IMA Journal of Management Mathematics (2009) **20**, 395–409
doi:10.1093/imaman/dpn035
Advance Access publication on November 13, 2008

Equilibrium replenishment in a supply chain with a single distributor and multiple retailers

KONSTANTIN KOGAN† AND YAEL PERLMAN
Department of Interdisciplinary Studies, Bar-Ilan University, Ramat-Gan, Israel

AND

SHARON HOVAV
Clalit Health Insurance, Department of Management, Tel-Aviv, Israel

[Received May 2007; accepted November 2007]

This paper addresses a problem encountered by a large-scale health service supply chain operating in a periodic review mode. Due to the vital nature of the products it provides, the number and timing of urgent orders are not limited. As a result, increasingly high transportation costs are incurred and the problem is to select an inventory replenishment (review) period that minimizes the transportation cost. Moreover, the supply chain involves multiple retailers which inevitably and independently respond to any change in replenishment policy since it may affect their inventory costs. Such a relationship results in a game between a distribution centre and retailers. Since the problem is intractable due to its scale and stochastic nature, we combine a game theoretic approach with an empirical analysis. We show that this system is predictable using equilibria and that the current replenishment equilibrium of the health service supply chain is close to the Nash solution. Numerical analysis shows that the transportation costs are cut if the distribution centre implements in reality its formal (Stackelberg) leadership by reducing the replenishment period. However, this does not coordinate the supply chain and greater system-wide savings are possible by increasing the replenishment period if the supply chain is vertically integrated or the parties cooperate.

Keywords: inventory replenishment; supply chain management; gaming.

1. Introduction

We consider a supply chain which comprises a single distributor and multiple retailers. Two supply modes, regular and urgent, characterize the system. A retailer places an order from the distribution centre at regular time intervals imposed by the distributor. Thus, the inventories are reviewed and replenished periodically. In case of a shortage, the retailer can place orders via the urgent mode. The distributor (the supplier) has ample capacity and her warehouse induces a constant inventory cost which is independent of the level of inventory handled. At the same time, transportation costs incurred by the supplier very much depend on the retailer's inventory policies. Thus, in this supply chain, the supplier seeks to find a replenishment period which minimizes the transportation cost related to both regular and urgent orders. The retailers, on the other hand, are interested in minimizing their inventory-related

†Email: kogank@mail.biu.ac.il

costs, thereby affecting the supplier's goal. This situation is naturally described by the game theoretic approach where the distributor competes with the retailers.

1.1 *Motivating example*

Clalit Health Services (Clalit), with an annual budget of NIS 13.2 [$3.3] billion, is the largest health care organization in Israel. More than 30,000 employees are engaged in providing highly advanced medical care to 55% of the Israeli population. With health care providers issuing 5 million prescriptions per year, Clalit's logistic operations deliver approximately 5000 types of items to the organization's 14 hospitals, 1380 primary and specialized clinics and 400 pharmacies. Annual operating costs amount to NIS 68.8 [$17.2] million while transportation costs stand at NIS 9.2 [$2.3] million.

Clalit successfully maintains steady operating costs but transportation costs have been increasing. Relative to operating costs, transportation costs reached 11.8% in 2003, 12.2% in 2004 and 13.3% in 2005. This increase is attributed to Clalit's willingness to dispatch urgent supplies when shortages arise. The transportation costs induced by urgent orders have reached 40% of the regular supply transportation costs (NIS 5.6 [$1.4] million per year) and have become a major management concern.

There is a number of causes for the increase. First of all, Clalit distribution centres handle regular supplies of products under a periodic review mode which allows the pharmacies to place as many urgent orders as needed. Secondly, the competition between health care providers in general and pharmacies specifically has resulted in increased frequency of deliveries—sometimes once a day or even twice a day. The main contributor to such a competition is a high concentration of the pharmacies in every city and thereby a significant likelihood that a customer who could not find a prescription drug in one pharmacy will simply go to another one (even if the need for the drug is not urgent), rather than wait for a day or two. The average value of the deliveries in these circumstances is about NIS 880 [$220] to NIS 1120 [$280] and includes only five to seven different items. Thirdly, since the pharmacies were being charged according to total shipments per month regardless of the transportation frequency, they were being encouraged to keep their inventories as low as possible.

In 2001, the distribution companies (which Clalit employed) changed this way of working since their profits were being erased and the distribution charges became frequency dependent. More specifically, the distributors began to distinguish between regular deliveries, i.e. those planned in advance and which are normally every 2 weeks, and urgent supplies, which must be carried out within a day or two. Naturally, higher charges were imposed on urgent supply deliveries.

The change in transportation charges had a decentralizing effect on the two-echelon supply chain. On the one hand, Clalit is interested in decreasing supply frequency since it is still responsible for about 60% of transportation costs. On the other hand, the competition as well as high inventory holding costs induce pharmacies to increase urgent order frequencies in response to less frequent regular supplies. Since urgent orders are more expensive, this significantly affects the overall transportation costs the health care company incurs. As a result, even though the pharmacies are a part of Clalit, the new transportation charges (imposed on Clalit by its transportation subcontractors) along with the privilege of urgent orders increase the impact the pharmacies have on the supply chain. This is to say, driven by their own goals of minimizing inventory costs, the pharmacies reduce the way that Clalit dominates the chain.

Today, Clalit's distribution centre directly supplies 350 pharmacies. About 75% of the pharmacies are supplied every 14 days, 15% once a month (every 4 weeks) and about 10% every 7 days. The pharmacies and the distribution centre are connected exclusively on an ongoing basis and orders are issued electronically. The major challenge of Clalit's distribution centre, which is a leader in the supply chain, is to determine and implement an optimal review or replenishment period that will lead to minimized overall transportation costs with respect to the best response of the pharmacies (i.e. the followers) in terms of regular and urgent orders. It is also important to understand the current position of the players in terms of leadership in the supply chain: does the distribution centre really dominate (thus implementing a Stackelberg strategy) or do the pharmacies succeed in imposing an independent (Nash) strategy on the chain by means of urgent orders? What are the losses and implications associated with the current position? These are the questions which motivated the research presented in this paper.

1.2 *Literature review*

Many authors have addressed various replenishment policies intended for either continuous or periodic inventory review. The choice of which review policy to use depends on the corresponding costs as well as practical and organizational considerations (see, e.g. Chiang & Gutierrez, 1996; Teunter & Vlachos, 2001; Rao, 2003; Bollapragada & Rao, 2006).

