11 Purchasing and Material Requirements Planning

Hartmut Stadtler

An indispensable part of an ERP system, Material Requirements Planning, also plays an important role in APS, because it

- Generates replenishment orders (production orders) for uncritical components and parts (operations) in a multi-stage production environment (Sects. 11.1 and 11.2)
- Provides access to a transactional ERP system and thus can initiate the execution of orders.

The typical tasks of purchasing are to analyze procurement markets, to negotiate the terms of trade with potential suppliers and finally to select suppliers and to place replenishment orders. Here, we are interested in the way APS can support the selection of suppliers and the decisions on order sizes, taking into account the specific cost functions of suppliers, which often allow for quantity discounts (Sect. 11.3). This may apply to input materials for production, indirect materials and articles of merchandise.

11.1 Basics of Material Requirements Planning

Material Requirements Planning (MRP) is regarded as the core engine of an ERP system, which calculates time-phased plans of secondary demands for components and parts based on a time series of primary demands (usually finished products). Time-phased secondary demands are a prerequisite for generating production or replenishment orders so that demands for finished products can be met in time with as little work-in-process and inventory as possible.

H. Stadtler (✉)
Institute for Logistics and Transport, University of Hamburg, Von-Melle-Park 5, 20146 Hamburg, Germany
e-mail: hartmut.stadtler@uni-hamburg.de

H. Stadtler, C. Kilger and H. Meyr (eds.), *Supply Chain Management and Advanced Planning*, Springer Texts in Business and Economics,
DOI 10.1007/978-3-642-55309-7_11,

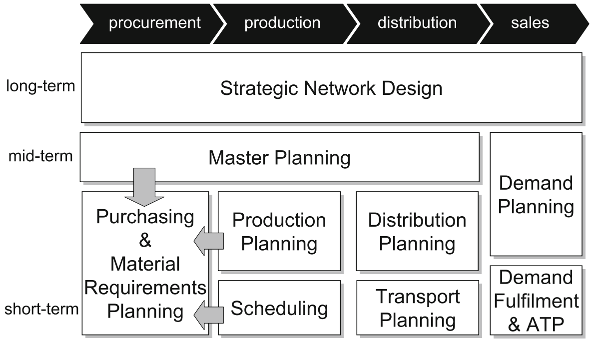


Fig. 11.1 Modules providing the input data (production quantities) for Purchasing and MRP

Although most appealing, this logic suffers from ignoring available capacities. Consequently, production orders may result in overloaded capacities and thus infeasibilities. Experience has shown that a two step procedure, i.e. first calculating all secondary demands and then balancing capacities by means of an ERP's capacity requirements planning (CRP) module, does not provide satisfactory solutions (for a further discussion of the drawbacks of ERP systems see Drexl et al. 1994 or Tempelmeier and Derstroff 1996).

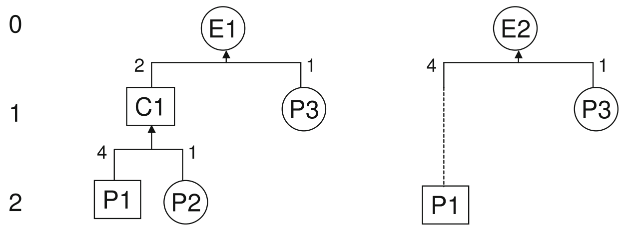
These drawbacks gave rise to develop APS, which do not separate the generation of secondary demands and capacity balancing. However, in order to reduce complexity, APS concentrate on operations to be performed on potential bottlenecks, which usually are only a small subset of all operations relating to factory orders. The time needed to execute non-bottleneck operations (including transport) in between two adjacent *critical operations* is taken into account by a fixed lead-time offset. Once plans have been generated for critical operations, the timing and quantities of non-critical operations can be calculated easily by making use of the standard MRP logic. This is the topic of the next subsection.

There are many textbooks that describe the MRP logic (e.g. Silver et al. 1998; Vollman et al. 1997). Thus we will only briefly describe the terms and the basic logic. More important is a discussion of issues occurring when using MRP in conjunction with an APS.

