Int. J. Production Economics 155 (2014) 361–379

Contents lists available at ScienceDirect

Int. J. Production Economics

journal homepage: www.elsevier.com/locate/ijpe

The dynamic lot-sizing problem with convex economic production costs and setups

Ramez Kian [a], Ülkü Gürler [a], Emre Berk [b,*]

[a] Department of Industrial Engineering, Bilkent University, Ankara, Turkey
[b] Department of Management, Bilkent University, Ankara, Turkey

ARTICLE INFO

Article history:
Received 19 June 2013
Accepted 4 February 2014
Available online 12 February 2014

Keywords:
Uncapacitated lot-sizing
Production plan
Non-linear production cost
Production functions
Heuristics

ABSTRACT

In this work the uncapacitated dynamic lot-sizing problem is considered. Demands are deterministic and production costs consist of convex costs that arise from economic production functions plus set-up costs. We formulate the problem as a mixed integer, non-linear programming problem and obtain structural results which are used to construct a forward dynamic-programming algorithm that obtains the optimal solution in polynomial time. For positive setup costs, the generic approaches are found to be prohibitively time-consuming; therefore we focus on approximate solution methods. The forward DP algorithm is modified via the conjunctive use of three rules for solution generation. Additionally, we propose six heuristics. Two of these are single-stepSilver–Meal and EOQ heuristics for the classical lot-sizing problem. The third is a variant of the Wagner–Whitin algorithm. The remaining three heuristics are two-step hybrids that improve on the initial solutions of the first three by exploiting the structural properties of optimal production subplans. The proposed algorithms are evaluated by an extensive numerical study. The two-step Wagner–Whitin algorithm turns out to be the best heuristic.

1. Introduction

In this paper, we consider the problem of dynamic lot-sizing in the presence of polynomial-type convex production functions and non-zero setup costs. The dynamic lot-sizing problem is defined as the determination of the production plan that minimizes the total (fixed setup, holding and variable production) costs incurred over the planning horizon for a single, storable item facing deterministic demands. The so-called classical dynamic lot-sizing problem was first analyzed by Wagner and Whitin (1958). They established that, in an optimal plan with positive fixed setup costs and linear production and holding costs, production is done in a period only if its net demand (actual demand less inventories) is positive, and a period's demand is satisfied entirely by production in a single period (that is, integrality of demand is preserved.) For linear production costs, extensions include Zangwill (1966), Blackburn and Kunreuther (1974), Lundin and Morton (1975), Federgruen and Tzur (1991), Wagelmans et al. (1992), Aggarwal and Park (1993), Azaron et al. (2009), Ganas and Papachristos (2005), Okhrin and Richter (2011) and Toy and Berk (2013). The fundamental properties of the optimal plans for linear costs hold for piecewise linear and concave cost structures, as well. For details on such results, we refer the reader to the reviews in Brahimi et al. (2006), Karimi et al. (2003), Jans and Degraeve (2007), Buschkühl et al. (2010) and Jans and Degraeve (2008). There is also a parallel stream of research that focuses on developing lot sizing heuristics based on simple stopping rules. (See Vollmann et al. (1997), Simpson (2001), and Jeunet and Jonard (2000) for a full list and review.) The advantages of such approximate solution methodologies are their ease-of-use, smoother production schedules and providing more intuition to practitioners about the fundamental trade-offs. Hence, the available commercial ERP software (*e.g.*, *SAP*) offers the well-known heuristics for the classical lot sizing problem as options for decision-makers in theirmanufacturing modules. These include the Silver–Meal and economic order quantity (EOQ) based heuristics among others (Silver and Meal, 1973; Harris, 1913; Erlenkotter, 1989).

Most of the existing works on the dynamic lot-sizing problem deal with linear and/or concave production functions rather than convex functions. For convex cost functions and zero setup costs, a parametric algorithm was developed by Veinott (1964) for the problem, which can be solved by an incremental approach satisfying each unit of demand as cheaply as possible. The algorithm has a computational complexity of$O(TD_{1,T})$ where T is the problem horizon length and $D_{1,T}$ stands for the total demand over the problem horizon. Works by Meyer (1977) and Khachian (1979)

* Corresponding author. Tel.: +90 312 2902413.
E-mail addresses: ramezk@bilkent.edu.tr (R. Kian), ulku@bilkent.edu.tr (Ü. Gürler), eberk@bilkent.edu.tr (E. Berk).

http://dx.doi.org/10.1016/j.ijpe.2014.02.006
0925-5273

render this problem solvable in strictly polynomial time. Our work differs from the existing literature in our main assumption about the structure of production costs. Specifically, we consider variable production costs in period t of the polynomial form $\sum_{n=1}^{m} w_t^n q_t^{r_t^n}$ where q_t denotes the quantity produced in the period, w_t^n and r_t^n are positive constants and m is the number of resources. The assumed non-linearity aims to capture the externalities in production activities that are encountered in a number of industrial settings as briefly discussed below:

(i) Productive assets require maintenance and repair activities over their lifetimes and almost all production processes generate undesirable wastes, which must be disposed of and/or whose negative ecological impact must be mitigated. As additional resources are required or legal penalty rates become progressive, the costs associated with such auxiliary activities exhibit a convex behavior. To the best of our knowledge, the only attempt to incorporate such non-linear costs in production planning is performed by Heck and Schmidt (2010) who proposed a heuristic which is a variant of the incremental solution approach in Veinott (1964).

(ii) Non-linear production functions also arise from production activities that use a number of substitutable resources such as materials, labor, machinery, capital, energy, *etc*. One of the most common production functions is the Cobb–Douglas production function, which was introduced at a macroeconomic level for the US manufacturing industries for the period 1899–1922 but has been widely applied to individual production processes at the microeconomic level, as well. For example, Shadbegian and Gray (2005) use the Cobb–Douglas production function to model production processes in the paper, steel and oil industries, Hatirli et al. (2006) to model agricultural production, and Kogan and Tapiero (2009) to model logistics/supply chain operations. The Cobb–Douglas production function assumes that multiple (m) resources are needed for output, Q and they may be substituted to exploit the marginal cost advantages. In general, it has the form $Q = A\prod_{i=1}^{m} x(i)^{\alpha(i)}$ where A is the technology level for the production process, $x(i)$ denotes the amount of resource i used and $\alpha(i) > 0$ is the resource elasticity. Assuming that resource i has a unit cost of $p(i)$, the total cost for output Q is given by wQ^r where $w = (1/r)A^{-r}\prod_{i=1}^{m}(p(i)/\alpha(i))^{\alpha(i)r}$ and $1/r = \sum_{i=1}^{m}\alpha^i$ (Heathfield and Wibe, 1987). The total elasticity parameter $1/r$ may be greater than (smaller than) or equal to 1 depending on whether there is diminishing (increasing) returns to resources, resulting in convex (concave) variable production costs. Despite its widespread occurrence, the impact of the Cobb–Douglas production function on dynamic lot-sizing problems has not been studied.

(iii) Another commonly used economic production function is the Leontieff function introduced by Leontieff (1947). Its main difference from the Cobb–Douglas function is that it assumes that resources are not substitutable but complementary. The applications include Haldi and Whitcomb (1967) for refining of petroleum and primary metals, Ozaki (1976) for large-scale assembly production, Lau and Tamura (1972) for ethylene production, and Nakamura (1990) for iron and steel production. The Leontieff production function has the form $Q = \min_i\{x(i)^{\alpha(i)}\}$ for a given set of resources where $x(i)$ denotes the amount of resource i used and $\alpha(i) > 0$ is the resource elasticity. Assuming that resource i has a unit cost of $p(i)$, the total cost for output Q is given by $\sum_{i=1}^{m} w^i Q^{1/\alpha(i)}$ where $w^i = p(i)$. Typically, it is assumed that $\alpha(i) \le 1$ so that the variable cost of production is convex in output. Similarly, there are no studies on the dynamic lot-sizing problem in the presence of Leontieff production functions.

The general structure for variable production costs assumed above subsumes the above three classes of costs of production externalities. For $m > 1$, each term $w_t^i q_t^{r_t^i}$ corresponds either directly to the cost of using resource i in a complementary fashion in order to produce q_t units in period t through a Leontieff-type production function or to the individual polynomial terms of the cost of efforts to mitigate the ecological impact. For $m = 1$, the only term $w_t^1 q_t^{r_t^1}$ corresponds to the effective cost of using *all* resources to produce q_t units in period t through a Cobb–Douglas type production function. To avoid confusion, we remind the reader that the above discussion of multiple resources is to motivate the form of the variable production cost functions. Once we have them, we focus on the production plan of the single item.

In this paper, we formulate the dynamic lot-sizing problem first as a mixed integer non-linear programming (MINLP) problem and obtain fundamental properties of the optimal solution. In particular, we characterize the optimal solution structure for the case of zero setup costs and establish the property that shows how the optimal solution for a T-period problem can be updated to give the solution for a $(T+1)$-period problem. This property leads us later to develop a forward dynamic programming (DP) formulation which obtains the optimal production plan in $O(T^2 2^T)$ run time in general. For positive setup costs, we also show that the same optimal production plan structure (consisting of G-class subplans) is retained when periods are pre-specified in which production is done. Based on this property, we modify the forward DP algorithm by means of three simple set-construction rules so that $O(T^2)$ computational complexity is achieved. This constitutes our benchmark algorithm for large sized problems. In addition, we propose six new heuristics for the lot sizing problem at hand. Heuristics $H1$ and $H2$ are based on stopping rules and variants of the Silver–Meal and EOQ based heuristics for the classical lot sizing problem. Heuristic $H3$ is a variant of the Wagner–Whitin solution that employs the forward DP algorithm while imposing demand integrality on the production quantities. The first three heuristics are single step heuristics. The remaining three heuristics, which we call the G-heuristics, are two-step hybrids that use the set of production periods of the solutions obtained by the first three heuristics and improve them via G-class production subplans.

An extensive numerical study establishes that a forward DP algorithm wherein production periods within generations are selected via simple rules provides a reasonably fast and efficient solution methodology. Among the proposed heuristics, the Wagner–Whitin heuristic ($H3$) performs best among the single step heuristics and the hybrid G-heuristics exploiting the optimal production plan structure outperform the single step heuristics significantly. The best heuristic among all those proposed turns out to be the hybrid one that improves on the Wagner–Whitin solution, namely, heuristic $H6$. These are followed in performance by the single step heuristic $H2$, which is based on the EOQ model, and the G-heuristic $H5$, which improves on that. The sensitivity analysis on the optimal solutions (obtained by the benchmark DP algorithm) reveals two fundamental tendencies which are in accordance with intuition. Higher production cost non-linearities and lower average unit production costs force production to be spread over a larger number of periods to exploit the marginal cost benefits. Thus, unlike the classical lot-sizing model with the non-speculative cost structure, production functions generate a tendency to produce in earlier periods when setup costs are zero. This results in *production smoothing* – production decisions in more periods with smaller quantities. Positive setup costs, on the other hand, introduce the batching tendency, as expected; for larger setup costs, larger production quantities emerge to compensate for a setup in a period. The interaction between these two tendencies is not always straightforward for particular cost parameter values but the fundamental trade-offs could be observed in all experiment instances. The production smoothing tendency revealed in our study is of interest from a practical perspective, as well; it supports the managerial attitudes toward dedicated facilities and high asset utilization rates in practice.

The remainder of the paper is organized as follows: Section 2 describes the model and provides the MINLP formulation. Section 3 presents the structural results on the optimal solution. In Section 4, we discuss possible solution approaches that can be applied to the problem at hand, formulate both backward and forward DP algorithms based on the fundamental structural properties of optimal solutions and develop our additional heuristics. We present our findings from an extensive numerical study in Section 5. Specifically, we compare the performance of the heuristics and the forward DP algorithm in terms of attained costs and corresponding computational times; and, discuss some sensitivity results. Finally, in Section 6 we briefly summarize our findings and suggest future research directions.

2. Model assumptions and formulation

We consider a single item. The length of the problem horizon, T, is finite and known. The demand amount in period t is denoted by d_t (t=1,…,T). All demands are non-negative and known, but may be different over the planning horizon. No shortages are allowed; that is, the amount demanded in a period has to be produced in or before its period. The amount of production in period t is denoted by q_t and is uncapacitated. Production quantities may be real-valued. Production in any period t incurs a fixed cost (of setup) $K_t(\geq 0)$ and a variable cost component. Variable cost of production is non-linear in q_t and is of the form: $\sum_{j=1}^{m} w_t^j q_t^{r_t^j}$, where w_t^j and r_t^j are non-negative constants. We assume that $r_t^j > (\leq) 1$ *for all* j, t resulting in convex (concave) variable production costs. Any period in which $q_t > 0$ is called a *production* period; otherwise, it is a *no-production* period. The inventory on hand at the end of period t is denoted by I_t; each unit of ending inventory in the period is charged a unit holding cost of h_t. Without loss of generality, the initial inventory level, I_0, is assumed to be zero. The objective is to find a production plan that determines the timing and amount of production (q_t) such that the total cost of production and holding over the horizon is minimized. For the sub-horizon consisting of periods $\{u, u+1, \ldots, v\}$ ($[u, v]$ in short), let $P_{u,v}$ denote the production planning problem, $D_{u,v} = d_u + d_{u+1} + \cdots + d_v$ denote the total demand, $Q_{u,v} = (q_u, \ldots, q_v)$ denote the production plan and $F_{u,v}$ denote the corresponding total cost.

We formulate the problem as a MINLP problem. This allows us to establish certain structural properties of the optimal solution. We can state problem $P_{u,v}$ formally as follows:

$$\min_{q_u,\ldots,q_v} \quad F_{u,v} = \sum_{t=u}^{v} \left[\left(K_t y_t + \sum_{j=1}^{m} w_t^j q_t^{r_t^j} + \left(\sum_{i=t}^{v} h_i \right) q_t \right) \right] - \sum_{t=u}^{v} h_t D_{u,t} \tag{1a}$$

$$\text{s.t.} \quad \sum_{i=u}^{t} q_i \geq D_{u,t}, \quad t \in \{u, \ldots, v\} \tag{1b}$$

$$q_t \geq 0, \quad t \in \{u, \ldots, v\} \tag{1c}$$

$$q_t \leq D_{t,v} y_t, \quad t \in \{u, \ldots, v\} \tag{1d}$$

$$y_t \in \{0, 1\}, \quad t \in \{u, \ldots, v\} \tag{1e}$$

where y_t denotes the binary variable for a setup. The first set of constraints (1b) ensure that all demands will be met and (1c) are nonnegativity constraints. The optimization problem at hand is finding $Q_{1,T}^* = (q_1^*, \ldots, q_T^*)$ and $F_{1,T}^*$ for $P_{1,T}$ over the horizon $[1, T]$, where we use (*) to indicate optimality for all entities. In the analysis that follows, we assume, for convenience, that production quantities are non-negative real numbers.

The nonlinear convex production costs are the key difference between our model and the classical well-known model introduced by Wagner and Whitin (1958) which is a Mixed integer Programming (MIP) model. The fundamental properties of the optimal solution for $r \leq 1$ are that, in an optimal plan, (i) production may occur in period t only if $I_{t-1} = 0$ and (ii) the entire demand in a period is covered by production in a single period (demand integrality is preserved) (Wagner and Whitin, 1958). For $r > 1$, these properties do not hold. This makes the production planning problem in the presence of convex production costs challenging and interesting. To illustrate this point, consider $P_{1,T}$ for the following simple example. For T=2, $K_t = K = 700$, $h_t = h = 1$, m=1, $w_t^1 = w = 0.01$, $r_t^1 = r = 2$ for $1 \leq t \leq T$ and $\mathbf{d} = (100, 300)$. As will be established later, the optimal plan for this problem gives $q_1^* = 175$ and $q_2^* = 225$. Note that neither of the two properties holds; $I_1^* \times q_2^* \neq 0$ and $0 < q_2^* < d_2$. In technical jargon, the feasible solution set is convex. A concave function attains its minimum over a convex set at an extreme point. Thus, whenever the cost functions in a lot sizing model is concave, the optimal solution lies on the extreme points. On the other hand, a convex function may attain its minimum in an interior point of the feasible region (as in the example above). Such an interior point solution is called a non-integral plan since the production quantity in each period is not exactly equal to the demand summed over one or more future periods. Our main contribution is to characterize such non-integral solutions (if any) and the related structural results which are provided in the next section.