Traditionally, and in contrast to our problem, most papers dealing with periodic review inventory systems operating under regular and urgent orders assume that the review period is predetermined (for further details see, e.g. Veinott, 1966; Whittmore & Saunders, 1977; Chiang & Gutierrez, 1996; Teunter & Vlachos, 2001; Chiang, 2003; Bylka, 2005). Specifically, Veinott (1966) and Whittmore & Saunders (1977) derive an optimal ordering policy only when regular and emergency lead times differ by one time unit. They focus on a situation in which supply lead times are a multiple of a review period. Chiang & Gutierrez (1996) and Chiang (2003) assume a relatively large predetermined review period so that the lead times can be shorter than the review period. (This assumption is similar to ours. In our study, since the retailers are located relatively close to the distributor, the lead time is short.) They develop optimal policies for regular and urgent orders at a periodic review. Teunter & Vlachos (2001) also presume that the lead times can be shorter than the review period whose duration is predetermined. A recent work by Bylka (2005) assumes that emergency orders arrive immediately (so that the lead time of a regular order is equal to 1). The measure of effectiveness is the total (or average per period) expected cost, which includes holding, shortages and both types of order costs. The typical feature of these studies is that they allow only a very restricted (normally one or two) number of urgent orders per review period.

In this paper, we combine the game theoretic approach with empirical studies which makes it possible to account for an unlimited number of urgent orders per period. To the best of our knowledge, there is no research related to the optimal review length for two supply modes. Moreover, only a few works (see Flynn & Garstka, 1997; Rao, 2003) analyse the optimal review or replenishment period for the single supply mode. Flynn & Garstka (1997) develop a model where every T periods a retailer observes the current stock level and places orders for the next T periods. They assume that the retailer orders a sequence of deliveries and distinguish between review and delivery intervals. The review period T that they find minimizes the average cost per period. Flynn and Garstka note that T should increase as order set-up cost increases; it decreases as the holding and shortage costs as well as the variance in demand increase. Rao (2003) compares two control policies: the periodic review (R, T) policy and the continuous review, reorder point (Q, r). He shows that an economic order interval from a deterministic analysis can provide a good approximation to the optimal T.

As mentioned above, we optimize the replenishment period in a game context to take into account conflicting goals set by the distributor, on the one hand, and multiple retailers, on the other. In such a game, the distributor is a leader, who specifies first the length of the review period T. The retailers respond with regular and urgent order quantities.

Game theory has been extensively applied to supply chain management. A vast body of literature is devoted to inventory coordination or stock-related games (for literature reviews see, e.g. Cachon &

Netessine, 2004; Leng & Parlar, 2005). However, there are only a relatively small number of papers that focus on supply chains comprising multiple retailers (see, e.g. Cachon, 2001b; Wang *et al.*, 2004). Specifically, Cachon (2001b) studies the competitive and cooperative selection of inventory policies and assumes that each location implements a continuous review policy, the demand for the product is Poisson distributed and the supplier serves the retailers on a first come first serve basis. He shows that while a Nash equilibrium for a set of reorder points exists, it does not necessarily lead to supply chain efficiency. Thus, a competitive solution need not coincide with the global optimum. Wang *et al.* (2004) study a system with one supplier and multiple retailers each with her own lead time and holding cost. Each echelon uses a base-stock policy. Because the players are not cooperative and care only for their own profit, the supply chain performance deteriorates. Several contracts for the system-wide optimal cooperation are introduced.

A different setting is studied in Cachon (2001a) where one retailer sells N products with stochastic demands and trucks with finite capacity are dispatched from a warehouse. There is a constant lead time from the warehouse to the retailer. Three policies for dispatching are considered: full-service periodic review, minimum quantity periodic review and continuous review. Cachon shows that continuous review is less costly if the warehouse is close to the retailers. When the lead time is long, the advantage is small.

In this paper, we focus on the replenishment period rather than the dispatching policies. The warehouse is relatively close by and the retailers are able to issue an urgent order at any time and as many times as needed. As a result, they are continuously disrupting the periodic replenishment strategy of the distributor, which can be viewed as a constant replenishment period policy with continuous supply adjustments. We show that the replenishment period shortens when (i) the supplier is the leader, i.e. the Stackelberg strategy is employed, and (ii) the transportation cost along with the elasticity of the transportation cost with respect to the replenishment period increases. All proofs of our theoretical results are relegated to a mathematical appendix. Our empirical results are based on a case study carried out in cooperation with Clalit. The transportation costs were obtained from a sample of 16 pharmacies. To study the influence of a periodic review cycle on the transportation costs, the replenishment period for these pharmacies was changed from the original 2 weeks to 3 and 4 weeks.

2. Problem formulation

Consider a distribution center of ample capacity which supplies products to N retailers at each replenishment period t of length T. The distributor has a large automated warehouse. The warehouse is never completely filled up while inventory handling operations incur negligible variable costs compared to the fixed cost of maintaining the warehouse. The cost of transportation to the retailers during period T, $C(T, Q_t)$, on the other hand, is significant and is incurred only by the distributor. The transportation cost depends on the period length T and total order quantity $Q_t = \sum_{n=1}^{N} q_t^n$, where q_t^n is the regular order of retailer n, $n = 1, \ldots, N$, at period t. We assume that urgent orders depend on both T and Q_t and thus affect $C(T, Q_t)$, which will be studied empirically.

Since the overall number of products each retailer orders is overwhelming, similar to supply contracts which specify the total purchase when dealing with multiple products (see, e.g. Anupinidi & Bassok, 1998), we consider an aggregate order over all items of retailer n, q_t^n, measured in monetary units. Various researchers report that aggregating data in about 150–200 points normally result in less than 1% error in estimating the total transportation costs (Ballou, 1992; Hause & Jamie, 1981). In addition to the regular orders, the distributor allows for special orders in case of emergency. Since these contingent orders involve small quantities, they do not affect the retailers' inventory costs. However, as mentioned above, they do affect the transportation cost of the distributor, $C(T, Q_t)$, since special, smaller capacity vehicles are employed to carry out urgent orders. This is to say, by increasing the length of period T or decreasing the frequency of supplies—both are the same—the distributor diminishes the transportation cost of regular orders. This, however, causes the retailers to boost urgent orders required for the entire period T, thereby inducing additional costly transportation costs for the distributor. As a result, although the supply chain is formally centralized, the situation reflects a classical non-cooperative game in which the distributor is a leader, who sets first the length of the regular review period T and the retailers respond with regular and urgent order quantities. The distributor's strategy in this game is referred to as the Stackelberg solution.

