Booking Limit Based Revenue Management Approaches for Customer-Value Oriented Make-to-Order Production

Nina Lohnert[(✉)] and Kathrin Fischer

Institute for Operations Research and Information Systems, Hamburg University of Technology, Am Schwarzenberg-Campus 4, 21073 Hamburg, Germany
nina.lohnert@tuhh.de

Abstract. In make-to-order (MTO) production, decisions are to be made about the jobs which are to be accepted and the sequence in which they are to be carried out. While in practice often rather simple rules like first-come-first-served (FCFS) are used, also strategies from the field of revenue management can be applied to achieve better results. In MTO not only the maximization of short-term profit should be focused on, but also the long-term perspective of performing good service in particular to valuable and returning customers is important. Therefore, in this work a booking-limit approach is combined with an order acceptance and scheduling model for a single machine environment to derive new strategies which take this aspect into account by defining different service levels to be strived at for the different customer segments. These strategies are tested on data settings with three customer segments. It turns out that a newly developed reversed nested booking limit approach (RNBL) leads to the best results regarding the conflicting aims of short-term profit maximization and customer satisfaction, whereas the classical partitioned booking limit (PBL) strategy is not recommendable.

Keywords: Revenue management · Make-to-order production ·
Booking limits · Order acceptance and scheduling · Nesting strategies

1 Introduction

These days, more and more companies use the make-to-order (MTO) principle in their production, in order to stay competitive. In MTO production, manufacturing only starts when an order has been received. This has the advantage of massively reducing the inventory of finished goods, but it leads to the difficulty that a decision has to be made regarding the orders which are to be accepted and regarding the time at which to carry them out, i.e. their scheduling. Acceptance of an order means to occupy production capacity, and the same capacity cannot be used for another, possibly more valuable order which might arrive later, and therefore has to be denied.

In this work, booking limits in combination with an order acceptance and scheduling model are used for making combined decisions on order acceptance and the subsequent scheduling. A tailored revenue management approach is developed, the basic idea of which has been presented in Lohnert and Fischer (2019). However, the

approach is modified, improved and extended here. Moreover, different booking limit strategies are developed and the results they lead to are presented.

First, some relevant background knowledge on manufacturing logistics and revenue management is briefly presented in Sects. 2 and 3. In Sect. 4, the problem definition is given, and in Sect. 5, the process of order acceptance and order scheduling is discussed and two quantitative models for these planning problems are presented. The combination of these two models and different strategies using the calculated booking limits leads to new solution procedures for the acceptance and scheduling problem. The results of the different strategies are compared in Sect. 6 using randomly generated case study data. The classical first-come-first-served (FCFS) strategy, which is often used in practice, is also included in the comparison as a reference strategy. Finally, in Sect. 7, conclusions and some directions for future research are given.

2 Theoretical Background

Production logistics is about organizing, planning, controlling and finally implementing and supervising the flow of materials and information in manufacturing companies (Gudehus 2009). In particular, capacity planning and production scheduling are important tasks in this field which will be considered by the approach presented in this work.

The major objectives that are pursued are often conflicting, as e.g. the goal of low inventory levels which leads to low costs, and the goal of service level maximization which is aimed at the maximization of customer satisfaction and the related maximization of revenue. By means of a constantly high inventory level, customers can be served quicker, but it comes with high inventory holding costs. Due to the development from a sellers' market towards a buyers' market, the customer is an important factor that is often focused on, which also leads to more variety in products to enhance customer satisfaction (Seeck 2010). As also Kersten et al. (2017) state, product individualization, cost pressure and complexity are the most important current trends in logistics, leading to changes in many companies and also in the field of production, where the individual customization of products has become more and more common.

MTO is particularly recommendable in cases where many different product varieties exist, as it helps to reduce inventories. However, this is also a disadvantage as there are no buffers, and hence any delays directly affect the customers. Still, the advantages outweigh the problems of MTO, and therefore, MTO production has become increasingly important over the past years (Barut and Sridharan 2005, Stevenson et al. 2005) and more and more companies apply this principle (Dan et al. 2018).

As these companies are confronted with uncertainty regarding the arrival of future orders, they often accept all arriving orders, including less profitable ones and more than they can actually fulfill within their limited production capacities. If employing any methods at all, they only use simple acceptance and scheduling principles like FCFS (Kalyan 2002). However, it will be shown in the following that methods from the field of revenue management can be used to allocate the scarce production resources in an improved way.

Revenue management (RM) is an approach for allocating limited, inflexible capacity in a revenue maximizing way (Guadix et al. 2010). Talluri and Van Ryzin (2005, p. 2) define RM as follows: "RM is concerned with (…) demand management decisions and the methodology and systems required to make them (…) with the objective of increasing revenues."

RM was first introduced in the flight industry, after the deregulation of the US flight market in the late 1970s, and in the meantime has been applied in many other areas as well, e.g. in the hotel and the car rental industry. An overview over the different application areas is given by Chiang et al. (2007).

Four different techniques are included in RM, namely price discrimination, capacity control or allocation, overbooking and dynamic pricing (Klein and Steinhardt 2008). In this work, only the first two instruments will be used.