First of all, we have to decide on the time series of primary demands to take as a starting point. These may be (see Fig. 11.1)

- Production quantities per period for (critical) product groups calculated in Master Planning (see Chap. 8)
- Production quantities per period for critical operations calculated in the Production Planning module or
- Critical production orders generated in the Scheduling module (see Chap. 10).

In case we look for the requirements of parts to be purchased from outside suppliers over a longer period of time (e.g. for negotiating contracts with suppliers or providing an outlook of expected part demands to suppliers), Master Planning will be the starting point. Note that demands for product groups have to be disaggregated into demands of respective products before starting the MRP logic.



Explanations:
– E1, E2 represent end products, C1 a component and P1, P2, P3 single parts
– single digits indicate production coefficients
– materials in circles are regarded as *critical*, materials in boxes as *uncritical*

Fig. 11.2 Bill of materials for end products E1 and E2 as well as low-level codes

For placing replenishment orders or for the timing of uncritical operations (production orders), either Production Planning or Scheduling will be the source of information. If Production Planning is chosen, demands per time bucket will result, while Scheduling will give the exact timing of the start of production orders. Hence, Scheduling best corresponds to a bucketless (continuous time axis) MRP, while the two former are best suited for a bucket oriented MRP logic. Both time axes are possible today (Vollman et al. 1997, p. 30). In the following, we assume Production Planning to be the starting point.

As additional data we will need:

- Bill of materials, indicating for each part number, which other part numbers are required as direct inputs
- Production coefficients indicating the quantity of each direct input part needed for one unit of a given part number
- Lead-times representing a fixed interval of time needed between releasing an order for a part number and its availability
- The inventory status, indicating for each part number, the (physical) stock at hand, scheduled receipts (i.e. outstanding orders and work-in-process), reservations, backorders and safety stock levels
- Low-level code (numbers).

A *low-level code* of a part number or operation corresponds to the longest path in the product structure starting with an end item and terminating in the respective part number. All parts visited along the path are counted yielding the level code. Due to the fact that a part number may be used in several product structures, the maximum has be taken for determining the low-level code. By definition, a low-level code "0" is attributed to end items (for an example see Fig. 11.2). Low-level codes have to be calculated preceding the bill of materials (BOM) explosion, i.e. the generation of secondary demands, to allow a pure sequential execution of calculations.

While in standard text books on MRP the level of detail for a BOM explosion is finished products, components or parts, the level of detail required in the context of APS is *operations*. Normally, several operations are required to transform input material(s) into a specific part. Some of these operations may be critical, i.e. they have to be performed on a potential bottleneck resource, some are uncritical. Consequently, we will have to combine the BOM with the routing of operations—sometimes called the bill of capacities (BOC) (Vollman et al. 1997, p. 128).

To ease understanding we will simplify matters (without loss of generality) by assuming that there is exactly one operation to a finished product, component or part.

11.2 Generation and Timing of Uncritical Orders

The generation of uncritical orders originating from production orders scheduled on bottleneck resources will be explained now by an example. Firstly, the data required—like the BOM—will be presented (see Fig. 11.2). Secondly, some remarks on the generation of a production plan will follow and thirdly, we will show how to derive orders for uncritical operations. Fourthly, a simplification is shown as proposed by APS vendors today.

E1 and E2 are completed on a highly utilized assembly line. Component C1 is produced in a manufacturing cell. Since the manufacturing cell is underutilized if only C1 is produced, surplus capacity has been sold to a partner company. The terms of the contract establish priorities for scheduling operation C1; hence, the manufacturing cell is no bottleneck. P1 is bought from an external supplier, while P2 and P3 are processed on an injection moulding machine which is a potential bottleneck, too.

Consequently, E1, E2, P2 and P3 are regarded critical operations for which a production plan is generated by the APS module Production Planning.

In addition to the data shown in Fig. 11.2 lead-time offsets are needed for each operation. For the example presented here we assume one period except for C1, which has a lead-time of two periods.

