Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 58-19 (2024) 427–432

Joint Discount and Replenishment Parametric Policies for Perishable Products

Edoardo Fadda * **Daniele Giovanni Gioia** ** **Paolo Brandimarte** * **Francesca Maggioni** ***

* *Dipartimento di Scienze Matematiche, Politecnico di Torino, Torino, Italy (e-mail: paolo.brandimarte, edoardo.fadda@polito.it).*
** *German Aerospace Center, Institute for the Protection of Terrestrial Infrastructures, Sankt Augustin, Germany (e-mail: daniele.gioia@dlr.de)*
*** *Dipartimento di Ingegneria Gestionale, dell'Informazione e della Produzione, Università degli Studi di Bergamo, Bergamo, Italy (e-mail: francesca.maggioni@unibg.it).*

Abstract: We consider a joint discount and replenishment problem in a discrete periodic review fashion for the sale of a perishable product, characterized by limited deterministic shelf life, replenishment lead times, and stochastic demand. Customers decide what to buy according to a linear discrete choice model, balancing price and perceived quality, uniquely determined by the residual shelf life. The decisions we consider are: How many new items to order, the age of the items to be discounted, and how much discount to offer. In this context, we compare a set of policies mixing the constant order policy and the base stock one with some easy discounting policies, optimizing their parameters using a simulation-based optimization framework. To evaluate their performance in terms of revenue and quantity of scraped items, we consider four realistic instances for a grocery retailer characterized by products of different shelf life and variance of demand. Experiments show that best results are achieved by a base stock policy that discounts products of different ages based on a threshold: If the quantity of the inventory of a given age is greater than a threshold it applies a discount, otherwise no discount is proposed. In the presented configurations, this policy increases the average reward compared to policies that do not discount.

Keywords: inventory control; perishable products; discounts; parametric policy function; consumer behaviour; simulation-based optimization.

1. INTRODUCTION

Managing perishable items in a retail setting is a challenging but relevant problem, in terms of profitability on the one hand, and food waste reduction on the other hand. The problem is strongly affected by uncertainty, mainly due to demand size and consumer attitude towards items of different characteristics: Customers can be more attracted by new items, thus implementing a *Last-In-First-Out* (LIFO) strategy, or they can be environmentally conscious and prone to cooperate and behave according to a *First-In-First-Out* (FIFO) scheme. Nevertheless, consumers may freely choose their preferred items in a retail setting, and inventory issuing cannot be tightly controlled. To encourage purchases of non-fresh items and reduce waste while improving profits, discounts might be introduced on older items.

Existing literature and practical experiences suggest that, in addition to influencing purchasing decisions, pricing serves as an effective strategy for enhancing the profitability of perishable products by affecting the demand (Şen, 2013; van Donselaar et al., 2016). Therefore, the retailer has not only to choose an ordering policy based on the system state but also to decide if and when a discount should be offered. In this paper, we explore these decisions in a periodic review setting. In particular, following the classification for pricing problem of perishable items in Elmaghraby and Keskinocak (2003), we are dealing with *deterministic* shelf life, with *replenishment*, considering *independent demand over time* and *delivery lead time*. *Myopic customers* are considered, i.e., customers purchasing the product if their utility is positive, without considering future prices (conversely, *strategic* customers take into account the future path of prices when making purchasing decisions). To model customers' utility, toward different product ages, we use a linear discrete choice model, which allows us to model vertical differentiation (Pan and Honhon, 2011). These characteristics define our problem as belonging to the *Replenishment - Independent demand - Myopic* class. Finally, we do not consider menu costs, i.e., there are no costs associated with the discount application (neither in terms of updating computer systems nor in terms of re-tagging items, printing new menus, etc.).

The presented approach allows for an analysis of the characteristics of a set of ordering and discounting rules, provides useful insight, and paves the way for more advanced techniques. The contributions of this paper are: *(i)* test different policies of join replenishment and discounting in

10.1016/j.ifacol.2024.09.249

a realistic simulation environment, *(ii)* quantify the value of using a discounting policy.

