple, critical repairables are kept on stock to serve a set of technical systems installed in a certain region; this set of technical systems is also called the installed base. We distinguish $N$ repairables SKU-s, which are numbered $1, \ldots, N$. These SKU-s occur in the configurations of the technical systems and are subject to failures. When a part fails in a technical system, a demand is immediately placed for a ready-for-use part of the same SKU at the stockpoint and the failed part is sent to the repair facility. If the requested ready-for-use part is on stock, then the demand is immediately fulfilled. Otherwise, the demand is backordered and fulfilled as soon as a ready-foruse part of the requested SKU becomes available from the repair facility. In the latter case, the
technical system where the part failed is down until a ready-for-use part is delivered. This is of course a situation we like to prevent.

Figure 1 gives a graphical representation of the system.


Figure 1: The spare parts supply system
The lifetime of the technical systems is assumed to be long (e.g., $20-30$ years). We therefore formulate a model with an infinite time horizon and policies for the inventory control and the job scheduling at the repair facility are analyzed via a steady-state analysis. Time is continuous and the time horizon is $[0, \infty)$. We assume that, for the total installed base, failures of an SKU occur according to a Poisson process with a constant rate. The failure rate for SKU $n$ is given by $\lambda_{n}$.

Let $S_{n} \geq 0$ be the number of spare parts of SKU $n$ that is taken on stock at time instant $t=0$. We assume that all failed parts can always be repaired. This implies that the inventory position of SKU $n$ always stays equal to $S_{n}$. This means that at any time instant $t>0$ it holds that $O H_{n}(t)+I R_{n}(t)-B O_{n}(t)=S_{n}$, where $O H_{n}(t), I R_{n}(t)$, and $B O_{n}(t)$ denote the physical stock on hand, the number of parts inside the repair facility (under repair or in the queue), and the number of backorders at time instant $t$. We call $S_{n}$ the basestock level for SKU $n$. The vector $\mathbf{S}=\left(S_{1}, \ldots, S_{n}\right)$ contains the base stock levels for all SKU-s.

The repair facility is modeled as a single exponential server, which repairs the failed parts of all SKU-s. All repair times are exponentially distributed and mutually independent. For all SKU-s, we assume the same repair rate $\mu$. To obtain a stable system, we assume $\mu>\sum_{n=1}^{N} \lambda_{n}$. operations at the repair facility are controlled by a scheduling policy $\pi$. This policy may use all available
information to select the failed part that is put into repair. For the repair facility we assume preemption, i.e. repair of a failed part can be interrupted (and resumed) at any time. Notice that, because of the assumption of exponential repair times, the remaining repair time of an interrupted repair has the same exponential distribution as a new repair.

For the cost, we distinguish holding cost and backorder cost. The holding cost consist of investment cost and storage cost. The procurement cost are made at time $t=0$ and are linear in the base stock level $S_{n}$. These cost may be translated into a linear cost that is paid per spare part per time unit. The storage costs are paid per time unit and they are linear in the physical stock on hand. We assume that the procurement costs dominate. Hence, the holding cost for SKU $n$ are taken equal to $h_{n} S_{n}$, where $h_{n}>0$ is the holding cost parameter of SKU $n$. We pay a penalty cost $b>0$ per time unit per technical system that is down because of a lack of spare parts. We assume that different backordered demands correspond to different technical systems. This implies that the total backorder cost can be measured via the total number of backordered parts. Let $E B O_{n}(\mathbf{S}, \pi)$ denote the mean number of backordered demands of SKU $n$ in steady state; this mean only depends on the steady-state number of failed parts of SKU $n$ in the repair facility, which is determined by the scheduling policy $\pi$, and the basestock level $S_{n}$. The average backorder cost for systems that are down due to a temporarily lack of spare parts of SKU $n$ is equal to $b E B O_{n}(\mathbf{S}, \pi)$ and the total average backorder cost is equal to $b \sum_{n=1}^{N} E B O_{n}(\mathbf{S}, \pi)$. Our objective is to choose the basestock levels $S_{n}$ and the scheduling rule $\pi$ such that the sum of the holding cost and the expected backorder cost is minimized. The mathematical formulation of this optimization problem is as follows:

$$
(P) \quad \min _{\mathbf{S}, \pi} \sum_{n=1}^{N}\left[h_{n} S_{n}+b E B O_{n}(\mathbf{S}, \pi)\right]
$$

The main difficulty in $(P)$ is the evaluation of $E B O_{n}(\mathbf{S}, \pi)$ for given basestock levels $S_{n}$ and a given scheduling rule $\pi$. Model $(P)$ is very similar to the non linear integer optimization problem formulated by Adan et al. (2009). In fact, their model is a specialization of our model, where they restrict the analysis to scheduling rules that can be expressed in terms of a mapping from every SKU to an integer-valued priority. This mapping is calculated once, and during execution their scheduling rule selects a failed part belonging to the SKU with the highest priority.

## 3 Analysis

In this section, we describe various methods to calculate base stock levels and/or scheduling rules. These methods serve as building blocks in our heuristics and algorithms to solve the integrated problem. We start with formulating the scheduling problem for given base stock levels as an MDP in order to obtain the optimal scheduling policy. Next, we introduce the EBT-rule and the M-rule for scheduling work at the repair facility. Then, we shortly describe the method of Adan et al.
(2009) for joint optimization of base stock levels and static priorities. We conclude this section with a description on how we have used discrete event simulation to obtain accurate estimates for the annual total cost for problem instances of real-life size and dynamic scheduling rules.

### 3.1 Optimal scheduling rule

In this subsection we derive the optimal scheduling rule at the repair facility for given base stock levels (i.e. solve $(P)$ for given $\mathbf{S}$ ). First, we show that the optimal scheduling decision in situations with at least one outstanding backorder is straightforward. Next, we formulate our problem as an MDP. Finally, we truncate the state space in order to arrive at an MDP with finite state space. Once this has all been done, we use relative value iteration to solve the Bellman equations for optimal control in an MDP with average cost. This gives use the optimal scheduling rule.

Now we take a closer look at the situation where there is at least one backorder (or equivalently: there is at least one SKU with negative net stock). Assume that SKU $n$ has negative net stock, and that SKU $m$ has non-negative net stock. Then, it seems attractive to prioritize SKU $n$ over SKU $m$ as SKU $n$ is causing backorder cost (at a rate of $b$ per time unit) whereas SKU $m$ is not. Under different assumptions, it might be that scheduling SKU $m$ before SKU $n$ might bring advantages in the future. In our setting however, all SKU-s have identical backorder cost and repair times and thus there is no reason to postpone SKU $n$ in favor of another SKU. We now make following statements regarding the optimal scheduling policy: (i) the optimal scheduling rule will always prefer an SKU with negative net stock over an SKU with a non-negative net stock, and (ii) the optimal scheduling rule may use any criterion to prioritize among failed parts belonging to SKU-s with negative net stock. Both properties can be proven easily using sample path arguments.