While lot-sizing plays a major role for critical operations, incurring setup times or setup costs on potential bottlenecks, this is generally negligible on non-bottlenecks. Since time is not scarce at non-bottlenecks, an hour saved by saving setup time is of no value. Hence, a lot-for-lot production, i.e. no lot-sizing, for non-critical operations is advisable. Exceptions may only occur in case of technological reasons relating to production or transport activities requiring some minimum quantity or integer multiple of a fixed amount to work properly (e.g. production in full tub loads). Often companies make use of fixed lot sizes based on the economic order quantity. Note that these lot sizes should not be regarded as strict instructions because even a significant deviation will increase total variable cost only marginally. For example assume an optimal lot size (Q) corresponding to a time between orders (TBO) of 5 weeks. Then we can choose a lot size in the range of $[0.25 \cdot Q, 4 \cdot Q]$ with an increase in total variable cost of at most 1 %. This result is based on the assumption that holding cost consist of the interest paid on the lot size stock and an interest rate of 10 % per year. Further findings including general formulas are presented in (Stadtler 2007b).

material	period	1	2	3	4	5
E1	demands	30	20	30	20	30
	starting inv.	*40*	10	-	-	-
	order	10	30	20	30	-
E2	demands	20	-	20	-	30
	starting inv.	*20*	-	-	-	-
	order	10	10	20	10	-

Fig. 11.3 Primary demands and production plans for E1 and E2 (in quantities per period; inventory abbreviated by inv.)

In contrast to lead-times used in an ERP system, which usually incorporate a large portion of waiting times, lead-times in the context of an APS pertaining to uncritical operations should only cater for production and transport activities. The reason is that, by definition, utilization rates of non-bottlenecks are low and thus a production order should find the resource empty in general. However, it seems wise to include "some" safety time into the lead-time offset of an uncritical operation being a *direct predecessor* of a critical operation. This will allow for some uncertainties in processing times and will make sure that a bottleneck resource, which governs the throughput of the whole supply chain, will not run empty. Another reason why an APS can do with smaller lead-times than an ERP system (and thus smaller planned throughput times) is due to the fact that lead-times in an ERP system also cater for its inability to take into account finite capacity checks of bottleneck resources when making the BOM explosion. However, in order to avoid an overlap of two adjacent operations—which might cause infeasibilities when it comes to Scheduling—an operation's minimum lead-time should be set to one period.

From these lead-times now *cumulated lead-times* have to be calculated relating two adjacent critical operations simply by adding the single lead-times of operations along the path (in the BOM) from the upstream critical operation to the downstream critical operation—excluding the lead-time of the upstream critical operation. Thereby, the *finishing* point (period) of the downstream critical operation is connected with the *finishing* point (period) of the upstream critical operation. Consequently, cumulated lead-times cover production times and transport activities in between two critical operations plus the lead-time of the downstream critical operation (e.g. cumulated lead-times for E1-P2, E1-P3 and E2-P3 are 3, 1, and 1 period(s), respectively). These cumulated lead-times, as well as (cumulated) production coefficients, primary demands and the inventory status of items, parts, and components form the input to Production Planning.

Figure 11.3 shows the primary demands for finished products E1 and E2 (critical operations) and resultant production orders to meet demands for the upcoming five periods, while taking into account a lead-time offset of one period (see solid arrows). This production plan has been generated assuming that operations E1 and E2 are produced on the same machine with a capacity of 40 units per period and that productions coefficients are "1". Note, that some demands are fulfilled from initial inventory (dashed arrows).