The paper is organized as follows: Section 2 provides the necessary literature background. Section 3 presents the mathematical model of the problem. Section 4 describes the set of parametric policies used in the computational experiments of Section 5. Finally, Section 6 concludes the paper and outlines future research directions.

2. LITERATURE REVIEW

The inventory management literature typically assumes that the price of a product is a static single price, exogenous to the inventory management problem. In contrast, the literature about dynamic pricing focuses solely on price variation. This paper has a vision in between, where the initial price is fixed but decisions to lower it when aging occurs to increase revenue and reduce waste are possible.

Being discounting a sub-problem of pricing, all the literature related to policies that jointly manage replenishment and dynamic pricing can be of potential interest. Nevertheless, since the literature about these policies is broad, we focus on reviewing papers with closer settings to ours, i.e., Chen and Simchi-Levi (2004), Chew et al. (2014), Sainathan (2013), and Chua et al. (2017).

Chen and Simchi-Levi (2004) deal with pricing and inventory control decisions for a single non-perishable product, where demand distribution depends on the product price, all shortages are backorders, and both fixed and variable ordering costs are considered. They show that in this setting it is optimal to replenish the on-hand inventory up to a certain level whenever it falls below a certain threshold while determining the price according to the inventory position at the beginning of each period.

Chew et al. (2014) propose a technique based on dynamic programming to compute the optimal order quantity and prices for a perishable product with a shelf life of two periods. Their computational results show that the total profit significantly increases when demand transfers between products of different ages are considered.

In the same setting, Sainathan (2013) shows that demand uncertainty, together with discounts, can make the sale of old products profitable. Moreover, they find that the benefit obtained from selling old items with constant decisions is much higher than the benefit from allowing all the decisions to vary.

Finally, Chua et al. (2017) measure the impact of consumer behavior and shelf life on discount decisions for old items, determining the discount amount, and establishing a replenishment policy. They find the optimal policy for an item of shelf life equal to two periods and consider a finite time horizon with no lead time. In particular, they prove that with given discounts, the optimal discounting policy is a threshold policy, and the optimal order quantity decreases according to the inventory of old units with a significant decrease at the threshold. They also show that, if a discount is considered, its value first increases and then decreases (eventually to zero) in the inventory of old units. Their results are proved up to instances with two periods while for bigger instances the course of dimensionality prevents finding optimal solutions.

In our study, we start to extend their work by considering greater shelf life, lead times, and an infinite time horizon. We use a discrete choice model for modeling customers' choices. Due to such realistic, but complex, features, we rely on heuristic policies.

3. PROBLEM STATEMENT

We consider a retailer who sells a single perishable product and makes daily decisions on the quantity of fresh items to order and possible discounts to apply.

The product is characterized by: Purchasing cost c, selling price p, a fixed discrete delivery lead time LT, and a deterministic discrete shelf life τ equal to SL at delivery time. Items with a residual shelf life τ have a perceived quality q_τ and when $\tau = 0$, they are scrapped. Quality is increasing in τ, i.e., $q_{\tau+1} \geq q_\tau$.

We divide the simulation into discrete time steps t, where each period corresponds to one day. At the beginning of day t, prices for old items are updated according to discounts (we call p_τ the price for items with residual shelf life τ). The number of items to be delivered in $l \in \{0, \dots, \mathsf{LT}-1\}$ days is called O_t^l, and the on-hand inventory of items with a residual life of $\tau \in \{1, \dots, \mathsf{SL}\}$ days at time t before sales is called I_t^τ.

When the shop opens, N_t heterogeneous customers enter the store. For the sake of simplicity, we assume N_t to be independent and identically distributed random variables. Each customer is characterized by a utility function that differently weights quality and price. We model their choice using a linear discrete choice model (Gioia et al., 2022, 2023). Therefore, the utility of customer n is:

$$U_{n\tau} = \theta_n q_\tau - p_\tau, \qquad (1)$$

where θ_n is a stochastic positive variable representing the consumer's valuation of the price-quality ratio of the item with residual shelf life τ.

