Value-based $Q(s, S)$ policy for joint replenishments

Søren Glud Johansen\textsuperscript{a}, Anders Thorstenson\textsuperscript{b,}\textasteriskcentered

September 6, 2006

\textsuperscript{a} Department of Operations Research, University of Aarhus, Ny Munkegade, Bldg. 530, DK-8000 Aarhus C, Denmark, sgj@imf.au.dk

\textsuperscript{b} Logistics/SCM Research Group, Aarhus School of Business, Fuglesangs Allé 4, DK-8210 Aarhus V, Denmark, ath@asb.dk

Abstract

We improve the recently proposed $Q(s, S)$ policy for the stochastic joint replenishment problem. In our value-based $Q(s, S)$ policy, item inventories are reviewed continuously when the total demand since the last replenishment is close to $Q$. At each such review instant a new replenishment order is issued, if the expected cost of ordering immediately according to the $(s, S)$ policy is less than the expected cost of deferring the order until the next demand or until the level $Q$ is reached.

We use simulation to evaluate our policy. Applying the value-based $Q(s, S)$ policy to a standard set of 12-item numerical examples from the literature, the long-run average cost of the best known solution is reduced by approximately 1%. Further examples are also investigated and in some cases for which the cost structure implies a high service level, the cost reduction exceeds 10% of the cost of the pure $Q(s, S)$ policy.

\textit{Keywords} : Inventory; Joint replenishment; Markov chain; Optimization; Stochastic demand.

1 Introduction

This paper studies the joint replenishment problem encountered when opportunities exist for coordination of the replenishments of different items. The system considered consists of $n$ items facing Poisson demands. The demand rate for item $i$ is $\lambda_i$. A
setup cost $K$ is incurred for every joint order, and in addition an item-specific setup
cost $k_i$ is incurred for every order of item $i$. Each replenishment of item $i$ requires
a constant lead time $L_i$. Costs are charged at rates $h_i$ and $b_i$ per unit of item $i$ on
stock and backorder, respectively. A fixed shortage cost $\pi_i$ is incurred every time a
unit of item $i$ is backlogged.

The stochastic joint replenishment problem has been studied extensively in the
literature. The can-order policy was suggested long ago (Balintfy, 1964). It has
been shown not to be the optimal policy for this problem (Ignall, 1969) and it has
been demonstrated that it may be hard to find an exact, optimal policy within
this class of policies (Schultz and Johansen, 1999). However, methods exist to
compute good can-order policies, both in a continuous-review setting (Federgruen et
al., 1984; Melchiors, 2002) and in a periodic-review setting (Johansen and Melchiors,
2003). The $QS$ policy, specified below, was also suggested fairly early (Renberg and
Planche, 1967) and has been further investigated (Pantumsinchai, 1992). Moreover,
periodic-review heuristics using coordinated base-stock policies have been studied
(Atkins and Iyogun, 1988). More recently, the periodic-review $P(s, S)$ policy based
on optimal $(s, S)$ policies for a common, optimized review period has been proposed
(Viswanathan, 1997).

The demand-reporting $Q(s, S)$ policy was also proposed by Viswanathan (1997).
It works as follows: The total demand of all items is reviewed continuously, while
item inventories are only reviewed when the total demand since the last replenish-
ment has reached the level $Q$. At the latter review instants, all items $i$ with inventory
position less than or equal to the reorder point $s_i$ are ordered up to the level $S_i$.
The $QS$ policy is a special case of the $Q(s, S)$ policy in which $s_i = S_i - 1$ for all
items $i$. An optimal $Q(s, S)$ policy can be computed by the algorithm suggested by
Nielsen and Larsen (2005). For most of the standard examples from the literature,
the $Q(s, S)$ policy as computed from this algorithm seems to outperform the alterna-
tive policies. For further applications of the $Q(s, S)$ policy, see also Brønmo (2005;
Brønmo’s maiden name was Nielsen.). In this paper we suggest an improvement to
the $Q(s, S)$ policy by evaluating the economic consequences of deviating from it.

Using the improved policy, which we denote the value-based $Q(s, S)$ policy, item
inventories are reviewed continuously when the total demand since the last replen-
ishment is close to $Q$. At each such review instant a new replenishment order is
issued, if the expected cost of ordering immediately according to the $(s, S)$ policy
is less than the expected cost of deferring the order until the next demand or until
the level $Q$ is reached. It is straightforward to determine the expected cost incurred
until the next replenishment order is planned to be issued. To this expected cost
we add the relative value of the system’s state when the replenishment order is issued. The relative value is determined from a Markov chain and is computed by a value-iteration algorithm. In the Markov chain the state of each item is not specified as the inventory position (or some derived measure thereof), but as the number of replenishment cycles since the item was last ordered. This might limit the size of the state space considerably. The size is also limited if some items are included in every replenishment order.

A motivation for exploring the value-based $Q(s, S)$ policy stems from the aggregate nature of the parameter $Q$ used in the pure $Q(s, S)$ policy. It implies simply adding units of demand for items which may have quite different cost and demand structures. Even if demand is expressed in monetary units this will not in general reflect the full economic consequences related to issuing a replenishment order. By contrast, our approach is similar in spirit to the heuristic approach suggested in Tijms (2003, Subsection 7.5), whereby one first determines a good estimate for the relative values of the system’s state at a decision instant and then applies a single policy-improvement step. This approach is based on the observation that the first few improvement steps in a policy-iteration procedure in many cases capture most of the potential improvement benefits. Accordingly, our approach is also closely related to the one used in Axsäter (2006) to make consecutive decisions on whether to utilize an option for emergency replenishments. A further source of inspiration is Adelman (2004), who uses dual prices from a large-scale linear program to approximate the optimal value function of a stochastic, discounted cost problem in which joint inventory replenishments are embedded. From a practical point of view, our value-based policy offers a natural mechanism for adjusting the pure $Q(s, S)$ policy to the current state of the inventory system.

