Dynamic expediting of an urgent order with uncertain progress

Luca Bertazzi
Univerity of Brescia, Italy
Riccardo Mogre
Durham University, UK

November 8, 2017

Abstract

A supplier manages an urgent order with uncertain progress for which her client has set a deadline for its completion. The supplier observes in real time the order progress and chooses dynamically the effort level, such as manpower level, to expedite the order. Her problem is to identify expediting policies to minimise her expected cost, given by two costs in trade-off with each other: an effort level cost and a one-time penalty cost for late completion of the order. We formulate this problem using a discrete stochastic dynamic programming framework and obtain an optimal expediting policy for it.

By conducting a worst-case analysis, we show that decreasing the level of flexibility, interpreted as the supplier's ability to update effort levels more often, may lead to a large increase in the costs of managing the order. We refine the problem formulation by modelling the case in which the supplier takes into account the negative effects of late order completion on her client using a penalty cost charged every period the order is delayed. We find an optimal policy for this case, by solving two dynamic programming problems sequentially. For both problems, we show that in presence of certain assumptions, there is an optimal expediting policy for which effort levels are non-increasing in the order progress. Finally, using a simulation study based on a car seat assembly case, we compare the performance of the two policies.

Keywords: supply chain management; tracking technologies; supply chain risk management; stochastic dynamic programming; worst-case analysis

1 Introduction

Suppliers often receive urgent orders which are critical for their business. For example, automotive manufacturers, faced with sudden surges in consumer demand, may in turn place urgent orders of components to their suppliers (Reed and Simon, 2010). Similarly, aerospace manufacturers may anticipate the delivery of planes in an effort to please airline companies, relying on their suppliers to urgently deliver the components for production (Wilhelm, 2016). Both these examples show that suppliers, by managing urgent orders effectively, can positively influence future business relations with their clients. Suppliers face the problem of meeting deadlines for orders whose progress is uncertain, because of disruptions and productivity problems. To manage such orders effectively, suppliers can monitor them and dynamically choose whether to allocate extra resources to accelerate their progress. Auto-id technologies can facilitate this procedure, known as dynamic expediting (Selko, 2008). In practice, dynamic expediting decisions are often dealt without codified rules because the time pressure leads to unstructured decision-making processes.

Managing urgent orders effectively is critical to suppliers' business. We argue that suppliers can achieve effective urgent-order management by making dynamic expediting decisions in a structured manner. For this reason, we aim at providing guidance on how to manage urgent orders using dynamic expediting. In this article we consider orders arising from urgent requests made by clients, as in Chevalier et al. (2015). However, some of the guidance we provide could also be useful when scheduling problems or production uncertainties trigger regular orders to become urgent.

We study the novel problem of devising dynamic expediting policies for an urgent order. Dynamic expediting has been mostly analysed by studies investigating inventory policies in multi-stage supply chains. Our approach differs from these studies in two main aspects. First, these studies consider expediting as an action to safeguard companies from demand variability, but generally not from lead-time variability, as we do in here. Second, these studies do not take into account deadlines for order deliveries, an element which we consider in our study. In Section 2 we present previous contributions that are relevant for this study.

In our problem a supplier processes an urgent order, such as an indivisible lot of components. This lot is indivisible by nature or because of a managerial choice; for example the supplier keeps all components together to facilitate their tracing. The supplier already made lot-sizing and reorder decisions and only chooses order expediting policies. The order progress is not constant, but random, because of disruptions, efficiency, and yield problems. The supplier, with the support of tracking technologies, monitors the order progress periodically. Based on this information, she decides the effort to invest in the order. A personal communication between the authors and the general manager of sales and marketing for Sanden International Europe highlights that suppliers in the automotive industry commonly employ this practice to manage urgent orders (Coulson, 2016).

More effort invested corresponds to higher costs and, in expectation, to higher order progress accomplished. For chiefly manual tasks, we find reasonable to assume that the supplier can change dynamically the effort invested in the order. The supplier can increase the effort level by diverting personnel working on other orders. Therefore, the supplier can estimate the effort cost as the opportunity cost of these workers, implicitly assessing the effect of expediting on other orders. We assume that the supplier manages the urgent order with priority over all other orders because it is business-critical. For this reason, we argue that the supplier makes expediting decisions on the urgent order independently from the progress of other orders. Therefore, we do not explicitly model the effect of expediting decisions on the other orders. Finally, if the urgent order has not finished by a deadline agreed with a client, the supplier faces an actual or estimated penalty cost. Our problem description is general and has applications in various contexts. We illustrate how to operationalise it for two cases rather different from each other: the assembly of car seats and the production of heavy machinery. Car seat assembly takes place on mostly manual production lines. In this context, the order is a lot of product and its progress could be measured by the number of units assembled from that lot. When calculating the order progress, we could ignore the work-in-progress, which is negligible because of the short processing times. Operators scan completed products using auto-id technologies, automatically updating the count of units assembled. The car seat assembler could increase the effort level by assigning more than one assembly line to the production of the order. In this context, the effort level could be the number of lines assigned to the production of the lot. For heavy machinery assembly, the order could be a single machine. Assuming that the production process could be divided in tasks with equal work content, the order progress could be measured by the number of tasks completed. Operators scan the machine with auto-id technologies each time a task is completed, updating the count of the tasks performed. In this context, the effort level could be the number of operators assigned to perform a task.

In Section 3, we formally formulate this problem. Clients often do not specify penalty costs for urgent orders. For this reason, in Section 5 we refine the problem formulation by modelling the case in which the supplier takes into account in her penalty function the negative effects of late order completion on her client. We solve both problems and investigate how the sequence of effort levels changes in the order progress, proving some interesting monotonicity results for the optimal policies. In this study, we assume that the supplier tracks the order progress and updates the effort level at each time period. However, for some activities and tasks, especially if these are partially automated, it may not be possible for the suppliers to be so flexible in changing the effort levels. Cognisant of this limitation of our study, in Section 4 we discuss the value of flexibility, which we define as the organisation's ability to change effort levels more often. In Section 6, we introduce a simulation study based on a car seat assembly case. In this study, we compare the performance of the policy obtained in Section 3 against the one of the policy obtained in Section 5. In Section 7 we include the concluding remarks. All the proofs of the results are in the Appendix.

2 Relevant work

This research contributes to the literature on expediting decisions and especially to those studies analysing how expediting could mitigate operational risks arising from lead-time variability. We review previous research on expediting in make-to stock and make-to-order systems. Then, we discuss how our model relates to previous studies.

In make-to-stock systems, the use of expediting upgrades units in transit to faster delivery modes, triggering urgent deliveries. The majority of make-to-stock literature on urgent deliveries focuses on emergency orders, additional orders usually placed later than regular orders to avoid inventory shortages. Most of these studies analyze the use of express orders in response to demand variability, with fewer contributions investigating their use in response to lead-time variability, such as the study of Kouvelis and Li (2008). They assume that real-time order progress information is available to decision maker, a setting similar to the one we study.

