

# Joint Inventory and Distribution Strategy for Online Sales with a Flexible Delivery Option

Muzaffer Alim<sup>a,b,\*</sup>, Patrick Beullens<sup>c</sup>

<sup>a</sup>*Southampton Business School, University of Southampton*

<sup>b</sup>*Technology Faculty, Batman University, Turkey*

<sup>c</sup>*Mathematical Sciences and Southampton Business School and CORMSIS, University of Southampton*

---

## Abstract

This paper develops a strategy to jointly optimize the inventory and distribution for an online sales firm. The firm has to decide how to distribute the products from its warehouse to customers: this can either be done by using a company-owned vehicle, or by outsourcing to a third-party transportation company. The online sales environment includes a flexible delivery option that gives a discount to customers in return. This option is offered when the inventory level in the warehouse is lower than a threshold level. Customers accepting flexible delivery pay a deposit at the time they place the order and pay the remaining reduced price at the time of delivery. By offering the flexible delivery option, the firm aims to reduce the cost of distribution to the customers as well as postpone the timing of paying an outside supplier for stock replenishment. Additionally, this allows the firm to use on hand stock effectively to respond to more urgent customer requests. As the timing of cash-flows are dependent on the customer behaviour and the inventory and distribution strategy, the profit function is the Net Present Value of future cash-flows. We analyze the benefit of flexible delivery to the firm and perform sensitivity analysis with respect to various parameters. The profitability of flexible delivery depends on price setting and customer behaviour. Flexible delivery, in this model, has great potential to reduce transport distances and emissions when firms use their own vehicles.

*Keywords:* Continuous review inventory, Demand postponement, Net Present Value,

---

\*I am corresponding author

*Email addresses:* `ma4e13@soton.ac.uk` (Muzaffer Alim), `P.Beullens@soton.ac.uk` (Patrick Beullens)

## 1. Introduction

Online sales have been pioneered in the retail sector of products that can be transported as small parcels by postal or courier services. The strong competition this has generated has led to closures of many traditional stores in the high street. Online sales channels allow customers to save on the time and effort needed to select and order desirable products but do introduce a time lag between order placement and delivery. Can this disadvantage turn into an advantage for both customers and online firm?

The difference between customers' willingness to wait is recognized by several online retailers by offering customers a choice between delivery options, where price discounts are given when customers accept increased (uncertainty on) delivery time. Delivery options range from fast, say within 24 or 48 hours, to delivery at any time within a given time period of e.g. 7 days. We refer to the latter as a *flexible delivery option*. If greater flexibility to plan the delivery allows the online retailer to make costs savings, providing a financial incentive to customers giving the firm this increased planning flexibility could thus make economic sense.

The use of online sales channels has become increasingly important for businesses who sell larger-sized products. In the UK, for example, there now exist a fair number of online sales companies offering products and materials for home improvement projects. This may range from kitchen and bathroom units and appliances to materials for plumbing, electrics, and joinery projects, servicing both professional tradesmen and DIY enthusiasts. The bulkier or heavier products require transport in specialized (company-owned) vehicles for the delivery to customer locations. The model developed and analyzed in this paper is particularly aimed at achieving a better understanding of the impact of flexible delivery for these types of businesses.

The time lags between ordering and delivery will vary with the sector. In the online sales of books, delivery options are typically restricted to hours or days. In the home improvement

sector, fast delivery may be within one week, and a flexible delivery may be any time within several weeks. Another difference is that customers ordering home improvement products may, instead of having to pay the full price when ordering, only be required to pay a deposit. The residual payment due is then settled on the day of delivery. Given that this payment may occur several weeks later, it is thus of interest to model the impact of the timing of payments on the profitability of the online business.

The costs that arise from meeting online sales naturally depend on how the supply chain is organized. In this paper, we assume that the products are stored at the online firm's single depot, from which 'last-mile' transportation needs to be organized to customer locations within a sales region surrounding the depot<sup>1</sup>. See also Figure 1. The online retailer can make use of its own vehicle to make deliveries, or alternatively outsource to a third party for a fixed price per product (or unit of volume or weight). In addition to delivery costs, costs also arise from the inventory stored at the depot and order costs for the replenishment from an upstream supplier. In this paper, we aim to get more insight into how inventory and transportation costs can be minimized in this setting, and how this depends on other system characteristics such as the proportion of customers that can be induced to select flexible delivery.

As the value and timing of both revenue and cost cash-flows will be affected by the choices made by customers and the firm, the minimization of average costs only is not an appropriate objective. We develop, therefore, the model with the objective of maximizing the Net Present Value (NPV) of the firm's future cash-flows. The usefulness of this approach for the study of production-inventory systems was first demonstrated in Grubbström (1980), see also Beullens and Janssens (2011).

---

<sup>1</sup>This situation can also represent the case of a firm operating one single online sales website, but adopting a decentralized organization of its supply chain into disjoint service regions, where each sales region is serviced by a single depot.

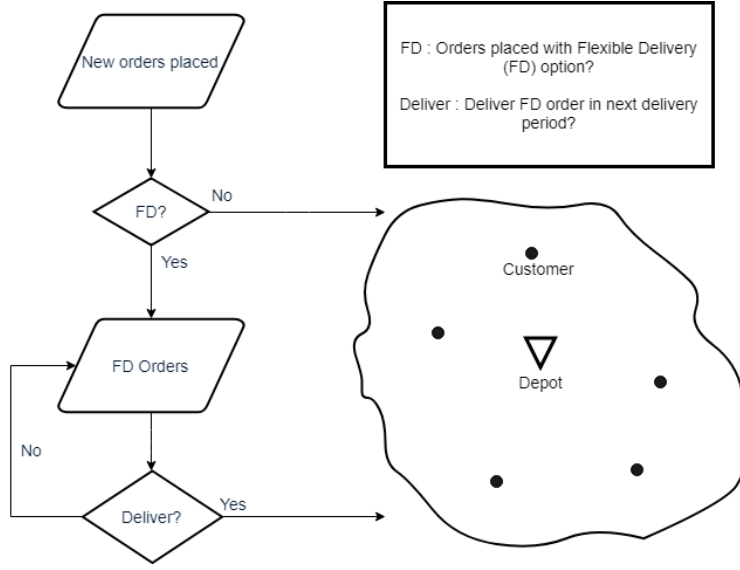


Figure 1: Process flow of order placement and delivery.

## 2. Related research

Considering the different ingredients discussed above, this study is of relevance to the research on supply chain inventory and distribution management, and in particular to the use of price incentives for shifting demand across time. We summarize the core literature in Table 1.

Pricing strategies are widely applied in supply chain literature for various purposes. In the area of revenue management, this approach has been studied in great detail, see e.g. Quante et al. (2009); Agatz et al. (2013). This area of research is primarily concerned with maximizing the revenue extractable from a capacitated resource with often a fixed cost structure. This area of research finds application in the booking of (airline, rail, ferry) travel, hotel rooms, and car rentals. Pricing strategies will affect both total demands as well as the timing of when demand is announced.

Many papers in the inventory literature consider pricing strategies as a mean to increase the total demand or speed up the demand rate, rather than shift demand, as in the studies on price-dependent demand by Venegas and Ventura (2018), Papachristos and Skouri (2003),

Table 1: Related Literature

Authors	Dependent Demand	Financial Incentive	Time Value of Money	Deposit	Demand Shifting	Delivery
Tiwari et al. (2018a)	P, DS	TC	✓	×	×	×
Li et al. (2017)	P, ED	TC	✓	✓	×	×
Venegas and Ventura (2018)	P	QD	×	×	×	×
Tiwari et al. (2018b)	P, DS	×	×	×	×	×
Feng et al. (2017)	P, DS, ED	×	✓	✓	×	×
Teng et al. (2016)	ED	×	×	✓	×	×
Zhang et al. (2014)	×	TC, QD	✓	×	×	×
Xu et al. (2017)	×	PD	×	×	AB	×
Zhuang et al. (2017)	×	PD	✓	×	AB	×
Wu and Wu (2015)	P	PD	×	×	DP	×
Lappas et al. (2017)	×	×	×	×	×	IRP
Archetti et al. (2017)	×	×	×	×	×	IRP
<b>This Paper</b>	×	PD	✓	✓	DP	✓

P:Price, DS:Displayed Stock, ED:Expiry Date, PD: Price Discount, TC:Trade Credit  
QD:Quantity Discount, AB:Advance Booking, DP:Demand Postponement  
IRP: Inventory Routing Problem

Wee (2001), Dye et al. (2007), Tiwari et al. (2018b) and quantity discounts by Zhang et al. (2014), Taleizadeh and Pentico (2014), Shin and Benton (2007), Wee (1999). Dynamic pricing in the presence of inventory considerations is another area of application extensively discussed in Elmaghraby and Keskinocak (2003).

Using financial incentives to cause demand to shift across time is an area of research that has been mainly analyzed within the contexts of advance booking (or pre-orders) and demand postponement, respectively. Advance booking programs entice customers to place their demand earlier by offering them (financial) incentives, see e.g. Tang et al. (2004) and Li and Zhang (2013). According to Xu et al. (2017), the success of online retail channels has made it easier for companies to obtain such advance bookings. The benefits of using advance booking systems include that it could greatly improve the demand forecast for new products, as in Tang et al. (2004), or could lead to a better match between demand and

capacity availability, as in Zhuang et al. (2017).