In price discrimination, different customer segments are built according to the customers' willingness to pay. The product is offered to the customers in the different segments at different prices, but these different prices do not result from different production costs, i.e. the product is mainly identical. This kind of segmentation and price differentiation leads to larger sales and, hence, higher revenues, as the customers' different willingness to pay is exploited (Rehkopf 2006).

Price discrimination is often combined with capacity allocation methods. There are two types of capacity control approaches, the quantity-based and the revenue-based approach. In this work, the quantity-based approach is used in which the restricted capacity is reserved for the different customer segments, resulting in so-called booking limits. Arriving customers are accepted as long as there is capacity available in the respective segment. The segments are ranked in a hierarchical manner, e.g. according to the revenue that can be achieved by selling a unit from this segment to a customer (Talluri and Van Ryzin 2005).

As with partitioned booking limits high-ranked demand may be rejected if the respective booking limit has been exhausted, although there is still capacity available in other (lower-ranked) booking limits, the nested booking limit approach has been suggested as an alternative. This approach allows the higher-ranked classes to use the capacity of the lower-ranked ones (Talluri and Van Ryzin 2005), as long as there is still capacity available in such a lower-ranked class.

3   Revenue Management in Make-to-Order Production

3.1   Requirements for Applicability and Special Characteristics

In order for RM methods to be applicable, certain requirements have to be fulfilled. These requirements can be categorized into those regarding the capacity and those regarding the demand. The requirements and their fulfillment in MTO are also discussed in e.g. Harris and Pinder (1995), Spengler and Rehkopf (2005) and Rehkopf and Spengler (2005).

Capacities have to be inflexible, i.e. limited and fixed, and perishable. This is the case, since production capacities usually cannot be expanded, at least not in the short-run and not without significant investments and as an unused hour of production

capacity cannot be stored but is lost if it has not been used. The manufactured product could be stored, however. But a so-called external factor, in MTO production the customer and his/her specific order, has to be involved and hence, production in advance is impossible. Furthermore, there should be high fixed costs – which is the case in MTO as production capacities are usually expensive – and low marginal costs. The latter assumption is not fulfilled, since variable costs are not negligible in the case of MTO production. However, RM can still be applied when based on contribution margins instead of revenue, as e.g. Spengler and Rehkopf (2005) state, and this is also confirmed by other authors (e.g. Hintsches 2012, Klein and Steinhardt 2008).

Requirements regarding the demand are that demand should be heterogeneous, leading to different willingness to pay and to the possibility of market segmentation, and stochastic, i.e. fluctuating over the time (Talluri and Van Ryzin 2005). In MTO production, the ordered products, the times at which orders are placed and the due dates can differ and are not known in advance, hence these assumptions are fulfilled (Sucky 2009). Furthermore, the possibility of prior bookings should be given which leads to the decision to either accept the booking and hence block capacity in the future, or reject the booking, i.e. reserving capacity for potentially more attractive future bookings. As orders in MTO production often are accepted much earlier than they are actually carried out, this requirement is fulfilled (Harris and Pinder 1995). Additionally, historical data should be available, which also offers the option of forecasting demand (Talluri and Van Ryzin 2005). This is the case in MTO production due to the usage of, e.g., ERP systems. Moreover, there should be a standardized production program. Although in MTO production the products are not standardized, the required production capacity per product is.

Since the relevant requirements for the application of RM are fulfilled, MTO production is a promising field for the application of RM. However, there are important differences in comparison to other areas in which RM is usually applied.

As mentioned, marginal costs are not negligible here, as it is the case, e.g. in the flight industry, where an additional customer hardly leads to any additional costs. While in this case, revenue maximization can be used as an approximation of profit maximization, in MTO production the objective must be changed to maximization of the contribution margins, as stated above (Spengler et al. 2007). Since variable costs can in principle be allocated directly to products in MTO production, the cost allocation process is facilitated, in contrast to the service industry where such an allocation is usually not possible. However, it leads to an increased complexity of RM in MTO production compared to classical application areas since sophisticated cost accounting methods are required to capture the variable costs in sufficient detail (Rehkopf and Spengler 2005).

What is even more important is the fact that prices can be agreed individually with the customers. In order to develop strong customer relationships, those customers which are more important for the company, i.e. those who place orders on a regular basis, should be offered lower prices than one-time or new customers. Consequently, the value of the demand decreases with increasing (long-term) customer value which in turn results from the customer's contribution to the company's long-term monetary objectives (see e.g., von Martens and Hilbert (2011) for details on customer value). Valuable customers should receive a particularly high service level, as rejecting one of

their orders might lead to a loss of many more orders in the future; therefore, it can be a reasonable objective to aim at maximization of the minimal service level of valuable customer segments (Sucky 2009).

In contrast to most of the other application areas, the planning horizon is infinite in MTO production as it is the case in the hotel industry. However, there is the particularity that while the customer is promised a certain delivery date when his order is accepted, production remains flexible until it actually gets started. Moreover, in contrast to other RM problems, delays and hence late deliveries are possible, but usually they are penalized (Barut and Sridharan 2005).

