INVENTORY MANAGEMENT IN A \((Q, r)\) INVENTORY MODEL WITH TWO DEMAND CLASSES AND FLEXIBLE DELIVERY

SHUO HUANG, BO LIAO AND DAVID J. ROBB

School of Economics and Management
Tsinghua University
Haidian District, Beijing 100084, P. R. China
\{huangsh; liaob.05; david\}@sem.tsinghua.edu.cn

Received October 2009; revised March 2010

Abstract. This paper considers a single-product inventory system that serves both online and traditional customers. Demands from both channels follow independent stationary Poisson processes. Traditional customers have their demands fulfilled upon arrival if the retailer has stock on hand; otherwise the demand is lost. However, online customers place their orders in advance, and delivery is flexible in the sense that early shipment is allowed. In particular, an online order placed at time \(t\) should be fulfilled by time \(t + T\), where \(T\) is the customer’s demand lead time [1]; late fulfillment incurs a time-dependent backorder cost. A \((Q, r)\) replenishment policy is used and replenishment lead times are assumed to be constant. We develop an approximation for the expected annual cost for the retailer, and compare analytically and simulated results. The optimal parameters of the system are derived by minimizing the expected annual cost. We illustrate the model with numerical examples, and discuss the sensitivity of the results to variables such as demand lead time and the split between online and traditional orders.

Keywords: Inventory model, Multi-channel, Demand lead time

1. Introduction. The Internet has become an important retail channel in the past decade. In 2004, online sales comprised about 5.5% of all retail sales excluding travel [2]. In most product categories, a bricks-and-clicks structure, where one company has both the Internet and a traditional presence, is becoming the dominant form of the Internet participation. Large, traditional retailers such as Wal-Mart, Staples and Best Buy have websites that sell their products in significant volumes [3]. Meanwhile, some traditional online-only companies are expanding their business to retail stores. For example, Dell has installed kiosks in shopping malls and sells its computers through Costco [4]. Such firms now have two different channels to reach their customers, and face challenges in effectively managing inventory given the distinct features of different channels.

The objective of this paper is to study the inventory management problems under a bricks-and-clicks structure. We consider a model where a retailer sells a single product in both online and traditional channels. The customers in different channels exhibit different aversions to waiting. The retailer replenishes its stock according to a continuous review \((Q, r)\) policy. The objective is to make replenishment decisions to minimize the retailer’s long-run average costs. We show how this problem can be solved and present numerical examples to analyze how certain variables affect the expected costs of the retailer.

There is extensive research on inventory models with multiple demand classes [6-9]. However, few of them consider a combination of backorders and lost sales in multiple demand classes, especially under a \((Q, r)\) policy. Many papers in this stream do not deal with inventory issues [10-12], or simply assume a base stock policy and exponential replenishment lead time [5,13]. One reason for this assumption is that it enables a Markov
process analysis. However, fixed ordering costs do exist and a \((Q, r)\) policy is widely used in practice. To the best of our knowledge, we are the first to consider such a case in the prevalent \((Q, r)\) inventory model employed in retailing. We also employ an assumption of constant lead time; since in a retailing environment, replenishment lead time between suppliers and retailers is usually specified, and has little variation.

Our model also directly contributes to the stream of research on multiple demand classes/channels by taking the demand-lead-time of the online customers into consideration. In practice, online customers often place their orders in advance and early shipment is allowed. Though research on demand-lead-time has traditionally taken the exact delivery assumption for granted [1,14], the work of [15] argues that flexible delivery arrangements are prevalent in many instances where firms sell directly to consumers.

In this paper, we develop and evaluate a tractable model for bricks-and-clicks retailers, and incorporate more realistic assumptions to assist retailers make better inventory control decisions. Sensitivity analysis on free waiting time shows its impact on average inventory cost. We also evaluate how a change in the online demand ratio affects cost – providing additional insights for managers. Although this problem is motivated by inventory problems faced by bricks-and-clicks retailers, we consider our inventory system structure and solution to be applicable to a wide range of industry settings. For example, since our model combines two ways to handle shortages under a \((Q, r)\) policy, it is applicable to all three applications described in the work of [5], i.e., a retailer facing both loyal customers and occasional walk-in customers, a physical-and-online store, and an original equipment manufacturer satisfying demand using both emergency and replenishment demands from local warehouses.

