

University of Groningen

## Planning of Combined Make-to-Order and Make-to-Stock Production

Beemsterboer, Bartholomeus Jacobus

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*

Publisher's PDF, also known as Version of record

*Publication date:*

2016

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Beemsterboer, B. J. (2016). *Planning of Combined Make-to-Order and Make-to-Stock Production*. [Thesis fully internal (DIV), University of Groningen]. University of Groningen, SOM research school.

### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

## Chapter 2

# Hybrid MTO-MTS production planning: an explorative study

***Abstract.** Research on the planning and control of combined make-to-order/make-to-stock or hybrid production systems often takes a typical MTO or MTS perspective. We examine the benefits of a hybrid planning approach without priority for either MTO or MTS. We develop a Markov Decision Process model for a two-product hybrid system to determine when to manufacture MTS and MTO products. Contrary to earlier studies with this approach, this study includes a positive lead time for MTO products. We characterize optimal policies and show how decisions should be based on both inventory level and backlog state of MTO products. Especially discriminating between states with and states without backlog of MTO orders is shown to be important in determining whether to increase MTS stock. Savings of up to 65% are achieved compared to policies that prioritize MTS or MTO.*

### 2.1 Introduction

Technological developments in manufacturing systems increasingly allow companies to manufacture different products on the same production facility, as indicated by Iravani et al. (2012), Altendorfer & Minner (2014), and others. Also more hybrid production systems have emerged, combining make-to-order (MTO) and make-to-stock (MTS) production. Low-valued, standardized products with regular demand will often be made to stock, allowing demand to be satisfied instantly. For high-valued and/or customized products with irregular demand, stocking can be expensive or even impossible and these are typically produced to order (see e.g. Van Donk (2001)).

Planning of a hybrid production system is not straightforward, because planning and control decisions are significantly different for MTO and MTS production. MTO products must be manufactured before a prespecified due date and production planning and control focuses on the timing and/or sequencing of the operations. MTS production aims at prevention of stock-outs, whilst limiting the inventory holding costs. Therefore, performance criteria in MTS production planning are usually cost-based.

The planning and control of production systems have received considerable attention in the literature. However, as Soman et al. (2004) indicate, most research in this field assumes either pure MTS or pure MTO. Research on the planning and control of hybrid production systems is limited, and often still takes the perspective of either the MTO or the MTS products as a starting point. For instance, Chang et al. (2003) and Wu et al. (2008) modify existing planning methods designed for MTO products by letting a separate MTS process generate 'MTS-orders', and Soman et al. (2006) modify existing planning methods for MTS products by replacing the run-out time of MTS products by the remaining lead time allowance of MTO products. Carr & Duenyas (2000) and Iravani et al. (2012) consider contract manufacturers who produce on an MTS basis for contracted and hence more important customers, and on an MTO basis via other sales channels. They analyze basic two-product models in which the lead time allowance of MTO is not included, and use a Markov Decision Process (MDP) to determine whether MTO or MTS should be produced and whether new MTO orders should be accepted and rejected, based on the MTS inventory level and the amount of MTO orders. Despite the latter studies providing more insight into optimal hybrid policies, we notice a lack of an approach that fully does right to MTS and MTO characteristics. We discuss the literature on the planning of hybrid production systems more extensively in Section 2.2.

In this paper, we investigate the benefits of a hybrid planning method for hybrid production systems, i.e. a planning method that takes the state of both product types into account in deciding which product to manufacture. The main difference with earlier contributions is that our study is not driven by either an MTO or an MTS perspective. Moreover, we allow a positive lead time for MTO production, as typically applies in practice. We follow the modeling approach of Carr & Duenyas (2000) and Iravani et al. (2012) and propose a simple model with two products, one manufactured on an MTO basis and one manufactured on an MTS basis. We use a discrete-time Markov Decision Process to find which product should be manufactured in each period based on the MTS inventory level,

the amount of MTO orders in the system, and the remaining lead time allowance of these orders. We provide insight into the optimal policies that result from the model, where the consideration of lead times turns out to be important, and compare these with two benchmark policies that both prioritize one of the two product types. We show that significant savings can be obtained, which motivates the development of sophisticated planning procedures for real-life hybrid production cases.

The remainder of this paper is organized as follows. Section 2.2 gives an overview of the relevant literature. Section 2.3 formulates the model of the production system as a Markov Decision Process and Section 2.4 describes the resulting optimal policy structure. Section 2.5 compares the performance of our optimal planning procedure with two policies that are based on one of the two production systems. Conclusions and directions for future research are presented in Section 2.6.

## 2.2 Literature review

We provide an overview of the literature related to hybrid production systems and we focus on the planning related contributions, in particular on contributions with MDPs of two-product hybrid production systems. Regarding non-hybrid production systems, we refer the reader to Stevenson et al. (2005) for a review on planning concepts for MTO production and to Sox et al. (1999) and Winands et al. (2011) for reviews on the Stochastic Economic Lot Scheduling Problem, a common approach for MTS production.

A substantial part of the research on hybrid production focuses on the decision which items to produce to stock and which to order. Williams (1984), Arreola-Risa & DeCroix (1998), Rajagopalan (2002), Kaminsky & Kaya (2009), and Altendorfer & Minner (2014) propose quantitative frameworks classifying products as MTO or MTS and determining related production decisions such as base stock levels and/or batch sizes for MTS products. Van Donk (2001) discusses the MTO/MTS classification problem from a qualitative viewpoint for the food processing industry. Soman et al. (2007), Kerkkänen (2007), and Perona et al. (2009) conduct case studies in which they determine the MTO/MTS classification for the companies. Cattani et al. (2010) address this problem from the perspective of factory design. They consider the choice between separate MTO and MTS factories, a hybrid factory and an MTS factory, and only a hybrid factory, and analyze

several trade-offs between these choices. Besides these two scenarios, the authors focus on the decision how many orders for non-contracted customers to accept. Other contributions investigate the benefits and/or performance of hybrid production systems. Nguyen (1998), for instance, approximates fill rates and average inventory levels using a queueing theory approach. Martínez-Olvera (2009) evaluates the influence of hybrid business models on the performance of a supply chain using a system dynamics simulation model. The above studies took a strategic perspective on hybrid production. When it comes to the planning and control, existing approaches often take the perspective of either the MTO or the MTS products as a starting point, as indicated in Section 2.1.