The purpose of demand postponement is to convince customers to shift delivery to a later date. The concept was first introduced in Iyer et al. (2003), who studied the value of postponement of a fraction of demand as to allow the firm to procure additional capacity and reduce overall stockouts. Wu and Wu (2015) use demand postponement to create a capacity buffer for urgent demand, leading to a reduced expected risk of shortages on urgent demand in addition to overall increased efficiency of capacity usage. Tang (2006) use the strategy as a means to mitigate supply chain disruptions.

The effect of customers choosing for the flexible delivery option in our model is comparable to demand postponement by enticing customers to switch from the fast delivery option towards a delivery at a later date. This study, however, differs from the previous literature as decisions need to be made as to when to deliver to which customer based on the joint consideration of inventory and distribution costs.

In literature, the integrated inventory and distribution (as routing) problems are classified as Inventory Routing Problems (IRP). We refer to Coelho et al. (2014) for a detailed review. Broadly speaking, the main aim in the IRP is to decide when to deliver how much to each customer (inventory management), and simultaneously to decide on the optimal allocation and sequencing of customer delivery tasks to individual vehicles (routing), while aiming to minimize the total inventory and vehicle routing costs over a time horizon. Some studies in the IRP also consider time windows for the delivery as decision variables (Lappas et al., 2017; Archetti et al., 2017) which allows the delivery to be made within a period also called flexible delivery. As most IRP formulations are NP-hard, it is known that both computational time and solution structure are sensitive to small variations in the data, making it a cumbersome task to derive general insights about the expected long-term performance of the system. Examining the potential of the flexible delivery option in the online system environment, when modelled into an IRP framework, would hence become a challenging task.

In this paper, we wish to develop an approach to help quantifying the strategic value of the flexible delivery option in an online sales context, and how this may depend on certain key features of the system. For this purpose we do not require knowledge about the exact locations of customer requests nor the exact routing solutions of the vehicle. Therefore, we will use in this paper the continuous approximation methodology, a descriptive (approximate) modelling technique developed by Carlos Daganzo, see also Daganzo (2005). The approach can give insight into the relative efficiency of different logistics system design options based on concise data summaries and with less computational efforts than the IRP framework, see also Daganzo and Newell (1985).

In seeking a good trade-off between model complexity and insight, we will also consider an inventory system as a single item with constant demand rate. Even when demand is variable from period to period, it is known since Peterson and Kelle (1985) that if the coefficient of variation remains below 0.2, then the constant demand rate assumption is a very good approximation.<sup>2</sup> Once a satisfactory policy with respect to discount-postponement and distribution has been identified, with for example the approach developed in this paper, then applying this policy into an IRP, which can more accurately account for actual demand patterns of individual products, would certainly allow for further fine-tuning and planning of day-to-day operations, and is an appealing yet challenging idea for further research.

Our model and analysis are also different from previous literature on demand postponement in that it considers the impact of the logistics strategy on the timing of payments in an NPV framework. Permissible delay in payment when a certain amount is ordered (Chung et al., 2005), selling price- dependent demand (Dye et al., 2007), trade credits (Tiwari et al., 2018a), inflation (Tiwari et al., 2016), advance cash credit (Li et al., 2017), deposit payment

---

<sup>2</sup>Furthermore, it is known that when ordering  $n$  products, each with their own demand rate  $y_i$  and unit holding cost  $h_i$ , and there is a fixed set-up cost per delivery, then it is optimal to always order these products jointly from the supplier, see also Silver et al. (1998). It is thus fair for strategic analysis to substitute  $n$  products ordered from the same supplier by a single hypothetical aggregate product, which will also reduce variability of demand.



(Teng et al., 2016) and backorders with a deposit (Ghiami and Beullens, 2016) are examples in which payment structures are shown to affect the inventory decisions.

The mechanism of price discounts in combination with deposit values we adopt in this study was first introduced in Ghiami and Beullens (2016), who used this mechanism in a production-inventory system with backorders. The impact of deposit value in that situation was limited. In our inventory-distribution model with the flexible delivery option, we find that its impact is more pronounced.

The paper is further organized as follows. Section 3 introduces the description of the system. In Section 4, the optimization problem is formulated. Section 5 develops properties and presents the algorithm. Section 6 provides insights derived from a variety of numerical experiments. Conclusions and further research areas are presented in Section 7.

### 3. Description of the System

#### 3.1. General characteristics

We consider a single product ordered by customers via an online store at a given constant demand rate  $y$ . Upon placing their order, customers may be given a choice between the fast (or normal) delivery option and a flexible delivery option. In the fast delivery option, the customer pays the full sales price  $p$  when placing the order. In the flexible delivery option, the customer receives a small discount  $r$  on the full sales price, but pays upon placing the order a pre-agreed deposit  $g$ , while paying the remainder  $p - r - g$  upon the time of actual delivery. We note that in most online systems we would have  $g = p - r$ . However, in applications where customers' financial trustworthiness can be verified, a reduced deposit value  $g < p - r$  may help in convincing more customers accepting the flexible delivery option, in particular in cases where the possible waiting time may become larger. Furthermore, we assume that more customers will choose the flexible delivery option the larger the discount value  $r$ . Part of the aims of the study is to investigate how system performance is affected

by various different choices for  $r$  and  $g$  values.

The performance of the system is further affected by decisions about how to place replenishment orders from an upstream supplier. In line with EOQ-type model characteristics, we consider that in steady-state a constant order quantity  $yT$  will be placed at equidistant moments in time, where  $T$  denotes the inventory cycle time measuring the time between two consecutive orders arriving in the retailer's warehouse. In addition, system performance is further determined by the choices made about how the demand orders are delivered to customer locations. The choice here is between performing the distribution by a company-owned vehicle, and the option of outsourcing the delivery to a third party. The company can also choose to outsource the delivery of only part of the demand and make this variable over time. The model constructed will maximize the Net Present Value of this system by choosing the optimal joint replenishment and distribution strategy.

## Parameters

$y$	Constant demand rate per unit time $y > 0$
$p$	Sales price per unit of the product $p > w$
$w$	Cost price of a unit product
$r$	Discount amount per unit of product, $0 \leq r < p - w$
$g$	Deposit paid per unit of product, $g < p - r$
$S$	Setup cost per replenishment
$I(t)$	Inventory level at time $t$
$U$	Maximum waiting time promised to customers
$\Delta t$	Time of a delivery period
$A$	Size of the total distribution region
$\rho$	Average distance from a customer in distribution region to the depot
$n(k)$	Number of customers need to be served in the $k^{th}$ delivery period
$d(k)$	Average distance travelled in the $k^{th}$ delivery period
$c(k)$	Total distribution cost of the $k^{th}$ delivery period
$c_u(k)$	Distribution cost per unit product in the $k^{th}$ delivery period when delivered by the company-owned vehicle
$c_o$	Transport cost per unit product charged by third-party
$\beta$	Ratio of demand accepting the flexible delivery option, $0 < \beta < 1$
$\alpha$	Opportunity cost of capital rate
$\gamma$	Vehicle operation cost per mile

### Decision Variables

- $T$  Inventory cycle time,  $T = K\Delta t$  with  $K$  integer,  $\Delta t \ll T$ ,  $T > 0$
- $FT$  Time period during which flexible delivery option is not available,  $0 \leq F \leq 1$
- $J(k)$  Binary variable to decide whether to use own vehicle (1) or not (0)

### 3.2. Inventory system

An important feature not yet discussed is that the flexible delivery option is not necessarily always made available to customers. That is, the flexible delivery option is only made available within a time period  $(1 - F)T$  of a replenishment inventory cycle. The reason for this is to allow the replenishment cycle  $T$ , should this be optimal, to become (much) larger than the maximum time  $U$  that customers may have to wait when choosing the flexible delivery option. To minimize waiting time for customers, it is best to have the flexible delivery option made available at the end of each inventory cycle<sup>3</sup>.

Figure 2 illustrates the inventory level as a function of time for the first inventory cycle. It is assumed that the system starts at time 0 with no outstanding demand orders. At time  $FT$ , the flexible delivery is introduced for a time  $(1 - F)T$ . During this period, it is assumed that customers choose the flexible option at a rate of  $\beta y$  (further discussed in Section 3.4). Delivery of these orders will be postponed until after the next replenishment order has arrived at time  $T$ . Note that the first replenishment order only needs to cover the demand during the first cycle  $T$  for those demands to be delivered with the normal delivery option. From the second cycle and onwards, the replenishment order also needs to cover

---

<sup>3</sup>This will also maximize the NPV of revenues and holding costs. The NPV of distribution costs alone (see Figure 3), however, would be maximized by starting each cycle with the flexible delivery time window instead. Distribution costs are typically small, however, so overall the adopted model in this paper will be financially better. For high values  $g = p - r \approx p$ , it may be better to have the first inventory cycle different, i.e. being of length  $T_1 \neq T$  in which the flexible delivery option is offered throughout. The second and all future inventory cycles would then be of length  $T$  and follow the structure as in Figure 3. We do not pursue this possibility in this paper.

the postponed demand but is otherwise identical to the first cycle. The system then repeats itself at infinitum.

