OR Spectrum (2009) 31:337–360
DOI 10.1007/s00291-008-0122-y

REGULAR ARTICLE

An integrated approach to inventory and flexible capacity management subject to fixed costs and non-stationary stochastic demand

Tarkan Tan · Osman Alp

Published online: 24 January 2008

Abstract In a manufacturing system with flexible capacity, inventory management can be coupled with capacity management in order to handle fluctuations in demand more effectively. Typical examples include the effective use of temporary workforce and overtime production. In this paper, we discuss an integrated model for inventory and flexible capacity management under non-stationary stochastic demand with the possibility of positive fixed costs, both for initiating production and for using contingent capacity. We analyze the characteristics of the optimal policies for the integrated problem. We also evaluate the value of utilizing flexible capacity under different settings, which enable us to develop managerial insights.

Keywords Inventory · Production · Stochastic models · Capacity management · Flexible capacity

1 Introduction

A crucial problem that manufacturing companies face is how to cope with volatility in demand. For make-to-stock environments, holding safety stocks is the traditional remedy for handling the variability in demand. If demand is non-stationary, as it is in seasonal business environments, then adjusting production capacity dynamically

T. Tan (✉)
Department of Technology Management, Technische Universiteit Eindhoven,
Den Dolech 2, Paviljoen F-07, 5612AZ Eindhoven, The Netherlands
e-mail: t.tan@tue.nl

O. Alp
Department of Industrial Engineering, Bilkent University, 06800 Ankara, Turkey
e-mail: osmanalp@bilkent.edu.tr

is another possible tool. While there is an extensive amount of literature on both of these measures, our aim is to contribute to the relatively limited research that considers both at the same time, which may be necessary especially if the demand is both non-stationary and stochastic. Consequently, in this paper we consider a periodic review make-to-stock production environment under non-stationary stochastic demand.

In most of the traditional production/inventory literature, either an infinite production capacity is assumed or a given finite capacity is considered as a constraint rather than as a decision variable. We relax this assumption in the sense that the flexible capacity level in each period is to be optimized, as well as the amount of production. Capacity can be defined as the total productive capability of all productive resources such as workforce and machinery. These productive resources can be permanent or contingent. We define permanent capacity as the maximum amount of production possible in regular work time by utilizing internal resources of the company, such as existing workforce level on the steady payroll or the machinery owned or leased by the company. Total capacity can be increased temporarily by acquiring contingent resources, which can be internal or external, such as hiring temporary workers from external labor supply agencies, subcontracting, authorizing overtime production, renting work stations, and so on. We refer to this additional capacity acquired temporarily as the contingent capacity. Flexible capacity management refers to adjusting the total production capacity in any period with the option of utilizing contingent resources in addition to the permanent ones.

Capacity decisions can occur in all hierarchies of decision making: strategic, tactical, and operational. Examples of each type of decision are determining how many production facilities to operate, determining the permanent capacity of a facility, and making contingent capacity adjustments. Our focus is on the operational level. For ease of exposition, we refer to the workforce capacity setting in general. More specifically, we use the temporary (contingent) labor jargon to refer to capacity flexibility. Consequently, the problem we consider can be viewed as one where the production is mostly determined by the workforce size, permanent and contingent.

There exists a significant usage of flexible workforce in many countries. For example, 6.6% of the active labor force of the Netherlands was composed of flexible workers (temporary, standby, replacement, and such other workers) in 2003 (Beckers 2005), and 32% of the employees in the manufacturing industry were regularly working overtime in 2004 (Beckers and Siermann 2005). US Bureau of Labor Statistics indicates that in February 2005 there were 14.8 million flexible workers (independent contractors, on-call workers, temporary help agency workers, and workers provided by contract firms) constituting 10.7% of total employment (US Bureau of Labor Statistics 2006). Aside from the workers with alternative work arrangements as indicated above, contingent workers (temporary jobs) accounted for 4.1% of the total US employment. In March 2006, 7.9% of the active labor force in Turkey was composed of contingent workers (Turkish Statistical Institute 2006).

Temporary workers can be hired in any period and they are paid only for the periods they work, whereas permanent workers are on a payroll. According to our experience, the lead time of acquiring temporary workers is generally short for jobs that do not require high skill levels. It takes as little as 1 or 2 days to acquire temporary workers from the external labor supply agencies, whereas in some cases temporary labor acquisition is actually practically immediate. In some developing countries, abundant temporary workers that are looking for a job and the companies that are in need of temporary labor gather in some venues early in the morning and the companies hire workers that they are going to make use of that very day.

Changing the level of permanent capacity as a means of coping with demand fluctuations, such as hiring and/or firing permanent workers frequently, is not only very costly in general, but it may also have many negative impacts on the company. In case of labor capacity, the social and motivational effects of frequent hiring and firing makes this tool even less attractive. Utilizing flexible capacity, such as hiring temporary workers from external labor supply agencies, is a means of overcoming these issues, and we consider this as one of the two main operational tools of coping with fluctuating demand, along with holding inventory. Nevertheless, long-term changes in the state of the world can make permanent capacity changes unavoidable. Consequently, we consider the determination of the permanent capacity level as a tactical decision that is made at the beginning of a finite planning horizon and not changed until the end of the horizon. This decision is kept out of the scope of this study since we focus only on the operational decisions. We refer the reader to Alp and Tan (2008) for the problem of determining the optimal permanent capacity level.

While flexibility provides many benefits, it comes at a cost. The nominal variable cost of the flexible resources are likely to be higher than that of the permanent ones. In the workforce context, the productivity of temporary workers is probably less than that of permanent workers. Finally, there may be fixed costs associated with ordering contingent capacity, such as the costs of contacting external labor supply agencies and training costs. Another type of fixed costs in the problem environment that we introduce might be the costs associated with initiating production in each period, such as setup costs, overhead costs, and the like.

In many environments, it is possible that the contingent capacity that can be acquired in any period is limited, which is the case when—for example—overtime production is the only tool for temporarily increasing production capacity, due to reasons such as unavailability of temporary workers or special skill requirements. Such environments also fall within our setting, as we allow for an upper bound in the amount of contingent capacity that can be acquired in any period. In the context of overtime production, overtime workers would correspond to contingent capacity, and initiating an overtime production would correspond to ordering contingent capacity subject to a fixed cost associated with facilities, managerial overhead, and the like.

Our contributions is this paper can be summarized as follows:

- We show the equivalence of the problem that we consider with the classical capacitated production/inventory problem which has a piecewise-linear production cost function.
- We characterize the optimal ordering policy when the setup costs are negligible. When the setup costs are positive, we derive the optimal policy for the single period problem. For the multi-period problem, we discuss some properties of the optimal solution through numerical studies.
- Finally, we build several managerial insights as to the use and value of capacity flexibility.

The rest of the paper is organized as follows. We present a review of relevant literature in Sect. 2 and present our dynamic programming model in Sect. 3. The optimal policy for the integrated problem is discussed in Sect. 4, and the value of utilizing flexible capacity is analyzed in Sect. 5. We summarize our conclusions and suggest some possible extensions in Sect. 6.

2 Related literature

The problem that we deal with has interactions with a number of related fields: (i) integrated production/capacity management, (ii) workforce planning and flexibility, and (iii) capacitated production/inventory models. Instead of providing a detailed literature survey on each of these fields, we cite examples of related work from each of them and discuss some of the similarities and differences between those problems and the one that we consider.

An excellent survey of strategic capacity management problems focusing mainly on the capacity expansion decisions is presented by Van Mieghem (2003). The author discusses prominent issues in formulating and solving various capacity management problems.

Atamtürk and Hochbaum (2001) deal with an integrated capacity and inventory management problem under a finite planning horizon and deterministic demand, where trade-offs between capacity expansions, subcontracting, production, and inventory holding are exploited. Angelus and Porteus (2002) also deal with an integrated problem for a short-life-cycle product where the demand has a stochastically increasing and then decreasing structure. Authors show that the optimal capacity level follows a target interval policy. In another work that deals with integrated problems, Dellaert and De Kok (2004) show that integrated capacity and inventory management approaches outperform decoupled approaches.

Hu et al. (2004) deal with an environment similar to ours: There is a fixed permanent production capacity, but it can be increased temporarily by using contingent capacity. Unlike ours, those problems are modelled on a continuous-time framework with a demand rate that is Markov-modulated, and no fixed costs are considered. Tan and Gershwin (2004) also deal with a similar problem. The differences are the existence of several subcontracting opportunities with different cost and capacity structures and the demand being dependent on the lead time distribution during out-of-stock periods. Yang et al. (2005) deal with a production/inventory system under uncertain capacity levels and the existence of outsourcing opportunities. In their model, there is no fixed cost of production but a fixed cost is associated with outsourcing. The decision epochs of producing with internal resources and determining the quantity of outsourcing are distinguished in such a way that the outsourcing decision is made only after the capacity is observed, production is materialized, and the demand is observed. By using this model, the authors characterize the optimal outsourcing policy under certain conditions and elaborate on the optimal policy of production with internal resources.

