Contents lists available at ScienceDirect

Int. J. Production Economics

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

Modeling and simulation of order-driven planning policies in build-to-order automobile production

Thomas Volling *, Thomas S. Spengler

Institute of Automotive Management and Industrial Production, Technische Universität Braunschweig, Katharinenstraße 3, D-38106 Braunschweig, Germany

ARTICLE INFO

ABSTRACT

In adopting build-to-order order fulfillment systems, automotive companies strive to better synchronize their production output with market demand. This essentially gives rise to a new paradigm in production planning. Since all business is linked to customer orders, the operational performance is substantially determined by order-driven planning. Therefore, a clear understanding of the associated planning tasks, order promising and master production scheduling, as well as their dynamic interaction is essential. Based on the analysis of the decision situation of order-driven planning in build-to-order settings, we provide a framework comprising separate interlinked quantitative models for order promising and master production scheduling. The focus of the contribution is on the modeling and evaluation of both models in a dynamic setting. The approach is evaluated by means of a simulative analysis using empirical data from the automotive industry. Conclusions regarding the potentials of such systems with respect to customer service, the leveling of resource utilization, and holding are presented.

1. Introduction

Much has been changing since Henry Ford's "we believe […] that no factory is large enough to make two kinds of products" (Ford, 1988, p. 82). With their Scion brand Toyota joined the race for offering customers an ever increasing product variety; a trend which has been characterizing the automotive industry throughout the last decades (Lee et al., 2005; Pibernik, 2005). This development has been recognized to be driven by two factors. On the demand side, customization is driven by the improved competitive position of companies which address individual customer's needs (Kotler, 1989). On the supply side, customization strategies have been significantly promoted – if not been made possible at all – by advances in product design and manufacturing as well as information technology (Da Silveira et al., 2001; Jiao et al., 2003). Based on these advances it became possible to postpone the point of product differentiation after customer orders are received, by combining standardized modules within final assembly (Andres, 2006). The corresponding strategy is referred to as build-to-order (BTO) production.

The main challenge companies have to address when pursuing BTO production lies in mastering the increased exposure to demand variability (Holweg et al., 2005). In contrast to make-to-stock order fulfillment systems, where inventory is used to hedge against short-term demand variability and to facilitate an economic mode of production, BTO directly links production activities to market dynamics. Reasons for these dynamics are variations in the timing and specifications of customer requests (i.e. the demand sequence), the resulting model mix (demand structure), and the aggregated demand per period (demand level). Considering the limited flexibility of automobile production systems in terms of production and procurement capabilities, the synchronized adjustment of capacity with the volatility of the market environment is not viable. Instead, adequate control concepts are needed to match the supply of resources with the demand for products (Alford et al., 2000; Holweg and Pil, 2004, p. 105). These encompass policies for the determination of due dates in response to customer requests as well as for the consolidation of these quoted requests into production plans that allow for an economic mode of production. The associated planning tasks are referred to as order-driven planning (ODP).

The increasing interest in BTO production is likewise reflected by the fact that leading business software providers have developed commercial applications to support ODP for BTO production as integral part of their Advanced Planning and Scheduling (APS) Systems (Abraham, 2002; SAP, 2005a). In order to fully benefit from such systems, however, models are required, which provide decision support for ODP. To this end, the performance of such models in dynamic settings, as it is the case in BTO automobile

* Corresponding author.
  E-mail addresses: t.volling@tu-braunschweig.de (T. Volling),
t.spengler@tu-braunschweig.de (T.S. Spengler).

production, needs to be well understood. Despite this fact there is very limited work on this subject.

The aim of this paper is to model and simulate ODP policies for BTO automobile production. The contribution is threefold. First, we provide an in-depth analysis of the decision situation of ODP and derive a framework comprising separate modules for order promising and master production scheduling. Second, we develop mathematical models for both modules and introduce three planning policies which result from the combination of these models. The third contribution is that we perform a simulation study to quantify the potentials of each planning policy with respect to customer service, the leveling of demand, and holding. To validate the modeling approach, we analyze the controlling influence of the relevant structural model elements. The most important feature is that we explicitly address the dynamic interaction of the ODP planning modules in model development and evaluation. The remainder is organized as follows: In Section 2, we will review prior work. The problem setting is characterized in Section 3. Based on that, models will be presented in Section 4 and analyzed using discrete event simulation in Section 5. The findings will be summarized in Section 6.

2. Literature review

Work on the performance of order fulfillment systems in dynamic environments is limited. Recent papers have been published by Holweg et al. (2005) as well as Brabazon and MacCarthy (2006). Subject of the former contribution is the highly aggregated analysis of BTO order fulfillment systems using a system dynamical approach. The latter investigates potentials and risks of an order fulfillment concept entitled virtual-BTO. This concept essentially allocates customer requests to (pre-) specified products planned for production or kept in inventory. The authors set up a discrete event simulation model for the case of the automotive industry. In contrast to the problems studied in these studies, our concern is not on order fulfillment when capacity is represented by products which have been fully specified based on forecasts, but on the explicit modeling of bottleneck capacity. Individually configured customer requests are thus checked against the unused capacities upon their entry into the system. This goes along with a capacity oriented approach of hierarchical production planning instead of material oriented procedures.

Regarding decision support for ODP, two research streams can be distinguished. Within the first stream the focus is on the analysis of the customer interaction, i.e. the promising of customer requests. A second stream has evolved from work on assembly line balancing and sequencing. The focus of which is to develop decision models for the master production scheduling of orders. Both streams are briefly reviewed in the following.

The objective of work on *order promising* is to support decisions of whether to reject or quote customer requests. Two methodological approaches can be distinguished. In *batch order promising* orders are collected over a certain time span, and are subsequently processed simultaneously. Numerous approaches have been provided (Chen et al., 2001, 2002; Jeong et al., 2002; Pibernik, 2005). These models seek to maximize contribution margin with respect to holding, production, and procurement costs as well as 'soft' costs such as penalties for late delivery or low capacity utilization. A central finding is that, if previously promised due dates and quantities are treated as constraints for subsequent planning cycles, good results can be obtained for rather large batching intervals. A reduction of the intervals results in the significant deterioration of the planning performance (Chen et al., 2001). A similar approach is introduced by Fleischmann and Meyr (2004). To avoid the drawbacks of short batching intervals, the authors propose an integrated assessment of order promising and master production scheduling. They however do not elaborate on the consequences in terms of modeling and performance. A drawback with batch approaches is that they do not support the interaction between customer and company. *Real-time approaches* therefore build on an ad-hoc assessment of the ability to deliver at a certain point in time. A recent review is given by Moses et al. (2004). Optimization based real-time approaches have, to our knowledge, only been presented for academic examples (e.g. Raaymakers et al., 2000; Wester et al., 1992; Kate, 1994). The objective is to quickly determine detailed schedules in job shop environments. Due to the complexity of the problem, the scalability of the approach is limited. One exception is Robinson and Carlson (2007). The focus is on the dynamic pegging of material in multi level production. Resource constraints are not incorporated into the model. To facilitate a responsive order promising mechanism, which can be used as an order winning factor within the sales processes, we develop an optimization based real-time approach. The structure of the objective function is somewhat similar to the batch models discussed above. Yet, to avoid the downside of long response times attributed to extended batching intervals, as well as the myopic performance of short ones, we integrate the model into a framework, which allows for asynchronous re-planning. This framework consists of distinct models for order promising and master production scheduling. Resource constraints are explicitly taken into account.

There is a rich body of literature considering questions of *master production scheduling*. The underlying structure of our model is thereby similar to models used for order-driven production planning in mixed model scenarios. Based on Hindi and Ploszajski (1994), Bolat (2003) provides solution procedures for the problem concerning the selection of orders to be produced in the upcoming period out of a pool of previously quoted orders. The objective of the resulting multi-dimensional knapsack problem is to minimize delivery date dependent costs. For a similar decision situation, Ding and Tolani (2003) present a lexicographic goal programming approach. Unlike these approaches, we do not assume orders to be exogenously given, but explicitly model order promising and master production scheduling as distinct, interdependent planning functions. Therefore, we are firstly able to capture the impact of production planning routines on the responsiveness and reliability of the order fulfillment system and, vice versa, that of order promising decisions on the performance of production planning. Boysen (2009) presented a model that integrates aspects of sequencing and optimization-based batch order promising. The objective of the approach is to minimize delivery dependent costs for a mixed model scenario. Further model formulations are presented, which incorporate order selection decisions, capacity adjustments and sequencing. In contrast to our approach, the focus is on a static decision situation without any information dynamics.