We now formulate the problem as an MDP. First, we define the state $\mathbf{x}=\left\{x_{1}, \ldots, x_{N}, x_{N+1}\right\}$ as an $N+1$ dimensional vector that holds the maximum of zero and the net stock level for all SKU-s, and the sum of all backordered items. Let $z_{n}$ be the net stock of SKU $n$, then we obtain:

$$
x_{i}= \begin{cases}\max \left(0, z_{i}\right) & 1 \leq i \leq N  \tag{1}\\ \sum_{j=1}^{N}\left[\max \left(0,-z_{j}\right)\right] & i=N+1\end{cases}
$$

This state definition (and in particular the fact that we do not keep track of negative net stock levels of individual SKU-s) helps us to limit the number of states to a minimum. In this state definition we have exploited the two properties we have derived for the optimal scheduling policy. For the state space $\mathcal{S}$, it holds that $\mathcal{S}=\mathcal{S}_{1} \cup \mathcal{S}_{2}$ with $\mathcal{S}_{1}=\left\{\left(x_{1}, \ldots, x_{N}, 0\right) \mid x_{i} \in\left\{0,1, \ldots, S_{i}\right\}, i=1, \ldots, N\right\}$, and $\mathcal{S}_{2}=\left\{\left(x_{1}, \ldots, x_{N}, x_{N+1}\right) \mid x_{N+1} \in \mathbb{N} \wedge x_{i} \in\left\{0,1, \ldots, S_{i}\right\} \wedge \exists i\left(x_{i}=0\right), i=1, \ldots, N\right\}$. Set $\mathcal{S}_{1}$ consists of all states in $\mathcal{S}$ with no backorders, whereas $\mathcal{S}_{2}$ consists of all states in $\mathcal{S}$ with at least one backorder. Because the number of aggregated outstanding backorders is unbounded, the state
space is infinite and we need to truncate the state space in order to use relative value iteration. State space truncation is discussed at the end of this subsection.

Because both interfailure times and repair times are exponentially distributed and mutually independent, all information that is relevant for making the optimal decision at a certain point in time is contained in the state $\mathbf{x}$. This implies that we only need to make a scheduling decision when a repair job is completed, or a new part failure occurs.

We now define the action space $\mathcal{A}$, and the set of admissible actions $\mathcal{A}(\mathrm{x})$ for all states $\mathrm{x} \in \mathcal{S}$. The action space for our model is $\mathcal{A}=\left\{a_{0}, \ldots, a_{N}, a_{N+1}\right\}$. Here, action $a_{n}(n \leq n \leq N)$ represents the decision to repair a part of SKU $n$, action $a_{N+1}$ stands for the decision to repair a (random) part from the backorder queue, and action $a_{0}$ stands for the decision to repair no part at all. Because it only makes sense to repair parts that have failed, we have that $\mathcal{A}(\mathbf{x})=\left\{a_{0}\right\}$ for all states $\mathbf{x}$ with $x_{i}=S_{i} \forall i$. Furthermore, $\mathcal{A}(\mathbf{x})=\left\{a_{N+1}\right\}$ for all $\mathbf{x} \in S_{2}$. And $\mathcal{A}(\mathbf{x})=\left\{a_{n} \mid 1 \leq n \leq N \wedge x_{n}<S_{n}\right\}$ for all other states.

We can now calculate the optimal expected cost per time unit, and the optimal action for every state $\mathbf{x}$ by solving the Bellman equations for an optimal policy for a Continuous Time MDP with infinite horizon and average cost (see Bertsekas, 1995, p. 268):

$$
\begin{equation*}
\mathbf{h}^{\star}(\mathbf{x})=\min _{a \in \mathcal{A}(\mathbf{x})}\left[g(\mathbf{x}, a) \tau(\mathbf{x}, a)-\lambda^{\star} \tau(\mathbf{x}, a)+\sum_{\mathbf{y} \in \mathcal{S}} p_{\mathbf{x y}}(a) \mathbf{h}^{\star}(\mathbf{y})\right] \forall \mathbf{x} \tag{2}
\end{equation*}
$$

In this system of equations, $\mathbf{h}^{\star}$ is the optimal differential vector, $\mathbf{x}$ and $\mathbf{y}$ are system states, and $\lambda^{\star}$ represents the optimal expected cost per time unit. Furthermore, $g(\mathbf{x}, a)$ represents the cost per time unit if control $a$ is applied in state $\mathbf{x}$. In our setting, this means that $g(\mathbf{x}, a)=b x_{N+1}$.

In the Bellman equation, the transition probabilities from state $\mathbf{x}$ to state $\mathbf{y}$ under control $a$ are denoted with $p_{\mathbf{x y}}(a)$ and the transition period lengths with $\tau(\mathbf{x}, a)$. Before we give the transition probabilities, we introduce the transition operators $T^{+}: \mathcal{S} \times \mathbb{N} \rightarrow \mathcal{S}$ and $T^{-}: \mathcal{S} \times \mathbb{N} \rightarrow \mathcal{S}$. Operator $T^{+}(\mathbf{x}, n)$ returns the state we end up with when we start in $\mathbf{x}$ and a part of SKU $n$ becomes ready for use. On the other hand, operator $T^{-}(\mathbf{x}, n)$ returns the state we end up with when we start in $\mathbf{x}$ and a part of SKU $n$ fails. Using the transition operators $T^{+}$and $T^{+}$, and defining $\mathbf{e}_{k}$ as the unit vector with a 1 on position $k$, we can write the transition probabilities $p_{\mathbf{x y}}(a)$, and the transition
period lengths $\tau(\mathbf{x}, a)$ as follows:

$$
\begin{align*}
& p_{\mathbf{x y}}\left(a_{n}\right)=\left\{\begin{array}{lll}
\frac{\mu}{\mu+\sum_{i=1}^{N} \lambda_{i}} & \text { if } n=N+1 & \text { and } \mathbf{y}=\mathbf{x}-\mathbf{e}_{\mathbf{N}+\mathbf{1}} \\
\frac{\mu}{\mu+\sum_{i=1}^{N} \lambda_{i}} & \text { if } 1 \leq n \leq N & \text { and } \mathbf{y}=T^{+}(\mathbf{x}, n) \\
\frac{\lambda_{q}}{\mu+\sum_{i=1}^{N} \lambda_{i}} & \text { if } 1 \leq n \leq N & \text { and } \mathbf{y}=T^{-}(\mathbf{x}, q) \\
\frac{\lambda_{q}}{\sum_{i=1}^{N} \lambda_{i}} & \text { if } n=0 & \text { and } \mathbf{y}=T^{-}(\mathbf{x}, q) \\
0 & \text { otherwise }
\end{array}\right.  \tag{3}\\
& \tau\left(\mathbf{x}, a_{n}\right)=\quad \begin{cases}\frac{1}{\sum_{i=1}^{N} \lambda_{i}} & \text { if } n=0 \\
\frac{1}{\mu+\sum_{i=1}^{N} \lambda_{i}} & \text { if } n \neq 0\end{cases}
\end{align*}
$$

Now all elements in (2) have been specified and we can try to solve the system of equations, yielding the optimal policy and the optimal expected annual total cost. An appropriate method for solving Bellman equations for average cost problems is relative value iteration. Before we can apply this procedure, we need to truncate the state space $\mathcal{S}$ in order to make it finite. We now describe how we have truncated the state space:
For every non-empty subset $J$ of SKU-s we calculate a threshold $L_{J}$ such that the probability that the total number of backorders exceeds $L_{J}$ and all SKU-s that do not belong to $J$ have positive net stock, does not exceed a small epsilon (say $10^{-12}$ ). For the part of the state space where $x_{i}=0$ for $i \in J$ and $x_{i}>0$ for $i \notin J$, we truncate the state space by not creating states where the aggregated number of backorders is bigger than the threshold $L_{J}$.