If at least a combination of quality and price generates positive utility, the customer picks one unit of the item that maximizes it. In formula, customer n picks the item of residual shelf life τ if

$$\tau \in \arg\max_{\tau : I_t^\tau > 0} \left[\theta_n q_\tau - p_\tau\right]. \qquad (2)$$

Lost sales costs are not considered due to the complexity of tracking them in brick-and-mortar grocery retail settings.

If no discount is applied and items of all ages are available, customers prefer new items to older ones. This is due to a vertical differentiation entailed by the linear discrete choice model between items of different ages (Gioia et al., 2022). The effect of discounts for a product with fixed price $p = 6$, $\mathsf{SL} = 2$, $q_2 = 20$, $q_1 = 10$, and $\theta_n \sim \text{beta}(2, 3)$ is represented in Fig. 1. Since $\mathsf{SL} = 2$, we can either have new products ($\tau = 2$) or old item ($\tau = 1$). The red dash-dotted (green continuous) lines represent the utility in Eq. (1) for different types of customers of the new (old) item. Discounting old items leads to a vertical shift of the green continuous line (the green dashed line in the plot) which has two effects. On the one hand, it opens a new share

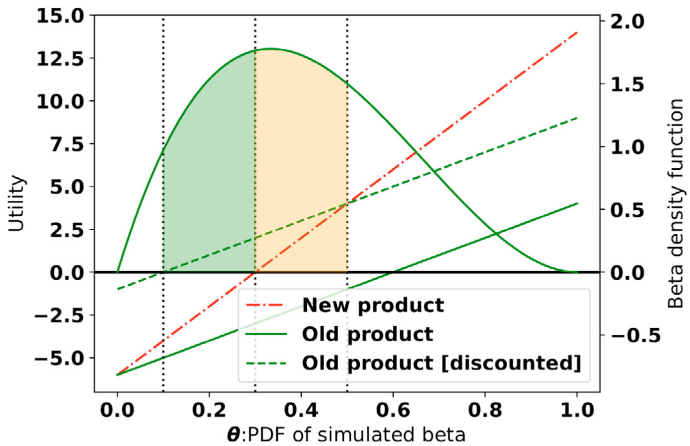

Fig. 1. Graphical representation of the effect of the discount on old items.

of the market, providing a positive utility for customers who are not willing to pay the price for fresh items (this is represented as the green area). In other words, discounts are useful to exploit the share of customers not originally interested in the full-price product. On the other hand, discounting creates a new price-quality combination that, for some customers, can be better than the new and more remunerative product (the yellow area), leading to the cannibalization of part of the profit coming from new discounted items [this is an example of *age-based* substitution (Gioia et al., 2023)].

At the end of the day, the retailer examines whether any product has expired (scrapping it accordingly), updates the inventory residual shelf-lives, and decides

- the quantity of product to order $x_t \in \mathcal{X} \subset \mathbb{N}$
- the discount $\delta_t^\tau \in \Delta$ to apply to items with residual life $\tau = 1, \dots, \mathsf{SL}$.

Both sets $\mathcal{X}$ and Δ are discrete sets, where $\mathcal{X}$ contains the possible orders [retailers buy batches of products (Broekmeulen and van Donselaar, 2019)], and Δ contains the possible discounts.

The dynamics of the system are represented in Figure 2.