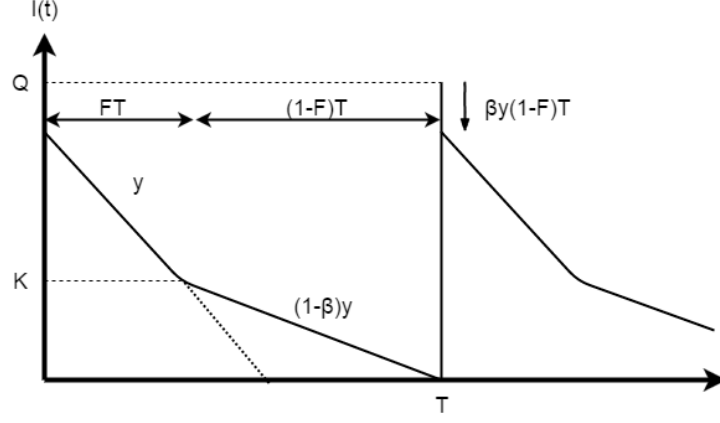


Figure 2: Inventory Flow with Constant  $\beta$

### 3.3. Delivery system

Distribution to customer locations is organized at discrete moments in time ( $\Delta t, 2\Delta t, \dots, k\Delta t, \dots$ ). Let  $n(k)$  denote the number of customers to be served in the delivery period  $k$  and  $n(1)$  is the number of customers at the first delivery.

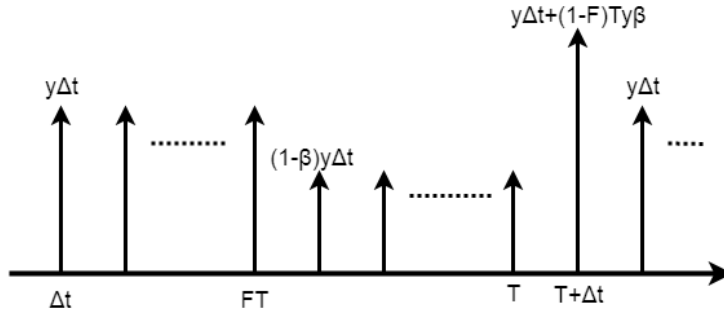


Figure 3:  $n(k)$  values within an inventory cycle

Fig. 3 illustrates the number of customers to be served relative to the inventory cycle time (from Fig. 2). When the delivery occurs at a time when we do not offer the flexible delivery option, the number of demands to be delivered equals  $y\Delta t$  which is due to the regular demand rate. If the delivery occurs in the time interval during which flexible delivery is

offered, the demand rate reduces to  $(1 - \beta)y$  and the number of deliveries is  $(1 - \beta)y\Delta t$ . When it is in the interval of  $(FT, FT + \Delta t)$ , then both types of demand (not postponed and postponed) have to be distributed at the same delivery. Also, at the first delivery moment, after any replenishment order has arrived (except for the first cycle shown by  $n(1)$ ), the delivery of postponed demand from the previous cycle has to occur. Therefore the initial delivery after the order arrival serves usual customers plus the postponed customers. We can thus quantify  $n(k)$  and  $n(1)$  as follows (with mod indicating the modulo operator):

$$n(1) = \begin{cases} y\Delta t & \text{for } \Delta t \leq FT \\ y\Delta t - y\beta(\Delta t - FT) & \text{for } FT < \Delta t < FT + \Delta t \\ (1 - \beta)y\Delta t & \text{for } FT + \Delta t \leq \Delta t \leq T \end{cases} \quad (1)$$

$$n(k) = \begin{cases} y\Delta t & \text{for } \Delta t < k\Delta t \mod T \leq FT \\ y\Delta t - y\beta(k\Delta t \mod T - FT) & \text{for } FT < k\Delta t \mod T < FT + \Delta t \\ (1 - \beta)y\Delta t & \text{for } FT + \Delta t \leq k\Delta t \mod T \leq T \\ n(1) + (1 - \beta)Ty\beta & \text{for } k\Delta t \mod T = \Delta t \text{ and } k > 1 \end{cases} \quad (2)$$

For the delivery of  $n(k)$ , the company can choose between using a company-owned vehicle, or outsourcing to a third party. When the company makes its own delivery, the distribution cost includes fuel, driver wages, vehicle depreciation and (un)loading costs. Among these, we focus on the costs directly related to the travelled distance. The average distance travelled in the  $k^{th}$  delivery is approximated in (3). This model is derived from the continuous approximation method developed in Daganzo and Newell (1985), where  $\zeta(n(k)) \approx 0.57$  and  $\gamma$  is the vehicle operation cost per mile including fuel costs and driver's wages (at average vehicle speed). The main assumption underlying (3) is that the  $n(k)$  customer locations are

independently drawn from a uniform distribution over the service region of area  $A$ . The first term measures twice the average distance  $\rho$  between the warehouse and a customer location. The second term measures the shortest distance to visit all  $n(k)$  customer locations. The corresponding unit distribution cost is  $c_u(k)$  as in (4).

$$d(k) \approx 2\rho + \zeta(n(k))\sqrt{An(k)} \quad (3)$$

$$c_u(k) \approx \frac{\gamma d(k)}{n(k)} \quad (4)$$

If the warehouse is located inside the service region, Danganzo's formula can be simplified to  $d(k) \approx \zeta\sqrt{A(n(k) + 1)}$ , where  $\zeta \approx 0.71$ . Without much loss of generality, we will adopt this setting here.

An external distribution partner is offering a cost of  $c_o$  per unit item for delivery. The unit delivery cost of the company's own vehicle depends on the demand intensity across the service area. If the unit cost is less than the outsourcing cost then the company delivers the product itself. The decision variable  $J(k)$  is thus as given by (5), and the total distribution cost of the  $k^{th}$  delivery as in (6).

$$J(k) = \begin{cases} 1 & \text{for } c_u(k) < c_o \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

$$c(k) = n(k) \left[ J(k)c_u(k) + (1 - J(k))c_o \right] \quad (6)$$

### 3.4. Customer's acceptance of flexible delivery

It is in line of expectation that more customers will be induced to accept the flexible delivery option the higher the discount and the lower the maximum waiting time and deposit value. Pentico et al. (2015) and Tiwari et al. (2016) formulate the customer responses for

the partial backlogging as inversely proportional with the waiting time. Similarly, Wu et al. (2014) formulate the demand rate as dependent on the downstream credit period while Feng et al. (2017) assume that demand rate is proportional with the retailer selling price. All the studies above formulate the demand rate as an exponential function. We will examine the case that the fraction  $\beta$  is influenced by the discount amount  $r$ , maximum waiting time  $U$ , and deposit value  $g$  according to the following relationship:

$$\beta(r, U, g) = e^{-\frac{\theta(U, \frac{g}{p-r})}{r}}, \quad (7)$$

where  $\theta(U, g/(p-r))$  is a factor that measures the *customer resistance* to the discount offer, which increases with  $U$  and  $g/(p-r)$ . According to (7),  $\beta$  grows faster and faster towards 1 with the increase in  $r$ , and approaches 1 in the limit for large  $r$ . Note, however, that  $r$  will always remain finite as  $r < p - g$ . For given  $U$ ,  $g$  and  $r$  values,  $\beta(r, U, g) = \beta$  is thus assumed to be a constant.

#### 4. Model Development

As the value and the timing of revenues and cost cash-flows are affected by the choices made by customers and the firm, we develop the model with the objective to maximize the Net Present Value of the firm's future cash-flows. We summarize the main modelling assumptions:

- Demand occurs at a constant rate and individual customer orders are unit sized.
- Demands which accept the discounts are postponed to the first delivery after the replenishment order from the supplier is received.
- Customers who choose fast delivery pay the full price upon placing the order online.



- Customers who accept flexible delivery pay a deposit when placing the order, but pay the discounted remainder at the delivery time epoch in which the delivery is made.
- Initial inventory is zero and the first replenishment occurs at  $t = 0$ .
- There is no constraint on the replenishment order size and no quantity discounts are offered by the external supplier.
- The replenishment lead-time is zero (or any non-zero constant value).
- The supplier is paid upon the arrival of the replenishment order in the firm's depot <sup>4</sup>.
- Out-of-pocket holding cost is assumed small and will be ignored but we consider the opportunity cost of capital invested in the stock.
- The retailer depot is located within the customer service region.
- When the third-party transportation company is used to deliver in a period, the cost incurred is paid out at the end of that delivery period <sup>5</sup>.
- When using the company-owned vehicle for transport in a delivery period, costs are incurred at the end of the delivery period.
- Continuous compounding is used for the net present value analysis.
- The time horizon is infinite.

---

<sup>4</sup>This is in line with the implicit assumption adopted in classic inventory models, see Beullens and Janssens (2014), who also show how to adapt the model to incorporate different payment structures. If, for example, the supplier wishes to be paid a deposit of  $\epsilon w$  per unit of product ordered, and the lead-time is  $L$ , then the model developed in this paper is valid by replacing the unit cost  $w$  by  $\epsilon w e^{\alpha L} + (1 - \epsilon)w$ . This will make the products more costly to hold in inventory and may affect the numerical values of optimal decision variables and annuity stream profits, but will not affect the main insights in this paper about the trade-offs analyzed.

<sup>5</sup>See also the previous footnote.