Let us illustrate this procedure on an example with two SKU-s. The system state then consists of three dimensions: $\mathbf{x}=\left(\max \left[0, z_{1}\right], \max \left[0, z_{2}\right], \max \left[0,-z_{1}\right]+\max \left[0,-z_{2}\right]\right)$. First, consider the situation where $x_{1}=0, x_{2}>0$, and $x_{3}>0$. Because we have argued that an optimal scheduling rule will prefer SKU-s with a negative net stock over SKU-s with a non-negative net stock, we know that the optimal scheduling rule will select parts of SKU 1 for repair. This means that the number of failed parts of SKU 1 is now equal to the queue length in an $M|M| 1$ system with demand rate $\lambda_{1}$ and service rate $\mu$. Standard queuing theory tells us that this queue length is geometrically distributed with parameter $\rho=\lambda / \mu$. The threshold $L_{1}$ can now be obtained from the formula for the cumulative geometric distribution. We can repeat this procedure for the situation where $x_{1}>0, x_{2}=0$, and $x_{3}>0$ to calculate $L_{2}$ and for the situation where $x_{1}=0, x_{2}=0$, and $x_{3}>0$ to calculate $L_{1,2}$. We only need to replace the demand rate $\lambda_{1}$ for the $M|M| 1$ analysis with $\lambda_{2}$, and $\left(\lambda_{1}+\lambda_{2}\right)$ respectively. We can now write the truncated state space for this example with 2 SKU-s
as the union of 4 sub state spaces:

$$
\begin{aligned}
\mathcal{S}= & \{0\} \times\left\{1, \ldots, S_{2}\right\} \times\left\{0,1, \ldots, L_{1}\right\} \cup \\
& \left\{1, \ldots, S_{1}\right\} \times\{0\} \times\left\{0,1, \ldots, L_{2}\right\} \cup \\
& \{0\} \times\{0\} \times\left\{0,1, \ldots, L_{1,2}\right\} \\
& \left\{1, \ldots, S_{1}\right\} \times\left\{1, \ldots, S_{2}\right\} \times\{0\}
\end{aligned}
$$

A graphical illustration of the (truncated) state space of the example discussed above is shown in Figure 2. This figure shows the transition rates for the underlying Markov process for base stock level vector $(1,2)$, and the scheduling policy that selects the SKU with most failed parts in the repair queue (in case of ties SKU-1 is selected).


Figure 2: Truncated state space
In this subsection we have described a method on how to obtain the optimal scheduling rule. In the remainder of this paper we denote the optimal policy with $\pi^{O P T}$. Since the state space grows exponentially in the number of SKU-s, this method only works for sufficiently small problem instances. In the remainder of this paper we therefor focus on heuristics and algorithms that can be applied to any problem instance.

### 3.2 EBT-rule

The Equalization of Backorder Times rule (EBT rule) tries to maximize the expected time until the next backorder. To realize this, the EBT rule selects the SKU with the shortest expected time until the next backorder (assuming that no new ready-for-use part are made available). Let us denote the EBT rule with $\pi^{B}$ and let $\mathbf{z}$ be the vector of net stock levels, then we can write:

$$
\begin{equation*}
\pi^{B}(\mathbf{z})=\arg \min \left\{\frac{\left(z_{n}+1\right)}{\lambda_{n}}\right\} \tag{4}
\end{equation*}
$$

Our EBT-rule is inspired on the Equalization of Runout Times rule (ERT-rule) proposed by Hax and Meal (1973) for MTS systems. What makes the EBT-rule (and the ERT-rule) attractive is that it is an intuitive and simple rule that honors the dynamic nature of the scheduling problem.

In Section 4, we will show that in general the EBT rule performs very well. However, a disadvantage of the EBT-rule is that it does not take into account the variability in the part failures process. We can illustrate this with following example that constitutes a motivation for the second heuristic rule:

Example 1 Take a problem with two SKU-s, SKU-1, and SKU-2, base stock level vector $\left(S_{1}, S_{2}\right)=$ $(1,99)$, failure rate vector $\left(\lambda_{1}, \lambda_{2}\right)=(1,100)$, and utilization rate $\rho$. It is easy to verify that in situations where SKU-1 has net stock 0, the EBT-rule will tell to repair a part of SKU-2 independent of the net stock of SKU-2. It is obvious that this is not a smart decision when the net stock of SKU-2 is equal to 98 . The simple reason is that there is an immediate danger of getting a backorder on SKU-1, whereas there are so many part of SKU-2 in stock that it should not hurt spending one repair time on SKU-1. Of course this argument does not only hold for a net stock level of 98 for SKU-2, but for a wide range of net stock levels. In particular if we let $\rho$ go to 0 , we can show that the optimality gap for the EBT-rule can get arbitrarily large.

### 3.3 M-rule

To capture the problem described in the previous example, we have developed the myopic scheduling rule (M-rule). In Figures Figures 3(a), 3(b), and 3(c) we have shown the optimal scheduling rule, the EBT-rule, and the M-rule for the example described above. The corresponding expected number of backorders are $0.000063,0.001155$, and 0.000063 respectively.

The idea behind the myopic scheduling rule (denoted with $\pi^{M}$ ) is that it selects an SKU such that the expected backorder cost in the near future are minimized. Suppose that we are at time $t=t_{0}$ and we need to make a decision on which part to put into repair. Ignoring preemption, it is clear that we cannot impact the backorder cost until the repair job (that we need to select now) is completed. All SKU-s have the same repair time distribution with mean equal to $1 / \mu$. In our myopic rescheduling rule we try to minimize the expected backorder cost rate (i.e. the expected


Figure 3: Scheduling policies for Example 1
number of backorders) at time $t=t_{0}+1 / \mu$. This is equivalent with choosing the SKU which has the largest probability of reaching a negative net stock level at time $t=t_{0}+1 / \mu$.

Let us define $P_{\xi}(Y>y)$ as the probability that $Y>y$ where $Y$ follows a Poisson distribution with mean $\xi$. Our myopic scheduling rule then reads as:

$$
\begin{equation*}
\pi^{M}(\mathbf{z})=\arg \max \left\{P_{\frac{\lambda_{n}}{\mu}}\left(Y>z_{n}\right)\right\} \tag{5}
\end{equation*}
$$

Our myopic scheduling rule is inspired on the well-known myopic scheduling rule as proposed by Perez and Zipkin (1997) for MTS systems.

### 3.4 Procedure for the joint calculation of static priorities and base stock levels

In this subsection we summarize the method presented in Adan et al. (2009) to simultaneously optimize base stock levels and static priorities. We use this method: (i) to calculate the base stock level vectors for the dynamic scheduling rules $\pi^{B}$ and $\pi^{M}$, and (ii) as a reference point when examining the performance of the dynamic scheduling rules. A very important building block in their work is an analytical procedure, based on steady state behavior for the two-queue preemptive priority model, to evaluate the expected number of backorders under given base stock levels, and given static repair priorities. The procedure is described in detail in Sleptchenko et al. (2005).