2.1 *The distributor's problem*

The distributor's problem is to minimize her expected transportation cost per unit time:

$$\min_T J_{\mathrm{d}} = \lim_{K\to\infty} \frac{1}{KT} E\left[\sum_{t=1}^{K} C(T, Q_t)\right] \tag{1}$$

$$\text{s.t. } T \geqslant 0.$$

Note that although order Q_t is the total result of retailer decisions at period t, the length of the period T is independent of t, as the distributor adopts a constant-period review policy.

2.2 *The retailer's problem*

Let d_{it}^n be the customer demand rate for retailer n at ith time unit of period t. The demand is random and characterized at each time unit i by probability density $f_n(\cdot)$ and cumulative distribution $F_n(\cdot)$ with mean μ_n and standard deviation σ_n. Denote the demand for T time units at period t as

$$d_t^{Tn} = \sum_{i=1}^{T} d_{it}^n \tag{2}$$

and its density and cumulative functions as $f_{nT}(\cdot)$ and $F_{nT}(\cdot)$, respectively, with mean $T\mu_n$ and standard deviation $\sqrt{T}\sigma_n$.

The retailer n problem is to minimize her expected inventory costs per unit time:

$$\min_{\{q_t^n\}} J_{\mathrm{r}}^n = \lim_{K\to\infty} \frac{1}{KT} E\left[\sum_{t=1}^{K} h_n^+(X_t^n) + h_n^-(X_t^n)^-\right], \tag{3}$$

where
X_t^n is the retailer n inventory level at the beginning of period t,
$(X_t^n)^+ = \max\{0, X_t^n\}$ and $(X_t^n)^- = \max\{0, -X_t^n\}$,
h_n^+ and h_n^- are the unit surplus and backlog costs per time unit, respectively.

The inventory dynamics are described by the following balance equation:

$$X_{t+1}^n = X_t^n + q_t^n - d_t^{nT}, \quad q_t^n \geqslant 0, \quad t = 1, 2, \ldots. \tag{4}$$

In what follows, we derive the Stackelberg strategy by first solving the retailer's problem and then substituting the solution into the distributor's problem to find an equilibrium review period. This strategy corresponds to the fact that the distributor first sets a constant replenishment period, then in response the retailers set regular and urgent orders at each period.

3. Optimal retailer's response

According to the Stackelberg strategy, the best retailer n response is sought for a given T. This is accomplished by calculating the expectation in (3):

$$J_{\mathrm{r}}^{n}=\lim_{K\to\infty}\frac{1}{KT}\sum_{t=1}^{K}\left[\int_{-\infty}^{X_t^n+q_t^n}h_n^{+}(X_t^n+q_t^n-D)\,f_{nT}(D)\mathrm{d}D-\int_{X_t^n+q_t^n}^{\infty}h_n^{-}(X_t^n+q_t^n-D)\,f_{nT}(D)\mathrm{d}D\right]. \tag{5}$$

Next, applying the first-order optimality condition to a single-period term of (5) (i.e. differentiating a single term of (5) and setting it at zero), we obtain the single-period newsvendor-type optimal policy

$$F_{nT}(X_t^n+q_t^n)=\frac{h_n^{-}}{h_n^{-}+h_n^{+}}. \tag{6}$$

Denote the base-stock value by s_{nT},

$$X_t^n+q_t^n=s_{nT}, \tag{7}$$

such that $F_{nT}(s_{nT})=\frac{h_n^{-}}{h_n^{-}+h_n^{+}}$. Retailer n then orders up to this stock level s_{nT} if the current level of inventory is less than s_{nT}, otherwise she does not order at all. Assuming that initial inventory is less than or equal to s_{nT}, we observe that if this single-period solution (which is myopic as it ignores possible effects of the other periods) is applied at each period, then $X_{t+1}^n\leqslant s_{nT}$ for each t. This argument is then used in the following theorem (detailed proof can be found in, e.g. Zipkin, 2000).

THEOREM 1 The myopic stationary base-stock policy with base-stock level s_{nT} is optimal for multi-period problem (3–4).

We thus determined the best retailer's n response, $q_t^n=s_{nT}-X_t^n$, to any replenishment period T set by the distributor. In practice, retailers are not always able to calculate their unit backlog costs. However, the value of $w=\frac{h_n^{-}}{h_n^{-}+h_n^{+}}$ which is referred to as the service level (i.e. $1-w$ is the probability of backlog $P(d_t^{Tn}>s_{nT})=1-w$) is frequently used by the management as a goal to be met. The higher the goal (the service level) set, the greater the base-stock level and thus the lower the risk of backlogs induced by random demands.

4. Stackelberg strategy

Given initial inventory levels, the retailers' orders are deterministic at the first period. Therefore, calculating the expectation in (3) we find that

$$J_{\mathrm{d}}=\lim_{K\to\infty}\frac{1}{KT}E\left[\sum_{t=1}^{K}C(T,Q_t)\right]=\lim_{K\to\infty}\frac{1}{KT}\left\{C(T,Q_1)+\sum_{t=2}^{K}E[C(T,Q_t)]\right\}. \tag{8}$$

Let $f_Q(\cdot)$ and $F_Q(\cdot)$ be the total order Q_t density and cumulative distributions, respectively. Then, (8) can be presented as

$$J_{\rm d} = \lim_{K\to\infty} \frac{1}{KT} \left\{ C(T, Q_1) + \sum_{t=2}^{K} \int_{-\infty}^{\infty} C(T, \xi_t)\, f_Q(\xi_t) {\rm d}\xi_t \right\}. \tag{9}$$