#### 4.1. Inventory level and order quantity

For  $\Delta t \ll T$ , the inventory level  $I(t)$  as a function of time over one inventory cycle follows with sufficient accuracy the pattern as in Figure 2. The change of inventory level over any cycle can thus be described by the following differential equation:

$$\frac{dI(t)}{dt} = \begin{cases} -y & \text{for } 0 \leq t \leq FT \\ -(1 - \beta)y & \text{for } FT \leq t \leq T \end{cases} \quad (8)$$

the solution of which is given by:

$$I(t) = \begin{cases} I(0) - yt & \text{for } 0 \leq t \leq FT \\ I(FT) - (1 - \beta)y(t - FT) & \text{for } FT \leq t \leq T \end{cases} \quad (9)$$

Since backorders are not allowed, we must have  $I(t) \geq 0$ . As it is unnecessary to keep stock that is not used,  $I(T) = 0$  for each inventory cycle. In the first cycle, the order quantity  $Q_1 = I_1(0)$  needs to cover the demand that needs to be fulfilled in that cycle:

$$Q_1 = yFT + (1 - \beta)y(1 - F)T. \quad (10)$$

The first term is the demand in the period  $(0, FT)$ , and the second term is the demand in period  $(FT, T)$ . In all subsequent cycles, the order quantity  $Q = I(0)$  covers the amount as in (10) plus the postponed demand from the previous cycle  $\beta y(1 - F)T$ , which gives:

$$Q = Q_1 + \beta y(1 - F)T = yT. \quad (11)$$

#### 4.2. Annuity streams of cash-flows

Since the time horizon is infinite, the appropriate objective function is given by the Annuity Stream (AS) of the relevant incoming and outgoing cash-flows. The relevant cash-flows for the revenues and inventory costs are shown in Fig.4. (The cash-flows of the distribution system are not shown in this figure.)

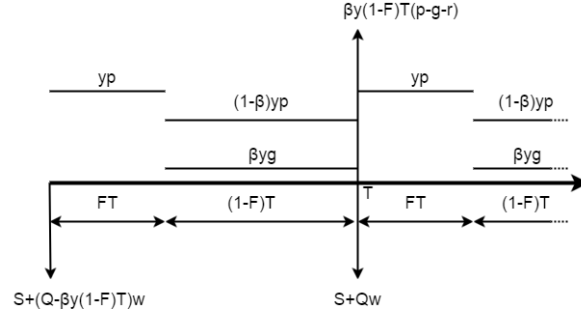


Figure 4: Cash-flows of revenue streams and inventory system costs

At the start of each inventory cycle, the firm incurs a set-up cost  $S$  independent of order size. In all but the first order cycle, the firm has to pay  $wQ = wyT$  for the products replenished, see (11). The total replenishment cost is then  $C = S + wyT$ . For the first order cycle, since  $Q_1 = Q - \beta y(1 - F)T$ , the total replenishment cost can be written as:

$$C_1 = S + w(Q - \beta y(1 - F)T) = C - C_0^-. \quad (12)$$

The annuity stream of replenishment costs are thus given by:

$$\begin{aligned} ASC &= \sum_{i=0}^{\infty} (\alpha C e^{-i\alpha T}) - \alpha C_0^- = \alpha C \frac{1}{1 - e^{-\alpha T}} - \alpha C_0^- \\ &\approx C \left( \frac{1}{T} + \frac{\alpha}{2} \right) - \alpha C_0^- \end{aligned} \quad (13)$$

The approximated value on the right-hand-side is derived from the Maclaurin expansion of the exponential terms in  $\alpha T$ , where second and higher orders of  $\alpha$  are ignored.

The postponed demand from previous periods is satisfied by on-hand inventory and these customers pay the discounted and deposit-subtracted price as given in (14). Not having postponed demand at  $t = 0$ , this revenue stream starts at  $t = T$ , thus the annuity stream value is as given in (15).

$$R_1 = \beta y(1 - F)T(p - g - r) \quad (14)$$

$$\begin{aligned} ASR_1 &= e^{-\alpha T} \left( \sum_{i=0}^{\infty} \alpha R_1 e^{-i\alpha T} \right) = \alpha R_1 \frac{e^{-\alpha T}}{1 - e^{-\alpha T}} \\ &\approx R_1 \left( \frac{1}{T} - \frac{\alpha}{2} \right) \end{aligned} \quad (15)$$

Between 0 and  $FT$  in each inventory cycle, sales occur at demand rate  $y$  for a price of  $p$ . The present value of this revenue for the first cycle is given in (16) and the equivalent annuity stream (from all cycles) in (17).

$$R_2 = \int_0^{FT} pye^{-\alpha t} dt = \frac{py}{\alpha} \left( 1 - e^{-\alpha FT} \right) \quad (16)$$

$$\begin{aligned} ASR_2 &= \sum_{i=0}^{\infty} \alpha R_2 e^{-i\alpha T} = py \left( \frac{1 - e^{-\alpha TF}}{1 - e^{-\alpha T}} \right) \\ &\approx py \left( F - \frac{\alpha TF^2}{2} + \frac{\alpha TF}{2} \right) \end{aligned} \quad (17)$$

At time  $FT$ , the flexible delivery option becomes available. The demand rate requiring fast delivery reduces to  $(1 - \beta)y$  and generates a revenue stream at the price  $p$ . The demand rate  $\beta y$  corresponding to the flexible delivery customers generates a deposit  $g$  in this interval. Present value of this revenue between  $FT$  and  $T$  is given in (18), and corresponding annuity stream function in (19).

$$\begin{aligned}
R_3 &= \int_{FT}^T ((1-\beta)yp + \beta yg) e^{-\alpha t} dt \\
&= \frac{((1-\beta)yp + \beta yg)}{\alpha} \left( e^{-\alpha FT} - e^{-\alpha T} \right)
\end{aligned} \tag{18}$$

$$\begin{aligned}
ASR_3 &= \sum_{i=0}^{\infty} \alpha R_3 e^{-i\alpha T} = ((1-\beta)yp + \beta yg) \left( \frac{e^{-\alpha TF} - e^{-\alpha T}}{1 - e^{-\alpha T}} \right) \\
&\approx ((1-\beta)yp + \beta yg) \left( 1 - F + \frac{\alpha TF^2}{2} - \frac{\alpha TF}{2} \right)
\end{aligned} \tag{19}$$

The distribution cost over time depends on the number of customers served at each delivery  $n(k)$ , and on the choice of distribution strategy, producing different possible values for  $c(k)$  as given by (6). The structure of the cash-flows of the distribution system is given by the reflection about the y-axis of Figure 3, and replacing the  $n(k)$  by  $c(k)$  values. The present value of the distribution cost in every cycle but the first is thus given by (20). In the first cycle, we do not have postponed demand from previous periods. So instead of delivering a postponed demand, we only need to consider  $c(1)$ . This correction is used in the annuity stream function as given by (21).

$$C_{dist} = c\left(\frac{T}{\Delta t} + 1\right)e^{-\Delta t} + \sum_{k=2}^{T/\Delta t} c(k)e^{-k\Delta t} \tag{20}$$

$$\begin{aligned}
ASC_{dist} &= \sum_{i=0}^{\infty} \alpha C_{dist} e^{-i\alpha T} - \alpha \left( c\left(\frac{T}{\Delta t} + 1\right) - c(1) \right) e^{-\Delta t} \\
&\approx C_{dist} \left( \frac{1}{T} + \frac{\alpha}{2} \right) - \alpha c\left(\frac{T}{\Delta t} + 1\right) e^{-\Delta t} + \alpha c(1) e^{-\Delta t}
\end{aligned} \tag{21}$$

The annuity stream of total profit over an infinite time horizon is the sum of all revenues

minus the setup, unit and distribution costs:

$$ASTP(F, T) = ASR_1 + ASR_2 + ASR_3 - ASC - ASC_{dist}$$

Our objective is to maximize the annuity stream of the total profit while keeping maximum waiting time under a certain limit.

$$\begin{aligned} & \text{maximize} && ASTP(F, T) \\ & \text{subject to} && (1 - F)T \leq U \end{aligned}$$

## 5. Properties and algorithm

To facilitate further analysis as well as algorithm design, we present some properties related to the impact of flexible delivery on the replenishment pattern and distribution strategy, respectively.

### 5.1. Special Case: Inventory profit only

In this section, we only consider the inventory system. This situation may arise in the case where the firm has outsourced the delivery to a third party for an agreed total price based on an agreed annual demand volume and paid out according to a fixed payment structure that is independent of the actual delivery schedules. Flexible delivery will then only affect the replenishment strategy and we can focus on only the inventory system related profits arising from the cash-flows depicted in Figure 4. The relevant profit function is given by  $ASTP$  as above but where  $ASC_{dist}$  is a constant and can thus be ignored.

**Lemma 1.** *The objective function  $ASTP(F, T)$  is a concave function under positive setup cost.*

Based on Lemma 1, we conclude that the  $F$  and  $T$  values which make the partial derivatives zero are optimal.

**Proposition 1.** *Let the  $T_{F \leq 1}$  and  $T_{F=1}$  be the optimal  $T$  values in the NPV model where the flexible delivery option is offered and not offered, respectively. Then:*

$$T_{F \leq 1} \geq T_{F=1}$$

Proposition 1 shows that the cycle time for the model without a flexible delivery being offered is a lower bound for the model with the flexible delivery option.

**Lemma 2.** *For the flexible delivery option to be profitable, the discount amount should satisfy the condition:*

$$r \leq \frac{\sqrt{\frac{8S\alpha w}{y}}}{2 - \sqrt{\frac{2S\alpha}{yw}}} \quad (22)$$

Lemma 2 calculates the value of  $r$  which makes  $F = 1$ . This indicates that when the discount amount reaches this level, it is no longer profitable to offer the flexible delivery option.

Note that the above results are only valid when the distribution cost is ignored.