In terms of the long-run, average cost performance the value-based $Q(s, S)$ policy dominates the pure $Q(s, S)$ policy. The price for this dominance is in terms of an increased computational burden. In cases where the size of the state space is overwhelmingly large, the value-based policy is not applicable, because relative state values cannot be computed. The numerical results obtained show improvements for all examples in which the value-based policy was computed. The relative magnitude of the improvements ranges from approx. 1% in a standard 12-item example from Atkins and Iyogun (1988) to more than 10% in some 3-item examples with relatively high service level requirements.

The rest of the paper is organized as follows: In Section 2 we specify the model for our value-based $Q(s, S)$ policy. Evaluation and comparison of alternative policies by means of simulation are contained in Section 3. The numerical examples studied are
taken directly from or are modified from examples found in the literature. Finally, our conclusions are presented in Section 4.

2 Modeling

In Subsection 2.5 we show how the expressions, needed to determine the optimal (or a near-optimal) \( Q(s, S) \) policy, can be used to compute our value-based \( Q(s, S) \) policy. The first two subsections present the ingredients used in Subsection 2.3 to describe how Nielsen and Larsen (2005) suggest to determine the optimal \( Q(s, S) \) policy. Subsection 2.4 explains how to compute relative state values when a replenishment order is issued. These state-values are needed in Subsection 2.5 for the value-based \( Q(s, S) \) policy.

2.1 Expected costs incurred between demand instants

The total demand of all items is Poisson with rate \( \mu = \sum_{i=1}^{n} \lambda_i \) and, at all instants \( t \), \( \alpha_i = \frac{1}{\mu} \) is the probability that the next demand is for item \( i \). Because excess demand is backlogged and the lead time of each replenishment order for item \( i \) is a constant \( L_i \), the net inventory of item \( i \) at time \( t + L_i \) is distributed as \( IP_i(t) - D_i(L_i) \), where \( IP_i(t) \) is the inventory position of item \( i \) at instant \( t \) and \( D_i(L_i) \) is the generic variable for the lead-time demand of item \( i \). Hence, \( D_i(L_i) \) has the Poisson distribution with mean \( \lambda_i L_i \). Moreover, it is a standard procedure to specify the expected holding and penalty costs of item \( i \) from any instant \( t \) to the next demand instant as the expected such costs incurred from instant \( t + L_i \) to the next demand instant after \( t + L_i \) (Federgruen and Schechner, 1983). We use this observation below in Eq. (1) at any demand instant.

Suppose that the inventory position of item \( i \) is \( x_i \) just after a considered demand instant. Then the expected costs \( \gamma_i(x_i) \) of item \( i \) incurred between the considered instant and the next demand instant can be computed recursively:

\[
\gamma_i(x_i) = \begin{cases} 
\frac{h_i}{\mu} (\lambda_i L_i - x_i) + \frac{n_i \alpha_i}{\mu}, & x_i = 0, -1, -2, \ldots \\
\gamma_i(x_i - 1) + \frac{h_i}{\mu} - \frac{b_i + b_i}{\mu} \Pr\{D_i(L_i) \geq x_i\}, & x_i = 1, 2, 3, \ldots \\
-\frac{n_i \alpha_i}{\mu} \Pr\{D_i(L_i) = x_i - 1\}, & x_i = 1, 2, 3, \ldots 
\end{cases}
\]

2.2 Long-run average cost of item \( i \)

For an initial inventory position \( x_i \) of item \( i \), we denote by \( G_{i,Q}(x_i) \) the expected holding and penalty costs of this item incurred until the total demand of all items
reaches $Q$. Hence $G_{i,0}(x_i) = 0$ and recursively for $Q = 1, 2, \ldots$,
\[
G_{i,Q}(x_i) = \gamma(x_i) + \alpha_i G_{i,Q-1}(x_i - 1) + (1 - \alpha_i) G_{i,Q-1}(x_i).
\]
(2)

If item $i$ is controlled by the $(s_i, S_i)$ policy when the total demand since the last replenishment order has reached the level $Q$, then Nielsen and Larsen (2005) have shown that the long-run average cost per unit time for this item is
\[
g_i(Q, s_i, S_i) = \frac{\mu}{Q} k_i + \sum_{j=0}^{S_i-s_i-1} m_i,Q(j) G_i,Q(S_i - j).
\]
(3)

Here $m_i,Q(0) = \frac{1}{1-f_i,Q(0)}$ and
\[
m_i,Q(j) = m_i,Q(0) \sum_{y=1}^{j} f_i,Q(y) m_i,Q(j - y), \quad j = 1, 2, \ldots,
\]
(4)

where
\[
f_i,Q(y) = \binom{Q}{y} (\alpha_i)^y (1 - \alpha_i)^{Q-y}, \quad y = 0, 1, \ldots, Q,
\]
(5)

is the binomial probability mass function (pmf) of the demand of item $i$ during a time period where the total demand of all items is $Q$.

For fixed $Q$, an optimal reorder point $s^*_i(Q)$ and an optimal order-up-to level $S^*_i(Q)$ for item $i$ can be computed by the fast algorithm of Zheng and Federgruen (1991). By $g_i(Q)$ we denote $g_i(Q, s^*_i(Q), S^*_i(Q))$. Note that the assumption of Poisson demand can easily be generalized to compound Poisson demand.