In the workforce planning and flexibility field, Holt et al. (1960) in their seminal work present models that exploit the trade-off between keeping large permanent workforce levels capable of satisfying peak season demands and frequent adjustment of the workforce level to cope with fluctuations. Indeed, this very idea of aggregate production planning problem constitutes the essence of our problem too; nevertheless, we consider non-stationary stochastic demand, unlike the deterministic demand assumption of aggregate production planning models. Wild and Schneeweiss (1993) analyze and compare four "instruments" to cope with fluctuating demand when the capacity is defined in terms of work force level: variation of monthly working time, use of overtime, employment of temporary workers, and use of leased work force. A hierarchical model based on dynamic programming is presented for making rational decisions on the selective use of these instruments.

Milner and Pinker (2001) deal with the design of contracts between firms and external labor supply agencies for hiring long-term and temporary workers under supply and demand uncertainty in a single period environment. In a related work, Pinker and Larson (2003) consider the problem of managing permanent and contingent workforce levels under uncertain demand in a finite planning horizon where inventory holding is not allowed. The sizes of regular and temporary labor are decision variables that are fixed throughout the planning horizon, but the capacity level may be adjusted by setting the number of shifts for each class of workers.

The papers by Federgruen and Zipkin (1986) and Kapuscinski and Tayur (1998) are two examples of the research stream on capacitated production/inventory problems with stochastic demand, where no fixed costs of production exist. In this case, it is shown that base-stock type policies are optimal. Gallego and Scheller-Wolf (2000) consider the fixed cost of production under a similar environment and partially characterize the optimal policy. Our model is an extension and a generalization of this stream of research, in which we provide the explicit solution of a single period problem.

Finally, the special case of the problem that we consider with no fixed costs is shown in Sect. 3 to translate into the classical capacitated production/inventory problem with a piecewise-linear, convex production cost. When the capacity constraint is neglected, Karlin (1958) characterizes the optimal policy for the case of a strictly convex production cost function as state-dependent order-up-to type . For the single-period problem with a piecewise-linear (hence non-strict) convex production cost and infinite capacity, Porteus (1990) discusses that the optimal policy is of order-up-to type, where the order-up-to level is piecewise-linear increasing in initial inventory level.

3 Model formulation

In this section, we present a finite-horizon dynamic programming model to formulate the problem under consideration. Unmet demand is assumed to be fully backlogged. The relevant costs in our environment are inventory holding and backorder costs, unit cost of permanent and contingent capacity, fixed cost of production, and fixed cost of ordering contingent capacity, all of which are non-negative. We assume that raw material is always available, and that the lead time of production and acquiring contingent capacity can be neglected. We allow for an upper bound on the supply of contingent workers.

We consider a production cost component which is a linear function of permanent capacity in order to represent the costs that do not depend on the production quantity (even when there is no production), such as the salaries of permanent workers. That is, each unit of permanent capacity costs c_p per period, and the total cost of permanent capacity per period is $U \times c_p$, for a permanent capacity of size U, independent of the production quantity. We do not consider material-related costs in our analysis. In order to synchronize the production quantity with the number of workers, we redefine the "unit production" as the number of actual units that an average permanent worker can produce; that is, the production capacity due to U permanent workers is U "unit"s per period. We also define the cost of production by temporary workers in the same unit basis, where the cost for flexible workers is related to their productivity. In particular, let c_c' be the hiring cost of a temporary worker per period, and let c_c'' denote all other relevant variable costs associated with production by temporary workers per period. It is possible that the productivity rates of permanent and temporary workers are different. Let γ be the average productivity rate of temporary workers, relative to the productivity of permanent workers; that is, each temporary worker produces γ units per period. Assuming that this rate remains approximately the same in time, the unit production cost by temporary workers, c_c, can be written as $c_c = (c_c' + c_c'')/\gamma$. It is likely that $0 < \gamma < 1$, but the model holds for any $\gamma > 0$.

Under these settings, it turns out that the production quantity of a period, Q_t, is sufficient to determine the number of temporary workers that need to be ordered in that period, m_t, for any given level of permanent capacity, via $m_t = [(Q_t - U)^+/\gamma]$, ignoring integrality, where $(\cdot)^+$ denotes the value of the argument inside if it is positive and assumes a value of zero otherwise. There also exists an upper bound, V', on the size of contingent capacity that can be ordered at any period, so that $V = V'/\gamma$ defines this limit in terms of the permanent-worker equivalent "unit"s. In the context of overtime production, V would be defined as a proportion of U.

Let K_p denote the fixed cost of production and K_c denote the fixed cost of ordering contingent capacity. K_p is charged whenever production is initiated, even if the permanent workforce size is zero and all production is due to temporary workers. Therefore, it is never optimal to order contingent capacity unless permanent capacity is fully utilized. On the other hand, K_c is charged only when temporary workers are ordered, independent of the amount.

Consequently, the problem translates into the classical capacitated production/inventory problem with a production capacity of $U + V$ and a piecewise-linear production cost (made up of labor costs), which is neither convex nor concave under positive fixed costs. See Fig. 1 for an illustration. Note that when K_p and K_c are both zero, this function is convex.

The order of events is as follows. At the beginning of each period t, the initial inventory level, x_t is observed, the production decision is made and the inventory level is raised to $y_t \leq x_t + U + V$ by utilizing the necessary capacity means; that is, if $y_t \leq x_t + U$ then only permanent capacity is utilized, otherwise a contingent capacity of size $m_t = [(y_t - x_t - U)^+/\gamma]$ is hired on top of full permanent capacity usage. At the end of period t, the demand d_t is met/backlogged, resulting in $x_{t+1} = y_t - d_t$. We denote the random variable corresponding to the demand in period t as $W(t)$ and its distribution function as $G_t(w)$. Consequently, denoting the minimum cost of operating the system from the beginning of period t until the end of the planning horizon as $f_t(x_t)$, we use the following dynamic programming formulation to solve the integrated

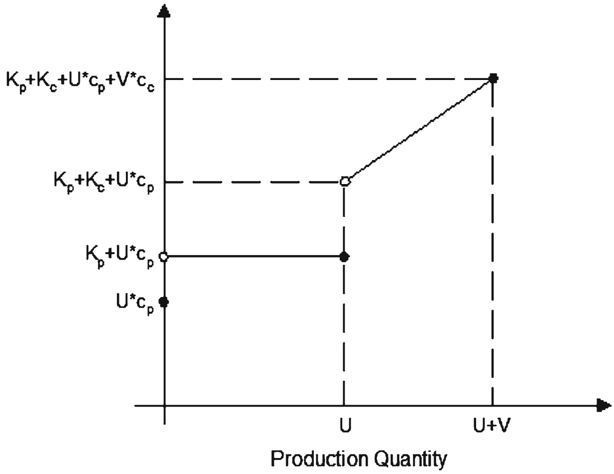

Fig. 1 Production cost function under positive fixed costs

capacity and inventory management problem (CIMP):

$$\text{(CIMP):} \quad f_t(x_t) = Uc_p + \min_{y_t: x_t \le y_t \le x_t + U + V} \{K_p \delta(y_t - x_t) + K_c \delta(y_t - x_t - U) + [y_t - x_t - U]^+ c_c + L_t(y_t) + \alpha E\left[f_{t+1}(y_t - W_t)\right]\} \quad \text{for } t = 1, 2, \ldots, T$$

where $L_t(y_t) = h \int_0^{y_t} (y_t - w) \mathrm{d}G_t(w) + b \int_{y_t}^{\infty} (w - y_t) \mathrm{d}G_t(w)$ is the regular loss function, and $\delta(\cdot)$ is the function that attains the value 1 if its argument is positive, and zero otherwise. We assume the ending condition to be $f_{T+1}(x_{T+1}) = 0$. We note that the optimal contingent capacity usage is independent of c_p, because the permanent capacity is fixed and cannot be changed, and hence the costs of holding that capacity is "sunk".

Remark 1 CIMP reduces to a capacitated production/inventory problem with a concave production cost when $V = 0$ or $K_c \to \infty$ or $c_c \to \infty$.