## 5.2. Determining the distribution strategy

In principle, the distribution strategy needs to decide on the optimal choice of  $c(k)$  for  $k = 1, 2, 3, \dots$ . For any values of  $T$  and  $F$ , however, the sum of distribution costs as given by (20) within an inventory cycle is of a structure in which we recognise up to four different demand intensity situations, see (2) and Figure 3. The unit distribution cost for the company when it would execute the delivery by itself is decreasing by the demand intensity on the service area. We can see this from (4): the numerator increases with the square root of  $n(k)$ , while the denominator increases by  $n(k)$ .

Let  $c_u^1$  and  $c_u^2$  be the minimum and maximum value of the unit distribution cost when using the companies' own vehicle, for any values of  $T$  and  $F$ . Given the above,  $c_u^1$  is thus the unit distribution cost that would arise from servicing the highest possible demand intensity. As seen from Figure 3, the highest demand intensity would arise in the period just after the replenishment:

$$y\Delta t + (1 - F)Ty\beta \leq y\Delta t + Uy\beta, \quad (23)$$

where the upper bound follows from constraint  $(1 - F)T \leq U$ . Likewise,  $c_u^2$  would arise in the period when flexible delivery is offered and when the demand level is:

$$(1 - \beta)y\Delta t. \quad (24)$$

**Lemma 3.** *For any values of  $T$  and  $F$ :*

1. *Distribution is fully outsourced if*

$$c_o \leq c_u^1 = \gamma\zeta \sqrt{\frac{A}{y\Delta t + Uy\beta}} \quad (25)$$

2. *Distribution is always conducted by the company if*

$$c_o \geq c_u^2 = \gamma\zeta \sqrt{\frac{A}{(1 - \beta)y\Delta t}} \quad (26)$$

3. *Distribution is partially outsourced if*

$$c_u^1 \leq c_o \leq c_u^2, \quad (27)$$

where  $J(k)$  is decided according to (5) for every  $k$  in an inventory cycle with given  $F$  and  $T$  values.



Lemma 3 will be used in the solution algorithm.

### 5.3. General Case: Solution Algorithm

The decision variables are  $T$  and  $F$ , and  $J(k)$  as given by (5). The main Algorithm 2 performs an exhaustive search over  $T$  and  $F$ , and calls Algorithm 1 which decides on the optimal distribution strategy  $J(k)$  over an inventory cycle of length  $T$  and given  $F$  and such that it maximises the objective function ASTP.

The algorithm assumes that  $T$  is restricted to an integer multiple of  $\Delta t$ . If we have the fast delivery distribution every  $x$  days, then  $\Delta t = x/365$ , and a maximum waiting time of three weeks corresponds to  $U = 21/365$ , for example. We restrict the search for an optimal  $T$  value to a maximum of  $1/\Delta t$ , or one year, while the search for optimal  $F$  values over 100 possible fractions of  $T$ . These settings can be easily adjusted if a more refined search is desired. However, the algorithm implemented in C++ on a normal PC solves all numerical examples we investigated at a sufficient level of accuracy while keeping computational time at less than 1 second for each instance.

If the model only needs to consider the inventory system, then:

$$T_{EOQ} = \lfloor \frac{T_{F=1}}{\Delta t} \rfloor,$$

due to Proposition 1, and the call to  $ATSP(F, T)$  in Algorithm 2 simply needs to evaluate this function. In case the inventory-distribution system is to be optimised,  $T_{EOQ} = 1$  and Algorithm 1 is used to evaluate  $ATSP(F, T)$ .

---

**Algorithm 1:** Calculation of Profit Function

---

```
1: return  $ASTP(F, T)$ 

2: if  $c_o \leq c_u^1$  then
3:    $J(k) = 0 \quad \forall k \in [1, \frac{1}{\Delta t}]$ 
4: else if  $c_o \geq c_u^2$  then
5:    $J(k) = 1 \quad \forall k \in [1, \frac{1}{\Delta t}]$ 
6: else
7:   for  $k = 1$  to  $\frac{1}{\Delta t}$  do
8:     if  $c_o \leq c_u(k)$  then
9:        $J(k) = 0$ 
10:    else
11:       $J(k) = 1$ 
12:    end if
13:  end for
14: end if

15: Calculate  $ASTP(F, T)$  with given  $F, T$  and  $J(k)$  values
```

---

---

**Algorithm 2:** Exhaustive Search Algorithm for NPV Model

---

```
1: return  $T^*, F^*, \maxprofit$ 

2: initialize;  $\maxprofit, T_{EOQ}, U$ 

3: for  $k \geq \frac{T_{EOQ}}{\Delta t}$  to  $\frac{1}{\Delta t}$  do

4:   for  $j = 0$  to 100 do

5:      $T = k\Delta t$ 

6:      $F = 0.01j$ 

7:     if  $(1 - F)T \leq U$  then

8:       Call  $ASTP(F, T)$ 

9:       if  $ASTP(F, T) \geq \maxprofit$  then

10:         $T^* = T$ 

11:         $F^* = F$ 

12:         $\maxprofit = ASTP(F, T)$ 

13:      end if

14:    end if

15:  end for

16: end for
```

---

## 6. Numerical Experiments

We examine the impact of the flexible delivery option and choice of distribution strategy on the performance of the inventory-distribution system for a set of instances. Parameters are set at the following value ranges:  $y = 1,000$  per year;  $S = 50$  or  $80$ ;  $p = 2w, 1.5w$ , or  $1.3w$ ;  $\alpha = 0.2$  or  $0.1$  (per year);  $\Delta t = 4/365$  (fast delivery within 4 days);  $U = 3$  or  $6$  weeks (maximum waiting time for flexible delivery option). Other parameters are set as further specified.

First, we examine the special case of only considering inventory profits. We then consider the joint inventory and distribution system in which we optimally decide when to outsource

the distribution. We end with the situation in which outsourcing is not available.

### 6.1. Special case: Inventory profits only

We consider here the special case introduced in Section 5.1. We compare the performance of a system offering flexible delivery ( $F \leq 1$ ) to a system in which this option is not available ( $F = 1$ ). In each case, we use the model to determine the optimal inventory system strategy that maximizes the AS profits  $ASTP$  for the firm. The percentage difference is calculated as:

$$\delta = 100(ASTP_{F \leq 1} - ASTP_{F=1})/ASTP_{F=1} \quad (28)$$

The higher value of  $\delta$ , the higher the benefit obtained from using the flexible delivery option. The difference is tested with various customers reactions and setup costs. Although we consider  $r = 1\%p$ , its maximum value  $r_{max}$  beyond which the flexible delivery option cannot be profitable is also calculated (Lemma 2). A summary of results is given in Table 2.

Table 2: Impact of flexible delivery ( $p = 1.3w$ ,  $w = 10$ ,  $r = 0.01p$ ,  $g = 0$ )

Parameters	$F = 1$		$F \leq 1$ (U=3 weeks)			$F \leq 1$ (U=6 weeks)			$r_{max}(\%p)$
	T	$\beta$	T	F	$\delta(\%)$	T	F	$\delta(\%)$	
$\alpha = 0.1, S = 50$	0.32	0.1	0.32	0.82	0.10	0.32	0.63	0.16	2.47
		0.5	0.33	0.82	0.54	0.35	0.66	0.86	
		0.9	0.35	0.83	0.99	0.40	0.70	1.63	
$\alpha = 0.1, S = 80$	0.40	0.1	0.41	0.86	0.13	0.40	0.70	0.23	3.13
		0.5	0.42	0.86	0.67	0.44	0.73	1.16	
		0.9	0.42	0.86	1.22	0.46	0.74	2.15	
$\alpha = 0.2, S = 50$	0.22	0.1	0.23	0.74	0.26	0.23	0.48	0.36	3.51
		0.5	0.23	0.74	1.30	0.26	0.54	1.95	
		0.9	0.26	0.77	2.37	0.29	0.59	3.74	
$\alpha = 0.2, S = 80$	0.28	0.1	0.28	0.79	0.31	0.29	0.59	0.50	4.47
		0.5	0.30	0.80	1.60	0.31	0.62	2.57	
		0.9	0.30	0.80	2.90	0.33	0.64	4.80	

As can be observed, the customers' acceptance rate  $\beta$  greatly affects the benefits of flexible delivery. Although customers do not pay deposits ( $g = 0$ ) and thus delay paying

until delivery takes place, increased profits still result from adopting a larger maximum delivery window  $U$ . Profits increase at higher values of the opportunity cost of capital  $\alpha$ . With higher set-up costs, cycle times  $T$  increase but the waiting time constraint  $U$  will limit the period in which the flexible delivery is offered.

Fig. 5 and Fig. 6 show the impact on the relative performance of the flexible delivery option system at different marginal profit values, when  $\alpha = 0.2$ ,  $S = 80$ ,  $w = 10$ ,  $g = 0$ , and  $r = 0.01p$ . The results clearly show that the flexible delivery option is more effective for low profit margins and is then also more sensitive to the maximum waiting time.