Chang et al. (2003) and Wu et al. (2008) address hybrid production planning from an MTO perspective. They propose planning concepts with order release and dispatching rules. In both papers, the systems generate 'MTS orders' when the work-in-progress is sufficiently low, such that these fill the production capacity that would otherwise be lost due to lack of MTO orders. If an MTS order is generated, the product to be manufactured is selected based on the run-out time. Hadj Youssef et al. (2004) investigate priority rules for MTO and MTS product classes in a simple model. MTS demand is replenished according to a lot-for-lot policy, so each demand results in an 'MTS-order'. The authors compare two priority rules: first come first served (FCFS), and priority for the MTO products, and investigate the impact of model parameters on the performance of the two rules. Sox et al. (1997) compare the performance of three policies: one using FCFS, one selecting products based on the run-out time, and one that is a mixture of the other two. The latter two policies regard MTO products as MTS products with a zero run-out time, so that these have priority over MTS products with a positive inventory level. Chang & Lu (2010) analyze the performance of a hybrid production system that manufactures standard products on an MTS basis and customized products by performing an additional customization operation to finished standard products on an MTO basis. The authors obtain optimal base-stock levels, under the assumption that the customization operations are prioritized above MTS standard product replenishments.

Several other studies propose planning and control approaches from an MTS perspective. Soman et al. (2006) investigate the performance of four existing MTS planning procedures in the context of a hybrid production environment: the method known as the Economic Manufacturing Quantity and the methods proposed by Vergin & Lee (1978), Leachman & Gascon (1988), and Fransoo (1992). Soman et al. include MTO orders in the calculations of these methods by replac-

ing the stock run-out time by the remaining time until the due date of the MTO orders. Federgruen & Katalan (1999) analyze the impact of adding an MTO item to an MTS production system that is controlled by a fixed production sequence and base-stock levels. They compare two possible planning adaptations: prioritizing the MTO item, and inserting the MTO product one or more times in the predetermined production sequence. Zhang et al. (2013) develop a multi-server queuing model of a hybrid production system with two product types (MTO and MTS), two groups of machines (flexible and MTS) and two modes (hybrid, where the flexible machines produce MTO, and MTS, where all machines produce MTS). The model is used to obtain optimal values for the number of flexible machines, the number of machines dedicated to MTS, and the two inventory level thresholds at which the group of flexible machines switch from MTS to MTO and vice versa. Hence, the switching decision is solely based on the MTS inventory level; the MTO queue length is not taken into account.

Carr & Duenyas (2000), Iravani et al. (2012), and Gupta & Wang (2007) consider control policies of hybrid production systems of contract manufacturers, which are compelled to manufacture a part of their products on an MTS basis because long-term contracts with some customers require timely deliveries. The firms also sell through other sales channels, for which they produce on an MTO basis, and for which demands can be rejected. Although the MTO products are usually more profitable, the MTS products are more important from a planning perspective because of the high penalties for late deliveries. Carr & Duenyas and Iravani et al. provide optimal production and admission policies for such systems using simple, two-product models, modeled as Markov Decision Processes. Both ignore MTO lead times and use a backlogging cost proportional to the waiting time in order to penalize late deliveries. Hence, MTO orders are 'late' from the moment they arrive. As the authors indicate, the MTO product in their model can be considered an MTS product with backorders and an order-up-to-level of zero. Their state spaces are two-dimensional, allowing them to characterize optimal policies resulting from the MDPs by two switching curves, each splitting the state space into two areas. One curve indicates for which states to switch between accepting and rejecting the MTO orders and another curve indicates for which states to switch between producing MTS and producing MTO. Besides a few differences that are not relevant for our study, the models of Carr & Duenyas and Iravani et al. differ in their assumption on the consequence of stock-outs of the MTS products. Carr & Duenyas assume that sales are lost, while Iravani et al. assume backlogging. Gupta & Wang consider a more extended discrete-time MDP which includes a lead time allowance. They compare two modes, one

in which all customers are served on an MTO basis, and one in which contracted customers are served on an MTS basis. They obtain optimal results for instances with a lead time allowance of two periods and provide heuristic solutions for instances with longer lead times.

From a modeling perspective, we follow the approach of Carr & Duenyas (2000) and Iravani et al. (2012). However, our study is not motivated by a specific type of companies and we subsequently do not consider MTS more important than MTO. Also, our research is focused on the selection of the product that is manufactured in a period, whereas these authors emphasize the MTO order admission decision and the dynamics between the two decisions. Finally and most importantly, we include an MTO lead time allowance in our model, whereas Carr & Duenyas and Iravani et al. do not.

## 2.3 Model formulation

In this section, we model a hybrid production system. In Subsection 2.3.1, we provide the framework of the model, which we then formulate as a Markov Decision Process in Subsection 2.3.2. A summary of the notations is given in Table 2.1.

### 2.3.1 Production system

We consider a system that manufactures two products, one on an MTO basis and one on an MTS basis, and it may produce only one of these at a time. MTS products are stored in inventory, from which demands are satisfied directly. Demands are lost when no inventory is available. MTO is produced after an order has been placed and should be delivered within a limited lead time. Late MTO orders are penalized, but should still be delivered.

