Coordinated Replenishment Game and Learning Under Time Dependency and Uncertainty of the Parameters

Stefanny Ramirez[1] · Laurence H. van Brandenburg[1] · Dario Bauso[2,3]

Abstract

This research proposes a periodic review multi-item two-layer inventory model. The main contribution is a novel approach to determine the can-order threshold in a two-layer model under time-dependent and uncertain demand and setup costs. The first layer consists of a learning mechanism to forecast demand and forecast setup costs. The second layer involves the coordinated replenishment of items, which is analysed as a Bayesian game with uncertain prior probability distribution. The research builds on the concept of the (S, c, s) policy, which is extended to the case of uncertain and time-dependent parameters.

Keywords Coordination game · Inventory control · Can-order policy · Bayesian game · Distributionally robust optimization

1 Introduction

This paper studies a periodic review multi-item two-layer inventory model to determine the can-order threshold of a (S, c, s) policy in a coordinated replenishment problem. The (S, c, s) policy was first introduced by Balintfy [3] and is useful in situations where the demand is stochastic and where the setup costs are high.

In the (S, c, s) policy, an order is placed to take the inventory level to the order-up-to threshold S. A replenishment occurs either when the inventory level is lower than the must-order threshold s or when the inventory level is lower than the can-order threshold c and

---

✉ Stefanny Ramirez
  s.g.ramirez.juarez@rug.nl

1   ENTEG, Faculty of Science and Engineering, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands

2   Jan C. Willems Center for Systems and Control, ENTEG, Faculty of Science and Engineering, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands

3   Dipartimento di Ingegneria, Università di Palermo, 90128 Palermo, Italy
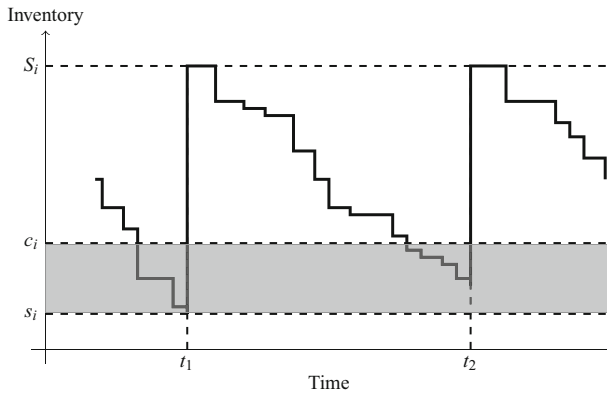
Fig. 1 Inventory evolution over time; in the grey area coordinated replenishment may occur

there is at least another item being ordered. The inventory evolution of a generic item i over time under the (S, c, s) policy is illustrated in Fig. 1.

Various algorithms [4, 8, 18, 23, 29, 32, 33] and heuristics [6, 22, 34] to determine the thresholds of the (S, c, s) policy are established and compared in the literature [12, 19, 21, 26, 27]. Federgruen et al. [8] proposed an algorithm to determine the optimal (S, c, s) policy in a multi-item inventory system with compound Poisson demands. In the work of Zheng [35], the author provides a method to determine the thresholds in a single-item continuous review inventory system with Poisson demand. Similarly, in this work we assume that the demand and discount opportunity of each item evolves according to two independent Poisson processes. However, we consider that the rate parameters of these processes are unknown and change over time.

In practice, demand is often assumed to be uncertain [2, 10, 20, 23–26, 28, 31] and demand forecasting models are widely developed in order to cope with uncertainty [9]. In this research, as the demand follows a Poisson distribution, we apply a Poisson regression model to learn the future demand [5]. This technique is commonly used to forecast a count variable that follows a Poisson distribution and it ensures that the estimated parameter takes values in the set of nonnegative integers $\mathbb{Z}_+$. On the other hand, the setup costs can also be uncertain [1]. In this case, we apply an exponential smoothing method to forecast the values of the setup cost at each time t [11, 17]. The uncertainty in the parameters highlights the versatility of the proposed method and its link to data-driven approaches. Indeed, the method is effective even when we do not know a priori the values of the parameters. Furthermore, considering setup costs as unknown parameters allows to accommodate real-life scenarios where it is common that market conditions generate changes in transportation cost.

In this paper, we focus on determining how the optimal can-order level in a periodic review system with multiple items and discount opportunities can be obtained, while learning uncertain setup cost and uncertain demand. We assume that each item is managed by an independent decision maker. In addition, assuming that the decision-maker of each item does not have information about the demand of any other item, we translate this coordinated replenishment problem into a Bayesian game, where each item is considered as a player. The theory about Bayesian games was introduced by Harsanyi [14–16]. However, in a Bayesian game all the players need to know the prior probability distribution of the demand. In this work, we assume that the players do not know the exact prior probability distribution, but each player knows the set of possible probability distributions. To be protected against the

uncertainty of not knowing the real probability distribution, we apply distributionally robust optimization [30] to obtain the optimal can-order level that minimizes the expected cost.

The main contributions of this paper are:

- We present a new model to describe a coordinated replenishment game which involves a learning mechanism to forecast uncertain and time-dependent parameters. This model is based on Bayesian games and distributionally robust optimization to accommodate the uncertainty in the probability distributions.
- We prove that the cost function is convex with respect to the can-order threshold and that it is concave with respect to the set of probability distributions used for the robust optimization problem. Furthermore we prove that, in terms of Bayesian games, finding an optimal can-order level that minimizes the cost function corresponds to finding a strategy at a Bayesian Nash equilibrium.
- We provide a new method to determine the optimal-can order level that minimizes the cost function.
- We conclude with a numerical analysis to corroborate our theoretical results.

This paper is organized as follows. In Sect. 2, we introduce the model for a specific item and we extend the model to multi-item joint replenishment. In Sect. 3, we analyse the coordinated replenishment model presented in Sect. 2 as a Bayesian game. In Sect. 4, we present our main results. In Sect. 5, we provide a numerical analysis. Finally, in Sect. 6 we draw conclusions and discuss future works.

2 Problem Statement and Model

Consider consecutive time intervals $[t, t + 1)$ for all $t = 0, 1, \ldots$, and let the set of items $N = \{1, 2, \ldots n\}$ be given. Let us assume that each item $i \in N$ is managed by an independent decision maker. For each specific item $i \in N$, we also assume that the customers arrive one at a time. The arrival of a customer indicates the occurrence of a demand event. The cumulative demand in a time interval $[t, \tau]$, where $\tau \in [t, t + 1)$, is given by the number of customers that arrive during that period of time. Let $d_i(\tau) \in \mathbb{Z}_+$ be the cumulative demand of item i from time t to τ, where $\mathbb{Z}_+$ represents the set of nonnegative integers. The value of the cumulative demand $d_i(\tau)$ is reset at the beginning of each time interval $[t, t + 1)$. It means that when $\tau = t^+$, $d_i(\tau) = 0$, where t^+ denotes the first time instant in the interval $[t, t + 1)$. We are interested in determining the cumulative demand that is observed at the end of each discrete-time interval. Let us denote by $D_i(t+1) \in \mathbb{Z}_+$ the observed cumulative demand of item i from time t to $t + 1$. In addition, let us denote by $(t + 1)^- \in [t, t + 1)$ the last time instant in the interval $[t, t + 1)$. The cumulative demand at time $t + 1$ is given by $D_i(t + 1) = d_i((t + 1)^-)$.

It is assumed that $d_i(\tau)$ evolves according to a continuous time Poisson process with rate $\lambda_i(t) \in \mathbb{R}_+$, where $\lambda_i(t)$ represents the average number of demand events that occur during the time interval $[t, t + 1)$, and $\mathbb{R}_+$ denotes the set of positive real numbers. However, at the beginning of each time interval the true value of the rate parameter $\lambda_i(t)$ is unknown and has to be estimated. Let us denote by $\hat{\lambda}_i(t)$ the estimated parameter of the Poisson process. This parameter is learned from past data by applying a Poisson regression model. In this Poisson regression model, the estimated parameter depends on the historical observed cumulative demand $D_i(t)$, as well as on the historical data of the minor and major setup costs, denoted by $a_i(t)$ and $A(t)$, respectively. The minor setup cost $a_i(t) \in \mathbb{R}_+$ is the cost that has to be paid when item i is ordered jointly with at least one other item $j \in N$, $j \neq i$. On the other

hand, the major setup cost $A(t) \in \mathbb{R}_+$ corresponds to the cost that has to be paid when an order is placed and no other items are ordered at the same time.

Before explaining the Poisson regression model, let us introduce the forecast cumulative demand of item i at time $t+1$, denoted by $\hat{D}_i(t+1) \in \mathbb{Z}_+$. We assume that the forecast cumulative demand follows a Poisson distribution with rate $\hat{\lambda}_i(t)$. Let $\mathscr{D}_i(t)$ be the discrete-time random variable of the forecast cumulative demand. The corresponding Poisson distribution has the following probability mass function:

$$Pr(\mathscr{D}_i(t) = \varkappa) = Pr(\varkappa) = \frac{e^{-\hat{\lambda}_i(t)}\hat{\lambda}_i(t)^{\varkappa}}{\varkappa!}, \quad \varkappa \in \mathbb{Z}_+. \tag{1}$$

It is important to remember that the expected value and variance of a Poisson distribution correspond to the value of the rate parameter. Furthermore, we consider that the forecast cumulative demand is given by the value of the estimated parameter $\hat{\lambda}_i(t)$, namely $\sigma^2(\mathscr{D}_i(t)) = \mathbb{E}(\mathscr{D}_i(t)) = \hat{\lambda}_i(t) = \hat{D}_i(t+1)$.

For a specific item i, let us assume that we have historical data of the observed demand and the setup costs. The data consist of m independent observations of the cumulative demand $D_i(t)$, as well as of the major and minor setup costs $A(t)$ and $a_i(t)$, respectively, for all $t = 0, 1 \ldots, m$. In a Poisson regression model, $D_i(t)$ is the dependent variable and $(A(t), a_i(t))$ are the linear independent regressors that determine $D_i(t)$ at each time t. The Poisson regression model in this case is defined by the following equations:

$$\begin{cases} Pr(D_i(t)|(A(t), a_i(t))) = \frac{e^{-\lambda_i(t)}\lambda_i(t)^{D_i(t)}}{D_i(t)!}, \\ \lambda_i(t) = \lambda_i(A(t), a_i(t), \beta_1, \beta_2) = e^{A(t)\beta_1 + a_i(t)\beta_2}, \end{cases} \tag{2}$$

where $Pr(D_i(t)|(A(t), a_i(t)))$ is the conditional probability distribution of the observed cumulative demand given the independent variables $A(t)$ and $a_i(t)$, and the unknown parameters $\beta_1, \beta_2 \in \mathbb{R}$ called regression coefficients. Observe that this conditional probability distribution is also a Poisson distribution, where the parameter $\lambda_i(t)$ is a continuous function that depends on the variables $(A(t), a_i(t))$ and also on the unknown parameters β_1 and β_2.

Let us denote by $\beta = (\beta_1, \beta_2)$ the vector of unknown parameters and by $\hat{\beta} = (\hat{\beta}_1, \hat{\beta}_2)$ its corresponding vector of estimated parameters. In a Poisson regression model, these parameters are estimated by *maximum likelihood estimation (MLE)*. Given m independent observations, the log-likelihood function is given by

$$\begin{aligned} ln\mathscr{L}(\beta) &= ln \prod_{k=1}^{m} \frac{e^{-\lambda_i(k)}\lambda_i(k)^{D_i(k)}}{D_i(k)!} \\ &= ln \prod_{k=1}^{m} \frac{e^{-[e^{A(k)\beta_1 + a_i(k)\beta_2}]}\left[e^{A(k)\beta_1 + a_i(k)\beta_2}\right]^{D_i(k)}}{D_i(k)!} \\ &= \sum_{k=1}^{m} \left\{ [A(k)\beta_1 + a_i(k)\beta_2]D_i(k) - e^{A(k)\beta_1 + a_i(k)\beta_2} - ln(D_i(k)!) \right\}. \end{aligned} \tag{3}$$

To maximize the log-likelihood function, we differentiate (3) with respect to $\beta = (\beta_1, \beta_2)$ and we set the derivative equal to zero. The estimated parameters $\hat{\beta}_1$ and $\hat{\beta}_2$ are obtained from solving the set of nonlinear equations:

$$\begin{cases} \frac{\partial[ln\mathscr{L}(\beta)]}{\partial\beta_1} = \sum_{k=1}^{m} A(k)\left[D_i(k) - e^{A(k)\beta_1 + a_i(k)\beta_2}\right] = 0, \\ \frac{\partial[ln\mathscr{L}(\beta)]}{\partial\beta_2} = \sum_{k=1}^{m} a_i(k)\left[D_i(k) - e^{A(k)\beta_1 + a_i(k)\beta_2}\right] = 0. \end{cases} \tag{4}$$
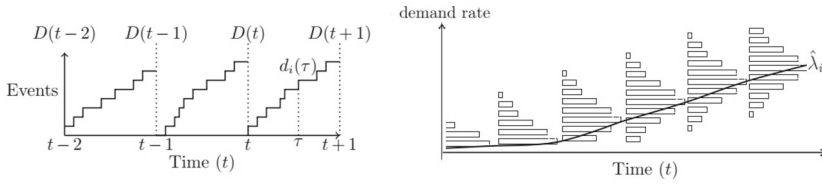
Fig. 2 Demand evolves according to a Poisson process (left). The observed cumulative demand at discrete time $D(t)$ is used to estimate the Poisson rate in the interval $[t, t+1]$ using a Poisson regression model (right)

Note that finding an analytical solution for $\hat{\beta}_1$ and $\hat{\beta}_2$ is prohibitive. However, it is common to apply iterative algorithms to obtain the values of the parameters that maximize the log-likelihood function. The most common method that statistical programs use to find the solution for the parameters is the Newton–Raphson iterative method. By applying this method convergence is guaranteed, as the log-likelihood function (3) is globally concave.