2.3 Optimal $Q(s, S)$ policy

The long-run average cost per unit time of the $Q(s, S)$ policy with item $i$ controlled at replenishment instants by $s^*_i(Q)$ and $S^*_i(Q)$ is
\[
C(Q) = \frac{K \mu}{Q} + \sum_{i=1}^{n} g_i^*(Q).
\]
(6)

It cannot be concluded that $C(Q)$ is the minimum average cost for $Q$ because, at the time of an order opportunity, it may happen that none of the $n$ items have inventory positions less than or equal to their reorder points. However, this is very unlikely if $Q$ is fixed properly. Therefore, Nielsen and Larsen (2005) suggest to fix $Q^*$ by searching for a $Q$ value minimizing $C(Q)$. Our way of searching for $Q^*$ is essentially the following: An initial $Q$ value is obtained by using the procedure in Atkins and Iyogun (1988) to compute a common review interval which is then multiplied by $\mu$. Next, with two additional, appropriately chosen, initial values the golden section search procedure (Press et al., 1989, Section 10.1) is applied to compute $Q^*$. 

5
2.4 Relative state-values when a replenishment order is issued

According to the optimal $Q(s, S)$ policy, the replenishment orders are issued when the total demand since the last replenishment order has reached the level $Q^*$ and, for $i = 1, 2, \ldots, n$, item $i$ is controlled by the $(s_i^*(Q^*), \bar{s}_i^*(Q^*))$ policy. Then the demand $Y_i$ of item $i$ between replenishment instants has the pmf specified in Eq. (5) with $Q = Q^*$. By $Y^j_i$ we denote the sum of $j$ mutually independent values of $Y_i$ and by $f^j_{i,Q^*}(y)$ and $F^j_{i,Q^*}(y)$ we denote the pmf and the cumulative distribution function, respectively, of $Y^j_i$. Define $\Delta_i = S^*_i(Q^*) - s^*_i(Q^*) - 1$ and $F^0_{i,Q^*}(y) = 1$, $y = 0, 1, \ldots$

It appears obvious to identify the state of the inventory system at an instant when a replenishment order is issued as the $n$-dimensional vector $z$ with component $z_i = 0$ if item $i$ is included in the order. If not, $z_i = S^*_i(Q^*) - x_i$ where $x_i$ is the actual inventory position of item $i$. However, the number of such states is very large unless $n$ and/or $\Delta_i$ are/is small. In order to make the state space tractable, we divide the items into groups indexed by $j$, where $j = 0$ refers to items $i$ for which $F_{j,Q^*}(\Delta_i) < \epsilon_1$. A positive $j$ refers to items $i$ for which

$$\frac{F^{j+1}_{i,Q^*}(\Delta_i)}{F^j_{i,Q^*}(\Delta_i)} < \epsilon_1 \leq \frac{F^{j}_{i,Q^*}(\Delta_i)}{F^{j-1}_{i,Q^*}(\Delta_i)}.$$  \hspace{1cm} (7)

We let $\text{Group}(i)$ denote $j$ for the group of item $i$. In our numerical examples we have set $\epsilon_1 = 0.001$. In Subsection 3.2 we have restricted $\text{Group}(i)$ to be at most two.

We specify the state just after each order instant as a vector $r = (r_1, r_2, \ldots, r_n)$, where component $r_i$ is the current consecutive number of order instants at which item $i$ was not ordered. We assume as approximations for items $i$ that

$$0 \leq r_i \leq \text{Group}(i), \ i = 1, 2, \ldots, n, \hspace{1cm} (8)$$

and, when $r_i > 0$, that the inventory position is $S^*_i(Q^*) - y$ with pmf

$$\phi^{r_i}(y) = \frac{f^r_{i,Q^*}(y)}{F^r_{i,Q^*}(\Delta_i)}, \ y = 0, 1, \ldots, \Delta_i.$$  \hspace{1cm} (9)

We let $R$ be the set of all $n$-dimensional vectors $r$, which have integer components satisfying Cond. (8).

By $P_{r,r'}$ we denote the probability of a transition from state $r$ just after one order instant to state $r'$ just after the next order instant. The set of states, which can be reached from state $r \in R$ after one transition, is $R(r) = \{r' \in R \mid P_{r,r'} > 0\}$. Note from Cond. (8) that if state $r$ has the property $r_i = \text{Group}(i)$, then $r'_i = 0$ for any
state \( r' \in R(r) \). If \( r_i < \text{Group}(i) \), then either \( r'_i = 0 \) or \( r'_i = r_i + 1 \) for the states \( r' \in R(r) \). Hence, it is straightforward but tedious to compute \( P_{r,r'} \) in terms of the multinomial distribution specifying the demands during a time period, where the total demand of all items is \( Q^* \), of all items \( i \) with the property \( r_i < \text{Group}(i) \).

With state \( r \) we associate the expected cost

\[
c_r = K + \sum_{r' \in R(r)} P_{r,r'} \sum_{i=1}^{n} \hat{c}_{r'_i},
\]

where

\[
\hat{c}_{r'_i} = \begin{cases} 
  k_i + G_{i,Q^*}(S_i^*(Q^*)), & r'_i = 0 \\
  \sum_{y=0}^{\Delta_i} \phi_{i,y} G_{i,Q^*}(S_i^*(Q^*) - y), & r'_i > 0.
\end{cases}
\]

When being in state \( r \) just after an order instant, the expected future costs over \( N \) order instants can be computed recursively for \( N = 1, 2, \ldots \)

\[
V_N(r) = c_r + \sum_{r' \in R(r)} P_{r,r'} V_{N-1}(r'), \ r \in R,
\]

where \( V_0(r') = 0 \). Motivated by the value-iteration algorithm (Tijms 2003, Section 6.6), we suggest to stop the computations as soon as

\[
M_N - m_N < \epsilon_2,
\]

where \( M_N = \max_{r \in R} \{ V_N(r) - V_{N-1}(r) \} \) and \( m_N = \min_{r \in R} \{ V_N(r) - V_{N-1}(r) \} \). With \( N \) determined in this way it can be concluded that \( \frac{m_N + M_N}{2} \) is close to \( C(Q^*)\frac{Q^*}{\mu} \). Over a remaining horizon where the total demand is \( NQ^* \) units, we let \( V_N(r) \) be our evaluation of being in state \( r \) just after an order instant. If the total demand over the remaining horizon is reduced by \( q \) units, then we reduce the state value by \( C(Q^*)\frac{Q^*}{\mu} \).