Overall, it can be concluded that there are some important differences of RM in MTO production compared to classical RM, and hence RM techniques and procedures need to be adapted accordingly in order to be applicable to MTO problems.

3.2   Literature Review

Harris and Pinder (1995) are the first to study RM in assemble-to-order (ATO) production. Kalyan (2002) states that RM in production is similar to RM in the airline industry, and concentrates particularly on the determination of bid prices. Spengler et al. (2007) point out that the complexity of RM in MTO production, especially in the steel industry, is higher than in other application areas. This might be a reason why there are only a few publications yet which study RM in MTO production.

The majority of the publications on RM in MTO considers the iron and steel industry (see Hintsches 2012, Hintsches et al. 2010, Rehkopf and Spengler 2005, Spengler and Rehkopf 2005, Spengler et al. 2007). In most cases, revenue-oriented capacity allocation is used, e.g. a capacity allocation based on opportunity costs or on bid prices (see e.g., Talluri and Van Ryzin 2005 for details on bid prices). Bid prices are also used for MTO problems in other industries, see, e.g., Elimam and Dodin (2001), Kalyan (2002) and Herde (2018).

A first contribution considering capacity allocation problems in MTO production is Balakrishnan et al. (1996). Barut and Sridharan (2005) develop a heuristic approach for making capacity-allocations and hence order acceptance decisions, while Kuhn and Defregger (2004) and Defregger and Kuhn (2007) use a Markov-chain based approach. Another approach using quantity-oriented capacity planning is developed by Kumar and Frederick (2007) for the home construction industry.

Sucky (2009) derives booking limits for order acceptance decisions and studies different effects of profit and service level maximization as well as concepts for deriving a compromise for these two objectives. Below, a quantity-based approach is presented which includes ideas from Sucky (2009) and builds up on the capacity allocation model by Klein and Steinhardt (2008) and on the order acceptance and scheduling model by Thevenin et al. (2016), but extends these approaches in different directions.

4   Problem Definition

In this work, the order acceptance and scheduling process of a make-to-order driven company in a single machine environment is studied. The materials which are required for production have to be ordered from suppliers. Two different types of delivery are available: The standard delivery and an express delivery which is faster but comes with extra costs. These costs also depend on the size of the order. Furthermore, for each job a due date and a deadline are given. If a job is finished after the due date, it is delayed and the customer will not pay the full price, e.g. because of contract penalties. However, no job may be completed later than its deadline which is also known.

It is assumed that the company under consideration has been operating in the market for a while already, and therefore has a partly known customer base. The customers are assigned to different segments. Customers who have a long-term framework contract with the company are assigned to segment A. They have a high customer value for the company as they are returning customers, but due to their contract they get discounts on the price they pay per order, leading to a lower (short-term) value of their demand for the company. If a customer does not have such a contract but has already ordered sometimes at this company, the customer is assigned to segment B. These customers do not get any special conditions. Hence, the value of a customer from segment B can be described as "medium", and the same is true for the respective demand value. Finally, new customers are assigned to segment C. Since their value for the company cannot be evaluated yet, the customer value is considered to be low. However, the demand value is the highest of the customer segments, since the new customers' negotiation power is weakest and therefore they have to pay the highest prices. Moreover, also the orders as such are grouped into different categories according to their lead time, i.e. into short, medium and long orders.

The planning problem at hand is to decide which orders to accept, and when and in which sequence to carry them out in order to maximize profit while violating the aspired service levels for the different customer groups as little as possible.

5   Solution Approach

With the approach presented below, the order acceptance and the ensuing job scheduling is optimized as follows: A preliminary decision will be made by means of the booking limits which result from the capacity allocation model presented in Sect. 5.1. If the order is preliminarily accepted, it will be determined by the order acceptance and scheduling model (OAS) (see Sect. 5.2) if the order can actually be fulfilled before its deadline. If the order is to be accepted according to the OAS-model, the customer is informed accordingly. Finally, in practice also the customer can withdraw his order, e.g. if he is not satisfied with the offered delivery time. However, this aspect is ignored in the following. This order acceptance process will be conducted in specified time intervals, e.g. once per day, as new orders arrive at the company over time. Figure 1 shows this process for a single job j.
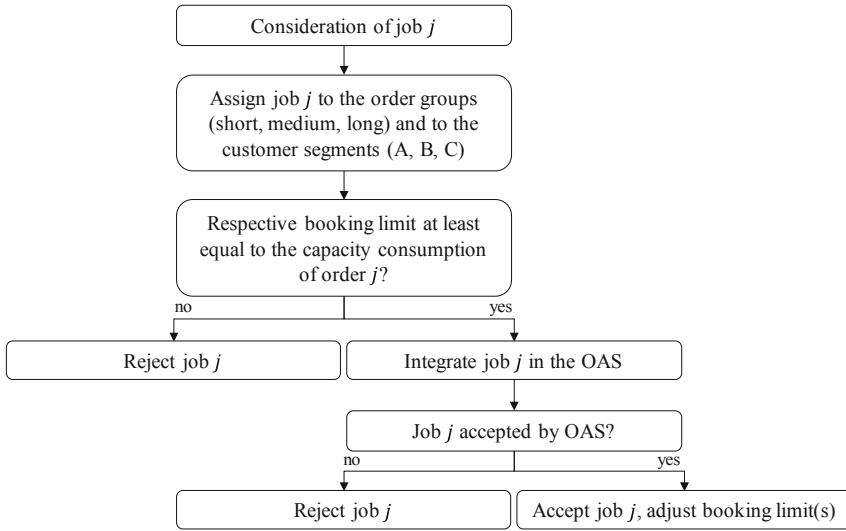
Fig. 1. Order acceptance process

5.1   Capacity Allocation Model

Using a capacity allocation model, which is based on Klein and Steinhardt (2008), optimal partitioned capacity-oriented booking limits can be determined in a first step.