3. Problem setting

In the remainder, we consider ODP in BTO automobile production. We distinguish a real-time order promising (OP) model and a master production scheduling (MPS) model, executed asynchronously based on rolling horizons. Both models are interlinked by information flows such that we obtain the planning system illustrated in Fig. 1.

The process cycle of ODP is as follows. Incoming customer requests are processed individually upon their arrival within the OP procedure (1). OP results into a production period for the specific configuration of the order. This is done by taking into
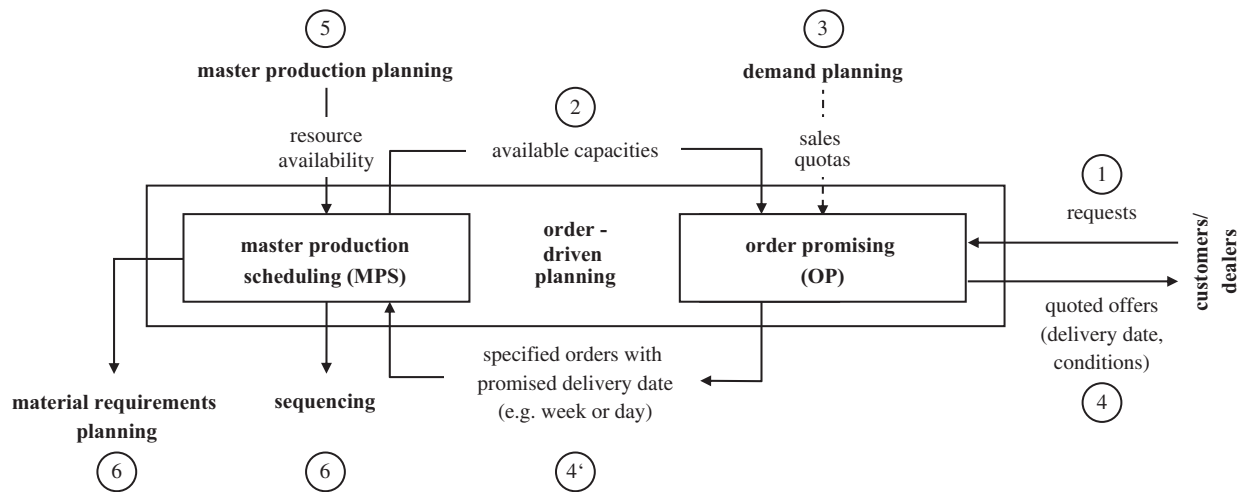
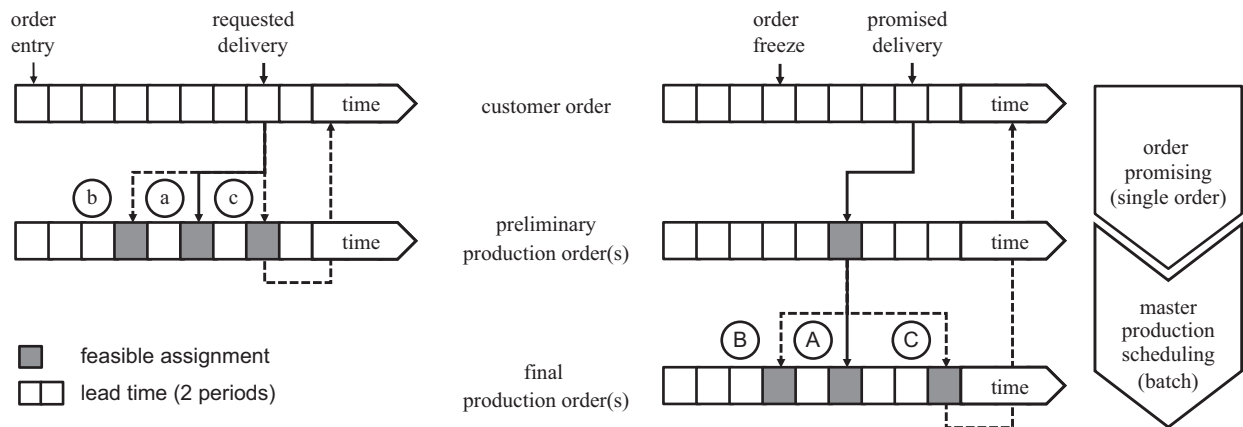Fig. 1. Customer service view of order-driven planning.

Fig. 2. Structure of the costs of the assignment: response delay (left) and confirmed fill delay (right).

account the lead time, the currently available capacities (2) and sales quotas (3), which are used to coordinate the different sales channels of automotive OEMs (Meyr, 2004). The quoted offer and the delivery date that results from the production period are returned to the requester (4). In case the offer is accepted, a preliminary production order is generated and passed on to MPS (4′).

The objective of MPS is to coordinate production, procurement and sales on the short-run in order to facilitate an efficient mode of production (Vollmann et al., 2005, p. 169). This is done by taking into account the set of accepted orders with specific configurations and quoted due dates (4′) as well as aggregate capacity constraints, which are originating from the subordinate master production planning and are used for the mid-term coordination of production and sales (5). MPS results in instructions for downstream planning tasks such as material requirements and sequencing planning (6) as well as in updated data regarding unused capacities, which is again transferred to OP (2). As follows, MPS defines the factual production period of the preliminary production orders. Since the decision situation of MPS is different to that of OP, the factual period does not necessarily coincide with the one derived from the promised due date. In this case the order is either produced earlier than requested or it is delayed. The consequences in terms of customer service are elaborated on in more detail below. MPS planning is

executed asynchronously based on rolling horizons. Accordingly, only the first period of each planning horizon is put into practice and passed on to downstream planning. The other ones are updated with respect to new information available (i.e. the newly promised orders) in the course of consecutive planning cycles.

Considering the illustrated decision situation, two dimensions of customer service can be distinguished. The difference between the requested delivery date and the quoted date is referred to as *response delay*. Since the requested date is given exogenously, the OP scheme can be employed to influence this performance indicator by adequately setting the period quoted to the customer. In order to assure the feasibility of this quotation, a preliminary production order needs to be generated which reserves the capacities necessary to serve the request. This fact is illustrated by the left hand side of Fig. 2. Case (a) corresponds to an assignment according to the lead time of the production order. The order would thus be completed just in time and the requested date could be confirmed. If capacities do not allow for the assignment indicated by case (a), the order could be filled by production in advance of the requested period (b). In this case, the finished product needs to be held on inventory for the time remaining, if the customer cannot be convinced of an earlier delivery. Lastly, the production of the order could be delayed (c). This would increase flexibility in production but would also require

convincing the customer of a delivery later than requested. The response delay can thus be both, negative (b) and positive (c).

A second dimension of customer service arises from the fact that there typically are some periods between the order entry and the point in time, when the order is ultimately passed on to the production system for completion, the so called order freeze. The order might therefore be re-assigned, if this allows for a more efficient mode of production. This is done in the course of MPS, which in contrast to OP takes into account all orders that have been quoted up to that point in time (batch approach). As a consequence of the re-assignment three further cases can be distinguished. Case (A) represents the default case where the preliminary assignment is kept unchanged. In case (B) the order is moved forward in time. In this case the finished product has to be stored until delivery or the previously quoted delivery is changed. Case (C) results in the order being delivered late. The time span between the quoted and the factual delivery is referred to as *confirmed fill delay*. The distinction between both service measures is necessary, since requested delivery dates are typically quite flexible in the automotive industry. Having once quoted a period, however, there is a contractual agreement. Deviating from this quotation might result in severe penalties. Against this background we restrict the analysis to the use of holding, if the MPS assignment is earlier than that of OP.