The remainder of the paper is organized as follows. In Section 2 we discuss the relevant literature. Section 3 introduces the model. Section 4 describes algorithms for deriving the mathematical expression of the expected total annual cost. Numerical examples and concluding remarks are presented in Sections 5 and 6.

2. Literature Review. There exists an extensive literature considering inventory systems with multiple demand classes. The work of [17] provides a high-level categorization of the previous literature. Our model falls into the research stream of continuous review inventory control policy, but combines lost sales and backorders.

The work of [6] appears to be the first to study a continuous review \((Q, r)\) inventory model with two demand classes; they assume a critical-level policy and backordering of excess demand. The work of [9] extends the model in the work of [6] by relaxing the assumption of at most one outstanding order. Differentiating two demand classes by shortage costs, they introduce a threshold clearing mechanism to fill backorders. Using this mechanism they develop algorithms to calculate the optimal ordering and rationing parameters, and demonstrate the effectiveness of their model numerically. The work of [18] also analyzes a \((Q, r)\) inventory model with two demand classes, but differ from the work of [6] and [9] by considering a lost sales environment. They also ration inventory between different classes by applying a critical level policy.

All of the above papers differentiate demand classes based on the priority of the customers and apply a critical level policy. Our model differs in that we differentiate demand classes based on how excess demands are handled. Furthermore, our focus is on evaluating system performance given a \((Q, r)\) policy, and selecting optimal parameter values. Wang et al. [16] use delivery lead-time to differentiate demand classes. Similar to our model, they develop system performance evaluations for a given policy, although they consider a base-stock policy for inventory replenishment. Chen [8] considers multiple demand segments with differing aversions to shipping delays. This work develops an optimal pricing
and replenishment strategy for a multi-stage supply chain with backlogging – for which he proves the optimality of an echelon base stock policy.

Related work includes the multiple channel supply chain literature modeling coordination between traditional and internet channels. Cattani et al. [3] provide a review of some of the early work. These papers either analyze the dual-channel design problem by modeling the price and/or service interactions between upstream and downstream echelons [10,11], or analyze a supply chain’s performance in equilibrium [12]. However, they do not deal with inventory issues, the theme of our model.

Several studies have examined inventory issues in a dual-channel supply chain. The work of [19] considers a two-echelon dual-channel inventory model in which stock is held at both the manufacturer’s warehouse and a retail store. They assume a one-for-one policy and discuss insights based on parametric analysis. The work of [13] studies several rationing policies in which the firm either blocks or backlogs orders for lower priority customers when inventory falls to or below a certain level. Given that the customers are willing to wait, they formulate the inventory model as a Markov Process and show that their rationing policies can improve the overall performance of the firm, compared to FCFS policies. However, these papers assume base stock policies and exponential production lead times, which may not be practical in some retailing environments. In contrast, we assume a \((Q,r)\) inventory policy and a constant supply lead time, which are quite common in a retailing environment. Seifert et al. [20] model the inventory decisions for a supply chain comprised of a manufacturer and \(N\) retailers and derive optimal base stock levels. Their model is similar to ours in that the retail customers demand instantaneous availability and the retailer incurs lost sales if the product is not in stock; whereas internet customers are willing to wait. However, they assume a base stock policy and demands in the internet channel are ultimately lost after a fixed delay.