We can model the system as a sequential optimization problem (Powell, 2022), in which the state of the system is $[O_t^{\mathsf{LT}-1}, \dots, O_t^0, I_t^{\mathsf{SL}-1}, \dots, I_t^1]$. Then, the decision maker decides the number of items to order (x_t), and the discounts to apply (δ_t^τ). Once these decisions are fixed, the order O_t^0 arrives at the shop generating the new post decision state $[x_t, O_t^{\mathsf{LT}-1}, \dots, O_t^0 = I_t^{\mathsf{SL}}, I_t^{\mathsf{SL}-1}, \dots, I_t^1]$ and customers enter the shop. By calling S_t^τ the number of sold items with residual shelf life τ, the retailer gets a *reward*

$$R_t = \sum_{\tau=1}^{\mathsf{SL}} p(1 - \delta_t^\tau) S_t^\tau - c x_t. \tag{3}$$

Then, the inventory becomes

$$I_{t+1}^{\tau-1} = I_t^\tau - S_t^\tau \qquad \forall\, \tau \geq 1, \tag{4}$$

and orders shift accordingly

$$O_{t+1}^{l-1} = O_t^l \qquad \forall\, l \in \{1, \dots, \mathsf{LT}\}. \tag{5}$$

In contrast with several studies, we consider an infinite time horizon problem and our goal is to maximize the average profit.[1]

4. SOLUTION APPROACHES

All the proposed solution approaches fall under the umbrella of *parametric policy function approximation*, i.e., parametric functions that map a state to an action (Powell, 2022). While many replenishment policies are well known for perishable (Haijema and Minner, 2016), there are fewer discounting policies. The most used ones in practice are:

- Fixed percentage discount applied to all items with shelf life less than SL (Chen et al., 2021).
- Periodic clearance sales which are used to sell items quickly and make room for new ones (Chen et al., 2021).
- Tiered discounting which implements different discount levels based on the age of the product (Chen et al., 2021).
- Last day discount which applies the discount the last day before expiration (Broekmeulen and van Donselaar, 2019).

Within the wider context of pricing policies, an important policy is the *base stock list price policy* (BSLP) (Zabel, 1970). This policy is characterized by a base stock level and a list of prices. If the inventory level is below the base stock level, it is increased to the base stock level and the standard price is charged. Otherwise, nothing is ordered, and a price discount is offered. This policy is shown to be optimal under certain conditions in a finite horizon in Zabel (1970). We generalize the discounting policies and the BSLP and we join them with the replenishment policies by defining the following policies:

- `COPs`: Orders a constant quantity and discounts all the items older than τ_0 by an amount equal to δ. Therefore, this policy requires 3 parameters to be set.
- `COPtr`: Orders a constant quantity and applies a discount δ_τ to products with residual shelf life τ if I_t^τ is greater than a threshold ψ_τ. Therefore, this policy requires $1+2(\mathsf{SL}-1)$ parameters to be set (we assume no discount on fresh items).
- `BSs`: Orders
$$x_t = \max[z - \sum_{l=0}^{\mathsf{LT}-1} O_t^l - \sum_{\tau=1}^{\mathsf{SL}-1} I_t^\tau, 0], \tag{6}$$
where z is the parameter of the base stock level. Moreover, it applies discount as in `COPs`, thus this policy requires 3 parameters to be set.
- `BStr`: Orders according to (6) and it discounts as `COPtr`, thus requiring $1 + 2(\mathsf{SL} - 1)$ parameters to be set.

To assess the value of discounting, we also consider the base stock policy and the constant order policy with no discount. We call them `BS`, and `COP`, respectively.

[1] Note that there is no fixed order cost since we assume that orders are made so often that the logistic provider will deliver products each day.