Two sets are required for the model formulation. The set K contains the customer segments k, whereas Q includes the order groups q. It is important here that the combination of one element of each set is a product in the sense of RM. For each such product a booking limit x_{qk} will be determined which indicates the capacity (measured in time units) which is reserved for the respective product. For each product a demand forecast, denoted by DF_{qk}, is given for the planning horizon. Since the actual demand is unknown, it is assumed that the forecast DF_{qk} corresponds well to the actual demand and therefore reflects the orders to be expected. By the parameters aCM_{qk}, the average contribution margins for the products are given. These values are derived from historical data.

Within the planning horizon, there is a fixed production capacity CAP available on the machine. The capacity consumption of a product is specified by the order group q and mLT_q denotes the maximum lead time for this group. It is important to consider the maximum lead time (and not the average) since with this approach the real (at this point unknown) lead time cannot be underestimated and hence, the determined solution will never be infeasible.

In contrast to Klein and Steinhardt (2008), here customer segments and the corresponding aspired service levels α_k play an important role in the capacity allocation model. The service levels α_k indicate the percentage of orders of the respective customer segment k that should be accepted. Since the maximization of contribution margins does not align with the maximization of the service level (Sucky 2009), a compromise solution has to be found for these conflicting aims. The approach taken

here is an objective weighting approach, where deviations dev_k from the aspired service levels α_k are weighted with penalty costs $Pdev_k$ and contribute negatively to the objective function which otherwise consists of contribution margin maximization.

The capacity allocation model for the determination of optimal partitioned booking limits x_{qk} can be formulated as follows:

$$maximize \sum_{q \in Q} \sum_{k \in K} aCM_{qk} \cdot \frac{x_{qk}}{mLT_q} - \sum_{k \in K} Pdev_k \cdot dev_k \qquad (1)$$

$$\sum_{q \in Q} \sum_{k \in K} x_{qk} \leq CAP \qquad (2)$$

$$\frac{x_{qk}}{mLT_q} \leq DF_{qk} \qquad \forall q \in Q, k \in K \qquad (3)$$

$$\sum_{q \in Q} \frac{x_{qk}}{mLT_q} \geq (\alpha_k - dev_k) \cdot \sum_{q \in Q} DF_{qk} \qquad \forall k \in K \qquad (4)$$

$$dev_k \geq 0 \qquad \forall k \in K \qquad (5)$$

$$x_{qk} \geq 0 \qquad \forall q \in Q, k \in K \qquad (6)$$

As stated above, the objective function (1) maximizes the sum of the contribution margins minus the penalty costs for deviations from the aspired service levels. Note that the quotient of the booking limit x_{qk} and the maximum lead time mLT_q gives the product-oriented booking limit (number of product units to be produced). Constraint (2) guarantees that the sum of the booking limits x_{qk} does not exceed the available capacity CAP. With restrictions (3), it is ensured that the product-oriented booking limits do not exceed the demand forecasts DF_{qk}. By means of constraints (4), deviations dev_k from the aspired service levels α_k are captured. Finally, constraints (5) and (6) guarantee the non-negativity of the variables dev_k and x_{qk}.

5.2   Order Acceptance and Scheduling Model

The OAS-model is based on the work of Thevenin et al. (2016). However, the model is adapted to the specific planning situation as, for example, the aspired service levels α_k of the different customer segments k are taken into account. Since there is a link between this model and the capacity allocation model, some of the notation from the capacity allocation model is also used in the OAS-model.

All arriving orders j are included in the set J. By means of the indicator parameters L_{kj} and $\overline{L}_{qj}$, it is captured to which customer segment k and to which order group q job j belongs. If job j originates from customer segment k, L_{kj} is set to 1, and otherwise to 0. In the same manner it can be captured whether the job j is from order group q ($\overline{L}_{qj} = 1$) or not ($\overline{L}_{qj} = 0$).

Furthermore, there are several subsets of J. It is important to remember that the OAS-model has to be solved several times (e.g. once per day in the planning horizon). Hence, decisions which have been made in the past must be recorded and remain valid. Therefore, all jobs j which have already been accepted have to be added to the subset H

of accepted orders, and all jobs j which have been rejected either by the booking limits or by the OAS-model have to be added to the subset A of rejected orders. Hence, only the jobs j which arrive at the day of the execution are not assigned to any subset of J. Moreover, the sets $J0$ and Jn contain one dummy order each, one for the beginning and one for the end of the scheduling.