In the literature on expediting in make-to-stock systems, we note that some of these studies consider expediting as a modelling assumption to allow all orders to be delivered on time, as in Huggins and Olsen (2003). These contributions are not closely related to our research. Instead, we review those studies explicitly determining expediting policies based on the current information on the supply chain, including inventory levels, demand or supply conditions. These studies determine expediting policies in single-stage, two-stage and multi-stage supply chains.

Bookbinder and Çakanyildirim (1999), Gallego et al. (2007) and Chiang (2010) analyse expediting decisions in single-stage make-to-stock systems operating according to an order-quantity/reorder point-type policy. Bookbinder and Çakanyildirim (1999) obtain optimal conditions for a system with constant demand and stochastic lead times, made endogenous by considering expediting factors. Gallego et al. (2007), for a system facing random demand, prove that, at optimality, inventory managers should expedite orders according to a threshold policy, assuming that the time between order reception and customer demand can be modelled using an Erlang distribution. Chiang (2010) also identifies an optimal threshold policy for a system in which allocating part of an outstanding order to a faster, non-zero lead-time expediting option is allowed. Fu et al. (2013) consider a news-vendor problem with many expediting options.

Minner et al. (2003) provide approximate solutions for a two-stage inventory system with one central depot and many retailers facing Mixed-Erlang demand. The depot could expedite outstanding orders in case it has insufficient stocks to satisfy retailers. However, the expediting outcome is not known in advance and is modeled after a stochastic process taking into account the number of orders expedited and the age of the orders.

Studies on expediting decisions in multi-stage supply chains introduce the idea of dynamic management of orders, in which order progess is checked at each time period and each stage and expediting decisions are made accordingly. Lawson and Porteus (2000) analyse a periodic-review multi-stage make-to-stock system facing random demand. They assume regular lead times between adjoining stages to be one week and expediting to be instant.

They make decisions from the most upstream stage to the one closest to consumer demand, allowing units to be shipped instantly many stages downstream in the supply chain. They assume expediting costs to be linear in the number of stages across which units are shipped. They show top-down base stock inventory policies to be optimal. In such policies, decisions of a particular stage are constrained by the decisions made in the stages farther up in the supply chain. Muharremoglu and Tsitsiklis (2003) and Kim et al. (2007) extend Lawson and Porteus model and find optimal base-stock inventory policies. Muharremoglu and Tsitsiklis (2003) allow expediting cost to be super-modular in the number of stages across which units are shipped. Kim et al. (2007) consider regular lead times stochastic but not independent across stages. In their model, at each time, they choose a single progression pattern at random for all orders in the supply chain. They assume expediting to be instant. Berling and Martínez-de Albeníz (2011) study expediting decisions in a continuous-time, continuous-stage multi-stage supply chain facing random demand. They model expediting choices as a continuous variable, the speed at which each unit is shipped down the supply chain. They show that the optimal speed of a given unit accelerates upstream and slows downstream in the supply chain.

In summary, only studies on expediting in multi-stage supply chains represent the process of dynamically managing lead-time of outstanding orders. They consider a single expediting option, except for Berling and Martínez-de Albeníz (2011). These studies do not take into account deadlines for order deliveries. They assume expediting costs as linear or super-modular in the number of stages. They consider regular lead-times as deterministic, except for Kim et al. (2007), and expedited lead-times as instantaneous, except for Berling and Martínez-de Albeníz (2011).

Studies analysing make-to-order systems consider expediting a faster but more expensive choice to regular processing options. These studies mostly use expediting as a modelling assumption necessary to allow all the orders to be delivered on time; they are not usually concerned with identifying policies for the dynamic lead-time management of orders. Nevertheless, these studies are more intricate in other dimensions than the contributions on inventory management, for example taking into account the process of setting deadlines for order deliveries. Plambeck and Ward (2008) and Çelik and Maglaras (2008) model queueing systems in presence of lead-time quotations and expediting. For an assemble-to-order system, Plambeck and Ward (2008) study how to set nominal component production rates, quote prices and maximum lead-times for products, and then, dynamically sequence orders for assembly and expedite components. Their proposed policy maximises the infinite-horizon expected discounted profit when the system processes a high volume of customers and expediting costs are large. Çelik and Maglaras (2008) consider a profit-maximising make-to-order manufacturer offering many products to a market of price- and delay-sensitive users. The manufacturer jointly uses dynamic pricing and lead-time quotation controls to manage demand in combination with sequencing and expediting. Arslan et al. (2001) focus more into depth on identifying expediting policies for make-to-order systems. They consider deterministic non-zero expediting order processing times and stochastic regular order processing times. They show that, for many continuous and discrete-time queues, the optimal policy takes the form of a (s, S) policy, where the system manager should expedite s units when the number of units backlogged reaches S.

Compared with studies on expediting in make-to-stock systems, we assume that orders have already been placed and we focus on the real-time management of remaining make-to-order activities of a single urgent order. Therefore, we assume constant demand and stochastic lead times, a setting similar to the one analysed for emergency orders by Kouvelis and Li (2008). This setting allows us to relax restrictive assumptions on costs, expediting options and lead times assumed by previous studies analysing expediting in make-to-stock systems. We consider many expediting options, whose outcome is random, and allow for general forms of expediting and penalty costs. Moreover, we focus on the case of lead-time management of an order in presence of a deadline. Studies on expediting in make-to-stock systems do not consider deadlines for orders, which are taken into account in studies on expediting in make-to-order systems. However, these studies do not study expediting as dynamic management of order lead times, as we do in here.

3 Formulating and solving the dynamic expediting problem

A supplier aims at minimising the expected cost of managing an urgent order. She is required to complete S tasks before the deadline D. Each task has the same work content. After tracking the order progress at each time $t = 0, 1, \ldots, D - 1$, she decides dynamically whether to expedite it.

We provide a stochastic finite horizon discrete dynamic programming formulation for the problem. The state of the system x_t is the order progress, that is, the number of tasks completed at time t. In the initial state, denoted with 0, the supplier has not made any progress on the order yet. In the final state, denoted with S, the supplier has completed all the work on the order. We denote with $\mathcal{S} = \{0, 1, \ldots, S\}$ the set of states. The supplier observes the progress x_t at time t and decides the effort level k to employ from t to $t+1$. The effort level could be the number of workers assigned to perform a task. The supplier chooses the effort level from the set of available levels $\mathcal{K} = \{1, \ldots, K\}$. If $k = 0$ the supplier does not deploy any resources. For each state $0 \leq x_t < S$ and time instant $0 \leq t \leq D-1$ the set of feasible effort levels is $\mathcal{E}(x_t) = \mathcal{K}$, while $\mathcal{E}(S) = \{0\}$. We note that it is possible to extend this formulation by having a different set of feasible effort levels for each order progress.