In addition to customer service, ODP determines the aggregated production plans. If the generated plans do not comply with the configuration of the production system, inefficiencies in terms of increased production and procurement costs result. Since production and procurement capacities are regarded as fixed for the planning horizon considered, a leveled utilization of the resources is to be achieved (Engel and Zimmermann, 1998). Leveling allows for efficiency in production and helps minimizing the variability which is propagated into the supply chain, reducing system costs therewith.

The decision situation of ODP is consequently characterized by multiple, potentially conflicting objectives, which are influenced by the dynamic interaction of OP and MPS. Moreover, the decision situations of OP and MPS differ significantly. OP has to process a large number of customer requests and generates an output that is directly perceived by customers. Major requirements to be met by OP policies therefore regard the responsiveness, which determines the latency in quoting customer requests, as well as the granularity of the generated quote, which may be the day or week of delivery. The situation is different, if MPS is considered. Since there is no output, which is directly perceived by the customer, response time requirements are less restrictive. However, all accepted orders need to be taken into consideration when constructing production plans. This results into a significantly higher complexity. In order to take the distinct characteristics of both planning tasks into account, the presented approach is based on separate models for OP and MPS. Against this background, the question will need to be addressed in the remainder, of how OP and MPS decision making can be coordinated such that an acceptable overall performance with respect to service and efficiency criteria is obtained in the considered setting of BTO production.

The presented framework is of generic nature. It is not relevant, whether the requesting party is a final customer or an intermediate dealer. In the automotive industry, however, the dominating sales channel is via dealers. Online sales directly to customers have not been found to account for a major share, yet (Holweg and Miemczyk, 2002). Accordingly, dealers are helping in translating the requirements of the customer into product characteristics and resource demand, respectively. This is usually done with the use of product configuration tools, dealers can operate while negotiating with the customer. With respect to the framework illustrated in Fig. 1, these tools generate the input needed for OP. We will therefore assume in the following that a dealer is formulating the requests and returned the quoted offers. In this, the dealer cooperates with the manufacturer such that both strive to minimize system costs. Moreover, we assume that all orders have been back-scheduled with respect to their production and distribution lead times. The analysis can thus be restricted to the start of production. We also assume that sales quotas are checked separately and do not need to be considered within ODP. Lastly, we restrict our analysis to one production system. This results in the models introduced in the following section.

4. Modeling order-driven planning

Time is divided into intervals t ($t \geq 1$) of equal length. The objective of OP is to assign randomly arriving customer requests $i(i = 1,\ldots,I)$ to planning periods $\tau(\tau = t,\ldots,t + T^{max})$ such that the associated costs $c_{i\tau}^{OP}$ are minimized. To generate consistent and feasible quotes, two constraint sets have to be considered. Firstly, capacities need to be sufficiently available to serve the request. Secondly, the assignment cannot precede the first disposable period t, which is the first period that has not yet been frozen. We use production coefficients a_{ir} to transform each order's product configuration into capacity requirements for each potential bottleneck resource r. These resources likewise comprise capacities of the production system as well as the availability of major components (heavy items), which are supplied just in time (Boysen, 2009; Meyr, 2004). The combined index set will be referred to as Ω.

In this model formulation the costs of the assignment $c_{i\tau}^{OP}$ are used to model the costs encountered depending on the quotation of a request. The costs of the assignment can be conceptualized as follows. It is certainly valued most, if the order is fulfilled as requested by the customer. If capacity constraints do not allow for such an assignment, the first option is to encourage the customer to agree with an earlier or later delivery. Assuming that any period other than the requested is of negative value to the customer, the dealer might need to offer appropriate incentives or encounter a loss of goodwill. This fact can be conceptualized by delivery dependent costs, which are increasing with the time span between the requested and the quoted period. The second option is to make use of holding. This is possible if capacities are sufficiently available to produce the request earlier than demanded. However, holding results in additional costs such as insurances fees and capital commitment costs. These costs are increasing with the time span between the feasible (early) period and the one quoted to the customer. Accordingly, the following cases can be distinguished. If the feasible period is later than the requested, the costs of the assignment $c_{i\tau}^{OP}$ equal the delivery dependent costs. If the feasible period is earlier than the requested, the costs of the assignment can be calculated as the sum of holding and delivery dependent costs. This case is illustrated in Fig. 3. Assuming linear cost functions, the optimal quotation policy is to either quote the feasible period – if the marginal delivery depended costs are lower than the marginal holding costs – and to quote the requested period vice versa. In the automotive industry the delivery date is typically quite flexible within the negotiation phase, while holding costs account for a significant share of total costs (Krog, 2006). We will therefore restrict the analysis to the case that the dealer will always quote the delivery derived from the production period.

The result of the OP procedure is a quoted planning period, or the corresponding delivery date, for a certain product configuration. Executing OP every time an order arrives will yield a set of
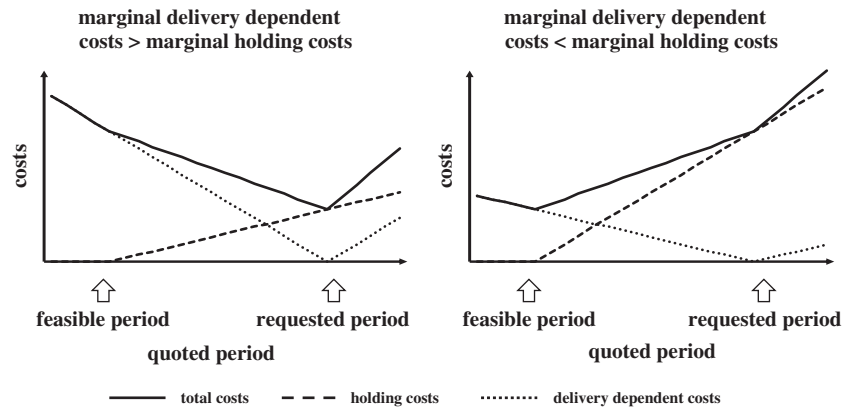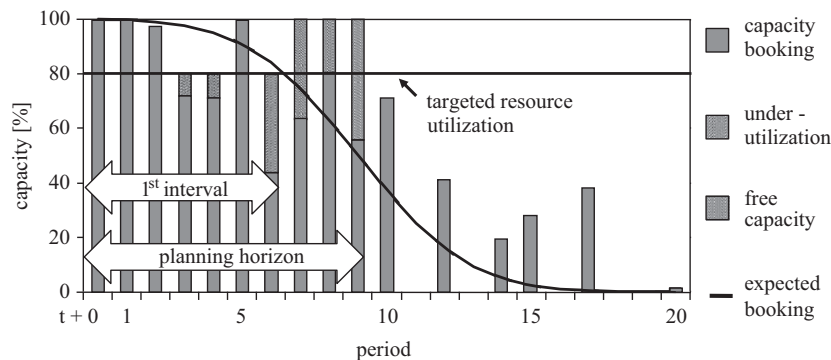
Fig. 3. Structure of the costs of the assignment.

Fig. 4. MPS decision situation.

preliminary production orders Ψ, which is specified in terms of product configuration and production period. This set is subsequently transferred to MPS.

MPS is executed based on rolling horizons, with each planning horizon covering T periods. Due to its interface function between the production system and sales, two criteria need to be considered. Firstly, MPS generates instructions for subsequent planning. One possibility to incorporate leveling is to minimize shortfalls on specified lower levels of capacity utilization $cap_{r\tau}^{min}$. These result from subordinate master production planning and determine efficient operating points of the production system (Holweg and Pil, 2004, p. 33, 4). The same holds true for just in time procurement. Here, the results of master production planning are used to fix average component supply quantities per period. The factual demand may vary in between certain levels around these average quantities. However, if the order quantity falls below the lower level agreed on with the supplier, the OEM has to provide the supplier with appropriate compensation. In the remainder let $ctp_{r\tau}^{-}$ denote the standardized shortfalls on the lower level of the capacity utilization or the minimum component supply per period.