The distribution of the total order, $f_Q(\xi_t)$, can be determined with the aid of Theorem 1. Indeed, from (4) and (7), we have

$$X^n_{t+1} = s_{nT} - d^{nT}_t, \quad X^n_{t+1} = s_{nT} - q^n_{t+1}$$

and thus

$$q^n_{t+1} = d^{nT}_t, \quad t = 1, 2, \ldots.$$

Consequently, the distribution of the optimal orders q^n_t, $t \geqslant 2$ (for a replenishment period of length T), is identical to the demand d^{nT}_t distribution. Since the demand distribution per time unit is assumed to be independent of time, i.e. d^{nT}_t is stationary, the distribution of the total order

$$Q_t = \sum_{n=1}^{N} q^n_t = \sum_{n=1}^{N} d^{nT}_t, \quad t \geqslant 2,$$

is stationary as well with mean $T\sum_{n=1}^{N}\mu_n$ and standard deviation $\sqrt{T\sum_{n=1}^{N}\sigma_n^2}$. Since the problem data are stationary, (9) simplifies to

$$\begin{aligned} J_{\rm d} &= \lim_{K\to\infty} \left\{ \frac{C(T, Q_1)}{KT} + \frac{1}{KT} \sum_{t=2}^{K} \int_{-\infty}^{\infty} C(T, \xi_t)\, f_Q(\xi_t) {\rm d}\xi_t \right\} \\ &= \lim_{K\to\infty} \left\{ \frac{C(T, Q_1) - E[C(T, Q)]}{KT} + \frac{1}{T} \int_{-\infty}^{\infty} C(T, \xi)\, f_Q(\xi) {\rm d}\xi \right\}. \end{aligned}$$

Assume that the probability of extremely high demands is negligible and a much stretched replenishment period results in enormous transportation costs due to urgent orders. This, along with the constraints $T \geqslant 0$ and $Q_t \geqslant 0$, implies that the solution sets for T and Q_t are compact. Consequently, the limit in the last expression results in

$$J_{\rm d} = \frac{1}{T} \int_{-\infty}^{\infty} C(T, \xi)\, f_Q(\xi) {\rm d}\xi. \tag{10}$$

The equilibrium is then obtained by assuming that $f_Q(\cdot)$ depends on T and by applying the first-order optimality condition with respect to T:

$$\int_{-\infty}^{\infty} \left(\frac{\partial [C(T, \xi)\, f_Q(\xi)]}{T\partial T} - \frac{C(T, \xi)\, f_Q(\xi)}{T^2} \right) {\rm d}\xi = 0. \tag{11}$$

Denoting $A = \int_{-\infty}^{\infty} (\frac{\partial [C(T,\xi) f_Q(\xi)]}{T\partial T} - \frac{C(T,\xi) f_Q(\xi)}{T^2}) {\rm d}\xi$ and a solution of (11) in T by α, we conclude with the following theorem.

THEOREM 2 If $\frac{\partial A}{\partial T} > 0$ and $\alpha \geqslant 0$, then the replenishment period $T^* = \alpha$ and the base-stock level $s^*_{nT} = s_{n\alpha}$ with order quantity $q^{n*}_t = s_{n\alpha} - X^n_t$ for $t = 1, 2, \ldots, n = 1, \ldots, N$ constitute a unique Stackelberg equilibrium in the distributor/retailer replenishment game.

5. Nash strategy

To compare the effect of leadership on the game between the supply chain parties, we next assume that there is no leader in the chain. This implies that the distributor and the retailers make their decisions simultaneously so that in contrast to the Stackelberg strategy, the distributor's objective function (1) is minimized as though the total order quantity Q_t (and hence, $f_Q(\cdot)$) does not depend on T. Then, applying the first-order optimality condition to the objective function (10), we obtain

$$\int_{-\infty}^{\infty}\left(\frac{\partial C(T,\xi)}{T\partial T}-\frac{C(T,\xi)}{T^2}\right) f_Q(\xi)\mathrm{d}\xi = 0. \tag{12}$$

Introducing $B=\int_{-\infty}^{\infty}\left(\frac{\partial C(T,\xi)}{T\partial T}-\frac{C(T,\xi)}{T^2}\right) f_Q(\xi)\mathrm{d}\xi$ and denoting a solution of (12) in T by β, we conclude with the following theorem.

THEOREM 3 If $\beta \geqslant 0$ and $\frac{\partial B}{\partial T} > 0$, then the replenishment period $T^* = \beta$ and the base-stock level $s^*_{nT} = s_{n\alpha}$ with order quantity $q^{n*}_t = s_{n\beta} - X^n_t$ for $t = 1, 2, \ldots$, $n = 1, \ldots, N$ constitute a Nash equilibrium in the distributor/retailer replenishment game.

6. Theoretical results for a normal distribution of the demand

As discussed above, independent optimization of the retailers' responses results in the distribution of the optimal orders q^n_t (for a replenishment period of length T) identical to the demand d^{nT}_t distribution. Assuming that the demand distribution is normal and its mean is greater than its three standard deviations (to ensure that the probability of negative demands is negligible), we observe that $f_{nT}(\cdot)$ and $F_{nT}(\cdot)$ are normal density and cumulative functions with mean $T\mu_n$ and standard deviation $\sqrt{T}\sigma_n$. Then the total optimal order, Q_t, is characterized by the normal distribution as well. Note that if the demand is independent at each time unit (i.e. stationary), then according to the central limit theorem, summation of independent demands over T time units tends to the normal distribution even if the demand at each time unit is not normal. In other words, the normality assumption of this section is not very restricting.

Our first observation with respect to the normal distribution is related to the retailers' objective functions. When $T = 1$, the standard deviation of demand d_n per time unit is σ_n. When T increases, the standard deviation reduces, $\frac{\sqrt{T}}{T}\sigma_n = \frac{\sigma_n}{\sqrt{T}}$. Therefore, similar to the pooling demand effect widely employed in supply chains, the retailer's expected inventory cost per time unit is a monotonically decreasing function of T (see Fig. 2). This is shown in the following proposition by utilizing the standard normal density function $\Phi(\cdot)$.

PROPOSITION 1 Let $f_{nT}(\cdot)$ be the normal density function with mean $T\mu_n$ and standard deviation $\sqrt{T}\sigma_n$. Then the greater the replenishment period T, the lower the retailer's expected inventory cost per period, J^n_{r}, so that $\frac{\partial J^n_{\mathrm{r}}}{\partial T} < 0$ and $\frac{\partial^2 J^n_{\mathrm{r}}}{\partial T^2} > 0$.

The following observation is related to the Nash solution. Considering now the distributor's cost J_{d}, the best response of the distributor, $T = T^R(s_{nT})$, as well as of retailer n, $s_{nT} = s^R_{nT}(T)$, we obtain the following properties.

PROPOSITION 2 Let $f_{nT}(\cdot)$ be the normal density function with mean $T\mu_n$ and standard deviation $\sqrt{T}\sigma_n$. Then the distributors cost and, hence, best distributor's response do not depend on retailer n base-stock level, $\frac{\partial T^R}{\partial s_{nT}} = 0$. On the other hand, the best retailer n response does depend on the replenishment period: the greater the replenishment period T, the larger the base-stock level $\frac{\partial s^R_{nT}(T)}{\partial T} > 0$.