3. Structural results

In this section, we present structural results on the optimal production plan for the dynamic lot-sizing problem $P_{u,v}$ introduced above. In particular, we introduce the key concept of a generation and related definitions; establish the decomposition properties for production subplans in terms of inventory levels and generations, and the characteristics of a production plan for a generation; and, based on these, we characterize the optimal production plan structure. For the special case of K=0, we also provide a planning horizon that rests on merging of generations as problem horizon extends. We begin with the definitions and key concepts.

Definition 1. In a given production plan, Q_{ij} for periods $\{i, \ldots, j\}$,

(1) period t is a regeneration point if $I_{t-1} = 0$;
(2) a sequence of periods $\{u, u+1, \ldots, v\}$, for $i \leq u \leq v \leq j$, is a generation, denoted by $\langle u, v \rangle$, if $I_{u-1} = I_v = 0$ and $I_t > 0$ for $t \in \{u, u+1, \ldots, v-1\}$;
(3) the production plan of a generation is called a production sequence.

Note that the definitions above are similar to those in Manne and Veinott (1967) and Florian and Klein (1971) with slight notational differences. Regeneration points (and, thereby, generations) play a central role in finding the optimal production plans in lot-sizing problems. Specifically, they allow us to partition the problem horizons and to independently solve for sub-problems. Florian and Klein (1971) have established this property for any cost structure. We re-state their result below.

Theorem 1 (*Inventory decomposition property*). *Suppose that the constraint*

$$I_k = 0 \quad \text{for some } k \in \{1, \dots, t-1\}, \tag{2}$$

is added to problem $P_{1,t}$, then, an optimal solution to the original problem can be found by independently finding solutions to the problems for the first k periods and for the last $t-k$ periods.

Inventory decomposition has direct implications on the structure of an optimal production plan. Based on this property, it suffices to consider only production sequences to find the optimal solution to problem $P_{u,v}$ as stated below.

Corollary 1 (*Generation decomposition property*). *An optimal production plan $Q^*_{u,v}$ for problem $P_{u,v}$ consists of production sequences which can be independently solved.*

Proof. By assumption, $I_{u-1}=0$. Clearly, in an optimal production plan, $I^*_v = 0$. If $I^*_t \neq 0$ for $t \in \{u, \dots, v-1\}$, then there is a single production sequence. Otherwise, $I_k=0$ for some $k \in \{u, \dots, v-1\}$. In this case, there are $k+1$ generations by definition. From Theorem 1, each generation can be solved as a sub-problem. Hence, the result. □

In the remainder of this section, we provide results on the characteristics of generations and optimal production sequences.

Lemma 1 (*Generation characteristics*). *For a generation $\langle u, v\rangle$,*

(i) $q_u = d_u \geq 0$ *if* $u=v$;
(ii) $\sum_{s=u}^{t} q_s > \sum_{s=u}^{t} d_s$ *for* $t \in \{u, u+1, \dots, v-1\}$ *if* $u<v$;
(iii) $q_u > 0$ *if* $u<v$;
(iv) $d_v > 0$ *if* $u<v$.

Proof. (i) Follows from (1b). (ii) By definition. That is, if $\sum_{s=u}^{t} q_s = \sum_{s=u}^{t} d_s$, then the generation would have ended at $v=t$, which contradicts the definition. (iii) Immediately follows from the previous two results. (iv) We prove the result by contradiction. Suppose that $d_v=0$. Then, the inventory balance equation of period v, $I_v = q_v + I_{v-1} - d_v$, implies $0 = q_v + I_{v-1}$, which is possible only if $q_v = I_{v-1} = 0$ due to the non-negativity of these variables. But this contradicts the definition of a generation, hence the result. □

The above lemma implies that a generation whose total demand is zero consists of a single no-production period, and that a generation with at least two periods can neither end with a zero-demand period nor start with a no-production period. Next, we present our results on the structure of the optimal production plan. In any production plan, there may be production and no-production periods. Given a production plan $Q_{u,v}$, let $S(Q_{u,v})$ denote the set of production periods. A special class of production plans forms the basis of the characterization of the optimal solution. We introduce this class below.

Definition 2. A production plan $Q_{u,v} = (q_u, \dots, q_v)$ is of class G if

$$\sum_{n=1}^{m} r_i^n w_i^n q_i^{r_i^n - 1} = \sum_{n=1}^{m} r_j^n w_j^n q_j^{r_j^n - 1} - \sum_{s=i}^{j-1} h_s \tag{3}$$

for any $i, j \in S(Q_{u,v})$ and $u \leq i < j \leq v$.

Now, we can give the fundamental results on the optimal production plan structure. (The proofs of the results in the remainder of this section are provided in Appendix.)

Theorem 2 (*Optimal production plan structure I*). *In an optimal production plan $Q^*_{1,T}$, for any generation $\langle u, v\rangle$,*

(i) $Q^*_{uv} = (d_u)$ *if* $1 \leq u = v \leq T$,
(ii) $Q^*_{u,v} = (D_{uv}, 0, \dots, 0)$ *if* $1 \leq u < v \leq T$ *and* $r_t^n \leq 1$ *for* $t \in [u, v]$,
(iii) $Q^*_{u,v} = (q^*_u, \dots, q^*_v)$ *is of class G if* $1 \leq u < v \leq T$ *and* $r_t^n > 1$ *for* $t \in [u, v]$,

The above result implies that it suffices to consider only those feasible production plans that are of class G in order to optimize the problem $P_{u,v}$ for any horizon $[u, v]$. We shall exploit this property when we develop our forward dynamic programming solution approach. Theorem 2 characterizes the relationship among the production quantities within a generation. Next, we establish the relationship between the production quantities of two consecutive generations in an optimal production plan.

Theorem 3 (*Optimal production plan structure II*). *If $r_t^n \geq 1$ for all n, t, in an optimal production plan, for generations $\langle u, v\rangle$ and $\langle v+1, v'\rangle$,*

$$\sum_{n=1}^{m} r_{v+1}^n w_{v+1}^n (q^*_{v+1})^{r_{v+1}^n - 1} \leq \sum_{n=1}^{m} r_l^n w_l^n q_l^* r_l^n - 1 + \sum_{i=l}^{v} h_i, \tag{4}$$

where, l is the last production period in $\langle u, v\rangle$.

The above theorem enables us to check whether a proposed bisecting of the sub-horizon $[u, v']$ can be optimal. So far, we have provided structural results of the optimal production plans for the general case that allows for non-zero fixed production (setup) costs. Next, we focus on the special case of $K_t = 0\ \forall t$, which enables us to obtain further results on the optimal production plans.

3.1. A special case: zero setup costs ($K_t = 0$)

Recall that, in the classical lot-sizing problem with the non-speculative cost structure ($c_t + h_t > c_{t+1}\ \forall t$), the optimal production plan consists of lot-for-lot productions in the absence of setup costs. This has two implications: (i) each period is one generation, and (ii) production is done only in periods of non-zero demand. In the presence of production functions, these results no longer hold. In particular, it is optimal to produce in every period within a generation $\langle u, v\rangle$ if $D_{uv} > 0$. This result follows from the property below.

Lemma 2. *If $r_t^n \geq 1$ and $K_t = 0\ \forall t$, in an optimal production plan, for generation $\langle u, v\rangle$, $q^*_j > 0$ if $q^*_t > 0$ for $u \leq t < j \leq v$.*

It follows from the lemma above that all periods within a generation are production periods provided that the total demand is positive and setup costs are negligible.

For convex production and zero setup costs, the optimal solution behaves in a particular way with respect to demand increases and horizon extensions. If the last period's demand is increased (all else being the same), then in the optimal production plan for the modified problem, (1) the number of generations cannot increase, and (2) the optimal solution to the original problem is retained up to a regeneration point obtained in the original problem. That is, only the last generation in the original solution may merge with previous ones to form a longer last generation in the modified problem's solution. If the problem horizon is extended, then, in the optimal solution, either the new period constitutes the (new) last generation in addition to those obtained in the original problem or the effect of extending the problem horizon is similar to a demand increase in the last period of the original problem. We formally state these properties in the following theorem.

Theorem 4 (*Planning horizon theorem*). *Given a problem $P_{1,t}$ with demands $\mathbf{d}_t = (d_1, \dots, d_t)$ and $r_i^n > 1$ and $K_i=0$ for $n=1,\dots,m$ and*

$i=1,\ldots,t$, suppose that the optimal production plan is $Q^*_{1,t} = Q^*_{t_1,t_2-1} \cup Q^*_{t_2,t_3-1} \cup \ldots Q^*_{t_k,t}$ where k denotes the number of generations in the plan and t_j denotes the regeneration points with $t_1=1$.

(i) *For a modified problem $\overline{P}_{1,t}$ with modified demands $\overline{\mathbf{d}}_{1,t} = (d_1,\ldots,d_{t-1},d_t+x)$ where $x>0$, the optimal production plan, $\overline{Q}^*_{1,t}$ is given as $Q^*_{t_1,t_2-1} \cup \ldots \cup Q^*_{t_{i-1},t_i-1} \cup \overline{Q}^*_{t_i,t}$ where $\overline{Q}^*_{t_i,t}$ denotes the (new) production sequence for the (new) last generation and $i \in \{1,\ldots,k\}$.*

(ii) *For problem $P_{1,t+1}$ with demands $\mathbf{d}_{t+1} = (d_1,\ldots,d_t,d_{t+1})$, the optimal production plan is $Q^*_{1,t+1} = Q^*_{t_1,t_2-1} \cup \ldots \cup Q^*_{t_{i-1},t_i-1} \cup \overline{Q}^*_{t_i,t+1}$ where $\overline{Q}^*_{t_i,t+1}$ denotes the (new) production sequence for the (new) last generation $i \in \{1,\ldots,k+1\}$ with $t_{k+1}=t+1$ if $r^n_{t+1}>1$ for $n=1,\ldots,m$ and $K_{t+1}=0$.*

An illustration of this property is given in the example in Table 1 as evolution of the optimal solution is depicted for successively longer problem horizons. As horizon extends from $T=7$ to $T=8$, the former set of generations is retained and the last generation is composed of the new period, whereas the last generation merges with three former ones as horizon further extends to $T=9$. Thus, the last generation in an optimal solution can only extend and its regeneration point can only shift toward the time origin. (See also $T=10,11$.) This theorem is of interest for settings where production plans may be done on a rolling horizon basis. In certain cases, the merging of the last generation with the previous ones may continue up to the first period. Unlike the classical lot-sizing problem, there exists no guaranteed partitioning of the problem horizon even for zero setup costs.

4. Solution algorithms and heuristics

The dynamic lot sizing problem with convex economic production functions can be solved in a number of ways: Direct application of the available generic optimizers on the given mixed integer nonlinear programming (MINLP) formulation; a backward dynamic programming (DP) formulation with inventory levels as states and time periods as stages; a forward DP formulation with exhaustive and heuristic search subroutines; and, heuristics specially developed for the problem at hand. We considered all of these approaches. Below, we discuss the particulars of each approach with its merits and disadvantages.

Problem $P_{u,v}$ is already formulated as an MINLP problem. Therefore, one option is to employ the commercially available solvers which have been developed for generic MINLP problems. In a preliminary unreported numerical study, we tested the suitability of such optimization packages. A direct application of the given MINLP formulation resulted in poor performance of the available solvers; sometimes no solution could be found at all. To overcome this, a possibility is to consider reformulations of the MINLP problem similar to those in Brahimi et al. (2006) making the problem more amenable to the available solvers. A small numerical study indicated that there is indeed room for improvement in the performance of the generic solvers with different reformulations. But, for large scale problems, we still encountered the difficulties of computational time and iteration limits. Another option is to obtain the optimal solution to problem $P_{1,T}$ by a general backward dynamic programming (DP) algorithm. To this end, define $J^T_t(I_t)$ as the minimum total cost under an optimal production plan for periods $t+1$ through T, where I_{t-1} is ending inventory as defined before and follows the recursion $I_t = I_{t-1}+q_t-d_t$ for all t. (We retain all other notation introduced previously.) Then

$$J^T_{t-1}(I_{t-1}) = \min_{q_t \geq \max(0,d_t-I_{t-1})} \left\{ K_t \mathbb{1}_{q_t>0} + h_t I_t + \sum_{n=1}^{m} w^n_t q^{r^n_t}_t + J^T_t(I_t) \right\}$$

$$t \in \{1,\ldots,T\} \tag{5}$$

with $\mathbb{1}_{q_t>0}$ denoting the indicator for a setup and the boundary condition in period T being $J^T_T(I_T)=0$ for all I_T. The optimal solution is found using the above recursion and $J^T_0(0)$ denotes the minimum cost over the problem horizon. The main difficulty with this backward DP algorithm is the curse of dimensionality. For real valued demands, implementing the above formulation requires discretization of ending inventories (and production quantities) with a suitable step-size, say, δ. Then the total memory requirement for the cost-to-go array is of size $[\sum_{t=1}^{T}\sum_{i=t}^{T} d_i]/\delta$. As the problem horizon extends, it becomes prohibitively high preventing its usage for large problems. However, it is possible to use the structural properties of optimal solutions and formulate the problem as a forward DP problem which we discuss next.

Generation decomposition property in Corollary 1 implies that an optimal plan for $P_{u,v}$ can be found by considering generations over $[u,v]$ which can be independently solved. This property forms the basis of the forward dynamic programming recursion which uses only the period information. The logic of the forward DP rests on partitioning any given problem. For any problem horizon t, we construct the feasible production plans by considering the last generation in the plan, $\langle i,t\rangle$, for some $i\in[1,t]$ and the best solution obtained for $[0,i-1]$ where period 0 denotes the time origin for convenience. Formally, we can state the forward DP algorithm as follows. Let f^*_t be the cost under an optimal production plan for $[1,t]$ given that $I_0=0$. Then, for $t=1,\ldots,T$, we have

$$f^*_t = \min_{1\leq i\leq t} \{f^*_{i-1}+g_{i,t}\}, \tag{6}$$

where $g_{i,t}$ is the cost associated with generation $\langle i,t\rangle$, $f^*_0 \equiv 0$ and f^*_T is the optimal cost for problem $P_{1,T}$. To find the optimal production sequence for generation $\langle i,t\rangle$, we search over the feasible production plans of class G as implied by Theorem 2. Specifically, we start with some production period set, S for the given generation $\langle i,t\rangle$ and solve for the positive production quantities that satisfy the condition for class G plans. (If the obtained production plan is not feasible, it is discarded as having infinite cost.) If necessary, we update the set S and find new production sequences until no further cost improvements are achieved. Recall that, if $K_t=0 \quad \forall t$, production is done in all periods within a generation except for one-period generations with zero demands. For this case, it suffices to choose the initial S as containing all of the periods $\{i,i+1,\ldots,t\}$ and no updating is necessary. Furthermore, as the algorithm progresses (as t is increased to $t+1$), from Theorem 4, instead of minimizing over $i\in[1,t]$, it is sufficient to consider only the regeneration points $\{t_1,\ldots,t_k\}\cup\{t\}$, where t_j's denote the regeneration points obtained for problem horizon t. The above algorithm is guaranteed to give the optimal solution for problem $P_{1,T}$ (*i.e.*, $f^*_t = F^*_{1,T}$). We provide the pseudo-code for the forward DP algorithm in Appendix. For zero setup fixed costs, it has a computational complexity of $O(T^2)$; in practice, this translates to the algorithm being able to solve large scale problems with 300 periods within a millisecond on a personal computer. For positive setup costs, however, production may not be done in all periods in a generation $\langle i,t\rangle$, and all 2^{t-i+1} possible sets must be considered for S as candidates for new production sequences. The forward DP algorithm that considers all these sets provides the optimal solution and has $O(T^2 2^T)$ run time complexity. But such an exhaustive search is prohibitively time-consuming rendering the exact solution by the given forward DP formulation impractical for $K>0$ and long problem horizons.