The second criterion to be regarded in MPS results from the interaction with OP. MPS updates the capacities used for promising newly arriving requests. If capacity shortages can be avoided, MPS positively influences the ability to serve newly arriving requests and therewith the response delay. This objective is equivalent to the leveling of the available capacity. Let $ctp_{r\tau}^{+}$ denote the available capacity, standardized with respect to the maximum capacity per resource and period $cap_{r\tau}^{max}$.

In the following we will introduce a segmented objective function to simultaneously incorporate both objectives into MPS.

The modeling approach is illustrated in Fig. 4. Given a stationary ordering process, the booking curve, which quantifies the expected share of orders placed for production in each period $t+\tau$ ($\tau \geq 0$), is strictly decreasing with τ (Meyr, 2004). Each realization of the stochastic ordering process will generally have some degree of variation to this ideal booking curve. In segmenting the MPS objective function, both leveling (1st objective) and the ability to serve newly arriving requests (2nd objective) can be addressed simultaneously. Within the first interval on average most orders have been placed. To include leveling into the objective function, the standardized shortfalls on the targeted resource utilization $ctp_{r\tau}^{-}$ are minimized. Within the second interval the inflow of new customer requests is expected to be higher. For this interval bottlenecks can be avoided by maximizing the unused capacity $ctp_{r\tau}^{+}$.

In addition to the first two, a third criterion is considered to model the coupling of OP and MPS. The costs $c_{i\tau}^{MPS}$ quantify the consequences of bringing forward or delaying an order, given a promised delivery date.

The contribution to leveling and the response delay can hardly be quantified in monetary terms. Therefore, we will make use of the non-monetary indicators $ctp_{r\tau}^{-}$ and $cap_{r\tau}^{max}$. The combination of the two non-monetary criteria and the costs of the assignment into a single objective function requires for the application of a method from multi-criteria decision making. Vast research is available on multi-criteria decision making (Steuer et al., 1996). The most straightforward implementation is the use of simple additive weighting (SAW), which we will adopt in the following. The idea is to use weighting factors to scale the relative importance of the multiple criteria. A complete list of symbols is given in Appendix A.

MPS results into an aggregated production plan. The assignments of the first period of the planning horizon are in the following passed on to sequencing and material requirements planning for further detailing. The particular orders are deleted from the set Ψ. All assignments other than those for the first period are of temporary nature. In the remainder OP will be executed every time an order enters the system (real-time processing) while MPS is executed at the end of each period based on rolling horizons (batch processing).

4.1. Master production scheduling

The mathematical program for MPS is given in (1)–(7). The aim of the model is to assign all orders that have been quoted for production in the planning horizon to production periods. For these orders the quoted period falls into the interval $[t, t+T-1]$. In this, the binary variables $x_{i\tau}^{MPS}$ are set to 1 if order i is assigned to period τ and 0 if otherwise.

$$\text{Minimize} \quad p^{leveling} \cdot \sum_{\tau=t}^{t+k} \sum_{r \in \Omega} ctp_{r\tau}^- - \sum_{\tau=t+k+1}^{t+T-1} \sum_{r \in \Omega} P^{service}(ctp_{r\tau}^+)$$

$$+ \sum_{\tau=t}^{t+T-1} \sum_{i \in \Psi} c_{i\tau}^{MPS} \cdot x_{i\tau}^{MPS} \tag{1}$$

$$\text{s.t.} \quad \frac{cap_{r\tau}^{min} + ctp_{r\tau} - cap_{r\tau}^{max}}{cap_{r\tau}^{min}} = ctp_{r\tau}^- \quad \forall r \in \Omega; \quad \tau = t,\ldots,t+k \tag{2}$$

$$\frac{ctp_{r\tau}}{cap_{r\tau}^{max}} = ctp_{r\tau}^+ \quad \forall r \in \Omega; \quad \tau = t+k+1,\ldots,t+T-1 \tag{3}$$

$$\sum_{i \in \Psi} x_{i\tau}^{MPS} \cdot a_{ir} = cap_{r\tau}^{max} - ctp_{r\tau} \quad \forall r \in \Omega; \quad \tau = t,\ldots,t+T-1 \tag{4}$$

$$\sum_{\tau=t}^{t+T-1} x_{i\tau}^{MPS} = 1 \quad \forall i \in \Psi \tag{5}$$

$$ctp_{r\tau}, ctp_{r\tau}^-, ctp_{r\tau}^+ \geq 0 \quad \forall r \in \Omega; \quad \tau = t,\ldots,t+T-1 \tag{6}$$

$$x_{i\tau}^{MPS} \in \{0,1\} \quad \forall i \in \Psi; \quad \tau = t,\ldots,t+T-1 \tag{7}$$

In order to reflect the divergent requirements discussed above, we implemented two intervals for the objective function: the first interval ranges from t to $t+k$ and the second one from $t+k+1$ to the end of the planning horizon, with $0 \leq k < T-1$. We did not constrain the assignment of the orders to any of these segments, as to take full advantage of the information available. To implement SAW we incorporated the weighting factor $p^{leveling}$ to assess deviations to the lower level of the targeted capacity utilization of resource r in period τ (leveling aspects) and the weighting function $P^{service}(\cdot)$ to assess the available capacity of resource r in period τ. The objective of the corresponding mixed integer program is to minimize the relative shortfalls on the targeted utilization levels $ctp_{r\tau}^-$, weighted by $p^{leveling}$ throughout the interval $[t, t+k]$, and to maximize the unused capacity $ctp_{r\tau}^+$, weighted by $P^{service}(\cdot)$ throughout the interval $[t+k+1, t+T-1]$. The third term of the objective function computes the order related costs of assigning order i to period τ as given by $c_{i\tau}^{MPS}$. These reflect costs that are incurred when the periods of MPS and OP do not coincide (i.e. holding, penalties).

The figures for the first interval result from standardizing shortfalls on the targeted capacity utilization $cap_{r\tau}^{min}$ according to inequalities (2). Those of the second interval are calculated as given by inequalities (3). Constraints (4) assure feasibility with respect to resource capacities, whereas $ctp_{r\tau}$ specifies the prevailing slack. The constraints (5) assure that each order is assigned to

one period, while constraints (6) and (7) define non-negativity and binary coding for the variables. Input from OP is considered by means of the order related costs $c_{i\tau}^{MPS}$, since they depend on the promised period. In return, updated information on the available capacity $ctp_{r\tau}$ is transferred to OP.

4.2. Order promising

The OP model is given by (9)–(13). Each order is individually assigned to a period $[t, t+T^{max}-1]$, such that the associated costs are minimized (9). In this, the length of the planning horizon T^{max} needs to be set sufficiently high, as to allow for at least one feasible assignment. Note that in contrast to MPS only a single order is processed at a time. Combinatorial effects can thus be disregarded. As such, a rule based model description could be used as well. Despite this fact, we make use of the mathematical programming formulation since it is better suited to indicate similarities and interfaces between both models.

The decision is modeled by the binary variables $x_{i\tau}^{OP}$, which are set to 1 if the order i is assigned to the particular period τ and 0, if otherwise. The objective function consists of two terms. The first one quantifies the costs of the assignment as introduced before. It represents the primary objective of OP. In addition to that, a second term is introduced in order to anticipate the calculus of MPS. It rewards a potentially improved leveling of the model-mix. This term is structurally identical to the one used within MPS and regards all periods that fall into the first interval of the MPS objective function. Since the costs of an improved leveling are not readily assessable, a second multi-criteria decision situation results. Again we make use of SAW. To this end, the relative contribution of the particular order to decrease shortfalls on the targeted capacity utilization $ctp_{r\tau}^-$ is weighted by the non-negative coefficient $p^{anticipation}$. Inequalities (10) incorporate resource constraints into the model. In this, the capacity available at the time of the order arrival $ctp_{r\tau}$ is updated with each order being processed. Constraints (11) are computing the contribution of the particular order to decrease shortfalls on the targeted capacity utilization $cap_{r\tau}^{min}$. Only those combinations of resources and periods are considered that exhibit a shortfall greater than 0 at the current execution of OP. These combinations are given by the set Θ

$$\Theta \in \{(r,\tau) | ctp_{r\tau} > cap_{r\tau}^{max} - cap_{r\tau}^{min}; \quad r \in \Omega; \quad \tau = t,\ldots,t+k\}. \tag{8}$$