On the other hand, at time $t+1$ the values of the major and minor setup cost are also unknown. The learning mechanism applied to forecast the setup costs is the exponential smoothing. Let us denote by $\hat{A}(t+1)$ and $\hat{a}_i(t+1)$ the forecast values of the major and minor setup costs at time $t+1$, respectively. These values are obtained from the following exponential smoothing model:

$$
\begin{cases}
\mathbb{E}(a_i(t)) = \hat{a}_i(t+1) = \alpha a_i(t) + (1-\alpha)\hat{a}_i(t), \\
\mathbb{E}(A(t)) = \hat{A}(t+1) = \gamma A(t) + (1-\gamma)\hat{A}(t),
\end{cases} \tag{5}
$$

where α and γ are the smoothing coefficients.

Finally, to forecast the value of the demand rate at time $t+1$ we substitute in (2) the estimated parameters $(\hat{\beta}_1, \hat{\beta}_2)$, and the forecast setup costs $(\hat{A}(t+1), \hat{a}_i(t+1))$ as follows:

$$
\mathbb{E}(\mathscr{D}_i(t+1)) = \hat{\lambda}_i(t+1) = e^{\hat{A}(t+1)\hat{\beta}_1 + \hat{a}_i(t+1)\hat{\beta}_2}. \tag{6}
$$

Figure 2(left) displays the continuous cumulative demand $d_i(\tau)$ which evolves according to a Poisson process and the cumulative demand $D_i(t)$ in the interval $[t-1, t)$. The displayed example repeats over three consecutive time periods. Once a Poisson regression model has been fit for the historical information up to time t, we can use this model to forecast the demand rate at time $t+1$ as in Fig. 2(right).

In the (S, c, s) policy, item i is ordered as a result of a discount opportunity when its inventory level is lower than its can-order threshold $c_i(t) \in \mathbb{R}_+$ and there is, at least, an order placed by the decision maker of another item $j \in N$, $j \neq i$. We assume that the discount opportunity events for item i are also modelled by a Poisson process with rate $\mu_i(t) \in \mathbb{R}_+$. However, similarly as in the rate parameter $\lambda_i(t)$, the discount opportunity rate is unknown at the beginning of each time interval $[t, t+1)$ and has to be estimated. Let us denote by $\hat{\mu}_i(t)$ the estimated rate parameter of the Poisson process that describes the discount opportunity events for item i. The discount opportunity rate can be calculated as follows:

$$
\hat{\mu}_i(t) = \sum_{j \in N, j \neq i} \frac{\hat{\lambda}_j(t)}{S_j(t) - s_j(t)} + \sum_{j \in N, j \neq i} \frac{\hat{\lambda}_j(t)}{S_j(t) - c_j(t)} \sum_{k \in N, k \neq j} \frac{\hat{\lambda}_k(t)}{S_k(t) - s_k(t)}. \tag{7}
$$

For the computation of the discount opportunity rate, in (7) we consider the possible events that can generate a discount opportunity. For a specific item i, the first term in (7) corresponds to the possibility that any other item $j \neq i$ is reordered separately. On the other hand, the
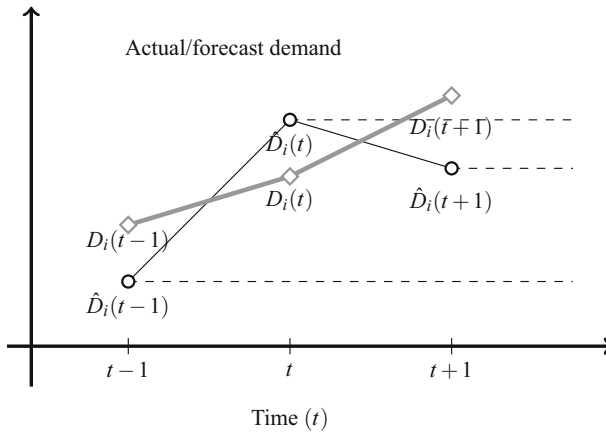
Fig. 3 One iteration of the receding horizon optimization; three samples of observed demand D_i (diamond), and forecast demand $\hat{D}_i$ (circles). Forecast is kept constant throughout the optimization horizon (dashed lines)

second term in (7) corresponds to the possibility that any item $j \in N$, $j \neq i$, is reordered jointly with a different item $k \in N$.

Note that the discount opportunity rate of item i depends on the demand rates of the other items in N. In this work, we assume that the demand Poisson processes for the different items are independent. In Sect. 3, we explain how each player i estimates the values of $\hat{\lambda}_j$ for all $j \neq i$, $j \in N$. It is important to observe that the demand and the discount opportunity are two independent Poisson processes; therefore, any possible event, which involves either the materializing of a demand or the occurrence of a discount opportunity, is determined by a Poisson process with rate $\lambda_i(t) + \mu_i(t)$.

At each time t, an optimization problem over an infinite time horizon is formulated to obtain the optimal thresholds of the can-order policy (S_i, c_i, s_i). In doing this the values of the demand parameters are regarded as constant and equal to the last forecast. This is illustrated in Fig. 3, where $\hat{D}_i(t)$ is the forecast demand at time t, which is regarded as a constant as indicated by the horizontal line. At each time t, a new $\hat{D}_i(t)$ is calculated and used in the new optimization problem.

The must-order $s_i(t) \in \mathbb{N}$ level can be obtained as

$$s_i(t) := \mathbb{E}(\mathscr{D}_i(t)) + k\sqrt{\sigma^2(\mathscr{D}_i(t))} = \hat{\lambda}_i(t) + k\sqrt{\hat{\lambda}_i(t)}. \tag{8}$$

The term k is a safety factor that has to be set to ensure that in a specified percentage of the possible cases shortage does not occur [13].

The order-up-to level $S_i(t) \in \mathbb{N}$ consists of the safety stock SS and the economic order quantity EOQ_i, namely $S_i := SS + EOQ_i$. For the economic order quantity, we have $EOQ_i := \sqrt{2A(t)\mathbb{E}(\mathscr{D}_i(t))/h_i}$, where h_i is the holding cost per unit time. Then, for the order-up-to level we obtain

$$S_i(t) := k\sqrt{\hat{\lambda}_i(t)} + \sqrt{\frac{2A(t)\hat{\lambda}_i(t)}{h_i}}. \tag{9}$$

The dynamics of the inventory level for a specific item i is characterized by a Markov decision-making process. Let us assume that the inventory level of item i at time t is y and
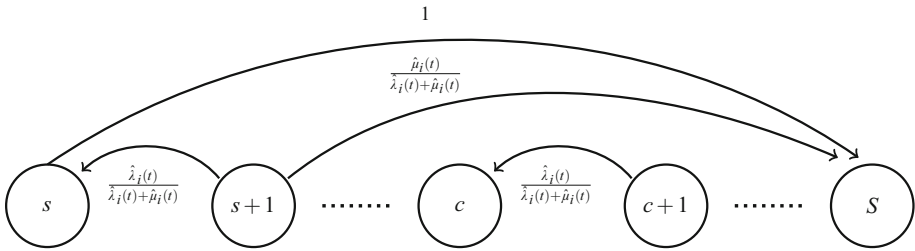
Fig. 4 Markov decision process: inventory levels (nodes) and transitions (arcs)

when the next event occurs, the inventory goes to level x, the transition probabilities $p_{yx}(t)$ between two inventory levels are given by

$$
p_{yx}(t) = \begin{cases}
1 & y \leq s_i, \quad x = S_i, \\
\dfrac{\hat{\lambda}_i(t)}{\hat{\lambda}_i(t)+\hat{\mu}_i(t)} & s_i < y \leq c_i, \quad x = y - 1, \\
\dfrac{\hat{\mu}_i(t)}{\hat{\lambda}_i(t)+\hat{\mu}_i(t)} & s_i < y \leq c_i, \quad x = S_i, \\
\dfrac{\hat{\mu}_i(t)}{\hat{\lambda}_i(t)+\hat{\mu}_i(t)} & c_i < y \leq S_i, \quad x = y, \\
\dfrac{\hat{\lambda}_i(t)}{\hat{\lambda}_i(t)+\hat{\mu}_i(t)} & c_i < y \leq S_i, \quad x = y - 1 \\
0 & \text{otherwise.}
\end{cases}
\tag{10}
$$

The transition probabilities in Eq. (10) describe a data-driven time varying Markov chain. The Markov decision process is illustrated in Fig. 4. The plot depicts transitions between inventory levels. When at time t the inventory is $s_i(t) + 1$ the transition can be to $s_i(t)$ (unit demand) with probability $\dfrac{\hat{\lambda}_i(t)}{\hat{\lambda}_i(t)+\hat{\mu}_i(t)}$ or to S_i (discount opportunity) with probability $\dfrac{\hat{\mu}_i(t)}{\hat{\lambda}_i(t)+\hat{\mu}_i(t)}$. When the inventory is $c_i + 1$ the transition can be to c_i (unit demand) with probability $\dfrac{\hat{\lambda}_i(t)}{\hat{\lambda}_i(t)+\hat{\mu}_i(t)}$ or it remains at the same level in case of discount opportunity with probability $\dfrac{\hat{\mu}_i(t)}{\hat{\lambda}_i(t)+\hat{\mu}_i(t)}$. The expected time between consecutive events is given by $1/(\hat{\lambda}_i(t) + \hat{\mu}_i(t))$.

Given the transition probabilities, the purpose is to determine the optimal policy that minimizes the long-run average cost for each item i. The long-run average cost is defined similarly as in the work of Zheng [35] and is given by

$$
g_i(S_i, c_i, s_i) := \frac{I_i(S_i) + \mathbb{E}(A)Q_i(S_i) + \mathbb{E}(a_i)(1 - Q_i(S_i))}{T_i(S_i)}.
\tag{11}
$$

In the above, assuming that the inventory level is S_i (we are at the beginning of a new replenishment period), we denote the expected time until the next order is placed by $T_i(S_i)$, the expected holding and penalty costs incurred until the next order is placed by $I_i(S_i)$, and the probability that the next order is triggered by a demand by $Q_i(S_i)$.

Let us denote by y the inventory level of item i at time t and by $G_i(y)$ its corresponding one-step inventory cost. In addition, let us denote by $\theta_i(t) = \dfrac{\hat{\lambda}_i(t)}{\hat{\lambda}_i(t)+\hat{\mu}_i(t)}$ the probability that the next event is the occurrence of a unitary demand. In the following result, we explain how to obtain an explicit expression for each term in (11).