In order to formulate our problem as a Markov Decision Process in the next subsection and in order to keep its formulation accessible, we make the following simplifying assumptions. We use a discrete time framework and we assume that the unit processing times of the products are equal to each other and given as one period. We assume that this processing time includes eventual machine setups. Moreover, MTS production can be interpreted as producing and MTS quantity equivalent to the time required for an average MTO order. In order to ensure finiteness of the state space, we will use bounded demand distributions and we define a maximum MTS inventory level  $I$  and a maximum amount of

**Table 2.1:** Notations

|                             |             |   |
|-----------------------------|-------------|---|
| <b>Demand parameters</b>    | $d_o$       | Average MTO demand per period   |
|                             | $d_s$       | Average MTS demand per period   |
|                             | $d_o^{max}$ | Maximum MTO demand per period   |
|                             | $d_s^{max}$ | Maximum MTS demand per period   |
| <b>Cost parameters</b>      | $q$         | Lateness costs (per unit, per period)                                   |
|                             | $b_o$       | MTO lost sales costs (per unit)   |
|                             | $b_s$       | MTS lost sales costs (per unit)   |
| <b>System parameters</b>    | $K$         | Maximum number of MTO orders  |
|                             | $L$         | MTO lead time allowance   |
| <b>Demand probabilities</b> | $p_o(j)$    | Probability of $j$ MTO demands in a period                              |
|                             | $p_s(j)$    | Probability of $j$ MTS demands in a period                              |
| <b>State space</b>          | $i$         | MTS inventory level   |
|                             | $k_l$       | # MTO orders in the system for $l$ periods<br>( $l = 0, \dots, L - 1$ ) |
|                             | $k_L$       | # late MTO orders   |
| <b>Other</b>                | $k$         | Total number of MTO orders  |

MTO orders in the system  $K$ . We assume that excess MTO demand is lost. The maximum inventory level  $I$  is not treated as a system control parameter, but set large enough so as not to affect the numerical results. The promised lead time for MTO orders is assumed fixed and given by  $L$  periods (not including the period in which the order arrives). With these simplifications, we retain the core of the problem so that we can obtain the insights we aim for.

Demands of MTO and MTS follow two independent truncated Poisson distributions with average demands (per period) denoted by  $d_o$  (MTO) and  $d_s$  (MTS) and maximum demands denoted by  $d_o^{max}$  (MTO) and  $d_s^{max}$  (MTS), respectively. We denote  $p_o(j)$  as the probability of  $j$  MTO demands and  $p_s(j)$  as the probability of  $j$  MTS demands in a period, respectively. By definition,

$$p_o(j) = \begin{cases} \frac{\lambda_o^j e^{-\lambda_o}}{j!} / \sum_{m=0}^{d_o^{max}} \frac{\lambda_o^m e^{-\lambda_o}}{m!} & \text{for } j = 0, \dots, d_o^{max} \\ 0 & \text{otherwise,} \end{cases}$$

and



$$p_s(j) = \begin{cases} \frac{\lambda_s^j e^{-\lambda_s}}{j!} / \sum_{m=0}^{d_s^{max}} \frac{\lambda_s^m e^{-\lambda_s}}{m!} & \text{for } j = 0, \dots, d_s^{max} \\ 0 & \text{otherwise.} \end{cases}$$

The demand rates  $\lambda_o$  and  $\lambda_s$  are selected such that the average demands  $d_o$  and  $d_s$  satisfy

$$d_o = \sum_{j=1}^{d_o^{max}} j p_o(j) \text{ and } d_s = \sum_{j=1}^{d_s^{max}} j p_s(j).$$

The order of events in each period is as follows. First, it is decided whether a unit of MTO, a unit of MTS, or nothing is manufactured. Second, demand occurs and demands are either accepted/satisfied or lost. Third, if a product was manufactured it becomes available; if this is an MTO product, it fulfills an order in the same period.

The objective is to minimize the average costs per period. These costs are composed of per period holding costs of MTS products, lateness costs of MTO products, and lost sales costs. We normalize the MTS holding cost to 1, i.e. the per period holding cost serves as the cost unit. Late MTO orders cost  $q$  per period and lost sales costs are  $b_o$  for MTO and  $b_s$  for MTS, respectively.

### 2.3.2 Markov Decision Process

We model the problem described above as a discrete-time Markov Decision Process. The state space  $\mathcal{S}$  consists of the MTS inventory level  $i$ , the numbers of accepted MTO orders  $k_l$ ,  $l = 0, \dots, L - 1$ , that are in the system for  $l$  periods (not including the period of arrival) and hence have  $L - l$  periods remaining until the due date, and the number of outstanding late orders  $k_L$ . A state  $s \in \mathcal{S}$  is described by the tuple  $s = (i, k_0, \dots, k_L)$ . For convenience, we define the MTO part of the state space,  $(k_0, \dots, k_L)$ , as the *order state* and we will refer to the MTS inventory level  $i$  as the *inventory level*. We further define  $k$  as the total number of MTO orders in the system, i.e.  $k = \sum_{j=0}^L k_j$ . Note that  $0 \leq k_l \leq d_o^{max}$  for  $l = 0, \dots, L - 1$ , and  $k \leq K$ .

The action space is denoted by  $\mathcal{A} = \{1, 2, 3\}$ , where  $a = 1$  indicates production of an MTO order,  $a = 2$  production of an MTS product, and  $a = 3$  idling. Action  $a = 1$  is not admissible when there are no orders at hand, i.e. in states  $(i, 0, \dots, 0)$  for  $i = 0, \dots, I$ . Action  $a = 2$  is not admissible when the inventory is full, i.e. in states  $(I, k_0, \dots, k_L)$  for all allowed combinations of  $k_0, \dots, k_L$ , but this becomes redundant for sufficiently large  $I$ .