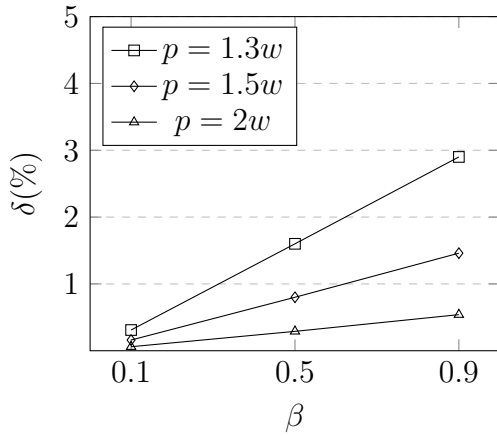


Figure 5: U=3 weeks

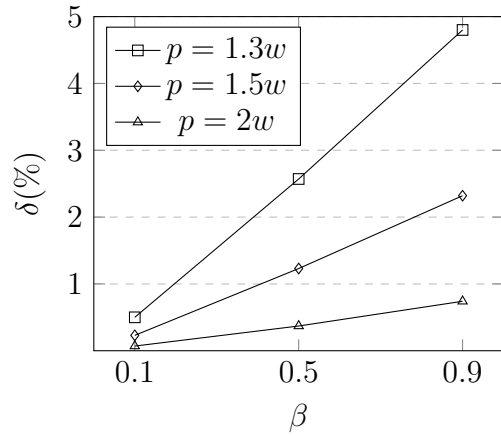


Figure 6: U=6 weeks

Another financial setting that can be examined in an NPV model is the impact of the deposit amount  $g$ . In many online sales environments, it is reasonable to assume that customers pay the full discounted price at the time of placing the order, which corresponds in our model with a deposit value  $g = p - r$ . It is intuitively clear that profits will improve with higher deposit values and in particular for larger  $U$  values. Table 3 presents results which show the impact of deposit value on system performance that confirms this.

## 6.2. General case: Inventory-Distribution system

In this section, we consider the general inventory-distribution system in which the company also determines the optimal distribution strategy. We calculate the impact of the

Table 3: Impact of deposit value ( $S = 80$ ,  $\alpha = 0.2$ ,  $p = 1.3w$ ,  $w = 10$ ,  $r = 0.01p$ )

Parameters	$\beta$	$F \leq 1$ (U=3 weeks)			$F \leq 1$ (U=10 weeks)		
		T	F	$\delta(\%)$	T	F	$\delta(\%)$
$g = 0$	0.1	0.28	0.79	0.31	0.29	0.40	0.56
	0.5	0.30	0.80	1.60	0.34	0.45	3.13
	0.9	0.30	0.80	2.90	0.37	0.49	6.15
$g = 0.5(p - r)$	0.1	0.28	0.79	0.34	0.29	0.32	0.90
	0.5	0.30	0.80	1.75	0.32	0.38	4.73
	0.9	0.30	0.80	3.18	0.35	0.43	8.85
$g = p - r$	0.1	0.28	0.79	0.38	0.29	0.35	1.20
	0.5	0.30	0.80	1.91	0.31	0.39	6.08
	0.9	0.30	0.80	3.47	0.31	0.39	11.03

flexible delivery option on the distribution costs as follows:

$$\delta_{dist} = 100(ASC_{dist}(F = 1) - ASC_{dist}(F \leq 1)/ASC_{dist}(F = 1), \quad (29)$$

where  $F = 1$  corresponds to the case in which the flexible delivery option is not offered. In each case, we use the model to find the joint optimal inventory-distribution strategy that maximizes the  $ASTP$  value and use the corresponding values obtained as given by (21) in the evaluation of (29).

Table 4 shows how flexible delivery affects distribution costs on the chosen instance at six different levels of the distribution cost  $c_o$  charged by the third party.

When  $c_o = c_u^1$ , the distribution is always outsourced (see Lemma 3). At first sight, one may perhaps not expect any improvements in the distribution costs from flexible delivery in that case. However, since the flexible delivery option affects the timing when we pay the third party for its distribution tasks, the NPV model does still show a fair benefit occurring. For example, with  $U = 3$  weeks, we would still expect a benefit of 2.5% from the flexible delivery option.

When  $c_o = c_u^2$ , distribution is never outsourced (Lemma 3). The benefits of flexible delivery now become larger. This can be expected since we will not only have a benefit from

postponing distribution costs (as in the case of  $c_o = c_u^1$ ), but also will reduce the actual average unit distribution cost per demand delivery over a cycle.

In the case of  $c_o = c_u(k)$ , we assume that  $c_o$  equals the company's unit delivery cost when delivering to fast delivery customers in periods in which the flexible delivery option is not offered. In this situation, we expect that only the deliveries including postponed demand are fulfilled by the company, while the delivery of normal demand customers in other periods can be outsourced. The results in Table 4 now show that flexible delivery will the highest potential to improve the firm's distribution costs, with values ranging between 8% and 22%.

Table 4: Impact of flexible delivery on distribution AS costs ( $S = 80, \alpha = 0.2, w = 10, r = 0.01p, \beta = 0.5, g = p - r$ )

Exp.	Unit Cost	Dist. Strategy	U=3 weeks			U=6 weeks		
			T	F	$\delta_{dist}$	T	F	$\delta_{dist}$
1	$c_o = c_u^1 = 0.05w$	Outsource	0.30	0.80	2.5	0.29	0.59	5.0
	$c_o = c_u^1 = 0.10w$		0.30	0.80	2.5	0.29	0.59	5.0
2	$c_o = c_u^2 = 0.05w$	Local delivery	0.30	0.80	3.3	0.29	0.59	7.6
	$c_o = c_u^2 = 0.10w$		0.27	0.78	3.5	0.26	0.54	8.4
3	$c_o = c_u(k) = 0.05w$	Partial Outsource	0.26	0.77	8.3	0.26	0.54	17.8
	$c_o = c_u(k) = 0.10w$		0.26	0.77	8.3	0.21	0.43	21.9

We now consider the impact of flexible delivery on the total system AS profits. The relative improvement from flexible delivery is calculated as in (28), but now including the variable distribution cost. Figure 7 illustrates the percentage improvement as a function of  $r$  for a setting that corresponds to the setting in the bottom row of Table 3, while the distribution cost parameters are set as the bottom row in Table 4, case  $c_o = 0.05w$ . Flexible delivery in a jointly optimized inventory-distribution system performs much better than when only accounting for inventory profits. At  $r = 1\%p$  and  $\beta = 0.5$ , for example, flexible delivery could improve the inventory profits by 1.91% (see Table 3), while in the jointly optimized inventory-distribution system this becomes 4.47%. Figure 7 also shows that, while a higher  $\beta$  brings more benefit, the maximum profitable discount value  $r$  increases only modestly with  $\beta$ .

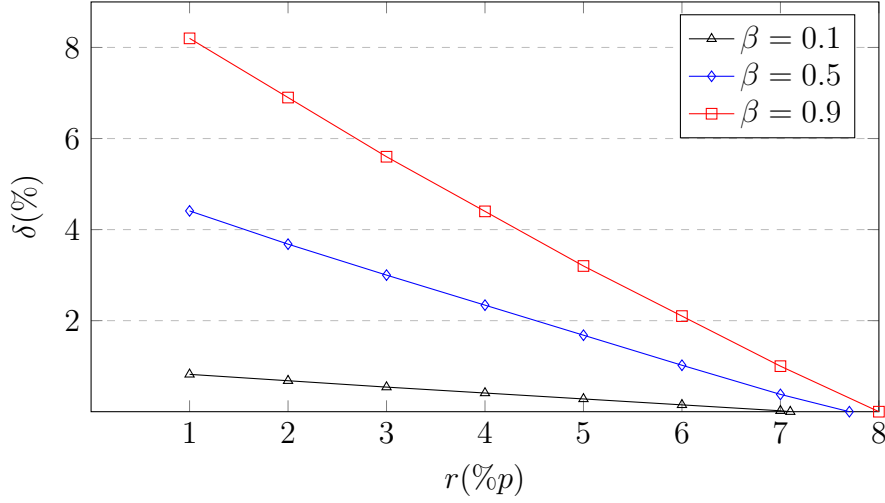


Figure 7: Impact of flexible delivery on total AS profits ( $p = 1.3w$ ,  $g = p - r$ ,  $S = 80$ ,  $c_u = 0.05w$ ,  $\alpha = 0.2$ ,  $U = 3$  weeks)

Figure 7 also illustrates that it may not always be worthwhile to offer more discount in order to get more customer acceptance of the flexible delivery option. Instead of offering 5% discount to get  $\beta = 0.9$ , for example, the  $\beta = 0.5$  case with 2% discount would achieve a higher overall profit for the firm. Firms wanting to apply the flexible delivery policy should thus consider the trade-off between customer acceptance and incentive structure. A low flexible delivery acceptance with low discount amount can sometimes be better than a higher acceptance with higher discounts.

In order to illustrate this trade-off more clearly, we end this section with an experiment in which we make use of the customer acceptance function defined in Section 3.4. We use the same parameter settings as used in Figure 7, but now let  $\beta$  be a function of  $r$  for three different values of the resistance function  $\theta = 0.1, 0.2, 0.5$  (for  $U = 3$  weeks and  $g = p - r$ ). We can now use the algorithm repeatedly over a range of  $r$  values to produce the results as illustrated in Fig. 8. As shown, the optimal discount would then be 2%, 3%, 4% when  $\theta = 0.1, 0.2, 0.5$ , respectively. At these optimal discount rates, customer acceptance is at the level  $\beta = 0.68, 0.60, 0.38$ , respectively (evaluate (7)). The higher the customer resistance, the higher the optimal discount value but the lower the customer acceptance rate and the



lower the total profit increase achievable from flexible delivery. When applying the flexible delivery option, firms should thus also aim to determine the customer acceptance function.