Lemma 1 *For the long-run average cost* (11)*, the following equations hold:*

$$T_i(S_i) := \frac{1}{\hat{\lambda}_i(t) + \hat{\mu}_i(t)} \sum_{x=0}^{(c_i-s_i-1)} \theta(t)^x + \frac{1}{\hat{\lambda}_i(t)} (S_i - c_i),$$

$$I_i(S_i) = \frac{1}{\hat{\lambda}_i(t) + \hat{\mu}_i(t)} \sum_{x=0}^{(c_i-s_i-1)} \theta(t)^x G_i(S_i - x) + \frac{1}{\hat{\lambda}_i(t)} \sum_{x=c_i+1}^{S_i} G_i(x),$$

$$Q_i(S_i) = \theta(t)^{(c_i-s_i)}.$$

Proof The proof is in the Appendix. □

Since the parameters involved in the computation of the long-run average cost depend on the thresholds $S_i(t)$, $c_i(t)$, and $s_i(t)$, the minimization of the long-run average cost with fixed $S_i(t)$ and $s_i(t)$ consists in finding the optimal value of $c_i(t)$. The problem we wish to solve can be formulated as follows:

For all $i = 1, \ldots, n$

$$\min_{c_i(t)\in\mathbb{N}} g_i(S_i(t), c_i(t), s_i(t)) := \frac{I_i(S_i(t)) + \mathbb{E}(A(t))Q_i(S_i(t)) + \mathbb{E}(a_i(t))(1 - Q_i(S_i(t)))}{T_i(S_i(t))},$$

(12)

where the expressions for $T_i(S_i)$, $I_i(S_i)$ and $Q_i(S_i)$ are as in Lemma 1.

In the same spirit as in Zheng [35], an algorithm to calculate the optimal can-order level for a single-item inventory system with Poisson demand and Markovian discount opportunities on the setup cost is presented in the rest of this paper. In this research, the long-run average cost formula is extended to a multi-item periodic review system, where the demand rate $\hat{\lambda}_i(t)$ is forecast by a Poisson regression model.

It is important to observe that the long-run average cost of a specific item i depends on the decisions made for the other items. When item i is reordered as a result of a discount opportunity the minor setup cost $a_i(t)$ is paid. Otherwise, whenever item i is reordered without a discount opportunity, the major setup cost $A(t)$ is paid. Nevertheless, the decision-maker of item i does not know the demand rate $\lambda_j(t)$ of any other item, for all $j \in N$, $j \neq i$. In the next section, we explain how the decision-maker of each item i can estimate the demand rate $\hat{\lambda}_j(t)$, for all $j \in N$, $j \neq i$ and how it affects the computation of the optimal can-order level that minimizes the long-run average cost (12). This estimation is based on the theory of Bayesian games as introduced by Harsanyi [14–16].

3 Bayesian Game

In this section, we analyse the coordinated replenishment model presented in Sect. 2 as an *incomplete game*, where each item $i \in N$ represents a player. A game is called incomplete when the players do not have complete information about the important parameters of the game. As it is explained in the previous section, at each time t a new observation of the demand is obtained based on which we estimate the future demand and setup costs. These estimated parameters are used in the Bayesian game and in the computation of the optimal can-order level $c_i^*(t)$ that minimizes the expected long-run average cost. For the sake of simplicity in the rest of this paper, we drop the dependence on time t.

First, let us introduce the control variable u_i for a specific item i. This variable indicates the possible amount to be reordered, which can be triggered either by the occurrence of

a demand event or by a discount opportunity, and it is determined by the corresponding (S_i, c_i, s_i) policy. For any inventory level y, this strategy can be defined formally as follows:

$$
u_i(y) = \begin{cases} S_i - y, & \text{if } y < s_i, \\ S_i - y, & \text{if } s_i < y < c_i \quad \text{and discount opportunity,} \\ 0, & \text{if } s_i < y < c_i \quad \text{and no discount opportunity,} \\ 0, & \text{if } y \geq c_i. \end{cases} \tag{13}
$$

Let us denote by $\mathscr{G}$ the incomplete game related to the initial coordinated replenishment problem. We consider this as an incomplete game since the only information available to player i is its own estimated parameter $\hat{\lambda}_i$. However, player i does not know the estimated parameter $\hat{\lambda}_j$ of any other player $j \in N$, $j \neq i$. Let $\hat{\lambda} = (\hat{\lambda}_1, \ldots, \hat{\lambda}_n)$ be the random vector that contains the demand rates of all the players. Also, let $\hat{\lambda}_{-i} = (\hat{\lambda}_1, \ldots, \hat{\lambda}_{i-1}, \hat{\lambda}_{i+1}, \ldots, \hat{\lambda}_n)$ be the random vector that contains the demand rates of all the players, except player i. The incomplete game $\mathscr{G}$ can be formally defined in normal form as follows:

$$
\mathscr{G} = \langle N, \{U_i\}_{i=1}^n, \{L_i\}_{i=1}^n, \{g_i(S_i, c_i, s_i)\}_{i=1}^n, \{P_i\}_{i=1}^n \rangle, \tag{14}
$$

where $N = \{1, \ldots, n\}$ is the set of players, U_i is set of strategies of player i, L_i is the set of possible values that the rate parameter $\hat{\lambda}_i$ can take, $g_i(S_i, c_i, s_i)$ is the long-run average cost of player i. $P_i = P_i(\hat{\lambda}_{-i}|\hat{\lambda}_i)$ is the conditional probability distribution that player i uses to describe the chance of having certain vector $\hat{\lambda}_{-i}$ given that player i knows the value of its estimated parameter $\hat{\lambda}_i$.

To determine the best strategy for player i, denoted by u_i^*, we need to define a *complete game* that is *Bayes-equivalent* to the incomplete game $\mathscr{G}$. This complete game is also called *Bayesian game*, and it is denoted by $\bar{\mathscr{G}}$. To obtain the Bayesian game, we have to assume that each player i knows not only its own estimated parameter $\hat{\lambda}_i$, but also the common prior probability distribution of $\hat{\lambda}$, represented by $\bar{P}(\hat{\lambda}) = \bar{P}(\hat{\lambda}_1, \ldots, \hat{\lambda}_n)$. Given the prior probability distribution $\bar{P}(\hat{\lambda})$, the complete game $\bar{\mathscr{G}}$ can be formally defined by

$$
\bar{\mathscr{G}} = \left\langle N, \{U_i\}_{i=1}^n, \{L_i\}_{i=1}^n, \{g_i(S_i, c_i, s_i, \hat{\lambda}_i)_{\hat{\lambda}_i \in L_i}\}_{i=1}^n, \bar{P} \right\rangle. \tag{15}
$$

Note that $\bar{\mathscr{G}}$ differs from $\mathscr{G}$ from the fact that the set of probability distributions $\{P_i\}_{i=1}^n$ is replaced by $\bar{P}$, which is supposed to be known by all players. The games $\mathscr{G}$ and $\bar{\mathscr{G}}$ are called Bayes-equivalent because $P_i(\hat{\lambda}_{-i}|\hat{\lambda}_i) = \bar{P}(\hat{\lambda}_{-i}|\hat{\lambda}_i)$, where $\bar{P}(\hat{\lambda}_{-i}|\hat{\lambda}_i)$ can be obtained from $\bar{P}(\hat{\lambda})$ using the Bayes' rule:

$$
\bar{P}(\hat{\lambda}_{-i}|\hat{\lambda}_i) = \frac{\bar{P}(\hat{\lambda})}{\sum_{L_{-i}} \bar{P}(\hat{\lambda})}. \tag{16}
$$

It is important to mention that in the Bayesian game instead of minimizing the long-run average cost $g_i(S_i, c_i, s_i)$ as in (12), we aim to minimize the conditional expected long-run average cost. Let us denote the conditional expected long-run average cost by $\mathbb{E}[g_i(S_i, c_i, s_i)|\hat{\lambda}_i]$, where

$$
\mathbb{E}[g_i(S_i, c_i, s_i)|\hat{\lambda}_i] = \sum_{\hat{\lambda}_{-i} \in L_{-i}} \bar{P}(\hat{\lambda}_{-i}|\hat{\lambda}_i) g_i(S_i, c_i, s_i). \tag{17}
$$

Let u_{-i}^* be the set of optimal strategies of all players except player i. In the rest of this paper, we are going to make use of the following definitions.

Definition 1 A *best-response* set in a Bayesian game is the set of all strategies that minimize the expected long-run average cost:

$$\mathscr{B}_i(u_{-i}) := \{u_i^* | \min_{c_i \in \mathbb{N}} \sum_{\hat{\lambda}_{-i} \in L_{-i}} \bar{P}(\hat{\lambda}_{-i}|\hat{\lambda}_i) g_i(S_i, c_i, s_i)\}. \tag{18}$$

The can-order level c_i takes values between the upper threshold S_i and the lower threshold s_i. Therefore, we have a finite set of options for the values of the can-order level. As a result, there is always an optimal strategy u^* in the best-response set $\mathscr{B}_i(u_{-i})$.

For compactness, let us denote the optimal strategy $u_i^* = (S_i, c_i^*, s_i)$. This strategy corresponds to the optimal can-order level c_i^* that minimizes the expected long-run average cost.

Definition 2 The strategy u_i^* is at *Bayesian Nash equilibrium*, if it holds for all item $i \in N$

$$\sum_{\hat{\lambda}_{-i} \in L_{-i}} \bar{P}(\hat{\lambda}_{-i}|\hat{\lambda}_i) g_i(u_i^*, u_{-i}^*) \leq \sum_{\hat{\lambda}_{-i} \in L_{-i}} \bar{P}(\hat{\lambda}_{-i}|\hat{\lambda}_i) g_i(u_i^*, u_{-i}). \tag{19}$$

Note that at a Nash equilibrium all players play a best-response, such that $u_i^* \in \mathscr{B}_i(u_{-i}^*)$ for all $i \in N$, and no player benefits from changing its reorder strategy.

The following result was borrowed from Harsanyi [15], and it states that if $u^* = (u_1^*, \ldots, u_n^*)$ is a Nash equilibrium strategy of the complete game $\bar{\mathscr{G}}$, then it is also a Bayesian equilibrium strategy for the incomplete game $\mathscr{G}$.

Theorem 1 *Let the incomplete game* (14) *and the Bayesian game* (15) *be given, such that* (15) *is Bayes-equivalent to* (14). *In order to any given n-tuple of strategies* $u = (u_1, ..., u_n)$ *be a Bayesian equilibrium point in the game* (14), *it is both sufficient and necessary that in the normal form of the game* (15) *this n-tuple u be an equilibrium point in Nash's sense.*

However, defining a prior probability distribution $\bar{P}$ that is known by all players can be very complicated. In this work we assume that the players do not know the exact prior probability distribution $\bar{P}$, but each player knows the set of possible probability distributions $\bar{P}(\hat{\lambda}_{-i}|\hat{\lambda}_i)$, denoted by $\mathscr{P}$. In order to be protected against the uncertainty of not knowing the real probability distribution we reframe our approach within the context of *distributionally robust optimization*. Distributionally robust optimization emerges from the idea of being protected against the ambiguity in the probability distribution by taking the worst-case scenario, which involves finding the probability distribution that maximizes the long-run average cost. As a result, we can formulate the following robust optimization problem:

$$\min_{c_i \in \mathbb{N}} \max_{\bar{P}(\hat{\lambda}_{-i}|\hat{\lambda}_i) \in \mathscr{P}} \mathbb{E}[g_i(S_i, c_i, s_i)|\hat{\lambda}_i] = \min_{c_i \in \mathbb{N}} \max_{\bar{P}(\hat{\lambda}_{-i}|\hat{\lambda}_i) \in \mathscr{P}} \sum_{\hat{\lambda}_{-i} \in L_{-i}} \bar{P}(\hat{\lambda}_{-i}|\hat{\lambda}_i) g_i(S_i, c_i, s_i). \tag{20}$$

Each player $i \in N$ can define its best strategy u_i^* by finding the optimal solution of (20) denoted by the point $(c_i^*, \bar{P}^*(\hat{\lambda}_{-i}|\hat{\lambda}_i))$. Observe that for fixed values of the upper and lower thresholds S_i and s_i, respectively, determining the optimal strategy that characterizes the control variable u_i consists first in finding the optimal probability distribution $\bar{P}(\hat{\lambda}_{-i}|\hat{\lambda}_i)$ that maximizes the long-run average cost $g_i(S_i, c_i, s_i)$ for each possible can-order threshold c_i. Then, we can find an optimal can-order threshold c_i that minimizes the expected long-run average cost $\mathbb{E}[g_i(S_i, c_i, s_i)|\hat{\lambda}_i]$ for all $\bar{P}(\hat{\lambda}_{-i}|\hat{\lambda}_i) \in \mathscr{P}$.

To find the optimal solution, it is important to define the correct set of possible probability distributions $\mathscr{P}$. We assume that $\mathscr{P}$ is a set of multi-variate Poisson probability distributions.