There are two important conclusions related to Proposition 2. The first conclusion is concerned with the supply chain performance and thereby the corresponding centralized supply chain. If the supply chain is vertically integrated with one decision maker responsible for setting both a replenishment period and a base-stock level for each retailer, then the centralized objective function is a summation of all costs involved:

$$J(T) = J_{\mathrm{d}} + \sum_{n} J_{\mathrm{r}^n}.$$

The distributor's cost J_{d} is independent of the base-stock level, as shown in Proposition 2. Therefore, applying the first-order optimality condition to $J(T)$ with respect to either q_t^n or s_{nT}, we obtain (6). This implies that the condition for the Nash base-stock level is identical to the system-wide optimality condition. Next, to find the system-wide optimality condition for the replenishment period, we differentiate $J(T)$ with respect to T, which, when taking into account (12) and (A.1), results in

$$\frac{\partial J(T)}{\partial T} = \int_{-\infty}^{\infty} \left(\frac{\partial C(T,\xi)}{T \partial T} - \frac{C(T,\xi)}{T^2} \right) f_Q(\xi) \mathrm{d}\xi - \frac{\sigma_n}{2\sqrt{T^3}} \left[\int_{-\infty}^{s_n^*} h_n^{+}(s_n^* - z)\Phi(z)\mathrm{d}z - \int_{s_n^*}^{\infty} h_n^{-}(s_n^* - z)\Phi(z)\mathrm{d}z \right] = 0. \tag{14}$$

Comparing (14) and (12), we find the following property.

PROPOSITION 3 Let $\frac{\partial B}{\partial T} > 0$ and $f_{nT}(\cdot)$ be the normal density function with mean $T\mu_n$ and standard deviation $\sqrt{T}\sigma_n$. The system-wide optimal replenishment period and base-stock level are greater than the Nash replenishment period and base-stock level, respectively.

Proposition 3 sustains a well-known observation that vertical competition causes the supply chain performance to deteriorate. Similar to the double marginalization effect, this happens because the retailers ignore the distributor's transportation cost by keeping lower, base-stock inventory levels. The distributor, on the other hand, ignores the retailer's inventory costs when choosing the replenishment period. Figure 2 illustrates the effect of vertical competition on the supply chain.

Theorem 3 derives conditions when Nash equilibria exist. The second property, which is readily derived from Proposition 2, is related to the uniqueness of the Nash solution.

PROPOSITION 4 Let $f_{nT}(.)$ be the normal density function with mean $T\mu_n$ and standard deviation $\sqrt{T}\sigma_n$. The Nash equilibrium (T^*, s_{nT}^*) determined by Theorem 3 is unique.

7. Empirical results and numerical analysis

7.1 *Closed-form equilibrium expressions under normal demands*

Let us denote $\mu = \sum_{n=1}^{N} \mu_n$ and $\sigma = \sqrt{\sum_{n=1}^{N} \sigma_n^2}$. Then, $\mu_Q = T\sum_{n=1}^{N} \mu_n = \mu T$ and $\sigma_Q = \sqrt{T\sum_{n=1}^{N} \sigma_n^2} = \sigma\sqrt{T}$. Thus, the total order probability density function is

$$f_Q(\xi) = \frac{1}{\sqrt{2\pi}\sigma_Q} \mathrm{e}^{-(\xi-\mu_Q)^2/2\sigma_Q^2} = \frac{1}{\sqrt{2T\pi}\sigma} \mathrm{e}^{-(\xi-\mu T)^2/2T\sigma^2}.$$

Guided by our empirical results, which are discussed below, we consider an exponential function of the transportation cost, $C(T, Q) = a\,\mathrm{e}^{bTQ}$. Substituting this into (11), we obtain an explicit closed-form expression for Stackelberg equilibrium replenishment period $T^* = \alpha$:

$$\int_{-\infty}^{\infty} \frac{a}{\sqrt{2\pi\sigma}} \left[\frac{\partial(\mathrm{e}^{b\alpha\xi} \alpha^{-0.5}\, \mathrm{e}^{-(\xi-\mu\alpha)^2/2\alpha\sigma^2})}{-} \frac{\mathrm{e}^{b\alpha\xi}\, \mathrm{e}^{-(\xi-\mu\alpha)^2/2\alpha\sigma^2}}{\alpha^{2.5}} \right] \mathrm{d}\xi = 0,$$

i.e.

$$\int_{-\infty}^{\infty} \frac{a}{\alpha^{1.5}\sqrt{2\pi\sigma}} \mathrm{e}^{b\alpha\xi-(\xi-\mu\alpha)^2 2\alpha\sigma^2} \left[b\xi + \frac{2(\xi-\mu\alpha)\mu}{2\alpha\sigma^2} + \frac{(\xi-\mu\alpha)^2}{2\alpha^2\sigma^2} - \frac{1}{\alpha} \right] \mathrm{d}\xi = 0. \qquad (15)$$

Similarly, substituting the normal distribution and exponential cost function into (11), we find an expression for the Nash equilibrium replenishment period $T^* = \beta$:

$$\int_{-\infty}^{\infty} \frac{a}{\sqrt{2\pi\sigma}} \mathrm{e}^{b\beta\xi-(\xi-\mu\beta)^2/2\beta\sigma^2} \left[\frac{b\xi}{\beta} - \frac{1}{\beta^2} \right] \mathrm{d}\xi = 0. \qquad (16)$$

Both (15) and (16) are numerically studied next.

7.2 *Empirical results*

The transportation costs were obtained from a sample of 16 pharmacies, which are being supplied every 14 days on a regular basis exclusively from Clalit's primary distribution centre. The base-stock policy was determined according to service-level definition and demand forecasts. Pharmacists place their orders using software that computes replenishment quantities for every item with respect to the base-stock level. The pharmacist electronically sends the completed order to the distribution centre for packing and dispatching. If there is a shortage or expected shortage before the next planned delivery, the pharmacist can send an urgent order to be delivered not later than two working days from the time of the order.

An external subcontractor (according to the outsourcing agreement) delivers the orders to the pharmacies. The contractor schedules the appropriate vehicle (trucks in case of regular orders and mini-trucks for urgent orders) according to the supply plans for the following day. Delivery costs depend on the type of the vehicle used (track or mini-track) and the number of pharmacies to be supplied with the specific transport.