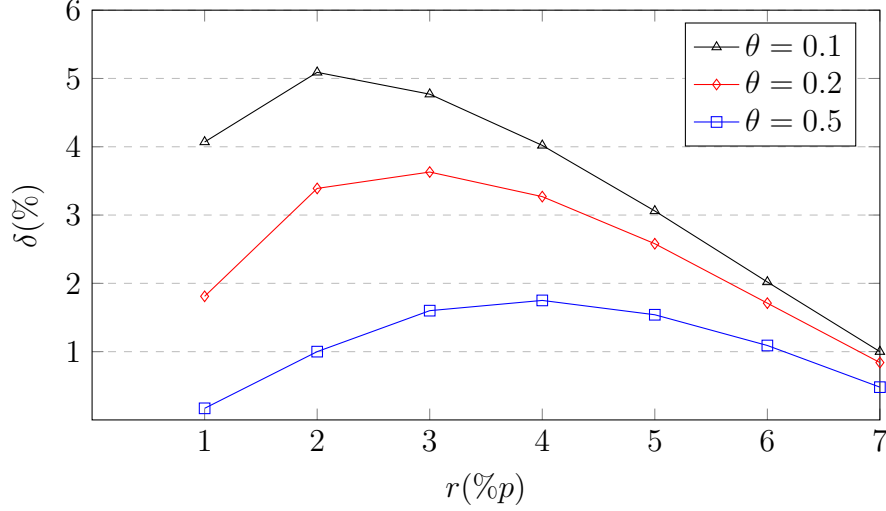


Figure 8: Effects of Discount by  $\beta(r)$  ( $p = 1.3w$ ,  $w = 10$ ,  $S = 80$ ,  $g = p - r$ ,  $U = 3$  weeks)

To summarize, we have shown that the flexible delivery option can improve system performance. The joint optimization of the inventory and distribution strategy shows larger benefits than only considering the inventory strategy, in particular when the company has the option to distribute the delivery between a third party and its own vehicle. The company is able to offer larger discounts to customers accepting flexible delivery when performing part of the distribution by itself. To ensure optimal overall performance of the system with flexible delivery, firms need to investigate how customers respond to the offer, in particular how acceptance of flexible delivery is a function of the discount offer, the maximum waiting time, and deposit value.

### 6.3. Impact flexible delivery on distance travelled

In this part, we illustrate the potential impact of flexible delivery on the average travelled distance when the distribution is solely executed by the firm. This measure may be of interest to firms to assess their impact on emissions from transportation. The performance measure

of interest is now:

$$\delta_{AVD} = 100(AVD_{F=1} - AVD_{F \leq 1})/AVD_{F=1}, \quad (30)$$

where  $AVD_{F=1}$  is the average distance travelled in the optimal solution that maximizes the AS profits of the joint inventory-distribution system in the case that no flexible delivery is offered, while  $AVD_{F \leq 1}$  is the average distance travelled in the optimal solution when the flexible delivery option is offered.

The numerical experiments are summarized in Table 5. As expected, travelled distance reduces for higher maximum waiting times and customer acceptance levels. Flexible delivery can thus not only help the firm to increase its profitability but may also help to reduce the impact on the environment. When customers are sensitive to this topic, it may help the firm to make customers aware of this positive environmental benefit and thereby increase customers' willingness to accept flexible delivery.

Table 5: Average travelled distances (  $c_u = 0.05w, p = 1.3w = 13, g = p - r, r = 0.01p, A = 50$  miles<sup>2</sup>,  $\gamma = 0.30$ )

Max Waiting Time	$F = 1$			$\beta$	$F \leq 1$			$\delta_{avgdist}(\%)$
	T	F	Dist		T	F	Dist	
3 weeks	0.29	1	2276.84	0.1	0.29	0.8	2267.81	0.396
				0.5	0.3	0.8	2196.59	3.525
				0.9	0.26	0.77	2018.86	11.330
6 weeks	0.29	1	2276.84	0.1	0.29	0.59	2255.68	0.930
				0.5	0.29	0.59	2095.03	7.985
				0.9	0.26	0.54	1720.36	24.441

## 7. Conclusions

Shifting the demand over time has become an increasingly studied strategy in supply chain management. Most studies have analysed demand postponement in a context of production resource capacity constraints. We have developed a model to study the value of flexible delivery in a continuous review inventory and distribution system. The majority of

studies have focused on static postponement schemes available to all customers at all times. In our case, we only offer the postponement at a time when the inventory level reaches a critical value, and incorporated not only the discount as a means to induce customers to accept flexible delivery, but also considered the deposit value and maximum waiting time as parameters that can affect customer acceptance. In order to see the impact of changes in the timing of cash-flows as a result of offering the flexible delivery option, we use an NPV model that is able to consider these financial aspects of the system. Numerical experiments under different parameters settings were presented to analyze the benefits of flexible delivery in comparison to a system in which this option is not offered.

The findings presented show that demand postponement through flexible delivery can bring significant benefits. With respect to inventory costs, more benefits can be expected when the marginal profit on a product remains modest, as then the optimization of the inventory costs is relatively more important in determining the firm's profits. At marginal value  $p = 1.3w$ , we find that profits from optimising the inventory system may be boosted with a few percentages by offering a discount on the sales price of a smaller to equal magnitude unless the firm's replenishment set-up cost and the opportunity cost of capital are very small. We have further shown how profitability depends on customer acceptance rate ( $\beta$ ), discount offered ( $r$ ), maximum waiting time ( $U$ ), and deposit value ( $g$ ).

Larger benefits from flexible delivery are observed when jointly optimising the inventory-distribution system. The benefit of flexible delivery to the distribution system depends on the price of outsourcing distribution relative to the firm's own delivery cost structure; the largest benefits in the distribution system can be expected when the firm finds it optimal to partially outsource delivery in delivery periods with low customer density. Total profits in the inventory-distribution system analysed can, *ceteris paribus*, more than double in comparison to the inventory system (i.e. with fixed distribution cost structure).

For given  $r$ ,  $U$  and  $g$ , benefits of the flexible delivery option generally increase with  $\beta$ , but

customer acceptance can be expected to increase with  $r$  but decrease with  $U$  and  $g$  relative to  $p - r$ . We have shown that the model can be used to determine the optimal discount value  $r$  when knowing the function  $\beta(r, U, g)$  for different customer resistance function values  $\theta(U, g)$ . Higher customer resistance values lead to higher optimal discount values, lower acceptance rates, and lower overall profits achievable. Extensions of this approach can be readily adapted to identify the optimal combination of values for all three parameters  $r$ ,  $U$ , and  $g$ . To ensure optimal overall performance of the system with flexible delivery, knowledge of  $\beta(r, U, g)$  is thus important. Further research into how firms can gather sufficient information to determine how customers respond to the flexible delivery option is warranted. When the firm only uses its own vehicle for transport, we have shown that flexible delivery can help to reduce transport distances and thus save on emissions.

The NPV methodology adopted in this paper also illustrates its benefits in showing how changes in the timing of payments can affect system performance. While Ghiami and Beullens (2016) found that deposit values  $g$  do not greatly affect performance in their production-inventory NPV model, they only tested values of deposits up to  $20\%p$ , which is reasonable in the context of backorders. In our context of online ordering, it is possible and in fact quite common that customers pay in full when placing the order. In our inventory-distribution system, the impact of  $g$  is thus found to be more significant. We also found that accounting for the timing of transportation costs is of importance in our system. Even when fully outsourced, i.e. when the total outgoing payments to the third party transportation company are unaffected by the distribution strategy, the change in the timing of these costs arising from flexible delivery decreased AS distribution costs by 2.5% for 3 weeks and 5% for a 6 weeks maximum waiting time. Models not based on the NPV methodology would not be able to identify such impacts of flexible delivery on system performance.

Further research in this area can investigate how system performance is affected by the customer acceptance function, i.e. by the shape of the function  $\beta(r, g, U)$ . It seems worth-

while to investigate how order cancellations and product returns should be incorporated into a model if their (relative) propensity is affected by the parameters of the flexible delivery option. Different possible flexible delivery schemes can also be investigated, as well as different types of distribution systems or cost structures. Since flexible delivery tends to increase replenishment order sizes, its value may further increase in situations where suppliers would offer quantity discounts, or offer a delay in payment if the order is higher than a threshold level, as in e.g. Chung et al. (2005) and Pal and Chandra (2014). Additional benefits of larger shipments may also lead to reductions in emissions from transportation between the firm and its supplier, as argued in Van Hoek (1999).

## References

- Agatz, N., Campbell, A. M., Fleischmann, M., van Nunen, J., Savelsbergh, M., 2013. Revenue management opportunities for internet retailers. *Journal of Revenue and Pricing Management* 12 (2), 128–138.
- Archetti, C., Fernández, E., Huerta-Muoz, D. L., 2017. The flexible periodic vehicle routing problem. *Computers & Operations Research* 85, 58–70.
- Beullens, P., Janssens, G., 2014. Adapting inventory models for handling various payment structures using net present value equivalence analysis. *International Journal of Production Economics* 157, 190–200.
- Beullens, P., Janssens, G. K., 2011. Holding costs under push or pull conditions—the impact of the anchor point. *European Journal of Operational Research* 215 (1), 115–125.
- Chung, K., Goyal, S., Huang, Y., 2005. The optimal inventory policies under permissible delay in payments depending on the ordering quantity. *International Journal of Production Economics* 95 (2), 203–213.
- Coelho, L. C., Cordeau, J.-F., Laporte, G., 2014. Thirty years of inventory routing. *Transportation Science* 48 (1), 119.
- Daganzo, C. F., 2005. *Logistics systems analysis*. Springer Science & Business Media, Berlin Heidelberg.
- Daganzo, C. F., Newell, G. F., 1985. Physical distribution from a warehouse: Vehicle coverage and inventory levels. *Transportation Research Part B: Methodological* 19 (5), 397–407.
- Dye, C.-Y., Ouyang, L.-Y., Hsieh, T.-P., 2007. Inventory and pricing strategies for deteriorating items with shortages: A discounted cash flow approach. *Computers & Industrial Engineering* 52 (1), 29–40.