Adan et al. also show that under a given assignment of SKU-s to priority classes, the optimal base stock levels follow from a kind of newsvendor equation. They use the procedure for evaluating the expected number of backorders, and the newsvendor equation for calculating optimal base stock levels in a heuristic that determines the static priorities. This heuristic consists of full enumeration among all ordered priority assignments followed by local search (here, an ordered priority assignment is an assignment of SKU-s to priority classes such that for each pair of SKU-s $(n, m)$, it holds that $n$ is assigned to the same or a higher priority class than $m$ if and only if $h_{n} \geq h_{m}$ ). Numerical experiments showed that their method for joint optimization of static priorities and base stock levels is fast and accurate. On a test bed with 2 priority classes and up to 15 SKU-s the largest optimality gap is only $1.1 \%$.

In their paper, Adan et al. have evaluated four algorithms for optimizing the static priorities. We use Algorithm 4 (i.e. full enumeration of ordered assignments followed by local search) because the authors show in their paper that this algorithm outperforms the other three. Finally, we need to decide on the number of priority classes to use. Adan et al. have carried out experiments with $2,3,4$, and 5 priority classes. Obviously, the best result were obtained using 5 priority classes. Nevertheless, we use 3 priority classes because Adan et al. show that using 5 instead of 3 priority classes leads to an average cost reduction of only $0.6 \%$ whereas the computational effort increase with almost a factor 100 for problem instances with 50 SKU-s.

In the remainder of this paper, we denote the static scheduling rule and the base stock levels obtained from this joint optimization with $\pi^{S P}$, respectively $S^{S P}$.

### 3.5 Discrete Event Simulation

Except for sufficiently small problem instances, we are not able to evaluate the dynamic scheduling rules $\pi^{B}$ and $\pi^{M}$ analytically. In order to make a fair comparison between these scheduling rules and other scheduling rules we use discrete event simulation and the method of batch means. This allows us to obtain accurate estimates of the average number of backorders and the average annual cost. For every problem instance that we want to run with one of the two dynamic scheduling rules, we simulate at least 1,050,000 part failure events. Each simulation run is divided in 21 sub-runs of equal length. The first sub-run is used as a warming-up period and is not used for calculating key performance indicators. For all other 20 sub-runs we calculate the average number of backorders and the average annual total cost. From these 20 samples we calculate sample mean and standard error of the average annual total cost. If the standard error is smaller than $0.33 \%$ of the sample mean, simulation is terminated. Otherwise, we double the number of part failure events we want to simulate and merge the current 20 sub-runs into 10 new ones by merging sub-runs 1 with 2 , 3 with $4, \ldots$, and 19 with 20 .

Under some mild assumptions (batch means are independent and normally distributed), the
$99.7 \%$ confidence interval for the average annual total cost is equal to $[\widehat{x}-3 \widehat{s}, \widehat{x}+3 \widehat{s}]$, where $\widehat{x}$ stands for the sample mean of the average annual total cost and $\widehat{s}$ stands for the standard error of the average annual total cost. Since we simulate until we have reached a point where $\widehat{s}<\frac{1}{300} \widehat{x}$, the $99.7 \%$ confidence interval for the average annual total cost is [ $0.99 \widehat{x}, 1.01 \widehat{x}$ ]. In plain words, this means that with a probability of more than $99 \%$ the average annual total cost as obtained from our simulation does not differ more than $1 \%$ from its true value.

## 4 Numerical Experiments

In this section, we investigate the performance of our two proposed dynamic scheduling rules. In Section 4.1, we compare our dynamic scheduling rules $\pi^{B}$ and $\pi^{M}$ with the optimal scheduling rule $\pi^{O P T}$ for problem instances with a small number of SKU-s and given base stock levels. In Section 4.2, we define a test bed that contains a wide range of problem instances of real-life size. In Section 4.3, we use this test bed to compare our dynamic scheduling rules $\pi^{B}$ and $\pi^{M}$ with the static scheduling rule $\pi^{S P}$ for base stock levels that have been optimized for $\pi^{S P}$.

### 4.1 Optimality gap of the dynamic scheduling rules

We first want to gain insights in the optimality gap of the two dynamic scheduling rules. For this purpose we set up a numerical experiment where we compare the expected number of backorders for the optimum scheduling rule $\pi^{O P T}$ with the expected number of backorders for the dynamic scheduling rules $\pi^{B}$ and $\pi^{M}$. The test bed we use in this experiment consists of problem instances where the base stock level vector $S$ is a given input (i.e. it is not subject to optimization). Because we are interested in the stand-alone performance of the scheduling rules, we experiment with various base stock level vectors. These base stock level vectors are obtained from a calculation scheme that requires specification of a target fill rate $\beta_{n}$ for each SKU. Using formulas (2.6), (2.8), and (3.1) in Adan et al. (2009), and assuming repair jobs are processed in FCFS order, we can calculate the minimum base stock level $S_{n}$ for every SKU $n$ such that the target fill rate $\beta_{n}$ is reached. We create base stock level vectors for our problem instances by varying the target fill rates $\beta_{n}$ in the test bed.

For all problem instances in our test bed the number of SKU-s is equal to 2 and the mean repair time is equal to 1 . We use a factorial design for our test bed. The failure rates $\lambda_{1}$ for SKU-1 and $\lambda_{2}$ for SKU-2 are chosen such that the quotient $\lambda_{2} / \lambda_{1}$ is equal to 1 , or 4 , and the utilization rate $\rho=\left(\lambda_{1}+\lambda_{2}\right) / \mu$ is equal to 0.8 . 0.9 , or 0.95 . To obtain appropriate base stock levels we use the procedure described above. The target fill rates $\beta_{1}$ and $\beta_{2}$ are taken from the set $\{0.80,0.95\}$. This gives us problem instances where the base stock levels are low ( $\beta_{1}=\beta_{2}=0.80$ ), problem instances where the base stock levels are high ( $\beta_{1}=\beta_{2}=0.95$ ), but also problem instances where the base
stock levels are unbalanced $\left(\beta_{1} \neq \beta_{2}\right)$. Through this set-up, we want to find out to what extent the base stock level vector has an impact on the performance of the scheduling rule.

The results of the experiment are shown in Table 1. In this table we have used $\Delta^{B}\left(\Delta^{M}\right)$ to denote the relative difference between the expected number of backorders under scheduling rule $\pi^{O P T}$ and the expected number of backorders under scheduling rule $\pi^{B}\left(\pi^{M}\right)$.