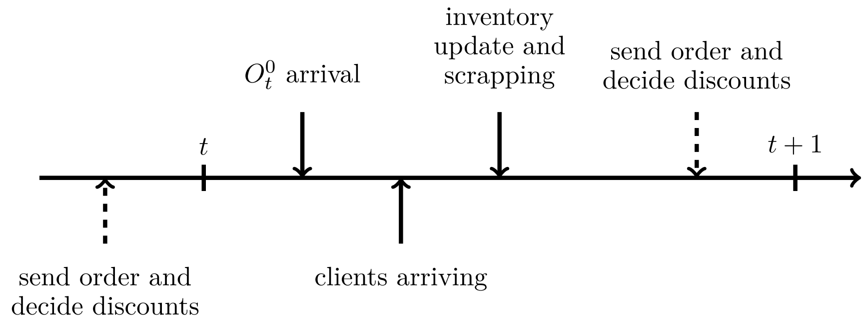

Fig. 2. Dynamics of the system

5. COMPUTATIONAL EXPERIMENTS

To test the policies in a realistic setting, parameters are based on data for perishables from empirical studies in grocery stores. In particular, according to van Donselaar et al. (2006); Broekmeulen and van Donselaar (2019), we consider orders in batches of 6 items, $\mathsf{LT} = 1$, and two possible values for SL, of 5 and 7 time periods. Instead, following ÓRiordan (1993), we set the margin of the product, computed by using the newsvendor ratio $\frac{p-c}{p}$, to ∼30%. Therefore, we set the price of the new product to be 6 and its cost to be 4. Finally, we define the set of possible discounts $\Delta = \{0\%, 15\%, 25\%, 50\%\}$.

The number of customers entering the shop is distributed according to a negative binomial distribution, widely used in market applications Ehrenberg (1959), and interpretable as a Poisson distribution with parameters distributed according to gamma or, as a compound Poisson process with geometrically distributed purchases quantity Agrawal and Smith (2015). Each customer entering the shop is characterized by a θ_n in Eq.(1) distributed according to a beta$(2, 3)$ (Gioia et al., 2023). We set the mean number of customers entering the shop to $\mu = 30$ and the standard deviation σ using the *coefficient of variation* ($cv = \frac{\sigma}{\mu}$). Since the literature provides cv for the sales, we transform them to obtain μ, σ of the consumers. In particular, in Eq. (1) of Broekmeulen and van Donselaar (2019) the authors suggest that the mean μ_w and standard deviation σ_w of weekly sales are linked by

$$\sigma_w = 0.7\mu_w^{0.77}. \tag{7}$$

Adapting Eq. (7) to a daily setting and assuming independent demand, we get $\sigma_d = 1.18\mu_d^{0.77}$. The mean of the demand is $\mu_d = \mathbb{E}[\sum_{n=1}^{N_t} Y_n]$, where Y_n is a Bernoulli random variable equal to 1 if customer n buys an item. By assuming that there are only new items, Eq. (1) implies that

$$\mathbb{P}[Y_n = 1] = \mathbb{P}[\theta_n \geq \frac{p_{\mathsf{SL}}}{q_{\mathsf{SL}}}].$$

Therefore, it holds that

$$\mathbb{E}[\sum_{n=1}^{N_t} Y_n] = \mathbb{E}\left[\mathbb{E}[\sum_{n=1}^{N_t} Y_n | N_t]\right] = \mu\left(1 - F_{\text{beta}(2,3)}\left(\frac{p_{\mathsf{SL}}}{q_{\mathsf{SL}}}\right)\right),$$

where $F_{\text{beta}(2,3)}$ is the cumulative distribution of the beta$(2, 3)$ random variable. With a similar calculation, we get the correspondent standard deviation

$$\sigma_d = \sqrt{(1 - F_{\text{beta}(2,3)})\left(\sigma^2 + \mu^2 F_{\text{beta}(2,3)}\right)}.$$

Inverting these two equations, we get $cv \approx 0.3$ (i.e., $\sigma = 9$) on the *consumers'* distribution. Furthermore, since high daily demands are correlated to higher correspondent daily standard deviations (Broekmeulen and van Donselaar, 2019) and such empirical estimations do not consider discounted items, we also test the policies in a more difficult setting, using $cv = 0.7$ (i.e., $\sigma = 21$).

The quality of the product at the different ages is set to 30, 29, 29.5, 28, 26, 24, 22 for the case with $\mathsf{SL} = 7$, and 30, 29, 28, 26, 24 for the case with $\mathsf{SL} = 5$.