Eq. (12) assure that the order is assigned to a period. Binary coding is subject to (13).

$$\text{Minimize} \quad \sum_{\tau=t}^{t+T^{max}-1} c_{i\tau}^{OP} \cdot x_{i\tau}^{OP} + p^{anticipation} \cdot \sum_{\tau=t}^{t+k} \sum_{r \in \Omega} ctp_{r\tau}^- \tag{9}$$

$$\text{s.t.} \quad x_{i\tau}^{OP} \cdot a_{ir} \leq ctp_{r\tau} \quad \forall r \in \Omega, \quad \tau = t,\ldots,t+T^{max}-1 \tag{10}$$

$$\frac{x_{i\tau}^{OP} a_{ir}}{cap_{r\tau}^{min}} = ctp_{r\tau}^- \quad \forall (r,\tau) \in \Theta \tag{11}$$

$$\sum_{\tau=t}^{t+T^{max}-1} x_{i\tau}^{OP} = 1 \tag{12}$$

$$x_{i\tau}^{OP} \in \{0,1\} \quad \forall \tau = t,\ldots,t+T^{max}-1 \tag{13}$$

4.3. Planning policies

Despite the fact, that OP and MPS are conceptualized as distinct decision models, both of them interact with respect to the achievement of service and efficiency objectives. As a

consequence issues of coordination arise. We will in the following discuss three planning policies:

a. The *baseline policy*, represented by a capacitated OP without anticipation term ($p^{anticipation} = 0$); MPS is not considered. The OP results are directly passed on to subsequent planning.
b. The *non-reactive policy* distinguishes between models for OP and MPS. Anticipation is not considered ($p^{anticipation} = 0$).
c. The third policy is characterized by the fact, that components of the MPS decision model are considered within the objective function of the OP model ($p^{anticipation} > 0$). Referring to (Schneeweiß, 2003, pp. 42) this corresponds to an *implicit anticipation policy*.

If the anticipation term is considered within the OP model, a multi-criteria decision situation results. An assignment which might be optimal with respect to the costs of the assignment, i.e. the first term, might not be chosen, if the anticipated benefit for the performance of the planning system is greater than the increase in costs. To this end, a benefit is given, if the positive effect with respect to the improved confirmed fill delay and the leveling objective is valued more than the additional costs of the assignment. However, since there is more information available at the time of MPS, these positive effects may only be anticipated. The solution of this trade-off is controlled by the parameter $p^{anticipation}$. In the following we will evaluate the three policies for ODP and the effect of the anticipation term using simulation.

5. Simulation experiments

Our simulation model captures the dynamic interaction of OP and MPS. For the analysis we chose the following approach. First, we investigate the validity of the proposed modeling approach. The modeling approach is regarded valid, if the models developed for OP and MPS are adequate to control the simulation response. If so, the models provide decision support for ODP. A second goal was then to analyze the performance of the presented planning policies. The special focus here was on the effect of the anticipation term.

Each simulation replication is performed in three steps. At first, a demand sequence is generated. More specifically, this sequence comprises a set of orders, which consist of a specific configuration, an order entry date, and a preferred delivery date. In a second step, this demand sequence is subject to the ODP policy to be analyzed. Accordingly, each order is at first promised individually and secondly assigned to a production period in the course of MPS planning executed on rolling horizons (if applicable). In a third step, the results are evaluated. For the implementation we used Plant Simulation and Lingo.

5.1. Experimental data

We consider an autoregressive moving-average (ARMA) demand process to derive the aggregated demand level d_τ of period τ according to $d_\tau = \lceil 50 + 0.8 \cdot (d_{\tau-1} - 50) - 0.1 \cdot \varepsilon_{\tau-1} + \varepsilon_\tau \rceil$ with $\varepsilon_i \sim N(0; 5)$. ARMA processes have been widely used to model time series and avoid the lack of internal consistency attributed to IID-figures (Makridakis et al., 1998; Law and Kelton, 2000). The brackets, $\lceil \rceil$, indicate that we rounded figures to integers. In a second step, we computed requested lead times for each order using a (truncated) normal distribution $N(10; 2)$. Using these lead times each request was back-scheduled to compute the order arrival time. In order to reflect the situation of the automotive industry we grouped the resources into option families. An empirical distribution with specific take rates $take_r$ was used for

each option family to derive the demand for the associated resources (the capacity coefficients a_{ir}). For the analysis we referred to empirical data as depicted in Appendix B (Holweg and Pil, 2004, p. 31). This data results in 576 product configurations.

Each simulation run covered 50 periods such that on average 2500 orders were processed. We set the maximum capacity per resource and period to $cap_{r\tau}^{max} = \lambda \cdot take_r \cdot E[d_\tau]$. In this, the parameter λ will be referred to as the capacity/demand ratio. The lower level of the targeted capacity utilization $cap_{r\tau}^{min}$ was set to 80 percent of the maximal capacity, rounded to integers. We assumed a linear cost function for $c_{i\tau}^{OP}$: each period earlier than the requested added one unit to the costs while each period later than the requested added five units. $c_{i\tau}^{MPS}$ was set equivalent to $c_{i\tau}^{OP}$ for earliness. Tardiness was not allowed for. Hence, we set the associated costs prohibitively high.

A piecewise linear term was implemented for the weighting of the second term of the MPS objective function using the non-negative coefficient $p^{service}$ (14). Accordingly, free capacity is incentivized up to a threshold value α. For the remainder we set α to 30 percent.

$$P^{service}(ctp_{r\tau}^+) = \begin{cases} p^{service} ctp_{r\tau}^+ & \text{for } ctp_{r\tau}^+ < \alpha \\ p^{service} \alpha & \text{else} \end{cases} \quad (14)$$

5.2. Model evaluation

We chose a differentiated approach to evaluate the performance of ODP in order to better understand the prevailing effects. Accordingly, we incorporated three performance measures instead of a single, aggregated one. The customer-oriented performance was evaluated by means of the average costs of the assignment in the course of OP. In addition to that, the under-utilization quantified by the cumulated figure for $ctp_{r\tau}^-$ was evaluated as measure for the compliance with resource-oriented objectives. Finally, we computed the average number of periods in which orders were held on inventory as a consequence of re-assignments within MPS.

Since MPS was executed based on rolling horizons, it would not be meaningful to evaluate the objective function values. We therefore restricted the analysis to the final implementations (i.e. each MPS execution's first period), which corresponds to the instructions passed on to sequencing as discussed in Section 3.

5.3. Experimental design

The experimental study was set up in two stages. The objective of the initial phase was to validate the modeling approach for each capacity/demand ratio. The models are regarded valid, if the model parameters (independent variables) are capable of controlling the simulation response, i.e. the performance of ODP (dependent variable). If this is the case, the parameters can be used to solve the trade-off between service and leveling such that decision support for ODP is provided. We set all parameters except the weighting coefficients used to implement SAW to constants. In particular the structural parameters such as the product definition, take rates, and production and procurement capacities were fixed. This replicates the decision situation in industry, where structural parameters are generally fixed for planning horizons considered within ODP. As a fourth experimental factor the capacity/demand ratio was added, since the performance of ODP was expected to depend on it. We chose values of different order of magnitude for the weighting factors. The resulting models follow the idea of lexicographic programming and therefore allow for a differentiated assessment of the

particular terms. The levels were set in accordance with the results of a pre-test. The factors are summarized in Appendix B.

In a second phase we focused on the comparative evaluation of the three planning policies introduced above. The objective is to assess the potential of more complex mechanisms to coordinate OP and MPS. The planning policy is determined by the weighting factor $p^{anticipation}$. Accordingly, we varied the level of this factor for different capacity/demand scenarios leaving the others unchanged.

This cumulated in 16 configurations (called scenarios in the following) to be analyzed in the first phase and 9 in the second. 50 replications were run for each scenario. Since we expected all performance measures to be correlated with the demand scenario (number of orders, preferred lead time, and model mix), we used a common random number (CRN) approach.