In the absence of reasonably fast exact solution methodologies, one may resort to approximate solutions. We develop an approximate version of the above forward DP algorithm, which will be used a benchmark. Additionally, we propose six heuristics for problem $P_{1,T}$, which we refer to as heuristics $H1$–$H6$. Heuristics $H1$ and $H2$ are based on stopping rules and variants of the Silver–Meal and EOQ based heuristics for the classical lot sizing problem. Heuristic $H3$ is a variant of the Wagner–Whitin solution that employs the forward DP algorithm while imposing demand integrality on the production quantities. The first three heuristics are single step heuristics. The remaining three heuristics, which we call the G-heuristics, are two-step hybrids that use the set of production periods of the solutions obtained by the first three heuristics and improve them via G-class production subplans. For all heuristics, we adopt the following notation. The solution for problem $P_{1,T}$ obtained under heuristic Hj consists of the set of production quantities denoted by $Q_T^{(j)}=\{q_1^{(j)},q_2^{(j)},\ldots,q_T^{(j)}\}$ and the index set of production periods for the problem horizon denoted by $\Omega_T^{(j)}$ in which period t is a production period if $q_t^{(j)}>0$ for $t=1,\ldots,T$, and results in the cost, $f_T^{(j)}=\sum_{t\in\Omega_T^{(j)}}K_t+\sum_{t=1}^{T}[h_tI_t+\sum_{n=1}^{m}w_t^n(q_t^{(j)})^{r_t^n}]$ with I_t as defined before. Below, we explain the construction and particulars of each heuristic in detail.

Heuristic $H1$ is similar in construction to the heuristic in Silver and Meal (1973) developed for the classical dynamic lot sizing problem. Under this heuristic, the beginning period of each generation is its sole production period. The generations themselves are obtained in a forward manner along the problem horizon by means of a stopping rule. A generation starting in period u terminates in period $u+\hat{l}(u)$ where

$$\hat{l}(u)=\max\left\{l:\frac{g_{u,u+l}^{(1)}}{l}\geq\frac{g_{u,u+l+1}^{(1)}}{l+1},u\leq l\leq T\right\}$$

with $g_{u,v}^{(1)}:=K_u+[\sum_{s=u}^{v-1}h_sD_{s+1,v}]+\left[\sum_{n=1}^{m}w_u^nD_{u,v}^{r_u^n}\right]$ being the cost associated with the periods $[u,v]$. The generation terminates at $\hat{l}(u)$ because the cost per period starts to increase after that. The solution algorithm starts with the initial period of the problem horizon.

Once the stopping rule is satisfied and $\hat{l}(1)$ is found, the production plan over $[1,\hat{l}(1)+1]$ is retained and the procedure is repeated for the remaining periods starting with period $\hat{l}(1)+1$ until the entire horizon is covered. The pseudo-code is provided in Appendix and has $O(T)$computational complexity. Under this heuristic, the quantity produced in period t is given as $q_t^{(1)}=D_{t,\hat{l}(t)}$ if $t\in\Omega_T^{(1)}$ and zero, otherwise. Then, we have $f_T^{(1)}=\sum_{i\in\Omega_T^{(1)}}g_{i,i+\hat{l}(i)}^{(1)}$. By design, with this heuristic, demand integrality is preserved in production quantities and each production period constitutes a generation start in the solution. The stopping rule computation differs from the classical Silver–Meal heuristic in order to incorporate the nonlinear production costs in our setting.

Heuristic $H2$ is based on a variant of the economic order quantity (EOQ) model which was developed by Harris (1913) for linear acquisition costs. To develop the heuristic, consider the following stylized continuous time counterpart for our production setting. Demand for the item is deterministic with a constant rate, $\overline{d}$ over an infinite problem horizon with stationary cost parameters. Production is done in lots of constant size $\tilde{Q}$ (because of infinite horizon) incurring costs nonlinear in the production quantity. The objective is to minimize the total cost rate $TC(\tilde{Q})=K\overline{d}/\tilde{Q}+h\tilde{Q}/2+[\sum_{n=1}^{m}w^n\tilde{Q}^{r^n}]\overline{d}/\tilde{Q}$ where K stands for the fixed setup cost and h for the unit holding cost rate. Let the minimum total cost rate be denoted by TC^* and the corresponding optimal lot size by $\tilde{Q}^*$. We have the following result.

Lemma 3. *The total cost rate $TC(\tilde{Q})$ is quasi-convex for $r_n\geq 1$ and has a unique minimizer $\tilde{Q}^*$ which solves*

$$K-h(\tilde{Q}^*)^2/2\overline{d}+\sum_{n=1}^{m}(1-r^n)w^n(\tilde{Q}^*)^{r^n-1}=0.$$

The proof rests on a standard optimization methodology and is provided in Appendix. Note that the above result reduces to the classical EOQ result for $r^n=1$ for $n=m=1$. For the general case, it does not yield a closed-form solution for $\tilde{Q}^*$ but the uniqueness of the solution allows for an efficient linear search for it. (For integer demands, it is possible to modify the expressions similar to Garcia-Laguna et al. (2010); but it has not been pursued herein.) Under heuristic $H2$ the production quantity in period t is found as $q_t^{(2)}=\min([D_{t,T}-I_{t-1}]^+,\max([d_t-I_{t-1}]^+,\tilde{Q}^*))$ for $1\leq t\leq T$ starting with $I_0=0$. The solution algorithm starts with the initial period of the problem horizon, and production quantities are obtained as one proceeds over the entire problem horizon. The pseudo-code for the algorithm is provided in Appendix and has $O(T)$ computational complexity. We have $f_T^{(2)}=\sum_{t\in\Omega_T^{(2)}}K_t+\sum_{t=1}^{T}[h_tI_t+\sum_{n=1}^{m}w_t^n(q_t^{(2)})^{r_t^n}]$. The condition on the net remaining total demands ($[D_{t,T}-I_{t-1}]^+$) ensures ending inventory to be zero. Unlike the above heuristic, demand integrality is not preserved under this heuristic.

Heuristics $H3$–$H6$ and the benchmark approximate DP employ the forward DP algorithm introduced above and obtain solutions by means of simple rules to construct the set S in a generation resulting in a possibly suboptimal solution. The approximate algorithm differs from the exact one only in its construction of S. The approximate forward DP that one would get has the advantage of providing solutions within reasonable times and the goodness of the solutions can be improved by developing efficient set-construction heuristics. Below, we explain the details of these heuristics.

Heuristic $H3$ is obtained by employing the forward recursive procedure in Eq. (6) while imposing the condition that demand integrality is preserved. Hence, for any generation $\langle u,v\rangle$ in the solution, we set the quantity produced in period i, $q_i^{(3)}=D_{u,v}$ for $i=u$ and $q_i^{(3)}=0$ for $u<i\leq v$ and search over all possible generations over the problem horizon. Let $g_{i,t}^{(3)}$ be the total cost of the subproblem $[i,t]$ which constitutes a single generation $\langle i,t\rangle$, $f_0^3\equiv 0$. Then, for $t=1,\ldots,T$, $f_t^{(3)}=\min_{1\leq i\leq t}\{f_{i-1}^{(3)}+g_{i,t}^{(3)}\}$ where $g_{i,t}^{(3)}=K_i+[\sum_{s=i}^{t-1}h_sD_{s+1,t}]+[\sum_{n=1}^{m}w_t^nD_{i,t}^{r_t^n}]$. Due to the imposition of demand integrality, this heuristic may be viewed as a version of the classical Wagner–Whitin solution methodology. It has the same computational complexity as the classical algorithm in Wagner and Whitin (1958) and it reduces to the solution in the classical setting, for $r_t^n=1$ for all n. In its implementation, the forward DP algorithm is employed wherein the production period set S for a generation $\langle u,v\rangle$ is constructed as consisting of only period u. Aside from being a viable approximate solution technique, heuristic $H3$ is important in that its performance illustrates the significance of demand splitting in the case of nonlinear production costs and the importance of class G production subplans.

Next, we introduce heuristics $H4$–$H6$ which exploit the G-class property of the production subplans. They work as follows. First, we obtain an initial (approximate) solution to the problem $P_{1,T}$ by one of the above three heuristics. Of this initial solution obtained via heuristic *Hj*, we take only the set of production periods $\Omega_T^{(j)}$, and use it as the given global set of production periods. That is, as we implement the forward DP algorithm, we construct the set S for the generation $\langle u, v\rangle$ using the subset of $\Omega_T^{(j)}$ corresponding to the problem subhorizon $[u, v]$. In practice, this amounts to simply reading off the indexes of the production periods, if any, in the set-construction subroutine. The rest of the algorithm is applied as before. Hence, *H4*–*H*6 are two-step improvement extensions of heuristics *H*1–*H*3. That is, *H4* takes $\Omega_T^{(1)}$ obtained by heuristic *H*1 and improves on it via class G subplans in accordance with Theorem 3, heuristic *H*5 takes $\Omega_T^{(2)}$ obtained by heuristic *H*2 and improves on it, and so forth. By construct, the use of the initial solutions implies that we construct the set S *a priori* and, hence, need only to consider a smaller fraction of class G subplans. This greatly reduces the computational effort. The approximate forward DP algorithm has $O(T^2)$ computational time complexity, given that $\Omega_T^{(i)}$ is provided as pre-processed data. We denote the usage of these S-construction heuristics in the pseudo-codes as instructions denoted by $\Omega_T^{(i)} \rightarrow S$. The performance of this group of heuristics depends, to some extent, on the performance of the initial approximate solution which gives $\Omega_T^{(i)}$. But, the significant improvements over the initial solutions indicate that developing the G-class subplans for generations is the main factor for obtaining good solutions.

Lastly, we consider another method of constructing the set S for a generation in the solution. In this method, we create the set S for each generation under consideration according to three set-construction rules used conjunctively as the algorithm proceeds over the problem horizon. (i) The first rule is a greedy exclusion rule. Initially, S contains all periods within the generation. One by one, each period (other than the first) is excluded in the updated S. The best is retained and the greedy improvement is repeated with the remaining periods until no further improvement. To avoid possible local optima, we also implemented a scatter search by means of updating S randomly as follows. (ii) The second rule is a randomized exclusion rule. This is the randomized version of the greedy exclusion rule. Initially, S is full. A period is randomly selected to be excluded from the updated S. This is repeated for n times. The best is retained and the greedy improvement is repeated with the remaining periods until no further improvement. (iii) Finally, a randomized inclusion rule. Initially, S contains only the first period of the generation. This corresponds to the solution in the classical dynamic lot-sizing problem. A period is randomly selected to be included in the updated S. This is repeated for n times. The best is retained and the greedy improvement is repeated with the remaining periods until no further improvement. The conjunctive use of these rules implies that, for a generation considered in the solution, set S that gives the minimum cost among all those constructed by the three rules is taken as the production period set for that generation. With the implementation of the S-construction subroutine using the above rules, the forward DP algorithm has a computational complexity of $O(T^4)$ in the presence of positive setup costs. Clearly, this algorithm cannot guarantee optimality for positive setups costs; however, our preliminary numerical tests (with problem horizon length of 100 periods) indicate that the suboptimality decreases significantly for long problem horizons with average deviations from the optimal (obtained by backward DP algorithm) of approximately 0.1%. Therefore, we adopted this solution algorithm as our *benchmark* solution methodology.

Before we proceed with our detailed numerical study, we illustrate the implementation of the proposed solution algorithms through a small example. We have $h_t = h = 0.1$, $m=1$, $w_t^1 = w = 0.01$, $r_t^1 = r = 2$, $K_t = K$ for all $t \in \{1, \ldots, T\}$, $T=12$, $K = \{0, 100\}$ and the demand vector, $\mathbf{d} = (50, 100, 0, 70, 80, 40, 45, 30, 80, 35, 250, 75)$. We assume that production quantities can be real numbers. In Table 1, we present the optimal production plans $Q_{1,i}^*$, the corresponding total cost $f_{1,i}^*$, the regeneration points in the optimal solution and the candidate solutions developed for problem $P_{1,i}$ as the DP progresses over the horizon length $i = 1, \ldots, T$ for $K=0$. Note that for zero setup costs, the forward DP is guaranteed to find the optimal. But, for $K > 0$, the forward algorithm does not guarantee the optimal solution. In Table 2, for different sub-problem horizon lengths i, we present the optimal production plan $Q_{1,i}^*$ and the corresponding total cost $F_{1,i}^*$ as obtained by the backward DP algorithm and the counterparts $\tilde{Q}_{1,i}$ and $\tilde{F}_{1,i}$ obtained by the forward DP employing with a discretization increment of $\delta=0.01$ units. As the forward algorithm partitions the problem into the last generation $\langle k+1, i\rangle$ and the sub-horizon $[1, k]$, it

Table 1
Forward dynamic programming algorithm solution ($m=1, w_t^1 = w = 0.01, h_t = h = 0.1, r_t^1 = r = 2, K_t = 0$ for all $t \in \{1, \ldots, T\}$).

i	$Q_{1,i}^*$	f_i^*	Regeneration points	Minimization search
1	{50}	25	{1}	$\{\mathbf{g_{1,1}}\}$
2	{72.5,77.5}	114.88	{1}	$\{\mathbf{g_{1,2}}, f_1^* + g_{2,2}\}$
3	{72.5,77.5}, {0}	114.88	{1,3}	$\{g_{1,3}, \mathbf{f_2^*} + \mathbf{g_{3,3}}\}$
4	{72.5,77.5}, {32.5,37.5}	142.75	{1,3}	$\{g_{1,4}, \mathbf{f_2^*} + \mathbf{g_{3,4}}, f_3^* + g_{4,4}\}$
5	{72.5,77.5}, {45,50,55}	197.38	{1,3}	$\{g_{1,5}, \mathbf{f_2^*} + \mathbf{g_{3,5}}, f_4^* + g_{5,5}\}$
6	{72.5,77.5}, {45,50,55}, {40}	213.38	{1,3,6}	$\{g_{1,6}, f_2^* + g_{3,6}, \mathbf{f_5^*} + \mathbf{g_{6,6}}\}$
7	{72.5,77.5}, {45,50,55}, {40}, {45}	233.63	{1,3,6,7}	$\{g_{1,7}, f_2^* + g_{3,7}, f_5^* + g_{6,7}, \mathbf{f_6^*} + \mathbf{g_{7,7}}\}$
8	{72.5,77.5}, {45,50,55}, {40}, {45}, {30}	242.63	{1,3,6,7,8}	$\{g_{1,8}, f_2^* + g_{3,8}, f_5^* + g_{6,8}, f_6^* + g_{7,8}, \mathbf{f_7^*} + \mathbf{g_{8,8}}\}$
9	{72.5,77.5}, {45,50,55}, {41.25, 46.25,51.25,56.25}	296.44	{1,3,6}	$\{g_{1,9}, f_2^* + g_{3,9}, \mathbf{f_5^*} + \mathbf{g_{6,9}}, f_7^* + g_{8,9}, f_6^* + g_{7,9}, f_8^* + g_{9,9}\}$
10	{72.5,77.5}, {45,50,55}, {41.25, 46.25,51.25,56.25}, {35}	308.69	{1,3,6,10}	$\{g_{1,10}, f_2^* + g_{3,10}, f_5^* + g_{6,10}, \mathbf{f_9^*} + \mathbf{g_{10,10}}\}$
11	{72.5,77.5}, {50,55,60, 65,70,75,80,85,90}	629.38	{1,3}	$\{g_{1,11}, \mathbf{f_2^*} + \mathbf{g_{3,11}}, f_5^* + g_{6,11}, f_{10}^* + g_{11,11}\}$
12	{72.5,77.5}, {50,55,60, 65,70,75,80,85,90}, {75}	685.63	{1,3,11,12}	$\{g_{1,12}, f_2^* + g_{3,12}, \mathbf{f_{11}^*} + \mathbf{g_{12,12}}\}$