2.5 Value-based \( Q(s, S) \) policy

Our value-based \( Q(s, S) \) policy modifies the \( Q(s, S) \) policy when it is expected that this policy can be improved by a simple modification of when the replenishment is issued. It is not very likely that an optimal \( Q(s, S) \) policy can be improved by immediately issuing a replenishment order, unless the total demand \( q \) remaining until the next planned order is small. Therefore, we suggest to review the inventory system continuously only when \( q \) is at or below some threshold \( Q_1 \). In our numerical examples we investigate different levels of \( Q_1 \) in relation to \( Q^* \). When \( 0 < q \leq Q_1 \), we evaluate just after each demand instant the expected cost of immediately ordering \( S_i^*(Q^*) - x_i \) units of items \( i \) with inventory position \( x_i \leq s_i^*(Q^*) \). We compare this...
expected cost to the expected cost of ordering in the same way just after the next demand. If the first option is best and if its expected cost is also less than the expected cost of ordering according to the optimal $Q(s, S)$ policy, then the option is implemented.

Let $\mathbf{1}(\cdot)$ denote the indicator function being one if its argument is true and zero if the argument is false. Suppose that the state of the inventory system was $\mathbf{r}$ at the last order instant and that the vector of inventory positions is $\mathbf{x}$ just after a considered demand instant, where the total demand remaining until the next planned order is $q \leq Q_1$. We evaluate the expected cost of ordering immediately as

$$V_{\text{now}}(\mathbf{r}, \mathbf{x}) = \sum_{i=1}^{n} \left[ k_i + G_{i,Q^*}(s_i^*(Q^*)) \right] \mathbf{1}(x_i \leq s_i^*(Q^*)) + G_{i,Q^*}(x_i) \mathbf{1}(x_i > s_i^*(Q^*)) + V_N(\mathcal{R}(\mathbf{r}, \mathbf{x})), \quad (13)$$

where $\mathcal{R}(\mathbf{r}, \mathbf{x})$ is the state vector for which component $i$ is

$$\mathcal{R}_i(\mathbf{r}, \mathbf{x}) = \begin{cases} 0, & x_i \leq s_i^*(Q^*) \\ \min \{ r_i + 1, \text{Group}(i) \}, & x_i > s_i^*(Q^*) \end{cases}. \quad (14)$$

The expected incremental cost of ordering just after the next demand rather than immediately is evaluated as

$$I(\mathbf{r}, \mathbf{x}) = \sum_{i=1}^{n} (G_{i,1}(x_i) + \alpha_i \cdot \begin{pmatrix} [k_i + G_{i,Q^*}(s_i^*(Q^*))] + \Delta V_N(\mathcal{R}(\mathbf{r}, \mathbf{x}); \mathbf{e}_i) - G_{i,Q^*}(x_i) \\
 \cdot \mathbf{1}(x_i = s_i^*(Q^*) + 1) \\
 + [G_{i,Q^*}(x_i - 1) - G_{i,Q^*}(x_i)] \cdot \mathbf{1}(x_i > s_i^*(Q^*) + 1) \end{pmatrix} - C(Q^*) \frac{1}{\mu}, \quad (15)$$

where $\mathbf{e}_i$ is the unit vector describing that the next demand is for item $i$ which, when $x_i = s_i^*(Q^*) + 1$, changes the state value upon ordering by

$$\Delta V_N(\mathcal{R}(\mathbf{r}, \mathbf{x}); \mathbf{e}_i) = V_N(\mathcal{R}(\mathbf{r}, \mathbf{x}) - \mathcal{R}_i(\mathbf{r}, \mathbf{x})\mathbf{e}_i) - V_N(\mathcal{R}(\mathbf{r}, \mathbf{x})). \quad (16)$$

We evaluate the expected cost of ordering according to the optimal $Q(s, S)$ policy as

$$V_{\text{postpone}}(\mathbf{r}, \mathbf{x}; q) = \sum_{i=1}^{n} \left( G_{i,q}(x_i) + [k_i + G_{i,Q^*}(s_i^*(Q^*))] \Pr \{ x_i \leq s_i^*(Q^*) + D_i(q) \} \\
+ G_{i,Q^*}(x_i) \Pr \{ x_i > s_i^*(Q^*) + D_i(q) \} \right) \\
+ E[V_N(\mathcal{R}(\mathbf{r}, \mathbf{x}; q)] - C(Q^*) \frac{2}{\mu}. \quad (17)$$
Here $\mathcal{D}_i(q)$ is the $i^{th}$ component of the vector $\mathcal{D}(q)$ of the demands during a period where the total demand of all items is $q$, and $\mathcal{R}(r, x; q)$ is the stochastic state vector for which component $i$ is

$$\mathcal{R}_i(r, x; q) = \begin{cases} 0, & x_i \leq s_i^*(Q^*) + \mathcal{D}_i(q) \\ \min\{r_i + 1, \text{Group}(i)\}, & x_i > s_i^*(Q^*) + \mathcal{D}_i(q). \end{cases}$$

Hence $\mathcal{D}(q)$ has an $n$-dimensional multinomial distribution. To compute the expected cost of postponement it suffices to focus on the multinomial distribution for the demands of the items $i$ with the property $\mathcal{R}_i(r, x) > 0$. If only few items have this property and $q$ is close to $Q_1$, then the restricted focus reduces the effort needed to compute the expected cost in Eq. (17) dramatically, but the effort remains much heavier than what is needed to compute the expected incremental cost in Eq. (15). Fortunately, the former computation occurs rarely compared to the latter.