At each time $t = 0, 1, \ldots, D-1$, the supplier pays an effort cost c_k, increasing in the effort levels $k \in \mathcal{K}$. Only after the deadline, if the supplier has not finished working on the order, she further bears an actual or estimated penalty cost, given by the function $U(x_D)$. The penalty cost function could include two elements: a fixed cost U_f and a variable cost, a linear or non-linear function of a unit cost U_v and the state of the system at the deadline x_D. A possible formulation of the penalty cost function is $U(S) = 0$ and $U(x_D) = U_f + U_v(S - x_D)$ for $x_D < S$ in which the unit cost U_v is multiplied by the uncompleted tasks at the deadline. In Section 5 we extend our formulation by refining the estimation of the penalty cost function for a supplier taking into account the effects of late order completion on her client.

The discrete random variable w_k is the order progress in the time period when the supplier employs a positive effort level k. The uncertain order progress is caused by disruptions and productivity problems in the supply process. w_k has expected value $E(w_k)$ and is defined as follows: $Pr\{w_k = 0\} > 0$ and for $d = 0, 1, \ldots,$ $0 \leq Pr\{w_k = d\} < 1$ and $\sum_d Pr\{w_k = d\} = 1$. The faster the effort level employed, the greater should be the order progress toward its delivery. Formally, if $w_{k'}$ and $w_{k''}$ are two discrete random variables with $k' < k''$, we assume that $w_{k''}$ first-order stochastically dominates $w_{k'}$, that is $\sum_{d=0}^{\tau} Pr\{w_{k'} = d\} \geq \sum_{d=0}^{\tau} Pr\{w_{k''} = d\}$ for any $\tau = 0, 1, \ldots$. If $w_{k''}$ first-order stochastically dominates $w_{k'}$, then $E(w_{k''}) \geq E(w_{k'})$ (Ross, 1996, 405–406). The transition probabilities from the state $i \in \mathcal{S}$ to the state $j \in \mathcal{S}$ when the supplier uses the effort level k are $P_{i,j}(k) = 0$ when $j < i$; $P_{i,j}(k) = Pr\{w_k = j - i\}$ when $j \geq i, j \neq S$ and $P_{i,S}(k) = 1 - \sum_{j=i}^{S-1} P_{i,j}(k)$.

We note that $P_{S,S}(k) = 1$.

At each time t, based on the order progress X_t, the supplier chooses an effort level $K_t \in \mathcal{E}(X_t)$. At time $t+1$, based on the realization W_{kt} of the random variable w_k at time t, the supplier observes the updated order progress X_{t+1}, given by the transition equation $X_{t+1} = \min\{X_t + W_{kt}, S\}$.

We denote as $\pi = \{\mu_0, \mu_1, \ldots, \mu_{D-1}\}$ a time-dependent policy for the problem in which μ_t is a function assigning an effort level to each state x_t. If the initial state is 0, the expected cost $J_\pi(0)$ over the finite-time horizon corresponding to the policy π is as follows:

$$J_\pi(0) = E\{\sum_{t=0}^{D-1} c_{\mu_t(X_t)} + U(X_D)\}. \tag{1}$$

The expected optimal cost for the problem is $J_{\pi^*}(0)$, with π^* the policy minimising (1).

We solve the problem using the finite-horizon dynamic programming algorithm based on Bertsekas (2005, 23) and described as follows.

1. $J_D^*(x_D) = U(x_D) \qquad x_D = 0, 1, \ldots, S$

2. Proceed backward in time by calculating the optimal cost-to-go $J_t^*(x_t)$ for $t = D-1, D-2, \ldots, 0$ from the system of equations as follows:

$$J_t^*(x_t) = \min_{k \in \mathcal{K}} \left[c_k + \sum_{x_{t+1}=x_t}^{S-1} P_{x_t, x_{t+1}}(k) J_{t+1}^*(x_{t+1}) \right] \qquad x_t = 0, 1, \ldots, S. \tag{2}$$

The optimal policy π^* determines the optimal effort level for each time t and state i, denoted with $k_t^*(i)$. We investigate how the optimal effort levels change in the state. Intuitively, for a given time period, the optimal effort levels should be non-increasing in the order progress. If a supplier follows this intuition, she employs faster effort levels earlier in the supply process, followed by slower effort levels when the order is near completion. This strategy not only safeguards the supplier from the risk for the supply process 'to get stuck' in its earlier stages but also avoids assigning fast effort levels when approaching the terminal state because these effort levels will be underutilised. We formalise this intuition in Theorem 1, in which we use ρ to denote a given value of the state j.

Theorem 1. *If $\sum_{j=\rho}^{\infty} P_{i,j}(k)$ is a sub-additive function on $\mathcal{S} \times \mathcal{K}$ for all $\rho \in \mathcal{S}$ and $U(x_D)$ is a non-increasing function in x_D, then there is an optimal policy such that $k_t^*(i)$ is non-increasing in i for $t = 1, \ldots, D-1$.*

We note that Theorem 1 requires an additional assumption on the random variable modelling the order progress. Example 1 in the Appendix shows that if $\sum_{j=\rho}^{\infty} P_{i,j}(k)$ is not a sub-additive function on $\mathcal{S} \times \mathcal{K}$ for all $\rho \in \mathcal{S}$, Theorem 1 does not always hold.

Next, we investigate how the optimal effort levels change as t increases. Intuitively, for a given state, the optimal effort levels should be non-decreasing in time. This intuition means that if the supplier does not make any progress on the order in the current time period, she employs an effort level in the next time period greater or equal than the one in the current time period. However, Example 2 in the Appendix shows that this property does not always hold. Instead, we show in Theorem 2 that, for a large enough penalty function, $J_t^*(s)$ is non-decreasing in time for any state $s \in \mathcal{S}$. This property means that if the supplier does not make any progress on the order in the current time period, she bears an optimal cost in the next time period greater or equal than the one in the current time period.

Theorem 2. *If, $J_{D-1}^*(s) \leq U(s)$ for all states $s \in \mathcal{S}$, then $J_t^*(s) \leq J_{t+1}^*(s)$ for any state $s \in \mathcal{S}$ and any time $t = 1, \ldots, D-1$.*

From (2) $J_{D-1}^*(s)$ can be written as: $\min_{k \in \mathcal{K}} \left[c_k + \sum_{x_D=s}^{S-1} P_{s,x_D}(k) U(x_D) \right]$. We choose the penalty function $U(x_D)$ to be non-increasing in x_D. This implies that $\sum_{x_D=s}^{S-1} P_{s,x_D}(k) U(x_D) \leq U(s)$ for any effort level $k \in \mathcal{K}$. Therefore, the condition in Theorem 2 is met if the effort costs c_k are sufficiently low relative to the penalty costs. For some activities and tasks, changing effort levels may require a certain setup time. We highlight that it is possible to extend the formulation of the problem discussed in this Section to the case with setup times by using time lags Bertsekas (2005, 35–36). In the next Section we discuss a related problem, with the aim of assessing the value of the supplier ability to change effort levels more often.

4 Assessing the value of flexibility

In Section 3 we assumed that the supplier tracks the order progress at each time period $t = 0, 1, \ldots D-1$ and modifies the effort level accordingly, by applying an optimal policy. However, for some activities and tasks, especially if these are partially automated, the supplier may not be able to change the effort level at each time period. This lack of flexibility leads to increase in the costs of managing the order. In this Section we aim at assessing the cost increase from the case in which the supplier can change the effort level at each time period (Case 1) against the case in which the supplier can change the effort level every $\eta \geq 2$ time periods (Case 2). We do so by conducting a worst-case analysis.