Let us denote by $MPoisson(\bar{\eta}, \hat{\lambda}_{-i})$ the probability distribution of a multi-variate Poisson random variable, where $\bar{\eta} \in \mathbb{R}_+^{1 \times (n-1)}$ represents the vector of rate parameters of each random variable $\hat{\lambda}_j \in \hat{\lambda}_{-i}$, $j \neq i$. Now we can formally define the set $\mathscr{P}$ as follows:

$$\mathscr{P} = \{MPoisson(\bar{\eta}, \hat{\lambda}_{-i}) | \bar{\eta} = (\eta_1, \eta_2, \dots, \eta_{i-1}, \eta_{i+1}, \dots, \eta_n) \in \mathbb{R}_+^{1 \times (n-1)}, j \neq i\}. \tag{21}$$

We assume that the probability distributions in the set $\mathscr{P}$ satisfy the Bayesian-equivalence $P_i(\hat{\lambda}_{-i}|\hat{\lambda}_i) = \bar{P}(\hat{\lambda}_{-i}|\hat{\lambda}_i)$. Under this assumption, we ensure that Theorem 1 is still satisfied. In addition, we assume that from the perspective of player i the rest of the players are homogeneous, namely the values of the parameters $\eta_j \in \mathbb{R}_+$ are the same for all player $j \neq i$. Let us denote this general parameter by $\eta \in \mathbb{R}_+$. For the sake of simplicity and without loss of generality, the random variables $\hat{\lambda}_j$, $j \in N$, are independent and identically distributed with parameter η. We have to solve the robust optimization problem to obtain a single optimal value $\eta^* \in \mathbb{R}_+$ for the multi-variate Poisson distribution. Based on these assumptions, the expected long-run average cost is explicitly computed as follows:

$$\mathbb{E}[g_i(\hat{\lambda}_{-i})|\hat{\lambda}_i] = \sum_{\hat{\lambda}_{-i} \in L_{-i}} exp(-(n-1)\eta) \frac{\eta^{\sum_{\hat{\lambda}_j \in \hat{\lambda}_{-i}} \hat{\lambda}_j}}{\prod_{\hat{\lambda}_j \in \hat{\lambda}_{-i}} \hat{\lambda}_j!} g_i(\hat{\lambda}_{-i}). \tag{22}$$

Note that in (22) we take the thresholds S_i, c_i, and s_i as fixed and for each term in the sum we compute the long-run average cost as a function of $\hat{\lambda}_{-i}$.

4 Main Results

In this section, we show that the expected long-run average cost $\mathbb{E}[g_i(S_i, c_i, s_i)|\hat{\lambda}_i]$ is convex on the can-order threshold c_i. Furthermore, we prove that finding an optimal threshold c_i corresponds to finding a Bayesian Nash equilibrium strategy.

To prove the convexity of the expected long-run average cost with respect to the can-order threshold c_i, first we are going to analyse the behaviour of the long-run average cost $g_i(S_i, c_i, s_i)$ for any possible vector $\hat{\lambda}_{-i}$.

Let us define

$$\omega_i = \frac{(S_i - c_i - 1)(1 - \theta)}{\sum_{x=1}^{c_i+1-s_i} \theta^x + S_i - c_i - 1}. \tag{23}$$

It is important to observe that ω_i is a positive fraction, namely $\omega_i \in (0, 1)$. Note that ω_i increases when the can-order level decreases.

Also let the following function be given

$$f_i(c_i, S_i) := \frac{\hat{\lambda}_i \mathbb{E}(a_i) + \sum_{x=c_i+1}^{S_i} G(x)}{S_i - c_i}. \tag{24}$$

The above is related to the cost of replenishing always with discount opportunity when the inventory reaches the can-order level.

Lemma 2 *For ω_i as in* (23)*, the following equation holds:*

$$g_i(S_i, c_i + 1, s_i) = \omega_i f_i(c_i + 1, S_i) + (1 - \omega_i)g_i(S_i, c_i, s_i). \tag{25}$$

Proof Let $\theta = \hat{\lambda}_i/(\hat{\lambda}_i + \hat{\mu}_i)$. Since the inventory level is at S_i and c_i must be less than S_i, for $T_i(y)$ we have

$$T_i(S_i) = \frac{1}{\hat{\lambda}_i + \hat{\mu}_i} \sum_{x=0}^{(c_i-s_i-1)} \theta^x + \frac{1}{\hat{\lambda}_i}(S_i - c_i). \tag{26}$$

$I_i(y)$ at an inventory level of S_i is given by

$$I_i(S_i) = \frac{1}{\hat{\lambda}_i + \hat{\mu}_i} \sum_{x=0}^{(c_i-s_i-1)} \theta^x G_i(S_i - x) + \frac{1}{\hat{\lambda}_i} \sum_{x=c_i+1}^{S_i} G_i(x), \tag{27}$$

where $S_i > c_i$.

Q_i at inventory S_i can be written as:

$$Q_i(S_i) = Q_i(y-1), \tag{28}$$

where $S_i > c_i$.

Let $\mathbf{N}_{g_i}$ and $\mathbf{D}_{g_i}$ be the numerator and denominator of the cost function in (12). From (26) we have that

$$\hat{\lambda}_i \mathbf{D}_{g_i} = \hat{\lambda}_i T_i(S_i) = \frac{\hat{\lambda}_i}{\hat{\lambda}_i + \hat{\mu}_i} \sum_{x=0}^{(c_i-s_i-1)} \theta^x + S_i - c_i$$

$$= \sum_{x=1}^{c_i-s_i} \theta^x + S_i - c_i.$$

From (27), we also have

$$\hat{\lambda}_i \mathbf{N}_{g_i} = \hat{\lambda}_i \mathbb{E}(a_i) + \hat{\lambda}_i \left(\mathbb{E}(A) - \mathbb{E}(a_i)\right) \theta^{c_i-s_i}$$

$$+ \sum_{x=1}^{c_i-s_i} \theta^x G(c_i - x + 1) + \sum_{x=c_i+1}^{S_i} G(x).$$

The variables, T_i, I_i and Q_i can be substituted into the objective function in (12), and this yields

$$g_i(S_i, c_i, s_i) = \frac{\hat{\lambda}_i \mathbf{N}_{g_i}}{\hat{\lambda}_i \mathbf{D}_{g_i}} = \frac{\hat{\lambda}_i \mathbb{E}(a_i) + \hat{\lambda}_i \left(\mathbb{E}(A) - \mathbb{E}(a_i)\right) \theta^{c_i-s_i}}{\sum_{x=1}^{c_i-s_i} \theta^x + S_i - c_i}$$

$$+ \frac{\left(\sum_{x=1}^{c_i-s_i} \theta^x G(c_i - x + 1)\right) + \sum_{x=c_i+1}^{S_i}(G(x))}{\sum_{x=1}^{c_i-s_i} \theta^x + S_i - c_i}. \tag{29}$$

Let us now substitute $c_i + 1$ in (29), which yields:

$$g_i(S_i, c_i + 1, s_i) = \frac{\hat{\lambda}_i \mathbb{E}(a_i) + \hat{\lambda}_i \left(\mathbb{E}(A) - \mathbb{E}(a_i)\right) \theta^{c_i+1-s_i}}{\sum_{x=1}^{c_i+1-s_i} \theta^x + S_i - c_i}$$

$$+ \frac{\sum_{x=1}^{c_i+1-s_i} \theta^x G(c_i - x + 2) + \sum_{x=c_i+2}^{S_i} G(x)}{\sum_{x=1}^{c_i+1-s_i} \theta^x + S_i - c_i}. \tag{30}$$

Note that for the last part of the numerator it holds

$$
\sum_{x=1}^{c_i+1-s_i} \theta^x G(c_i - x + 2) + \sum_{x=c_i+2}^{S_i} G(x)
$$

$$
= \sum_{x=2}^{c_i+1-s_i} \theta^x G(c_i - x + 2) + \theta \sum_{x=c_i+1}^{S_i} G(x) + (1-\theta) \sum_{x=c_i+2}^{S_i} G(x).
$$

Then, (30) can be rewritten as

$$
g_i(S_i, c_i + 1, s_i) = \frac{\hat{\lambda}_i \mathbb{E}(a_i) + \hat{\lambda}_i \left(\mathbb{E}(A) - \mathbb{E}(a_i)\right)\theta^{c_i+1-s_i}}{\sum_{x=1}^{c_i+1-s_i} \theta^x + S_i - c_i}
$$

$$
+ \frac{\sum_{x=2}^{c_i+1-s_i} (G(c_i - x + 2)) + \theta \sum_{x=c_i+1}^{S_i} G(x) + (1-\theta)\sum_{x=c_i+2}^{S_i} G(x)}{\sum_{x=1}^{c_i+1-s_i} \theta^x + S_i - c_i}.
$$

$$(31)$$

We also have

$$
(\mathbf{D}_{\omega_i}) \cdot (1 - \omega_i) g(S_i, c_i, s_i)
$$

$$
= \sum_{x=1}^{c_i+1-s_i} \left(\theta^x + S_i - c_i\right) g_i(S_i, c_i, s_i)\theta
$$

$$
= \hat{\lambda}_i \mathbb{E}(a_i)\theta + \hat{\lambda}_i \left(\mathbb{E}(A) - \mathbb{E}(a_i)\right)\theta^{c_i+1-s_i}
$$

$$
+ \theta \sum_{x=1}^{c_i-s_i} \theta^x G(c_i - x + 1) + \theta \sum_{x=c_i+1}^{S_i} G(x)
$$

$$
= \hat{\lambda}_i \mathbb{E}(a_i)\theta + \hat{\lambda}_i \left(\mathbb{E}(A) - \mathbb{E}(a_i)\right)\theta^{c_i+1-s_i}
$$

$$
+ \sum_{x=2}^{c_i+1-s_i} \theta^x G(c_i - x + 2) + \theta \sum_{x=c_i+1}^{S_i} G(x).
$$

$$(32)$$

Likewise, from (24) we have:

$$
(\mathbf{D}_{\omega_i}) \cdot \omega_i f_i(c_i + 1, S_i) = (S_i - c_i - 1)(1 - \theta) f_i(c_i + 1, S_i)
$$

$$
= (1 - \theta)\hat{\lambda}_i \mathbb{E}(a_i) + (1 - \theta) \sum_{x=c_i+2}^{S_i} G(x).
$$

$$(33)$$

From (31)–(33), we can conclude that $g_i(S_i, c_i + 1, s_i) = \omega_i f_i(c_i + 1, S_i) + (1 - \omega_i)g_i(S_i, c_i, s_i)$ and (25) is proved. □

A direct consequence of Lemma 2 is that $g_i(c_i + 1)$ is in the convex hull between $g_i(c_i)$ and $f_i(c_i + 1)$.

In the next result, we show that the optimal value of c_i is located at a local minimum of g_i for any possible vector $\hat{\lambda}_{-i}$, under continuous relaxation for c_i, namely $c_i \in \mathbb{R}_+$. Let the discrete difference operator be given $\frac{\partial}{\partial c_i}$, which is defined as

$$
\frac{\partial}{\partial c_i} g_i(S_i, c_i, s_i) := g_i(S_i, c_i + 1, s_i) - g_i(S_i, c_i, s_i).
$$

Lemma 3 *Let the can-order threshold be $c_i \in \mathbb{R}$ and in the interval $[s_i, S_i]$. Then, at a local minimum it must hold*

$$f_i(c_i + 1, S_i) - g_i(S_i, c_i, s_i) \geq 0. \tag{34}$$

Proof First, we obtain the derivative of $g_i(\cdot)$ as

$$\frac{\partial}{\partial c_i} g_i(S_i, c_i, s_i) = \omega_i \left(f_i(c_i + 1, S_i) - g_i(S_i, c_i, s_i) \right). \tag{35}$$

At a local minimum it must hold $\frac{\partial g_i(S_i, c_i, s_i)}{\partial c_i} \geq 0$. As ω_i is only zero when there is no demand, the above implies $f_i(c_i + 1, S_i) - g_i(S_i, c_i, s_i) \geq 0$ and (34) is proven. $\qquad \square$

In the next result, we establish the convexity of function $g_i(S_i, c_i, s_i)$ in the decision variable c_i for any possible values of $\hat{\lambda}_{-i}$.

Lemma 4 *The long-run average cost function is convex with respect to c_i, and the following inequality holds:*

$$g_i(S_i, \nu c_i + (1 - \nu)(c_i + x), s_i) \leq \nu g_i(S_i, c_i, s_i) + (1 - \nu) g_i(S_i, c_i + x, s_i), \tag{36}$$

where $\nu \in [0, 1]$ and $g_i(\cdot)$ is defined on positive real numbers.