4 Analysis

The characteristics of the problem and the optimal solution show significant differences depending on whether the fixed costs are strictly positive or not. Therefore, we analyse those two cases separately in the following two subsections:

4.1 Analysis with no fixed costs

In what follows we characterize the optimal policy for the problem without fixed costs. We start by defining the auxiliary functions: $J_t^p(y) = L_t(y) + \alpha E\left[f_{t+1}(y - W_t)\right]$

and $J_t^c(y|x) = J_t^p(y) + c_c(y - x - U)$. Then the integrated capacity and inventory management problem without fixed costs (CIMP-NF) can be stated as

$$\text{(CIMP-NF):} \qquad f_t(x_t) = Uc_p + \min_{y_t : x_t \le y_t \le x_t + U + V} \{J_t(y_t|x_t)\}$$

where

$$J_t(y|x) = \begin{cases} J_t^p(y) & \text{if } y \le x + U \\ J_t^c(y|x) & \text{if } y \ge x + U \end{cases}$$

for $t = 1, 2, \ldots, T$. Recalling that $f_{T+1}(x) = 0$, $J_T^p(y)$ is convex[1] in $y \in \Re$ since the loss function $L_T(y)$ is a convex function. $J_T^c(y|x)$ is also a convex function in $y \in \Re$ for a given value of x. Hence, the first order condition is sufficient for the minimization, where

$$\frac{\mathrm{d}J_T^p(y)}{\mathrm{d}y} = (h + b)G_T(y) - b \quad \text{and} \quad \frac{\mathrm{d}J_T^c(y|x)}{\mathrm{d}y} = (h + b)G_T(y) - b + c_c.$$

Let y_t^p and y_t^c be the minimizers of the functions $J_t^p(y)$ and $J_t^c(y|x)$, respectively, for $t = 1, 2, \ldots, T$. Then, assuming that G is an invertible function and $b \ge c_c$, we have

$$y_T^p = G_T^{-1}\left(\frac{b}{h + b}\right) \quad \text{and} \quad y_T^c = G_T^{-1}\left(\frac{b - c_c}{h + b}\right).$$

Theorem 1 *Optimal ordering policy of CIMP-NF at any period $t = 1, 2, \ldots, T$ is of state-dependent order-up-to type, where the optimal order-up-to level, $y_t^*(x_t)$, is*

$$y_t^*(x_t) = \begin{cases} x_t + U + V & \text{if } \quad x_t \le y_t^c - U - V \\ y_t^c & \text{if } y_t^c - U - V \le x_t \le y_t^c - U \\ x_t + U & \text{if } \quad y_t^c - U \le x_t \le y_t^p - U \\ y_t^p & \text{if } \quad y_t^p - U \le x_t \le y_t^p \\ x_t & \text{if } \quad y_t^p \le x_t \end{cases}.$$

Proof See Appendix.

We note that although y_t^c and y_t^p are independent of x_t, $y_t^*(x_t)$ is a function of x_t. The optimal order-up-to level as a function of x_t is illustrated in Fig. 2 for a single period problem instance.

4.2 Analysis with fixed costs

As discussed in Sect. 3, when the fixed costs in the problem, K_p and K_c, are positive, CIMP translates into a finite-horizon capacitated production/inventory problem with

[1] We note that $f_{T+1}(x) = 0$ is not necessary for the convexity of $J_T^p(y)$. As long as $f_{T+1}(x)$ is convex, this result and the other convexity-related results that we present in this paper hold.

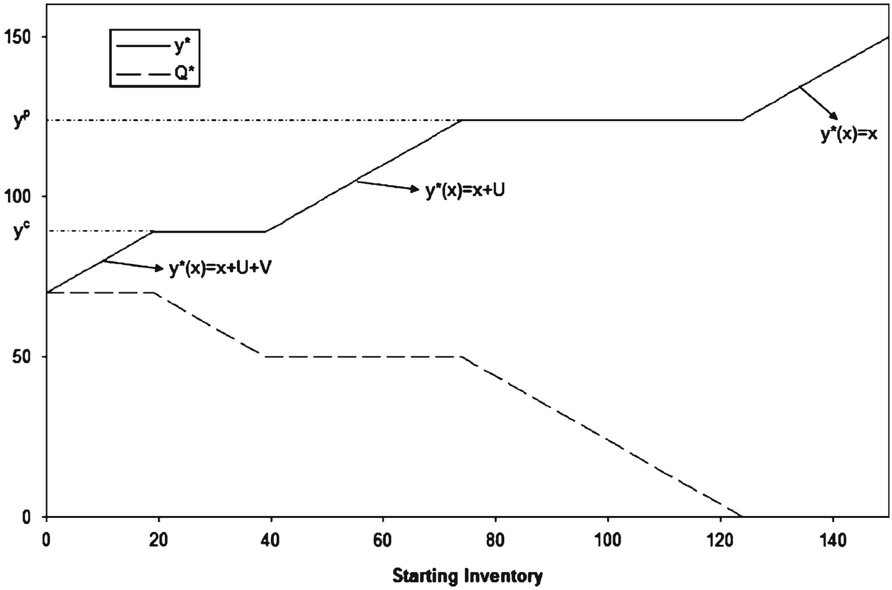

Fig. 2 Illustration of the optimal policy with no fixed costs

neither concave nor convex production costs. In what follows, we characterize the optimal ordering policies for the last period of the planning horizon (single-period problem) and discuss the complications of deriving general optimal policies for multiple periods. We show that the optimal policy for the single-period problem is a state-dependent (s, S)-type policy. Nevertheless, the order-up-to level has a particular structure. It turns out that the optimal order-up-to level attains two distinct scalar values in two separate regions of the starting inventory level x, one involving production with the permanent capacity only and the other involving production with the contingent capacity as well; it is optimal to produce with full permanent and contingent capacity when x is below a certain value and to produce with full permanent capacity at a specific region of x; and no production is necessary for sufficiently high levels of x.

We first note that, due to the convexity of the functions $J_T(y|x)$ and $J_T^p(y)$, there exist unique numbers $s^v(x)$, $s^c(x)$, $s^u(x)$, and s^p for a given value of x such that

$$\begin{aligned}
s^v(x) &= \min\{s : J_T(s|x) = J_T^c(x + U + V|x) + K_p + K_c\},\\
s^c(x) &= \min\{s : J_T(s|x) = J_T^c(y_T^c|x) + K_p + K_c\},\\
s^u(x) &= \min\{s : J_T^p(s) = J_T^p(x + U) + K_p\},\\
s^p &= \min\{s : J_T^p(s) = J_T^p(y_T^p) + K_p\}.
\end{aligned}$$

In order to simplify the exposition of the optimal policy in the last period, we introduce two auxiliary functions, $s(x)$ and $S(x)$, as follows:

$$s(x) = \begin{cases} \max\{s^u(x), s^v(x)\} & \text{if} \quad x \le y_T^c - U - V \\ \max\{s^u(x), s^c(x)\} & \text{if} \quad y_T^c - U - V \le x \le y_T^c - U \\ s^u(x) & \text{if} \quad y_T^c - U \le x \le y_T^p - U \\ s^p & \text{if} \quad y_T^p - U \le x \end{cases}$$

$$S(x) = \begin{cases} x + U + V & \text{if } s(x) = s^v(x) \\ y_T^c & \text{if } s(x) = s^c(x) \\ x + U & \text{if } s(x) = s^u(x) \\ y_T^p & \text{if } s(x) = s^p \end{cases}$$

Theorem 2 *The optimal ordering policy of the last period is*

$$y_T^*(x) = \begin{cases} S(x) & \text{if } x \le s(x) \\ x & \text{otherwise} \end{cases}.$$

Proof See Appendix.

In this optimal policy, there exists a state-dependent reorder level which is a function of the starting inventory level, x. This function, $s(x)$, takes the value of one of the critical levels $s^v(x)$, $s^c(x)$, $s^u(x)$, or s^p, where production only pays off when the starting inventory is below that level, due to the existence of fixed costs. Critical levels $s^v(x)$, $s^c(x)$, $s^u(x)$, and s^p can be considered as the "reorder" levels for production with full permanent and contingent capacity, production with partial contingent capacity, production with full permanent capacity, and production with idle capacity, respectively. Therefore, if the starting inventory level is below the reorder level, the optimal order-up-to level is given by $x + U + V$, y_T^c, $x + U$, or y_T^p depending on the value that $s(x)$ takes, either $s^v(x)$, $s^c(x)$, $s^u(x)$, or s^p, respectively. Otherwise, no orders are placed. The optimal policy is illustrated in Fig. 3.

The following result characterizes the structure of the reorder level function with respect to the starting inventory level, x.

Theorem 3 *$s(x)$ is non-decreasing in x.*

Proof See Appendix.

While it is more likely that the problem environment that we have discussed has a multi-period structure, there are also some problem environments where the single-period model is appropriate. When the demand for the product is mostly observed in a condensed time interval, such as the Christmas period, the single-period model is clearly relevant. As a matter of fact, one of our motivations in initiating this research came from a CD-producer which faces elevated demand during Christmas season, since CDs are popular gift items. The company utilizes temporary labor for this extra production, particularly for manual CD boxing and packaging, which is required for all CDs except for the "regular" CDs that are boxed in standard 1-CD jewel boxes. Another application would be a product selection problem where the item(s) to produce in a finite planning horizon must be selected. In that case, this tactical level decision could be handled by representing the total planning horizon as a single-period. Specifically, such a firm may be capable of producing a variety of items but is able to produce