The costs in state  $s$  if action  $a$  is taken, denoted as  $c^a(s)$ , are as follows. Writing  $[x]^+$  for  $\max\{x, 0\}$ ,

$$\begin{aligned} c^1(s) &= i + qk_L + b_s \sum_{j=0}^{d_s^{max}} (p_s(j)[j - i]^+) \\ &\quad + b_o \sum_{j=0}^{d_o^{max}} (p_o(j)[k + j - 1 - K]^+), \\ c^2(s) = c^3(s) &= i + qk_L + b_s \sum_{j=0}^{d_s^{max}} (p_s(j)[j - i]^+) \\ &\quad + b_o \sum_{j=0}^{d_o^{max}} (p_o(j)[k + j - K]^+). \end{aligned}$$

The first and second terms in both equations represent the holding and lateness costs, respectively. The third and fourth terms are, respectively, the expected MTS and MTO lost sales costs. For  $a = 1$ , one more MTO order can be accepted than for  $a = 2$  and  $a = 3$ , hence the difference in the expressions. The MTS lost sales costs are the same for all actions because MTS lost sales are based on the amount of inventory at the beginning of the period.

We next discuss the transition probabilities  $\pi^a(s, s')$ , denoting the probability that the state at the end of a period is  $s' = (i', k'_0, \dots, k'_L)$  given that the state at the beginning is  $s = (i, k_0, \dots, k_L)$  and action  $a$  is taken. In this section, we restrict the discussion to transitions to states in which  $i' > 0$  (for  $a = 2$ , we restrict to  $i' > 1$ ) and  $k' < K$ , where  $k'$  is defined as  $k' = \sum_{j=0}^L k'_j$ . For transitions to states with  $i' = 0$  and/or  $k' = K$ , the possibility of lost sales must be taken into account which leads to slightly different expressions, but the dynamics are essentially the

same. We refer to Appendix A for a complete overview of the transition probabilities.

We define  $l^* = \max\{l | k_l \geq 1\}$ ,  $k_l'' = k_l$  for  $l \in \{0, \dots, L\} \setminus l^*$ , and  $k_{l^*}'' = k_{l^*} - 1$ . Then, for the cases considered here, the transition probabilities, that we will explain in what remains of this section, are as follows.

$$\begin{aligned} \pi^1(s, s') &= p_o(k'_0)p_s(i - i') && \text{if } k'_l = k''_{l-1} \text{ for } l = 1, \dots, L - 1 \\ & && \text{and } k'_L = k''_L + k''_{L-1} \end{aligned} \quad (2.1)$$

$$\begin{aligned} \pi^2(s, s') &= p_o(k'_0)p_s(i - i' + 1) && \text{if } k'_l = k_{l-1} \text{ for } l = 1, \dots, L - 1 \\ & && \text{and } k'_L = k_L + k_{L-1} \end{aligned} \quad (2.2)$$

$$\begin{aligned} \pi^3(s, s') &= p_o(k'_0)p_s(i - i') && \text{if } k'_l = k_{l-1} \text{ for } l = 1, \dots, L - 1 \\ & && \text{and } k'_L = k_L + k_{L-1} \end{aligned} \quad (2.3)$$

$$\pi^a(s, s') = 0 \quad \text{otherwise}$$

The transition probabilities in (2.1) - (2.3) are written as the product of the probability of  $k'_0$  MTO demands and the probability of a transition to an inventory level of  $i'$ , which implies  $i - i' + 1$  MTS demands if a unit of MTS is produced (in (2.2)) and  $i - i'$  MTS demands otherwise (in (2.1) and (2.3)). The conditions of (2.1) - (2.3) state that the given probability applies to states for which the MTO orders all have a unit of their lead times deducted (unless they are already late), and that the new state counts one MTO order less if an MTO product is manufactured. The decrease of the lead time allowance is achieved by letting the number of MTO orders having waited for  $l$  periods in the new state take the value of the number of MTO orders having waited for  $l - 1$  periods of the previous state, for  $l = 1, \dots, L - 1$ . The number of late MTO orders increases by the number of MTO orders that had one period remaining. In (2.1), a unit of MTO is produced. As MTO orders differ only in their remaining lead time allowance, it is clearly optimal to produce them in a first come, first served fashion to realize the lowest lateness costs. Hence we satisfy one order with the smallest remaining lead time allowance, which is accounted for by using the order state  $(k''_0, \dots, k''_L)$  instead of  $(k_0, \dots, k_L)$  in the conditions.

We obtain optimal solutions for the MDP using a solver from the package ‘MDPtoolbox’, based on Value Iteration, in the computer program ‘R’.

## 2.4 Policy structure

This section characterizes the structure of optimal policies for the Markov Decision Process defined in the previous section using an illustrative numerical example. We also build on the insights obtained by Carr & Duenyas (2000) and Iravani et al. (2012) who characterize optimal policies for simpler systems without a lead time allowance for MTO orders.

The typical structure of the optimal policy can best be characterized by an illustrative example with a relatively small state space such that all states can be presented in a table.

**Example.** Let  $d_o = d_s = 0.43$ ,  $d_o^{max} = d_s^{max} = 2$ ,  $L = 2$ ,  $K = 4$ ,  $q = 5$ , and  $b_o = b_s = 500$ .

Table 2.2 presents the optimal policy of this example. The resulting policy does not produce MTS for inventory levels above 7 and hence states with inventory levels higher than 8 are transient and can be ignored. Before we discuss this policy, recall from Section 2.2 that Carr & Duenyas (2000) and Iravani et al. (2012) characterize optimal policies for hybrid production systems with one MTO and one MTS product. Their findings include the following. First, both find that the difference between producing MTS and producing MTO can be described by a ‘switching curve’. For lower inventory levels, it is optimal to produce MTS, while for larger inventory levels, MTO is produced. In the model of Carr & Duenyas, the curve also depends on the amount of MTO orders, where an increasing amount of orders leads to a ‘switch’ from MTS to MTO. Second, both find that idling is only optimal in states with no MTO orders.