Proof First note that $g_i(\cdot)$ is defined on positive real numbers as the costs are nonnegative. Now, take a value for c_i which is a local minimum for $g_i(\cdot)$ and denote it $\bar{c}_i$. Then the value of $g_i(S_i, c_i, s_i)$ on the interval range $H(\bar{c}_i, \epsilon) = [\bar{c}_i - \epsilon, \bar{c}_i + \epsilon]$ should be greater than or equal to $g_i(S_i, \bar{c}_i, s_i)$. This is the case, if and only if the second derivative of $g_i(\cdot)$ is greater than or equal to zero. From the first derivative in (35), we obtain

$$\frac{\partial^2 g_i(S_i, c_i, s_i)}{\partial c_i^2} = \frac{\partial \omega_i}{\partial c_i} f_i(c_i + 1, S_i) + \omega_i \frac{\partial f_i(c_i + 1, S_i)}{\partial c_i}$$
$$- \frac{\partial \omega_i}{\partial c_i} g_i(S_i, c_i, s_i) - \omega_i \frac{\partial g_i(S_i, c_i, s_i)}{\partial c_i}, \tag{37}$$

where ω_i is as in (23).

To simplify the expression of (37), the result of (35) is substituted in (37), which yields

$$\frac{\partial^2}{\partial c_i^2} g_i(S_i, c_i, s_i) = \omega_i \frac{\partial f_i(c_i + 1, S_i)}{\partial c_i} + \left(\frac{\partial \omega_i}{\partial c_i} - \omega_i^2 \right) (f_i(c_i + 1, S_i) - g_i(S_i, c_i, s_i))$$
$$= \omega_i \frac{\partial f_i(c_i + 1, S_i)}{\partial c_i} + \left(\frac{\partial \omega_i}{\partial c_i} - \omega_i^2 \right) \frac{1}{\omega_i} \underbrace{\frac{\partial}{\partial c_i} g_i(S_i, c_i, s_i)}_{\approx 0}. \tag{38}$$

Note that $\omega_i \geq 0$, $\omega_i^2 \geq 0$, $g_i(S_i, c_i, s_i) \geq 0$ and $f_i(c_i + 1, S_i) \geq 0$. We wish to show that $f_i(c_i + 1, S_i)$ is an increasing function, namely that $(\partial f_i(c_i + 1, S_i))/(\partial c_i)$ is positive. This holds true if $f_i(c_i + 2, S_i) \geq f_i(c_i + 1, S_i)$ is proven. If the previous inequality is true, then we have

$$\frac{\hat{\lambda}_i \mathbb{E}(a_i) + \sum_{x=c_i+3}^{S_i} G(x)}{S_i - c_i - 2} \geq \frac{\hat{\lambda}_i \mathbb{E}(a_i) + \sum_{x=c_i+2}^{S_i} G(x)}{S_i - c_i - 1}.$$

Furthermore, the above can be rewritten as

$$\left(\hat{\lambda}_i \mathbb{E}(a_i) + \sum_{x=c_i+3}^{S_i} G(x) \right) (S_i - c_i - 1) \ge \left(\hat{\lambda}_i \mathbb{E}(a_i) + \sum_{x=c_i+2}^{S_i} G(x) \right) (S_i - c_i - 2).$$

(39)

The previous condition yields

$$S_i \hat{\lambda}_i \mathbb{E}(a_i) - c_i \hat{\lambda}_i \mathbb{E}(a_i) - \hat{\lambda}_i \mathbb{E}(a_i)$$

$$+ S_i \sum_{x=c_i+3}^{S_i} G(x) - c_i \sum_{x=c_i+3}^{S_i} G(x) - \sum_{x=c_i+3}^{S_i} G(x)$$

$$\ge S_i \hat{\lambda}_i \mathbb{E}(a_i) - c_i \hat{\lambda}_i \mathbb{E}(a_i) - 2 \hat{\lambda}_i \mathbb{E}(a_i)$$

$$+ S_i \sum_{x=c_i+2}^{S_i} G(x) - c_i \sum_{x=c_i+2}^{S_i} G(x) - 2 \sum_{x=c_i+2}^{S_i} G(x),$$

which can be simplified as

$$S_i G(c_i + 2) - c_i G(c_i + 2) - 2G(c_i + 2) - \sum_{x=c_i+3}^{S_i} G(x) - \hat{\lambda}_i \mathbb{E}(a_i)$$

(40)

$$= S_i(c_i + 2)h_i - c_i(c_i + 2)h_i - 2(c_i + 2)h_i - \sum_{x=c_i+3}^{S_i} G(x) - \hat{\lambda}_i \mathbb{E}(a_i) \le 0.$$

Let us rewrite $\sum_{x=c_i+3}^{S_i} G(x) = \frac{1}{2} h_i (S_i - c_i - 2)(c_i + S_i + 3)$, then we have

$$\hat{\lambda}_i \mathbb{E}(a_i) \ge h_i (S_i - c_i - 2) \left(\frac{c_i}{2} + \frac{1}{2} - \frac{S_i}{2} \right),$$

(41)

where $h_i(S_i - c_i - 2) \ge 0$ and $(\frac{c_i}{2} + \frac{1}{2} - \frac{S_i}{2}) \le 0$, if $S_i \ge c_i + 1$ which holds true. Since $\mathbb{E}(a_i) > 0$ and $\hat{\lambda}_i > 0$ hold true, then (41) is satisfied and we can conclude that $f_i(c_i + 2, S_i) \ge f_i(c_i + 1, S_i)$ holds.

A direct implication of the above is that $(\partial f_i(c_i + 1, S_i))/(\partial c_i)$ is positive. This means that the first term in (38) for $\frac{\partial^2 g_i(S_i, c_i, s_i)}{\partial c_i^2}$ is greater than or equal to zero, namely:

$$\omega_i \frac{\partial f_i(c_i + 1, S_i)}{\partial c_i} \ge 0.$$

(42)

In order to prove convexity, the rest of $\frac{\partial^2 g_i(S_i, c_i, s_i)}{\partial c_i^2}$ must be positive or smaller than (42). Since from (35), we know that $\partial g_i(\cdot)/\partial c_i$ is proportional to $f_i(c_i + 1, S_i) - g_i(S_i, c_i, s_i)$ which becomes zero in a local minimum. Then in the neighbourhood of a stationary point $f_i(c_i + 1, S_i) - g_i(S_i, c_i, s_i)$ is relatively small. This makes the second term in (38) approximately equal to zero and therefore negligible. We can conclude that for any stationary point the second derivative is nonnegative and therefore function $g_i(\cdot)$ is convex on c_i. $\qquad\square$

Lemma 4 proves that $g_i(S_i, c_i, s_i)$ is convex on c_i; therefore, it is known that the optimal can-order level is at the same time a local and global minimum.

In the following result, we prove the convexity of the function $\mathbb{E}[g_i(S_i, c_i, s_i)|\hat{\lambda}_i]$ with respect to the decision variable c_i, for a given parameter η that characterizes the probability $\bar{P}(\hat{\lambda}_{-i}|\hat{\lambda}_i)$.

Theorem 2 *The expected long-run average cost function is convex with respect to c_i, and the following inequality holds:*

$$\mathbb{E}[g_i(S_i, \nu c_i + (1 - \nu)(c_i + x), s_i)|\hat{\lambda}_i] \leq \nu\mathbb{E}[g_i(S_i, c_i, s_i)|\hat{\lambda}_i] + (1 - \nu)\mathbb{E}[g_i(S_i, (c_i + x), s_i)|\hat{\lambda}_i],$$
(43)

where $\nu \in (0, 1)$ and $g_i(\cdot)$ is defined on positive real numbers.

Proof Similarly as in Lemma 4 we have to prove that for any c_i in the interval $[\bar{c}_i - \epsilon, \bar{c}_i + \epsilon]$, where $\bar{c}_i$ is the local minimum of $\mathbb{E}[g_i(\cdot)|\hat{\lambda}_i]$, the value of $\mathbb{E}[g_i(S_i, c_i, s_i)|\hat{\lambda}_i]$ is greater than or equal to $\mathbb{E}[g_i(S_i, \bar{c}_i, s_i)|\hat{\lambda}_i]$. This is satisfied when the second derivative of $\mathbb{E}[g_i(S_i, c_i, s_i)|\hat{\lambda}_i]$ is greater than or equal to zero.

The first derivative of $\mathbb{E}[g_i(S_i, c_i, s_i)|\hat{\lambda}_i]$ with respect to c_i is

$$\frac{\partial \mathbb{E}[g_i(S_i, c_i, s_i)|\hat{\lambda}_i]}{\partial c_i} = \frac{\partial}{\partial c_i} \sum_{\hat{\lambda}_{-i} \in L_{-i}} exp(-(n - 1)\eta) \frac{\eta^{\sum_{\hat{\lambda}_j \in \hat{\lambda}_{-i}} \hat{\lambda}_j}}{\prod_{\hat{\lambda}_j \in \hat{\lambda}_{-i}} \hat{\lambda}_j!} g_i(S_i, c_i, s_i)$$

$$= \sum_{\hat{\lambda}_{-i} \in L_{-i}} exp(-(n - 1)\eta) \frac{\eta^{\sum_{\hat{\lambda}_j \in \hat{\lambda}_{-i}} \hat{\lambda}_j}}{\prod_{\hat{\lambda}_j \in \hat{\lambda}_{-i}} \hat{\lambda}_j!} \frac{\partial}{\partial c_i} g_i(S_i, c_i, s_i).$$

Consequently, the second derivative is given by

$$\frac{\partial^2 \mathbb{E}[g_i(S_i, c_i, s_i)|\hat{\lambda}_i]}{\partial c_i^2} = \sum_{\hat{\lambda}_{-i} \in L_{-i}} exp(-(n - 1)\eta) \frac{\eta^{\sum_{\hat{\lambda}_j \in \hat{\lambda}_{-i}} \hat{\lambda}_j}}{\prod_{\hat{\lambda}_j \in \hat{\lambda}_{-i}} \hat{\lambda}_j!} \frac{\partial^2}{\partial c_i^2} g_i(S_i, c_i, s_i). \quad (44)$$

Based on Lemma 4, we know that $\frac{\partial^2}{\partial c_i^2} g_i(S_i, c_i, s_i) \geq 0$, and then each term in (44) is greater than or equal to zero. We can conclude that for any stationary point and for any parameter η, the second derivative is nonnegative and the function $\mathbb{E}[g_i(S_i, c_i, s_i)|\hat{\lambda}_i]$ is convex with respect to c_i. □

The set $\mathscr{P}$ defined by (21) has to be determined in such a way that the expected long-run average cost function is concave with respect to the parameter η. In the following lemma, we prove that such a set $\mathscr{P}$ exists.