Besides considering a dual-channel structure, we take the characteristics of online customers into account when we model the online channel, assuming that they are willing to accept a reasonable waiting time after placing their order. The concept of “free waiting time” of online customers is closely related to the advance demand information (ADI) literature. Most previous studies on ADI concentrate on the value of ADI in production-inventory systems, e.g., the work of [21] presents a detailed analysis of a single-stage make-to-stock queue with ADI. Based on the same basic model, the work of [22] studies a discrete-time make-to-stock queue, showing that base stock policies are near optimal for demand lead times below a threshold. For more general inventory systems, the work of [1] denotes the time from a customer’s order until the due date as the demand lead time, and demonstrates that demand lead times directly offset supply lead times. The work of [14] investigates optimal replenishment policies for a single-stage periodic review inventory system with ADI. Using numerical results, they show that under the optimal replenishment policy, ADI can lead to significant cost reductions. This work is later extended to the multi-stage case in the work of [23]. Marklund [24] extends the work of [1] to one-warehouse multiple-retailer divergent supply chains.

Most ADI literature assumes exact delivery, early fulfillment is not allowed. In a Just-in-time (JIT) purchasing environment, this assumption is reasonable. However, as online customers, we have probably experienced the following situation. After order placement, we are notified that delivery will take place no later than a particular date. Relatively few models consider allowing for early delivery – what the work of [15] terms “flexible delivery”. The work of [25] studies inventory management with a service level constraint under a flexible time-window fulfillment scheme. They use an \((s,S)\) policy and develop algorithms for finding the optimal parameters. The work of [15] extends the model in the work of [14] by allowing for flexible delivery. They consider both homogeneous and
heterogeneous customer demand lead times. In contrast to the work of [15], we consider a continuous review environment, and allow for two demand classes with online customers forming one of the classes.

3. The Model. We consider a retailer that carries inventory for a single product to serve both online and traditional customers. The retailer replenishes inventory with a continuous review \((Q, r)\) policy from an outside supplier with infinite capacity and constant lead time. Demand from channel \(i\) (\(i = 1\) for the traditional channel, \(i = 2\) for the online channel) follows a stationary independent Poisson process with rate \(\lambda_i\).

Customers from both channels are served on a first come first served (FCFS) basis as long as there is inventory on hand. If the inventory level has dropped to zero, demands from the traditional channel are lost, whereas online customers wait until the next replenishment to be served. Due to the convenience of online purchasing, customers may accept a reasonable waiting time after they have placed their order, \textit{i.e.}, a sufficiently short waiting time will not incur any shortage cost for the retailer. We assume that an online demand is charged a backorder cost only if its waiting period exceeds the maximum free waiting time of online demands in a backorder period, denoted by \(T > 0\). Obviously, a meaningful \(T\) should not exceed the retailer lead time, \(L\), otherwise the retailer can meet every online demand as it occurs and need not hold inventory for online customers, and the system will reduce to a simple system with only one traditional demand class. A graphical illustration of the system is shown in Figure 1.

![Figure 1. An inventory system with both traditional and online demand](image)

We further assume that there is at most one outstanding order. This assumption is quite common in inventory control literature considering lost sales. The difficulties in analyzing inventory models with lost sales and the usage of this assumption has been discussed by the work of [26]. Furthermore, the assumption is often satisfied in practice because order cycles are usually long compared to lead times [27,28].

To derive the expected cost per unit time, we need to describe how the retailer inventory level, \(I(t)\), changes with time. In Figure 2, a graphical illustration of an order cycle is presented, from the time when the inventory on hand reaches the reorder point \(r\) (at \(t = 0\)), to when the replenishment arrives \((t = L)\) and then until the inventory drops to \(r\) again.

The explanation of Figure 2 is as follows: The retailer issues an order of size \(Q\) to replenish its inventory whenever the inventory level drops to \(r\). Since we assume there is no more than one outstanding order, the inventory level of the retailer immediately before this replenishment order is equal to the inventory position. Denote this time point as time 0. After time 0, the inventory continues to deplete at rate \(\lambda = \lambda_1 + \lambda_2\) until time \(T(r)\), when all \(r\) units have been demanded and the inventory level is 0. From \(T(r)\), since there is no inventory on hand, traditional customers will leave without purchasing, whereas demands from the online channel will continue to reduce the inventory, \textit{i.e.}, inventory is depleted at rate \(\lambda_2\) after \(T(r)\). We denote the time period of lost sales in a cycle as \(T_{LS}\).
INVENTORY MANAGEMENT IN A \((Q, r)\) INVENTORY MODEL

Figure 2. Changes in inventory level in a dual-channel inventory system

The backorders of online demands only incur shortage penalty costs after waiting time \(T\), \textit{i.e.}, during period \(T_B\). The accumulation of backorders that incur shortage costs is shown in Figure 2 with the bold curve. The replenishment arrives at time \(L\), after which the inventory decreases again at rate \(\lambda\) until the reorder point \(r\), completing a typical cycle.