The acceptance decision of a job j is modeled by binary variables z_j ($z_j = 1$ if order j is accepted). Furthermore, for the decisions regarding the acceptance and scheduling of jobs some scheduling related parameters are necessary. The estimated lead time (including a buffer) of job j is given by the parameter LT_j. Furthermore, for each job j a regular release date R_j and an early release date $\bar{R}_j$ (resulting from material availability, i.e. the type of delivery which is chosen) is given. D_j denotes the due date and $\bar{D}_j$ represents the deadline of a job j. Both dates are defined by the customer. The variables b_j and f_j contain the times of the beginning and the completion of the respective job's production. Since the OAS-model includes sequence-dependent setup times, the setup time from order j to order i is represented by parameter ST_{ji} and the setup costs are given by parameter SC_{ji}. The sequence of orders can be determined by the binary variable y_{ji} which equals 1, if order j directly precedes order i. The individual sales prices minus several directly attributable costs are denoted by CM_j. Moreover, there are penalty costs. PE_q give the penalty costs for early start of production (per time unit) while the time is captured by variables e_j, and likewise the penalty costs PT_k are given for delayed completion (per time unit), and the number of time units a job is delayed are recorded by variables t_j. Finally, by the penalty cost parameter PU any unused production capacity (variable u) is punished, since a high capacity utilization of the perishable capacity is strived for.

Thus, the following order acceptance and scheduling model results:

$$maximize \sum_{j \in J} CM_j \cdot z_j - \sum_{j \in J0 \cup J} \sum_{i \in J \cup Jn: i \neq j} SC_{ji} \cdot y_{ji}$$
$$- \sum_{q \in Q} \sum_{j \in J} PE_q \cdot \bar{L}_{qj} \cdot e_j - \sum_{k \in K} \sum_{j \in J} PT_k \cdot L_{kj} \cdot t_j \qquad (7)$$
$$- \sum_{k \in K} Pdev_k \cdot dev_k - PU \cdot u$$

$$f_j = b_j + LT_j \qquad \forall j \in J \qquad (8)$$

$$t_j \geq f_j - D_j - \bar{D}_j \cdot (1 - z_j) \qquad \forall j \in J \qquad (9)$$

$$e_j \geq R_j - b_j - \bar{D}_j \cdot (1 - z_j) \qquad \forall j \in J \qquad (10)$$

$$b_j \geq f_i + ST_{ij} \cdot y_{ij} - \overline{D}_i \cdot (1 - y_{ij}) \qquad \forall j \in J \cup Jn, i \in J0 \cup J : i \neq j \qquad (11)$$

$$b_j \geq \overline{R}_j \qquad \forall j \in J \qquad (12)$$

$$f_j \leq \bar{D}_j \qquad \forall j \in J \qquad (13)$$

$$z_j - \sum_{i \in J \cup Jn: i \neq j} y_{ji} = 0 \qquad \forall j \in J0 \cup J \tag{14}$$

$$z_j - \sum_{i \in J0 \cup J: i \neq j} y_{ij} = 0 \qquad \forall j \in J \cup Jn \tag{15}$$

$$\sum_{j \in J \setminus A} L_{kj} \cdot z_j \geq (\alpha_k - dev_k) \cdot \sum_{j \in J} L_{kj} \qquad \forall k \in K \tag{16}$$

$$b_{n+1} - \sum_{j \in J} \bar{R}_j \cdot y_{0j} - \sum_{j \in J} LT_j \cdot z_j - \sum_{j \in J0 \cup J}^{n} \sum_{i \in J \cup Jn: i \neq j} ST_{ji} \cdot y_{ji} = u \tag{17}$$

$$z_j = 1 \qquad \forall j \in H \cup J0 \cup Jn \tag{18}$$

$$z_j = 0 \qquad \forall j \in A \tag{19}$$

$$b_{n+1} \leq CAP \tag{20}$$

$$dev_k \geq 0 \qquad \forall k \in K \tag{21}$$

$$b_j, e_j, f_j, t_j \geq 0 \qquad \forall j \in J \tag{22}$$

$$u \geq 0 \tag{23}$$

$$y_{ji} \in \{0, 1\} \qquad \forall j \in J0 \cup J, i \in J \cup Jn : i \neq j \tag{24}$$

$$z_j \in \{0, 1\} \qquad \forall j \in J \tag{25}$$

In the objective function (7) the sum of the contribution margins of all accepted orders minus the setup and the different penalty costs is maximized. With constraints (8), the beginning of production and the completion time of a job j are set into a fixed relation, depending on the time required for job j. If the completion of a job j is delayed or the start of production is early, this is captured by constraints (9) and (10). Constraints (11) and (12) specify the beginning of the production. For the start of production of a job j, the preceding job i has to be completed and the setup must have been carried out. Furthermore, constraints (12) guarantee that the processing of a job j cannot be started before the necessary materials have been delivered. By means of constraints (13) it is ensured that the deadline $\bar{D}_j$ is the latest possible time of completion of job j. Constraints (14) and (15) guarantee that every accepted order has exactly one preceding and exactly one subsequent job.