LLC	Operation		demand/order per period 1	2	3	4	5
0	E1	order	10	30	20	30	-
0	E2	order	10	10	20	10	-
1	C1	starting inv.	80	60	-	-	-
		gross dem.	20 (E1)	60 (E1)	40 (E1)	60 (E1)	-
		net dem.	-	-	40	60	-
		order	40	60	-	-	-
2	P1	starting inv.	200	-	-	-	-
		gross dem.	40 (E2)	40 (E2)	80 (E2)	40 (E2)	-
			160 (C1)	240 (C1)	- (C1)	- (C1)	-
		net dem.	-	280	80	40	-
		order	280	80	40	-	-

Explanations:
– LLC: low-level code
– inv.: inventory

Fig. 11.4 BOM explosion with pegging

Positive lead-times are the reason why there are no production orders for E1 and E2 in period five even though the forecast and planning horizon is five periods. Similarly, even for materials with a low-level code greater than “0” production orders cover a smaller interval of time. Consequently, utilization rates near the planning horizon should be interpreted with caution. Furthermore, it becomes clear that a reasonable planning horizon for Production Planning should at least cover the longest path, with respect to lead-times, from a final operation (finished product) to a part with no direct predecessor in the BOM. In our example, the longest path is E1-C1-P1 or E1-C1-P2, both with an overall lead-time offset of four periods. An appropriate planning horizon should also cover a (small) *frozen horizon* and some periods for decision making (e.g. for making lot-sizing decisions).

To keep our example small production plans for critical operations P2 and P3 are not exhibited here, because they don’t cause secondary demands. Now we are in the position of calculating the time-phased order sizes of uncritical operations C1 and P1.

Here, the logic of a time-phased BOM explosion (Orlicky 1975; Tempelmeier 2006) has to be slightly adapted. First, finished products (i.e. final operations) are always declared “critical”. Second, all orders for critical operations *and* possessing at least one uncritical direct predecessor (i.e. upstream) operation, are labeled with low-level code “0”. Now we can start with any operation belonging to low-level code “0” and derive the associated secondary demands for all its uncritical direct predecessor operations by multiplying a period’s order size (e.g. generated in Production Planning) by the production coefficient and placing it in the same time period; e.g. the order for operation C1, for 20 units, must be ready at the beginning of period 1 in order to be used for the assembly operation E1 in period 1 (see Fig. 11.4). In order to know which operation caused the secondary demand we further store its name—(see the operation’s names in brackets in Fig. 11.4). This identification is called *pegging* and can be most useful in the case that operations are not ready in time. Then, it is easy to see which orders are affected and thus specific counter actions can be initiated.

Once direct secondary demands have been calculated for all low-level code "0" operations, then secondary demands of low-level code "1" operations are complete. Next, we can calculate orders for any low-level code "1" operation and explode these into the secondary demands of its direct predecessors. This is only necessary for uncritical direct predecessors, because a production plan exists for the critical operations. (However, a BOM explosion into critical operations may also be useful in order to check the feasibility of the production plan. In case there is a mismatch of orders between the production plan and the BOM explosion, an alert should be generated automatically.)

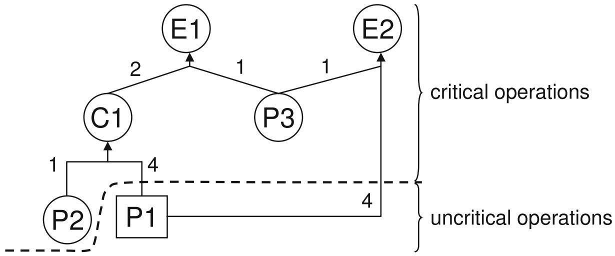
Before starting the BOM explosion, we will have to calculate net demands by netting gross demand with initial inventory. This logic may be more elaborate than shown in our example by considering safety stock requirements, outstanding orders and reservations, too. Given the net demands of an operation these have to be time-phased and assigned to an order period by taking into account the operation's lead-time offset (indicated by an arrow in Fig. 11.4). These tasks are repeated until all operations have been considered.

One may ask what reasons there are for generating an alert during the BOM explosion. Obviously, if we started from an infeasible production plan, e.g. with backlogging, then the BOM explosion would also generate alerts showing that some materials are not ready in time. At this stage a popular counter measure would be *expediting*, resulting in reduced lead-times. A second reason for a mismatch of a (feasible) production plan and the result of a BOM explosion may be that lead-times used in Production Planning are independent of the amount produced, while in a BOM explosion lead-times can be calculated based on the order size. Again, any discrepancy jeopardizing efficiency or feasibility should be shown to the decision maker by an alert.