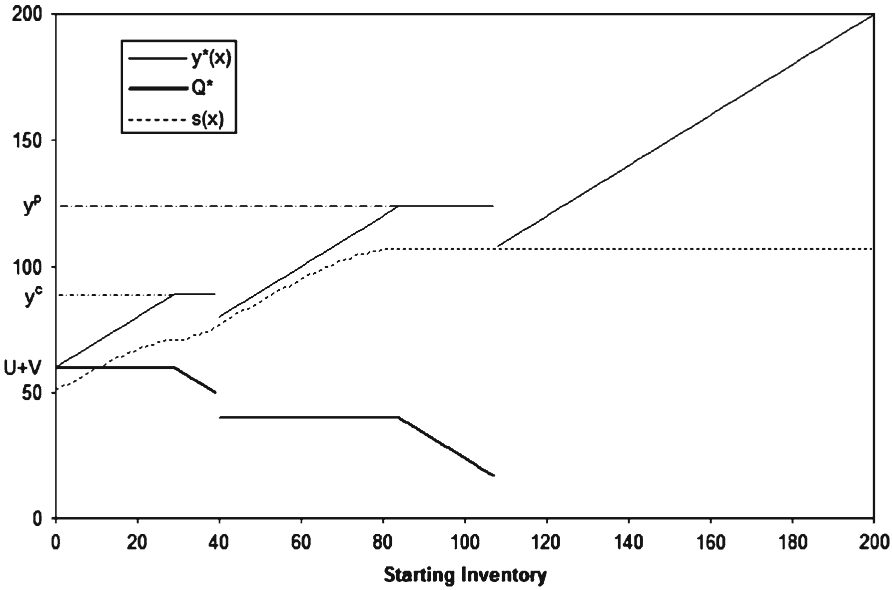

Fig. 3 Illustration of the optimal policy with fixed costs

only one or a small number of them due to scarce, possibly shared resources. Moreover, utilizing flexible resources such as overtime or temporary labor usage may be inevitable due to capacitated permanent productive resources. In such environments, product selection decisions might be based on the expected total costs of producing items (along with the associated revenues) depending on their probabilistic demand behaviors, starting inventories, productivity of permanent and temporary resources, operating costs, and availability of flexible resources. The single-period problem could be solved for all possible (combinations of) alternatives—with K_p denoting the total fixed costs of production changeovers from one planning horizon to the other—in order to make the decision as to how much of which item(s) should be produced.

The optimal policy of the last period cannot simply be generalized to multiple periods. For example, Scarf (1959) generalizes the optimality of (s, S) policy of the single-period in the standard production/inventory problem with a fixed cost to the multiple period case by making use of K-convexity in the inductive arguments. In our problem, even for the special case of $K_p = 0$, K_c-convexity does not necessarily hold. Similarly, quasiconvexity fails to be preserved from the last period backwards. For the other special case of $K_c = 0$, K_p-convexity or quasiconvexity does not necessarily hold either. Gallego and Scheller-Wolf (2000) consider a capacitated production/inventory problem under fixed costs of production, which is a special case of CIMP as mentioned in Remark 1. They state that the full characterization of the optimal policies for this special case is extremely difficult but the authors could generate four regions of the starting inventory level where optimal production decisions are characterized to a certain extent. We conjecture the optimal ordering policy of CIMP to be even more complicated than that of this special case.

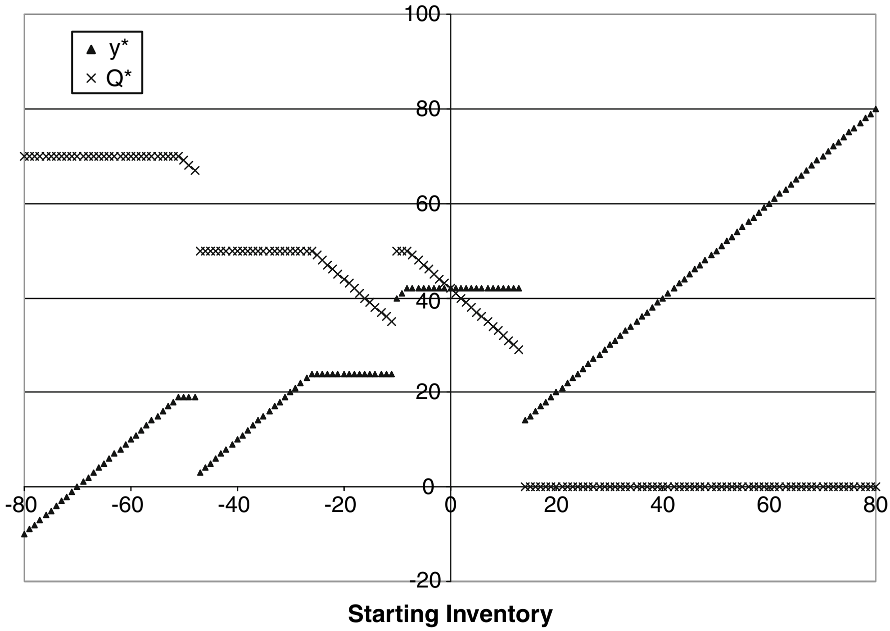

Fig. 4 Starting inventory versus optimal production quantity for three periods to go

We also observe that a monotonic relation between the starting inventory level (x) and the optimal production quantity in CIMP does not necessarily exist. An example of this result is depicted in Fig. 4, which is the result of a 3-period problem instance with $K_p = 30$, $K_c = 40$, $U = 50$, $V = 20$, $c_p = 1.5$, $c_c = 2.5$, $b = 5$, $h = 1$, where the demand is Poisson with an expected value of 20. We note that for this problem instance the optimal production quantity has a non-increasing structure in x up to $x = -11$, where the optimal production quantity is 35. However, when $x = -10$ the optimal decision is to produce 50 units due to complex dynamics involving the fixed costs and the permanent capacity, and this violates the monotonic structure. Specifically, when $x = -10$, the expected total cost of production with full permanent capacity of 50 in this period turns out to be less than that of producing 34 due to the opportunity of avoiding a prospective fixed cost in some future period.

Figure 4 also shows the optimal order-up-to levels with respect to x. We observe that there are three distinct values of optimal order-up-to levels on this figure other than x (no production), $x + U$ (full permanent capacity production), and $x + U + V$ (full permanent and contingent capacity production). Note that there exist only two distinct critical order-up-to points (other than x, $x + U$, and $x + U + V$) in the optimal policy stated for the last period.

5 Value of flexible capacity

Our purpose in this section is to investigate the value of utilizing flexible capacity and its sensitivity to change in the following system parameters: fixed cost of production, unit cost of permanent capacity, unit and fixed costs of contingent capacity, backorder cost, the size of permanent capacity, upper bound on the size of contingent capacity that can be acquired (which we refer to as "maximum contingent capacity" hereafter), and the variability of the demand. We compare a flexible capacity (FC) system with an inflexible one (IC), where the contingent capacity can be utilized in the former but cannot in the latter.

We define the (absolute) value of flexible capacity, VFC, as the difference between the optimal expected total cost of operating the IC system, ETC_{IC}, and that of the FC system, ETC_{FC}. That is, $\text{VFC} = \text{ETC}_{\text{IC}} - \text{ETC}_{\text{FC}}$. In order to reflect the relativity, we also define the relative value of flexible capacity as $\%\text{VFC} = \text{VFC}/\text{ETC}_{\text{IC}}$. We note that both VFC and %VFC are always non-negative.

We conduct some numerical experiments to reveal the sensitivity of VFC and %VFC with respect to changes in the rest of the system parameters. In the IC case of these numerical examples, we simply take $V = 0$ (see Remark 1). We solve CIMP for the following set of input parameters, unless otherwise noted: $T = 12$, $U = 10$, $V = 4$, $b = 5$, $h = 1$, $c_c = 2.5$, $c_p = 1.5$, $K_p = 0$, $K_c = 10$, $\alpha = 0.99$, and $x_1 = 0$. We set $K_p = 0$ to represent a production environment with negligible fixed cost of production and $V = 4$ to represent an upper bound on the contingent capacity that can be acquired (for example, in the form of producing overtime which increases the capacity by at most 40% or limited temporary worker supply). We consider demand that follows a seasonal pattern with a cycle of 4 periods, where the expected demand is 15, 10, 5, and 10, respectively. Note that the permanent capacity is the same as the average demand for this base set. We first assume that the demand has a Poisson distribution. In the results that we present, we use the term "increasing" (decreasing) in the weak sense to mean "non-decreasing" (non-increasing). We provide intuitive explanations to all of our results below and our findings are verified through several numerical studies. However, like any experimental result, one should be careful about generalizing them, especially for extreme values of problem parameters.

5.1 Effects of backorder and contingent capacity costs

We first test the value of flexibility with respect to the backorder cost by varying the value of b between 2 and 10. Table 1 verifies intuition in the sense that %VFC is higher when backorders are more costly, or equivalently, when higher service levels are targeted.