Rather than always ordering when $q = 0$, implying that issuing an order is prescribed by the optimal $Q(s, S)$ policy, our value-based $Q(s, S)$ policy postpones to issue an order until the expected incremental cost $I(r, x) > 0$ for the actual vector $x$ of inventory positions.

## 3 Numerical study

We shall examine various numerical examples. For each example we compute the pure $Q(s, S)$ policy by an algorithm programmed in Visual Basic for Applications (VBA), which is embedded in the Arena simulation software (Version 7.01.00, Kelton et al., 2004). In Arena we have also implemented a model, which makes use of prescriptions written in VBA, to simulate the average cost per unit time of the value-based $Q(s, S)$ policy with threshold $Q_1$ fixed as 1 plus the integer part of a desired fraction of $Q^*$. (The algorithm and simulation model can be obtained from the corresponding author upon request.)

All simulated average costs have 95% confidence intervals computed from 10 replications. Each replication runs the inventory system over 10,000 time units following a 100 time units long warm-up interval, initialized with no replenishment orders outstanding and with inventory position $S_i^*(Q^*)$ for item $i$.

### 3.1 Problems with 12 items

Atkins and Iyogun (1988) present a number of 12-item problems with values of $\lambda_i$, $k_i$ and $L_i$ listed in their Table 1. In that table the average costs of various policies are reported for a standard example. It has the same parameter values as in panel...
(d) with \( K = 150 \) in our Table 1. Most of the parameter values in this table are also investigated by Viswanathan (1997), Schultz and Johansen (1999), Melchiors (2002) and Nielsen and Larsen (2005). The last reference concludes that the average costs of all investigated examples are minimized by their \( Q(s, S) \) policy.

For the standard example, the pure \( Q(s, S) \) policy with \( Q^* = 195 \) computed by Nielsen and Larsen is the same as the policy computed by our algorithm programmed in VBA. The groups of the items are

\[
Group(i) = \begin{cases} 
0, & i \in \{1, 2, 3, 4, 5\} \\
1, & i = 6 \\
2, & i = 9 \\
3, & i \in \{7, 8, 10, 11, 12\}
\end{cases}
\]

and the number of states in the set \( R \) is \( 2 \cdot 3 \cdot 4^5 = 6144 \). The minimum number \( N \) satisfying Eq. (12) with \( \epsilon_2 = 0.001 \) is \( N = 103 \), and \( \frac{m_{103} + M_{103}}{2} \cdot \bar{Q}^* = 2251.5573 \) is very close to \( C(Q^*) = 2251.5564 \) from Eq. (6).

From Table 1 it can be concluded that our value-based policy dominates the hitherto best solutions which are the pure \( Q(s, S) \) policies. The improvements obtained are relatively small (less than 1.5% using the estimated mean values). However, it is interesting to note that by using information about the economic value of deviating from the pure \( Q(s, S) \) policy, the performance of the joint replenishment system can be further improved even for a set of well-studied examples.

In a further investigation of the result of different joint replenishment policies, Melchiors (2002) uses simulation to find the optimal can-order policy in an example with 12 identical items. The same example is investigated by Brønmo (2005, pages 94–111) in a paper entitled “A study of a combined can-order and \( Q(s, S) \) policy applied on the joint replenishment problem”. In that paper the pure \( Q(s, S) \) policy is shown to outperform the can-order policy in five out of six cases with different parameter values. Our value-based policy results in further slight decreases of the average costs in four out of the six cases. In the remaining two cases the number of states in the set \( R \) becomes overwhelmingly large. Because all items in these two cases belong to \( Group(i) = 3 \), the number of states becomes \( 4^{12} = 16,777,216 \). Even if \( Group(i) \) is restricted to be at most two, the number of states is impractically large (\( 3^{12} = 531,441 \)). Further restricting \( Group(i) \) to be at most one significantly reduces the number of states (\( 2^{12} = 4,096 \)) but makes the approximation induced by Cond. (8) too coarse and the model thus ill-specified.
Table 1
Average costs of various policies in four cases under variations of the setup cost $K$ in the 12-item problem introduced by Atkins and Iyogun (1988)

<table>
<thead>
<tr>
<th></th>
<th>Can-order</th>
<th>Periodic review</th>
<th>Demand reporting</th>
<th>Value-based</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Erlang$^1$</td>
<td>Compensation$^2$</td>
<td>$(R, T)^3$</td>
<td>$P(s, S)^4$</td>
</tr>
<tr>
<td>(a) $h_i = 2$, $b_i = 30$ and $\pi_i = 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$K = 50$</td>
<td>956.9±0.9</td>
<td>1029</td>
<td>944</td>
<td>928.9</td>
</tr>
<tr>
<td>$K = 100$</td>
<td>1022.9±0.6</td>
<td>1034</td>
<td>1005</td>
<td>992</td>
</tr>
<tr>
<td>$K = 150$</td>
<td>1067.5±0.9</td>
<td>1106</td>
<td>1046</td>
<td>1043</td>
</tr>
<tr>
<td>$K = 200$</td>
<td>1113.2±1.3</td>
<td>1159</td>
<td>1088</td>
<td>1085</td>
</tr>
<tr>
<td>(b) $h_i = 2$, $b_i = 0$ and $\pi_i = 30$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$K = 50$</td>
<td>1125.0±0.2</td>
<td>1109</td>
<td>1135</td>
<td>1121</td>
</tr>
<tr>
<td>$K = 100$</td>
<td>1197.0±0.4</td>
<td>1211</td>
<td>1198</td>
<td>1185</td>
</tr>
<tr>
<td>$K = 150$</td>
<td>1253.3±0.9</td>
<td>1289</td>
<td>1244</td>
<td>1241</td>
</tr>
<tr>
<td>$K = 200$</td>
<td>1306.5±1.2</td>
<td>1355</td>
<td>1289</td>
<td>1285</td>
</tr>
<tr>
<td>(c) $h_i = 6$, $b_i = 30$ and $\pi_i = 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$K = 50$</td>
<td>1540.3±1.0</td>
<td>1542</td>
<td>1570$^2$</td>
<td>1530$^2$</td>
</tr>
<tr>
<td>$K = 100$</td>
<td>1632.2±0.5</td>
<td>1660</td>
<td>1676</td>
<td>1623</td>
</tr>
<tr>
<td>$K = 150$</td>
<td>1724.6±0.5</td>
<td>1747</td>
<td>1733</td>
<td>1706</td>
</tr>
<tr>
<td>$K = 200$</td>
<td>1809.7±0.6</td>
<td>1821</td>
<td>1803</td>
<td>1778</td>
</tr>
<tr>
<td>(d) $h_i = 6$, $b_i = 0$ and $\pi_i = 30$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$K = 50$</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>$K = 100$</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>$K = 150$</td>
<td>2288.7±0.6</td>
<td>2313.0±0.8</td>
<td>2291</td>
<td>2267</td>
</tr>
<tr>
<td>$K = 200$</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