While the logic of the BOM explosion is rather simple, implementing the interface between the Production Planning module and the MRP module may be tricky. One issue is the generation and exchange of alerts between modules.

In order to avoid the complexity of an arbitrary mix of critical and uncritical operations some APS vendors propose a distinct separation: The final operation, resulting in a finished good, is always defined as critical. Also, any upstream operation can be defined as critical. However, a critical operation may never possess a direct uncritical downstream operation. This can best be illustrated by our example (Fig. 11.2) transformed into a Gozinto graph (Fig. 11.5). Here, a separation line divides operations into the set of critical operations and the set of uncritical operations.

The advantage is that Production Planning can be executed first, followed by the BOM (or BOC) explosion for uncritical operations—and one can be sure that both plans will match. Hence, an exchange of alerts between modules is unnecessary. Also, there is no need to calculate, maintain and use cumulated lead-times or cumulated production coefficients. The disadvantage is that some formerly uncritical operations now have to be declared as critical (e.g. C1), which increases the scope and efforts of Production Planning. Especially, if the most upstream



Explanations:
– E1, E2 represent end products, C1 a component and P1, P2, P3 single parts
– single digits indicate production coefficients
– materials in circles are regarded as *critical*, materials in boxes as *uncritical*
– the dashed line separates critical from uncritical operations

Fig. 11.5 Gozinto representation of the bill of materials with a separation line for the set of critical and the set of uncritical operations

LLC	Operation		demand/order per period 1	2	3	4	5
0	E1	order	10	10	20	10	-
0	C1	order	40	60	-	-	-
1	P1	starting inv.	200	-	-	-	-
		gross dem.	40 (E2)	40 (E2)	80 (E2)	40 (E2)	-
			160 (C1)	240 (C1)	- (C1)	- (C1)	-
		net dem.	-	280	80	40	-
		order	280	80	40	-	-

Explanations:
– LLC: low-level code
– dem.: demand
– inv.: inventory

Fig. 11.6 BOM explosion with pegging

operations are processed on a bottleneck resource then (nearly) all operations in the BOC have to be defined as critical.

Referring to our example, the generation of purchase orders for P1 now starts from production orders for E2 *and* C1 (see Fig. 11.6). For simplification purposes, we assume here that production orders for C1, generated by Production Planning, are equal to those derived by the BOM explosion (Fig. 11.4). Now, applying the BOM explosion for P1 provides the same results as before. The only difference is that computational efforts will be smaller, while they will be larger for Production Planning (not shown here).

Given that the production plan started from is feasible and no alerts have been generated during the BOM explosion, then all production orders for critical and uncritical operations are known and can be handed over for execution (at least for the upcoming period, see Chap. 4). The only exception are purchase orders to outside suppliers which may need further attention due to fixed ordering costs or quantity discounts—which will be dealt with next.

11.3 Quantity Discounts and Supplier Selection

Life cycle contracts are predominant today in many industries for the most important production input. Also, materials to be purchased and considered strategically important are usually procured from a supply chain partner. However, there are a number of additional materials, which are purchased from outside suppliers, where it may be economical to select a supplier and to decide on the order size in the short term and to make use of quantity discounts. These materials may be commodities used as direct production input, often classified as C items, as well as materials for maintenance, repair and overhaul (MRO). In the case of a commodity, quality is also defined by industry standards and there are usually a number of suppliers to choose from. Also, it can be assumed that the quantity to be purchased is rather low compared to the overall market volume so that availability is no problem. Examples are standard electronic components, like a capacitor, or office equipment bought with the help of an e-catalog.

In an abstract form the procurement decision incorporates the following features (Tempelmeier 2002): For each item to be purchased there is a time series of demands over a finite planning interval (e.g. see row “order” for item P1, Fig. 11.4). There may be one or several suppliers to choose from, each with specific costs. These costs will incur

- Supplier specific fixed ordering and procurement costs (including the transport of the consignment)
- Supplier specific quantity discounts (either all-units or incremental discounts).