In order to explore the effects of fixed and unit costs of contingent capacity on the value of flexibility, it suffices to revert to the definition of VFC and %VFC. Then, it follows that the value of flexible capacity increases as the contingent capacity becomes

Table 1 % Value of flexible capacity versus backorder cost under different maximum contingent capacity levels

b	3	4	5	6	7	8	9	10
%VFC ($V = 4$)	18.90	26.89	33.39	38.73	43.24	47.04	50.32	53.15
%VFC ($V = 10$)	20.62	29.07	36.00	41.66	46.45	50.48	53.93	56.93

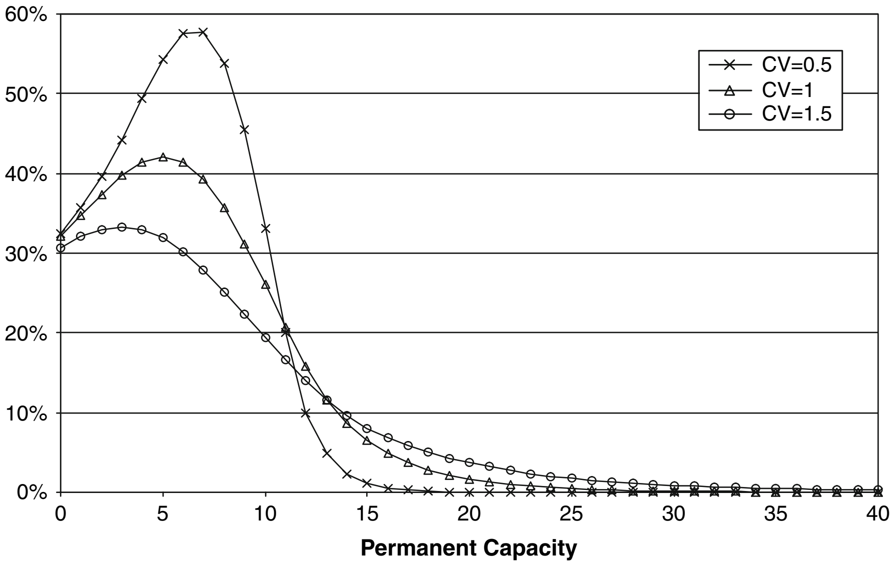

Fig. 5 % Value of flexible capacity versus permanent capacity size under different CVs when $V = 4$

less costly to utilize. It is easy to see why this relation holds: As c_c or K_c decreases while the other parameters are kept constant, $\mathrm{ETC_{IC}}$ remains the same, since flexible capacity is not utilized in this case, but $\mathrm{ETC_{FC}}$ decreases due to decreased costs. Consequently, both VFC and %VFC increases.

5.2 Effects of demand variability and the size of permanent capacity

In this section we assume Gamma distribution for demand, since testing the impact of change in demand variance for a given expected demand is not possible with a Poisson distribution. We also replicated the tests using Normal distribution (with smaller CV values). We obtained similar results, therefore we do not report them here. We refer to Figs. 5 and 6 for the illustration of our findings.

We first observe that VFC increases as the permanent capacity decreases, as can be expected, since the inadequacy of permanent capacity increases the requirement for contingent capacity. The value of flexibility becomes extremely high for low levels of permanent capacity, because of elevated backorder costs in the case of IC. Nevertheless, the monotonic behavior does not necessarily hold for %VFC, as shown in Fig. 5. This is because the rate of increase in $\mathrm{ETC_{IC}}$ as U decreases dominates the rate of increase in $\mathrm{ETC_{IC}} - \mathrm{ETC_{FC}}$ when the permanent capacity is extremely low, due to cumulative backordering even with the use of maximum contingent capacity. However, if the maximum contingent capacity (V) is high enough, then the behavior described for VFC also holds for %VFC, as shown in Fig. 6, where no limit on the supply of contingent capacity is assumed.

Next, we test the impact of changes in demand variance for a given expected demand. We vary the coefficient of variation (CV) values between 0.5 and 1.5. The

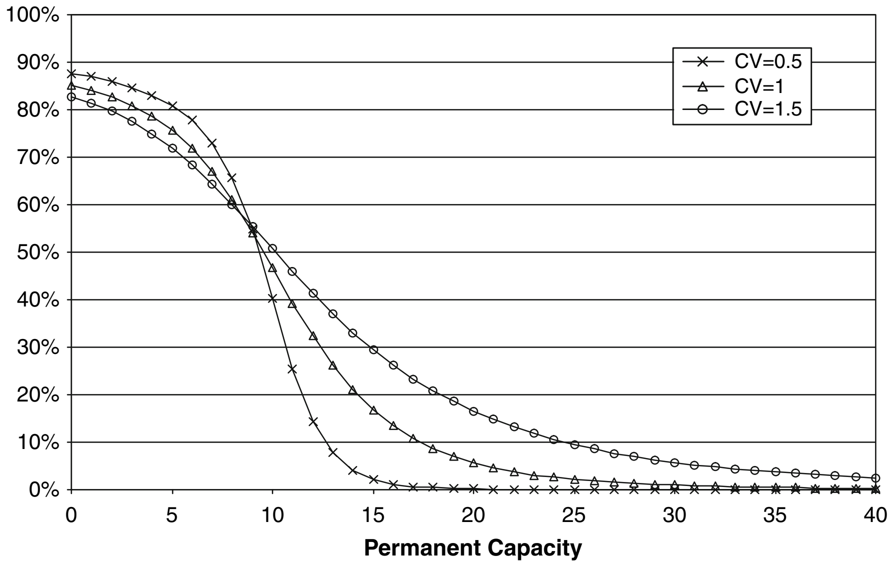

Fig. 6 % Value of flexible capacity versus permanent capacity size under different CVs when $V \rightarrow \infty$

results are surprising, in the sense that the value of flexible capacity does not always increase as variance increases. The reason for this is as follows. While both of the ETC_{IC} and ETC_{FC} terms do increase as demand variance increases, the increase in ETC_{IC} term is less than that in ETC_{FC} under low capacity. This is because most of the demand is backlogged anyway for the IC case, whereas the expected total cost increases more significantly for the flexible system. This behavior also holds for %VFC as shown in Figs. 5 and 6. (We note that the condition discussed here is sufficient for %VFC to decrease, but not necessary. This is because %VFC would decrease even when VFC increases, as long as the increase in VFC is in a lower rate than that in ETC_{IC}.) However, when the permanent capacity is sufficient to meet the demand on the average, the value of flexibility increases as demand variance increases.

Another observation that we make is that the value of flexible capacity continues to be very significant under high demand variance, even when the permanent capacity is much higher than average demand, especially when the maximum contingent capacity is high enough. For example, when $CV = 1.5$ and no limit on the supply of contingent capacity is assumed, %VFC equals 15.3% for $U = 20$, which is twice the size of average demand.

5.3 Effects of permanent capacity costs

As discussed in Sect. 3, the optimal contingent capacity usage is independent of c_p for any given value of U. Hence, VFC remains unchanged for all values of c_p when other problem parameters are kept constant. However, the relative value, %VFC, decreases for higher c_p, because both of the expected costs increase, and therefore ETC_{IC} does as well.

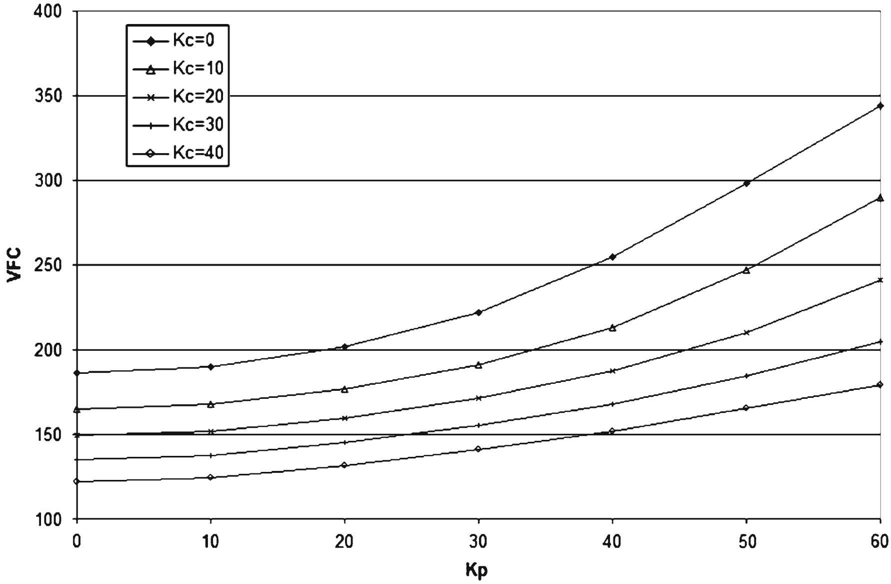

Fig. 7 Value of flexible capacity versus fixed costs

In order to investigate the effect of the fixed cost of production, K_p, we vary K_p between 0 and 60. We also vary the values of the fixed cost of contingent capacity (K_c) between 0 and 40, to illustrate the corresponding result in Sect. 5.1. Figure 7 depicts that VFC increases as K_p increases because flexibility becomes more crucial as initiating production becomes more costly.

5.4 Effects of maximum contingent capacity

As the maximum contingent capacity that can be acquired (V) increases, the value of flexibility also increases. This holds also for %VFC, as demonstrated in Fig. 8. We observe that introducing the use of contingent capacity has an immediate effect of decreasing overall operating costs considerably up to a certain point—the impact and duration of this effect depends on the permanent capacity size, as well as the values of other problem parameters—but above some value of V the value added by any additional contingent capacity becomes negligible. The marginal value of contingent capacity decreases as the size of the permanent capacity increases, because in such a case more production can be handled by permanent resources.