5.4. Results

We used analysis of variance (ANOVA) to identify factors statistically controlling the simulation response. In this, we conducted separate analysis for each performance measure and for each capacity/demand ratio. In order to account for the CRN approach chosen, we added a block variable entitled replication (Rardin and Uzsoy, 2001). The level of significance was chosen to 0.05. Accordingly, all effects are significant which exhibit a p-value lower than the Bonferroni's alpha family (0.05) divided by the number of statistical tests (8).

With respect to the average costs of the assignment, all main effects, except that of $p^{leveling}$ for $\lambda=1.2$, were found to be statistically significant (Table 1). An interaction effect between $p^{leveling}$ and $p^{service}$ was found for $\lambda=1$ while all others proofed to be insignificant. For the average periods orders were held on inventory all main effects but no interaction effect were found to be significant. Considering the under-utilization, we found more differentiated results. For the tight capacity scenario ($\lambda=1$) both MPS weighting coefficients have an effect on either figure, while for $\lambda=1.2$ only the weighting of the leveling term ($p^{leveling}$) exhibits a significant effect. The weighting of the anticipation term ($p^{anticipation}$) does not show any controlling influence in this regard. The high R^2 values indicate a good fit of the models.

Based on the results of the first phase we may conclude that the model parameters significantly control the trade-off between service and leveling. The simulation experiments thus support the validity of the proposed modeling approach.

The performance measures differ significantly with respect to the capacity/demand ratio. If capacity is tight ($\lambda=1$), the costs of the assignment of the baseline policy exceed those if $\lambda=1.2$ by factor 8. The opposite holds true for the under-utilization. This figure on average exceeds that for the tight scenario by factor 2.3. The interpretation of this finding is that it is more likely to achieve leveled schedules, when capacity is tight. Likewise there

is an increased probability of capacity shortages, such that orders can only be satisfied later than requested (resulting in increased assignment costs). On the contrary, holding is more likely to be used when capacity exceeds demand, since more orders might need to be re-assigned to level the schedule. On average, the figures for holding in case of $\lambda=1.2$ exceed those of the tight scenario by factor 2.3.

ODP is able to solve the trade-offs between service, under-utilization and holding, if the directions of the controlling effects complement each other. In line with the objectives of the analysis, ANOVA provides a statistical test to identify the controlling effects of the model parameters on the performance of ODP. To analyze the directions of these effects, additional analysis is necessary. The results of this analysis are summarized in Table 2. Accordingly, $p^{leveling}$ can be used to reduce under-utilization while $p^{service}$ controls the costs of the assignment. In both cases there is a trade-off with holding. The weighting of the anticipation term $p^{anticipation}$ may in turn contribute towards reduced holding, with the drawback of increased costs of the assignment. To summarize, the effects are not merely statistically significant, but also complement each other. The proposed models can thus be used to control the relevant trade-offs of ODP.

From a conceptual point of view, the anticipation term incorporates leveling aspects into OP and should thus be capable of reducing the amount of re-assignments within MPS. More complex ODP policies should therefore be more important, if there is excess capacity. In order to shed light on this aspect, we increased the resolution of the design for the second phase. As to limit the number or experiments, we set the parameters $p^{leveling}$ and $p^{service}$ to a constant ($p^{leveling}=p^{service}=10$). In this phase we compared the performance of the non-reactive policy ($p^{anticipation}=0$) and that of the implicit anticipation policy ($p^{anticipation}=6$ and $p^{anticipation}=10$) with the performance of the baseline policy and computed the 95% confidence interval for the expected mean of the difference for each performance measure (Fig. 5).

As can be seen from the simulation results, increasing the weighting factor contributes towards reduced holding. Accordingly, the number of average periods on inventory decreases

Table 2
Summary of the effects for $k=4$.

Criteria Measure	Leveling Under-utilization	Service Costs of the assignment	Holding Average periods on inventory
$p^{leveling}$	(−)	(−)[a]	(+)
$p^{service}$	(−)[a]	(−)	(+)
$p^{anticipation}$	( )	(+)	(−)

[a] $\lambda=1.0$.

Table 1
ANOVA results for $k=4$. The level of significance is 0.05. The R^2 values for the average costs of assignment are 0.97 for $\lambda=1$ and 0.97 for $\lambda=1.2$. Those for the average periods on inventory are 0.87 and those for the under-utilization are 0.95 for $\lambda=1$ and 0.97 for $\lambda=1.2$. The table is continued in Appendix B.

| λ	Source	Average costs of assignment			Average periods on inventory			Under-utilization		
		Sum of squares	F-value	p-value	Sum of squares	F-value	p-value	Sum of squares	F-value	p-value
1.0	$p^{leveling}$	0.31	9.68	0.00	0.96	43.71	0.00	964.92	90.37	0.00
	$p^{service}$	9.42	294.56	0.00	5.19	236.40	0.00	273.90	25.65	0.00
	$p^{anticipation}$	5.41	169.15	0.00	0.55	24.89	0.00	10.11	0.095	0.33
1.2	$p^{leveling}$	0.004	0.19	0.66	48.08	439.82	0.00	4,821	872.92	0.00
	$p^{service}$	0.65	31.62	0.00	3.56	32.57	0.00	116.42	2.11	0.15
	$p^{anticipation}$	18.31	890.67	0.00	4.49	41.09	0.00	39.77	0.72	0.34
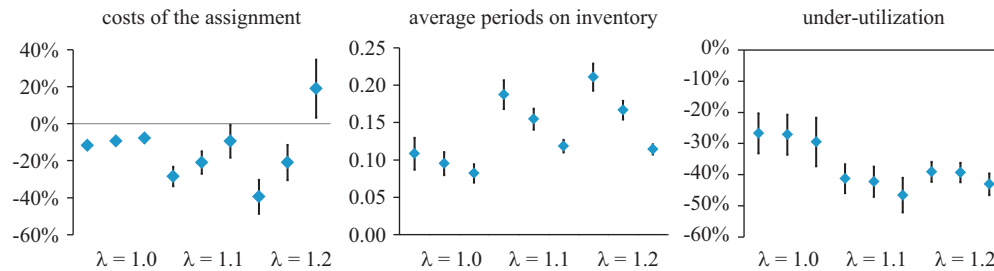
Fig. 5. Confidence intervals for the mean paired difference as compared to the baseline policy for $p^{leveling}=p^{service}=10$ and $k=4$. Figures except the average periods on inventory (measured in periods) indicate the standardized difference with respect to the baseline policy. The first figure of each capacity/demand ratio corresponds to $p^{anticipation}=0$, the second to $p^{anticipation}=6$, and the third to $p^{anticipation}=10$.

(Fig. 5). However, there is a trade-off with the costs of assignment. These increase with an increasing weighting of the anticipation term. Accordingly, the relative advantage over the baseline policy depicted in Fig. 5 decreases. This in particular holds true for high capacity/demand ratios. For $\lambda=1.2$ and $p^{anticipation}=10$ the costs in fact exceeds that of the baseline scenario. With respect to the under-utilization no significant effects were found. Accordingly, the findings from the simulation support the initial proposition: the weighting of the anticipation term can be used to control the trade-off between the costs of the assignment (OP) and holding, yet without significantly influencing the under-utilization.

From a systems perspective, the coordination of OP and MPS is improved, if the decrease in holding results into a net reduction in costs despite the relatively higher costs of the assignment.

A weighting factor of 6, which constitutes an implicit anticipation policy, yields significantly lower figures for expected holding if $\lambda>1.0$ as compared to the non-reactive policy ($p^{anticipation}=0$). For $\lambda=1.1$, there is no significant difference in the costs of the assignment nor in the under-utilization. The implicit anticipation policy therefore improves the coordination of the decisions towards the achievement of the performance figures. For $\lambda=1.2$ no generalized finding is possible without quantifying the costs of holding. This is due to the fact that there is a significant trade-off between holding and the costs of the assignment. For $\lambda=1.0$ no significant difference was found at all.