We denote with $k_t^*(x_t)$ the optimal effort level for each state $x_t \in \mathcal{S}$ at each time period $t \in T$. For each trajectory $\tau = \{d_0^\tau, d_1^\tau, \ldots, ...d_D^\tau\}$, in which d_t^τ is the realisation of the order progress at time t, the terminal state x_D is equal to $\min\{\sum_{t\in T} d_t^\tau, S\}$. Let $T' = \{0, 1, \ldots, D-1\}$ and T'' the subset of T' containing the time periods multiple of η. The cost in Case 1 is $z_\tau^1 = \sum_{t\in T'} c_{k_t^*(x_t)} + U_D(x_D)$ and the cost z_τ^2 in Case 2 is not greater than $\eta \sum_{t\in T''} c_{k_t^*(x_t)} + U_D(x_D)$, that is, the cost of the feasible solution obtained by applying the effort level $k_t^*(x_t)$ at time periods $t, t+1, \ldots, t+\eta-1$ for each $t \in T''$. We denote with $\mathcal{T}$ the set of trajectories used to evaluate the application of the optimal effort levels in Case 1 and in Case 2. The corresponding costs in Case 1 and Case 2 are $J_{\mathcal{T}}^1(0) = \frac{1}{|\mathcal{T}|}\sum_{\tau\in\mathcal{T}} z_\tau^1$ and $J_{\mathcal{T}}^2(0) = \frac{1}{|\mathcal{T}|}\sum_{\tau\in\mathcal{T}} z_\tau^2$, respectively. For $|\mathcal{T}| \to \infty$, these costs tend to the corresponding expected costs. Theorem 3 shows the worst-case performance of $J_{\mathcal{T}}^2(0)$ against $J_{\mathcal{T}}^1(0)$.

Theorem 3. *For each set of trajectories $\mathcal{T}$, $J_{\mathcal{T}}^2(0)$ is at most η times $J_{\mathcal{T}}^1(0)$ and the bound is tight for $\eta = 2$.*

Theorem 3 shows that decreasing the level of flexibility potentially leads to a large increase in the costs of managing the order.

5 Modelling the negative effects on clients

Contracts between suppliers and clients usually establish agreed monetary penalties for the late completion of regular orders. However, monetary penalties are far less common for urgent orders because they arise from short-notice client requests. In absence of monetary penalties for urgent orders, suppliers, when deciding how to allocate resources to expedite order progress, could take into account how late order completions disrupt their clients. In the cases from automotive and aerospace industries described in Section 1, urgent orders include components necessary for the assembly of products sold by the clients, such as cars and planes. For each time period the urgent order of components has not finished, the client cannot start product assembly and bears estimated or actual losses. Often these are potential revenues from product sales, which would have taken place if urgent orders had been delivered by the supplier and product assembly had been completed by the client.

In this Section, we provide an advanced formulation of the dynamic expediting problem in which the supplier takes into account her client's estimated or actual losses in presence of late order completion. Moreover, the policy obtained using this formulation identifies which efforts the supplier should use after the deadline. The policy obtained using the formulation in Section 3 does not provide such information.

We still consider the case of a supplier tracking the progress of an urgent order at time t and deciding dynamically the effort level to allocate to the order. The supplier aims at minimising the sum of effort level costs and penalty costs that she has to bear if she has not finished all tasks S, necessary to complete the order, before a deadline D. In Section 3 we formulated a problem in which a supplier dynamically allocates resources to order processing and faces a penalty cost $U(x_D)$ if the order has not finished by a deadline D. In here, we extend our model by explicitly considering how the supplier should allocate resources to accelerate the order progress also after the deadline D. In this latter formulation, the supplier faces penalty costs for each time unit the order has not finished. We formulate this problem using two integrated stochastic discrete dynamic programming models. Before time D we adapt the finite-horizon model described in Section 3. From time D the system is stationary, that is, time-independent, because the penalty cost is no longer a function of the deadline D. For this reason after D we can model the system using an infinite-horizon model. The infinite-horizon problem ends when all tasks are completed, corresponding to the state S. The presence of this cost-free termination state makes our problem a stochastic shortest path problem (Bertsekas, 2005, 403).

We provide the stochastic discrete dynamic programming formulation of the infinite-horizon problem. The state of the system is the order progress x, which is independent of time because the system is stationary. The state space is $\mathcal{S}_{x_D} = \{x \in \mathcal{S} : x \geq x_D\}$ with the state at the deadline x_D being the initial state of the infinite-horizon model. The supplier observes the progress x at time t and decides the effort level k to employ between t and $t+1$ from the set of available levels $\mathcal{K} = \{0, 1, \ldots, K\}$ with $k = 0$ corresponding to no resources deployed. For each state $0 \leq x < S$ the set of feasible effort levels is $\mathcal{E}(x) = \mathcal{K}$, while $\mathcal{E}(S) = 0$. At each time t, the supplier pays an effort cost c_k, increasing in the effort level $k \in \mathcal{K}$. Every time period the supply process has not finished, the supplier bears a fixed penalty cost u. The penalty per unit of time of delay u could be used to compensate the loss of production at the client's site. The definitions of the discrete random variables, transition equations and transitions probabilities are identical to those in Section 3. The immediate cost $g(i,k,j)$ of the infinite-horizon model for the transition from the state $i \in \mathcal{S}_{x_D}$ to the state $j \in \mathcal{S}_{x_D}$ when using the effort level k is defined as follows. If $i \neq S$ and $j \neq S$, then $g(i,k,j) = c_k + u$. If $i \neq S$ and $j = S$, then $g(i,k,S) = c_k$. If $i = S$, then $k = 0$, $j = S$ and $g(S,0,S) = 0$. The expected cost per stage is $\bar{g}(i,k) = \sum_{j \in \mathcal{S}_{x_D}} P_{i,j}(k) g(i,k,j)$, for $i \neq S$ and 0 otherwise. If μ is a function assigning an effort level to each state x, the constant sequence $\mu, \mu, \ldots, \mu$ is a stationary (time-independent) policy for the problem. We denote this policy with μ. For an order starting with progress x_D at the deadline D, we denote with $Y_\mu(x_D)$ its expected cost over the infinite-time horizon corresponding to the policy μ.

We obtain $Y_\mu(x_D)$ by taking the limit as the number of periods T goes to infinity.

$$Y_\mu(x_D) = \lim_{T \to \infty} E\{\sum_{t=D}^{T-1} \bar{g}(X_t, \mu(X_t)) | X_D = x_D\}. \tag{3}$$

We denote as $Y^*_\infty(x_D)$ the expected optimal cost. If $x_D = S$, then $Y^*_\infty(x_D) = 0$. If $x_D < S$, then $Y^*_\infty(x_D) = Y_{\mu^*}(x_D)$ with μ^* being the stationary policy minimising (3).