Constraints (16) are the service-level-restrictions. Any deviations from the service levels α_k will be determined through these constraints. By means of restriction (17), any unused production capacity will be identified and penalized in the objective function (7). The beginning b_{n+1} of the dummy order $n+1$, i.e. the end of the production time, is compared with the early release date $\bar{R}_j$ of the first scheduled order j. By subtracting the sum of all lead times LT_j of the accepted orders and the sum of the setup times ST_{ji}, the result gives the unused capacity u.

By constraints (18) it is ensured that every order from subset H and the dummy orders will be accepted. Consequently, every order from subset A is rejected by means

of constraints (19). Restriction (20) limits the planning horizon to its production capacity *CAP*. Finally, constraints (21) to (25) define the variables' definition areas.

5.3   Booking Limit Acceptance Strategies

The booking limits x_{qk} can be used in various ways in order to decide about the (preliminary) acceptance of bookings. One of them is the direct usage of the partitioned booking limits x_{qk} (PBL) calculated by the capacity allocation model. However, this has the disadvantage that bookings in higher valued classes cannot use the capacity of lower valued classes. As mentioned above, the nesting of booking limits is a possibility to get improved results. Three different nesting strategies will be explained in the following. For all nesting options considered it can be assumed that a short order which requires less time is more profitable per time unit than a medium one, and this in turn is more profitable than a long order.

One option is the nesting based on the demand value (NDV). For this approach, the products have to be ranked in a hierarchical manner according to their average contribution margins aCM_{qk}^* per time (and hence capacity) unit. Here, the superordinate criterion for the demand value of the product is the customer segment k, and the subordinate criterion is the order group q. Hence, a short order from a C-customer has the highest demand value, whereas a long order of from segment A has the lowest demand value. Consequently, a short order from a C-customer can use the booking limits of all other classes which gives preference to the – in terms of customer value less valuable – C-customers.

Another option is to nest the booking limits only within the customer segments (NWCS) in order to restrain, e.g., C-orders from using class A capacity. Thus, e.g. a short order of customer segment C can only access the booking limits $x_{medium,C}$ and $x_{long,C}$ and not those of the other customer segments.

Furthermore, so-called reverse nested booking limits (RNBL) can be used in order to pursue a higher customer satisfaction of the crucial customers, i.e. those from segment A. Therefore, again the customer segment k is the superordinate criterion, but with A as the highest ranked and C as the lowest ranked customer segment. The subordinate criterion is again the order group q. So in this case, a short order from an A-customer can use the booking limits of all other classes.

The nesting and the adjustments of the booking limits follow the rules of the standard nesting (Talluri and Van Ryzin 2005) and hence the nested booking limits can be determined by a summation of the partitioned booking limit x_{qk} itself and the booking limits of all accessible lower-ranked products. For a detailed explanation of the standard nesting process and the necessary adjustments in the case of the NDV strategy, see Lohnert and Fischer (2019).

If an order is accepted, the booking limit/s has/have to be adjusted by the actual capacity usage. Hence, the real (at this point known) lead time LT_j plus 5% of the maximum lead time mLT_q, which is used as an approximation of the sequence-dependent setup times, is subtracted from the respective booking limit/s. Further orders can only be accepted as long as the respective booking limit is sufficiently large.

6   Numerical Studies

Various case studies have been conducted in order to compare the four different booking limit acceptance strategies. Furthermore, FCFS, the most common acceptance approach in practice (Kalyan 2002), is considered for comparison. For all case studies, order data were randomly generated, but it was always ensured that the demand exceeds the supply and that the share of the orders of customer segment A makes up about 50% of the orders, whereas the share of orders originating from customer segment B contributes around 30% and the remaining orders are from new customers. Due to space limitations, the full data sets of these case studies and the complete set of rules used for their generation cannot be given here, but the data are available from the authors on request.

It is assumed that the demand forecast, hence the amount of expected orders of the different products, corresponds to the actual demand, i.e. the forecast is assumed to be perfect. However, this does not mean that the actual amount of capacity required is the same as predicted since the actual lead time LT_j might be smaller than the maximum lead time mLT_q. All penalty cost coefficients are set to fixed values, such that penalty costs which relate to a customer segment k are set to higher values, since they indicate that the customer is directly affected, compare to those which depend only on the order group q. $Pdev_k$ have even higher values than PT_k. The value of PU results from the objective function value of the capacity allocation model divided by the available capacity CAP, which is one month, i.e. 20 days, in all case studies.

For the different acceptance strategies, a recurring pattern over the various case studies was clearly visible. In the following, the results of one exemplary selected case study will be presented in order to illustrate the major findings. This case study includes 33 orders which are randomly generated according to the above-mentioned rules and arrive randomly during the planning horizon. Figure 2 shows the results of the different order acceptance strategies.
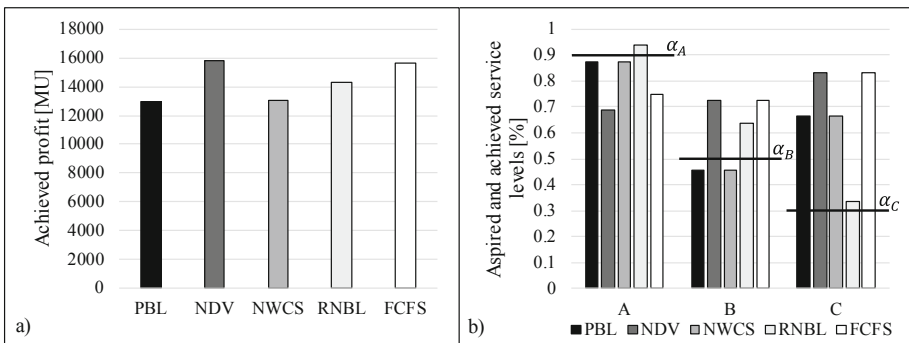