In the context of overtime production, V might be defined as a percentage of U, as stated in Sect. 3. For low values of permanent capacity, overtime production may be very valuable even when moderate proportions of the permanent resources are used for overtime production. In the problem instances presented in Fig. 8 for example, when $U = 8$ we observe that there exists a saving potential of 44% when V is set as 25% of U throughout the planning horizon. Even when operating with a permanent capacity that is sufficient to meet the expected demand, we observe in this particular example

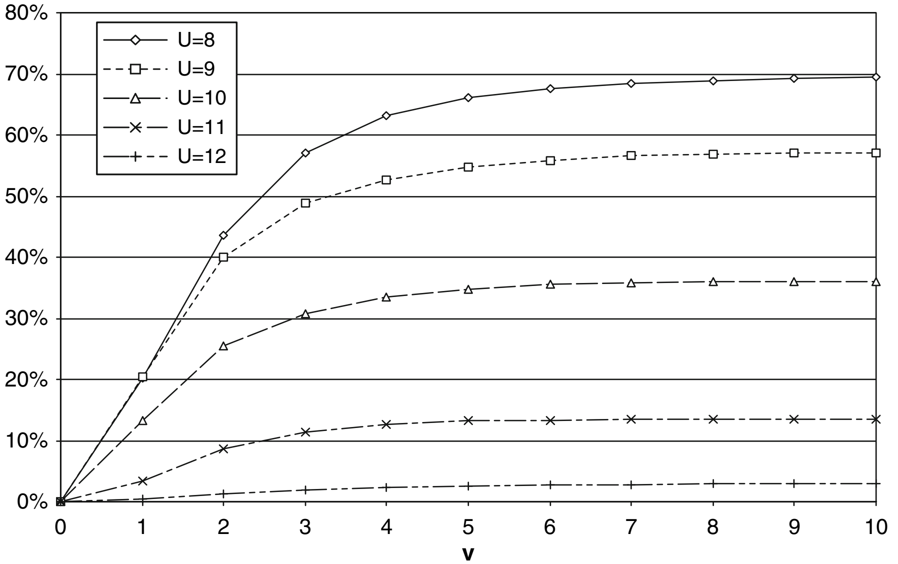

Fig. 8 % Value of flexible capacity versus maximum contingent capacity under different permanent capacity levels

that a 31% reduction in the expected operating costs is possible by utilizing 30% of the permanent resources for overtime production. Finally, we note that even though the value of overtime becomes less significant for larger U values, considerable savings can still be obtained even with low V values and a positive fixed cost associated with utilizing contingent capacity (%VFC is 9% when $U = 11$ and $V = 2$ in Fig. 8).

6 Conclusions and future research

In this paper, we consider the integrated problem of inventory and flexible capacity management under non-stationary stochastic demand with the possibility of positive fixed costs, both for initiating production and ordering contingent capacity. While a workforce capacity jargon is adopted in some parts of the paper, the model can be applied to any other bottleneck resource that defines the capacity as well, when there exists the option of increasing this capacity by making use of contingent resources. The amount of contingent capacity that can be acquired is allowed to be bounded from above, in order to represent situations such as limited availability of temporary workers supplied by the external labor supply agency, limited working space, and overtime production capacity. In that setting, our model and analysis also provide insights on—for example—the optimal usage of overtime production coupled with inventory management.

For the environment that we consider, the equivalence of this problem with the classical capacitated production/inventory problem which has a specified production cost function is shown. When the setup costs are negligible, the optimal policy depends on the level of starting inventory and it is a variant of base-stock policy, in the sense that there are two order-up-to levels in each period: one of them can be attained by utilizing contingent capacity, and the other can be attained by utilizing only permanent capacity. There is also a region of starting inventory level where full permanent capacity and no contingent capacity should be utilized.

When the fixed costs are positive, the optimal policy for the single-period problem is shown to be a variant of (s, S) type policies where the policy parameters are functions of the starting inventory level and available capacity limits. With the help of a problem instance, it is shown that the optimal policy for the single-period problem, as well as the monotonic relationship between the starting inventory level and the optimal production quantity, ceases to hold for the multi-period problem. This does not mean that the solution of the multi-period problem cannot be characterized; that is still an open question. Nevertheless, even a famous special case of this problem, the capacitated multi-period inventory/production problem with positive fixed cost, has not yet been completely characterized, which makes it likely that the optimal policy does not have a very simple structure.

Our computational analyses point out important directions for the use of flexibility. Flexibility turns out to be very important for some values of problem parameters: (i) lower costs of contingent capacity, (ii) higher costs of backorders, (iii) higher fixed costs of production, (iv) lower levels of permanent capacity, and (v) higher levels of maximum contingent capacity. Moreover, for businesses with demand volatility, the value of flexibility might be extremely high even under abundant permanent capacity levels. Under such circumstances, businesses should pursue establishing long-term contractual relations with third-party contingent capacity providers (such as external labor supply agencies). Such long-term agreements would bring significant operational cost savings. On the other hand, the opposite range of parameters yield relatively low value for flexibility, which is also very important from a managerial point of view. Under such circumstances, there does not exist enough motivation to invest in capacity flexibility, since the existing resources are sufficient for reasonable management of operations.

This research can be extended in several ways. One way to enrich the model is to relax our assumptions on the capacity usage, such as introducing uncertainty on contingent and/or permanent capacity, as well as to relax our assumption of zero lead times. Other extension possibilities are investigating interactions with material availability; determining the optimal permanent and maximum contingent capacity levels at the beginning of the problem horizon (given that reserving contingent capacity comes with a cost); incorporating intentional changes in capacity levels in strategic, tactical, or even operational level; exploring the structure of the optimal solution under seasonal demand; and developing an efficient heuristic for the multi-period problem with fixed costs.

Appendix

Proof of Theorem 1 First, we need the following preliminary results for the proof:

Lemma 1 $y_T^c \leq y_T^p$.

Proof $y_T^c \leq y_T^p$ since $\frac{b-c_c}{h+b} \leq \frac{b}{h+b}$ and $G_T(y)$ is a non-decreasing function. □

Lemma 2 *$J_T(y|x)$ is convex in y for any value of x. Moreover, y_T^* is the minimizer of $J_T(y|x)$ where*

$$y_T^* = \begin{cases} y_T^c & \text{if} \quad x \leq y_T^c - U \\ x+U & \text{if} \quad y_T^c - U \leq x \leq y_T^p - U \\ y_T^p & \text{if} \quad y_T^p - U \leq x \end{cases}.$$

Proof First note that $J_T^p(y) = J_T^c(y|x)$ at $y = x + U$. Moreover, $\frac{\mathrm{d}J_T^p(y)}{\mathrm{d}y} = L_T'(y) \leq L_T'(y) + c_c = \frac{\mathrm{d}J_T^c(y|x)}{\mathrm{d}y}$ for any y. For $y \leq x + U$, $J_T(y|x)$ is convex because $J_T^p(y)$ is convex. At $y = x + U$, $J_T(y|x)$ takes the form of another convex function $J_T^c(y|x)$ and since the derivative of the new function is greater than that of the previous, the convexity is not violated after the change of shape.

By using the definition and the convexity of $J_T(y|x)$ and Lemma 1, it can be observed that the function takes different forms in intervals $x \leq y_T^c - U$, $y_T^c - U \leq x \leq y_T^p - U$, and $y_T^p - U \leq x$ and the minimizer of the function in each of these intervals are y_T^c, $x + U$, and y_T^p respectively. □

Recall that $f_T(x) = Uc_p + \min_{y:x \leq y \leq x+U+V}\{J_T(y|x)\}$. Since Uc_p is a constant term we suppress it from the definition of $f_t(x)$ in what follows. Hence, $f_T(x)$ is determined by minimizing $J_T(y|x)$ over $x \leq y \leq x + U + V$. Therefore, due to Lemma 2 and the boundaries of the optimization, the optimal policy for the last period is as follows.

$$y_T^*(x) = \begin{cases} x+U+V & \text{if} \quad x \leq y_T^c - U - V \\ y_T^c & \text{if} \quad y_T^c - U - V \leq x \leq y_T^c - U \\ x+U & \text{if} \quad y_T^c - U \leq x \leq y_T^p - U \\ y_T^p & \text{if} \quad y_T^p - U \leq x \leq y_T^p \\ x & \text{if} \quad y_T^p \leq x. \end{cases}$$

In order this policy to hold for all other periods in the planning horizon, similar results as in Lemma 1 and 2 must hold for each of these periods: (i) $J_t(y|x)$ is convex over y for a given x, and (ii) $y_t^c \leq y_t^p$ for every period $t \leq T$.