$^1$ Schultz and Johansen (1999)

$^2$ Melchiors (2002). He reports a confidence interval for his simulated cost of the compensation policy only in case (d) with $K = 150$

$^3$ Atkins and Iyogun (1988)

$^4$ Viswanathan (1997)

$^5$ Our algorithm confirms the results (without the decimals) provided by Nielsen and Larsen (2005)

* Denotes that the average cost is not reported by the reference.
Table 2
Item-specific parameter values of the three-item example in Brønmo (2005)

<table>
<thead>
<tr>
<th></th>
<th>item 1</th>
<th>item 2</th>
<th>item 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>demand rate</td>
<td>$\lambda_1 = 0.5$</td>
<td>$\lambda_2 = 0.5$</td>
<td>$\lambda_3 = 0.25$</td>
</tr>
<tr>
<td>holding cost rate</td>
<td>$h_1 = 2$</td>
<td>$h_2 = 1$</td>
<td>$h_3 = 2$</td>
</tr>
<tr>
<td>backorder cost rate</td>
<td>$b_1 = 4$</td>
<td>$b_2 = 2$</td>
<td>$b_3 = 3$</td>
</tr>
<tr>
<td>shortage cost per unit</td>
<td>$\pi_1 = 30$</td>
<td>$\pi_2 = 20$</td>
<td>$\pi_3 = 20$</td>
</tr>
</tbody>
</table>

Table 3
Average costs of various policies for the examples with two and three items

<table>
<thead>
<tr>
<th></th>
<th>$Q(s, S)^\dagger$</th>
<th>Value-based</th>
<th>Optimal$^\ddagger$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$Q_1 = \text{Int}(\frac{Q^*}{30}) + 1$</td>
<td>$Q_1 = \text{Int}(\frac{Q^*}{2}) + 1$</td>
<td></td>
</tr>
<tr>
<td>(a) Two items</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k = 0$</td>
<td>15.46</td>
<td>15.24±0.11</td>
<td>15.19 ± 0.08</td>
</tr>
<tr>
<td>$k = 10$</td>
<td>17.83</td>
<td>17.56±0.10</td>
<td>17.41 ± 0.09</td>
</tr>
<tr>
<td>$k = 20$</td>
<td>19.81</td>
<td>19.47±0.08</td>
<td>19.28 ± 0.12</td>
</tr>
<tr>
<td>$k = 30$</td>
<td>21.43</td>
<td>21.13±0.12</td>
<td>20.95 ± 0.09</td>
</tr>
<tr>
<td>(b) Three items</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k = 0$</td>
<td>19.96</td>
<td>19.72±0.09</td>
<td>19.60 ± 0.10</td>
</tr>
<tr>
<td>$k = 10$</td>
<td>23.57</td>
<td>23.14±0.11</td>
<td>23.04 ± 0.08</td>
</tr>
<tr>
<td>$k = 20$</td>
<td>26.41</td>
<td>25.94±0.11</td>
<td>25.67 ± 0.17</td>
</tr>
<tr>
<td>$k = 30$</td>
<td>28.88</td>
<td>28.44±0.13</td>
<td>28.19 ± 0.16</td>
</tr>
</tbody>
</table>

The other parameter values are $K = 30$, $L_i = 2$ and as reported in Table 2.

$^\dagger$ Brønmo (2005, pages 91-92)

3.2 Problems with two and three items

Motivated by Ohno and Ishigaki (2001), Brønmo and Larsen in their paper “Development of the theoretical optimal policy for the joint replenishment problem: Implementation issues and comparisons with the optimal $Q(s, S)$ policy”, which is included in Brønmo (2005), have designed a modified value-iteration algorithm for computing the optimal policy for joint replenishment problems with a few items. We shall investigate their two- and three-item examples with the item-specific parameter values listed in Table 2. The joint setup cost is $K = 30$ and the constant lead time is $L_i = 2$, whereas the item-specific setup cost $k_i$ varies as indicated in Table 3. The example with two items in panel (a) is obtained by neglecting item 3.
The results reported in Table 3 indicate that the average-cost performance of the pure $Q(s, S)$ policy can again be improved with the value-based policy. On average, the lowest average costs reported for the value-based policy are less than 1% above the minimum average costs, whereas the pure $Q(s, S)$ policy has average costs exceeding the optimal costs by more than 3% on average. With the value-based policy the magnitude of the improvement evidently depends on the choice of the threshold value $Q_1$. For the last example in Table 3, Figure 1 shows the simulated average costs of the value-based policy for various values of $Q_1$. As expected, the choice of $Q_1$ exhibits a pattern of marginally decreasing benefits. A similar effect is also suggested by the results shown in Table 3. Hence, for the remaining parts of this study we have chosen to use $Q_1 = \text{Int}(\frac{Q^*}{2}) + 1$.