To conclude, the existence of excess capacity in fact increases the potential of coordination and therefore the impact of more complex ODP policies. However, no general conclusion can be drawn. The advantageousness depends on the value of the weighting parameters, which are subject to the cost structure of the specific industry setting. In the automotive industry, for instance, some OEMs face due date preferences which are rather weak, while system costs related to under-utilization as well as inventory are high. In this case the proposed modeling approach can be used to decrease under-utilization and holding significantly as compared to the baseline policy. To this end, high values should be chosen for $p^{leveling}$ and $p^{anticipation}$.

6. Discussion of results and conclusions

Within this study we developed models for ODP in BTO order fulfillment systems in the automotive industry and evaluated their potential using simulation. Although the scope of the analysis lies on rather general characteristics of automobile production, it captures the key elements of real world order fulfillment systems implemented in APS in both structural as well as logical respects. By using empirical data, insights into the dynamic interaction of OP and MPS could be obtained. These pave the way for a more thorough understanding of how to design both

planning functions. The major findings of the study can be summarized as follows:

- Conceptualizing ODP as a system of two interdependent planning tasks takes the divergent characteristics of the decision situation into consideration. For a setting that captures the key elements of automobile production, model configurations have been developed, which effectively control the performance of the ODP system.
- An anticipation term has been introduced to improve the coordination between OP and MPS. For the study considered, this term can be used to significantly reduce the need for re-assignments within MPS without negatively influencing under-utilization. Yet, there is an intrinsic trade-off with the costs of the assignment within OP.
- As follows, the potential of complex ODP policies is large, if the costs of under-utilization and that of holding are high as compared to the costs of the assignment within OP. In this case it is better to convince the customer of a delivery other than requested, if this allows for a more efficient mode of production or less holding.
- The potential of complex ODP policies likewise increases with the amount of excess capacity. This is due to the fact, that both the degrees of freedom as well as the necessity to level the schedule increase, if there is more capacity than demand.

Simulation has been used to analyze the complex interaction of OP and MPS in the context of BTO order fulfillment. The simulation model captures the key elements of order fulfillment systems used in the automotive industry. To run the simulation, a number of parameters needed to be specified. With the objective to provide in-depth insights into the performance of ODP policies we systematically varied the models and their configuration as well as the capacity/demand ratio. All structural and operational parameters other than that were set to constants. To generalize the insights provided above, future research on the impact of the parameters on the potential of ODP policies is required. In particular, we would expect three parameters to have a significant impact on the performance:

1. The product variety, given by the number and frequency of product options: In the simulation study we used empirical data on the product definition from one selected OEM. However, as reported by Pil and Holweg (2004), companies in the "automotive sector differ dramatically in the level of variety they offered". Future work is necessary to investigate the effect of different strategies of product variety on the potential of ODP.
2. The distribution of the requested lead time. Requested lead times vary in the automotive industry among countries and market segments (e.g. fleet vs. retail). Since longer lead times positively

influence the scope of planning decision within ODP, further work is necessary to generalize the findings reported above.

3. The span between the lower and upper level of the capacity utilization: An increase in the flexibility goes along with a decline in the relative importance of resource-oriented criteria. Since MPS incorporates resource-oriented criteria into planning, this might result into a generally reduced relevance of the MPS planning function as well. Further research is necessary to shed light on this issue.

Due to the reasons elaborated on above, we implemented a hierarchical approach. Incorporating more details from MPS into OP would result into a higher model complexity and thus extended quotation times. This would likewise increase the probability of two or more customer request entering the system simultaneously, which in turn causes further delays due to sequential processing or inconsistencies in the case of parallel processing (SAP, 2005b). In introducing separate models for OP and MPS, our approach guarantees for short response times, while ensuring feasible production plans. The potentials of the approach are illustrated by the results provided above. The decomposition of order-driven planning into two distinct planning tasks, however, gives rise to the question of how to adequately coordinate these interdependent decisions. Referring to the classification scheme introduced by Schneeweiß (2003) the coupling conditions employed can be described as non-reactive and implicit anticipation. Further work is necessary to evaluate the potentials that might be achieved by more sophisticated coupling mechanisms.

According to the analysis, the dynamic performance of order-driven planning routines is significantly controlled by the configuration of the mathematical programs. In using SAW, the developed objective functions allow for the simultaneous optimization of resource and customer-specific and thus originally conflicting objectives. From a conceptual point of view, the use of SAW gives rise to further questions. Firstly, we used identical weighting factors in our analysis. In practical applications there might be differences in the relative importance of resources or orders. This does for instance hold true, if it is not equally important to achieve leveled master schedules for all resources of the production system. As a consequence, the adaption of a more differentiated modeling of the weighting terms might be required. Since these parameters constitute exogenous input for the presented approach this does, however, not confine the applicability of the approach. Secondly and more importantly, the use of parameters within the models gives rise to the more fundamental question of how to identify parameter settings which adequately solve the trade-off between the performance measures for particular industrial settings. The main obstacle with this approach is that the evaluation of each configuration requires executing multiple simulation runs. Even if the analysis is restricted to discrete parameter values which are kept identical for all resources, a combinatorial optimization problem has to be solved. Therefore the availability of efficient search routines is essential. Techniques from simulation optimization or artificial intelligence seem to be promising.

While there is more work needed to exploit the full potentials of the approach, the paper shows that the presented framework and the developed models can be used to improve order-driven planning in BTO automobile production. This lays the basis for the advancement of order-driven planning systems for this industry.

Appendix A. List of symbols

Indices and index sets

i	customer requests/orders ($i = 1,2,\dots,I$)
t	subordinate planning periods ($t = 1,2,\dots,T^{max}$)

τ	planning periods considered for OP and MPS
r	resource index
Ω	index set of the resources
Ψ	index set of accepted orders within the MPS planning horizon
Θ	tuples of resources and periods $(r;\tau)$ that exhibit a shortfall to the lower level of the capacity utilization greater than 0

Parameters

T	length of planning horizon within MPS
T^{max}	length of planning horizon within OP
a_{ir}	production coefficient of order i with respect to resource r
$c_{i\tau}^{OP}$	costs of the assigning order i to period τ within OP
$c_{i\tau}^{MPS}$	costs of assigning order i to period τ within MPS
$cap_{r\tau}^{max}$	maximal available capacity of resource r in period τ
$cap_{r\tau}^{min}$	lower level for the capacity utilization of resource r in period τ
k	interval parameter for the MPS objective function ($0 \le k \le T-1$)
$p^{leveling}$	weighting factor for standardized deviations to the lower level of the capacity utilization
$P^{service}(\cdot)$	weighting function for the standardized capacity available $ctp_{r\tau}^{+}$
α	parameter used for the piecewise linear term within $P^{service}(\cdot)$
$p^{service}$	weighting factor used within $P^{service}(\cdot)$
$p^{anticipation}$	weighting factor for the anticipated contribution to the standardized deviations to the lower level of the capacity utilization $ctp_{r\tau}^{-}$
λ	capacity/demand ratio

Decision variables

$$x_{i\tau}^{OP} = \begin{cases} 1 & \text{if order } i \text{ is assigned to period } \tau(\text{OP}), \\ 0 & \text{else} \end{cases}.$$

$$x_{i\tau}^{MPS} = \begin{cases} 1 & \text{if order } i \text{ is assigned to period } \tau(\text{MPS}), \\ 0 & \text{else} \end{cases}.$$

$ctp_r\tau$	unassigned capacity of resource r in period τ
$ctp_{r\tau}^{-}$	standardized deviations to the lower level of the capacity utilization $cap_{r\tau}^{min}$ of resource r in period τ
$ctp_{r\tau}^{+}$	capacity available of resource r in period τ standardized with respect to the maximal available capacity $cap_{r\tau}^{max}$

Appendix B

See Tables 3–5.

Table 3
Factorial design (figures in brackets indicate second phase).

Factor	Denomination	Levels	Degrees of freedom
Weighting of the leveling term	$p^{leveling}$	10, 10000	1
Weighting of the service term	$p^{service}$	10, 10000	1
Weighting of the anticipation term	$p^{anticipation}$	0,(6),10	1 (2)
Capacity/demand ratio	λ	1.0, (1.1), 1.2	1 (2)

Table 4
Option families, options, and take rates.