Figure 11.7 illustrates the two most popular forms of quantity discounts.

Here, the supplier’s fixed ordering cost is depicted as “U” on the total acquisition cost axis. The x-axis represents the order quantity. There are three purchasing intervals, each with a specific price per unit. In the *all-units discount* case, the price charged for the last unit ordered also holds for the total order quantity. In an *incremental discount* case, only those units falling within a purchasing interval are charged with the corresponding price (see lower bounds Q_1 and Q_2 of purchasing intervals 2 and 3 in Fig. 11.7). In both cases it is wise to stick to one supplier and item per period and not to split the order, because this will result in the lowest total acquisition cost. Only if the amount ordered exceeds the maximum a supplier is able to procure (Q_3) another supplier will come into play.

In general, the demand of several periods will be combined when forming purchase orders in order to make use of attractive price reductions for a large quantity. Large order quantities usually result in holding stocks for some periods; thus, holding costs counteract savings due to quantity discounts.

Note that it might be difficult to specify an item’s “correct” holding cost per period because a large portion of the holding cost is interest on the capital employed. Since an item’s purchase price can change over time—especially if there are time-dependent, supplier-specific quantity discounts—one does not know in *advance* which items will be in inventory and at which price. One way to overcome this “problem” is to keep track of each item purchased, its purchase price, purchasing period and the period of consumption.

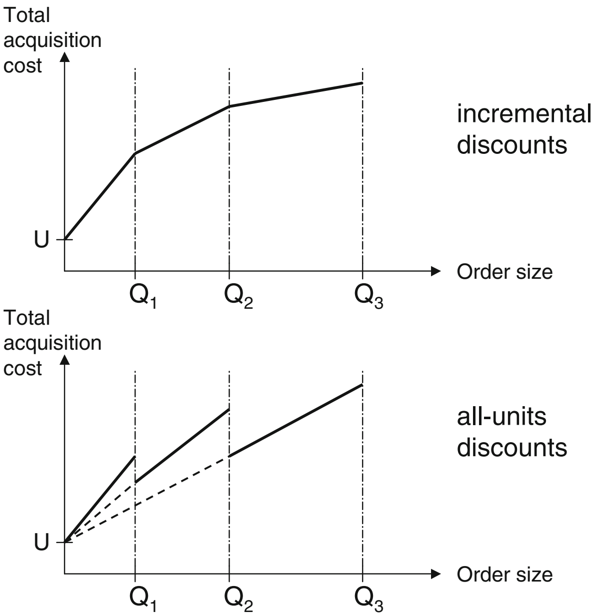


Fig. 11.7 Incremental discounts and all-units discount with three purchasing intervals

Table 11.1 Conditions for purchasing item P1 from two suppliers

Supplier s	Discount	Fixed cost U_s	$p_{1,s}$	$Q_{1,s}$	$p_{2,s}$	$Q_{2,s}$	$p_{3,s}$	$Q_{3,s}$
1	All-units	100	8.00	200	7.80	400	7.60	$+\infty$
1	Incremental	50	7.90	300	7.50	500	7.20	1,000

In a practical setting, one often has to take into account supplier-specific lead-times, delivery schedules or minimum order quantities. Also, if several items are bought from one supplier and procured by a single consignment, fixed ordering costs may be shared among these items. Even more, discounts may be granted for total purchases of a group of products (see Degraeve et al. 2005).

A simple example is constructed to illustrate the decision situation: Let us assume that item P1 can be purchased from two suppliers ($s = 1, 2$). One supplier is offering all-units and the other incremental discounts (Table 11.1). There are three purchasing intervals ($v = 1, 2, 3$) for each supplier s with prices $p_{v,s}$.

Some additional remarks are necessary regarding the time series of demands generated by the BOM explosion. Namely, we require a reasonable number of period demands covering a planning interval that allows for the exploitation of quantity discounts. Also, the first replenishment decision should not be influenced by the target inventory at the planning horizon (usually set to the safety stock level). A rough rule of thumb is a planning interval covering five ordering decisions (i.e. $5 \cdot \text{TBO}$).