By using the optimal policy of period T, it can be shown that

$$f_T(x) = \begin{cases} J_T^c(x+U+V|x) = c_c V + L_T(x+U+V) & \text{if} \quad x \le y_T^c - U - V \\ J_T^c(y_T^c|x) = c_c(y_T^c - x - U) + L_T(y_T^c) & \text{if} \quad y_T^c - U - V \le x \le y_T^c - U \\ J_T^p(x+U) = L_T(x+U) & \text{if} \quad y_T^c - U \le x \le y_T^p - U \\ J_T^p(y_T^p) = L_T(y_T^p) & \text{if} \quad y_T^p - U \le x \le y_T^p \\ J_T^p(x) = L_T(x) & \text{if} \quad y_T^p \le x, \end{cases}$$

and

$$\frac{df_T(x)}{dx} = \begin{cases} L_T'(x+U+V) & \text{if} \quad x \le y_T^c - U - V \\ -c_c & \text{if} \quad y_T^c - U - V \le x \le y_T^c - U \\ L_T'(x+U) & \text{if} \quad y_T^c - U \le x \le y_T^p - U \\ 0 & \text{if} \quad y_T^p - U \le x \le y_T^p \\ L_T'(x) & \text{if} \quad y_T^p \le x. \end{cases}$$

We first note that $f_T(x)$ is continuous. Next we check the continuity of $\mathrm{d}f_T(x)/\mathrm{d}x$. We have $\lim_{x \to (y_T^c - U - V)^-} \mathrm{d}f_T(x)/\mathrm{d}x = L_T'(y_T^c) = (h+b)G_T(y_T^c) - b = -c_c$, because $y_T^c = G_T^{-1}\left(\frac{b-c_c}{h+b}\right)$. We also have $\lim_{x \to (y_T^c - U)^+} \mathrm{d}f_T(x)/\mathrm{d}x = L_T'(y_T^c) = -c_c$. Finally, $\lim_{x \to (y_T^p - U)^-} \mathrm{d}f_T(x)/\mathrm{d}x = L_T'(y_T^p) = (h+b)G_T(y_T^p) - b = 0$, because $y_T^p = G_T^{-1}\left(\frac{b}{h+b}\right)$, and similarly, $\lim_{x \to (y_T^p)^+} \mathrm{d}f_T(x)/\mathrm{d}x = L_T'(y_T^p) = 0$. Since L is convex and is minimized at y_T^p, we conclude that $\mathrm{d}f_T(x)/\mathrm{d}x$ is continuous and non-decreasing, and consequently the functions $f_T(x)$ and $E[f_T(x)]$ are convex in x. Starting with period T and assuming that the optimal policy and the convexity of $E[f_k(x)]$ hold for all periods $k = T, T-1, \ldots, t+1$, it follows that $J_t(y|x)$ is convex over y for a given x. Let $j(y) = \mathrm{d}J_t^p(y)/\mathrm{d}y$ then $y_t^p = j^{-1}(0)$. Moreover, $y_t^c = j^{-1}(-c_c)$ since $\mathrm{d}J_t^c(y|x_t)/\mathrm{d}y = j(y) + c_c$. Hence, due to the convexity of function J_t^p, we have $y_t^c \le y_t^p$ and

$$y_t^*(x_t) = \begin{cases} x_t + U + V & \text{if} \quad x_t \le y_t^c - U - V \\ y_t^c & \text{if} \quad y_t^c - U - V \le x_t \le y_t^c - U \\ x_t + U & \text{if} \quad y_t^c - U \le x_t \le y_t^p - U \\ y_t^p & \text{if} \quad y_t^p - U \le x_t \le y_t^p \\ x_t & \text{if} \quad y_t^p \le x_t. \end{cases}$$

To conclude the proof, note that

$$f_t(x_t) = \begin{cases} J_t^c(x_t+U+V|x_t) = c_c V + J_t^p(x_t+U+V) & \text{if } x_t \le y_t^c - U - V \\ J_t^c(y_t^c|x_t) = c_c(y_t^c - x_t - U) + J_t^p(y_t^c) & \text{if } y_t^c - U - V \le x_t \le y_t^c - U \\ J_t^p(x_t+U) & \text{if } y_t^c - U \le x_t \le y_t^p - U \\ J_t^p(y_t^p) & \text{if } y_t^p - U \le x_t \le y_t^p \\ J_t^p(x_t) & \text{if } y_t^p \le x_t, \end{cases}$$

and

$$\frac{df_t(x_t)}{dx_t} = \begin{cases} j(x_t + U + V) & \text{if } x_t \le y_t^c - U - V \\ -c_c & \text{if } y_t^c - U - V \le x_t \le y_t^c - U \\ j(x_t + U) & \text{if } y_t^c - U \le x_t \le y_t^p - U \\ 0 & \text{if } y_t^p - U \le x_t \le y_t^p \\ j(x_t) & \text{if } y_t^p \le x_t. \end{cases}$$

We note that $\lim_{x_t \to (y_t^c - U - V)^-} \mathrm{d}f_t(x_t)/\mathrm{d}x_t = j(y_t^c) = j(j^{-1}(-c_c)) = -c_c$, and similarly, $\lim_{x_t \to (y_t^c - U)^+} \mathrm{d}f_t(x_t)/\mathrm{d}x_t = j(y_t^c) = -c_c$. Finally, $\lim_{x_t \to (y_t^p - U)^-} \mathrm{d}f_t(x)/\mathrm{d}x = j(y_t^p) = j(j^{-1}(0)) = 0$, and $\lim_{x_t \to (y_t^p)^+} \mathrm{d}f_t(x)/\mathrm{d}x = j(y_t^p) = 0$. Hence, $f_t(x_t)$ is a convex function over x_t. This completes the proof. □

Proof of Theorem 2 Uc_p is a constant term. Note that $s^u(x) \le x + U$ by definition of $s^u(x)$. We need to examine the following cases on the value of x.

Case I: $x \le y_T^c - U - V$ and $s^u(x) \le s^v(x) \le x + U + V$

$J_T(y|x)$ is non-increasing over $x \le y \le x + U + V$ since $x + U + V \le y_T^c$ which is the minimizer of $J_T^c(y|x)$ by Lemma 2. Therefore, the optimal course of action is either not to produce, to produce with full permanent capacity, or to produce with full permanent and contingent capacity. Since $s^u(x) \le s^v(x) \le x + U + V$, $J_T^p(x + U) + K_p = J_T(s^u(x)|x) \ge J_T(s^v(x)|x) = J_T^c(x + U + V|x) + K_p + K_c$. Hence, if production is viable, then it must be up to $x + U + V$. If $x \le s^v(x) = s(x)$, $J_T(x|x) \ge J_T(s^v(x)|x) = J_T^c(x + U + V|x) + K_p + K_c$ by the definition of $s^v(x)$. Therefore, $y_T^*(x) = x + U + V = S(x)$ since $s(x) = s^v(x)$. If $x > s^v(x) = s(x)$, $J_T(s^v(x)|x) \ge J_T(x|x)$ and therefore $y_T^*(x) = x$.

Case II: $x \le y_T^c - U - V$ and $s^v(x) \le s^u(x) \le x + U + V$

$J_T(y|x)$ is non-increasing over $x \le y \le x + U + V$ in this case, too. Since $s^v(x) \le s^u(x) \le x + U + V$, $J_T^c(x + U + V|x) + K_p + K_c = J_T(s^v(x)|x) \ge J_T(s^u(x)|x) = J_T^p(x + U) + K_p$. Hence, if production is viable, then it must be limited to $x + U$; ordering contingent labor does not pay off. If $x \le s^u(x) = s(x)$, $J_T(x|x) \ge J_T(s^u(x)|x) = J_T(x + U|x) + K_p$ by the definition of $s^u(x)$. Therefore, $y_T^*(x) = x + U = S(x)$ since $s(x) = s^u(x)$. If $x > s^u(x) = s(x)$, $J_T(s^u(x)|x) \ge J_T(x|x)$ and therefore $y_T^*(x) = x$.

Case III: $x + U \le y_T^c$ and $s^c(x) \le s^u(x) \le x + U$

$J_T(y|x)$ is minimized at y_T^c by Lemma 2 and therefore is non-increasing over $x \le y \le y_T^c$. Since $s^c(x) \le s^u(x)$, $J_T^c(y_T^c|x) + K_p + K_c = J_T(s^c(x)|x) \ge J_T(s^u(x)|x) = J_T^p(x + U) + K_p$. Hence, if production is viable, then it must be limited to $x + U$; ordering contingent capacity does not pay off. If $x \le s^u(x) = s(x)$, $J_T(x|x) \ge J_T(s^u(x)|x) = J_T^p(x + U|x) + K_p$ by the definition of $s^u(x)$. Therefore, $y_T^*(x) = x + U = S(x)$ since $s(x) = s^u(x)$. If $x \ge s^u(x) = s(x)$, $J_T(s^u(x)|x) \ge J_T(x|x)$ and therefore $y_T^*(x) = x$.

Case IV: $x + U \le y_T^c$ and $s^u(x) \le s^c(x)$

$J_T(y|x)$ is minimized at y_T^c by Lemma 2 and therefore is non-increasing over $x \le y \le y_T^c$. As $s^u(x) \le s^c(x)$, $J_T^p(x + U) + K_p = J_T(s^u(x)|x) \ge J_T(s^c(x)|x) = J_T^c(y_T^c|x) + K_p + K_c$. Hence, if production is viable, then it must be up to y_T^c.

If $x \leq s^c(x) = s(x)$, $J_T(x|x) \geq J_T(s^c(x)|x) = J_T(y_T^c|x) + K_p$ by the definition of $s^c(x)$. Therefore, $y_T^*(x) = y_T^c = S(x)$ since $s(x) = s^c(x)$. If $x \geq s^c(x) = s(x)$, $J_T(s^c(x)|x) \geq J_T(x|x)$ and therefore $y_T^*(x) = x$.