The discrete dynamic programming formulation of the finite-horizon problem is the finite-horizon model described in Section 3 in which the penalty cost is the optimal cost-to-go of the infinite-horizon model, that is $U(x_D) = Y^*_\infty(x_D)$.

We solve the infinite-horizon problem using the value iteration algorithm based on Bertsekas (2005, 413–414) and described as follows:

1. Set $b := 0$ and $Y_b(i)$ to a random value, $i = 1, 2, \ldots, S-1$.

2. Compute $Y_{b+1}(i) = \min_{k \in \mathcal{K}} \left[\bar{g}(i,k) + \sum_{j=i}^{S-1} P_{i,j}(k) Y_b(j)\right]$
for $i = 0, 1, \ldots, S-1$. Let $k^*_{b+1}(i)$ be the corresponding effort level.

3. If $Y_{b+1}(i) = Y_b(i)$, $\forall i = 0, 1, \ldots, S-1$, set $\mu^*(i) := k^*_{b+1}(i)$, for $i = 0, 1, \ldots, S-1$. Otherwise, go to step 2.

We solve the finite-horizon problem using the dynamic programming algorithm described in Section 3.

For the infinite-horizon problem we obtain a stronger version of Theorem 1.

Theorem 4. *There is an optimal policy for the infinite-horizon model such that, for each state $i' \in \mathcal{S}$ and $i \in \mathcal{S}$ with $x_D \leq i' < i < S$, $k^*_{i'} \geq k^*_i$.*

Theorem 4 states that, after the deadline D, the supplier could expedite the order optimally by using effort levels non-increasing in the order progress.

We note that the assumption of $U(X_D)$ being non-increasing in x_D in Theorem 1 is always true when $U(x_D) = Y^*_\infty(x_D)$ because of the definition of the cost-to-go $Y^*_\infty(x_D)$.

6 Simulation study

In Sections 3 and 5 we showed how to derive optimal policies to expedite an urgent order with uncertain progress by solving a dynamic programming problem. In Section 3 we obtained an optimal expediting policy for a finite-horizon model that ends at the deadline. We refer to this optimal policy as 'base-case policy'. In Section 5 we obtained an optimal expediting policy solving in sequence an infinite-horizon model and a finite-horizon model. This infinite-horizon model starts at the deadline and ends when the supplier has finished working on the order. We refer to this optimal policy as 'advanced policy'.

In this Section, we provide a simulation study based on a car seat assembly case. We use this study to compare the performance of the base-case policy against the one of the advanced policy.

Car seat assembly takes place on mostly manual assembly lines. Each assembly line usually includes 15 to 20 stations. Operators assemble one seat at a time. A typical assembly process is described as follows. First, operators assemble the upper cushioning, the headrest, the seat warmer system, the airbags and the metal spring framework. Then, operators assemble the lower cushioning, the height-adjustment mechanisms, the electrical connectors and the reclining springs. Finally, operators perform functional tests and iron the seats. Tsou and Chen (2005) provided a detailed description of the car assembly process.

We consider the case in which a plant manager has received an urgent order of 3 300 seats by an important client and has agreed to have it completed in 40 hours. In this context, the final state S is 3 300 and the deadline D is 40 hours. The plant includes five lines. In this context, the effort level k is the number of lines employed by the plant manager. The plant manager monitors the order progress every hour and decides how many lines to allocate to the order. The hourly order progress depends on the number of lines employed and follows a discrete uniform probability distribution. For one, two, three, four and five lines employed the support of the probability distributions of car seats per hour are as follows: $[64, 94]$, $[130, 190]$, $[195, 285]$, $[255, 385]$, and $[320, 480]$, respectively. We note that the corresponding discrete random variables satisfy the requirements detailed in Section 3. If the plant manager employs an assembly line for the production of an urgent order, the production of regular orders is delayed. This negative effect is more than linear in the number of lines used because using more lines could cause much longer delays to the production of regular orders. For example, when the plant manager employs all five lines to assemble the urgent orders, the production of regular orders has to be completely stopped. In this context, the negative effect is the effort cost c_k. For one, two, three, four and five lines employed the hourly effort costs are 60, 140, 240, 360, and 500 euros, respectively. Finally, if the plant manager does not complete the production of seats by the deadline, her client will not be able to start the car production process. The plant manager has estimated that this delay can cost her client euro 175/hour, which is the unit penalty cost u.

To assess the increase in cost ensuing from using the base-case policy instead of the advanced policy, we used a simulation that we implemented in C++. For this case, applying the advanced policy leads to an average cost over 1 000 replications of 2 583.6 euros. The base-case policy requires the definition of the penalty function $U(x_D)$. Let $\bar{d}_k$ be the expected progress associated to the effort level k. Assuming that the probability of choosing each effort level is the same and taking expectation over the effort levels, a good approximation for $Y^*_\infty(x_D)$ is $U(x_D) = E[(c_k + u)(S - x_D)/\bar{d}_k]$. Using this newly defined penalty function, we apply the base-case policy to our case and obtain an average cost over 1 000 replications of 2 590.1 euros. This approximation for our penalty function implicitly assumes that after the deadline the supplier employs each effort level k with the same probability until the order has finished. For each effort level k, in expectation, the time it takes for the order to be completed could be approximated by $(S - x_D)/\bar{d}_k$.

D [$hours$]	c_k [$euro$]	u [$euro\backslash hour$]	Average cost (advanced) [$euro$]	Cost increase (base case) [$euro$]	Interval (cost increase) [euro]
20	Low	175	2 950.2	14.4	[11.0, 17.7]
20	Low	1 225	2 957.0	30.5	[27.3, 33.8]
20	High	175	4 424.8	17.8	[13.0, 22.6]
20	High	1 225	4 436.4	38.5	[32.8, 44.3]
40	Low	175	2 583.6	6.5	[4.3, 8.8]
40	Low	1 225	2 584.4	12.3	[10.0, 14.6]
40	High	175	3 874.2	7.7	[4.2, 11.2]
40	High	1 225	3 877.7	15.4	[12.1, 18.7]

Table 1: Simulation comparison with the order progress following a discrete uniform distribution and confidence intervals at the 90% level of significance.

To obtain the expected penalty for each effort level k we multiply this time interval by the cost that the supplier bears at each period after the deadline before the order has finished, which is $c_k + u$.

To explore the sensitivity of the cost-increase to variation in the parameters, we conduct various simulation experiments. In our experiments the hourly order progress for each number of lines employed can follow a discrete uniform distribution or a Poisson distribution with the same mean. The deadline D can be 20 or 40 hours. The effort costs for one, two, three, four, and five lines employed can take low or high values. Low effort costs are 60, 140, 240, 360, and 500 euros, respectively. High effort costs are 90, 210, 360, 540, and 750 euros, respectively. The penalty costs can be 175 or 1 225 euros per hour. For each experiment and for each policy we conducted 1 000 replications. We used the common random number technique in the replications across the two policies. For each experiment we computed the difference between the costs accomplished using the base-case policy and those accomplished using the advanced policy. For this random variable we computed a paired–t confidence interval at the 90% of significance (Law, 2007, 552–554). We show the results for the order progress following discrete uniform and Poisson distributions in Tables 1 and 2, respectively.