Lemma 5 *Let $\eta^* \in \mathbb{R}_+$ be an optimal value that maximizes (22). There exists ε, $0 < \varepsilon < \eta^*$, such that in the interval $[\eta^* - \varepsilon, \eta^* + \varepsilon]$, the expected long-run average cost is concave with respect to η.*

Proof To prove the concavity of the expected long-run average cost with respect to η, we use the second derivative criteria. The first derivative is computed as follows:

$$\frac{\partial \mathbb{E}[g_i(S_i, c_i, s_i)|\hat{\lambda}_i]}{\partial \eta}$$

$$= \sum_{\hat{\lambda}_{-i} \in L_{-i}} g_i(S_i, c_i, s_i) exp(-(n - 1)\eta) \eta^{\sum_{\hat{\lambda}_j \in \hat{\lambda}_{-i}} \hat{\lambda}_j - 1} \left[\frac{\sum_{\hat{\lambda}_j \in \hat{\lambda}_{-i}} \hat{\lambda}_j - (n - 1)\eta}{\prod_{\hat{\lambda}_j \in \hat{\lambda}_{-i}} \hat{\lambda}_j!} \right].$$

By computing the second derivative and evaluating it in $\eta = \eta^*$, we obtain:

$$\frac{\partial^2 \mathbb{E}[g_i(S_i, c_i, s_i)|\hat{\lambda}_i](\eta^*)}{\partial \eta^2}$$

$$= \frac{1}{\eta^*} \sum_{\hat{\lambda}_{-i} \in L_{-i}} \frac{g_i(S_i, c_i, s_i) exp(-(n-1)\eta^*)\eta^{*\sum_{\hat{\lambda}_j \in \hat{\lambda}_{-i}} \hat{\lambda}_j - 1}}{\prod_{\hat{\lambda}_j \in \hat{\lambda}_{-i}} \hat{\lambda}_j!} \left[\left(\sum_{\hat{\lambda}_j \in \hat{\lambda}_{-i}} \hat{\lambda}_j - (n-1)\eta^* \right)^2 \right.$$

$$\left. - \left(\sum_{\hat{\lambda}_j \in \hat{\lambda}_{-i}} \hat{\lambda}_j - (n-1)\eta^* \right) - \eta^*(n-1) \right].$$

Observe that as η^* is the optimal value that maximizes the expected long-run average cost, the first two terms in the previous equation are equal to zero. Then, we have

$$\frac{\partial^2 \mathbb{E}[g_i(S_i, c_i, s_i)|\hat{\lambda}_i](\eta^*)}{\partial \eta^2} = -(n-1) \sum_{\hat{\lambda}_{-i} \in L_{-i}} \frac{g_i(S_i, c_i, s_i) exp(-(n-1)\eta^*)\eta^{*\sum_{\hat{\lambda}_j \in \hat{\lambda}_{-i}} \hat{\lambda}_j - 1}}{\prod_{\hat{\lambda}_j \in \hat{\lambda}_{-i}} \hat{\lambda}_j!}.$$

From the previous equation, we have that $\frac{\partial^2 \mathbb{E}[g_i(S_i,c_i,s_i)|\hat{\lambda}_i](\eta^*)}{\partial \eta^2} < 0$. Hence η^* is a local maximum. We can conclude that there exists an interval $[\eta^* - \varepsilon, \eta^* + \varepsilon]$ such that the expected long-run average cost is concave. $\qquad\square$

Remark 1 Note that as $c_i \in [S_i, s_i]$ and $\eta \in [\eta^* - \varepsilon, \eta^* + \varepsilon]$, both variables are contained in convex and compact sets. In addition, based on Theorem 2 and Lemma 5 we know that the expected long-run average cost is convex with respect to c_i and concave with respect to η, on their respective intervals. Therefore, according to John von Neumann's minimax theorem:

$$\min_{c_i \in \mathbb{N}} \max_{\bar{P}(\hat{\lambda}_{-i}|\hat{\lambda}_i) \in \mathscr{P}} \mathbb{E}[g_i(S_i, c_i, s_i)|\hat{\lambda}_i] = \max_{\bar{P}(\hat{\lambda}_{-i}|\hat{\lambda}_i) \in \mathscr{P}} \min_{c_i \in \mathbb{N}} \mathbb{E}[g_i(S_i, c_i, s_i)|\hat{\lambda}_i]. \qquad (45)$$

In the following result, we prove that finding an optimal can-order level c_i that minimizes (17) corresponds to finding a strategy at a Bayesian Nash equilibrium.

Theorem 3 *Let for all players* $i \in N$ *the upper threshold* S_i *and the lower threshold* s_i *be given. In addition for a given parameter* η *that maximizes the conditional expected long-run average cost, let the optimal can-order threshold* c_i^* *be obtained by:*

$$c_i^* := arg \min_{c_i \in \mathbb{N}} \sum_{\hat{\lambda}_{-i} \in L_{-i}} \bar{P}(\hat{\lambda}_{-i}|\hat{\lambda}_i) g_i(S_i, c_i, s_i). \qquad (46)$$

Then, the n-tuple $(c_1^*, c_2^*, \ldots, c_n^*)$ *is at a Bayesian Nash equilibrium of the Bayesian game* $\bar{\mathscr{G}}$.

Proof Based on Theorem 2, we know that the expected long-run average cost is convex on the can-order level c_i. Moreover, the optimal value c_i^* is a global minimum. Therefore, the can-order level c_i^* that minimizes the long-run average cost is unique.

By definition of best-response set in a Bayesian game $\mathscr{B}_i(u_{-i}^*)$, if we find the can-order level that minimizes $\mathbb{E}[g_i(S_i, c_i, s_i)|\hat{\lambda}_i]$ for a player $i \in N$, given the optimal strategy of the other players, then the optimal strategy of player i, u_i^* is in the set $\mathscr{B}(u_{-i}^*)$ defined by (18). Hence, the strategy is at a Bayesian Nash equilibrium. $\qquad\square$

Table 1 Setup cost and demand parameters

t	1	2	3	4	5	6	7	8	9
a_i	12	14	13	15	18	16	14	17	18
A	78	80	84	82	79	82	84	82	83
λ_i	10	10	5	4	7	8	5	5	10

	10	11	12	13	14	15	16	17	18	19
	20	17	18	21	23	20	25	23	24	24
	79	77	78	80	78	82	80	84	87	85
	8	9	8	11	12	13	16	16	15	17

	20	21	22	23	24	25	26	27	28	29
	19	18	17	16	17	18	20	21	23	24
	82	80	82	84	82	83	81	81	82	80
	6	9	10	7	7	13	11	13	12	16

	30	31	32	33	34	35	36	37	38	39
	23	20	21	20	20	21	21	18	16	12
	79	83	80	84	78	80	80	79	78	79
	14	11	11	7	13	11	12	12	10	11

5 Numerical Analysis

Let us consider a time interval of 39 months, where each month consists of 4 weeks. The simulated results are obtained by optimizing the can-order threshold for the minimum expected long-run average cost at every time interval. Consider a system of $N = 4$ items which are ordered at a single distribution centre. At every time period, we forecast the demand and setup costs of a single item i to obtain the optimal can-order threshold of the (S, c, s) policy.

The parameters are selected as follows. The demand is measured over 39 consecutive months. Each month consists of four weeks. The total number of weeks included in the simulations is 156. The lead time is 1 week. The holding costs h_i are €2 per month. We consider 4 items involved in the coordinated replenishment. The major setup cost $\hat{A}(1)$ is initially forecast to be €80. The minor setup cost $\hat{a}_i(1)$ is initially forecast to be €15. The safety factor k is 3. The forecast coefficients α, γ are both 0.2. The initial demand is forecast considering observed data for the first 12 months applying a Poisson regression model and every month we add four more new observations (one for each week) to forecast the demand for the next month. Finally, it is assumed that all items in the system face the same demand uncertainty for the sake of simplicity and without the loss of generality. Note that $S_i(t)$ and $s_i(t)$ can therefore be represented by $S(t)$ and $s(t)$, respectively, as the levels are assumed to be the same for all items. The demand is assumed to follow a Poisson distribution with mean equal to λ_i and the setup cost are assumed to follow a Gaussian distribution with mean equal to a_i and A, see Table 1.

Since it is a forecasting model of the can-order policy, the demands and setup costs for all weeks are forecast. Figure 5 displays the forecast $\hat{D}_i(t)$ and actual demand $D_i(t)$ over a 39-month period. Similar to the demand, the setup costs influence the control variables of the can-order policy and must therefore be redetermined at every time period.
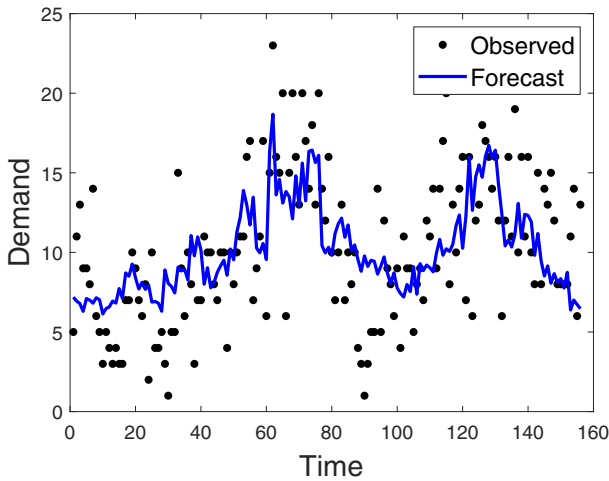
Fig. 5 Forecast demand $\hat{D}_i(t)$ (dots) and actual demand D_i (solid) over a 39-month period

In Algorithm 1, we present the pseudo-code used to obtain the optimal can-order level at each month. Each month, the events are forecast again and the control variables of the can-order policy are redetermined. The optimal can-order level is determined by increasing the value of c_i by one and calculating the corresponding expected long-run average cost for each probability distribution in the set $\mathscr{P}$ as in equation (21). If the expected long-run average cost calculated at $(c_i + 1)$ is higher than it was at the previous can-order level, the iterations stop and the value of the previous can-order level is the optimal value of c_i. As it is stated in Theorem 2, the expected long-run average cost is convex with respect to c_i. The optimal can-order level is at the same time a local and a global minimum. As a result of the behaviour of the long-run average cost and the iterative process of the algorithm, it is possible to say that it converges to the optimal can-order level c_i^* that minimizes the expected cost in the long run.

It is important to mention that for the Poisson regression model, we use the built-in MATLAB function *glmfit*. This function uses a method called weighted least squares to estimate the regression parameters $\hat{\beta}_1$ and $\hat{\beta}_2$. Weighted least squares is a generalization of generalized least squares. It takes into account the weights that determine how much each value influences the final parameter estimates. For the convergence of the weighted least squares MATLAB uses a Newton–Raphson iterative method capturing the first-order optimality conditions at the solution, resulting in strong local convergence rates.

The simulations of the expected long-run average cost are depicted in Fig. 6. The graph corroborates the results in Theorem 2 and Lemma 5, as it shows that the expected long-run average cost is a convex function on c_i and concave on η. For all consecutive time periods, the optimal can-order levels are listed in Table 2.

The optimal can-order levels which are determined in the previous step of the algorithm are used to simulate the inventory reordering policy and the related inventory evolution over time. Figure 7 shows that reordering also takes place at the can-order level. In addition, to analyse the stability of the algorithm, we run 100 simulations to compute the 95% confidence interval of the can-order level. The evolution over time of the can-order level and its 95% confidence interval is depicted in Fig. 8. In this plot, we can observe that the optimal values for the can-order level do not have a great variability as its confidence interval is small.

Algorithm 1 Algorithm to obtain the can-order level

Input: Simulation parameters as in Table 1, holding cost h_i, safety factor k, set of probability distribution $\mathscr{P}$ defined by the parameter η as in (21)
Output: Optimal can-order level c_i^* for each time period

1: Initialize forecast setup-costs $\hat{A}(1)$ and $\hat{a}_i(1)$
2: **for** every time instance **do**
3:      Generate new observed demand from a Poisson distribution with parameter λ_i from Table 1
4:      Generate new observed minor setup cost from a Gaussian distribution with mean a_i from Table 1
5:      Generate new observed major setup cost from a Gaussian distribution with mean A from Table 1
6:      Estimate the setup cost by applying exponential smoothing as in equation (5)
7:      Fit a Poisson regression model to obtain the regression coefficients $\hat{\beta}_1$ and $\hat{\beta}_2$
8:      Forecast the expected demand according to equation (6)
9:      Compute the must-order level s_i and order-up-to level S_i as in equations (8) and (9), respectively
10:      **for** every probability distribution in set $\mathscr{P}$ **do**
11:        **for** every inventory level c_i between s_i and S_i **do**
12:          **for** every vector $\hat{\lambda} \in \mathbb{R}^n$ **do**
13:            Compute the discount opportunity rate as in equation (7)
14:            Compute $T_i(S_i)$, $I_i(S_i)$, and $Q_i(S_i)$ as explained in Lemma 1
15:            Compute the expected long-run average cost $\mathbb{E}[g_i(S_i, c_i, s_i)|\hat{\lambda}_i]$ as in equation (17)
16:          **end for**
17:          **if** $\mathbb{E}[g_i(S_i, c_i + 1, s_i)|\hat{\lambda}_i] > \mathbb{E}[g_i(S_i, c_i, s_i)|\hat{\lambda}_i]$ **then**
18:            Save in a vector c_{opt} the value of c_i for every probability distribution in $\mathscr{P}$
19:            Save in a vector the value of the expected long-run average cost for the optimal c_i
20:            Break
21:          **end if**
22:        **end for**
23:      **end for**
24:      c_i^* is the value of the vector c_{opt} that solves the robust optimization problem (20)
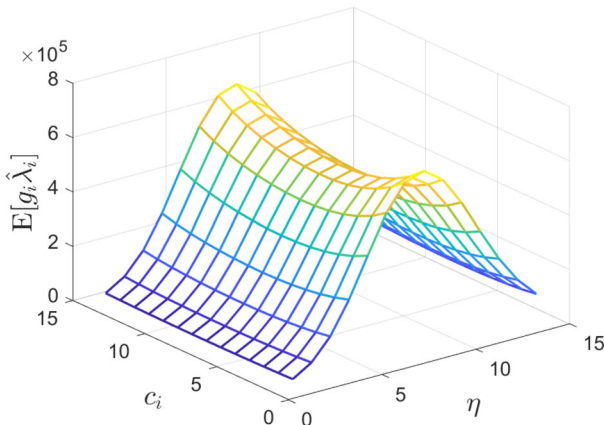25: **end for**