Table 2
Comparison of solutions for $P_{1,i}$ obtained by backward and forward dynamic programming algorithms, $Q^*_{1,i}, f^*_{1,i}$ and $\tilde{Q}^*_{1,i}, \tilde{f}^*_{1,i}$ $(m=1, w^1_t = w = 0.01, h_t = h = 0.1, r^1_t = r = 2, K_t = 100$ for all $t \in \{1,\dots,T\}, T=12)$.

i	$Q^*_{1,i}$	$f^*_{1,i}$	$\tilde{f}_{1,i}$	$\tilde{Q}_{1,i}$
1	{50}	125	125	{50}
2	{72.5,77.5}	314.88	314.88	{72.5,78.5}
3	{72.5,77.5}, {0}	314.88	314.88	{72.5,78.5}, {0}
4	{107,113,0,0}	461.88	463.88	{72.5,78.5}, {0}, {70}
5	{93.33,98.33,0,108,34,0}	621.83	627.75	{72.5,78.5}, {0}, {72.5,77.5}
6	{75,80,0,90,95,0}	701.5	701.75	{72.5,78.5}, {0}, {92.5,97.5,0}
7	{93.33,98.33,0,108,34,0}, {85,0}	798.59	804.46	{72.5,78.5}, {0}, {73.33,78.33,83.33,0}
8	{75,80,0,90,95,0}, {75,0}	860.75	861	{72.5,78.5}, {0}, {92.5,97.5,0}, {75,0}
9	{75,80,0,90,95,0}, {75,0}, {80}	1024.75	1025	{72.5,78.5}, {0}, {92.5,97.5,0}, {75,0}, {80}
10	{75,80,0,90,95,0}, {90,0,100,0}	1092	1092.25	{72.5,78.5}, {0}, {92.5,97.5,0}, {90,0,100,0}
11	{75,80,0,90,95,0}, {98.75,0,108.75,113.75,118.75}	1613.82	1614.06	{72.5,78.5}, {0}, {92.5,97.5,0}, {98.75,0,108.75,113.75,118.75}
12	{75,80,0,90,95,0}, {98.75,0,108.75,113.75,118.75}, {75}	1770.06	1770.31	{72.5,78.5}, {0}, {92.5,97.5,0}, {98.75,0, 108.75,113.75,118.75}, {75}

Table 3
Illustrative example showing solutions of heuristics $H1$–$H6$ $(m=1, w^1_t = w = 0.01, h_t = h = 0.1, r^1_t = r = 2, K_t = 100$ for all $t \in \{1,\dots,T\})$.

T	$Q^{(1)}_T$ Heuristic $H1$	$f^{(1)}_T$	$f^{(4)}_T$	$Q^{(4)}_T$ Heuristic $H4$
1	{50}	125	125	{50}
2	{50}, {100}	325	314.88	{72.5,77.5}
3	{50}, {100}, {0}	325	314.88	{72.5,77.5}, {0}
4	{50}, {100}, {0}, {70}	474	463.88	{72.5,77.5}, {0}, {70}
5	{50}, {100}, {0}, {70}, {80}	638	627.75	{72.5,77.5}, {0}, {72.5,77.5}
6	{50}, {100}, {0}, {70}, {120,0}	722	701.5	{75,80,0,90,95,0}
7	{50}, {100}, {0}, {70}, {120,0}, {45}	842.25	821.75	{75,80,0,90,95,0}, {45}
8	{50}, {100}, {0}, {70}, {120,0}, {75,0}	881.25	860.75	{75,80,0,90,95,0}, {75,0}
9	{50}, {100}, {0}, {70}, {120,0}, {75,0}, {80}	1045.25	1024.75	{75,80,0,90,95,0}, {75,0}, {80}
10	{50}, {100}, {0}, {70}, {120,0}, {75,0}, {115,0}	1117	1092	{75,80,0,90,95,0}, {90,,0,100,0}
11	{50}, {100}, {0}, {70}, {120,0}, {75,0}, {115,0},{250}	1842	1669.14	{88.57,93.57,0,103.57,108.57,0,118.57,0,128.57,0,138.57}
12	{50}, {100}, {0}, {70}, {120,0}, {75,0}, {115,0}, {325,0}	2280.75	1892.53	{99.29,104.29,0,114.29,119.29,0,129,29,0,139.29,0,149.29,0}
T	$Q^{(2)}_T$ Heuristic $H2$	$f^{(2)}_T$	$f^{(5)}_T$	$Q^{(5)}_T$ Heuristic $H5$
1	{50}	125	125	{50}
2	{96.82,53.18}	326.71	314.88	{72.5,77.5}
3	{95.35,54.65}, {0}	325.31	314.88	{72.5,77.5}, {0}
4	{95.74,95.74,0,28.51}	504.34	463.88	{72.5,77.5}, {0}, {70}
5	{96.08,96.08,0,96.08,11.77}	698.17	627.75	{72.5,77.5}, {0}, {72.5,77.5}
6	{95.86,95.86,0,95.86,52.42,0}	726.84	701.50	{75,80,0,90,95,0}
7	{95.74,95.74,0,95.74,95.74,0,2.03}	898.90	821.75	{75,80,0,90,95,0}, {45}
8	{95.50,95.50,0,95.50,95.50,0,32.99,0}	1010.52	960.75	{75,80,0,90,95,0}, {75,0}
9	{95.74,95.74,0 95.74,95.74,0,95.74,0,16.29}	1108.92	1024.75	{75,80,0,90,95,0}, {75,0}, {80}
10	{95.59,95.59,0 95.59,95.59,0,95.59,0,52.04,0}	1135.02	1092.00	{75,80,0,90,95,0}, {90,0,100,0}
11	{96.65,96.65,0 96.65,96.65,0,0,96.65,96.65,0,200.10}	1714.97	1657.07	{87.86,92.86,0,102.86,107.86,0,0,122.86,127.86,0,137.86}
12	{96.67,96.67,0 96.67,96.67,0,0,96.67,96.67,0,200.01}, {75}	1871.09	1813.32	{87.86,92.86,0,102.86,107.86,0,0,122.86,127.86,0,137.86}, {75}
T	$Q^{(3)}_T$ Heuristic $H3$	$f^{(3)}_T$	$f^{(6)}_T$	$Q^{(6)}_T$ Heuristic $H6$
1	{50}	125	125	{50}
2	{50}, {100}	325	314.88	{72.5,77.5}
3	{50}, {100}, {0}	325	314.88	{72.5,77.5}, {0}
4	{50}, {100}, {0}, {70}	474	463.88	{72.5,77.5}, {0}, {70}
5	{50}, {100}, {0}, {70}, {80}	638	627.75	{72.5,77.5}, {0}, {70}, {72.5,77.5}
6	{50}, {100}, {0}, {70}, {120,0}	722	701.50	{75,80,0,90,95,0}
7	{50}, {100}, {0}, {80}, {85,0}	814.75	804.46	{72.5,77.5}, {0}, {73.33,78.33,83.33,0}
8	{50}, {100}, {0}, {80}, {115,0,0}	880.75	863.46	{72.5,77.5}, {0}, {83.33,88.33,93.33,0,0}
9	{50}, {100}, {0}, {80}, {85,0}, {110,0}	1043.75	1031.38	{72.5,77.5}, {0}, {77.5,82.5,87.5,0,97.5,0}
10	{50}, {100}, {0}, {80}, {115,0,0}, {115,0}	1116.5	1098.88	{72.5,77.5}, {0}, {85,90,95,0,0,110,0}
11	{50}, {100}, {0}, {80}, {115,0,0}, {115,0}, {250}	1841.5	1680.79	{89.29,94.29,0,104.29,109.29,114.29,0,0,129.29,0,139.29}
12	{50}, {100}, {0}, {80}, {115,0,0}, {115,0}, {250}, {75}	1997.75	1837.04	{89.29,94.29,0,104.29,109.29,114.29,0,0,129.29,0,139.29}, {75}

results in some (globally suboptimal) local optima. Although the resulting production plans may differ significantly, the resulting maximum cost deviation from the optimal is about 0.95% for $i=5$ and less than 0.014% for $i=12$. As the problem horizon increases, the performance of the forward algorithm improves, as expected. For the illustrative example, only one of the set-construction rules (the greedy inclusion updating routine) has been used to find the best production sequence for the last generation. Based on similar preliminary studies, the other two rules (randomized search routines discussed above) have been developed and implemented which result in significant improvements within generations. Hence, they have been embedded to be used conjunctively in the benchmark solution algorithm for the numerical study. For the case of $K=100$, we provide the solutions obtained with the proposed eight additional heuristics in Table 3.

5. Numerical study

In this section, we present and discuss our findings in a numerical study.

For our numerical study, we considered a problem horizon of $T=300$ periods. Period demands are generated randomly from normal distribution with mean $\mu \in \{50, 100, 200\}$ and standard deviation $\sigma(=40)$; negative demand values have been replaced by zero demands. We denote the three demand patterns by $d1$, $d2$ and $d3$. All other system parameters are stationary. We set unit holding cost rate, $h_t = h = 1$ and setup cost is selected as a function of the mean demand rate, $K_t = K = [J^2/2]\mu$ where J may be viewed as a proxy for the average length of a production lot *if* production costs were linear as in the classical lot-sizing problem. We have $J \in \{0, 2, 3, 4, 5\}$ with $J=0$ corresponding to zero setup cost. The production cost structure was chosen with $m=1$ and $r_t^1 = r \geq 1$. This corresponds to the Cobb–Douglas type economic production function with convex costs. We selected $r \in \{1, 1.1, 1.5, 2.0, 2.2\}$; note that $r=1$ corresponds to the classical lot sizing setting used as a benchmark. To select the cost coefficient $w_t^1 = w$, we considered the variable cost of production per unit when a production quantity equals the average demand per period, $\overline{c}$ where $\overline{c} = [w\mu^r]/\mu = w\mu^{r-1}$. Then, letting $a = h/\overline{c}$, we have $w = h\mu/(a\mu^r)$ with $a \in \{0.02, 0.05, 0.1, 0.2\}$. Note that the resulting variable cost for a production quantity of q units is given by $[h\mu/a](q/u)^r$, and that, as a increases $\overline{c}$ decreases since we hold h equal to unity. Overall, our experimental set contains $120(=5 \times 4 \times 6)$ *parameter instances* for each of the three levels of demand mean. For visual displays, we used a shorthand notation to denote an experiment instance, $d_iK_ja_nr_s$, where, for example, $d3K2a1r4$ corresponds to the parameter values, $\mu=200$, $J=2$, $a=0.02$ and $r=2.0$. For each particular experiment instance we generated 10 demand stream *replications*. The average fraction of *zero* demand values in the generated replication streams for each demand pattern is approximately 11%, 0.67% and 0%, respectively with resulting means of 51.9, 100.4 and 200.7.

Next, we discuss our findings from our numerical study. We first report our findings on the relative performance of the proposed heuristics in comparison with the benchmark forward DP solution. Then, we provide our sensitivity analysis on the basis of the solutions obtained by the benchmark DP algorithm.

5.1. Comparison of heuristics

We conducted our numerical study to investigate the following: (i) percentage deviation from the benchmark minimum cost for each heuristic; (ii) dominance of heuristics among themselves; (iii) impact of the cost parameter values and demand patterns on performances of the heuristics. All heuristic comparisons have been performed on a static basis (when demands for the entire problem horizon are known at the beginning of the problem horizon.) For our numerical study, we considered the same experimental set described above. The rationale for this set has been the earlier performance studies for the classical problem setting; in particular, Simpson (2001).

We use as the benchmark the best solution to problem (P) by means of the forward DP solution algorithm discussed above. The total cost over the problem horizon under a particular heuristic is denoted by $f_T^{(j)}$ and the best solution with the forward DP is denoted by $\tilde{f}_T$. For computing the total cost under a particular heuristic for a problem instance, we used the corresponding algorithm provided in Appendix. For each experiment instance, we measure the performance of heuristic Hj in terms of percentage deviations from $\tilde{f}_T$ as follows:

$$\Delta_j\% = \frac{f_T^{(j)} - \tilde{f}_T}{\tilde{f}_T} \times 100$$

Table 4
Percentage deviation statistics for heuristics $H1$–$H3$.

Demand	Algorithm	Min	Max	Median	Average	$N(0)$	$N(-)$
$d1$	$H1$	1.23	89.23	17.97	24.43	0	0
	$H2$	0.08	65.03	6.24	12.47	0	0
	$H3$	0	65.03	5.01	12.76	124	0
$d2$	$H1$	0.79	28.2	10.72	11	0	0
	$H2$	0	19.9	3.24	4.71	8	0
	$H3$	0	19.9	1.58	4.16	161	0
$d3$	$H1$	0.5	7.59	3.18	3.33	0	0
	$H2$	0	7.4	1.69	1.99	36	0
	$H3$	0	5.37	0.44	1.04	235	0

We discuss the performance of heuristics $H1$–$H3$ and heuristics $H4$–$H6$ separately. For each heuristic, we report (i) the minimum, maximum, median and average percentage deviations, and (ii) the number of instances for which zero or negative deviations have been obtained for three different demand variance levels across all 1000 experiment instances. A negative deviation implies that a *better* solution has been found by the heuristic than the forward DP algorithm.