Experiments employing the advanced policy lead to less frequent changes in the effort levels than those employing the base-case policy. If the order progress follows a discrete uniform distribution, experiments employing the advanced policy and the base-case policy lead to average changes in the effort levels of 2.5 and 4.0, respectively. If the order progress follows a Poisson distribution, experiments employing the advanced policy and the base-case

D	c_k	u	Average cost (advanced)	Cost increase (base case)	Interval (cost increase)
[$hours$]	[$euro$]	[$euro\backslash hour$]	[$euro$]	[$euro$]	[euro]
20	Low	175	2 936.2	15.8	[13.3, 18.3]
20	Low	1 225	2 948.6	24.5	[19.3, 29.7]
20	High	175	4 401.3	21.1	[17.4, 24.8]
20	High	1 225	4 416.2	36.3	[31.0, 41.6]
40	Low	175	2 565.2	3.4	[1.6, 5.3]
40	Low	1 225	2 565.5	11.2	[9.5, 13.0]
40	High	175	3 848.2	2.5	[-0.1, 5.2]
40	High	1 225	3 847.4	16.7	[14.2, 19.2]

Table 2: Simulation comparison with the order progress following a Poisson distribution and confidence intervals at the 90% level of significance.

policy lead to average changes in the effort levels of 2.0 and 3.25, respectively. On average, 37% of the changes in the effort levels take place early in the simulation, that is, before $D/2$.

The base-case policy is worse than the advanced policy in every experiment. However, using the base-case policy instead of the advanced policy leads to small cost-to-go increases overall. This good results are also due to the initial simulation experiments we run to identify values for the base-case policy penalty function leading to a good approximation of the advanced policy penalty function. Experiments with the order progress following a Poisson distribution lead to larger average cost increases than those with the order progress following a discrete uniform distribution. The advanced policy works better than the base-case policy when the system is ‘under-stress’, that is, for shorter deadlines. Larger effort costs and penalty costs lead to larger average cost increases. As we consider comparable increases of effort and penalty costs, we note that the variability in penalty cost has a larger effect on the cost-to-go than the variability in the effort cost when the value of the penalty cost is small. The advanced policy leads to less frequent effort changes compared to the base-case policy. For this reason, the assembly plant manager could implement the advanced policy more easily than the base-case policy.

In summary, if the assembly plant manager calibrates the penalty function of the base-case policy using the penalty function of the advanced policy, she could use the ‘base-case policy’ with only small increases in the costs-to-go. However, the base-case policy does not provide information on which effort to use after the deadline. If the plant manager wants to know this information, she has to use the advanced policy.

7 Conclusions

Urgent orders are critical to clients. Therefore, suppliers managing these orders effectively can relevantly improve their customer relations. Using tracking technologies suppliers can monitor orders in real time, allowing them to identify in a timely manner delays and disruptions that could affect the order progress as in Mogre et al. (2014). Suppliers can use this real-time information to dynamically choose whether to allocate extra resources, such as manpower, to expedite orders with the aim of meeting tight deadlines set by clients.

In here, we studied the case of a supplier managing an urgent order with uncertain progress for which her client set a deadline for its completion. Her problem is to identify effort levels minimising her cost, given by two costs in trade-off with each other: an effort level cost, such as the cost of manpower, and a penalty cost charged for late order completion.

In Section 3, we formulated this problem using discrete stochastic dynamic programming and solved it by adapting the finite-horizon dynamic programming algorithm. In Theorem 2, we showed that, for a large enough penalty cost function, the optimal cost-to-go is non-decreasing in time. This property means that no order progress in the current time period leads the supplier to face an optimal cost-to-go in the next time period at least as large as the one in the current time period. In Theorem 1, we showed that, given certain assumptions on the order progress, there is an optimal policy such that effort levels are non-increasing in the order progress. Theorem 1 implies that a supplier employing an optimal policy is likely to use faster and more expensive effort levels earlier in the supply process, followed by slower and cheaper effort levels when the order is near completion.

Clients often do not specify penalty costs for urgent orders. For this reason, in Section 5 we extended the formulation of our problem to the case in which the supplier estimates her own penalty cost function by assessing the negative effects of a late order completion on her client. We found the optimal solution for this case by sequentially solving two dynamic programming problems, a finite-horizon problem modelling the order progress before the deadline, and an infinite-horizon problem modelling the order progress after the deadline. We solved the infinite-horizon problem using the value iteration algorithm. For the infinite-horizon problem we obtained a stronger version of Theorem 1. Theorem 4 states that after the deadline is elapsed, the supplier could manage the order using an optimal policy in which the effort levels are non-increasing in the order progress. Berling and Martínez-de Albeníz (2011) proved a similar monotonicity result for expediting policies in a sequential supply chain. They showed that the optimal speed of a given order should accelerate upstream and slow downstream.

In Section 6 we introduced a simulation study based on a car seat assembly case. This study showed that the performance of the policy obtained in Section 3, called 'base-case' policy is similar to the one of the policy obtained in Section 5, called 'advanced' policy, if the manager of the car seat assembly plant uses as penalty function a good approximation of the penalty cost function of the advanced policy. Nevertheless, the advanced policy, by leading to fewer changes in the effort levels compared to the base-case policy, could be easier to implement in practice.

In this study, we assumed that the supplier tracks the order progress and updates the effort level at each time period. However, for some activities and tasks, especially if these are partially automated, it may not be possible for the supplier to be so flexible in updating effort levels. To address this limitation of our study, in Section 4 we discussed the value of flexibility, which is the suppliers ability to update effort levels more often. By conducting a worst-case case analysis, in Theorem 3 we showed that decreasing the level of flexibility may lead to a large increase in the costs of managing the order.

Future work could extend our modelling framework to the dynamic management of expediting in projects, a setting similar to the one studied by Bregman (2009). This future study would face the complication that activities in projects are not only sequential, but arranged in networks.

Acknowledgement

The authors would like to thank Stefan Minner, Tava Olsen, Terry Williams, two anonymous referees and the editor for many constructive suggestions on earlier drafts of this work. The authors would also like thank Alastair Coulson for his kind participation in this research project.

References

Arslan, H., Ayhan, H., Olsen, T., 2001. Analytic models for when and how to expedite in make-to-order systems. IIE Transactions 33 (11), 1019–1029.

Berling, P., Martínez-de Albeníz, V., 2011. Optimal expediting decisions in a continuous-stage serial supply chain. Working paper, IESE.

Bertsekas, D., 2005. Dynamic programming and optimal control, 3rd Edition. Vol. 1. Nashua, NH, USA: Athena Scientific.

Bookbinder, J., Çakanyildirim, M., 1999. Random lead times and expedited orders in (q, r) inventory systems. European Journal of Operational Research 115 (2), 300–313.

Bregman, R., 2009. A heuristic procedure for solving the dynamic probabilistic project expediting problem. European Journal of Operational Research 192 (1), 125–137.