Fig. 6 Expected long-run average cost vs. can-order level for all probability distributions in the set $\mathscr{P}$

Table 2 Monthly optimal value of c_i

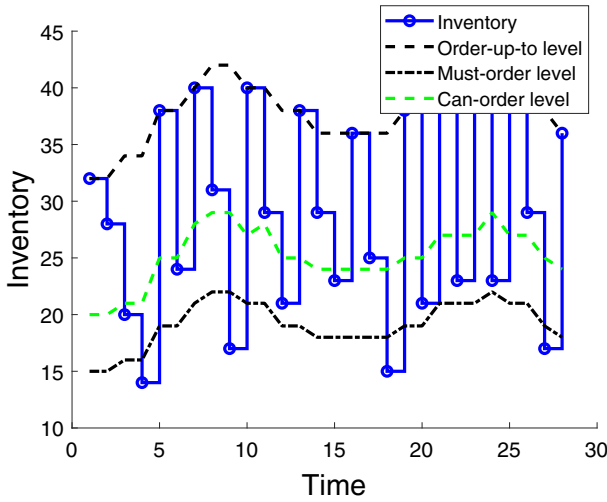| t	1	2	3	4	5	6	7	8	9	10	11	12	13	14
c_i	25	25	27	27	29	31	31	32	32	32	31	31	29	29
	15	16	17	18	19	20	21	22	23	24	25	26	27	28
	27	27	27	27	29	29	31	32	32	34	36	36	38	39

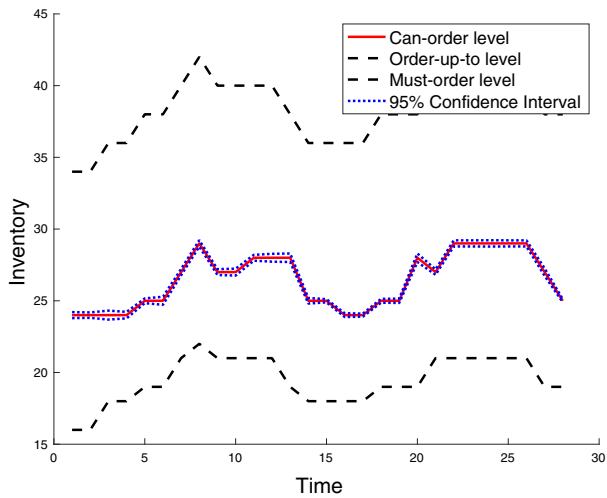**Fig. 7** Inventory evolution for the optimal (S, c, s) policy

Fig. 8 Time evolution of c_i^* (solid line) and its 95% confidence interval (dotted line)

6 Conclusions

In this research, we have developed a periodic review system with multiple items, with uncertain demand and uncertain setup costs. The problem consists in learning the uncertain demand and setup costs, while determining the values for the threshold of the can-order policy by minimizing the expected long-run average cost.

Motivations arise from two-layer time-dependent logistic systems. The first layer consists of a learning mechanism to obtain the uncertain input parameters. The setup cost and demand are uncertain, and their forecasts are run using exponential smoothing and Poisson regression, respectively. The second layer consists of a time dependent (S, c, s) policy where

the must-order, can-order and order-up-to levels are determined every time new input data are obtained. The proposed methodology is original in that the long-run average cost is optimized under the assumption that the parameters are time-dependent. We translate the initial coordinated replenishment problem into a Bayesian game, where the probabilities involved in this computation are unknown. In order to deal with the uncertainty in the probability distributions, we consider the worst-case scenario by making use of distributionally robust optimization.

The main result is the proof of convexity of the cost function with respect to the can-order level, according to which the can-order value can be optimized by solving a convex problem. In addition, we show that there exists a set of probability distributions where the cost function is concave with respect to the parameter of the corresponding distribution. Finally we show that finding an optimal can-order level that minimizes the expected cost function in the Bayesian game corresponds to finding a strategy at a Bayesian Nash equilibrium. The developed model is validated on a numerical example.

An interesting approach for future works is to analyse the model from a different game theory perspective, by focusing on minimizing the long-run average cost g_i for all the items at the same time applying, for example, the theory of strategic complements.

Appendix

In this section, we provide the proof of Lemma 1 presented in Sect. 2.

Proof of Lemma 1 Note that the expected time until the next order is placed follows a semi-Markov process that regenerates at the replenishment epochs. If the inventory level lays between $s_i < y \leq c_i$, then $T_i(y)$ is given by:

$$T_i(y) = \frac{1}{\hat{\lambda}_i(t) + \hat{\mu}_i(t)} + \frac{\hat{\lambda}_i(t)}{\hat{\lambda}_i(t) + \hat{\mu}_i(t)} T_i(y-1). \tag{47}$$

Here, the term $\frac{1}{\hat{\lambda}_i(t)+\hat{\mu}_i(t)}$ is the expected time interval between two consecutive events and is obtained assuming that the events "demand occurrence" and "discount opportunity" are independent and therefore the combination of the two events results in a Poisson process with rate equal to the sum of the rates of the two underlying events. The second term $\frac{\hat{\lambda}_i(t)}{\hat{\lambda}_i(t)+\hat{\mu}_i(t)} T_i(y-1)$ is the probability that the event is a "demand occurrence" which reduces the inventory level of one unit times the value of the same variable in the new inventory.

To obtain a general expression for $T_i(y)$, let us proceed with the following induction. Let us first consider the case where the inventory level is $s_i + 1$, which means that the inventory

level is one item above the must-order level. Namely, we have

$$T_i(s_i + 1) = \frac{1}{\hat{\lambda}_i(t) + \hat{\mu}_i(t)} + \frac{\hat{\lambda}_i(t)}{\hat{\lambda}_i(t) + \hat{\mu}_i(t)} T_i(s_i) = \frac{1}{\hat{\lambda}_i(t) + \hat{\mu}_i(t)}. \tag{48}$$

In the above, the time until the next order is placed $T_i(s_i) = 0$ as the inventory level has reached the must-order point. If the inventory level is two items above the must order level, namely $y = s_i + 2$, then the expected time until the next order is placed changes as follows:

$$T_i(s_i + 2) = \frac{1}{\hat{\lambda}_i(t) + \hat{\mu}_i(t)} + \frac{\hat{\lambda}_i(t)}{\hat{\lambda}_i(t) + \hat{\mu}_i(t)} T_i(s_i + 1) = \frac{1}{\hat{\lambda}_i(t) + \hat{\mu}_i(t)} \left(1 + \frac{\hat{\lambda}_i(t)}{\hat{\lambda}_i(t) + \hat{\mu}_i(t)} \right). \tag{49}$$

Analogously for the case of $y = s_i + 3$, we obtain

$$T_i(s_i + 3) = \frac{1}{\hat{\lambda}_i(t) + \hat{\mu}_i(t)} + \frac{\hat{\lambda}_i(t)}{\hat{\lambda}_i(t) + \hat{\mu}_i(t)} T_i(s_i + 2)$$

$$= \frac{1}{\hat{\lambda}_i(t) + \hat{\mu}_i(t)} \left(1 + \frac{\hat{\lambda}_i(t)}{\hat{\lambda}_i(t) + \hat{\mu}_i(t)} + \left(\frac{\hat{\lambda}_i(t)}{\hat{\lambda}_i(t) + \hat{\mu}_i(t)} \right)^2 \right). \tag{50}$$

For the case where $y > c_i$, $T_i(y)$ has a different structure. This can be seen by setting the inventory level to $y = c_i + 1$ and considering as next event only a demand occurring as discount opportunities will be ignored. This yields

$$T_i(c_i + 1) - T_i(c_i) = \frac{1}{\hat{\lambda}_i(t)}. \tag{51}$$

From the above, we have

$$(\hat{\lambda}_i(t) + \hat{\mu}_i)(T_i(c_i + 1) - T_i(c_i)) = 1 + \hat{\mu}_i T_i(c_i + 1) - \hat{\mu}_i T_i(c_i), \tag{52}$$

which can be rewritten as

$$(\hat{\lambda}_i(t) + \hat{\mu}_i)T_i(c_i + 1) = (\hat{\lambda}_i(t) + \hat{\mu}_i)T_i(c_i) + 1 + \hat{\mu}_i T_i(c_i + 1) - \hat{\mu}_i T_i(c_i). \tag{53}$$

Using (53) we obtain

$$T_i(c_i + 1) = \frac{1}{\hat{\lambda}_i(t) + \hat{\mu}_i(t)} + \frac{\hat{\mu}_i(t)}{\hat{\lambda}_i(t) + \hat{\mu}_i(t)} T_i(c_i + 1)\delta(1) + \frac{\hat{\lambda}_i(t)}{\hat{\lambda}_i(t) + \hat{\mu}_i(t)} T_i(c_i), \tag{54}$$

where $\delta(1)$ is the indicator function which takes value 1, if the inventory level is above c_i and 0 if the inventory level is less than or equal to c_i.

From this, it becomes clear that the solution to the renewal equation has the following general form:

$$T_i(y) := \frac{1}{\hat{\lambda}_i(t) + \hat{\mu}_i(t)} + \frac{\hat{\mu}_i(t)}{\hat{\lambda}_i(t) + \hat{\mu}_i(t)} T_i(y)\delta(y - c_i) + \frac{\hat{\lambda}_i(t)}{\hat{\lambda}_i(t) + \hat{\mu}_i(t)} T_i(y - 1). \tag{55}$$

Equation (55) has the following interpretation. Since the replenishment decision follows a Markov renewal process, a decision can be made in correspondence to two discrete events. One event corresponds to reducing the inventory of one unit, namely a unitary demand materializes, which occurs with probability $\frac{\hat{\lambda}_i(t)}{\hat{\lambda}_i(t) + \hat{\mu}_i(t)}$. A second event is referred to as

discount opportunity, which means that replenishment is triggered by another item which has reached the must-order point, and this occurs with probability $\frac{\hat{\mu}_i(t)}{\hat{\lambda}_i(t)+\hat{\mu}_i(t)}$. Combining both events, which we assume to occur with a Poissonian clock and independently then we obtain (55).

From the above, we then obtain the following compact formulation:

$$
T_i(y) := \begin{cases} \frac{1}{\hat{\lambda}_i(t)+\hat{\mu}_i(t)} \sum_{x=0}^{(y-s_i-1)} \theta(t)^x, & \text{if } s_i < y \le c_i, \\ \frac{1}{\hat{\lambda}_i(t)+\hat{\mu}_i(t)} \sum_{x=0}^{(c_i-s_i-1)} \theta(t)^x + \frac{1}{\hat{\lambda}_i(t)}(y-c_i), & \text{if } y > c_i, \end{cases} \tag{56}
$$

where $\theta(t) = \hat{\lambda}_i(t)/(\hat{\lambda}_i(t)+\hat{\mu}_i(t))$ and $\delta(x) = 1$ if $x \ge 1$, and $\delta(x) = 0$ if $u \le 1$.