We begin our discussion of the heuristic performances with their overall behavior. In Table 4, we report the performance statistics for heuristics $H1$–$H3$ for $1000(=10 \times 100)$ experiment instances for each of the three different demand patterns. We see that as the demand variance decreases (from $d1$ to $d3$), percentage deviations also decrease for all heuristics. All heuristics have left-skewed performance distributions for all demand variance levels. Heuristic $H3$ (that is, solving the problem in a forward DP algorithm while imposing demand integrality) turns out to be the best performer except for high variance levels on average. It is followed very closely by Heuristic $H2$. The high performance of heuristic $H3$ is due to the fact that the problem is solved optimally albeit under the restriction of demand integrality. Note that 564 $(=124+(161+8)+(235+36))$ out of $3000(=3 \times 1000)$ cases have resulted in zero deviations from the best solution with the forward DP ($N(0)$ column in Table 4), implying that demand integrality was preserved in the best forward DP solution for such instances. For the remaining instances, the deviations obtained under heuristic $H3$ may be viewed as the impact of not smoothing the production across successive periods within a generation. The second-best performance of heuristic $H2$ points to the fundamental trade-offs captured by Harris's formula; and, being a single step heuristic, its performance is excellent. The number of instances for which this heuristic resulted in the same solution as the forward DP also increases as demand variance decreases, as expected. The Silver–Meal heuristic (the original on which our versions are based) typically performs well in the classical setting. It is surprising that heuristic $H1$ did not do as well following $H2$ and $H3$ with a relatively large gap. As demand variance decreases, allowing for production smoothing seems to be counterproductive. This may explain the performance of heuristics $H2$ and $H3$ for d3; after all, both preserve demand integrality by construct. None of the heuristics in this group resulted in a solution better than the benchmark forward DP algorithm. (See $N(-)$ in Table 4.) Next, we look at sensitivity of heuristic performance with respect to the fixed setup cost K and the production cost nonlinearity measured through r. In Table 6, we present the minimum, maximum, median and average percentage deviations

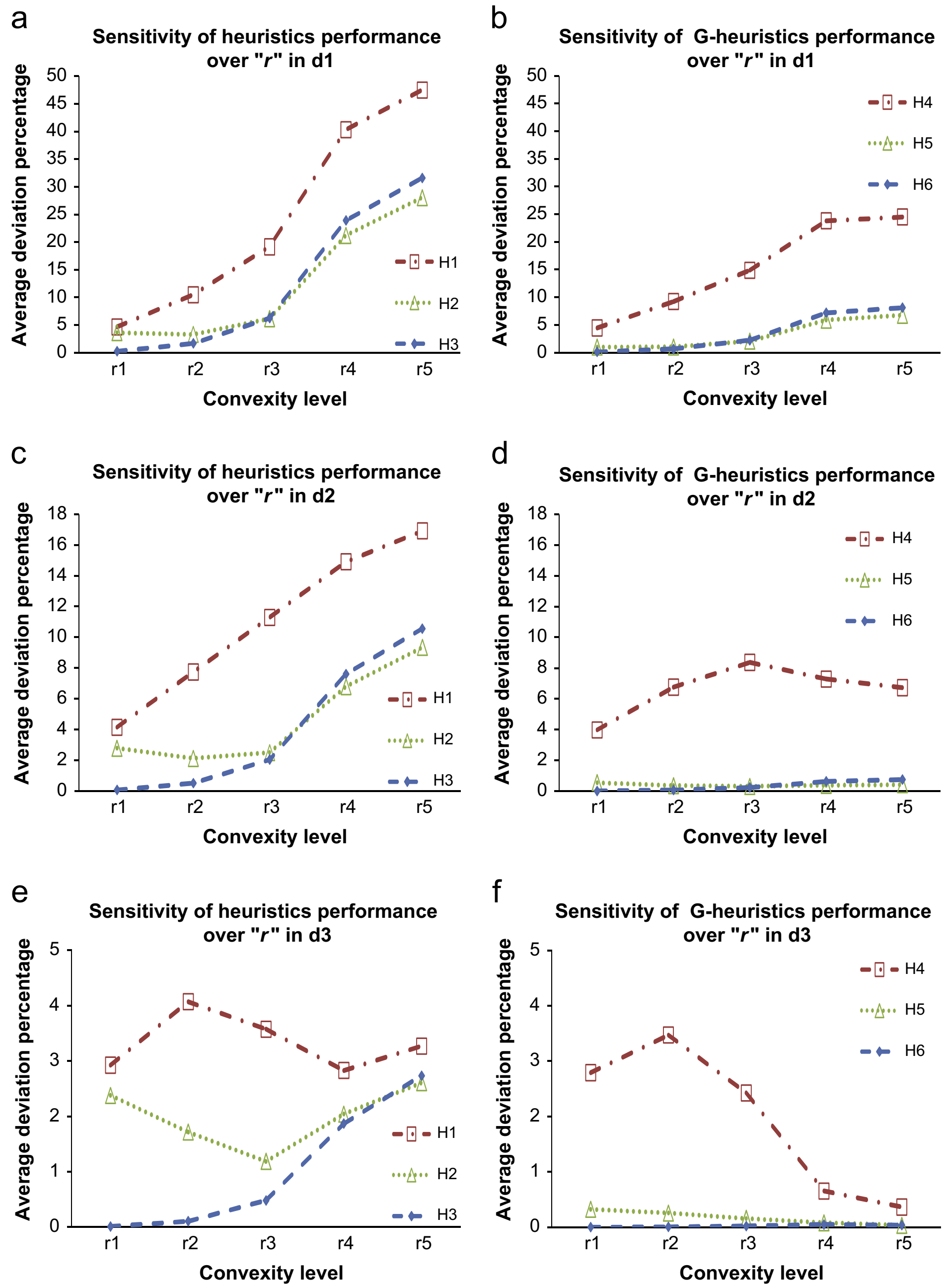

Fig. 1. Average percentage deviation of heuristics versus production cost convexity levels for $d1$, $d2$ and $d3$.

for the heuristics for different demand patterns and cost parameters. (See also Figs. 1 and 2 for pictorial depictions.) The performances of the heuristics are roughly similar for different measures (minimum, maximum percentage deviations, etc.); hence, we focus on the average percentage deviations. On average, the performances of heuristics $H1$–$H3$ deteriorate as demand variance increases. The performance also worsens as the nonlinearity in production cost (r) increases for moderate and high demand variances ($d1$ and $d2$). For low variance ($d3$), the performances of all heuristics lie within the 4% band with those of heuristics $H1$ and $H3$ being relatively less sensitive to r. The performance gap among the three heuristics gets smaller for large r values. With respect to the fixed setup cost, the performances of heuristics $H1$ and $H3$ improve as K increases for all demand patterns. The performance of heuristic $H2$ improves for moderate and high variance levels but worsens slightly for $d3$. Although, heuristic $H3$ was deemed to be the best performer in general, it does not do so well in comparison with $H2$ for large r values. As the production cost becomes more and more nonlinear, this heuristic starts to underperform especially as demand variance increases. This is due to the fact that heuristic $H2$ allows for demand splitting while heuristic $H3$ cannot smooth the production over successive periods. Heuristics $H2$ and $H3$ have similar performances for low and moderate values of fixed setup cost, but the former performs slightly better for large K values. Once again, this indicates that the variant of Harris's formula captures the fundamental trade-offs. Worst case performance is also of theoretical and practical interest. In terms of maximum percentage deviations, heuristic $H1$ always results in the maximum percentage deviations. Heuristic $H3$ clearly dominates $H2$ for small nonlinearity in production cost but their performance gaps decrease as r and K get large. Although $H3$ is the best performer

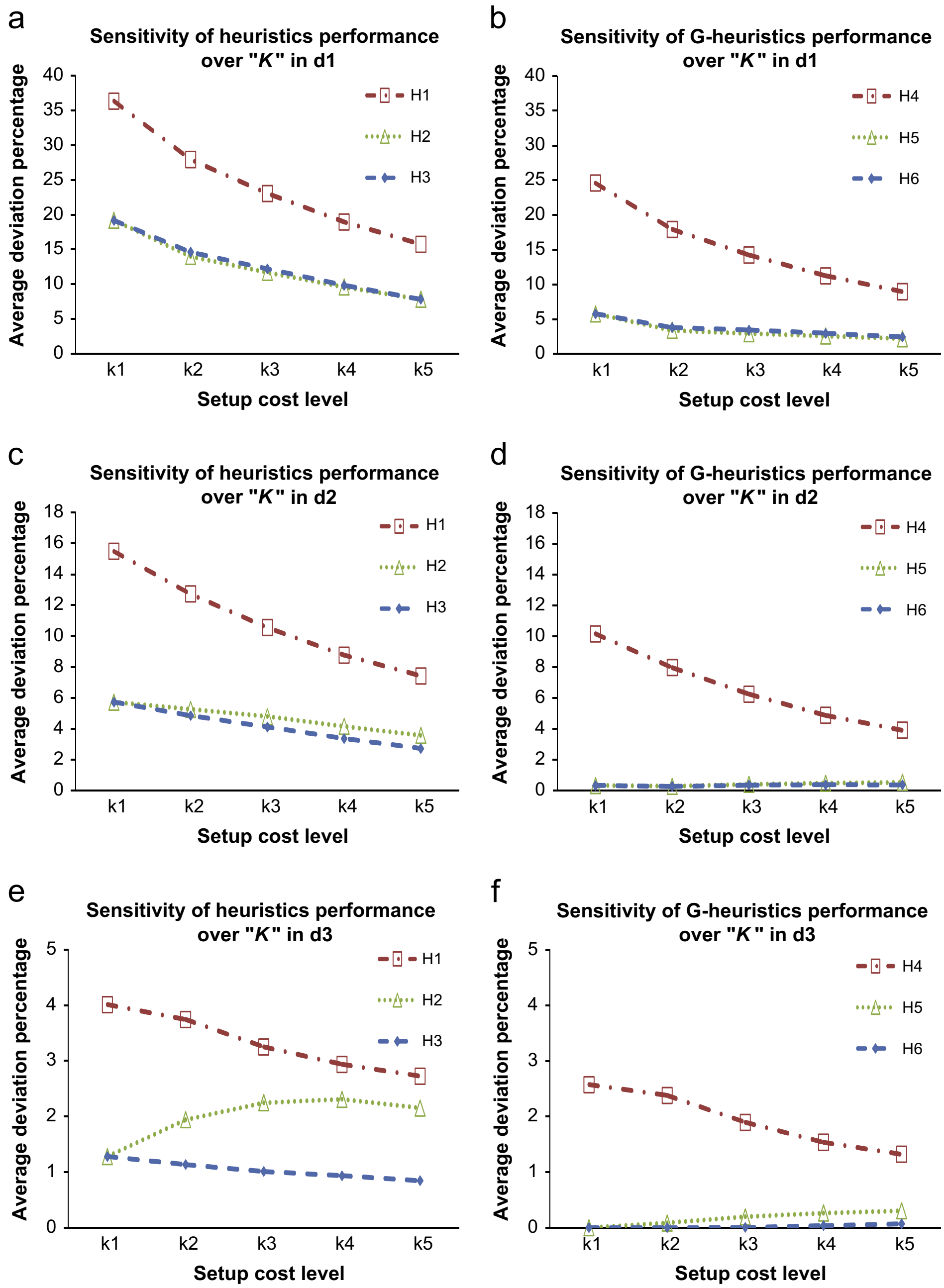

Fig. 2. Average percentage deviation of heuristics versus setup cost levels for *d*1, *d*2 and *d*3.

in this group and may result in zero deviation from the benchmark solution in a large portion of the solutions, it is still important to point out that imposing demand integrality may cause very large deviations (up to 65%) in certain cases. The performance deteriorates rapidly with large production cost nonlinearity and low setup costs. As we discuss below for the G-class heuristics, exploiting the optimal generation structures for given production period decisions does improve the solutions significantly. The last group of heuristics (*H*4–*H*6) results in significant improvements over the first group. In some cases, the average improvements are about 20-fold. Actually, the large performance difference between the heuristic groups directly implies the importance of class G production plans in a solution and points to the impact of production cost non-linearity. In Table 5, we report the performance statistics for this group for $1000(=10\times100)$ experiment instances for each of the three different demand patterns. Note that in some experiment instances, heuristics *H*5 and *H*6 obtained

Table 5
Percentage deviation statistics for heuristics *H*4–*H*6.

Demand	Algorithm	Min	Max	Median	Average	$N(0)$	$N(-)$
*d*1	*H*4	1.09	44.17	13.06	15.39	0	0
	*H*5	0.06	18.87	1.8	3.36	0	0
	*H*6	−0.01	18.87	1.73	3.68	157	2
*d*2	*H*4	0.79	15.84	6.78	6.63	0	0
	*H*5	−0.01	2.28	0.29	0.4	65	5
	*H*6	−0.01	2.37	0.19	0.34	223	3
*d*3	*H*4	0	7.04	1.37	1.94	114	0
	*H*5	0	1.29	0	0.17	636	0
	*H*6	0	0.65	0	0.02	808	0

better solutions than the benchmark forward DP rendering minimum deviations negative. Also, the number of instances for which the forward DP solution was obtained increases with this group;

Table 6
Percentage deviation statistics of heuristics $H1$–$H3$.

Heuristic	Statistic	Production cost exponent levels					Setup cost levels				
		$r1$	$r2$	$r3$	$r4$	$r5$	$k1$	$k2$	$k3$	$k4$	$k5$
$H1$	Min	1.23	1.48	2.92	8.77	9.65	8.59	2.23	1.23	1.25	1.24
($d1$)	Max	12.99	25.76	40.07	74.28	89.23	89.23	83.29	79.88	74.21	68.06
	Median	3.09	9.65	20.38	42.87	49.27	30.19	21.94	18.83	14.14	9.65
	Average	*4.67*	*10.5*	*19.14*	*40.31*	*47.47*	*36.36*	*27.95*	*23.06*	*18.96*	*15.75*
$H2$	Min	0.08	0.51	1	1.6	1.64	0.08	0.3	0.25	0.4	0.65
	Max	11.81	9.41	17.54	48.4	65.03	65.03	57.72	53.51	48.26	41.64
	Median	2.12	2.87	4.98	23.07	29.91	11.39	7.66	6.28	4.94	4.46
	Average	*3.61*	*3.3*	*6.14*	*21.21*	*28.03*	*19.19*	*13.98*	*11.7*	*9.6*	*7.83*
$H3$	Min	0	0	0.03	2.26	3.49	0.08	0	0	0	0
	Max	1.94	6.86	17.54	48.4	65.03	65.03	58.47	55.12	50.89	46.09
	Median	0.01	1.08	6.03	25.47	33.35	11.39	7.01	5.29	3.75	2.15
	Average	*0.25*	*1.68*	*6.28*	*23.91*	*31.6*	*19.19*	*14.64*	*12.21*	*9.85*	*7.83*
$H1$	Min	0.79	1.12	2.05	4.74	5.15	8.06	2.17	0.93	0.79	0.97
($d2$)	Max	11.15	17.48	21.35	25.67	28.2	28.2	28.03	27.56	25.63	24.95
	Median	2.42	7.94	11.72	15.56	16.99	14.8	12.58	10.14	7.67	6.19
	Average	*4.14*	*7.74*	*11.29*	*14.9*	*16.91*	*15.49*	*12.73*	*10.56*	*8.77*	*7.43*
$H2$	Min	0	0.14	0.52	0.75	0.94	0	0.19	0.12	0.51	0.18
	Max	8.86	6.66	5.06	14.76	19.9	19.9	18.91	17.78	16.09	14.19
	Median	2.09	1.3	2.51	7.1	10.08	3.05	3.65	3.77	3.12	2.46
	Average	*2.78*	*2.12*	*2.5*	*6.79*	*9.34*	*5.73*	*5.26*	*4.79*	*4.16*	*3.57*
$H3$	Min	0	0	0	0.77	1.41	0	0	0	0	0
	Max	0.44	1.82	5.06	14.76	19.9	19.9	19.14	18.29	17.2	15.9
	Median	0	0.27	1.92	7.99	10.85	3.05	2.48	1.85	1.24	0.8
	Average	*0.06*	*0.5*	*2.04*	*7.6*	*10.56*	*5.73*	*4.85*	*4.11*	*3.36*	*2.71*
$H1$	Min	0.5	0.69	0.88	0.79	1.06	0.79	0.83	0.71	0.58	0.5
($d3$)	Max	7.32	7.59	6.2	5.05	6.36	7.32	7.02	7.59	6.83	5.58
	Median	1.92	4.33	3.81	2.81	3.19	3.89	3.59	3.01	2.8	2.76
	Average	*2.92*	*4.07*	*3.57*	*2.83*	*3.27*	*4.01*	*3.74*	*3.25*	*2.94*	*2.72*
$H2$	Min	0	0	0.08	0.46	0.41	0	0.02	0.03	0.17	0.09
	Max	7.4	6.43	4.29	3.94	5.37	5.37	7.4	7.07	6.58	6.7
	Median	1.84	0.3	0.77	2.18	2.55	0.53	1.27	2.05	1.91	1.69
	Average	*2.38*	*1.72*	*1.18*	*2.04*	*2.61*	*1.28*	*1.94*	*2.25*	*2.31*	*2.15*
$H3$	Min	0	0	0	0.36	0.81	0	0	0	0	0
	Max	0.09	0.39	1.26	3.94	5.37	5.37	5.17	4.94	4.64	4.31
	Median	0	0.06	0.45	1.79	2.6	0.53	0.47	0.36	0.47	0.34
	Average	*0.01*	*0.11*	*0.48*	*1.87*	*2.74*	*1.28*	*1.13*	*1.01*	*0.93*	*0.84*