To estimate the influence of a periodic review cycle on the transportation costs (planned and urgent deliveries), the replenishment period for the 16 pharmacies was changed from the original 2 weeks to 3 and 4 weeks. This resulted in a total of 18 replenishment cycles representing 34 working weeks. Monthly sales of the selected pharmacies varied from NIS 200,000 [$50,000] to NIS 544,000 [$136,000]. Each order that was sent from a pharmacy was reported, and each transport, with every delivery on it, including invoices that were paid to vehicle contractor was reported. The data, processed with SPSS non-linear regression analysis, indicate that the resultant parameters of the transportation cost function are $a = 4463$, $b = 0.0000163$, while both parameters of the selected cost function are statistically significant at levels of P value much smaller than the accustomed maximum 5% (0.05) requirement. Specifically, the significance levels of a and b are 3.0653×10^{-12} and 1.08651×10^{-09}, respectively.

7.3 *Numerical analysis*

The goal of our numerical analysis is to check whether this supply chain is predictable using equilibria and how it is affected by the distributor's leadership. In other words, we compare the objective functions (1) and (3), as well as the effect on the overall supply chain (the sum of (1) and (3)). Specifically, with distributor leadership, its expected cost (10) is $J_{d1} = \frac{1}{\alpha}\int_{-\infty}^{\infty} C(\alpha,\xi)\, f_Q(\xi)\mathrm{d}\xi$ (see Theorem 2), while without leadership (see Theorem 3) it is $J_{d2} = \frac{1}{\beta}\int_{-\infty}^{\infty} C(\beta,\xi)\, f_Q(\xi)\mathrm{d}\xi$. Since α is found by minimizing the entire objective function J_{d1}, while β assumes that the normal probability function is independent of the period T, the distributor obviously is better off if she is the leader and therefore decides first rather than when the decision is made simultaneously (no leaders).

Similarly, retailer n expected cost under the distributor leadership is

$$J_{r1}^{n} = \frac{1}{\alpha}\left[\int_{-\infty}^{s_{n\alpha}} h_n^{+}(s_{n\alpha} - D)\, f_{n\alpha}(D)\mathrm{d}D - \int_{s_{n\alpha}}^{\infty} h_n^{-}(s_{n\alpha}^{-} D)\, f_{n\alpha}(D)\mathrm{d}D\right],$$

while under no leadership it is

$$J_{r2}^{n} = \frac{1}{\beta}\left[\int_{-\infty}^{s_{n\beta}} h_n^{+}(s_{n\beta}^{-} D)\, f_{n\beta}(D)\mathrm{d}D - \int_{s_{n\beta}}^{\infty} h_n^{-}(s_{n\beta}^{-} D)\, f_{n\beta}(D)\mathrm{d}D\right].$$

The numerical results of our empirical studies show that the current equilibrium of Clalit's supply chain, which is an outcome of many adjustments it has undergone during many years of operations, is close to and positioned in between both the Stackelberg and the Nash equilibria. This is in contrast to the skepticism of many practitioners who believe that a theoretical equilibrium is hardly attainable in real life. Specifically, the equilibrium replenishment period under equal competition is about 16 days, the current replenishment period is 14 days and the equilibrium under the distributor's leadership is 11 days. Figure 1 presents the equilibria over the distributor's transportation cost function.

The Stackelberg equilibrium demonstrates the power the distributor can harness as a leader. The economic implication of harnessing the distributor's power is about 20 NIS per day [$4 per day] for the sampled supply volumes. The annual significance (which is an estimate based on extrapolating our results), in terms of the overall supply chain, is NIS 1.4 [$0.35] million or 14% of the total delivery costs. Interestingly enough, the current equilibrium is closer to the Nash replenishment period rather than to the Stackelberg which sustains Clalit's managerial intuition that its distribution centres do not succeed in taking full advantage of their power over the pharmacies. There are also pragmatic reasons/benefits for the current period being 14 days rather than 11 (according to the Stackelberg strategy) related to the order for each pharmacy being placed on the same day of the working week.

Figure 2 presents the results of the calculation for the supply chain as a whole, i.e. including the retailers' inventory management costs and the distributor's transportation costs.

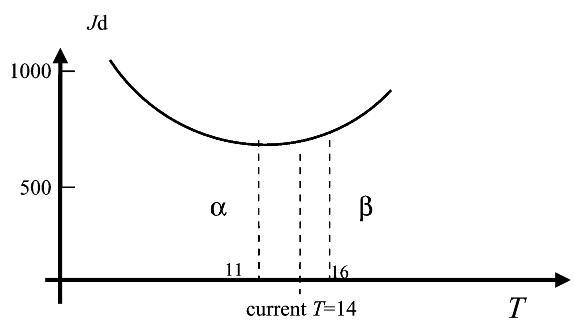

FIG. 1. The distributor's transportation cost, J_d (NIS), as a function of T (days) ($\beta = 16$ and $\alpha = 11$—Nash and Stackelberg equilibrium replenishment periods, respectively).

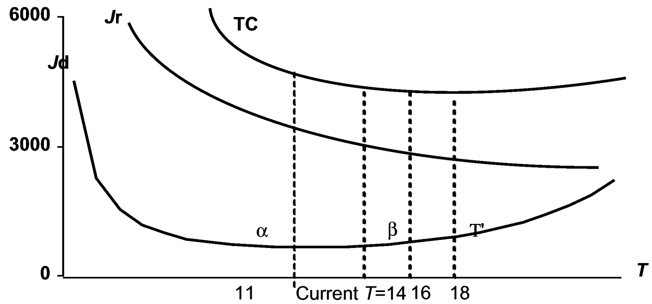

FIG. 2. Supply chain costs (TC—total supply chain cost, NIS; J_r—total retailers cost, NIS; J_d—distributor's transportation cost, NIS; T'—system-wide optimal period, days; $\beta = 16$ and $\alpha = 11$—Nash and Stackelberg equilibrium replenishment periods, respectively).