An explicit expression for $T_i(S_i)$ in (11) is then

$$
T_i(S_i) := \frac{1}{\hat{\lambda}_i(t)+\hat{\mu}_i(t)} \sum_{x=0}^{(c_i-s_i-1)} \theta(t)^x + \frac{1}{\hat{\lambda}_i(t)}(S_i - c_i). \tag{57}
$$

To obtain the expected holding and penalty costs incurred until the next order is placed similarly to the calculation of $T_i(y)$, we first consider the case where the inventory level is between $s_i < y \le c_i$. Then, $I_i(y)$ is given by:

$$
I_i(y) = \frac{G_i(y)}{\hat{\lambda}_i(t)+\hat{\mu}_i(t)} + \frac{\hat{\lambda}_i(t)}{\hat{\lambda}_i(t)+\hat{\mu}_i(t)} I_i(y-1), \tag{58}
$$

where $G_i(x)$ is the one-step inventory cost and is assumed to be quasi-convex with $\lim_{|u|\to\infty} G_i(x) = \infty$. All cases of linear holding and penalty costs, and fixed cost for each backorder item are included in the same spirit as in the work of Federgruen and Zheng [7]. To see how to obtain the above equation, first the inventory level is set to $y = s_i + 1$, which yields

$$
I_i(s_i + 1) = \frac{G_i(s_i + 1)}{\hat{\lambda}_i(t)+\hat{\mu}_i(t)} + \frac{\hat{\lambda}_i(t)}{\hat{\lambda}_i(t)+\hat{\mu}_i(t)} I_i(s_i) = \frac{G_i(s_i + 1)}{\hat{\lambda}_i(t)+\hat{\mu}_i(t)}. \tag{59}
$$

To determine the above for $y = s_i + 2$, we can write

$$
\begin{aligned}
I_i(s_i + 2) &= \frac{G_i(s_i + 2)}{\hat{\lambda}_i(t)+\hat{\mu}_i(t)} + \frac{\hat{\lambda}_i(t)}{\hat{\lambda}_i(t)+\hat{\mu}_i(t)} I_i(s_i + 1) \\
&= \frac{1}{\hat{\lambda}_i(t)+\hat{\mu}_i(t)} \left(G_i(s_i + 2) + \frac{\hat{\lambda}_i(t)}{\hat{\lambda}_i(t)+\hat{\mu}_i(t)} G_i(s_i + 1) \right).
\end{aligned} \tag{60}
$$

The above expression contains a combination of the result of (59) and a new part. The previous inventory level, $s_i + 1$ is reduced by one with probability $\frac{\hat{\lambda}_i(t)}{\hat{\lambda}_i(t)+\hat{\mu}_i(t)}$, as depicted in the Markov decision process in Fig. 4.

For $y = s_i + 3$, we then have

$$
\begin{aligned}
I_i(s_i + 3) &= \frac{G_i(s_i + 3)}{\hat{\lambda}_i(t)+\hat{\mu}_i(t)} + \frac{\hat{\lambda}_i(t)}{\hat{\lambda}_i(t)+\hat{\mu}_i(t)} I_i(s_i + 2) \\
&= \frac{1}{\hat{\lambda}_i(t)+\hat{\mu}_i(t)} (G_i(s_i + 3) + \frac{\hat{\lambda}_i(t)}{\hat{\lambda}_i(t)+\hat{\mu}_i(t)} G_i(s_i + 2) + \frac{\hat{\lambda}_i(t)^2}{(\hat{\lambda}_i(t)+\hat{\mu}_i(t))^2} G_i(s_i + 1)).
\end{aligned}
$$

In the case where $y > c_i$, $I_i(y)$ presents a similar structure as $T_i(y)$, namely:

$$(I_i(c_i + 1) - I_i(c_i)) = \frac{G_i(c_i + 1)}{\hat{\lambda}_i(t)}, \tag{61}$$

where $G_i(c_i + 1)$ divided by $\hat{\lambda}_i(t)$ is the expected average holding and penalty cost until the next demand occurs. The above can be rewritten as

$$\frac{\hat{\lambda}_i(t)}{\hat{\lambda}_i(t) + \hat{\mu}_i(t)}(I_i(c_i + 1) - I_i(c_i)) = \frac{G_i(c_i + 1)}{\hat{\lambda}_i(t) + \hat{\mu}_i(t)}. \tag{62}$$

From the above, we obtain that the solution to the renewal equation is given by:

$$I_i(y) := \frac{G_i(y)}{\hat{\lambda}_i(t) + \hat{\mu}_i(t)} + \frac{\hat{\mu}_i(t)}{\hat{\lambda}_i(t) + \hat{\mu}_i}I_i(y)\delta(y - c_i) + \frac{\hat{\lambda}_i(t)}{\hat{\lambda}_i(t) + \hat{\mu}_i(t)}I(y - 1). \tag{63}$$

Also in compact form we can write

$$I_i(y) := \begin{cases} \frac{1}{\hat{\lambda}_i(t) + \hat{\mu}_i(t)} \sum_{x=0}^{(y-s_i-1)} \theta(t)^x G_i(y - x), & \text{if } s_i < y \leq c_i, \\ \frac{1}{\hat{\lambda}_i(t) + \hat{\mu}_i(t)} \sum_{x=0}^{(c_i-s_i-1)} \theta(t)^x G_i(y - x) + \frac{1}{\hat{\lambda}_i(t)} \sum_{x=c_i+1}^{y} G_i(x), & \text{if } y > c_i, \end{cases}$$

where $\theta = \hat{\lambda}_i(t)/(\hat{\lambda}_i(t) + \hat{\mu}_i(t))$ and again $\delta(x) = 1$ if $x \geq 1$, and $\delta(x) = 0$ if $x \leq 1$.

An explicit expression for $I_i(S_i)$ in (11) is then

$$I_i(S_i) = \frac{1}{\hat{\lambda}_i(t) + \hat{\mu}_i(t)} \sum_{x=0}^{(c_i-s_i-1)} \theta(t)^x G_i(S_i - x) + \frac{1}{\hat{\lambda}_i(t)} \sum_{x=c_i+1}^{S_i} G_i(x). \tag{64}$$

Finally, denote the probability that an order is triggered by a demand by $Q_i(y)$. When the inventory level is $y > c_i$, no replenishment decision is made and therefore $Q_i(y)$ is given by

$$Q_i(y) = Q_i(y - 1). \tag{65}$$

Note that in the region of $s_i < y \leq c_i$, $Q_i(y)$ is given by:

$$Q_i(y) = \theta_i(t)^{(y-s_i)}. \tag{66}$$

Observe that the latter is the probability that $y - s_i$ unitary demand events occur consecutively. When this happens, the inventory level reaches the must-reorder point and a reorder must occur.

To see how to obtain (66), let $y = s_i + 1$, which results in the following:

$$Q_i(s_i + 1) = \theta(t)Q_i(s_i) = \theta(t) = \frac{\hat{\lambda}_i(t)}{\hat{\lambda}_i(t) + \hat{\mu}_i(t)}. \tag{67}$$

In the above, we use the fact that $Q_i(s_i) = 1$, as an order is placed directly when the inventory level is equal to the must-order level s_i. By repeating the same reasoning for $y = s_i + 2$, we obtain

$$Q_i(s_i + 2) = \theta(t)Q_i(s_i + 1) = \left(\frac{\hat{\lambda}_i(t)}{\hat{\lambda}_i(t) + \hat{\mu}_i(t)}\right)^2. \tag{68}$$

For $y = s_i + 3$, we obtain:

$$Q_i(s_i + 3) = \theta(t)Q_i(s + 2) = \left(\frac{\hat{\lambda}_i(t)}{\hat{\lambda}_i(t) + \hat{\mu}_i(t)}\right)^3. \tag{69}$$

From the above, we derive that the solution to the renewal equation can be written as

$$Q_i(y) := Q_i(y-1)\delta(y-c_i) + (1 - \delta(y-c_i))\theta_i(t)^{(y-s_i)}. \tag{70}$$

In compact form, we can write

$$Q_i(y) := \begin{cases} \theta(t)^{(y-s_i)}, & \text{if } s_i < y \le c_i, \\ Q_i(y-1), & \text{if } y > c_i. \end{cases} \tag{71}$$

An explicit expression for $Q_i(S_i)$ in (11) is then

$$Q_i(S_i) = \theta(t)^{(c_i - s_i)}. \tag{72}$$

$\square$

References

1. Alumur SA, Nickel S, Saldanha-da Gama F (2012) Hub location under uncertainty. Transp Res Part B Methodol 46(4):529–543
2. Axsäter S, Zhang WF (1999) A joint replenishment policy for multi-echelon inventory control. Int J Prod Econ 59(1–3):243–250
3. Balintfy JL (1964) On a basic class of multi-item inventory problems. Manag Sci 10(2):287–297
4. Buchbinder N, Kimbrel T, Levi R, Makarychev K, Sviridenko M (2013) Online make-to-order joint replenishment model: primal-dual competitive algorithms. Oper Res 61(4):1014–1029
5. Colin A, Pravin K (2013) Regression analysis of count data. Cambridge University Press, Cambridge
6. Eynan A, Kropp DH (2007) Effective and simple eoq-like solutions for stochastic demand periodic review systems. Eur J Oper Res 180(3):1135–1143
7. Federgruen A, Zheng YS (1992) An efficient algorithm for computing an optimal (r, q) policy in continuous review stochastic inventory systems. Oper Res 40(4):808–813
8. Federgruen A, Groenevelt H, Tijms HC (1984) Coordinated replenishments in a multi-item inventory system with compound poisson demands. Manag Sci 30(3):344–357
9. Fildes R, Nikolopoulos K, Crone SF, Syntetos AA (2008) Forecasting and operational research: a review. J Oper Res Soc 59(9):1150–1172
10. Forsberg R (1995) Optimization of order-up-to-s policies for two-level inventory systems with compound poisson demand. Eur J Oper Res 81(1):143–153
11. Gardner ES Jr (2006) Exponential smoothing: the state of the art part ii. Int J Forecast 22(4):637–666
12. Goyal SK, Satir AT (1989) Joint replenishment inventory control: deterministic and stochastic models. Eur J Oper Res 38(1):2–13
13. Grafarend EW (2006) Linear and nonlinear models: fixed effects, random effects, and mixed models. de Gruyter, Berlin
14. Harsanyi JC (1968) Games with incomplete information played by "bayesian" players, i–iii: Part i. the basic model. Manag Sci 50:1804–1817
15. Harsanyi JC (1968) Games with incomplete information played by "bayesian" players, i–iii: Part ii. bayesian equilibrium points. Manag Sci 14:320–334
16. Harsanyi JC (1968) Games with incomplete information played by "bayesian" players, i–iii: Part ii. part iii. the basic probability distribution of the game. Manag Sci 14:486–502
17. Hyndman R, Koehler AB, Ord JK, Snyder RD (2008) Forecasting with exponential smoothing: the state space approach. Springer, Berlin
18. Ignall E (1969) Optimal continuous review policies for two product inventory systems with joint setup costs. Manag Sci 15(5):278–283
19. Johansen SG, Melchiors P (2003) Can-order policy for the periodic-review joint replenishment problem. J Oper Res Soc 54(3):283–290

20. Kayiş E, Bilgiç T, Karabulut D (2008) A note on the can-order policy for the two-item stochastic joint-replenishment problem. IIE Trans 40(1):84–92
21. Khouja M, Goyal S (2008) A review of the joint replenishment problem literature: 1989–2005. Eur J Oper Res 186(1):1–16
22. Kiesmüller GP (2010) Multi-item inventory control with full truckloads: a comparison of aggregate and individual order triggering. Eur J Oper Res 200(1):54–62
23. Larsen C (2009) The q (s, s) control policy for the joint replenishment problem extended to the case of correlation among item-demands. Int J Prod Econ 118(1):292–297
24. Li Q, Zheng S (2006) Joint inventory replenishment and pricing control for systems with uncertain yield and demand. Oper Res 54(4):696–705
25. Minner S, Silver EA (2007) Replenishment policies for multiple products with compound-poisson demand that share a common warehouse. Int J Prod Econ 108(1–2):388–398
26. Nielsen C, Larsen C (2005) An analytical study of the q (s, s) policy applied to the joint replenishment problem. Eur J Oper Res 163(3):721–732
27. Pantumsinchai P (1992) A comparison of three joint ordering inventory policies. Decis Sci 23(1):111–127
28. Petrovic D, Xie Y, Burnham K, Petrovic R (2008) Coordinated control of distribution supply chains in the presence of fuzzy customer demand. Eur J Oper Res 185(1):146–158
29. Porras E, Dekker R (2008) A solution method for the joint replenishment problem with correction factor. Int J Prod Econ 113(2):834–851
30. Scarf K, Arrow SK (1958) A min–max solution of an inventory problem. Stanford University Press, Redwood City, pp 201–209
31. Silver EA (1973) Three ways of obtaining the average cost expression in a problem related to joint replenishment inventory control. Nav Res Logist Q 20(2):241–254
32. Silver EA (1974) A control system for coordinated inventory replenishment. Int J Prod Res 12(6):647–671
33. Tsai CY, Tsai CY, Huang PW (2009) An association clustering algorithm for can-order policies in the joint replenishment problem. Int J Prod Econ 117(1):30–41
34. Yao MJ (2010) A search algorithm for solving the joint replenishment problem in a distribution center with warehouse-space restrictions. Int J Oper Res 7(2):45–60
35. Zheng YS (1994) Optimal control policy for stochastic inventory systems with markovian discount opportunities. Oper Res 42(4):721–738