Case V: $y_T^c \leq x + U \leq y_T^p$

$J_T(y|x)$ is minimized at $x + U$ by Lemma 2 and therefore is non-increasing over $x \leq y \leq x + U$. If $x \leq s^u(x) = s(x) \leq x + U$, $J_T(x|x) \geq J_T(s^u(x)|x) = J_T^p(x + U) + K_p$ by the definition of $s^u(x)$. Therefore, $y_T^*(x) = x + U = S(x)$ since $s(x) = s^u(x)$. If $x \geq s^u(x) = s(x)$ then $y_T^*(x) = x$.

Case VI: $y_T^p \leq x + U$

$J_T(y|x)$ is minimized at y_T^p by Lemma 2 and therefore is non-increasing over $x \leq y \leq y_T^p$. If $x \leq s^p = s(x) \leq y_T^p$, then $J_T(x|x) \geq J_T(s^p|x) = J_T^p(y_T^p) + K_p$ by the definition of s^p. Therefore, $y_T^*(x) = y_T^p = S(x)$ since $s(x) = s^p$. If $x \geq s^p = s(x)$ then $y_T^*(x) = x$. □

Proof of Theorem 3 First note that $J_T(y|x) \geq J_T(y|x + \Delta)$ for all y and $\Delta > 0$, since

- For $y < x + U$: $J_T(y|x) = J_T(y|x + \Delta)$
- For $x + U \leq y < x + \Delta + U$: $J_T(y|x) = J_T^p(y) + c_c(y - x - U) \geq J_T^p(y) = J_T(y|x + \Delta)$
- For $x + \Delta + U \leq y$: $J_T(y|x) = J_T^p(y) + c_c(y - x - U) \geq J_T^p(y) + c_c(y - x - \Delta - U) = J_T(y|x + \Delta)$

We need to examine the following cases on the values of x and $s(x)$:

Case I $s^u(x) \leq s^v(x)$, $x + U + V \leq y_T^c \leq y_T^p$

Let $\Delta > 0$ be such that $x + \Delta + U + V \leq y_T^c$. $J_T(s^v(x)|x) = J_T(x + U + V|x) + K_p + K_c \geq J_T(x + \Delta + U + V|x) + K_p + K_c \geq J_T(x + \Delta + U + V|x + \Delta) + K_p + K_c = J_T(s^v(x + \Delta)|x + \Delta)$. Therefore

$$J_T(s^v(x)|x) \geq J_T(s^v(x + \Delta)|x + \Delta).$$

We have four possible cases for the values of $s^v(x)$ and $s^v(x + \Delta)$ with respect to $x + U$ and $x + \Delta + U$:

Case I-a $s^v(x) \leq x + U$, $s^v(x + \Delta) \leq x + \Delta + U$

$J_T(s^v(x)|x) = J_T^p(s^v(x)) \geq J_T(s^v(x + \Delta)|x + \Delta) = J_T^p(s^v(x + \Delta))$ since J_T^p is convex and $x \leq x + \Delta \leq y_T^p$. This implies that $s^v(x + \Delta) \geq s^v(x)$.

Case I-b $s^v(x) \leq x + U$, $s^v(x + \Delta) \geq x + \Delta + U$

$s^v(x + \Delta) \geq s^v(x)$ directly follows.

Case I-c $s^v(x) \geq x + U$, $s^v(x + \Delta) \geq x + \Delta + U$

Let $s' = \min\{s : J_T(s|x + \Delta) = J_T(x + U + V|x + \Delta) + K_p + K_c\}$. Since $J_T(y|x) = J_T(y|x + \Delta) + c_c\Delta$ for all $y \geq x + \Delta$, we must have $s' = s^v(x)$. By noting that $x + \Delta + U + V \leq y_T^c$,

$$J_T(x + \Delta + U + V|x + \Delta) + K_p + K_c \leq J_T(x + U + V|x + \Delta) + K_p + K_c$$
$$J_T(s^v(x + \Delta)|x + \Delta) \leq J_T(s'|x + \Delta)$$
$$s^v(x + \Delta) \geq s' = s^v(x).$$

Case I-d: $s^v(x) \geq x + U$, $s^v(x + \Delta) \leq x + \Delta + U$

One can show that $s' \leq s^v(x+\Delta)$ by following the steps of Case I-c. We have $s' \geq s^v(x)$ since $J_T^p(y) \geq J_T^c(y|x)$ for all $y \leq x+U$ and since J_T^p and J_T^c are convex. Then the result follows.

Case II $s^v(x) \leq s^u(x)$, $x+U \leq y_T^c$

$s(x) = s^u(x) \leq x+U$ and $s^u(x) \leq x+U$ by definition. J_T^p is convex and minimized at y_T^p. For $x+U < x+U+\Delta \leq x+U+V \leq y_T^c \leq y_T^p$, we have

$$\begin{aligned} J_T^p(x+U)+K_p &\geq J_T^p(x+U+\Delta)+K_p \\ J_T(x+U|x)+K_p &\geq J_T(x+\Delta+U|x+\Delta) \\ J_T(s^u(x)|x) &\geq J_T(s^u(x+\Delta)|x+\Delta) \\ s^u(x) &\leq s^u(x+\Delta) \end{aligned}$$

since $J_T(s^u(x)|x) = J_T(s^u(x)|x+\Delta)$.

Case III $s^u(x) \leq s^c(x)$, $x+U \leq y_T^c$

$s(x) = s^c(x)$. This case can be proved similar to Case I.

Case IV $s^c(x) \leq s^u(x)$, $x+U \leq y_T^c$

$s(x) = s^u(x)$. This case can be proved similar to Case II.

Case V $y_T^c \leq x+U \leq y_T^p$

$s(x) = s^u(x)$. This case can be proved similar to Case II.

Case VI $y_T^p \leq x+U$

s^p is independent of x therefore result directly follows.

References

Alp O, Tan T (2008) Tactical capacity management under capacity flexibility. IIE Trans 40:221–237

Angelus A, Porteus EL (2002) Simultaneous capacity and production management of short-life-cycle, produce-to-stock goods under stochastic demand. Manage Sci 48:399–413

Atamtürk A, Hochbaum DS (2001) Capacity acquisition, subcontracting, and lot sizing. Manage Sci 47:1081–1100

Beckers I (2005) Fewer flexible workers. Web Magazine, Statistics Netherlands. http://www.cbs.nl/en-gb/menu/themas/arbeid-inkomen-sociale-zekerheid/arbeidsmarkt/publicaties/artikelen/2005-1616-wm.htm. Accessed on 9 March 2006

Beckers I, Siermann C (2005) Over one third of jobholders are regularly working overtime. Web Magazine, Statistics Netherlands. http://www.cbs.nl/en-gb/menu/themas/arbeid-inkomen-sociale-zekerheid/arbeidsmarkt/publicaties/artikelen/2005-1749-wm.htm. Accessed on 9 March 2006

Dellaert N, De Kok T (2004) Integrating resource and production decisions in a simple multi-stage assembly system. Int J Produc Econ 90:281–294

Federgruen A, Zipkin P (1986) An inventory model with limited production capacity and uncertain demands *i*. the average cost criterion. Math Oper Res 11:193–207

Gallego G, Scheller-Wolf A (2000) Capacitated inventory problems with fixed order costs: some optimal policy structure. Eur J Oper Res 126:603–613

Holt CC, Modigliani F, Muth JF, Simon HA (1960) Planning production, inventories and work force. Prentice-Hall, Englewood Cliffs

Hu JQ, Vakili P, Huang L (2004) Capacity and production management in a single product manufacturing system. Ann Oper Res 125:191–204

Kapuscinski R, Tayur S (1998) A capacitated production-inventory model with periodic demand. Oper Res 46:899–911

Karlin S (1958) Optimal inventory policies for the arrow-harris-marschak dynamic model. In: Arrow K, Karlin S, Scarf H (eds) Studies in the mathematical theory of inventory and production. Stanford University Press, Stanford, California pp 135–154

Milner JM, Pinker EJ (2001) Contingent labor contracting under demand and supply uncertainty. Manage Sci 47:1046–1062

Pinker EJ, Larson RC (2003) Optimizing the use of contingent labor when demand is uncertain. Eur J Oper Res 144:39–55

Porteus EL (1990) Stochastic inventory theory. In: Heyman DP, Sobel MJ (eds) Handbooks in OR & MS, vol. 2., Elsevier, North-Holland pp 605–652

Scarf H (1959) The optimality of (S, s) policies in the dynamic inventory problem. In: Arrow K, Karlin S, Suppes H (eds) Mathematical methods in the social sciences. Stanford University Press, Stanford, California pp 196–202

Tan B, Gershwin SB (2004) Production and subcontracting strategies for manufacturers with limited capacity and volatile demand. Ann Oper Res 125:205–232

Turkish Statistical Institute (2006) Labour force survey, June 2006. http://www.turkstat.gov.tr Accessed on 20 June 2006

US Bureau of Labor Statistics (2006) Contingent and alternative employment arrangements. http://www.bls.gov/cps/ Accessed on 10 June 2006

van Mieghem JA (2003) Capacity management, investment, and hedging: review and recent developments. Manuf Serv Oper Manag 5:269–302

Yang J, Qi X, Xia Y (2005) A production-inventory system with markovian capacity and outsourcing option. Oper Res 53:328–349

Wild B, Schneeweiss C (1993) Manpower capacity planning - a hierarchical approach. Int J Produc Econ 30–31:95–106