Çelik, S., Maglaras, C., 2008. Dynamic pricing and lead-time quotation for a multiclass make-to-order queue. Management Science 54 (6), 1132–1146.

Chevalier, P., Lamas, A., Lu, L., Milnar, T., 2015. Revenue management for operations with urgent orders. European Journal of Operational Research 240 (2), 476–487.

Chiang, C., 2010. An order expediting policy for continuous review systems with manufacturing lead-time. European Journal of Operational Research 203 (2), 526–531.

Coulson, A., 2016. Discussion on dynamic expediting of urgent orders. [email] (Personal communication, 1 December 2016).

Fu, K., Xu, J., Miao, Z., 2013. Newsvendor with multiple options of expediting. European Journal of Operational Research 226 (1), 94–99.

Gallego, G., Jin, Y., Muriel, A., Zhang, G., Yildiz, V., 2007. Optimal ordering policies with convertible lead times. European Journal of Operational Research 176 (2), 892–910.

Huggins, E., Olsen, T., 2003. Supply chain management with guaranteed delivery. Management Science 49 (9), 1154–1167.

Kim, C., Klabjan, D., Simchi-Levi, D., 2007. Optimal expediting policies for an inventory system with stochastic lead time under radio frequency identification. Working paper, MIT.

Kouvelis, P., Li, J., 2008. Flexible backup supply and the management of lead-time uncertainty. Production and Operations Management 17 (2), 184–199.

Law, A., 2007. Simulation modeling & analysis, 4th Edition. New York, NY, USA: McGraw-Hill.

Lawson, D., Porteus, E., 2000. Multistage inventory management with expediting. Operations Research 48 (6), 878–893.

Minner, S., Diks, E., De kok, A., 2003. A two-echelon inventory system with supply lead time flexibility. IIE Transactions 35 (2), 117–129.

Mogre, R., Wong, C., Lalwani, C., 2014. Mitigating supply and production uncertainties with dynamic scheduling using real-time transport information. International Journal of Production Research 52 (17), 5223–5235.

Muharremoglu, A., Tsitsiklis, J., 2003. Dynamic leadtime management in supply chains. Working paper, MIT.

Plambeck, E., Ward, A., 2008. Optimal control of a high-volume assemble-to-order system with maximum leadtime quotation and expediting. Queueing Systems 60 (1-2), 1–69.

Puterman, M., 2005. Markov decision processes. Discrete stochastic dynamic programming, 2nd Edition. New York, NY, USA: John Wiley & Sons.

Reed, J., Simon, B., 2010. Global car sectors rebound runs out of gas. *Financial Times*, [online]. Available at: http://www.ft.com/cms/s/0/d9dba6d0-885f-11df-aade-00144feabdc0.html[Accessed 22 April 2016].

Ross, S., 1996. Stochastic processes, 2nd Edition. New York, NY, USA: John Wiley & Sons.

Selko, A., 2008. Rfid used to expedite production of vehicles for u.s. defense department. *Industry Week*, [online]. Available at:

http://www.industryweek.com/planning-amp-forecasting/rfid-used-expedite-production-vehicles-us-defense-department[Accessed 22 April 2016].

Tsou, J.-C., Chen, J.-M., 2005. Case study: quality improvement model in a car seat assembly line. Production Planning & Control 16 (7), 681–690.

Wilhelm, S., 2016. Boeing could accelerate 777x delivery to please emirates. *Puget Sound Business Journal*, [online]. Available at: http://www.bizjournals.com/seattle/news/2016/03/11/boeing-could-accelerate-777x-delivery-to-please.html[Accessed 22 April 2016].

Appendix

Proof of Theorem 1. We prove the result by using the sufficient condition provided in Puterman (2005, 107–108). In particular, the five conditions required in Theorem 4.7.4 are satisfied as:

1. c_k is constant in i for all k, so it is non-increasing in i for all k.

2. $\sum_{j=\rho}^{\infty} P_{i,j}(k) = \sum_{j=\rho}^{\infty} Pr\{w_k = j - i\} = \sum_{d=\rho-i}^{\infty} Pr\{w_k = d\}$ is non-decreasing in i for all ρ and k.

3. c_k is a sub-additive function on $\mathcal{K}$ (any function with one variable is both a super-additive and a sub-additive function by definition).

4. $\sum_{j=i+\rho}^{\infty} P_{i,j}(k)$ is a sub-additive function on $\mathcal{S} \times \mathcal{K}$ for all $\rho \in \mathcal{S}$ by assumption.

5. $U(x_D)$ is a non-increasing function in x_D by assumption.

$\square$

Example 1. *This example shows that, if $\sum_{j=\rho}^{\infty} P_{i,j}(k)$ is not a sub-additive function on $\mathcal{S} \times \mathcal{K}$ for all $\rho \in \mathcal{S}$, the property of Theorem 1 does not always hold. Let $S = 5$ and $K = 2$, with $c_1 = 10$ and $c_2 = 50$, $t = D - 1$, $Y_D^*(0) = 200$, $Y_D^*(1) = 160$, $Y_D^*(2) = 130$, $Y_D^*(3) = 80$, $Y_D^*(4) = 10$ and $Y_D^*(5) = 0$. Suppose that $Pr\{w_1 = 1\} = 1$, while $Pr\{w_2 = 2\} = 1$. Set $s^- = 0$, $s^+ = 2$, $k^- = 1$, $k^+ = 2$ and $\rho = 4$. Since*

$$\sum_{j=\rho}^{\infty}[P_{s^-,j}(k^-)+P_{s^+,j}(k^+)] = \sum_{d=4}^{\infty} Pr\{w_1 = d\}+\sum_{d=2}^{\infty} Pr\{w_2 = d\} = 0+1 > 0+0 =$$

$$= \sum_{d=4}^{\infty} Pr\{w_2 = d\} + \sum_{d=2}^{\infty} Pr\{w_1 = d\} = \sum_{j=\rho}^{\infty}[P_{s^-,j}(k^+) + P_{s^+,j}(k^-)],$$

then $\sum_{j=\rho}^{\infty} P_{i,j}(k)$ is not a sub-additive function on $\mathcal{S} \times \mathcal{K}$ for all $\rho \in \mathcal{S}$. The optimal effort level at state 0 is $k_{D-1}^*(0) = 1$, as $Y_{D-1}^*(0) = \min\{c_1 + Y_D^*(1), c_2 + Y_D^*(2)\} = \min\{170, 180\} = 170$. Instead, the optimal effort level at state 1 is $k_{D-1}^*(1) = 2$, as $Y_{D-1}^*(1) = \min\{c_1 + Y_D^*(2), c_2 + Y_D^*(3)\} = \min\{140, 130\} = 130$.