Fig. 2. Results of acceptance strategies in comparison: (a) profit, (b) service levels

As can be seen in Fig. 2a, the highest profit, i.e. the sum of the contribution margins, calculated only including actual variable costs and excluding penalty costs

$Pdev_k$ and PU, can be achieved by using the NDV strategy. With roughly 1% less, the second highest value in terms of profit can be achieved by using FCFS. This is quite a good performance of FCFS, taking into account that the resulting scheduling of the orders is not optimal. In the other case studies, FCFS also leads to the second highest profit, but with a larger difference compared to the highest value which is always achieved with NDV. With the RNBL strategy, a somewhat lower value can be gained. PBL and NWCS both attain the same result and in all cases the lowest profit.

However, it is crucial to also compare the achieved service levels for the different customer segments (Fig. 2b), since these different segments are of different relevance for the future long-term success of the company. As pointed out above, it is important to satisfy A-customers in order to fulfill the framework contract and therefore stabilize the customer base. New customers (segment C) are less important compared to the other segments. The aspired service levels α_k are therefore set to 90% for A-customers, 50% for B-customers and 30% for C-customers.

The best result regarding the service levels can be achieved with the RNBL strategy, as the service level for customer segment A is above 90% and also the aspired service levels for segment B and C are achieved by the RNBL strategy. The second highest service level for segment A can be achieved by PBL and NWCS. However, customer segment B has a rather low service level when those acceptance strategies are used, in contrast to RNBL which leads to quite a good value also for customer segment B, whereas it does not accept many customers from segment C. While in this case study, PBL and NWCS lead to the exactly same result, in other cases NWCS performed better in terms of profit achieved, without decreasing the service level compared to PBL. Hence, NWCS dominates PBL.

NDV and FCFS achieve the highest profit but they reach those good results because they accept most of the orders of customer segment C (high demand value, but low customer value). Especially FCFS does not take any service level requirements into account at all. Hence, FCFS is not a useful strategy when this long-term objective is of importance, but it performs rather well when only the profit is considered, as it leads to a high usage of available capacities.

It should be noticed that capacity utilization differs for the different strategies. The sum of all partitioned booking limits x_{qk} always equals the available capacity CAP but since the actual lead time LT_j might be smaller than the maximum respective lead time mLT_q, it is possible that the booking limits are not completely exhausted during the planning horizon. Thus, capacity can remain unused. In the case study considered here, the highest level of capacity utilization (98%) is reached by FCFS, closely followed by the acceptance approach NDV. With RNBL, a capacity utilization of 91% was reached. PBL and NWCS lead to a capacity usage of 80%.

Overall, the different case studies showed that the trade-off between profit maximization and the maximization of the service levels of the important customers can be optimized when the RNBL acceptance strategy is used. This strategy achieves the highest service level for customers in segment A and at the same time enables a rather high profit, while making good use of the available capacity.

7 Conclusions and Outlook

Since the customization of products leads to an enormous increase of variant diversity, more and more companies use the concept of MTO production in order to stay competitive. However, due to stochastic and volatile demand, optimal order acceptance decisions are rather difficult to make. Therefore, the application of appropriate RM strategies may help in facilitating and optimizing these decisions.

In this work, a capacity allocation model and an OAS-model are combined to develop different strategies for making order acceptance and scheduling decisions. The application of those strategies and the classical FCFS approach to different case studies shows that a new strategy using RNBL leads to the best compromise between the maximization of total short-term profit and the maximization of the service level of crucial customers which will contribute to the long-term success of the company. Hence, the RNBL approach is particularly suitable for MTO problems.

In future work, a larger number of cases has to be studied within a simulation study framework to confirm that the results stated above are independent of the underlying data. Moreover, the relation between different cost parameters, especially the penalty costs for the deviation from the aspired service levels, has to be studied in sensitivity analyses in order to investigate their influence on the acceptance of jobs. Furthermore, the performance of the different booking limit acceptance strategies should be tested when demand differs significantly from its forecast. Moreover, it should be studied how the long-term effects of the achieved service levels, in particular with respect to customers with low demand but high customer value, could be measured. In the approach presented here it is assumed that a higher service level will lead to higher long-term profits; however, this might be studied empirically and an appropriate approach for modeling this aspect might be built.

Also in the scheduling process more aspects could be considered like for instance more machines and the possibility of using overtime. Hence, RM in MTO is a very promising field for future research.