For the three-item example in Table 3 with $k = 10$ as a base case we have investigated further the effect of applying the value-based policy compared to using the pure $Q(s, S)$ policy. The base case was first varied by doubling and reducing to zero, respectively, the parameters $k$, $b_i$, and $\pi_i$ while keeping the other parameters fixed. Then the parameters $K$ and $L_i$ were doubled and halved, while $k$ was doubled and reduced to zero as before and the other parameters were fixed. This results in a total of 50 parameter combinations (including the base case, but excluding the cases in which both penalty costs are zero). Again, the value-based policy dominates the pure $Q(s, S)$ policy. The relative reduction of the average cost for these 50 cases ranges from about 1% to about 5%.

However, as shown in Table 2, the ratio between the backorder cost rate and the holding cost rate in the three-item example is only 2, 2, and 1.5 for the three items, respectively. This might result in unrealistically low inventory service levels, particularly when the unit shortage cost is low. Therefore, we have investigated another set of parameter variations for the three-item problem. Table 4 shows the design and the results of this investigation. To accommodate requirements for high service levels, when the unit shortage costs $\pi_i$ are fixed at zero, the backorder cost rates $b_i$ were increased considerably. Using the holding cost rates $h_i$ from the base case, the backorder cost rates were chosen so that the ratio between backorder cost and holding cost rates was either 16, 32, or 64 for all three items in panel (a). In addition, three different levels for the joint setup cost were analyzed by also doubling and quadrupling the value of $K$ in the base case. Finally, heterogeneity among the items was further increased by varying the holding cost rate $h_3$ for item 3 between panels (a), (b), and (c).
Table 4
Average costs of the optimal $Q(s, S)$ policy and the value-based $Q(s, S)$ policy (with 95% confidence intervals) for various examples with three items

<table>
<thead>
<tr>
<th></th>
<th>$b = (32, 16, 32)$</th>
<th>$b = (64, 32, 64)$</th>
<th>$b = (128, 64, 128)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(a)$ $h_3 = 2$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$K = 30$</td>
<td>$31.61$</td>
<td>$Q^* = 10$</td>
<td>$34.49$</td>
</tr>
<tr>
<td></td>
<td>$28.95 \pm 0.16$</td>
<td>$8.4%$</td>
<td>$31.29 \pm 0.15$</td>
</tr>
<tr>
<td>$K = 60$</td>
<td>$34.79$</td>
<td>$Q^* = 13$</td>
<td>$37.80$</td>
</tr>
<tr>
<td></td>
<td>$32.23 \pm 0.16$</td>
<td>$7.4%$</td>
<td>$34.73 \pm 0.16$</td>
</tr>
<tr>
<td>$K = 120$</td>
<td>$39.69$</td>
<td>$Q^* = 16$</td>
<td>$43.22$</td>
</tr>
<tr>
<td></td>
<td>$37.38 \pm 0.17$</td>
<td>$5.8%$</td>
<td>$40.01 \pm 0.19$</td>
</tr>
<tr>
<td>$(b)$ $h_3 = 4$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$K = 30$</td>
<td>$35.32$</td>
<td>$Q^* = 8$</td>
<td>$39.13$</td>
</tr>
<tr>
<td></td>
<td>$32.36 \pm 0.10$</td>
<td>$8.4%$</td>
<td>$35.13 \pm 0.12$</td>
</tr>
<tr>
<td>$K = 60$</td>
<td>$39.23$</td>
<td>$Q^* = 13$</td>
<td>$43.22$</td>
</tr>
<tr>
<td></td>
<td>$36.62 \pm 0.18$</td>
<td>$6.6%$</td>
<td>$39.78 \pm 0.20$</td>
</tr>
<tr>
<td>$K = 120$</td>
<td>$44.57$</td>
<td>$Q^* = 15$</td>
<td>$49.35$</td>
</tr>
<tr>
<td></td>
<td>$41.65 \pm 0.25$</td>
<td>$6.5%$</td>
<td>$45.61 \pm 0.12$</td>
</tr>
<tr>
<td>$(c)$ $h_3 = 8$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$K = 30$</td>
<td>$40.37$</td>
<td>$Q^* = 6$</td>
<td>$45.95$</td>
</tr>
<tr>
<td></td>
<td>$36.57 \pm 0.21$</td>
<td>$9.9%$</td>
<td>$41.93 \pm 0.24$</td>
</tr>
<tr>
<td>$K = 60$</td>
<td>$45.28$</td>
<td>$Q^* = 9^1$</td>
<td>$51.01$</td>
</tr>
<tr>
<td></td>
<td>$40.77 \pm 0.21$</td>
<td>$10.0%$</td>
<td>$46.35 \pm 0.23$</td>
</tr>
<tr>
<td>$K = 120$</td>
<td>$51.32$</td>
<td>$Q^* = 13$</td>
<td>$58.21$</td>
</tr>
<tr>
<td></td>
<td>$47.86 \pm 0.17$</td>
<td>$6.7%$</td>
<td>$53.31 \pm 0.29$</td>
</tr>
</tbody>
</table>

The parameter values $\lambda_i$, $h_1$ and $h_2$ are specified in Table 3, whereas $\pi_i = 0$. The percentages are the mean relative cost reductions obtained from using the value-based $Q(s, S)$ policy rather than the optimal $Q(s, S)$ policy.