| Input parameters |  |  |  |  | EBO |  |  | opt gap |  |
| :--- | :--- | :--- | :--- | :---: | :---: | :---: | :---: | :---: | :---: |
| $\rho$ | $\lambda_{2} / \lambda_{1}$ | $\left(\beta_{1}, \beta_{2}\right)$ | $\mathbf{S}$ | $\pi^{O P T}$ | $\pi^{B}$ | $\pi^{M}$ | $\Delta^{B}$ | $\Delta^{M}$ |  |
| 0.80 | 1 | $(0.80,0.80)$ | $(4,4)$ | 0.7023 | 0.7023 | 0.7023 | $0.0 \%$ | $0.0 \%$ |  |
| 0.80 | 1 | $(0.80,0.95)$ | $(4,8)$ | 0.2948 | 0.2948 | 0.2948 | $0.0 \%$ | $0.0 \%$ |  |
| 0.80 | 1 | $(0.95,0.80)$ | $(8,4)$ | 0.2948 | 0.2948 | 0.2948 | $0.0 \%$ | $0.0 \%$ |  |
| 0.80 | 1 | $(0.95,0.95)$ | $(8,8)$ | 0.1185 | 0.1185 | 0.1185 | $0.0 \%$ | $0.0 \%$ |  |
| 0.80 | 4 | $(0.80,0.80)$ | $(2,6)$ | 0.7014 | 0.7021 | 0.7116 | $0.1 \%$ | $1.5 \%$ |  |
| 0.80 | 4 | $(0.80,0.95)$ | $(2,12)$ | 0.1871 | 0.1878 | 0.1898 | $0.3 \%$ | $1.5 \%$ |  |
| 0.80 | 4 | $(0.95,0.80)$ | $(4,6)$ | 0.4675 | 0.4679 | 0.4779 | $0.1 \%$ | $2.2 \%$ |  |
| 0.80 | 4 | $(0.95,0.95)$ | $(4,12)$ | 0.1187 | 0.1193 | 0.1246 | $0.5 \%$ | $5.0 \%$ |  |
| 0.90 | 1 | $(0.80,0.80)$ | $(9,9)$ | 1.3639 | 1.3639 | 1.3639 | $0.0 \%$ | $0.0 \%$ |  |
| 0.90 | 1 | $(0.80,0.95)$ | $(9,15)$ | 0.7250 | 0.7250 | 0.7250 | $0.0 \%$ | $0.0 \%$ |  |
| 0.90 | 1 | $(0.95,0.80)$ | $(15,9)$ | 0.7250 | 0.7250 | 0.7250 | $0.0 \%$ | $0.0 \%$ |  |
| 0.90 | 1 | $(0.95,0.95)$ | $(15,15)$ | 0.3852 | 0.3852 | 0.3852 | $0.0 \%$ | $0.0 \%$ |  |
| 0.90 | 4 | $(0.80,0.80)$ | $(4,13)$ | 1.5150 | 1.5155 | 1.5280 | $0.0 \%$ | $0.9 \%$ |  |
| 0.90 | 4 | $(0.80,0.95)$ | $(4,24)$ | 0.4756 | 0.4758 | 0.4796 | $0.0 \%$ | $0.9 \%$ |  |
| 0.90 | 4 | $(0.95,0.80)$ | $(7,13)$ | 1.1065 | 1.1068 | 1.1213 | $0.0 \%$ | $1.3 \%$ |  |
| 0.90 | 4 | $(0.95,0.95)$ | $(7,24)$ | 0.3466 | 0.3468 | 0.3519 | $0.0 \%$ | $1.5 \%$ |  |
| 0.95 | 1 | $(0.80,0.80)$ | $(4.6)$ | 3.3287 | 3.3287 | 3.3287 | $0.0 \%$ | $0.0 \%$ |  |
| 0.95 | 1 | $(0.80,0.95)$ | $(4,12)$ | 1.7088 | 1.7088 | 1.7088 | $0.0 \%$ | $0.0 \%$ |  |
| 0.95 | 1 | $(0.95,0.80)$ | $(7,17)$ | 1.7088 | 1.7088 | 1.7088 | $0.0 \%$ | $0.0 \%$ |  |
| 0.95 | 1 | $(0.95,0.95)$ | $(17,30)$ | 0.8772 | 0.8772 | 0.8772 | $0.0 \%$ | $0.0 \%$ |  |
| 0.95 | 4 | $(0.80,0.80)$ | $(7,26)$ | 3.5035 | 3.5036 | 3.5130 | $0.0 \%$ | $0.2 \%$ |  |
| 0.95 | 4 | $(0.80,0.95)$ | $(7,48)$ | 1.1335 | 1.1335 | 1.1366 | $0.0 \%$ | $0.2 \%$ |  |
| 0.95 | 4 | $(0.95,0.80)$ | $(13,26)$ | 2.5754 | 2.5755 | 2.5837 | $0.0 \%$ | $0.3 \%$ |  |
| 0.95 | 4 | $(0.95,0.95)$ | $(13,48)$ | 0.8332 | 0.8333 | 0.8359 | $0.0 \%$ | $0.3 \%$ |  |

Table 1: Optimality gap of $\pi^{B}$ and $\pi^{M}$

Table 1 shows that the ERT-rule has excellent performance over all 24 problem instances; the average optimality gap is only $0.04 \%$ and the maximum optimality gap is just $0.5 \%$. The M-rule
also performs well with an average optimality gap of $0.5 \%$ and a maximum optimality gap of $5.0 \%$. It is a little surprise that the EBT-rule performs better than the M-rule on this test bed, as we showed that the optimality gap of the EBT-rule can be made arbitrarily large. This might be an indication that the average-case behavior of the EBT-rule is good, although its worst-case behavior is bad. We investigate this in the next subsection. In the table we can also see that for both scheduling rules the optimality gap gets larger when the utilization rate gets smaller, or when the failure rates are different.

### 4.2 Test Bed

We use a factorial design for our test bed with a similar setup as in Adan et al. (2009). The number of SKU-s is chosen equal to 15,25 , and 50 . The lowest holding cost ( $h^{\text {min }}$ ) is taken equal to 1,10 , and 100 whereas we fix the highest holding cost $\left(h^{\max }\right)$ at 1000 . Hence, the ratio between the highest and the lowest holding cost varies from 10 to 1000 , which are common ratios for spare parts. Regarding the dependence between holding cost parameters and failure rates, we consider a hyperbolic relation because this usually gives a close match to what is seen in practice. We choose the failure rates $\lambda_{n}$ for all SKU-s as independent samples from a uniform distribution on [1, 100]. The holding cost parameters $h_{n}$ are obtained by the following function:

$$
\begin{equation*}
h_{n}=\max \left[h^{\min }, a \frac{110}{10+\lambda_{n}}+b+\xi_{n}\right] \tag{6}
\end{equation*}
$$

In this function, $a=1 / 9\left(h^{\max }-h^{\min }\right)$ and $b=10 / 9 h^{\min }-1 / 9 h^{\max }$, and $\xi_{n} \in U[-v, v]$ with $v=1 / 40\left(h^{\max }-h^{\min }\right)$. Further, the mean repair time $1 / \mu$ is chosen such that the utilization rate $(\rho)$ of the repair facility is equal to $0.70,0.82,0.90$, and 0.95 , respectively. For the penalty cost parameter $b$, we choose the values 1000,10000 , and 100000 .

In Table 2, we have summarized our choices for all parameters. The total number of all possible combinations for these parameters is $3 \times 3 \times 1 \times 1 \times 4 \times 3=108$. We have generated 5 sets of values for $\lambda_{n}$ and $h_{n}$, as there are uniform distributions involved in the generation of these values. This gives us in total $108 \times 5=540$ instances.