Since the stochastic problem is ergodic, the estimate of the expected value can be calculated by simulation over a sufficiently long horizon (set to 70000 time periods in the experiments). To provide an estimation with reasonable accuracy, we apply two precautions. First, we compute the average using the observations after the first 1000. In such a way, we remove the effect of the initial transient where the inventory is empty because of the lead time. Second, we use a 100-period sliding window to decide whether the estimate is sufficiently accurate. If the difference between the maximum and minimum value of the estimated mean in that window is less than 0.5%, we stop the simulation. Tests have been implemented in Python generalizing the code available in Gioia et al. (2023).

To compute the parameters of the policies we use simulation-based optimization techniques, broadly used in practical inventory control problems (Deng et al., 2023). In particular, the parameters of `COP`, `BS`, `COPs`, and `BSs` are optimized using an enumerative search of the parameter space since, due to the characteristics of the problem, they can be computed in a reasonable amount of time. This enables us to get more accurate results without heuristic optimization algorithm uncertainties. Instead, the parameters of `COPtr` and `BStr` are optimized using the *bayes_opt* library (Nogueira, 2014). We use 50 initial points for pure exploration and 100 steps of Bayesian optimization. We do not modify the standard setting of the library. The selection of the heuristic optimization method for policies with a large number of parameters and the accuracy of reward estimates are hyperparameters of the proposed configurations.

For all the policies, we select a subset of possible parameter values following the bounds of Haijema and Minner (2016). In particular, we search the base stock level by considering the values in $[0, (\mathsf{SL}+\mathsf{LT})F^{-1}_{\text{n.b.}}(\frac{p-c}{p})]$, where $F^{-1}_{\text{n.b.}}$ is the inverse of the cumulative distribution function of the number of customer entering the shop (i.e., the negative binomial). Instead, we search for the constant order in the interval $[0, z]$, where z is the base stock level (we use the base stock level of `BSs` for `COPs` and the one of `BStr` for `COPtr`).

We report the results for the different settings in Table 1. In each row, we report the results of a technique and we add * if parameter optimization has been done by enumeration.

The average reward decreases as cv increases, regardless of the policy. This pattern holds for both $\mathsf{SL} = 5$ and for $\mathsf{SL} = 7$. This is due to the difficulty in handling demand variance which leads all the methods to reduce the quantity ordered also reducing the probability of having a big quantity of old items. Moreover, when there is a high variance, lower orders reduce the risk of having many fresh products that jeopardize the sales of older ones.

In contrast to the cv effect, the change in shelf life increases the profit for all the settings: Having more periods available augments the flexibility in selling items. This effect is bigger for $cv = 0.7$ than for $cv = 0.3$, where the average profit increase reaches an average of 3.5% against the 1.9% when $cv = 0.3$. This difference is due to the longer expiration time that enables the agents to better deal with the variance of the number of customers.

The change in SL has different effects also on the quantity of scrapped items: For `COP`, `COPs`, and `COPtr` the quantity of scrapped items decreases as the shelf life increases, while it remains almost constant for the base stock policies.

By comparing constant order policies against base stock ones, we notice that the latter achieves the best results. This is due to the possibility of ordering less when both on-order and inventory are high. This lead also to a reduced quantity of scrapped items. More in detail, the differences in average reward between these two types of policies decrease as cv increases while it remains almost the same for the two shelf-lives.

For $cv = 0.3$ and $\mathsf{SL} = 5{,}7$, `COPs` does not apply any discounts, while `COPtr` applies discounts increasing the average profit of 8.6%. This means that sometimes, applying a constant discount independently of the inventory quantity is not beneficial. Instead, for $cv = 0.7$, `COPs` applies discounts as well, improving the performance of the `COP` policy by 2.7% and by 3.6% for $\mathsf{SL} = 5$ and $\mathsf{SL} = 7$, respectively. Concerning base stock policies `BSs` performs better than `BS` by 2.4%, while `BStr` performs better than `BS` by 8.8%, on average. The finding that threshold policies perform better than the other ones is an empirical verification that confirms the results obtained in Chua et al. (2017).