- Elmaghraby, W., Keskinocak, P., 2003. Dynamic pricing in the presence of inventory considerations: research overview, current practices, and future directions. *IEEE Engineering Management Review* 31 (4), 47–47.
- Feng, L., Chan, Y.-L., Cárdenas-Barrón, L. E., 2017. Pricing and lot-sizing policies for perishable goods when the demand depends on selling price, displayed stocks, and expiration date. *International Journal of Production Economics* 185, 11–20.
- Ghiami, Y., Beullens, P., 2016. Planning for shortages? net present value analysis for a deteriorating item with partial backlogging. *International Journal of Production Economics* 178, 1–11.
- Grubbström, R. W., 1980. A principle for determining the correct capital costs of work-in-progress and inventory. *International Journal of Production Research* 18 (2), 259–271.
- Iyer, A. V., Deshpande, V., Wu, Z., 2003. A postponement model for demand management. *Management Science* 49 (8), 983–1002.
- Lappas, P. Z., Kritikos, M. N., Ioannou, G. D., 2017. A two-phase solution algorithm for the inventory routing problem with time windows. *Journal of Mathematics and System Science* 7 (9).
- Li, C., Zhang, F., 2013. Advance demand information, price discrimination, and preorder strategies. *Manufacturing & Service Operations Management* 15 (1), 57–71.
- Li, R., Chan, Y.-L., Chang, C.-T., Cárdenas-Barrón, L. E., 2017. Pricing and lot-sizing policies for perishable products with advance-cash-credit payments by a discounted cash-flow analysis. *International Journal of Production Economics* 193, 578–589.
- Pal, M., Chandra, S., 2014. A periodic review inventory model with stock dependent demand, permissible delay in payment and price discount on backorders. *Yugoslav Journal of Operations Research* 24 (1), 99–110.
- Papachristos, S., Skouri, K., 2003. An inventory model with deteriorating items, quantity discount, pricing and time-dependent partial backlogging. *International Journal of Production Economics* 83 (3), 247–256.
- Pentico, D. W., Toews, C., Drake, M. J., 2015. Approximating the eoq with partial backordering at an exponential or rational rate by a constant or linearly changing rate. *International Journal of Production Economics* 162, 151–159.
- Quante, R., Meyr, H., Fleischmann, M., Jan 2009. Revenue management and demand fulfillment: matching applications, models, and software. *OR Spectrum* 31 (1), 31–62.  
URL <https://doi.org/10.1007/s00291-008-0125-8>
- Shin, H., Benton, W., 2007. A quantity discount approach to supply chain coordination. *European Journal of Operational Research* 180 (2), 601–616.

- Silver, E. A., Pyke, D. F., Peterson, R., 1998. Inventory management and production planning and scheduling. Vol. 3. Wiley New York.
- Taleizadeh, A. A., Pentico, D. W., 2014. An economic order quantity model with partial backordering and all-units discount. *International Journal of Production Economics* 155, 172–184.
- Tang, C., Rajaram, K., Alptekinoglu, A., Ou, J., 2004. The benefits of advance booking discount programs: Model and analysis. *Management Science* 50 (4), 465–478.
- Tang, C. S., 2006. Perspectives in supply chain risk management. *International Journal of Production Economics* 103 (2), 451–488.
- Teng, J.-T., Cárdenas-Barrón, L. E., Chang, H.-J., Wu, J., Hu, Y., 2016. Inventory lot-size policies for deteriorating items with expiration dates and advance payments. *Applied Mathematical Modelling* 40 (19–20), 8605–8616.
- Tiwari, S., Cárdenas-Barrón, L. E., Goh, M., Shaikh, A. A., 2018a. Joint pricing and inventory model for deteriorating items with expiration dates and partial backlogging under two-level partial trade credits in supply chain. *International Journal of Production Economics* 200, 16–36.
- Tiwari, S., Cárdenas-Barrón, L. E., Khanna, A., Jaggi, C. K., 2016. Impact of trade credit and inflation on retailer's ordering policies for non-instantaneous deteriorating items in a two-warehouse environment. *International Journal of Production Economics* 176, 154–169.
- Tiwari, S., Jaggi, C. K., Gupta, M., Cárdenas-Barrón, L. E., 2018b. Optimal pricing and lot-sizing policy for supply chain system with deteriorating items under limited storage capacity. *International Journal of Production Economics* 200, 278–290.
- Van Hoek, R. I., 1999. From reversed logistics to green supply chains. *Supply Chain Management: An International Journal* 4 (3), 129–135.
- Venegas, B. B., Ventura, J. A., 2018. A two-stage supply chain coordination mechanism considering price sensitive demand and quantity discounts. *European Journal of Operational Research* 264 (2), 524–533.
- Wee, H. M., 1999. Deteriorating inventory model with quantity discount, pricing and partial backordering. *International Journal of Production Economics* 59 (1–3), 511–518.
- Wee, H., L. S., 2001. Replenishment and pricing policy for deteriorating items taking into account the time-value of money. *International Journal of Production Economics* 71 (1–3), 213–220.
- Wu, J., Ouyang, L.-Y., Cárdenas-Barrón, L. E., Goyal, S. K., 2014. Optimal credit period and lot size for deteriorating items with expiration dates under two-level trade credit financing. *European Journal of Operational Research* 237 (3), 898–908.

- Wu, Z., Wu, J., 2015. Price discount and capacity planning under demand postponement with opaque selling. *Decision Support Systems* 76, 24–34.
- Xu, H., Gong, Y. Y., Chu, C., Zhang, J., 2017. Dynamic lot-sizing models for retailers with online channels. *International Journal of Production Economics* 183, 171–184.
- Zhang, Q., Dong, M., Luo, J., Segerstedt, A., 2014. Supply chain coordination with trade credit and quantity discount incorporating default risk. *International Journal of Production Economics* 153, 352–360.
- Zhuang, W., Chen, J., Fu, X., 2017. Joint dynamic pricing and capacity control for hotels and rentals with advanced demand information. *Operations Research Letters* 45 (5), 397–402.

## Appendix A. Proof of Lemma 1

Taking the second partial derivatives of  $ASTP(F, T)$  with respect to  $F$  and  $T$  yields

$$\begin{aligned}\frac{\partial^2 ASTP}{\partial F^2}(F, T) &= -\alpha\beta(p - g)Ty < 0, \\ \frac{\partial^2 ASTP}{\partial T^2}(F, T) &= \frac{-2S}{T^3} < 0,\end{aligned}$$

The  $p - g$  is always positive from the definition, then  $ASTP(F, T)$  is concave while the  $S$  positive.

## Appendix B. Proof of Prop 1

Recall that the Lemma 1 shows that the objective function is concave. Then, the variables obtained by partial derivatives are optimal. The optimal values of  $T$  are;

$$T = \frac{\partial ASTP}{\partial T}(F, T) = \sqrt{\frac{2S}{\alpha y(w - \beta(1 - F)(2w + r + (1 - F)(p - g))}}$$

$$T_{F \leq 1} \geq T_{F=1}$$

$$\sqrt{\frac{2S}{\alpha y(w - \beta(1 - F)(2w + r + (1 - F)(p - g))}} \geq \sqrt{\frac{2S}{\alpha wy}}$$



$$2w + r + (1 - F)(p - g) \geq 0$$

$w, r$  and  $p - g$  are always non-negative and  $F \leq 1$  which left hand side of the equation is always greater than or equal to 0.

## Appendix C. Proof of Lemma 2

The proof is based on the  $ASTP(F, T)$  function when  $F = 1$ . The  $r$  value can be easily derived from the partial derivatives of  $ASTP(F, T)$ .

$$\begin{aligned} F = \frac{\partial ASTP}{\partial F}(F, T) &= -\frac{2\alpha T w + (\alpha r - 2\alpha p + 2\alpha g) T - 2r}{(2\alpha p - 2\alpha g) T} = 1 \\ -\frac{w}{p - g} - \frac{(\alpha T - 2)r}{2\alpha(p - g)T} &= 0 \\ r &= \frac{2\alpha w T}{2 - \alpha T} \end{aligned}$$

$$\begin{aligned} T = \frac{\partial ASTP}{\partial T}(F, T) &= \sqrt{\frac{2S}{\alpha y(w - \beta(1 - F)(2w + r + (1 - F)(p - g)))}} \\ T &= \sqrt{\frac{2S}{\alpha y w}} \text{ when } F=1 \end{aligned}$$

When we solve the two equations of  $r$  and  $T$  then the discount amount which makes  $F = 1$  is;

$$r = \frac{\sqrt{\frac{8S\alpha w}{y}}}{2 - \sqrt{\frac{2S\alpha}{yw}}}$$

## Appendix D. Proof of Lemma 3

The findings on Lemma 3 are obtained based on the number of demand  $n(k)$  and unit distribution cost  $c_u$  functions (described in Section 3.3). The unit distribution cost with the maximum number of demand is taken as the minimum cost while it is maximum for the minimum number of demand.