Other notation employed includes: \(h\) denotes unit holding cost per unit time at the retailer, \(b\) is unit backorder cost per unit time at the retailer, \(\pi\) stands for unit lost sales cost, \textit{i.e.}, the lost profit (including any lost goodwill) per unit, \(K\) and \(c\) represent fixed ordering cost and variable unit purchasing cost, respectively.

Our goal is to obtain \(Q\) and \(r\) so as to minimize the average cost per unit time, \(C\), where

\[
C(Q, r) = \frac{E[TC]}{E[T_c]} \tag{1}
\]

where \(TC\) is the total cost accumulated within a cycle. Here

\[
E[TC] = K + cQ + hE[I_T] + \pi E[LS] + bE[B_T] \tag{2}
\]

where \(I_T\) is the time-weighted inventory in a cycle, \(LS\) is the number of lost sales in a cycle, and \(B_T\) is the time-weighted backorders in a cycle. The first and second parts in Equation (2) are easy to obtain. In the next section, we develop an approximate method to derive the remaining components of the expected cost in a cycle.

We adopt the notation employed in [26] to denote the Poisson probability

\[
p(j; \lambda t) = \frac{(\lambda t)^j e^{-\lambda t}}{j!} \tag{3}
\]

and its tail probability

\[
P(j; \lambda t) = \sum_{i=j}^{\infty} p(i; \lambda t) \tag{4}
\]

4. Analysis. As in Figure 2, denote \(T_{LS}\) as the time period of lost sales in a cycle. Since no online order will be lost, \(T_{LS}\) equals the time period for lost sales in the traditional
channel in a cycle. Thus

\[ T_{LS} = \begin{cases} 
0 & \text{when } T(r) \geq L, \\
L - T(r) & \text{when } T(r) < L,
\end{cases} \tag{5} \]

where \( T(r) \) is the time for \( r \) units to be depleted in both channels. Notice that \( T(r) \) is Erlang distributed with parameters \( r \) and \( \lambda \), its probability density function is \( \lambda p(r-1; \lambda t) \). Thus

\[
E[T_{LS}] = \int_0^L (L-t)f(t)dt = \int_0^L \lambda(L-t)p(r-1, \lambda t)dt = LP(r, \lambda L) - \frac{r}{\lambda} LP(r+1, \lambda L) \tag{6}
\]

Then the expected number of lost sales per cycle can be obtained as

\[
E[LS] = \lambda_1 E[T_{LS}] = \lambda_1 LP(r; \lambda L) - \frac{\lambda_1}{\lambda} r P(r+1; \lambda L) \tag{7}
\]

On average all items ordered are consumed in a single cycle, the satisfied demands per unit time can be denoted as \( Q/E[T_c] \). This number is also equal to the incoming demands minus lost sales, i.e., \( \lambda - E[LS]/E[T_c] \). By equating these two expressions we obtain

\[
E[T_c] = \frac{Q + \lambda_1 E[T_{LS}]}{\lambda} \tag{8}
\]

Denote the time-weighted backorders per cycle as \( B_T \). Due to the free waiting time period \( T \), backorder costs are only incurred in time interval \( [T(r), t] \) and when \( t \) falls into a certain interval, \( T(r) < t < L - T \). The expression for \( B_T \) can thus be formulated as

\[
B_T = \int_{T(r)}^{L-T} D_2(t - T(r))dt \tag{9}
\]

where \( D_2(t - T(r)) \) denotes online demands arising during the period \( [T(r), t] \).