Table 11.2 Expected demands for item P1 resulting from BOM explosion and Demand Planning

	Demand/order period				
Source of demand	1	2	3	4	5
BOM explosion	280	~~80~~	~~40~~	–	–
Demand forecast	–	280	240	240	280
Expected demands	280	280	240	240	280

Table 11.3 Purchasing plan from two suppliers

	Order quantity per period from supplier				
Sourcing from supplier	1	2	3	4	5
1	–	–	–	520	–
2	800	–	–	–	–

To keep our example small, we will do with five periods. Here, the demands calculated (see Fig. 11.4) suffer from the effect of the lead-time offset, i.e. there are no demands at all in period five while for periods three and four secondary demands are missing resulting from future production of item C1. Hence, it is recommended to switch to demand forecasts (see Chap. 7) for periods with incomplete secondary demands (periods two to five in our example). Still, one should check whether existing secondary demands for these periods are in line with demand forecasts. Resulting demands are shown in Table 11.2.

The only data missing is the interest rate to be used for capital employed within the supply chain which is assumed 2.5 % per period.

The optimized purchasing plan (Stadtler 2007a) shows that the first order should be placed in period 1 from the second supplier with an order quantity of 800 units while the second order is placed with the first supplier in period four with an order quantity of 520 units (Table 11.3). The total cost within the planning interval comes to 10,333.25 [MU] (monetary units). Here, holding costs sum up to 201.25 [MU] (including interest on fixed ordering costs), fixed purchasing costs are 150 [MU] and variable purchasing cost are 9,982 [MU].

Some APS vendors provide a separate purchasing module for exploiting quantity discounts. This may be particularly appealing for commercial enterprises and for the procurement of MRO items in general. In the case that procurement decisions incur quantity discounts and resulting costs have a strong impact on the overall cost situation of a production unit, it may be advisable to declare respective items as “critical” and to include procurement decisions into the module Production Planning (assuming that corresponding cost functions can be modeled and solved there). If procurement decisions have to cover a longer planning horizon, one might even consider including these items at the Master Planning level.

In summary, the automation of the procurement process by means of an APS module can streamline the traditional, labor intensive tasks of procurement, especially in a B2B environment. Optimized procurement decisions can further reduce holding and total acquisition costs by exploiting quantity discounts and selecting suppliers in the best way possible.

References

Degraeve, Z., Roodhooft, F., & van Doveren, B. (2005). The use of total cost of ownership for strategic procurement: A company-wide management information system. *Journal of the Operational Research Society, 56*, 51–59.

Drexl, A., Fleischmann, B., Günther, H.-O., Stadtler, H., & Tempelmeier, H. (1994). Konzeptionelle Grundlagen kapazitätsorientierter PPS-Systeme. *Zeitschrift für betriebswirtschaftliche Forschung, 46*, 1022–1045.

Orlicky, J. (1975). *Material requirements planning*. New York: McGraw-Hill.

Silver, E., Pyke, D., & Peterson, R. (1998). *Inventory management and production planning and scheduling* (3rd ed.). New York: Wiley.

Stadtler, H. (2007a). A general quantity discount and supplier selection mixed integer programming model. *OR Spectrum, 29*(4), 723–745.

Stadtler, H. (2007b). How important is it to get the lot size right? *Zeitschrift für Betriebswirtschaft, 77*, 407–416.

Tempelmeier, H. (2002). A simple heuristic for dynamic order sizing and supplier selection with time-varying demand. *Production and Operations Management, 11*(4), 499–515.

Tempelmeier, H. (2006). *Material-Logistik – Modelle und Algorithmen für die Produktionsplanung und -steuerung und das Supply Chain Management* (6th ed.). Berlin: Springer.

Tempelmeier, H., & Derstroff, M. (1996). A lagrangean heuristic for multi-item multi-level constrained lotsizing with setup times. *Management Science, 42*, 738–757.

Vollman, T., Berry, W., & Whybark, D. (1997). *Manufacturing planning and control systems* (4th ed.). New York: Irwin/McGraw-Hill.