Joint effects of replenishment and discounting are very different for the two types of policies: The threshold discounting policy achieves better results if mixed with base stock policies than the constant order ones.

Concerning the policies implemented, we also notice that the general discounting features are independent of SL and cv: They all discount the items in two or three periods from the end of the shelf life. This is because in the initial periods, it is more likely to sell the items since on the one hand, their quality is not very low, on the other hand, the probability of ending fresher items is bigger. This general observation is deeply influenced by quality evolution: If the quality remains almost the same there is little value in making the discount while if the quality reduces a lot, shrinking the market for the product, applying discounts even in earlier stages is beneficial.

Concerning orders, all the policies order quantities of the product below the average number of arrivals due to the low newsvendor ratio and the presence of customers with negative utility. Moreover, if there are no fresh products, many customers will buy the old products without the need for discounting.

Finally, we also mention that the last-day discount policy and the fixed percentage discount one of Section 4 are special cases of `COPs` and `BSs`.

6. CONCLUSIONS

In this paper, we focus on comparing different policies in a setting characterized by a single perishable product with deterministic shelf life, independent demand over time, delivery lead time, and myopic customers with utilities modeled according to a linear discrete choice model. Similar to Chew et al. (2014), we show that deciding on replenishment and discounting is beneficial. Moreover, we found that in the proposed realistic setting, the best results are achieved by a policy that manages the inventory by using a base stock policy and that discounts the products of different ages based on a threshold: If the quantity of the inventory of a given age is greater than a threshold it applies a discount, otherwise no discount is proposed. This policy increases the average reward compared to policies that do not apply any discount.

Future work will tackle the problem in many other ways, such as *(i)* a static robust approach, where we assume that unknown parameters are contained within an uncertainty set, and we select a policy maximizing long-term average profit in the worst-case sense, *(ii)* a model-based learning approach, in which we assume a model structure and we apply a learning strategy to estimate the parameters, possibly starting from prior knowledge, following a Bayesian inventory management strategy, and *(iii)* a model-free learning approach, where we use a reinforcement learning strategy and try to improve a decision policy online. These three strategies need not be mutually exclusive. For instance, a static robust approach might provide an initial policy for online learning.

REFERENCES

Şen, A. (2013). A comparison of fixed and dynamic pricing policies in revenue management. *Omega*, 41(3), 586–597. doi:10.1016/j.omega.2012.08.001.

Agrawal, N. and Smith, S.A. (2015). *Retail Supply Chain Management: Quantitative Models and Empirical Studies*. Springer US. doi:10.1007/978-1-4899-7562-1.

Broekmeulen, R.A. and van Donselaar, K.H. (2019). Quantifying the potential to improve on food waste, freshness and sales for perishables in supermarkets. *In-*

	Avg Reward				Avg Scrapped			
	SL = 5		SL =7		SL =5		SL =7	
Policy	cv=0.3	cv=0.7	cv=0.3	cv=0.7	cv=0.3	cv=0.7	cv=0.3	cv=0.7
COP *	34.8	28.4	35.5	29.3	2.1	1.5	1.6	0.8
COPs *	34.8	29.1	35.5	30.4	2.5	0.8	1.5	0.6
COPtr	37.8	35.0	38.5	35.0	1.9	0.9	1.5	0.3
BS *	40.6	33.8	41.2	35.3	0.4	0.9	0.3	0.8
BSs *	41.9	34.5	42.4	35.6	0.1	0.9	0.1	0.1
BStr	45.3	35.4	45.8	37.8	0.2	0.9	0.3	0.6

Table 1. Average reward and average quantity of scrapped items for COP, COPs, COPtr, BS, BSs, and BStr in different settings.

ternational Journal of Production Economics, 209, 265–273. doi:10.1016/j.ijpe.2017.10.003.

Chen, X. and Simchi-Levi, D. (2004). Coordinating inventory control and pricing strategies with random demand and fixed ordering cost: The infinite horizon case. *Mathematics of Operations Research*, 29(3), 698–723. doi:10.1287/moor.1040.0093.