References

Balakrishnan, N., Sridharan, V., Patterson, J.W.: Rationing capacity between two product classes. Decis. Sci. **27**(2), 185–214 (1996)

Barut, M., Sridharan, V.: Revenue management in order-driven production systems. Decis. Sci. **36**(2), 287–316 (2005)

Chiang, W.-C., Chen, J.C., Xu, X.: An overview of research on revenue management: current issues and future research. Int. J. Revenue Manage. **1**(1), 97–128 (2007)

Dan, B., Gao, H., Zhang, Y., Liu, R., Ma, S.: Integrated order acceptance and scheduling decision making in product service supply chain with hard time windows constraints. J. Industr. Manage. Optim. **14**(1), 165–182 (2018)

Defregger, F., Kuhn, H.: Revenue management for a make-to-order company with limited inventory capacity. OR Spectrum **29**(1), 137–156 (2007)

Elimam, A.A., Dodin, B.M.: Incentives and yield management in improving productivity of manufacturing facilities. IIE Trans. **33**(6), 449–462 (2001)

Guadix, J., Cortés, P., Onieva, L., Muñuzuri, J.: Technology revenue management system for customer groups in hotels. J. Bus. Res. **63**(5), 519–527 (2010)

Gudehus, T.: Comprehensive Logistics, 1st edn. Springer, Heidelberg (2009)

Harris, F.H.deB., Pinder, J.P.: A revenue management approach to demand management and order booking in assemble-to-order manufacturing. J. Oper. Manag. **13**(4), 299–309 (1995)

Herde, F.: Revenue-Management-Ansatz für eine Annahmesteuerung kundenspezifischer Regenerationsaufträge komplexer Investitionsgüter. Springer, Wiesbaden (2018)

Hintsches, A.: Dynamische Kapazitätssteuerung bei kundenindividueller Auftragsproduktion in der stahlverarbeitenden Industrie, 1st edn. Gabler Verlag, Wiesbaden (2012)

Hintsches, A., Spengler, T.S., Volling, T., Wittek, K., Priegnitz, G.: Revenue management in make-to-order manufacturing: case study of capacity control at ThyssenKrupp VDM. Bus. Res. **3**(2), 173–190 (2010)

Kalyan, V.: Dynamic customer value management: asset values under demand uncertainty using airline yield management and related techniques. Inf. Syst. Front. **4**(1), 101–119 (2002)

Kersten, W., Seiter, M., von See, B., Hackius, N., Maurer, T.: Chancen der digitalen Transformation: Trends und Strategien in Logistik und Supply Chain Management. DVV Media Group GmbH (2017)

Klein, R., Steinhardt, C.: Revenue Management. Springer Verlag (2008)

Kuhn, H., Defregger, F.: Revenue Management in der Sachgüterproduktion. WiSt-Wirtschaftswissenschaftliches Studium **33**(5), 319–324 (2004)

Kumar, S., Frederick, J.L.: Revenue management for a home construction products manufacturer. J. Revenue Pricing Manag. **5**(4), 256–270 (2007)

Lohnert, N., Fischer, K.: Einsatz des Revenue Managements in der Make-to-Order-Produktion. In: Schröder, M., Wegner, K. (eds.) Logistik im Wandel der Zeit – Von der Produktionssteuerung zu vernetzten Supply Chains: Festschrift für Wolfgang Kersten zum 60. Geburtstag, pp. 235–262. Gabler Verlag (2019)

Rehkopf, S.: Revenue Management-Konzepte zur Auftragsannahme bei kundenindividueller Produktion: Am Beispiel der Eisen und Stahl erzeugenden Industrie. Springer Verlag, Wiesbaden (2006)

Rehkopf, S., Spengler, T.: Revenue management in a make-to-order environment. In: Operations Research Proceedings 2004, pp. 470–478. Springer, Heidelberg (2005)

Seeck, S.: Erfolgsfaktor Logistik. Gabler Verlag (2010)

Spengler, T., Rehkopf, S.: Revenue Management Konzepte zur Entscheidungsunterstützung bei der Annahme von Kundenaufträgen. Zeitschrift für Planung & Unternehmenssteuerung **16**(2), 123–146 (2005)

Spengler, T., Rehkopf, S., Volling, T.: Revenue management in make-to-order manufacturing – an application to the iron and steel industry. OR Spectr. **29**(1), 157–171 (2007)

Stevenson, M., Hendry, L.C., Kingsman, B.G.: A review of production planning and control: the applicability of key concepts to the make-to-order industry. Int. J. Prod. Res. **43**(5), 869–898 (2005)

Sucky, E.: Revenue Management bei Auftragsfertigung – Ein Ansatz zur Kapazitätssteuerung. In: Voß, S., Pahl, J., Schwarze, S. (eds.) Logistik Management, pp. 77–96. Physica Verlag (2009)

Talluri, K.T., Van Ryzin, G.J.: The Theory and Practice of Revenue Management. Springer, Heidelberg (2005)

Thevenin, S., Zufferey, N., Widmer, M.: Order acceptance and scheduling with earliness and tardiness penalties. J. Heuristics **22**(6), 849–890 (2016)

von Martens, T., Hilbert, A.: Customer-value-based revenue management. J. Revenue Pricing Manag. **10**(1), 87–98 (2011)