$1188(=157+223+808)$ instances for heuristic $H6$, $701(=65+636)$ instances for $H5$ and 114 instances for $H4$. The number of zero percentage deviation solutions increases as the demand variance gets smaller. In this group, two subgroups emerge: The pair of heuristics $H5$ and $H6$, and $H4$. The former pair dominates the latter in all performance measures by a relatively large margin. Heuristics $H5$ and $H6$ perform almost equally well; the magnitude of deviations is small when they differ most for low demand variance. In Table 7, we present the performance statistics with respect to r and K. (See also Figs. 1 and 2 for pictorial depictions.) With respect to nonlinearity in production cost, the behavior of the heuristics are not monotone. However, overall, there is a tendency for the performances to deteriorate as r gets large. Likewise, demand variance impacts the performances negatively. An interesting observation is that, with large demand variance, r impacts performances negatively whereas with low demand variance, all heuristics tend to converge to the benchmark. (See Fig. 1(f).) The effect of the fixed setup cost K is similar to that observed for the first group but with smaller deviations. For Heuristics $H5$ and $H6$, the improvement of performance for smaller demand variances becomes more pronounced than those of their counterparts $H2$ and $H3$. Lastly, we should mention the computational efficiency achieved by the proposed heuristics. As discussed above, the backward DP formulation which guarantees optimality is prohibitively slow and memory-inefficient for practical use. The computational times statistics measured in seconds for each heuristic and the benchmark forward DP algorithm for different demand patterns are presented in Table 8. As expected, heuristics $H1$–$H3$ have similar and the smallest run times, and $H4$–$H6$ have relatively larger and similar times but are still reasonably fast. The benchmark forward DP algorithm solves on average in roughly four to six minutes. Note that the times for pre-processing (arising from solving by $H1$–$H3$) needed for $H4$–$H6$ are included in the computational times. The statistics provided for the heuristics have been obtained for a 2.3 GHz processor whereas those for the benchmark have been obtained for a 3.3 GHz processor. To conclude, our numerical comparison of the heuristics reveals that (i) imposing demand integrality (using $H3$) may result in large deviations, especially in the presence of large production nonlinearities and low setup costs, (ii) a variant of Harris's formula that captures the fundamental trade-offs among setup costs, inventory holding costs and production costs provides a quick and reasonably good heuristic ($H1$), (iii) construction of G-class subplans (using heuristics $H4$–$H6$) is essential in developing heuristic solutions for the dynamic lot sizing problem in the presence of production cost nonlinearities, (iv) the computational time improvements through the proposed heuristics are justifiably significant, and finally (v) among all the proposed heuristics, $H6$ and, then, $H5$ are, by far, the best ones.

Table 7
Percentage deviation statistics of heuristics $H4$–$H6$.

Heuristic	Statistic	Production cost exponent levels					Setup cost levels				
		$r1$	$r2$	$r3$	$r4$	$r5$	$k1$	$k2$	$k3$	$k4$	$k5$
$H4$ ($d1$)	Min	1.09	1.32	1.93	3.5	4.04	7.98	2.22	1.23	1.13	1.09
	Max	12.94	22.12	31.05	41.25	44.17	44.17	40.22	38.75	35.25	32.28
	Median	2.88	8.35	15.64	26.45	26.33	24.84	17.02	12.16	7.59	4.88
	Average	*4.47*	*9.23*	*14.87*	*23.83*	*24.52*	*24.59*	*17.92*	*14.23*	*11.23*	*8.95*
$H5$	Min	0.06	0.08	0.24	0.3	0.45	0.07	0.06	0.06	0.07	0.13
	Max	5.22	3.3	7.15	15.29	18.87	18.87	15.08	13.76	12.78	10.82
	Median	0.64	0.84	1.41	6.06	6.66	4.65	1.44	1.61	1.5	1.33
	Average	*1.02*	*1.05*	*2.07*	*5.88*	*6.77*	*5.75*	*3.37*	*2.92*	*2.56*	*2.18*
$H6$	Min	−0.01	0	0	0.47	0.75	0.07	0	−0.01	−0.01	0
	Max	1.11	3.3	7.15	15.29	18.87	18.87	15.47	15.05	13.77	13.45
	Median	0	0.36	1.89	7.8	8.77	4.65	1.79	1.78	1.22	0.73
	Average	*0.13*	*0.68*	*2.24*	*7.21*	*8.12*	*5.75*	*3.78*	*3.44*	*2.95*	*2.43*
$H4$ ($d2$)	Min	0.79	1.09	1.28	1.89	1.95	4.43	2.17	0.93	0.79	0.95
	Max	11.15	15.84	15.3	12.88	11.46	15.84	14.35	14.58	13.46	11.5
	Median	2.29	6.43	9.19	7.53	6.68	10.02	7.89	6.18	4.17	2.76
	Average	*3.98*	*6.76*	*8.38*	*7.29*	*6.73*	*10.15*	*7.96*	*6.25*	*4.87*	*3.91*
$H5$	Min	0	0	0	0	−0.01	0	−0.01	−0.01	−0.01	−0.01
	Max	2.28	1.68	1.21	1.17	1.56	1.56	1.24	1.96	2.16	2.28
	Median	0.38	0.13	0.21	0.32	0.35	0.21	0.17	0.25	0.35	0.38
	Average	*0.55*	*0.36*	*0.31*	*0.37*	*0.42*	*0.33*	*0.27*	*0.4*	*0.48*	*0.54*
$H6$	Min	0	0	−0.01	0	−0.01	0	−0.01	0	0	−0.01
	Max	0.08	0.29	0.89	2	2.37	1.56	1.32	1.61	2.24	2.37
	Median	0	0.05	0.2	0.56	0.69	0.21	0.17	0.19	0.21	0.17
	Average	*0.01*	*0.07*	*0.24*	*0.63*	*0.75*	*0.33*	*0.27*	*0.35*	*0.37*	*0.37*
$H4$ ($d3$)	Min	0.5	0.69	0.7	0	0	0	0	0	0	0
	Max	7.04	6.19	4.32	2.1	1.54	7.04	6.13	6.19	5.35	4.04
	Median	1.83	3.89	2.47	0.54	0.25	1.86	2.57	1.6	1.24	1.13
	Average	*2.79*	*3.47*	*2.43*	*0.65*	*0.36*	*2.58*	*2.38*	*1.9*	*1.54*	*1.32*
$H5$	Min	0	0	0	0	0	0	0	0	0	0
	Max	1.29	1.24	1.03	0.76	0.51	0.04	0.99	1.16	1.09	1.29
	Median	0.26	0	0	0	0	0	0	0	0.16	0.27
	Average	*0.32*	*0.26*	*0.16*	*0.08*	*0.04*	*0*	*0.09*	*0.2*	*0.26*	*0.31*
$H6$	Min	0	0	0	0	0	0	0	0	0	0
	Max	0.04	0.09	0.19	0.5	0.65	0.04	0.02	0.12	0.5	0.65
	Median	0	0	0	0	0	0	0	0	0	0
	Average	*0*	*0.01*	*0.02*	*0.05*	*0.04*	*0*	*0*	*0.01*	*0.04*	*0.07*

Table 8
Execution time statistics for the entire experiment set measured in seconds. ('0' indicates run time less than 1 ms.)

	$H1$	$H2$	$H3$	$H4$	$H5$	$H6$	DP
$d1$							
Min	0	0	0.07	0.16	0.36	0.21	0
Max	0.02	0.02	0.34	13.48	17.68	13.54	359.64
Median	0	0	0.08	0.47	1.77	0.64	254.26
Average	0	0	0.1	1.03	2.39	1.27	207.87
$d2$							
Min	0	0	0.08	0.16	0.38	0.23	0
Max	0.02	0.02	0.36	13.7	8.69	7.43	329.89
Median	0	0	0.09	0.56	1.62	0.69	284.87
Average	0	0	0.11	0.92	1.79	1.01	230.56
$d3$							
Min	0	0	0.08	0.16	0.39	0.24	0
Max	0.03	0.02	0.26	3.32	3.86	4.85	373.74
Median	0	0	0.08	0.53	1.78	0.7	254.31
Average	0	0	0.09	0.61	1.67	0.79	240.08

5.2. Sensitivity analysis

We study the impact of experiment parameters through (i) the average minimum total cost over the horizon, (ii) the average number of generations over the problem horizon, (iii) the average percentage of production periods, and (iv) the average percentage of generations of certain types. All findings are based on the benchmark solutions.

Average minimum total cost. The minimum total costs behave as expected. (See Fig. 3(a) and (b) for an illustration of costs averaged over replications and all r and a values and selected two K, values.) As K decreases and a increases, the minimum total cost decreases. The impact of non-linearity in variable production costs (measured through r) becomes more pronounced with larger K. The parameters a and r interact in the variable production cost component. That is, the same cost can be obtained by distinct pairs of a and r; smaller a and r in one, and larger a and r in the other. Furthermore, the average total cost is not monotone in r.

Average number of generations. In Fig. 3(c) and (d), we illustrate the number of generations in the optimal solution for $P_{1,T}$ averaged over the 10 replications for all r and a values and $K=\{K1, K5\}$. We begin our discussion with zero setup costs. For $K1$ and linear variable production cost ($r0$), the Lot-for-Lot (LFL) solution is obtained. In this case, the solution is insensitive to a values as total production cost over the horizon is constant, $(h/a)D_{1,T}$. As r increases, the solution deviates from LFL, splitting some production quantities over a number of periods in order to exploit the marginal cost structure. This smoothes out the

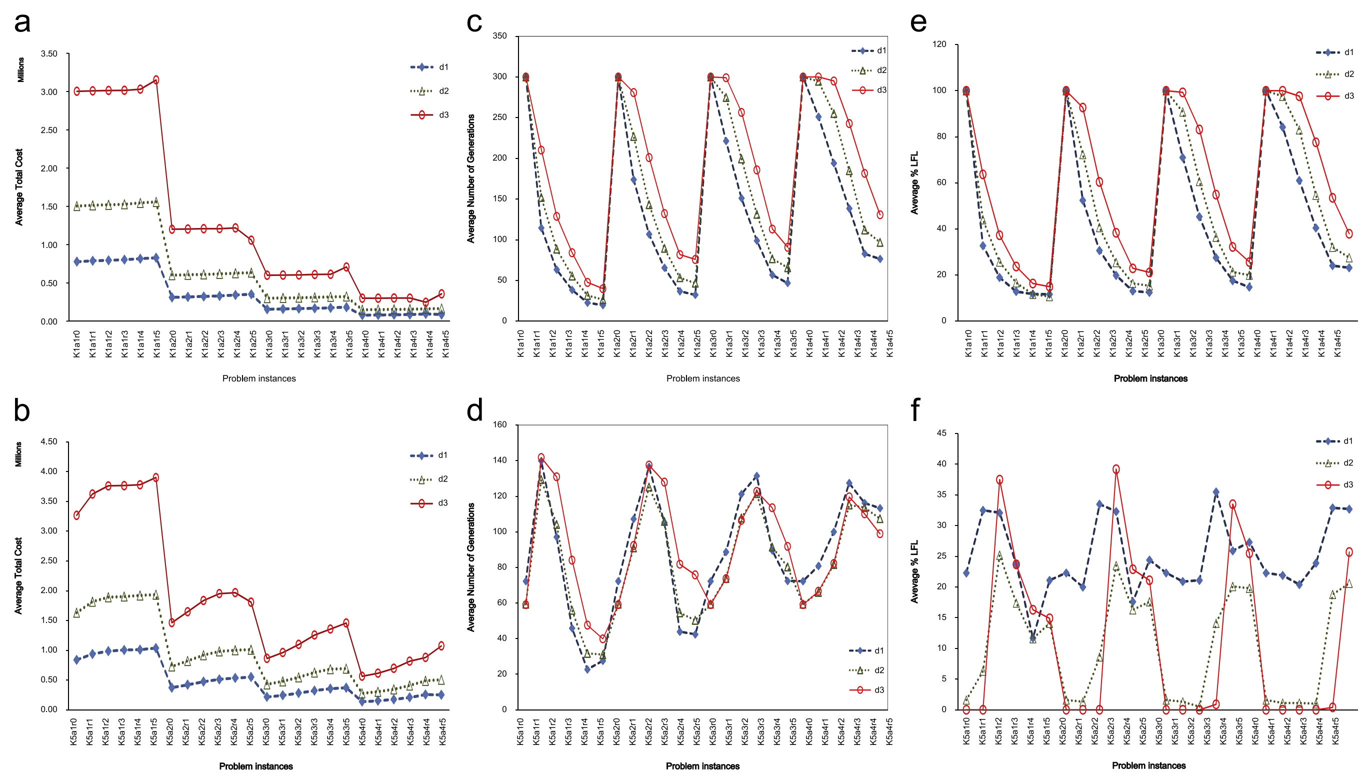

Fig. 3. Impact of system parameters on average total cost ((a), (b)), average number of generations ((c), (d)), and average percentage of LFL generations ((e), (f)) for $K1$ and $K5$ over 10 replications.

production plan resulting in longer generations on average. As demand's coefficient of variation increases, this effect increases. Increasing a, which decreases cost for a production quantity allowing for larger production quantities at the same cost, counters the advantages of smoothing out. As setup costs increase, batching effects dominate extending the generations lengths resulting in a smaller average number of generations in the solution. (Compare overall tendencies for $K1$ and $K5$.) However, the interaction between the setup cost and the variable production cost component is not simple. As K increases, on one hand, it is desirable to increase the quantity q_t in a production period to compensate for the fixed cost; on the other hand, increasing quantities result in a tendency to split/smooth out production as observed before.

The percentage of production periods. The percentage of production periods is of interest from a number of perspectives: It stands for the number of setups needed over a problem horizon; it can be used as a proxy for the average quantity produced over a problem horizon; and finally, it is a measure for utilization of production assets which is of managerial concern. For zero setup costs and $r0$, there is production only in periods with non-zero demands. As $r(>1)$ increases, it becomes more advantageous to do a new setup due to increasing marginal production costs. Thus, smoothing out of production occurs resulting in larger percentages of production periods. For positive setup costs, batching effects emerge reducing the percentage but the effect of r is as for zero setup costs. For a given $r(>1)$, the percentage of production periods decreases as a increases. This arises from the fact that lower a values allow for producing in larger quantities at the same cost levels. The effect of a becomes more pronounced with increasing fixed setup costs, as K exacerbates batching effects. As the demand patterns change from $d1$ to $d3$, the percentage of production periods increases. This is due to two reasons. First, the average demand per period μ increases, rendering a new setup more economical to do than to incur increasing marginal production costs. Second, the variability of demand (measured either as fraction of zeros or coefficient of variation) decreases that allows for smoother production plans.