The term ‘curve’ is inappropriate for our model, because our state space is multi-dimensional contrary to Carr & Duenyas and Iravani et al. who have two-dimensional state spaces. We instead define the notion of *switching levels* for describing optimal policies for our model. A switching level is defined for each MTO order state  $(k_0, \dots, k_L)$  as the lowest inventory level for which the system does not produce MTS. For this level, the system ‘switches’ to MTS when the inventory level drops below it, or ‘switches’ to MTO/idling when the inventory



level reaches it from below. If the switching level is 0, then for each possible inventory level, the system does not produce MTS. In inventory control terms, the switching level could be interpreted as the order-up-to level or base stock level (see Axsäter (2007)), but in our case dependent on the outstanding MTO orders.

Note that for the example considered, the switching level is uniquely described by the total number of MTO orders,  $k$ , and the number of periods until the first MTO order becomes late,  $L - \max\{j : k_j > 0\}$ . This allows us to describe the resulting policy in terms of the switching levels depending on these two values, which we have done in Table 2.3. We remark that, although in our experiments the switching levels are uniquely described by these two values in almost every case considered, there are exceptions.

**Table 2.3:** Optimal policy of the example in terms of switching levels

| # MTO orders | Remaining periods |   |   |   |
|--------------|-------------------|---|---|---|
|              | 3 or more         | 2 | 1 | 0 |
| 0            | 8                 |   |   |   |
| 1            |                   | 4 | 3 | 3 |
| 2            |                   | 3 | 3 | 3 |
| 3            |                   |   | 2 | 2 |
| 4            |                   |   | 0 | 0 |

Table 2.3 leads to the following observations regarding the optimal policy for this example. A striking observation is that there are considerable differences of the switching levels between states with and without MTO orders: if there are no MTO orders, the switch is between MTS and idling; while if there are orders, the switch is between MTS and MTO; and producing MTO is a much more attractive alternative to producing MTS than idling. These observations are made for most parameter settings we considered and Carr & Duenyas (2000) and Iravani et al. (2012) make similar observations. We remark, however, that our sensitivity study in Section 2.5 will reveal some exceptional situations. We further observe from Table 2.3 that the switching level is non-increasing in the amount of MTO orders on hand and that the switching level decreases as MTO orders become closer to their due date. However, even though these observations are intuitive and observed for most parameter settings, they do not apply in some extreme settings.

## 2.5 Numerical comparison

In order to see the advantage of choosing optimally between producing MTO or MTS in every system state, we compare the optimal hybrid solutions with two policies that prioritize either MTS or MTO. We conduct two sets of numerical experiments in which we vary the demand rates and the cost parameters from their base case values (to be discussed), respectively. Before turning to these experiments, however, we first introduce two benchmark policies and briefly discuss the base case.

**MTO Priority.** This policy gives priority to MTO, in the sense that it manufactures an MTO product whenever one is available. If there are none, the system chooses optimally between producing MTS and idling.

**MTS Priority.** This policy gives priority to MTS as soon as the inventory level drops below a certain threshold value. Above this value, the system produces MTO orders when available. Hence, MTS Priority considers a fixed switching level, independent of the order state. This switching level is selected optimally.

Note that policies favoring MTO production have been proposed by, for instance, Sox et al. (1997), Federgruen & Katalan (1999), Hadj Youssef et al. (2004) and Chang & Lu (2010). Policies prioritizing MTS have been used as reference policies for instance by Iravani et al. (2012). Furthermore, using such ‘extreme’ benchmark policies allows us best to determine the maximum potential savings compared to ‘standard’ policies, and under what circumstances those savings are particularly large. For the ease of presentation, we will refer to the optimal and benchmark policies as Hybrid, MTO Priority, and MTS Priority, respectively, from now on.

The reasons for selecting our base case settings are as follows. First, because of the known ‘curse of dimensionality’ of Markov Decision Processes (see e.g. Bellman (1957)), we must control the size of the state space, which is mainly determined by the values of  $L$ ,  $K$  and (especially)  $d_o^{max}$ , whilst retaining key problem characteristics. We select a lead time allowance of  $L = 4$ , such that accepted orders in general do not require immediate production before becoming late. We select  $d_o^{max} = 2$ , being the smallest value for which the MTO demand can still exceed the production capacity in one period, and set  $d_s^{max} = 2$  for consistency. We select  $K = 10$ , so that MTO lost sales cannot occur without a considerable amount of lateness. The base demand values are selected as  $d_o = d_s = 0.45$ , providing a potential utilization of 0.9 in which both product types are equally represented.

We select the base cost parameters as  $q = 5$  and  $b_o = b_s = 500$ . The resulting optimal policies of our experiments focus on avoiding MTO lateness rather than avoiding MTO lost sales, and MTS lost sales are less than 2% for all cases.

When certain parameters are varied in a set of experiments, other parameters keep the above base values. We discuss the first set of experiments, in which we vary the demand parameters, in Subsection 2.5.1. Subsection 2.5.2 discusses the second set, in which we vary the cost parameters.

### 2.5.1 Experiments with varying demand parameters

Because the demand rates affect the results through both the total demand rate and the ratio of the demand rates for the two product types, we select the rates indirectly by selecting values of total demand rate and an MTO/MTS ratio. For the total demand rate, we select the values 0.6, 0.8, 0.9, 0.95 and 1. These values provide a sufficient range in order to observe the effects of an increase in the total demand rate. For the MTO/MTS ratio, we select the levels 1/9, 1/3, 1/1, 3/1, and 9/1, varying from mostly MTS to mostly MTO in a symmetric fashion. The corresponding actual demand rates are shown in Table 2.4. Note that the median values yield the base demand,  $d_o = d_s = 0.45$ .