In Fig. 2, the total costs for the Stackelberg, current and Nash strategies as well as the system-wide optimal (global) solution appear as dots on the total costs curve. From this diagram, it is easy to observe the effect of the total inventory-related cost on the entire system performance. Specifically, we can see that if the supply chain is vertically integrated or fully centralized and thus has a single decision maker who is in charge of all managerial aspects, the system-wide optimal replenishment period is 18 days versus the current equilibrium of 14 days. The significance of this gap (which agrees with Proposition 4) is that more than 3 million NIS could be saved if the system were vertically integrated. If the distributor attempts to locally optimize (the Stackelberg strategy), this would lead to annual savings in transportation costs of only 1.4 million NIS. However, the significance of such an optimization for the supply chain as a whole is a 'loss' of 8 million NIS. This is the price to be paid if the supply chain is either decentralized or operates as a decentralized system.

Finally, to understand the sensitivity of the results, particularly of the replenishment periods to the transportation cost, we next conduct a sensitivity analysis.

7.4 *Sensitivity analysis*

Elasticity ζ of the transportation cost with respect to the replenishment period T, $\zeta = bQT$, shows the relative change in transportation cost when the length of the replenishment period changes by 1%. An important question relates as to what extent the transportation costs should increase, or the characteristic parameter b (elasticity per \$1 and 1 day) amplify so that the distributor will need to reconsider her supply policy from periodic review towards continuous review.

Specifically, if the elasticity per \$1 and 1 day changes, say three times from $b = 0.0000163$ to $b = 0.0000489$, then the Nash period falls 1.77 times and the Stackelberg period decreases 1.83 times, attaining 7 and 5 days, respectively (see Fig. 3 for detailed sensitivity of the equilibrium periods with respect to b). On the other hand, in terms of the elasticity of the transportation cost, the change required is less than that for b. We observe this by first plugging into the equation for ζ,

$$\zeta = bQT = 0.0000163QT,$$

the current average value of Q and current T, to find that $\zeta = 0.4$. Next, we find that $\zeta = 0.54$ when setting $b = 0.0000489$ (if we increase it three times), T equal to the corresponding Nash value β from Fig. 3 and Q equal to the total average order when $T = \beta$. Thus, if b increases three times, the elasticity with respect to the Nash period increases only 1.62 times. Similarly, we find for the Stackelberg

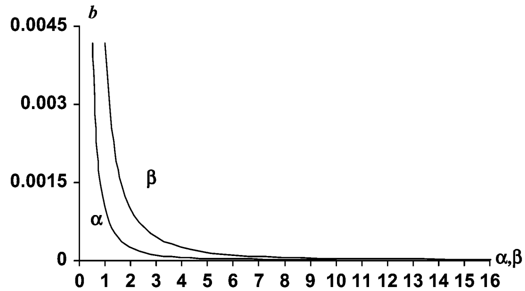

FIG. 3. Sensitivity analysis: b as a function of β and α.

period, i.e. $T = \alpha$, that $\zeta = 0.27$ and the elasticity increases only 0.84 times. This is to say that when reconsidering review policy from periodic to continuous, the transportation costs must significantly increase, while the elasticity with respect to the replenishment period does not need to increase to such an extent.

8. Conclusions

This research was motivated by increasingly high transportation costs incurred by a large health service provider which is part of a supply chain consisting of multiple retailers (pharmacies) and a distribution centre. The costs were attributed to unlimited urgent orders that the retailers could place in the system. Management's approach to handling this problem was to reduce the replenishment period or even reconsider the policy from periodic to continuous-time review. The latter in the current conditions would simply imply daily (regular) product deliveries.

For the case of a normal demand distribution, we show that similar to the pooling demand effect widely employed in supply chains, the retailer's expected inventory cost per time unit is a monotonically decreasing function of the replenishment period. Furthermore, using the game theoretic approach we find that the distributor's cost and, hence, optimal distributor's response do not depend on the retailer's base-stock level. On the other hand, the optimal retailer's response does depend on the replenishment period: the greater the replenishment period, the larger the base-stock level. As a result, the Nash equilibrium is unique and the system-wide optimal replenishment period and base-stock level are greater than the Nash replenishment period and base-stock level, respectively.

Due to the large scale of the supply chain, its stochastic character and urgent mode of orders, the choice of an optimal replenishment period is an intractable problem. To overcome this in our study, we combined the game theoretic approach with an empirical analysis. As a result, closed-form expressions for Nash and Stackelberg solutions were derived. A numerical analysis of these solutions shows that if a distributor imposes her leadership on the supply chain, i.e. acts as the Stackelberg leader, then the replenishment equilibrium period is reduced. This makes it possible to cut high transportation costs. However, these costs must be much higher in order for the distributor to reconsider her review policy from periodic to continuous. Moreover, both theoretical and empirical studies show that if instead an imposed leadership on the supply chain, it is vertically integrated or the parties cooperate, then the potential savings in overall costs are much greater. In such a case, the replenishment period must increase towards the system-wide optimal period rather than decrease or transform into a continuous review policy. Thus, in the short run, imposing leadership by reducing the replenishment period may cut high transportation costs. However, in the long-run, greater savings are possible if, e.g. the vendor managed inventory approach is adopted by the retailers or imposed on the retailers by the health provider. In such a case, a distribution centre will decide when and how to replenish inventories and thus the system will become vertically integrated with respect to transportation and inventory considerations. This illustrates the economic potential in cooperation and a total view of the whole supply chain.

REFERENCES

ANUPINIDI, R. & BASSOK, Y. (1998) Approximations for multiproduct contracts with stochastic demands and business volume discounts: single-supplier case. *IIE Trans.*, **30**, 723–734.

BALLOU, R. H. (1992) *Business Logistics Management*. Englewood Cliffs, NJ: Prentice-Hall.

BOLLAPRAGADA, R. & RAO, U. S. (2006) Replenishment planning in discrete-time, capacitated, non-stationary, stochastic inventory systems. *IIE Trans.*, **38**, 605–620.

BYLKA, S. (2005) Turnpike policies for periodic review inventory model with emergency orders. *Int. J. Prod. Econ.*, **93**, 357–373.

CACHON, G. (2001a) Managing a retailer's shelf space, inventory, and transportation. *Manuf. Serv. Oper. Manage.*, **3**, 211–229.

CACHON, G. (2001b) Stock wars: inventory competition in a two-echelon supply chain with multiple retailers. *Oper. Res.*, **49**, 658–674.

CACHON, G. & NETESSINE, S. (2004) Game theory in supply chain analysis. *Handbook of Quantitative Supply Chain Analysis: Modeling in the E-Business Era* (D. Simchi-Levi, S. D. Wu & Z. Shen eds). Kluwer.

CHIANG, C. (2003) Optimal replenishment for a periodic review inventory system with two supply modes. *Eur. J. Oper. Res.*, **149**, 229–244.