Percentage of different types of production sequences. In our problem, a production sequence may have four forms: (i) A single period in which no production is done (a generation with zero demands). (ii) A single period in which the production equals that period's demand (LFL). (iii) A single production period followed by a number of no-production periods; the production quantity covers the total demand in the generation (Wagner–Whitin type). Finally, (iv) a combination of production and no-production periods, to which we refer as non-Wagner–Whitin type. In regard to the occurrence of these types, the overall findings can be summarized as follows. As fixed setup cost K increases, production quantities tend to increase to compensate for the fixed cost. This results in longer generations as more demands of future periods are covered by production in a particular period. As r increases, the tendency to split production over periods (production smoothing) increases. This also results in longer generations on average. Finally, as the ratio of holding cost to mean production cost per unit a increases, production quantities tend to increase since larger lots can be produced at the same cost. The result is again longer generations on average. The two fundamental tendencies – batching and production smoothing – manifest themselves in the production sequence types. When the batching effect is more dominant, more LFL or Wagner–Whitin type sequences occur in the optimal production plan. When production smoothing effect is more dominant, the optimal production plan consists of more non-Wagner–Whitin type production sequences of class G.

We report the percentage of LFL production sequences over all generations in the optimal solution for $P_{1,T}$ averaged over the 10 replications for all r and a values and selected two K values in Fig. 3(e) and (f). For zero setup costs and $r0$, all production sequences are LFL, as expected. As $r(>1)$ increases, production smoothing tendency increases and generations with class G production sequences emerge resulting in decreasing LFL percentages. As a increases, variable production cost decreases allowing for larger production quantities for a given $r(>1)$ value. Hence, we see non-decreasing LFL percentages for a given $r(>1)$ as a increases. As K increases, it is desirable to increase the quantity qt in a production period to compensate for the fixed cost; hence, LFL structure becomes less desirable resulting in lower percentages of LFL generations. Likewise, increasing quantities result in a tendency to split/smooth out production as observed before. The confounded effect can easily be seen throughout the figure. Comparing the figure, we see a striking increase in the cases with zero LFL generations. For zero setup costs, percentage of LFL generations increases as the demand pattern changes from $d1$ to $d2$ to $d3$; but, the same does not hold for positive K. We observe wider ranges with respect to $r(>1)$ as demand pattern changes from $d1$ to $d3$. The same behavior appears also as a increases, especially for large setup costs ($K5$).

6. Conclusion

In this work, we considered the dynamic lot-sizing problem for a single item with deterministic demands and non-linear production costs that arise from well-known economic production functions (*e.g.*, Cobb–Douglas and Leontieff functions) and externalities in the production activity such as pollution control efforts. The optimal solution with convex production functions exhibits behavior dissimilar to that in the classical lot-sizing problem. In particular, it is now possible to produce in a period even if its net demand is zero, and to produce part of the demand of a production period at an earlier time. We characterized the structure of the optimal policy for zero and positive setup costs. We provided further results that enabled us to develop a forward DP algorithm which guarantees optimality with $O(T^2 2^T)$ run time complexity for the problem with horizon length T. For the case of zero setup costs, it reduces to $O(T^2)$ complexity. The fundamental property of the optimal solutions is that production subplans exhibit a specific structure – herein referred to as G-class subplans. This property is retained for positive setup cots, as well. Based on this property, a version of the forward DP algorithm for positive setup costs was developed that employs three simple rules conjunctively to generate production sequences has computational complexity of $O(T^4)$ and it performed well in numerical tests. We also proposed six heuristics to solve the problem. The first three of them are single step heuristics. The remainder are two-step heuristics that improve on an initial solution obtained by the first group. They exploit the G-class production subplans and outperform the first group significantly. The best single step heuristic is a variant of the Wagner–Whitin solution algorithm in which demand integrality is imposed on the production quantities. The best G-class heuristic is its improved version. The computational time improvements through the proposed heuristics are justifiably significant. In our numerical study, we also investigated the sensitivity of the optimal production plans to the non-linearity in production functions, the average unit cost of production and setup costs. Our findings revealed the fundamental trade-offs between the batching and production smoothing tendencies. Production smoothing that is observed in our problem has important managerial implications. In an ERP (enterprise resources planning) setting, the same level resource requirements such as labor are also smoothed out. Also higher level requirements become less lumpy (more uniform) allowing for simpler and steady delivery schedules and possible cost advantages. Capacity utilization rates also increase. Although we have considered an uncapacitated problem, production smoothing also has a positive impact on investment needs as capacity requirements per period are reduced. Overall, our model and results obtained for a single echelon may be viewed as a building block for analysis of such and other managerial issues in richer, multiple echelon settings.

Acknowledgements

This research is partially funded by the Research Grant #110M307 and Graduate Scholarship Program for International Students Program 2215 by the Scientific and Technological Research Council of Turkey (TUBITAK).

Appendix

Proof of Theorem 2. (i) Directly follows from the definition of a generation. (ii) If $r_t^n \le 1$ for $t \in [u,v]$, then the production costs are concave. The problem reduces to the classical lot-sizing problem and the result follows. (iii) From Theorem 1, $Q^*_{u,v}$ can be found by solving $P_{u,v}$ independently. The proof rests on obtaining the optimal solution for the sub-problem $P_{u,v}$. Similar to Eq. (1a), we can write the total cost over the generation $\langle u,v\rangle$, g_{uv} as follows:

$$g_{uv} = \sum_{t \in S(Q^*_{uv})} K_t + \sum_{t=u}^{v}\left[\left(\sum_{i=t}^{v} h_i\right) q_t + \sum_{n=1}^{m} w_t^n q_t^{r_t^n}\right] - \sum_{t=u}^{v} h_t D_{u,t}$$

Note that g_{uv} is convex in the production quantities, and that the feasible region defined by Eqs. (1b) and (1c) is regular. Constructing the Lagrangian $\mathcal{L}_{uv}$ for $P_{u,v}$, we have

$$\mathcal{L}_{uv} = g_{uv} - \sum_{t=u}^{v}\left[\lambda_t\left(\sum_{i=u}^{t} q_i - \sum_{i=u}^{t} d_i\right) + \mu_t q_t\right]$$

where λ_t and μ_t denote the shadow prices of the constraints. From the first order optimality (Karush–Kuhn–Tucker) conditions, we have, for $t \in S(Q^*_{uv})$,

$$\mu_t q_t^* = 0, \tag{7}$$

$$\lambda_t\left(\sum_{i=u}^{t} q_i^* - \sum_{i=u}^{t} d_i\right) = 0, \tag{8}$$

$$\left.\frac{\partial \mathcal{L}}{\partial q_t}\right|_{q_t = q_t^*} = \sum_{n=1}^{m} r_t^n w_t^n q_t^* r_t^{n-1} + \sum_{i=t}^{v}(h_i - \lambda_i) - \mu_t = 0. \tag{9}$$

Since in a generation, $I_t = (\sum_{i=u}^{t} q_i - \sum_{i=u}^{t} d_i) > 0$ for $t \in \{u, \dots, v-1\}$, (8) implies that

$$\lambda_t^* = 0, \quad t \in \{u, \dots, v-1\}. \tag{10}$$

Substituting (10) into (9) we find

$$\sum_{n=1}^{m} r_t^n w_t^n q_t^* r_t^{n-1} = \lambda_v - \sum_{i=t}^{v} h_i + \mu_t \quad \forall t \in S(Q^*_{uv}) \tag{11}$$

Now, $q_t^* > 0$ for $\forall t \in S(Q^*_{uv})$ together with (7) implies that $\mu_t^* = 0$ for $t \in S(Q^*_{uv})$. Hence, writing (11) for $i<j$, $i,j \in S(Q^*_{uv})$ gives $\sum_{n=1}^{m} r_i^n w_i^n q_i^* r_i^{n-1} = \lambda_v - \sum_{s=i}^{v} h_s$ and $\sum_{n=1}^{m} r_j^n w_j^n q_j^* r_j^{n-1} = \lambda_v - \sum_{s=j}^{v} h_s$. Equating the two expressions via λ_v, $\sum_{n=1}^{m} r_i^n w_i^n q_i^* r_i^{n-1} = \sum_{n=1}^{m} r_j^n w_j^n q_j^* r_j^{n-1} - \sum_{s=i}^{j-1} h_s$. Clearly, this is a production plan of class G. Hence, the result. □

Proof of Theorem 3. Let $Q^*_{u,v'}$ be the optimal production plan for $[u,v']$. If $q^*_{v+1} = 0$, the result follows immediately. Otherwise, consider the modified feasible production plan $Q'_{u,v'} = (q'_u, \dots, q'_{v'})$ such that $q'_l = q^*_l + \epsilon$, $q'_{v+1} = q^*_{v+1} - \epsilon$ and $q'_t = q^*_t$ for $t \in \{u, u+1, \dots, l-1, l+1, \dots, v, v+2, \dots, v'\}$ where $\epsilon > 0$. Due to the optimality of $Q^*_{u,v'}$ we have

$$\sum_{i=u}^{v'}\left[\sum_{n=1}^{m} w_i^n q_i^* r_i^n + h_i I_i\right] \le \sum_{i=u}^{l-1}\left[\sum_{n=1}^{m} w_i^n q_i^* r_i^n + h_i I_i\right] + \sum_{n=1}^{m} w_l^n (q_l^*+\epsilon)^{r_l^n} + \sum_{i=l}^{v} h_i(I_i+\epsilon) + \sum_{n=1}^{m} w_{v+1}^n (q_{v+1}^*-\epsilon)^{r_{v+1}^n} + h_{v+1}I_{v+1} + \sum_{i=v+2}^{v'}\left[\sum_{n=1}^{m} w_i^n q_i^* r_i^n + h_i I_i\right]$$

which implies $\sum_{n=1}^{m}[w_{v+1}^n q_{v+1}^* r_{v+1}^n - w_{v+1}^n (q_{v+1}^*-\epsilon)^{r_{v+1}^n}] \le \sum_n^m [w_l^n (q_l^*+\epsilon)^{r_l^n} - w_l^n q_l^* r_l^n] + (h_l+\cdots+h_v)\epsilon$.

Dividing both sides by ϵ and taking the limit $\epsilon \to 0$, we get

$$\sum_{n=1}^{m} w_{v+1}^n \frac{d}{dq_{v+1}} q_{v+1}^{r_{v+1}^n}\big|_{q_{v+1}=q_{v+1}^*} \le \sum_{n=1}^{m} w_l^n \frac{d}{dq_l} q_l^{r_l^n}\big|_{q_l=q_l^*} + (h_l+\cdots+h_v)$$

which becomes

$$\sum_{n=1}^{m} r_{v+1}^n w_{v+1}^n q_{v+1}^* r_{v+1}^n - 1 \le \sum_{n=1}^{m} r_l^n w_l^n q_l^* r_l^{n-1} + (h_l+\cdots+h_v). \quad \square$$

Proof of Lemma 2. Proof by contradiction. We first establish that $q^*_{t+1} > 0$. Suppose that in the subplan $Q_{u,v}$, $q_t > 0$ and $q_{t+1} = 0$. We will show that this subplan can be improved. To do so, consider the feasible subplan $Q'_{u,v} = (q'_u, \ldots, q'_v)$ with $q'_t = q_t - \epsilon$, $q'_{t+1} = \epsilon$ and $q'_i = q_i$ for $i \in \{u, \ldots, v\} \backslash \{t, t+1\}$ such that $0 < \epsilon < \min\{I_t, q_t, 1\}$. By definition of a generation, I_t is positive and q_t is positive by assumption. Therefore, such a positive ϵ which guarantees the feasibility of $Q'_{u,v}$ always can be found. We denote the corresponding costs of these two subplans by π and π'

$$\tilde{\pi}-\pi' = \left[\sum_{j=1}^{m} w_t^j q_t^{r_t^j} + h_t I_t\right] - \left[\sum_{j=1}^{m} w_t^j (q_t-\epsilon)^{r_t^j} + h_t(I_t-\epsilon) + \sum_{j=1}^{m} w_{t+1}^j \epsilon^{r_{t+1}^j}\right]$$
$$= \sum_{j=1}^{m} w_t^j (q_t^{r_t^j} - (q_t-\epsilon)^{r_t^j}) + h_t\epsilon - \sum_{j=1}^{m} w_{t+1}^j \epsilon^{r_{t+1}^j}$$
$$\ge h_t\epsilon - \sum_{j=1}^{m} w_{t+1}^j \epsilon^{r_{t+1}^j} = \epsilon\left(h_t - \sum_{j=1}^{m} w_{t+1}^j \epsilon^{r_{t+1}^j - 1}\right).$$

The inequality above follows from nonnegativity of the parenthetical term in the former expression. Letting $\overline{w}_{t+1} = \max_j\{w_{t+1}^j\}$ and $\overline{r}_{t+1} = \min_j\{r_{t+1}^j\}$, the last expression is positive for any $\epsilon < (h_t/\overline{w}_{t+1})^{1/(\overline{r}_{t+1}-1)}$ if h_t is positive. If h_t=0, then

$$\pi-\pi' = \sum_{j=1}^{m}[w_t^j q_t^{r_t^j}] - \sum_{j=1}^{m}[w_t^j (q_t-\epsilon)^{r_t^j} + w_{t+1}^j \epsilon^{r_{t+1}^j}]$$
$$= \sum_{j=1}^{m} w_t^j [q_t^{r_t^j} - (q_t-\epsilon)^{r_t^j}] - \sum_{j=1}^{m} w_{t+1}^j \epsilon^{r_{t+1}^j}.$$

Consider the function $f(x) = q^x - (q-\epsilon)^x$. Derivative of this function with respect to x is $f'(x) = \ln(q)q^x - \ln(q-\epsilon)(q-\epsilon)^x$ which is always positive for $(q > \epsilon)$. Therefore, $f(x)$ is an increasing function of x for $(q > \epsilon)$. Let $\lfloor x \rfloor$ be the greatest integer equal to or less than x. Then, $w_t^j [q_t^{r_t^j} - (q_t-\epsilon)^{r_t^j}] \ge w_t^j [q_t^{\lfloor r_t^j \rfloor} - (q_t-\epsilon)^{\lfloor r_t^j \rfloor}]$ for $j = 1, \ldots, m$. Suppose $\epsilon < 1$. Then, $w_{t+1}\epsilon^{r_{t+1}^j} < w_{t+1}\epsilon^{\lfloor r_{t+1}^j \rfloor}$ for all j, as well. Therefore,

$$\pi-\pi' \ge \sum_{j=1}^{m} w_t^j [q_t^{\lfloor r_t^j \rfloor} - (q_t-\epsilon)^{\lfloor r_t^j \rfloor}] - \sum_{j=1}^{m} w_{t+1}^j \epsilon^{\lfloor r_{t+1}^j \rfloor}$$
$$= \sum_{j=1}^{m} w_t \left[q_t^{\lfloor r_t^j \rfloor} - q_t^{\lfloor r_t^j \rfloor} + \binom{\lfloor r_t^j \rfloor}{1} q_t^{\lfloor r_t^j \rfloor - 1}\epsilon + q_t^{\lfloor r_t^j \rfloor - 2} - \binom{\lfloor r_t^j \rfloor}{2} q_t^{\lfloor r_t^j \rfloor - 2}\epsilon^2 + \cdots \pm \epsilon^{\lfloor r_t^j \rfloor} \right] - \sum_{j=1}^{m} w_{t+1}^j \epsilon^{\lfloor r_{t+1}^j \rfloor}$$
$$\ge \sum_{j=1}^{m} [M_1^j \epsilon - M_2^j (\epsilon^2 + \epsilon^3 + \cdots + \epsilon^{\max(\lfloor r_t^j \rfloor, \lfloor r_{t+1}^j \rfloor)})]$$
$$\ge \sum_{j=1}^{m}\left[M_1^j\epsilon - M_2^j \frac{\epsilon^2}{1-\epsilon}\right]$$

where $M_1^j = w_t^j \lfloor r_t^j \rfloor q_t^{\lfloor r_t^j \rfloor - 1} (>0)$ and $M_2^j = \max_{2 \le i \le \lfloor r_t^j \rfloor} \left\{ w_t^j \binom{\lfloor r_t^j \rfloor}{i} q_t^{\lfloor r_t^j \rfloor - i} + w_{t+1}^j \right\} (>0)$ for $j = 1, \ldots, m$. The last expression is positive for $\epsilon < \min_j\{(M_1^j)/(M_1^j + M_2^j)\}$.