Our golden section search procedure finds two local minima:

$^1$ $Q^* = 12$ with $C(12) = 45.34$

$^2$ $Q^* = 13$ with $C(13) = 58.39$. 
Examining the results in Table 4 it is evident that the magnitude of the cost reductions is significantly higher for this set of parameter variations than for any of the previously analyzed cases. The relative cost reductions obtained from using the value-based $Q(s, S)$ policy rather than the optimal $Q(s, S)$ policy are now in the range from 5.8% to 12.3% (using the estimated means from the simulations to calculate the percentages). We also observe from Table 4 that when $K$ is increased, then as expected $Q^*$ always increases, and in most cases the relative cost reduction decreases. Furthermore, when the shortage cost vector $b$ is multiplied by scalars 2 and 4, respectively, then there is a tendency toward a lower $Q^*$ and a higher relative cost reduction.

Note from the results reported in Table 4 that our golden section search procedure in two of the cases finds a local minimum, which is not the global minimum found by complete enumeration. The two cases indicate the need in general for an effective search algorithm that finds the global minimum cost policy. As a final comment, we observe from the simulations in Table 4 that when $K$ is increased, then as expected $Q^*$ always increases, and in most cases the relative cost reduction decreases. Furthermore, when the shortage cost vector $b$ is multiplied by scalars 2 and 4, respectively, then there is a tendency toward a lower $Q^*$ and a higher relative cost reduction.

Note from the results reported in Table 4 that our golden section search procedure in two of the cases finds a local minimum, which is not the global minimum found by complete enumeration. The two cases indicate the need in general for an effective search algorithm that finds the global minimum cost policy. As a final comment, we observe from the simulations in Table 4 that when $K$ is increased, then as expected $Q^*$ always increases, and in most cases the relative cost reduction decreases. Furthermore, when the shortage cost vector $b$ is multiplied by scalars 2 and 4, respectively, then there is a tendency toward a lower $Q^*$ and a higher relative cost reduction.

Note from the results reported in Table 4 that our golden section search procedure in two of the cases finds a local minimum, which is not the global minimum found by complete enumeration. The two cases indicate the need in general for an effective search algorithm that finds the global minimum cost policy. As a final comment, we observe from the simulations in Table 4 that when $K$ is increased, then as expected $Q^*$ always increases, and in most cases the relative cost reduction decreases. Furthermore, when the shortage cost vector $b$ is multiplied by scalars 2 and 4, respectively, then there is a tendency toward a lower $Q^*$ and a higher relative cost reduction.

Note from the results reported in Table 4 that our golden section search procedure in two of the cases finds a local minimum, which is not the global minimum found by complete enumeration. The two cases indicate the need in general for an effective search algorithm that finds the global minimum cost policy. As a final comment, we observe from the simulations in Table 4 that when $K$ is increased, then as expected $Q^*$ always increases, and in most cases the relative cost reduction decreases. Furthermore, when the shortage cost vector $b$ is multiplied by scalars 2 and 4, respectively, then there is a tendency toward a lower $Q^*$ and a higher relative cost reduction.

4 Conclusions

We have designed an algorithm for improving the recently proposed $Q(s, S)$ policy for the stochastic joint replenishment problem. The improvement is obtained by considering the economic value of deviating from the pure $Q(s, S)$ policy. In our value-based $Q(s, S)$ policy, item inventories are reviewed continuously when the total demand since the last replenishment is close to $Q$. At each review instant an order is issued according to the item-specific $(s, S)$ policies, if the expected cost of ordering is less than the expected cost of deferring the ordering until the next demand or until the total demand level $Q$ has been reached.

We have evaluated the proposed policy by simulation. Our value-based policy dominates the pure $Q(s, S)$ policy in all numerical examples tested. In a set of three-item examples, for which the ratio between the backorder cost rate and the holding cost rate is relatively high, the cost reduction using the value-based policy is in the range of 6-12% of the long-run average cost for the pure $Q(s, S)$ policy.

In the Markov chain, used in our value-based policy to determine relative values of the inventory system, the state of each item has not been specified as the inventory state of the system.
position (or some similar measure). Instead the state is represented by the number of replenishment cycles since the item was last ordered. This limits the size of the state space considerably. However, this is not always the case and then our approach for the value-based $Q(s, S)$ policy suffers from Bellman’s *curse of dimensionality*. To overcome this difficulty, a viable approach might be to apply the concept of a value-based policy to the $P(s, S)$ policy investigated by Viswanathan (1997). For such a policy the value function decomposes naturally in the items, which would be helpful in limiting the size of the state space.

**References**


Atkins D, Iyogun PO. Periodic versus ‘can–order’ policies for coordinated multi–item inventory systems. Management Science 1988;34(6); 791–796.


Federgruen A, Schechner Z. Cost formulas for continuous review inventory models with fixed delivery lags. Operations Research 1983;31(5); 957–965.


Nielsen C, Larsen C. An analytical study of the $Q(s, S)$ policy applied to the joint replenishment problem. European Journal of Operational Research 2005;163(3); 721–732.

Ohno K, Ishigaki T. A multi–item continuous review inventory system with compound Poisson demand. Mathematical Methods of OR 2001;53(1); 147-165.


Schultz H, Johansen SG. Can-order policies for coordinated inventory replenishment with Erlang distributed times between ordering. European Journal of Operational Research 1999;113(1); 30–41.


Viswanathan S. Periodic review $(s, S)$ policies for joint replenishment inventory systems. Management Science 1997;43(10); 1447–1454.

Average cost

Figure 1: Simulated average costs per unit time and the lower and upper limits of their 95% confidence intervals for various value-based $Q(s, S)$ policies in the three-item example with $k = 30$ in Brønmo (2005, pages 81–93) for which $Q^* = 15$. 