**Table 2.4:** Experimental design: demand rates (MTO \ MTS)

| MTO/MTS<br>ratio | Total demand rate |           |             |               |           |
|------------------|-------------------|-----------|-------------|---------------|-----------|
|                  | 0.6               | 0.8       | 0.9         | 0.95          | 1         |
| 1/9              | 0.06\0.54         | 0.08\0.72 | 0.09\0.81   | 0.095\0.855   | 0.1\0.9   |
| 1/3              | 0.15\0.45         | 0.2\0.6   | 0.225\0.675 | 0.2375\0.7125 | 0.25\0.75 |
| 1/1              | 0.3\0.3           | 0.4\0.4   | 0.45\0.45   | 0.475\0.475   | 0.5\0.5   |
| 3/1              | 0.45\0.15         | 0.6\0.2   | 0.675\0.225 | 0.7125\0.2375 | 0.75\0.25 |
| 9/1              | 0.54\0.06         | 0.72\0.08 | 0.81\0.09   | 0.855\0.095   | 0.9\0.1   |

Before discussing the performance of the policies, we first consider the switching levels. Recall from Section 2.4 that, for most settings that we consider, the optimal switching level is mostly affected by whether or not there are one or more MTO orders in the system and less by the number of orders and their arrival times. Therefore, Table 2.5 shows the switching level of the order state  $(0, 0, 0, 0, 0)$  with no orders (left part) versus the order state  $(1, 0, 0, 0, 0)$  with 1 order remaining that was received in the most recent period (right part). We indeed observe large differences in the hybrid policy between the switching levels



**Table 2.5:** Switching levels in two order states, depending on the demand rates

| MTO/MTS ratio | Order state (0,0,0,0) |     |     |      |    | Order state (1,0,0,0) |     |     |      |   |
|---------------|-----------------------|-----|-----|------|----|-----------------------|-----|-----|------|---|
|               | Total demand rate     |     |     |      |    | Total demand rate     |     |     |      |   |
|               | 0.6                   | 0.8 | 0.9 | 0.95 | 1  | 0.6                   | 0.8 | 0.9 | 0.95 | 1 |
|               | <b>Hybrid policy</b>  |     |     |      |    | <b>Hybrid policy</b>  |     |     |      |   |
| 1/9           | 5                     | 7   | 10  | 13   | 19 | 4                     | 5   | 6   | 5    | 5 |
| 1/3           | 4                     | 6   | 9   | 13   | 19 | 4                     | 5   | 5   | 5    | 4 |
| 1/1           | 4                     | 5   | 8   | 12   | 19 | 3                     | 4   | 4   | 4    | 3 |
| 3/1           | 3                     | 4   | 6   | 10   | 16 | 3                     | 3   | 3   | 3    | 3 |
| 9/1           | 2                     | 2   | 4   | 7    | 12 | 2                     | 2   | 2   | 2    | 2 |
|               | <b>MTO Priority</b>   |     |     |      |    | <b>MTO Priority</b>   |     |     |      |   |
| 1/9           | 5                     | 8   | 11  | 13   | 19 | 0                     | 0   | 0   | 0    | 0 |
| 1/3           | 5                     | 8   | 11  | 14   | 20 | 0                     | 0   | 0   | 0    | 0 |
| 1/1           | 5                     | 8   | 11  | 14   | 20 | 0                     | 0   | 0   | 0    | 0 |
| 3/1           | 3                     | 6   | 9   | 12   | 18 | 0                     | 0   | 0   | 0    | 0 |
| 9/1           | 2                     | 4   | 6   | 8    | 13 | 0                     | 0   | 0   | 0    | 0 |
|               | <b>MTS Priority</b>   |     |     |      |    | <b>MTS Priority</b>   |     |     |      |   |
| 1/9           | 5                     | 6   | 7   | 7    | 5  | 5                     | 6   | 7   | 7    | 5 |
| 1/3           | 4                     | 5   | 5   | 5    | 4  | 4                     | 5   | 5   | 5    | 4 |
| 1/1           | 3                     | 4   | 4   | 4    | 3  | 3                     | 4   | 4   | 4    | 3 |
| 3/1           | 3                     | 3   | 3   | 3    | 2  | 3                     | 3   | 3   | 3    | 2 |
| 9/1           | 2                     | 2   | 2   | 2    | 2  | 2                     | 2   | 2   | 2    | 2 |

of order states with and without MTO orders, for most settings. Only for low total demand rates (0.6 or 0.8), the differences are relatively small. In these cases, the demand is so low that it is not necessary to produce a large safety stock when there are no orders, as this is often the case and the lead time allowance provides a buffer in terms of time. Additional tests have shown that these differences are indeed larger for a zero lead time allowance.

Table 2.5 also allows for comparing optimal switching levels between the three policies. The switching level of MTO Priority is always higher than or equal to the switching level of the hybrid policy, apparently because MTO Priority provides less opportunity to manufacture MTS products in other order states (with orders present). Note that the switching levels for this benchmark are 0 in all states with MTO orders, by definition. When considering the switching levels of MTS Priority, we observe that these are always in between the two levels shown for the hybrid policy but in general closer to the lower levels that the hybrid policy selects for order states with MTO orders. As this benchmark prioritizes MTS also when there are many MTO orders, it is not necessary to restock to high inventory levels in order to prepare for busier times, and MTS Priority does not