Hence, by choosing any positive ϵ less than $\min\{I_t, q_t, 1, (h_t/\overline{w}_{t+1})^{1/(\overline{r}_{t+1}-1)}, \min_j\{(M_1^j)/(M_1^j+M_2^j)\}\}$ the subplan $Q_{u,v}$ can always be improved and, hence, it is not optimal. Having established the result for t and $t+1$, it can be extended to the remaining periods similarly by induction over periods $t+2$ to v. $\square$

Proof of Theorem 4.

(i) We construct the optimal production plan of $\overline{P}_{1,t}$ by changing $Q^*_{1,t}$. First we construct the new G-class production subplan for $[t_k, t]$, $\overline{Q}_{t_k,t}$. If $\sum_{n=1}^{m} r_{t_k-1}^n w_{t_k-1}^n q^*_{t_k-1}(r_{t_k-1}^n - 1) + h_{t_k-1} \ge \sum_{n=1}^{m} r_{t_k}^n w_{t_k}^n \overline{q}_{t_k}^{(r_{t_k}^n-1)}$, then $\overline{Q}^*_{1,t}$ is optimal with i=k. Otherwise, we can improve it by transferring some portion of the total demand of $[t_k, t]$, say ϵ, to the period t_k. However, this results in $\sum_{n=1}^{m} r_{t_k-2}^n w_{t_k-2}^n q^*_{t_{k-2}} r_{t_k-2}^n - 1 + h_{t_k-2} \le r_{t_k-1}^n w_{t_k-1}^n (q^*_{t_k-1}+\epsilon)^{r_{t_k-1}^n - 1}$ which implies that it can be improved again. By a similar argument, transferring some positive portion of x to all periods within $[t_{k-1}, t_k - 1]$ gives a better objective cost. We continue this procedure in a backward way until we reach to period t_i such that after augmenting $q^*_{t_i}$ to $\overline{q}_{t_i} = q^*_{t_i} + \epsilon_{t_i}$ we still have the optimality inequality of Theorem 3, that is, $\sum_{n=1}^{m} r_{t_i-1}^n w_{t_i-1}^n q^*_{t_i} r_{t_i-1}^n - 1 + h_{t_i-1} \ge \sum_{n=1}^{m} r_{t_i}^n w_{t_i}^n \overline{q}_{t_i}^{r_{t_i}^n - 1}$ and no further improvement can be made. Hence optimal augmentation of the old production quantities gives a new G-class production subplan and it stops in one of the t_i, $i \in \{1, \ldots, k\}$.

(ii) If d_{t+1} is such that $\sum_{n=1}^{m} r_t^n w_t^n q_t^* r_t^n - 1 + h_t \ge \sum_{n=1}^{m} r_{t+1}^n w_{t+1}^n d_{t+1}^{r_{t+1}^n - 1}$, the given plan $Q^*_{1,t+1}$ is optimal with $i = k+1$ and the new period is itself a generation. Otherwise, the rest of the proof follows from part (i) by considering $x = d_{t+1} - \tilde{d}$ where $\tilde{d}$ solves $\sum_{n=1}^{m} r_t^n w_t^n q_t^* r_t^n - 1 + h_t = \sum_{n=1}^{m} r_{t+1}^n w_{t+1}^n (\tilde{d})^{r_{t+1}^n - 1}$. $\square$

Proof of Lemma 3. The proof rests on standard optimization techniques. We first solve for the extrema that satisfy the first order condition. We have $(d/dQ)TC(Q) = TC'(Q) = -KD/Q^2 + h/2 - [\sum_{n=1}^{m} w^n Q^{r^n}]D/Q^2 + [\sum_{n=1}^{m} r^n w^n Q^{r^n-1}]D/Q = 0 \Rightarrow K + \sum_{n=1}^{m}(1-r^n)w^n Q^{*r^n} = (h/2D)Q^{*2}$

$$\Rightarrow KD = \frac{h}{2}Q^{*2} - D\sum_{n=1}^{m}(1-r^n)w^n Q^{*r^n} \tag{12}$$

where Q^* denote(s) the extremum(a). Considering the second derivative evaluated at the extrema, we have

$$\frac{d}{dQ}TC'(Q) = TC''(Q) = 2KD/Q^3 + \left[\sum_{n=1}^{m}(r^n-2)(r^n-1)w^n Q^{r^n-3}\right]D$$

where at the critical point(s) of TC since (12) holds, we have

$$TC''(Q^*) = hQ^{*2} - 2D\left[\sum_{n=1}^{m}(1-r^n)w^n Q^{*r^n}\right]$$

$$
\begin{aligned}
&+D\left[\sum_{n=1}^{m}(r^n-2)(r^n-1)w^nQ^{r^n-3}\right]\\
&=hQ^{*2}+D\left[\sum_{n=1}^{m}(-2+2r^n+r^{n2}-3r^n+2)w^nQ^{*r^n}\right]\\
&=hQ^{*2}+D\left[\sum_{n=1}^{m}r^n(r^n-1)w^nQ^{*r^n}\right]>0.
\end{aligned}
$$

So the function is convex in its all critical points and therefore, is a quasi-convex. Therefore, it has a unique extremum point. □

Algorithm 1. Forward DP and heuristic algorithms

```
Require: Problem instance, Algorithm name
Ensure: Computes solution for P_{1,T}
function FORWARD_zeroSETUP()
  for i=1 to T do
    f_i = f_{i-1} + g_{i,i}
    RegP(i) ← i                        ▹ RegP(i) = first period of the generation including i
    clrp ← i                           ▹ clrp = candidate last regeneration point
    while clrp > 1 and feasible_class-G_exists do
      if class-G sequence for the segment (clrp, i) is feasible then
        f_i = f_{clrp-1} + g_{clrp,i}
        RegP(i) = clrp
      else
        feasible_class-G_exists ← False
      end if
      clrp − 1 ← RegP(clrp − 1)
    end while
  end for
end function

function H_1()
  while i ≤ T do
    l ← −1
    repeat
      l ← l+1
    until ( g^1_{i,i+l} / l ≤ g^1_{i,i+l+1} / (l+1) ) or (i+l ≥ T)
    q_i ← D_{i,i+l}
    i ← i+l+1
  end while
end function

function H_2()
  Q̃* ← arg min_{Q̃>0} TC(Q̃)
  Ds ← D_{1,T}: Total remaining demand
  for i=1 to T do
    update net demand: d̂
    q_i = min(Ds, max(d̂_i, Q̃*))
    Ds ← Ds − q_i
  end for
end function

function H_3()
  f_0 ← 0
  for i=1 to T do
    f_i = f_{i-1} + g^3_{i,i},  q_i ← D_{i,i}
    for j=i−1 down to 1 do
      if f_{j-1} + g^3_{j,i} < f_i then f_i ← f_{j-1} + g^3_{j,i},  q_i ← D_{i,j}
    end for
  end for
end function

function INCLUSIONEXCLUSION (j, i)
for n=1 to 3 do                        ▹ method selection : n
```

```
      for nn=1 to Nrand do                                          ▹ number of randomizations : Nrand
        if n=1,2 then
          S={j,…,i}
        else
          S←{j}
        end if
        Q_{j,i} ← class-G sequence induced by S
        if Q_{j,i} is infeasible then
          Q_{j,i} ← (d_j, d_{j+1}, …, d_i)
        end if
        g_{ji} ← cost of Q_{j,i}
        repeat
          Improvement_Observation ← False
          switch n
            case 1 klow=j, kup=i                                    ▹ greedy exclusion heuristic
            case 2 klow=Random index from {j,…,i}\S, kup=klow       ▹ random exclusion heuristic
            case 3 klow=Random index from S, kup=klow               ▹ random inclusion heuristic
          for k=klow to kup do
            if k∉S and n=1,2 S'=S ∪ {k}
            else S'=S\{k}
            end if
            Q'_{j,i} ← class-G sequence induced by S'
            g'_{j,i} ← cost of Q'_{j,i}
            if Q'_{j,i} is feasible and g'_{j,i} < g_{ji} then
              candidate ← j, Improvement_Observation ← True
              g_{ji} ← g'_{j,i}, Q_{j,i} ← Q'_{j,i}
            end if
          end for
          S ← S'
        until Improvement_Observation=False
      end for
    end for
  end function

  function G-Heuristics_and_ForwardDP (algName)
    f_0 ← 0                                                         ▹ algName ∈ {H_4, H_5, H_6, ForwardDP}
    if algName = H_k then                                           ▹ H_k ∈ {H_4, H_5, H_6}
      Call H_{k−3}                                                  ▹ H_{k−3} ∈ {H_1, H_2, H_3}
      Ω ← production period indices
    end if
    for i=1 to T do
      f_i = f_{i−1} + g_{i,i}
      for j=i−1 down to 1 do
        if algName ≠ ForwardDP
          S ← indices in [j,i] ∩ Ω
          Q_{j,i} ← class-G sequence induced by S
        else Call InclusionExclusion(j,i)
        end if
        if Q_{j,i} is infeasible then g_{j,i} ← ∞ else g_{j,i} ← cost of Q_{j,i}
         if f_{j−1} + g_{j,i} < f_i then f_i ← f_{j−1} + g_{j,i}
      end for
    end for
  end function
```

References

Aggarwal, A., Park, J.K., 1993. Improved algorithms for economic lot size problems. Oper. Res. 41 (3), 549–571.

Azaron, A., Tang, O., Tavakkoli-Moghaddam, R., 2009. Dynamic lot sizing problem with continuous-time Markovian production cost. Int. J. Prod. Econ. 120 (2), 607–612.

Blackburn, J.D., Kunreuther, H., 1974. Planning horizons for the dynamic lot size model with backlogging. Manag. Sci. 21 (3), 251–255.

Brahimi, N., Dauzere-Peres, S., Najid, N.M., Nordli, A., 2006. Single item lot sizing problems. Eur. J. Oper. Res. 168 (1), 1–16.

Buschkühl, L., Sahling, F., Helber, S., Tempelmeier, H., 2010. Dynamic capacitated lot-sizing problems: a classification and review of solution approaches. OR Spectr. 32 (2), 231–261.

Erlenkotter, D., 1989. An early classic misplaced: Ford W. Harris's economic order quantity model of 1915. Manag. Sci. 35 (7), 898–900.

Federgruen, A., Tzur, M., 1991. A simple forward algorithm to solve general dynamic lot sizing models with n periods in 0 (n log n) or 0 (n) time. Manag. Sci. 37 (8), 909–925.

Florian, M., Klein, M., 1971. Deterministic production planning with concave costs and capacity constraints. Manag. Sci. 18 (1), 12–20.
Ganas, I., Papachristos, S., 2005. The single-product lot-sizing problem with constant parameters and backlogging: exact results, a new solution, and all parameter stability regions. Oper. Res. 53 (1), 170–176.
Garcia-Laguna, J., San-Jose, L.A., Cardenas-Barron, L.E., Sicilia, J., 2010. The integrality of the lot size in the basic EOQ and EPQ models: applications to other production-inventory models. Appl. Math. Comput. 216 (5), 1660–1672.
Haldi, J., Whitcomb, D., 1967. Economies of scale in industrial plants. J. Polit. Econ. 75 (4), 373–385.
Harris, F.W., 1913. How many parts to make at once. Factory, Mag. Manag. 10 (2) 135–136, 152.
Hatirli, S.A., Ozkan, B., Fert, C., 2006. Energy inputs and crop yield relationship in greenhouse tomato production. Renew. Energy 31 (4), 427–438.
Heathfield, D., Wibe, S., 1987. An Introduction to Cost and Production Functions. Humanities Press International, NJ.
Heck, M., Schmidt, G., 2010. Lot-size planning with non-linear cost functions supporting environmental sustainability. In: Networked Digital Technologies, pp. 1–6.
Jans, R., Degraeve, Z., 2007. Meta-heuristics for dynamic lot sizing: a review and comparison of solution approaches. Eur. J. Oper. Res. 177 (3), 1855–1875.
Jans, R., Degraeve, Z., 2008. Modeling industrial lot sizing problems: a review. Int. J. Prod. Res. 46 (6), 1619–1643.
Jeunet, J., Jonard, N., 2000. Measuring the performance of lot-sizing techniques in uncertain environments. Int. J. Prod. Econ. 64 (1–3), 197–208.
Karimi, B., Fatemi Ghomi, S., Wilson, J., 2003. The capacitated lot sizing problem: a review of models and algorithms. Omega 31 (5), 365–378.
Khachian, L.G., 1979. A polynomial algorithm in linear programming. Soviet Math. Dokl. (English transl.) 20, 191–194.
Kogan, K., Tapiero, C.S., 2009. Optimal co-investment in supply chain infrastructure. Eur. J. Oper. Res. 192 (1), 265–276.
Lau, L.J., Tamura, S., 1972. Economies of scale, technical progress, and the nonhomothetic Leontief production function: an application to the Japanese petrochemical processing industry. J. Polit. Econ. 80 (6), 1167–1187.
Leontieff, W., 1947. Introduction to a theory of the internal structure of functional relationships. Econometrica, 361–373.
Lundin, R.A., Morton, T.E., 1975. Planning horizons for the dynamic lot size model: Zabel vs. protective procedures and computational results. Oper. Res. 23 (4), 711–734.
Manne, A.S., Veinott, A.F., 1967. Optimal plant size with arbitrary increasing time paths of the demand.
Meyer, R., 1977. A class of nonlinear integer programs solvable by a single linear program. SIAM J. Control Optimiz. 15 (6), 935–946.
Nakamura, S., 1990. A nonhomothetic generalized Leontief cost function based on pooled data. Rev. Econ. Stat., 649–656.
Okhrin, I., Richter, K., 2011. The linear dynamic lot size problem with minimum order quantity. Int. J. Prod. Econ. 133 (2), 688–693.
Ozaki, I., 1976. Industrial structure and employment. Dev. Econ. 14 (4), 341–365.
Shadbegian, R.J., Gray, W.B., 2005. Pollution abatement expenditures and plant-level productivity: a production function approach. Ecol. Econ. 54 (2), 196–208.
Silver, E.A., Meal, H.A., 1973. A heuristic for selecting lot size requirements for the case of a deterministic time varying demand rate and discrete opportunities for replenishment. Prod. Inventory Manag. J. 14 (2), 64–74.
Simpson, N.C., 2001. Questioning the relative virtues of dynamic lot sizing rules. Comput. Oper. Res. 28 (9), 899–914.
Toy, A.O., Berk, E., 2013. Dynamic lot sizing for a warm/cold process: heuristics and insights. Int. J. Prod. Econ. 145 (1), 53–66.
Veinott, A.F., 1964. Production planning with convex costs: a parametric study. Manag. Sci. 10 (3), 441–460.
Vollmann, T.E., Berry, W.L., Whybark, D.C., 1997. Manufacturing Planning and Control Systems, 4th ed. Irwin/McGraw-Hill, Homewood, IL.
Wagelmans, A., Van Hoesel, S., Kolen, A., 1992. Economic lot sizing: an $o(n \log n)$ algorithm that runs in linear time in the Wagner–Whitin case. Oper. Res. 40 (1 (Suppl. 1)), S145–S156.
Wagner, H.M., Whitin, T.M., 1958. Dynamic version of the economic lot size model. Manag. Sci. 5 (1), 89–96.
Zangwill, W.I., 1966. A deterministic multi-period production scheduling model with backlogging. Manag. Sci. 13 (1), 105–119.