Given that the probability density function of \( T(r) \) is \( f(s) = \lambda p(r-1; \lambda s) \), the expected time-weighted backorders per cycle, \( E[B_T] \), is

\[
E[B_T] = E \left[ \int_{T(r)}^{L-T} D_2(t - T(r))dt \right] = \int_0^{L-T} \lambda_2(t-s)dt f(s)ds
\]

\[
= \frac{\lambda_2(L-T)^2}{2} P(r; \lambda(L-T)) - \frac{\lambda_2(L-T)}{\lambda} r P(r+1; \lambda(L-T)) + \frac{\lambda_2 r + 1}{2\lambda^2} P(r+2, \lambda(L-T)) \tag{10}
\]

We divide the derivation of the expected time-weighted inventory held per cycle, \( E[I_T] \), into two parts: prior to the replenishment time, denoted by \( E[I_B] \), and after the replenishment time, denoted by \( E[I_A] \). Defining \( D(t) \) as the total demand during \((0, t)\), with \( I(t) \) denoted as the inventory on hand at time \( t \), then

\[
E[I_B] = \int_0^L E[I(t)]^+ dt = \int_0^L E[r - D(t)]^+ dt
\]

\[
= \int_0^L \left\{ \sum_{x=0}^{r-1} (r-x)p(x; \lambda t) \right\} dt = \sum_{j=0}^{r-1} \frac{(r-j)}{\lambda} P(j+1; \lambda L) \tag{11}
\]

In the following steps, we deduce the expected time-weighted inventory held per cycle after the replenishment time, \( E[I_A] \). Let \( I(L) \) be the inventory level just before a replenishment order arrives. According to our assumption, \( I(L) \) can only be distributed from \( r - Q \) to \( r \). Recall that \( T(r) \) is the time for \( r \) demands to arrive from time \( 0 \). If \( T(r) > L \), \( I(L) \) is greater than \( 0 \); otherwise for a given \( T(r) = t \leq L \), the probability that \( I(L) \) equals \( j \) \((j \leq 0)\) equals the probability that there are \( k \) \((k \geq r)\) demands coming from
both channels in \([0, L]\) and \(-j\) of the last \(k - r\) demands come from the online channel. Thus, we have

\[
\Pr[I(L) = j] = \begin{cases} 
 p(r - j; \lambda L), & j = 1, 2, \ldots, r \\
 \sum_{k=-j}^{\infty} p(r + k; \lambda L) \left( \frac{k}{\lambda} \right)^{k+j} \left( \frac{\lambda_2}{\lambda} \right)^{-j}, & j = -(Q - r), \ldots, -1, 0
\end{cases}
\]

Now, let \(Z = Q + I(L)\) be the inventory level just after a replenishment order arrives. The state space of \(Z\) is \(\{r, r + 1, \ldots, r + Q\}\). Denote \(T_i\) as the inter-arrival time of the demands (including traditional and online demands). Then \(T_i, i = 1, 2, 3, \ldots\) are i.i.d. exponential random variables with mean \(1/\lambda\). The time-weighted inventory held from the replenishment time until the next ordering time is then given by

\[
I_A = \begin{cases} 
 0 & \text{if } Z = r \\
 ZT_1 + (Z - 1)T_2 + \ldots + (r + 1)T_{z-r} & \text{if } Z > r
\end{cases}
\]

Hence

\[
E[I_A] = \frac{1}{\lambda} \sum_{z=r+1}^{r+Q} \Pr[Z = z](z + (z - 1) + \ldots + (r + 1))
\]

\[
= \frac{1}{2\lambda} \sum_{z=r+1}^{r+Q} (\Pr[Z = z](z^2 + z - r(r + 1)))
\]

\[
= \frac{1}{2\lambda} \left( \sum_{z=r+1}^{r+Q} (z^2 + z) \Pr[Z = z] - r(r + 1) \Pr[Z > r] \right)
\]

\[
= \frac{1}{2\lambda} \left( \sum_{j=r+1-Q}^{r} [(j + Q)^2 + (j + Q)] \Pr[I(L) = j] - r(r + 1) \Pr[I(L) > r - Q] \right)
\]

Now, we can obtain the total expected holding cost per cycle: \(hE[I_T] = h(E[I_A] + E[I_B])\).