### 4.3 Comparison of the scheduling rules for large instances

In this subsection we describe a numerical experiment to examine the performance our proposed solution for large problem instances. In our proposed solution we use the base stock level vector $S^{S P}$ that follows from the method for joint optimization of base stock levels and static priorities. For the real-time scheduling of repair jobs, we use dynamic scheduling rule $\pi^{B}$ or $\pi^{M}$. We compare these two solution methods with the traditional static priorities solution method as described in proposed by Adan et al. (2009) and summarized in Section 3.4.

| Name of parameter | No. of <br> values | Values |
| :--- | :---: | :---: |
| Number of SKU-s | 3 | $\{15,25,50\}$ |
| Lowest holding cost $\left(h^{\text {min }}\right)$ | 3 | $\{1,10,100\}$ |
| Highest holding cost $\left(h^{\text {max }}\right)$ | 1 | $\{1000\}$ |
| Failure rate $\left(\lambda_{n}\right)$ | 1 | $\lambda_{n} \in U[1,100]$ |
| Holding cost $\left(h_{n}\right)$ | $h_{n}=h\left(\lambda_{n}\right)($ see equation (6)) |  |
| Utilization rate $(\rho)$ | 4 | $\{0.70,0.82,0.90,0.95\}$ |
| Penalty cost $(b)$ | 3 | $\{1000,10000,100000\}$ |

Table 2: Parameter choices for the test bed

Since all three solution methods use the same base stock level vector (denoted with $S^{S P}$ ), we can unambiguously identify a solution method by its scheduling rule. This means that in the remainder of this subsection we talk about solution methods $\pi^{S P}, \pi^{B}$, and $\pi^{M}$. We compare these three solution methods on all 540 problem instances in the test bed that has been described in the previous subsection. In contrast to Section 4.1, we are here not so much interested in the isolated performance of the scheduling rule, but more in the performance of the base stock level vector and the scheduling rule together. As a consequence we mainly focus on the average total cost and less on the number of expected backorders.

However, to get some more insights in the performance of the two dynamic scheduling rules, we first examine the reductions in the average number of backorders for all 540 problem instances. It is interesting to see if the reductions in the number of expected backorders we can achieve using a dynamic scheduling rule, are correlated with the fraction of the backorder cost in the total cost. Let us define $w_{b}$ as the fraction of the backorder cost in the total cost in the solution method $\pi^{S P}$. In Figures 4(a) and 4(c), we have shown how the reduction in the average number of backorders depends on $w_{b}$. The two figures clearly show that there is a very strong (negative) correlation. We see that the reductions in the number of expected backorders ranges from less than $5 \%$ (typically for $w_{b} \geq 0.60$ ) to more than $50 \%$ (typically for $w_{b} \leq 0.20$ ). Interestingly enough, the strong correlation almost disappears if we look at the reductions in annual total cost; see Figures 4(b) and 4(d). Figures 4(a)-4(d) also show that for realistic problem instances as considered in this section, the difference between the two scheduling rules is pretty small.

In order to get a better feeling on how the reductions in the average number of backorders and the total annual cost are distributed, we have shown in Figure 4.3 the cumulative improvement distribution for the scheduling rule that is based on equalizing backorder times a $\pi^{B}$ and the myopic scheduling rule $\pi^{M}$. The figures show for example that in approximately $80 \%$ of all cases, we achieve a total cost reduction of $6 \%$ for the EBT-rule $\left(\pi^{B}\right)$, and $7 \%$ for the M-rule $\left(\pi^{M}\right)$. Here we see that


Figure 4: Performance improvements as function of $w_{b}$
the two dynamic scheduling policies show similar performance. We only see that the EBT-rule has almost twice as many cases where the total cost could not be reduced with more than $5 \%$ than the M-rule. Interestingly enough, there is even one problem instance where $\pi^{B}$ does worse than $\pi^{S P}$; see Figure 4(b).

We conclude this section with an investigation on how the reductions in total cost depend on the four factors in our test bed: the number of SKU-s $(N)$, the utilization rate ( $\rho$ ), the minimum inventory holding cost $\left(h^{\min }\right)$, and the penalty cost $(b)$. The results can be found in Tables 6(a)$6(\mathrm{~d})$. These figures show that the reductions in total cost clearly increase as the utilization rates $\rho$ gets bigger. This holds for $\pi^{B}$ and $\pi^{B}$. The other three factors only show a weak correlation. The tables also show that for each subset of the 540 problem instances $\pi^{M}$ does (slightly) better than $\pi^{B}$.


Figure 5: Cumulative improvement distribution

|  |  |  | 15 | 25 | 50 |  |  |  |  |  | 0.70 | 0.82 | 0.90 | 0.95 |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| $\pi^{E B T}$ | vs | $\pi^{S P}$ | $9.0 \%$ | $9.0 \%$ | $8.8 \%$ |  |  | $\pi^{E B T}$ | vs | $\pi^{S P}$ | $5.8 \%$ | $8.2 \%$ | $10.0 \%$ | $11.8 \%$ |
| $\pi^{M}$ | vs | $\pi^{S P}$ | $9.7 \%$ | $9.8 \%$ | $9.4 \%$ |  | $\pi^{M}$ | vs | $\pi^{S P}$ | $6.0 \%$ | $9.0 \%$ | $10.6 \%$ | $12.3 \%$ |  |

(a) Rel. diff. as function of $N$
(b) Rel. diff. as function of $\rho$

|  |  |  | 1 | 10 | 100 |
| :--- | :---: | :---: | :---: | :---: | :---: |
| $\pi^{E B T}$ | vs | $\pi^{S P}$ | $8.7 \%$ | $8.6 \%$ | $9.5 \%$ |
| $\pi^{E B T}$ | vs | $\pi^{S P}$ | $9.4 \%$ | $9.3 \%$ | $10.1 \%$ |

(c) Rel. diff. as function of $h^{m i n}$

|  |  |  | 1000 | 10000 | 100000 |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\pi^{E B T}$ | vs | $\pi^{S P}$ | $8.3 \%$ | $10.5 \%$ | $8.0 \%$ |
| $\pi^{E B T}$ | vs | $\pi^{S P}$ | $8.3 \%$ | $10.7 \%$ | $9.9 \%$ |

(d) Rel. diff. as function of $b$

Figure 6: Relative difference between annual total cost obtained by different scheduling rules

## 5 Conclusions

We studied the use of dynamic scheduling rules in a repairable spare parts system consisting of one repair shop and one stockpoint. We made simplifying assumptions for the repair shop and repair times. We proposed two dynamic scheduling rules; one that tries to equalize the time until the next backorder (EBT-rule) and a myopic one that tries to minimize the expected backorder cost over a small future time horizon (M-rule).

We showed that the average optimality gap is small $(<1 \%)$ for both scheduling rules on a set of 'reasonable' problem instances. But we also showed that the optimality gap for the EBT-rule can be made arbitrarily large. We proposed a solution method for optimizing base stock levels and real-time scheduling decisions that consists of two steps. In the first step, we use the procedure
for joint calculation of base stock levels and static priorities as proposed in Adan et al. (2009) to obtain the base stock levels. In the second step, we replace the static priority rule by the EBT-rule or the M-rule.

We have compared our solution approach with the traditional static priority solution method on a test bed of 540 realistic problem instances. We showed that the average annual total cost over all 540 problem instances can be reduced with 8.9 \% when applying the EBT-rule and with $9.6 \%$ when applying the M-rule. These cost savings are only caused by reductions in the average backorders as the inventory holding costs remain unchanged. We showed that the cost reductions do not depend on the fraction of backorder cost in the total cost, but that they are positively correlated with the utilization rate.

For future research, we suggest investigate the opportunities to optimize the base stock levels for dynamic scheduling rules. Another direction for further research is to examine the robustness of base stock optimization procedures and scheduling rules against forecast errors.