Resource r	Option family	Option family name	Options			Take rates $take_r$		
1…3	1	Transmission	Manual	Automatic	–	0.92	0.08	0
4…6	2	Engine	Gasoline	Diesel	Turbo diesel	0.50	0.45	0.05
7…9	3	Chassis	Three door	Five door	–	0.60	0.40	0
10…12	4	Trim	Standard	Special	GTI	0.82	0.10	0.08
13…15	5	ABS	True	False	–	0.66	0.34	0
16…18	6	Air condition	True	False	–	0.66	0.34	0
19…21	7	Sunroof	True	False	–	0.25	0.75	0
22…24	8	Lateral airbag	True	False	–	0.08	0.92	0

Table 5
ANOVA results (continued).

λ	Source	Average costs of assignment			Average periods on inventory			Under-utilization		
		Sum of squares	F-value	p-value	Sum of squares	F-value	p-value	Sum of squares	F-value	p-value
1.0	$p^{leveling} \times p^{service}$	0.318	9.96	0.002	0.079	3.58	0.059	136.45	2.47	0.117
	$p^{leveling} \times p^{anticipation}$	0.019	0.59	0.444	0.028	1.27	0.261	77.55	1.40	0.237
	$p^{service} \times p^{anticipation}$	0.068	2.12	0.147	0.099	4.53	0.034	3.18	0.06	0.810
	Model	330.44	184.52	0.000	49.85	40.56	0.000	6.81	106.72	0.000
	Replication	314.90	196.94	0.000	42.95	39.14	0.000	62.505	117.08	0.000
1.2	$p^{leveling} \times p^{service}$	0.019	0.92	0.337	0.126	1.15	0.284	42.13	3.95	0.048
	$p^{leveling} \times p^{anticipation}$	0.036	1.75	0.187	0.236	2.16	0.143	7.72	0.72	0.396
	$p^{service} \times p^{anticipation}$	0.080	3.91	0.049	0.250	2.29	0.131	4.79	0.45	0.503
	Model	78.446	68.15	0.000	243.31	39.75	0.000	572.99	185.28	0.000
	Replication	59.350	57.75	0.000	186.57	34.14	0.000	524.41	189.92	0.000

References

Abraham, S.J., 2002. Aspiring for Automotive Excellence: The 5-Day Car.

Alford, D., Sackett, P., Nelder, G., 2000. Mass customisation—an automotive perspective. International Journal of Production Economics 65 (1), 99–110.

Andres, M., 2006. Die optimale Varianz. Brandeins 1, 65–69.

Bolat, A., 2003. A mathematical model for selecting mixed models with due dates. International Journal of Production Research 41 (5), 897–918.

Boysen, N., 2009. Production planning of mixed-model assembly lines. Production Planning & Control 20 (5), 455–471.

Brabazon, P.G., MacCarthy, B., 2006. Fundamental behaviour of virtual-build-to-order systems. International Journal of Production Economics 104 (2), 514–524.

Chen, C., Zhao, Z., Ball, M.O., 2001. Quantity and due date quoting available to promise. Information Systems Frontiers 3 (4), 477–488.

Chen, C., Zhao, Z., Ball, M.O., 2002. A model for batch advanced available-to-promise. Production and Operations Management 11 (4), 424–440.

Da Silveira, G., Borenstein, D., Fogliatto, F.S., 2001. Mass customization. International Journal of Production Economics 72 (1), 1–14.

Ding, F., Tolani, R., 2003. Production planning to support mixed-model assembly. Computers Industrial Engineering 45 (3), 375–392.

Engel, C., Zimmermann, J., 1998. Objectives for order-sequencing in automobile production. In: Drexl, A., Kimms, A. (Eds.), Beyond Manufacturing Resource Planning. Springer, Berlin, pp. 307–331.

Fleischmann, B., Meyr, H., 2004. Customer orientation in advanced planning systems. In: Dyckhoff, H., Lackes, R., Reese, J. (Eds.), Supply Chain Management and Reverse Logistics. Springer, Berlin, pp. 297–321.

Ford, H., 1988. Today and Tomorrow (special reprint). Productivity Press, Portland, OR.

Hindi, K.S., Ploszajski, G., 1994. Formulation and solution of a selection and sequencing problem in car manufacture. Computers Industrial Engineering 26 (1), 203–211.

Holweg, M., Disney, S.M., Hines, P., Naim, M.M., 2005. Towards responsive vehicle supply. Journal of Operations Management 23 (5), 507–530.

Holweg, M., Miemczyk, J., 2002. Logistics in the three-day car age. International Journal of Physical Distribution and Logistics Management 32 (10), 829–850.

Holweg, M., Pil, F.K., 2004. The Second Century, Cambridge, Mass. MIT Press.

Jeong, B., Sim, S., Jeong, H., Kim, S., 2002. An available-to-promise system for TFT LCD manufacturing in supply chain. Computers Industrial Engineering 43 (1), 191–212.

Jiao, J., Ma, Q., Tseng, M.M., 2003. Towards high value-added products and services: mass customization and beyond. Technovation 23 (10), 809–821.

Kotler, P., 1989. From mass marketing to mass customization. Planning Review 17 (5), 10–13.

Krog, E., 2006. Das kundengesteuerte Unternehmen von der Bestellung bis zur Auslieferung, vol. 3. Branchenforum Automobil-Logistik, Stuttgart.

Law, A.M., Kelton, W.D., 2000. Simulation Modeling and Analysis. McGraw-Hill, Boston.

Lee, H., Peleg, B., Whang, S., 2005. Toyota: Demand Chain Management.

Makridakis, S.G., Wheelwright, S.C., Hyndman, R.J., 1998. Forecasting: Methods and Applications. Wiley, New York.

Meyr, H., 2004. Supply chain planning in the German automotive industry. OR Spectrum 26 (4), 447–470.

Meyr, H., 2004. Die Bedeutung von Entkopplungspunkten für die operative Planung von Supply Chains. Zeitschrift für Betriebswirtschaft 73 (9), 941–962.

Moses, S., Grant, H., Gruenwald, L., Pulat, S., 2004. Real-time due-date promising by build-to-order environments. International Journal of Production Research 42 (20), 4353–4376.

Pibernik, R., 2005. Advanced available-to-promise. International Journal of Production Economics 93 (1), 239–252.

Pil, F.K., Holweg, M., 2004. Linking Product variety to order-fulfillment strategies. Interfaces 34 (5), 394–403.

Rardin, R.L., Uzsoy, R., 2001. Experimental evaluation of heuristic optimization algorithms. Journal of Heuristics 7 (3), 261–304.

Raaymakers, W.H., Bertrand, J.M., Fransoo, J.C., 2000. Using aggregate estimation models for order acceptance in a decentralized production control structure for batch chemical manufacturing. IIE Transactions 32 (10), 989–998.

Robinson, A.G., Carlson, R.C., 2007. Dynamic order promising: real time ATP. International Journal of Integrated Supply Management 3 (3), 283–301.

SAP, 2005a. "Managing the e-supply chain" in der Automobilindustrie mit mySAP Automotive Walldorf.

SAP 2005b. Capable To Promise, OSS Notes 426563.

Schneeweiß, C.A., 2003. Distributed decision making. Springer, Berlin.

Steuer, R.E., Gardiner, L.R., Gray, J., 1996. A bibliographic survey of the activities and international nature of multiple criteria decision making. Journal of Multi Criteria Decision Analysis 5 (3), 195 217.

Vollmann, T.E., Berry, W.L., Whybark, D.C., Jacobs, F.R., 2005. Manufacturing Planning and Control Systems for Supply Chain Management. McGraw-Hill, New York, NY.

Wester, F.A.W., Wijngaard, J., Zijm, W.H.M., 1992. Order acceptance strategies in a production-to-order environment with setup times and due-dates. International Journal of Production Research 30 (6), 1313–1326.

ten Kate, H.A., 1994. Towards a better understanding of order acceptance. International Journal of Production Economics 37 (1), 139–152.