The optimal control parameters \(Q\) and \(r\) can be determined by solving the following nonlinear integer optimization problem:

\[
\begin{align*}
\min & \quad C(Q, r), \\
\text{subject to} & \quad Q, r > 0 \text{ and integer.}
\end{align*}
\]

5. **Numerical Results.** Our objective in this section is to draw managerial insights based on a numerical analysis of our model. We first test the effectiveness of our method via comparison with simulation results. In the first stage of the experiments, we fix the financial parameters as follows: \(K = 100; c = 10; h = 3; b = 4; \pi = 25\). We set \(T = 1\) and study short and long lead times, respectively \((L = 2\) and 7\). Three demand scenarios are considered in the comparison, *i.e.*, low, medium and high demand rates with \(\lambda = 0.2, 2, 10\). In each scenario, we first assume that the demands from two channels are identical. For each instance, by employing the EOQ as an approximate \(Q\) and adjusting \(r\) from 0 to \(Q - 1\), we compare our analytical results for the expected average cost with the simulation. The results are shown in Table 1, and Figures 3 and 4.

As shown in the numerical examples, the deviations in our comparisons under low and medium demand scenarios are quite small, especially when the lead time is short. However, when the lead time is long, the gap between the simulation and analytical expected costs is increasing in the reorder point \(r\) (see Figures 3 and 4). Thus, the deviations in expected cost tend to increase when the demand rate is medium or high, and the reorder point is high. The increased gap associated with increase in \(L\) and \(r\) results from the assumption
Table 1. Expected total costs for low demand rate cases ($\lambda_1 = \lambda_2 = 0.1, Q = 4$)

<table>
<thead>
<tr>
<th>$r$</th>
<th>Short lead time ($L = 2$)</th>
<th>Long lead time ($L = 7$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Simulated</td>
<td>Analytical Difference (%)</td>
</tr>
<tr>
<td>0</td>
<td>13.52</td>
<td>13.50</td>
</tr>
<tr>
<td>1</td>
<td>16.31</td>
<td>16.31</td>
</tr>
<tr>
<td>2</td>
<td>19.28</td>
<td>19.30</td>
</tr>
<tr>
<td>3</td>
<td>22.28</td>
<td>22.30</td>
</tr>
</tbody>
</table>

Figure 3. Expected total costs for medium demand rate cases

Figure 4. Expected total costs for high demand rate cases

of at most one outstanding order. In the worst case ($\lambda_1 = \lambda_2 = 5, L = 7, r = 25$), that is, with the highest demand rate, longest lead time, and highest reorder point, the expected cost derived from the analytical method leads to a 20.51% deviation from the simulation result. However, considering that the lead time for fast moving products is usually very
short, and the reorder point is not likely to be very high, our analytical method is still effective in most real situations. Thus, we base the remaining numerical studies on the analytical model.

In the remaining numerical examples we employ the following parameters: $K = 100; c = 10; h = 3; b = 4; \lambda = 3; L = 7$. Figure 5 shows the optimal decision variables $Q^*$ and $r^*$ for different values of $\pi$ and $\lambda_1/\lambda$. We set $\lambda_1/\lambda = 0.1, 0.5, \text{ and } 0.9$, respectively. The upward trend of both $Q^*$ and $r^*$ is expected since as the cost of lost sales increases, it is reasonable to hold more inventory to shorten the period of lost sales. Another finding is that the increase in both $Q^*$ and $r^*$ with $\pi$ is more sudden for larger $\lambda_1/\lambda$, i.e., both $Q^*$ and $r^*$ are more sensitive to changes in lost sale cost when demand from the traditional channel constitutes the majority of total demand. This finding is intuitive since only demand losses are only possible from the traditional channel and the lost sales cost is usually much higher than the backorder cost, and therefore has a more significant influence on $Q$ and $r$. However, when the unit lost sales cost reaches an extremely high level, both $Q^*$ and $r^*$ tend to stabilize as the likelihood of a lost sale is extremely low in such circumstances.