CHIANG, C. & GUTIERREZ, G. J. (1996) A periodic review inventory system with two supply modes. *Eur. J. Oper. Res.*, **94**, 389–403.

DEBREU, D. (1952) A social equilibrium existence theorem. *Proc. Natl. Acad. Sci. U.S.A.*, **38**, 886–893.

FLYNN, J. & GARSTKA, S. (1997) The optimal review period in a dynamic inventory model. *Oper. Res.*, **45**, 736–750.

HAUSE, R. G. & JAMIE, K. D. (1981) Measuring the impact of alternative market classification systems in distribution planning. *J. Bus. Logist.*, **2**, 1–31.

LENG, M. & PARLAR, M. (2005) Game theoretic applications in supply chain management: a review. *Inf. Syst. Oper. Res.*, **43**, 187–220.

RAO, U. S. (2003) Properties of the periodic review (R, T) inventory control policy for stationary, stochastic demand. *Manuf. Serv. Oper. Manage.*, **5**, 37–53.

TEUNTER, R. & VLACHOS, D. (2001) An inventory system with periodic regular review and flexible emergency review. *IIE Trans.*, **33**, 625–635.

VEINOTT JR, A. F. (1966) The status of mathematical inventory theory. *Manage. Sci.*, **12**, 745–777.

WANG, H., GUO, M. & EFSTATHIOU, J. (2004) A game theoretical cooperative mechanism design for two-echelon decentralized supply chain. *Eur. J. Oper. Res.*, **157**, 372–388.

WHITTMORE, A. S. & SAUNDERS, S. (1977) Optimal inventory under stochastic demand with two supply options. *SIAM J. Appl. Math.*, **32**, 293–305.

ZIPKIN, P. H. (2000) *Foundations of Inventory Management*. New York: McGraw-Hill.

Appendix

Proof of Theorem 2: The proof immediately follows from the first-order optimality condition (8) and Theorem 1. The equilibrium is unique if objective function (1) is strictly convex in T, i.e. $\frac{\partial^2 J_{\rm d}}{\partial T^2} > 0$, which is ensured by $\frac{\partial A}{\partial T} > 0$. □

Proof of Theorem 3: First note that $\frac{\partial B}{\partial T} > 0$ ensures convexity of the distributor's cost, while the newsvendor type of objective (5) is evidently convex in q_t^n as well. If these conditions hold, then there must exist at least one simultaneous solution of the systems (12) and (6) (see, e.g. Debreu, 1952), which if positive constitutes a Nash equilibrium for the distributor/retailer replenishment game. □

Proof of Proposition 1: With respect to Theorem 1, retailer n expected cost is

$$J_{\mathrm{r}}^n = \frac{1}{T}\left[\int_{-\infty}^{s_{nT}} h_n^+(s_{nT}^- D)\, f_{nT}(D)\mathrm{d}D - \int_{s_{nT}}^{\infty} h_n^-(s_{nT}^- D)\, f_{nT}(D)\mathrm{d}D\right].$$

Using the fact that $f_{nT}(D) = \frac{1}{\sqrt{T}\sigma_n}\Phi\left(\frac{D-T\mu_n}{\sqrt{T}\sigma_n}\right)$ and introducing a new variable z,

$$z = \frac{D - T\mu_n}{\sqrt{T}\sigma_n},$$

as well as a standardized base-stock level $s_n^* = \frac{s_{nT}-T\mu_n}{\sqrt{T}\sigma_n}$, the expected cost takes the following form:

$$J_{\mathrm{r}}^n = \frac{\sigma_n}{\sqrt{T}}\left[\int_{-\infty}^{s_n^*} h_n^+(s_n^* - z)\Phi(z)\mathrm{d}z - \int_{s_n^*}^{\infty} h_n^-(s_n^* - z)\Phi(z)\mathrm{d}z\right]. \tag{A.1}$$

Differentiating this expression with respect to T, we immediately observe that $\frac{\partial J_{\mathrm{r}}^n}{\partial T} < 0$ and $\frac{\partial^2 J_{\mathrm{r}}^n}{\partial T^2} > 0$, as stated in this proposition. □

Proof of Proposition 2: First note that neither J_{d} nor its derivative, which is the left-hand side of (11), denoted by B, explicitly depends on s_{nT}. Furthermore, according to Theorem 1, no matter what base-stock level s_{nT} we choose, the quantity that retailer n orders has the same distribution, which depends only on demand. Thus, a given replenishment period T, $f_Q(\cdot)$, does not depend on the base-stock policy s_{nT} employed. This is to say that J_{d} and B are independent of s_{nT}. However, if B does not depend on s_{nT}, then the best distributor's response $T = T^R(s_{nT})$ does not depend on s_{nT}, i.e. $\frac{\partial T^R}{\partial s_{nT}} = 0$.

The best retailer's response is determined with the standardized base-stock level, $s_{nT} = s_{nT}^R(T) = T\mu_n + \sqrt{T}\sigma_n s_n^*$ (see Proposition 1), and thus,

$$\frac{\partial s_{nT}^R(T)}{\partial T} = \mu_n + \frac{1}{2\sqrt{T}}\sigma_n s_n^* > 0,$$

as stated in the proposition. □

Proof of Proposition 3: Let us substitute T in (12) with the Nash period $T = \beta$. Then, the first term in (12) vanishes as it is identical to B from (11), while the second term is negative, i.e. $\frac{\partial J(\beta)}{\partial T}<0$. Since both $\frac{\partial B}{\partial T} > 0$ and $-\frac{\partial^2 J_{\mathrm{r}}^n}{\partial T^2} > 0$ (see Proposition 1), $\frac{\partial J(T)}{\partial T}$ increases if T increases and thus (12) holds only if the system-wide optimal period $T > \beta$.

Finally, it is shown in Proposition 2 that $\frac{\partial s_{nT}}{\partial T} > 0$, i.e. if $T > \beta$, then $s_{nT} > s_{n\beta}$. □

Proof of Proposition 4: The proof immediately follows from Proposition 2 and Theorem 3. Indeed, the two best response curves $T = T^R(s_{nT})$ and $s_{nT} = s_{nT}^R(T)$ can intersect only once if $\frac{\partial T^R}{\partial s_{nT}} = 0$ and $\frac{\partial s_{nT}^R(T)}{\partial T} > 0$, i.e. a solution determined by Theorem 3 is unique. □