Chen, Y., Mehrotra, P., Samala, N.K.S., Ahmadi, K., Jivane, V., Pang, L., Shrivastav, M., Lyman, N., and Pleiman, S. (2021). A multiobjective optimization for clearance in walmart brick-and-mortar stores. *INFORMS Journal on Applied Analytics*, 51(1), 76–89. doi:10.1287/inte.2020.1065.

Chew, E.P., Lee, C., Liu, R., Hong, K.s., and Zhang, A. (2014). Optimal dynamic pricing and ordering decisions for perishable products. *International Journal of Production Economics*, 157, 39–48. doi:10.1016/j.ijpe.2013.12.022.

Chua, G.A., Mokhlesi, R., and Sainathan, A. (2017). Optimal discounting and replenishment policies for perishable products. *International Journal of Production Economics*, 186, 8–20. doi:10.1016/j.ijpe.2017.01.016.

Deng, Y., Zhang, X., Wang, T., Wang, L., Zhang, Y., Wang, X., Zhao, S., Qi, Y., Yang, G., and Peng, X. (2023). Alibaba realizes millions in cost savings through integrated demand forecasting, inventory management, price optimization, and product recommendations. *INFORMS Journal on Applied Analytics*, 53(1), 32–46. doi:10.1287/inte.2022.1145.

Ehrenberg, A.S. (1959). The pattern of consumer purchases. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 8(1), 26–41. doi:10.2307/2985810.

Elmaghraby, W. and Keskinocak, P. (2003). Dynamic pricing in the presence of inventory considerations: Research overview, current practices, and future directions. *Management Science*, 49(10), 1287–1309. doi:10.1287/mnsc.49.10.1287.17315.

Gioia, D.G., Felizardo, L.K., and Brandimarte, P. (2022). Inventory management of vertically differentiated perishable products with stock-out based substitution. *IFAC-PapersOnLine*, 55(10), 2683–2688. doi:10.1016/j.ifacol.2022.10.115.

Gioia, D.G., Felizardo, L.K., and Brandimarte, P. (2023). Simulation-based inventory management of perishable products via linear discrete choice models. *Computers & Operations Research*, 157, 106270. doi:10.1016/j.cor.2023.106270.

Haijema, R. and Minner, S. (2016). Stock-level dependent ordering of perishables: A comparison of hybrid base-stock and constant order policies. *International Journal of Production Economics*, 181, 215–225. doi:10.1016/j.ijpe.2015.10.013.

Nogueira, F. (2014). Bayesian Optimization: Open source constrained global optimization tool for Python. URL https://github.com/bayesian-optimization/BayesianOptimization.

ÓRiordan, D. (1993). Retail gross margins: Some international comparisons. *International Journal of Retail & Distribution Management*, 21(4). doi:10.1108/09590559310145536.

Pan, X.A. and Honhon, D. (2011). Assortment planning for vertically differentiated products. *Production and Operations Management*, 21(2), 253–275. doi:10.1111/j.1937-5956.2011.01259.x.

Powell, W.B. (2022). *Reinforcement learning and stochastic optimization*. John Wiley & Sons, Nashville, TN.

Sainathan, A. (2013). Pricing and replenishment of competing perishable product variants under dynamic demand substitution. *Production and Operations Management*, 22(5), 1157–1181. doi:10.1111/poms.12004.

van Donselaar, K., Van Woensel, T., Broekmeulen, R., and Fransoo, J. (2006). Inventory control of perishables in supermarkets. *International journal of production economics*, 104(2), 462–472.

van Donselaar, K., Peters, J., de Jong, A., and Broekmeulen, R. (2016). Analysis and forecasting of demand during promotions for perishable items. *International Journal of Production Economics*, 172, 65–75. doi:10.1016/j.ijpe.2015.10.022.

Zabel, E. (1970). Monopoly and uncertainty. *The Review of Economic Studies*, 37(2), 205–219. doi:10.2307/2296413.