## References

- Adan, I.J.B.F., Sleptchenko, A., Van Houtum G.J., 2009. Reducing costs of spare parts supply systems via static priorities. Asia-Pacific Journal of Operational Research, 26, 559-585.
- Albright, S.C. Gupta, A., 1992. Steady-state approximation of a multiechelon multi-indentured repairable-item inventory system with a single repair facility. Naval Research Logistics, 40, 479-493.
- Bertsekas, D.P., 1995. Dynamic Programming and Optimal Control, Vol. I, Athena Scientific, Belmont, MA.
- Caggiano, K.E., Muckstadt J.A., Rappold J.A., 2006. Integrated real-time capacity and inventory allocation for reparable service parts in a two-echelon supply system. Manufacturing and Service Operations Management, 8, 292-319.
- Gross, D., Miller, D.R., Soland, R.M., 1983. A closed queueing network model for multi-echelon repairable item provisioning. IIE Transactions, 15, 344-352.
- Hausman, W.H., Scudder, G.D., 1982. Priority scheduling rules for repairable inventory systems. Management Science, 11, 1215-1232.
- Hax, A. C., Meal, H.C., 1975. Hierarchical Integration of Production Planning and Scheduling. In: Studies in Management Sciences, Vol. I, Logistics, M. A. Geisler (ed.), North-Holland-American Elsevier, New York.
- Johnson, M. E., Scudder, G., 1999. Supporting Quick Response Through Scheduling of Make-to-Stock Production/Inventory Systems. Decision Sciences, 30, 441-467.
- Kat, B., Avsar Z.M., 2009. Using aggregate fill-rate for dynamic scheduling of multi-class systems. Annals of Operations Research, DOI:10.1007/s10479-009-0535-2.
- Muckstadt, J.A., 1973. A Model for a Multi-Item, Multi-Echelon, Multi-Identure Inventory System. Management Science, 20, 472-482.
- Perez, A.P., Zipkin, P., 1997. Dynamic scheduling rules for a multiproduct make-to-stock queue. Operations Research, 45, 919-930.
- Pyke, D.F., 1990. Priority repair and dispatch policies for repairable-item logistics systems. Naval Research Logistics, 37, 1-30.
- Rustenburg, W.D., Van Houtum, G.J., Zijm, W.H.M., 2001. Spare parts management at complex technology-based organizations: An agenda for research. International Journal of Production Economics, 71, 177-193.
- Rustenburg, W.D., van Houtum, G.J., Zijm, W.H.M., 2003. Exact and approximate analysis of multiechelon, multi-indenture spare parts systems with commonality. In: Shanthikumar, J.G., Yao, D.D., Zijm, W.H.M. (Eds.), Stochastic Modeling and Optimization of Manufacturing Systems and Supply Chains. Boston: Kluwer, 143-176.
- Sherbrooke, C.C., 1968. METRIC: Multi-echelon technique for recoverable item control. Operations Research, 16, 122-141.
- Sherbrooke, C.C., 2004. Optimal Inventory Modeling of Systems: Multi-Echelon Techniques. Dordrecht, The Netherlands: Kluwer.
- Slay, F.M., 1984, VARY-METRIC: An Approach to Modeling Multi-Echelon Resupply when the Demand Process is Poisson with Gamma Prior, Report AF301-3, Logistic Management Institute, Washington, D.C.
- Sleptchenko, A., Van der Heijden, M.C., Van Harten, A., 2002. Effects of finite repair capacity in multi-echelon, multi-indenture service parts supply systems, International Journal of Production Economics, 79, 209-230.
- Thonemann, U.W., Brown, A.O., Hausman, W.H., 2002. Easy quantification of improved spare parts inventory policies. Management Science, 48, 1213-1225.
- Van Harten, A., Sleptchenko, A., 2003. On Markovian multi-class, multi-server queueing. Queueing Systems, 43, 307-328.
- Winands, E.M.M., Adan, I.J.B.F., Van Houtum, G.J., 2011. The stochastic economic lot scheduling problem : a survey. European Journal of Operational Research, ..., submitted / in press.
- Zijm, W.H.M., Avsar, Z.M., 2003. Capacitated two-indenture models for repairable item systems. International Journal of Production Economics, 8182, 573-588.

| nr. Year Title | Author(s) |
| :---: | :---: |
| 3292010 Reducing costs of repairable spare parts supply systems via dynamic scheduling | H.G.H. Tiemessen, G.J. van Houtum |
| 3282010 <br> Identification of Employment Concentration and Specialization Areas: Theory and Application | F.P. van den Heuvel, P.W. de Langen, K.H. van Donselaar, J.C. Fransoo |
| 3272010 A combinatorial approach to multi-skill workforce scheduling | Murat Firat, Cor Hurkens |
| 3262010 Stability in multi-skill workforce scheduling | Murat Firat, Cor Hurkens, Alexandre Laugier |
| 3252010 |  |
| Maintenance spare parts planning and control: A framework for control and agenda for future research | M.A. Driessen, J.J. Arts, G.J. v. Houtum, W.D. Rustenburg, B. Huisman |
| 3242010 |  |
| Near-optimal heuristics to set base stock levels in a two-echelon distribution network | R.J.I. Basten, G.J. van Houtum |
| $3232010$ <br> Inventory reduction in spare part networks by selective throughput time reduction | M.C. van der Heijden, E.M. Alvarez, J.M.J. Schutten |
| 3222010 $\qquad$ service-contract differentiation | E.M. Alvarez, M.C. van der Heijden, W.H. Zijm |
| 3212010 Heuristics for Multi-Item Two-Echelon Spare Parts Inventory Control Problem with Batch Ordering in the Central Warehouse | B. Walrave, K. v. Oorschot, A.G.L. Romme |
| 3202010 Preventing or escaping the suppression mechanism: intervention conditions | Nico Dellaert, Jully Jeunet. |
| 3192010 <br> Hospital admission planning to optimize major resources utilization under uncertainty | R. Seguel, R. Eshuis, P. Grefen. |
| 3182010 Minimal Protocol Adaptors for Interacting Services <br> 3172010 Teaching Retail Operations in Business and Engineering Schools | Tom Van Woensel, Marshall L. Fisher, Jan C. Fransoo. <br> Lydie P.M. Smets, Geert-Jan van Houtum, Fred Langerak. |