**Table 2.6:** Savings % of hybrid policy against benchmarks (best benchmark in bold) depending on the demand rates

| MTO/MTS ratio       | Total demand rate |             |             |             |            |
|---------------------|-------------------|-------------|-------------|-------------|------------|
|                     | 0.6               | 0.8         | 0.9         | 0.95        | 1          |
| <b>MTO Priority</b> |                   |             |             |             |            |
| 1/9                 | 7.1               | <b>12.1</b> | <b>9.8</b>  | <b>6.4</b>  | <b>2.2</b> |
| 1/3                 | 18.5              | 21.6        | <b>16.8</b> | <b>10.5</b> | <b>4.3</b> |
| 1/1                 | 23.1              | 31.3        | <b>22.6</b> | <b>14.2</b> | <b>6.5</b> |
| 3/1                 | 23.6              | 34.1        | <b>24.8</b> | <b>15.3</b> | <b>6.3</b> |
| 9/1                 | 17.5              | 30.3        | <b>19.6</b> | <b>9.7</b>  | <b>2.6</b> |
| <b>MTS Priority</b> |                   |             |             |             |            |
| 1/9                 | <b>2.9</b>        | 12.8        | 31.0        | 43.9        | 51.1       |
| 1/3                 | <b>2.6</b>        | <b>19.3</b> | 40.0        | 49.1        | 52.8       |
| 1/1                 | <b>0.4</b>        | <b>21.5</b> | 42.4        | 49.9        | 51.8       |
| 3/1                 | <b>3.0</b>        | <b>16.3</b> | 38.8        | 47.1        | 49.1       |
| 9/1                 | <b>1.0</b>        | <b>8.5</b>  | 30.2        | 38.3        | 39.4       |

discriminate between order states with and without orders by definition, hence the relatively low switching levels.

Now consider the effects of the demand rate on the switching levels. Regarding the composition of the demand (shown vertically in Table 2.5), we observe that an increasing MTO part of the demand in general leads to lower switching levels for all policies, which is intuitive. Less MTS demand implies less need for safety stock. For increasing total demand rate levels, the effects are not unambiguous. First consider the hybrid policy. In the order state with no orders, the switching level is non-decreasing in the total demand rate, which again is intuitive: the heavier the load on the system, the more important it is to avoid idle time and instead produce MTS products in advance. In order states with MTO orders in the system, however, two effects can be observed. As the total demand rate increases from 0.6 to 0.8, the switching levels increase, as expected; but above 0.9, the switching levels decrease for an increasing total demand rate. This can be explained as follows. For a total demand rate close to 1, the policy is unable to satisfy all demand (on time). As a result, the policy focuses on the product type for which lost sales (for MTS) or the combination of lost sales and lateness (for MTO) is the most expensive. In the parameterization used here, the unit lost sales costs are equal, but for MTO, lateness costs are incurred, leading to a focus on MTO by decreasing the switching level. This effect outweighs the effect of an increasing demand rate which would favor higher switching levels.

Table 2.6 gives the savings percentages of the hybrid solution compared to the two benchmark policies for the specified values of the demand rates. Compared to MTO Priority, the hybrid policy provides savings ranging from 1% to a considerable 34%. The savings are especially large in what are arguably the most realistic settings where the total demand rate is moderate (0.8 to 0.9) and both MTO and MTS demand have a reasonable share of the demand. Table 2.5 showed that in these cases, the difference between the switching levels of the hybrid and the MTO Priority policy with no orders present are larger than in other cases. Hence, the hybrid policy benefits a lot from its ability to differentiate switching levels for the different order states, implying a much lower safety stock in the order state without MTO orders. The savings against MTO Priority are also relatively high for a low total demand rate. This is explained by the lead time allowance, which the hybrid policy uses as a natural time buffer, but MTO Priority does not. Finally, we observe that the savings decrease as the total demand rate increases, and that the smallest savings are observed when the total demand rate is 1. In these cases, the hybrid policy selects lower switching levels for states with MTO orders, and a higher switching level for the state without MTO orders, similar to the predetermined structure of MTO Priority, such that MTO Priority better resembles the hybrid policy. As a result, the savings decrease as the total demand rate increases.

Compared to MTS Priority, the hybrid policy achieves savings of about 50% if the demand is very high (0.95 to 1). As we have seen in Table 2.5, MTS Priority requires relatively low switching levels in these cases, in order to focus on MTO and prevent lateness and MTO lost sales costs. These low switching levels, however, also apply to the order state with no orders, leading to many (unnecessary) costly MTS lost sales. For lower total demand rates, the savings against MTS Priority decrease to a few per cent. In these cases, the switching levels of the hybrid policy do not vary that much over the different order states, and so MTS Priority performs better. The MTO/MTS ratio plays less of a role in determining the relative performance for MTS Priority.

If we compare to the best of the two benchmarks, the benefits range from about 1% to a considerable 25%. High savings are specifically achieved for a total demand rate of 0.9. For very low demand rates, MTS Priority provides a good alternative to the hybrid policy because it benefits from the time buffer that the lead time allowance offers. For higher total demand rates, MTO Priority provides a more reasonable alternative to the hybrid policy because also the hybrid policy focuses on MTO for these high demand rates, as explained. For a total demand rate of 0.9, none of these two effects is sufficiently present to provide a good al-

**Table 2.7:** Savings % of hybrid policy against benchmarks (best benchmark in bold) depending on the cost parameters

|     |       | MTO Priority |             |             | MTS Priority |             |             |
|-----|-------|--------------|-------------|-------------|--------------|-------------|-------------|
| $q$ | $b_o$ | $b_s$        |             |             | $b_s$        |             |             |
|     |       | 250          | 500         | 1000        | 250          | 500         | 1000        |
| 2.5 | 250   | <b>25.5</b>  | 34.8        | 43.1        | 27.4         | <b>22.8</b> | <b>21.6</b> |
|     | 500   | <b>25.2</b>  | <b>31.9</b> | 39.3        | 36.7         | 32.2        | <b>29.2</b> |
|     | 1000  | <b>25.0</b>  | <b>30.9</b> | <b>35.1</b> | 49.8         | 47.8        | 42.0        |
| 5   | 250   | <b>17.2</b>  | <b>24.7</b> | <b>33.8</b> | 41.3         | 36.5        | 34.2        |
|     | 500   | <b>17.2</b>  | <b>22.6</b> | <b>31.3</b> | 47.1         | 42.4        | 39.5        |
|     | 1000  | <b>17.2</b>  | <b>22.3</b> | <b>28.2</b> | 55.9         | 53.3        | 48.7        |
| 10  | 250   | <b>12.5</b>  | <b>14.5</b> | <b>24.1</b> | 59.8         | 54.2        | 51.5        |
|     | 500   | <b>12.5</b>  | <b>14.5</b> | <b>21.6</b> | 62.5         | 57.8        | 53.7        |
|     | 1000  | <b>12.5</b>  | <b>14.5</b> | <b>19.5</b> | 64.9         | 62.8        | 58.9        |

ternative to the hybrid policy, leading to poor performance of both benchmarks. Logically, the savings from using the hybrid policy are also especially large when both product types constitute a significant part of the total demand.