Example 2. *This example shows that $k_t^*(s)$ can be greater than $k_{t+1}^*(s)$. Let $S = 5$ and $K = 2$, with $c_1 = 26$ and $c_2 = 50$, $J_D^*(0) = 5$, $J_D^*(1) = 4$, $J_D^*(2) = 3$, $J_D^*(3) = 2$, $J_D^*(4) = 1$ and $J_D^*(5) = 0$. Suppose that $Pr\{w_1 = 1\} = 1$, while $Pr\{w_2 = 2\} = 1$. Let us focus on state $s = 3$. At time $t = D-1$, $k_{D-1}^*(3) = 1$, as $J_{D-1}^*(3) = \min\{c_1 + J_D^*(4), c_2 + J_D^*(5)\} = \min\{27, 50\} = 27$. Moreover, $k_{D-1}^*(4) = 1$ as $J_{D-1}^*(4) = \min\{c_1 + J_D^*(5), c_2 + J_D^*(5)\} = \min\{26, 50\} = 26$. At time $t = D-2$, $k_{D-2}^*(3) = 2$ as $J_{D-2}^*(3) = \min\{c_1 + J_{D-1}^*(4), c_2 + J_{D-1}^*(5)\} = \min\{52, 50\} = 50$.*

Proof of Theorem 2. We obtain the proof by backward induction. The statement is true for $t = D-1$ by assumption. Let us assume that it is true at time $t+1$, that is $Y_{t+1}^*(s) \leq Y_{t+2}^*(s)$. Then, $Y_t^*(s) = \min_k[c_k + \sum_{d=0}^{S-s} Pr\{w_k = d\}Y_{t+1}^*(s+d)] \leq \min_k[c_k + \sum_{d=0}^{S-s} Pr\{w_k = d\}Y_{t+2}^*(s+d)] = Y_{t+1}^*(s)$. □

Proof of Theorem 3. For each trajectory $\tau \in \mathcal{T}$, $z_\tau^1 = \sum_{t\in T'} c_{k_t^*(x_t)} + U_D(x_D) \geq \sum_{t\in T''} c_{k_t^*(x_t)} + U_D(x_D)$. Moreover, $z_\tau^2 \leq \eta \sum_{t\in T''} c_{k_t^*(x_t)} + U_D(x_D) \leq \eta \sum_{t\in T'} c_{k_t^*(x_t)} + U_D(x_D) \leq \eta z_\tau^1$. Hence,

$$J_{\mathcal{T}}^2(0) = \frac{1}{|\mathcal{T}|}\sum_{\tau\in\mathcal{T}} z_\tau^2 \leq \frac{1}{|\mathcal{T}|}\sum_{\tau\in\mathcal{T}} \eta z_\tau^1 = \eta J_{\mathcal{T}}^1(0).$$

The following instance allows us to prove that the bound is tight for $\eta = 2$. Let $S = 4$, $D = 2$ and $K = 2$, with $c_1 = \epsilon << 1$ and $c_2 = 2$, $U_D(0) = 6$, $U_D(1) = 4$, $U_D(2) = 2$, $U_D(3) = \epsilon$ and $U_D(4) = 0$. Suppose that $Pr\{w_1 = 1\} = 1$, while $Pr\{w_2 = 3\} = 1$. Let $\mathcal{T}$ be composed of just one trajectory τ with $d_0^\tau = 3$ and $d_1^\tau = 3$.

First, we compute the optimal policy at time $t = 1$.

$x_1\backslash k$	1	2	$J_1^*(x_1)$	$k_t^*(x_1)$
0	$c_1 + U_D(1) = 4 + \epsilon$	$c_2 + U_D(3) = 2 + \epsilon$	$2 + \epsilon$	2
1	$c_1 + U_D(2) = 2 + \epsilon$	$c_2 + U_D(4) = 2$	2	2
2	$c_1 + U_D(3) = 2\epsilon$	$c_2 + U_D(4) = 2$	2ϵ	1
3	$c_1 + U_D(4) = \epsilon$	$c_2 + U_D(4) = 2$	ϵ	1

Then, we compute the optimal policy at time $t = 0$.

$x_0\backslash k$	1	2	$J_0^*(x_1)$	$k_t^*(x_1)$
0	$c_1 + J_1^*(1) = 2 + \epsilon$	$c_2 + J_1^*(3) = 2 + \epsilon$	$2 + \epsilon$	1/2

Consider now the trajectory τ. In Case 1, $J_{\mathcal{T}}^1(0) = 2 + \epsilon$. In Case 2, if we apply $k_t^*(x_1) = k_t^*(x_2) = 2$, then $J_{\mathcal{T}}^2(0) = 4$. Therefore, $2J_{\mathcal{T}}^1(0) = 2(2 + \epsilon) \to 4 = J_{\mathcal{T}}^2(0)$ for $\epsilon \to 0$.

□

Proof of Theorem 4. Consider any trajectory from X_D to S such that the states i' and i, with $X_D \leq i' < i < S$, are visited. Suppose that the corresponding optimal effort levels $k^*_{i'}$ and k^*_i are such that $k^*_{i'} < k^*_i$. We prove that this policy, referred to as μ_1, is dominated by the one, referred to as μ_2, in which the effort level k^*_i is applied at state i' and $k^*_{i'}$ is applied at state i. The part of the total expected cost concerning the states i' and i is $C_{\mu_1} = c_{k^*_{i'}} + u \sum_{j=i'}^{S-1} P_{i',j}(k^*_{i'}) + c_{k^*_i} + u \sum_{j=i}^{S-1} P_{i,j}(k^*_i)$ in the first policy, while it is $C_{\mu_2} = c_{k^*_i} + u \sum_{j=i'}^{S-1} P_{i',j}(k^*_i) + c_{k^*_{i'}} + u \sum_{j=i}^{S-1} P_{i,j}(k^*_{i'})$ in the second. Since $P_{i',j}(k) = Pr\{W_k = j - i'\}$ for any k and $\sum_{d=0}^{\tau} Pr\{W_{k'} = d\} \geq \sum_{d=0}^{\tau} Pr\{W_{k''} = d\}$ for any $k' < k''$ and for any $\tau = 0, 1, \ldots$, then $\sum_{j=i'}^{S-1} P_{i',j}(k^*_{i'}) - \sum_{j=i}^{S-1} P_{i,j}(k^*_{i'}) = \sum_{d=0}^{S-i'-1} Pr\{W_{k^*_{i'}} = d\} + \sum_{d=0}^{S-i'-2} Pr\{W_{k^*_{i'}} = d\} + \ldots + \sum_{d=0}^{S-i} Pr\{W_{k^*_{i'}} = d\} \geq \sum_{d=0}^{S-i'-1} Pr\{W_{k^*_i} = d\} + \sum_{d=0}^{S-i'-2} Pr\{W_{k^*_i} = d\} + \ldots + \sum_{d=0}^{S-i} Pr\{W_{k^*_i} = d\} = \sum_{j=i'}^{S-1} P_{i',j}(k^*_i) - \sum_{j=i}^{S-1} P_{i,j}(k^*_i)$. Therefore, if the strict inequality holds, then $J_{\mu_1}(0) > J_{\mu_2}(0)$ and therefore μ_1 cannot be optimal. Otherwise, if equality holds, both μ_1 and μ_2 are optimal policies. Therefore, there exists an optimal policy such that $k^*_{i'} \geq k^*_i$.

$\square$