| 3162010 Design for Availability: Creating Value for Manufacturers and Customers | Pieter van Gorp, Rik Eshuis. |
| :---: | :---: |
| $3152010 \frac{\text { Transforming Process Models: executable }}{\text { rewrite rules versus a formalized Java program }}$ | Bob Walrave, Kim E. van Oorschot, A. Georges L. Romme |
| 3142010 <br> Ambidexterity and getting trapped in the suppression of exploration: a simulation model |  |
| 3132010 A Dynamic Programming Approach to MultiObjective Time-Dependent Capacitated Single Vehicle Routing Problems with Time Windows | S. Dabia, T. van Woensel, A.G. de Kok |
| Tales of a So(u)rcerer: Optimal Sourcing <br> 3122010 Decisions Under Alternative Capacitated Suppliers and General Cost Structures | Osman Alp, Tarkan Tan |
| In-store replenishment procedures for perishable <br> 3112010 inventory in a retail environment with handling costs and storage constraints | R.A.C.M. Broekmeulen, C.H.M. Bakx |
| 3102010 The state of the art of innovation-driven business models in the financial services industry | E. Lüftenegger, S. Angelov, E. van der Linden, P. Grefen |
| 3092010 Design of Complex Architectures Using a Three | R. Seguel, P. Grefen, R. Eshuis |
| $3082010 \begin{aligned} & \text { Effect of carbon emission regulations on } \\ & \text { transport mode selection in supply chains }\end{aligned}$ | K.M.R. Hoen, T. Tan, J.C. Fransoo, G.J. van Houtum |
| $3072010 \frac{\text { Interaction between intelligent agent strategies }}{\text { for real-time transportation planning }}$ | Martijn Mes, Matthieu van der Heijden, Peter Schuur |
| 3062010 Internal Slackening Scoring Methods | Marco Slikker, Peter Borm, René van den Brink |
| 3052010 <br> Vehicle Routing with Traffic Congestion and Drivers' Driving and Working Rules | A.L. Kok, E.W. Hans, J.M.J. Schutten, W.H.M. Zijm |
| $3042010 \frac{\text { Practical extensions to the level of repair }}{\text { analysis }}$ | R.J.I. Basten, M.C. van der Heijden, J.M.J. Schutten |
| Ocean Container Transport: An Underestimated 3032010 and Critical Link in Global Supply Chain Performance | Jan C. Fransoo, Chung-Yee Lee |
| Capacity reservation and utilization for a 3022010 manufacturer with uncertain capacity and demand | Y. Boulaksil; J.C. Fransoo; T. Tan |
| 3002009 Spare parts inventory pooling games | F.J.P. Karsten; M. Slikker; G.J. van Houtum |
| $2992009 \frac{\text { Capacity flexibility allocation in an outsourced }}{\text { supply chain with reservation }}$ | Y. Boulaksil, M. Grunow, J.C. Fransoo |
| 2982010 An optimal approach for the joint problem of leve of repair analysis and spare parts stocking | R.J.I. Basten, M.C. van der Heijden, J.M.J. Schutten |
| Responding to the Lehman Wave: Sales <br> 2972009 Forecasting and Supply Management during the Credit Crisis | Robert Peels, Maximiliano Udenio, Jan C. Fransoo, Marcel Wolfs, Tom Hendrikx |
| 2962009 An exact approach for relating recovering | Peter T. Vanberkel, Richard J. Boucherie, |


|  | surgical patient workload to the master surgical schedule | Erwin W. Hans, Johann L. Hurink, Wineke A.M. van Lent, Wim H. van Harten |
| :---: | :---: | :---: |
| $2952009$ | An iterative method for the simultaneous optimization of repair decisions and spare parts stocks | R.J.I. Basten, M.C. van der Heijden, J.M.J. Schutten |
| 2942009 | Fujaba hits the Wall(-e) | Pieter van Gorp, Ruben Jubeh, Bernhard Grusie, Anne Keller |
| 2932009 | Implementation of a Healthcare Process in Four Different Workflow Systems | R.S. Mans, W.M.P. van der Aalst, N.C. Russell, P.J.M. Bakker |
| 2922009 | Business Process Model Repositories Framework and Survey | Zhiqiang Yan, Remco Dijkman, Paul Grefen |
| 2912009 | Efficient Optimization of the Dual-Index Policy Using Markov Chains | Joachim Arts, Marcel van Vuuren, Gudrun Kiesmuller |
| 2902009 | Hierarchical Knowledge-Gradient for Sequential Sampling | Martijn R.K. Mes; Warren B. Powell; Peter I. Frazier |
| 2892009 | Analyzing combined vehicle routing and break scheduling from a distributed decision making perspective | C.M. Meyer; A.L. Kok; H. Kopfer; J.M.J. Schutten |
| 2882009 | Anticipation of lead time performance in Supply Chain Operations Planning | Michiel Jansen; Ton G. de Kok; Jan C. Fransoo |
| 2872009 | Inventory Models with Lateral Transshipments: A Review | Colin Paterson; Gudrun Kiesmuller; Ruud Teunter; Kevin Glazebrook |
| 2862009 | Efficiency evaluation for pooling resources in health care | P.T. Vanberkel; R.J. Boucherie; E.W. Hans; J.L. Hurink; N. Litvak |
| 2852009 | A Survey of Health Care Models that Encompass Multiple Departments | P.T. Vanberkel; R.J. Boucherie; E.W. Hans; J.L. Hurink; N. Litvak |
| 2842009 | Supporting Process Control in Business Collaborations | S. Angelov; K. Vidyasankar; J. Vonk; P. Grefen |
| 2832009 | Inventory Control with Partial Batch Ordering | O. Alp; W.T. Huh; T. Tan |
| 2822009 | Translating Safe Petri Nets to Statecharts in a Structure-Preserving Way | R. Eshuis |
| 2812009 | The link between product data model and process model | J.J.C.L. Vogelaar; H.A. Reijers |
| 2802009 | Inventory planning for spare parts networks with delivery time requirements | I.C. Reijnen; T. Tan; G.J. van Houtum |
| 2792009 | Co-Evolution of Demand and Supply under Competition | B. Vermeulen; A.G. de Kok |
| $2782010$ | Toward Meso-level Product-Market Network Indices for Strategic Product Selection and (Re)Design Guidelines over the Product LifeCycle | B. Vermeulen, A.G. de Kok |
| 2772009 | An Efficient Method to Construct Minimal Protocol Adaptors | R. Seguel, R. Eshuis, P. Grefen |


| 2762009 Coordinating Supply Chains: a Bilevel Programming Approach | Ton G. de Kok, Gabriella Muratore |
| :---: | :---: |
| $2752009 \frac{\text { Inventory redistribution for fashion products }}{\text { under demand parameter update }}$ | G.P. Kiesmuller, S. Minner |
| Comparing Markov chains: Combining 2742009 aggregation and precedence relations applied to sets of states | A. Busic, I.M.H. Vliegen, A. Scheller-Wolf |
| 2732009 Separate tools or tool kits: an exploratory study | I.M.H. Vliegen, P.A.M. Kleingeld, G.J. van Houtum |
| An Exact Solution Procedure for Multi-Item Two2722009 Echelon Spare Parts Inventory Control Problem with Batch Ordering | Engin Topan, Z. Pelin Bayindir, Tarkan Tan |
| $2712009 \text { Distributed Decision Making in Combined }$ | C.M. Meyer, H. Kopfer, A.L. Kok, M. Schutten |
| Dynamic Programming Algorithm for the Vehicle 2702009 Routing Problem with Time Windows and EC Social Legislation | A.L. Kok, C.M. Meyer, H. Kopfer, J.M.J. Schutten |
| 2692009 Similarity of Business Process Models: Metics and Evaluation | Remco Dijkman, Marlon Dumas, Boudewijn van Dongen, Reina Kaarik, Jan Mendling |
| $2672009 \frac{\text { Vehicle routing under time-dependent travel }}{\text { times: the impact of congestion avoidance }}$ | A.L. Kok, E.W. Hans, J.M.J. Schutten |
| 2662009 Restricted dynamic programming: a flexible framework for solving realistic VRPs | J. Gromicho; J.J. van Hoorn; A.L. Kok; J.M.J. Schutten; |

Working Papers published before 2009 see: http://beta.ieis.tue.nl


[^0]:    *Corresponding author, E-mail: hti@zurich.ibm.com, tel: + 41447248469 , fax: +41447248964