## 2.5.2 Experiments with varying cost parameters

In a second set of experiments, we vary the cost parameters. We consider three levels of parameters  $q$ ,  $b_o$ , and  $b_s$ , obtained from the base case by selecting half of the base value, the base value, or twice the base value. This leads to  $q = 2.5; 5; 10$  and  $b_o = b_s = 250; 500; 1000$ . All other parameters are equal to their base values.

Table 2.7 shows the outcomes. Again, the table reveals considerable savings, ranging from 12.5% to about 65%, and the largest savings are observed compared to MTS Priority. The savings compared to MTO Priority are especially large for lower values of  $q$  and  $b_o$  and for high values of  $b_s$  (although  $b_o$  has little effect on the savings in some cases). By contrast, the savings against MTS Priority are large for high values of  $q$  and  $b_o$  and low values of  $b_s$ . This is intuitive, as a policy with a fixed priority for either product will perform relatively better if that product is more important from a cost perspective.

Table 2.8 shows the switching levels of all three policies for the specified values of the cost parameters in the order state  $(0, 0, 0, 0, 0)$ . These reveal the following. Varying  $q$  and  $b_o$  has no effect on the switching level of MTO, because this

**Table 2.8:** Switching levels in the order state with no orders, depending on  $q$ ,  $b_o$  and  $b_s$ 

|     |       | Hybrid |     |      | MTO Priority |     |      | MTS Priority |     |      |
|-----|-------|--------|-----|------|--------------|-----|------|--------------|-----|------|
| $q$ | $b_o$ | $b_s$  |     |      | $b_s$        |     |      | $b_s$        |     |      |
|     |       | 250    | 500 | 1000 | 250          | 500 | 1000 | 250          | 500 | 1000 |
| 2.5 | 250   | 6      | 7   | 7    | 9            | 11  | 13   | 3            | 4   | 4    |
|     | 500   | 6      | 7   | 8    | 9            | 11  | 13   | 3            | 4   | 4    |
|     | 1000  | 6      | 7   | 9    | 9            | 11  | 13   | 3            | 4   | 4    |
| 5   | 250   | 7      | 8   | 9    | 9            | 11  | 13   | 3            | 4   | 4    |
|     | 500   | 7      | 8   | 9    | 9            | 11  | 13   | 3            | 4   | 4    |
|     | 1000  | 7      | 8   | 9    | 9            | 11  | 13   | 3            | 4   | 4    |
| 10  | 250   | 8      | 9   | 10   | 9            | 11  | 13   | 3            | 4   | 4    |
|     | 500   | 8      | 9   | 10   | 9            | 11  | 13   | 3            | 4   | 4    |
|     | 1000  | 8      | 9   | 11   | 9            | 11  | 13   | 2            | 3   | 4    |

benchmark chooses between MTS and idling only and MTO costs do not play any role in that decision. For the hybrid policy and MTS Priority, larger MTO costs lead to a stronger focus on MTO, i.e. lower switching levels. The effect of  $q$  is greater than that of  $b_o$ , because, for many settings, the lateness costs stimulate the policies to manufacture MTO products before lost sales come into view. Finally, for all policies, an increasing value of  $b_s$  leads to a stronger focus on MTS, i.e. larger switching levels, which is intuitive.

## 2.6 Conclusion

Existing hybrid production planning procedures generally focus on either MTO or on MTS and pay little attention to the state of the other product type. We studied a hybrid production system with both product types and showed that substantial benefits can be obtained if the state of both product types is taken into account in scheduling decisions. Up to 65% of cost savings are obtained compared to planning methods that prioritize either MTO or MTS, and up to 25% if we compare to the best of the two.

For most situations, the optimal scheduling decisions differ especially between states with and without MTO orders. If there are no MTO orders, the production capacity should be used to produce a considerable amount of stock to buffer against future demand, while if there are MTO orders, producing these orders may be preferred over replenishing stock. Interestingly, we found that

this difference vanishes if the total demand is very low. This finding relates to the presence of lead times in our model that can serve as a time buffer, which has been verified by additional tests.

A numerical sensitivity study revealed that the benefits of a hybrid planning approach are especially large in case the total demand is around 90% of the production capacity and both demand types account for at least 25% of the total demand. Under these circumstances, the demand is so large that the benefits of building a large safety stock when there are no MTO orders are substantial, while it is not large enough to allow a simple policy that selects the product for which lost sales are the most expensive to be a good alternative. As a result, the hybrid policy benefits most from its ability to differentiate its switching levels over the different order states.

Analyzing a simple, two-product system allowed us to fully describe optimal policies and make its benefits transparent, but comes with certain limitations. We have not been able to distinguish between multiple MTO or MTS products and/or corresponding variations in processing and lead times, we have not distinguished working stations (e.g. job shops), and we did not consider machine setups when switching from one product type to another. Future research could be directed at the development of heuristic planning methods for hybrid production systems in more complex settings, and evaluating these using, for instance, a simulation-based approach. Another promising direction for further research is the possibility to control the batch size of MTS products. When the machines require a setup time for MTS batches, a trade-off arises between small batches, which are finished more quickly, and larger batches, which use the production capacity more efficiently. The magnitude of savings that we observed for the optimal, hybrid policy compared to MTO (priority) or MTS (priority) approaches, certainly warrants further research and the development of specialized planning procedures for combined MTO/MTS production systems.