![Figure 5. Optimal decision variables](image)

Figures 6 shows the results of sensitivity analysis for different values of $\lambda_1/\lambda$. Again, we set $\lambda_1/\lambda = 0.1, 0.5 \text{ and } 0.9$. For higher $\lambda_1/\lambda$, the expected cost increases more quickly. This is understandable since with more demand coming from the traditional channel, the influence of $\pi$ (which is only incurred when traditional demand is not met) will be higher. When $\pi$ is very large, the cost differences among different demand scenarios are distinct and relatively insensitive to changes in $\pi$. This arises from the optimal inventory policy for large $\pi$ having a very short period of lost sales, thus the lost sales cost constitutes only a small portion of the total expected cost and the impact of $\pi$ is thus weak.

Figure 7 shows that the expected cost is more sensitive to $T$ when $\pi$ is small. This is because when the lost sales cost is small, the dominant parts of the total expected cost are the holding cost and the backorder cost, which are both significantly affected by the length of the free waiting time, $T$. That is, with a longer $T$, the holding cost and the backorder cost, constituting the main part of the total expected cost, decrease quickly. However, when $\pi$ is large (in our example larger than 60), the lost sales cost comprises the main part of the total cost, hence the impact of $T$ on the total expected cost is much
Figure 6. Expected cost for different values of $\lambda_1/\lambda$

less significant. In addition, the high lost sales cost will result in an optimal policy with higher $Q^*$ and $r^*$, which also weakens the affect of $T$ since the period of stock-out time is very short in these cases – sometimes even shorter than $T$-rendering $T$ ineffective in reducing cost. This explains why the cost differences for different $T$ values are very small when $\pi$ is large.

Figure 7. Expected cost for different values of $T$

6. Concluding Remarks. In this paper, a tractable inventory model for bricks-and-clicks retailers is developed and analyzed. We assume a constant lead time and the commonly found $(Q, r)$ inventory policy. We allow for a period of free waiting time for online customers, a practice widely accepted by online customers due to the convenience of online purchasing. An algorithm is developed to evaluate the system. Numerical
examples validate the effectiveness of the theoretical results when seeking to minimize expected cost. The algorithm performs particularly well when lead time is not too long.

Inventory systems with multiple channels are becoming more common with the growth of e-business. How inventories should be managed in such systems is a critical managerial decision. For given parameter settings our model can be employed to determine the order size and re-order point for an item, by solving (15), which could be conducted using a spreadsheet optimization package or a two-dimensional grid search, e.g., with the grid size reflecting realistic values. Our model can be used to analyze the impacts of different parameters on the optimal inventory policy and the expected cost. For example, when the cost of lost sales exceeds a particular level, an increase in the online demand proportion helps to reduce the expected cost, and this cost-reduction can be enhanced if the free waiting period for online customers is extended. This offers insights for managers to take measures to make online shopping more appealing, i.e., to encourage traditional customers to shift to online channels. The savings derived from our analytical results can be used to determine the price or discount to online customers. Another way to decrease the expected cost is to increase the free waiting time of online customers. Here the savings derived from our analysis can be used to determine the compensation for customers waiting longer.

Our work has limitations in not allowing for stochastic leadtimes (and thus potential for order crossover) nor the possibility of more than one order outstanding at any point in time. With the relatively short leadtimes found in many e-business/retail environments we believe these assumptions are not especially problematic.

Our work can be extended in several ways, e.g., by considering pricing strategy [8,30], which would be especially pertinent were online demand to be modelled as a function of waiting time. One could also develop heuristic inventory policies for the multiple channel system, perhaps expanded to include joint retailer-supplier costs, and allow for lost sales and backorder costs for both traditional and online sales. In this regard it should be noted that fuzzy models [31,32] are increasingly being developed to address such problems.

Acknowledgment. Shuo Huang’s work on this paper was supported by the National Natural Science Foundation of China, under Grant Nos. 70601018, 71072016 and 70890082. David Robb’s work on this paper was supported by the National Natural Science Foundation of China, under Grant No. 70890082. The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES