Chapter 12

PRODUCTION PLANNING OPTIMIZATION MODELING IN DEMAND AND SUPPLY CHAINS OF HIGH-VALUE CONSUMER PRODUCTS

Benoit Montreuil

Abstract This chapter is about production planning optimization modeling in the production centers in a demand and supply chain manufacturing, distributing and selling high value consumer products. First, it contrasts demand and supply chain alternatives in terms of collaboration, agility, customer-centricity and personalization offering, with a focus on the implications for production planning optimization. Second, it introduces a comprehensive production planning optimization model applicable to a large variety of centers in demand and supply chains. Third, it contrasts the production planning optimization model instance required as a demand and supply chain is transformed from a rigid and pushy implementation to integrate more collaboration, customer-centricity, agility and personalization. It also puts in perspective the importance of production planning optimization knowledge and technology. Finally it draws conclusive remarks for the research and professional communities.

1. Introduction

This chapter aims to clearly demonstrate that defining and modeling the production planning optimization problems of manufacturing centers in a demand and supply chain is an important activity which depends highly on the collaboration, agility, customer-centricity and personalization offering implemented through the demand and supply chain, as well as on the production planning optimization knowledge and technology available.

In order to contain complexity while insuring widespread representativeness, the chapter deals strictly with the demand and supply chain of manufacturers of high-value products such as vehicles, computers and equipment, sold to consumers in a large geographical region through a network of dealers. Furthermore, it focuses on the production planning optimization of the centers where the products are assembled. Finally, the emphasis is on problem modeling rather than on solution methodologies.

First the chapter presents a comprehensive description of alternate demand and supply chains, and the implications on production planning at the centers assembling the high value products. Second, it introduces a comprehensive production planning model encompassing a large variety of demand and supply chains. Third, it contrasts the production planning optimization model required when a demand and supply chain is transformed from a rigid and pushy mass production and distribution oriented implementation to integrate more collaboration, customer-centricity, agility and personalization. It also puts into perspective the importance of knowledge and technology. Fourth, it draws conclusive remarks for both the research and professional communities.

1.1 Contrasting alternative demand and supply chains

Demand and supply chains define the networks and processes through which demand and supply are expressed, realized and managed by an enterprise and its partners, all the way from suppliers to final customers. The demand and supply chain of an enterprise can take multiple forms, ranging widely in terms of customer-centricity, agility, collaboration and personalization capabilities.

In order to illustrate the spectrum of potential alternatives, Figures 12.1 and 12.2 synthesize two alternative demand and supply chains for an enterprise developing, manufacturing and distributing high value seasonal products to customers, such as recreational vehicles. In both cases the business sells products to several hundred thousand end-user clients spread throughout a large geographical region such as North America or Europe.

1.2 Mass production and distribution oriented demand and supply chain

In Figure 12.1, the demand and supply chain is built according to a mass production and distribution paradigm. It produces a mix of a few hundred standard products in a centralized factory with limited agility, requiring significant setups when its assembly center switches from one product to another. Each product is assembled from thousands of parts, components and modules. The enterprise has selected to produce some

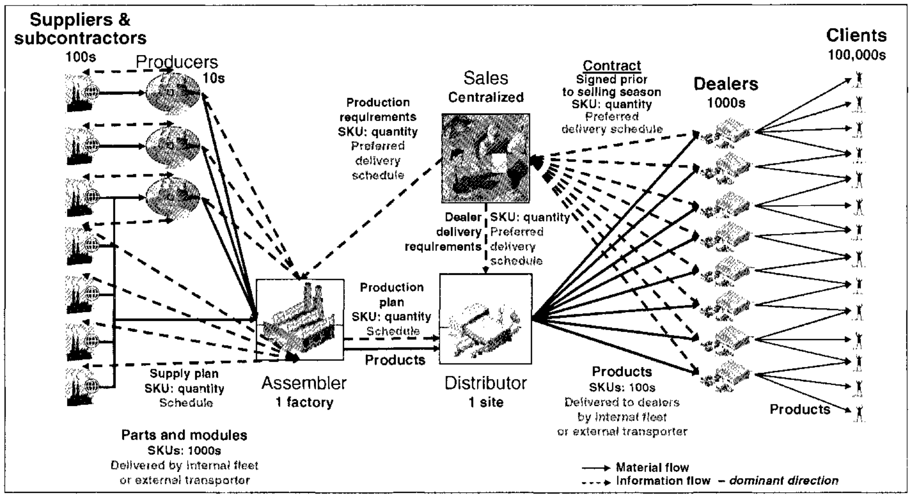

Figure 12.1. Demand and supply chain of a manufacturer with limited agility offering standard products and imposing pre-season sales contracts to dealers

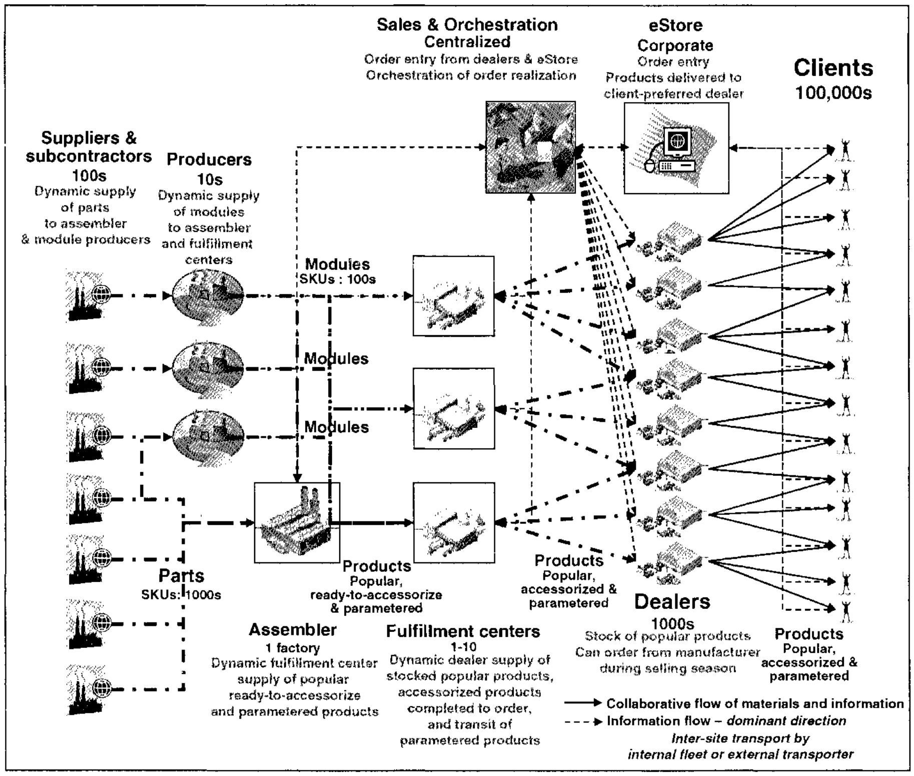

Figure 12.2. Demand and supply chain of an agile manufacturer offering both popular and personalized products with dynamic dealer supply

of these in one of its tens of internal production centers. The others are supplied from external suppliers and subcontractors.

The enterprise sells its products to a network of over a thousand dealers who have the responsibility of selling them to final customers. The dealers are independent businesses, not owned by the enterprise. Several months prior to the selling season, the enterprise forces each dealer to sign a contract stipulating how many units of each product it is buying.

Once all sales to dealers are known, the enterprise knows all production requirements for every product. This allows the factory to establish a master product assembly plan, deciding the sequence of products to be assembled through the entire production season. This plan can be very precise. Illustratively, it may state that from June 15th at 14:00 to June 17th at 10:00, the assembly center is planning to assemble 42 units of product 123 using a single eight-hour shift per day, with a takt time of 15 minutes per product unit. At 10:00 begins a period of 45 minutes, corresponding to three 15-minute cycles, required to change over to product 46 which is the next to be assembled.

The master assembly plan is transposed into an optimized supply plan for every part from every supplier and subcontractor. The supply plans take into account the cost structure, ordering constraint, and lead time speed and reliability of the supplier or subcontractor.

Once products are assembled, the enterprise assigns them to dealers. It optimizes the transportation of the products to their assigned dealers, taking into consideration its internal vehicle fleet and/or its external transporters. The dealers receive their ordered products prior to the heart of the selling season. They must attempt to satisfy clients as best as possible from their available product stock since the enterprise does not allow any reordering after their initial order.

The enterprise imposes such constraints to dealers and clients due to the generalized lack of agility through its supply chain. The assembly center requires significant setup times and costs when switching from a product to another. Its network of internal component/part/module production centers and external suppliers and subcontractors generally does not have the capability and capacity necessary to operate without the stability and visibility offered by the pre-season contract system.

In one variant of this rigid demand and supply chain, the pre-season contract with each dealer gives the enterprise full freedom in deciding when it is to ship the ordered products to the dealer, as long as each receives showcase products early on and the remainder prior to the selling season peak. In another variant, the enterprise is more collaborative with the dealers and lets them stipulate preferred target dates for receiving each unit. In a limited accountability version, the enterprise simply states that it will try to satisfy these targets as closely as possible. In an alternative version of this variant, the enterprise may offer dealers rebates proportional to the deviation between the delivered date and the target date for each ordered product.

1.3 Personalized, customer-centric, collaborative and agile demand and supply chain

Figure 12.2 depicts a much more customer-centric and agile demand and supply chain. Here the enterprise is geared to deliver a personalization offer to customers (Montreuil and Poulin, 2004; Poulin et al., 2004). It offers popular products, expected to be available off-the-dealer-shelves or to be delivered within a few days to the client through its selected dealer. It also offers two types of personalized products: accessorized products and parametered products. Accessorized products are assembled from ready-to-accessorize products to which are added personalized sets of modules. Parametered products are selected by customers through the setting of parameters or options. The personalized products, either accessorized or parametered, are promised to be delivered in a specific number of days to the client through its selected dealer. The order-to-delivery time is promised to be shorter for accessorized products than for parametered products. All products can be ordered by customers either at a dealership or through the web-based eStore operated by the enterprise. In the latter case, the client selects a dealer where he wants the product delivered and where he wants after-sales service.

The main factory has the mandate to assemble standard products, ready-to-accessorize products and parametered products. It is more agile and has lower changeover time from one product to the next. The completion of personalized products from modules and ready-to-personalize products is performed in one of the few fulfillment centers strategically distributed throughout the territory. These fulfillment centers are highly agile, capable of finishing the personalized products on a first-come-first-serve basis with no changeover time from one product to another. The fulfillment centers also serve as transhipment points for parametered products. Both the factory and the fulfillment centers operate during the selling season.

The demand and supply chain has a collaborative nature. From the demand side, on one hand, dealers are allowed to reorder as often as they want. On the other hand, they are asked to collaborate by providing regular forecast updates on their forthcoming demand. The forecasts allow the enterprise to speculatively assemble standard products and ready-to-personalize products. The speculative stocks allow the enterprise to offer faster delivery, especially during the selling season peak where demand may exceed production capacity. The opportunity to build speculative stocks may also permit the enterprise to smooth its production, especially its manpower and supply requirements. From the supply side, the chain exploits collaborative exchange of information, plans and constraints between the partners.

1.4 Production planning implications

In the pushy demand and supply chain of Figure 12.1, the enterprise imposes dealer pre-season contracts and thereafter operates mostly in deterministic high visibility mode. The enterprise is thus free of the demand uncertainty during the production season. It faces demand uncertainty once the selling season starts. Its only reactive mechanism relies on a combination of publicity and rebates as the assembly center is closed and assembled products are already shipped to specific dealers. At the end of the season, it faces the sales results, including a percentage of unsold product units at most dealerships. These unsold units will have to be offered at discount price during the next selling season, cannibalizing the new vehicle market.

In the demand and supply chain of Figure 2, the enterprise is facing a much lower visibility and a much higher uncertainty, as dealers decide to order whenever they want, and clients may similarly order whenever they want directly on the web. High availability and fast and reliable delivery are promised. This implies that the enterprise may regularly find itself with a few-day order booking, having to take decisions based on forecasts. To compensate, the much more agile chain is such that there is much less pressure to produce and order in large lots, which allows faster reaction to occurring events. However this agility is never perfect and numerous constraints may still have to be taken into account.

Production planning, the focus of this chapter, is a seasonal event at the assembly factory of Figure 1. The master assembly plan covers the entire production season in a mostly deterministic fashion. The only probabilistic events are engineering changes, machine breakdowns, quality problems, manpower strikes and supplier lateness. When these occur, they are taken care of by local adjustments to the master schedule. However, the enterprise constantly works at reducing their occurrence by better controlling its chain.

In Figure 12.2, assembly production planning occurs both at the main factory and at fulfillment centers. In both cases the planning horizon is much shorter than when operating according to Figure 12.1. In fact the master plan is an ongoing creation, constantly rejuvenating itself in light of recent events, systematically advancing through a short rolling planning horizon. Beyond the planning horizon lies the forecasting and resource planning horizon, mostly necessary to smooth production and deal with long lead time suppliers. Constant work by the enterprise to become more agile, internally as well as through the network of external suppliers and subcontractors, aims at reducing the need to rely on forecasts. Yet, especially in seasonal markets such as is the focus of the demand and supply chains studied in this chapter, there remains a dependency on some forecasting, mostly related to decisions to build or not anticipatory stock to smooth future production and help to fulfill demand in peak periods.

Both cases share many common features relative to production planning. They are also distinct relative to many others due to differences in customer centricity, agility and personalization offering. Yet, it is important to recognize that they reflect extreme situations. Real situations often lie in between these two extremes, often gradually evolving from the more rigid type to the more agile, customer-centric and personalized type. In light of such insights, section two introduces a comprehensive problem formulation which sustains both cases. The formulation allows dealing with each case by appropriately selecting sets of constraints, variables and by setting parameters.

In both cases the production planning problem definition is highly dependent on the level of consciousness of the decision makers about the impact of the planning on the operations of other stakeholders in the network and about the impact of these stakeholders on the global feasibility and optimality of the production plan. The presentation of the problem formulation in Section 2 is structured to highlight this phenomenon.

2. Formulating the production planning problem

This section provides a comprehensive formulation of the production planning problem in an assembly center driven by a stable takt time establishing the pace of finished products output from the center. The center may consist of a single assembly line or an assembly tree composed of sub-lines recursively feeding a master line. The main decision consists of determining which product to assemble in each takt time slot and when to switch from one product to another, which often involves changeover operations requiring time off in each station, consequently creating a production gap in the center output. This is illustrated in Figure 12.3. In the short extract displayed, it is shown that 33 units of product B are to be assembled, followed by at least 16 units of product M. A unit is to be produced every 15 minutes according to the takt

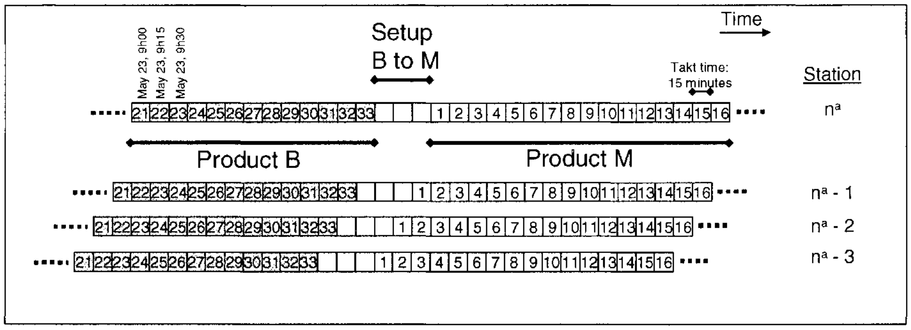

Figure 12.3. Illustrating the production plan of the assembly center

time. For example, unit 21 of product B is to be finished at 9:00 on May 23rd while the 22nd unit is to be finished at 9:15. Between products B and M, the center requires three slots of 15-minute takt time to put in effect the product changeover. Figure 12.3 also shows the forward shift in time of the plans associated with the last station, then the second to last, and so on up to the first station. For example, the 23rd unit of product B is to be finished at 9:30 in the last station, at 9:15 in the second-to-last station, and so on.

In real settings corresponding to Figure 12.1, the production plan (alternatively named master assembly plan thereafter) may readily comprise six months of production, roughly about 120 active days. Often, such centers operate one or two shifts. Assuming a single 8-hour shift per day and a takt time of three minutes, this cumulates to about 19,200 time slots available for assembling a product or making a changeover. Assuming 200 products, this means an assignment matrix of 200 by 19,200, which involves the assignment of 3,840,000 entries in the matrix. These entries are the key decision variables in the problem formulation presented here. This should make it clear that the production planning problem in such a context is a large scale problem.

It should be understood that when the production lot sizes are known to be large, reducing the overall potential number of production runs, then an alternative modeling framework based on start and end times for each production run may be more economical in terms of the number of variables and constraints. However, with the intended goal of the formulation presented here to sustain the full spectrum of possibilities relative to production agility, then the time slot assignment modeling framework is preferred.

The problem formulation is presented below in a modular fashion. First is listed the entire set of sets, indices, parameters and variables.

Second, the objective function is presented from an overall perspective, shown to be adapted depending on problem scoping options. These options depend both on the level of network stakeholder consciousness and collaboration, and on the type of demand and supply chain. Then are iteratively introduced the constraint sets associated with modeling features depending on problem scoping options.

2.1 Sets, indices, parameters and variables

The formulation being comprehensive in nature, it encompasses a large number of decision variables and parameters, requiring a significant number of sets and indices to permit its coherent representation. These are listed hereafter. An effort has been made to have the identifiers as meaningful as possible, however the sheer number of them has forced to use such tricks as superscripts to control complexity.

Sets.

$\boldsymbol{A}$: Set of production/assembly stations composing the production center

$\boldsymbol{B_{sr}}$: Set of time buckets to be used for planning critical resource r of capacitated supplier s, each defined through a starting time t^{s}_{srb} and a finish time t^{e}_{srb}

$\boldsymbol{C_p}$: Set of products p' requiring a nonzero changeover time when switching from product p to product p' ($e_{pp'} > 0$)

$\boldsymbol{C_{pa}}$: Set of products p' requiring a positive number of workers at assembly station a to perform the changeover work during the nonzero changeover time when switching from product p to product p'

$\boldsymbol{M}$: Set of all modules m

$\boldsymbol{M_{sr}}$: Set of modules m whose supply requires a positive amount of critical resource r of supplier s

$\boldsymbol{N^{sf}_p}$: Set of cost segments n for speculative stock of product p at the end of the planning horizon

$\boldsymbol{P}$: Set of all products

$\boldsymbol{P_{ma}}$: Set of products requiring module m at assembly station a

$\boldsymbol{R_s}$: Set of critical resources r of capacitated supplier s

$\boldsymbol{S^c}$: Set of capacitated suppliers s

$\boldsymbol{T}$: Set of time periods, linearly sequenced from 0 to t^l

$\boldsymbol{T^a}$: Set of assignable time periods, linearly sequenced from 1 to t^l

$\boldsymbol{T^m}$: Set of time periods at which a change in manpower is allowed

$\boldsymbol{T^s}$: Set of allowed supply time periods, linearly sequenced from t^{ss} to t^l

$\boldsymbol{T^w}$: Set of working time periods, linearly sequenced from t^{sw} to t^l

$\boldsymbol{U}$: Set of cost segments for unused time slots in the assembly center

W: Set of worker types
Z: Set of geographical zones in the dealership network

Indices.

a: A station in the production center
b: A planning time bucket
m: A module (component, part, etc.)
n: A linear cost segment of speculative product inventory, cost increasing with n
p, p': A product (when equal to zero, it means "no product")
r; A constraining resource
t: A time period of duration equal to the takt time of the production center
t^l: The last time period
w: A worker type
z: A geographical zone in the dealership network

Parameters.

$c^c_{pp't}$: Actualized marginal changeover cost when the production center switches from making product p to making product p' at time t
c^e_t: Actualized expected unit cost for not finishing a product in time t
c^{im}_{mt}: Actualized unit inventory cost for module m at time t
c^{ip}_{pt}: Actualized unit inventory cost for product p at time t
c^{om}_{mt}: Actualized cost for ordering module m from its supplier at time t, including administration and transport
c^o_t: Actualized marginal cost of opening the production center at time slot t as perceived from the end of the production center
c^{qm}_{mt}: Actualized unit purchasing cost for module m from its supplier at time t
c^{s-}_{pt}: Actualized unit cost per deviation from minimal safety stock target for product p at time t
c^{sf}_{pn}: Actualized unit cost of speculative product stock cost at the end of the planning horizon for product p, in cost segment n
c^{sl}_{srb}: Actualized marginal cost for using critical resource r of capacitated supplier s during bucket b
c^{sl+}_{srb}: Actualized marginal cost for exceeding the average load on critical resource r of capacitated supplier s during bucket b
c^{sl-}_{srb}: Actualized marginal cost for underachieving the average load on critical resource r of capacitated supplier s during bucket b
c^T_u: Actualized expected marginal cost for not using a number of available time slots in cost segment u

c^{v}_{zt}: Actualized cost for a round-trip transport to zone z departing at time t

c^{w}_{wt}: Actualized marginal cost per period for each worker of type w at time t

c^{w+}_{wt}: Actualized marginal cost for adding a worker of type w at time t

c^{w-}_{wt}: Actualized marginal cost for removing a worker of type w at time t

d_{pzt}: Preferred cumulative deliveries of product p in zone z at time t, summed over the individual preferences of each dealer in zone z

$e_{pp'}$: Number of time periods during which the changeover from product p to product p' requires to stall production at each station in the center

e^{T}_{u}: Number of unused time slots belonging to cost segment u

f^{v}: Number of vehicles in the fleet

i^{s}_{pt}: Target safety stock for product p at time t

i^{sf}_{pn}: Maximum inventory allowed in cost segment n for product p

l_{a}: Time lag between station a and the end of the production center, time between the end of production for a product at station a and its exit of the production center

l^{d}: Required time lag between the production completion of product p at the factory and its availability at the distribution center for delivery to dealers

l^{m}_{ma}: Required time lag between the delivery of module m from its supplier and its use in assembly station a

l^{r}_{srb}: Maximum load to be imposed on critical resource r of capacitated supplier s during planning time bucket b

l^{ra}_{srb}: Average load of critical resource r of capacitated supplier s over the planning horizon, adjusted to the length of time bucket b

l^{s}_{m}: Lead time from order to delivery of module m by its supplier to the factory

n^{a}: Number of assembly stations

n^{p}: Number of products

n^{f}_{p}: Number of cost segments for anticipatory stock of product p at the end of the planning horizon

$n^{\min}_{wt}$: Minimal allowed number of workers of type w at time t

$n^{\max}_{wt}$: Maximal possible number of workers of type w at time t

o_{p}: Total order for product p

q_{mpa}: Quantity of modules m required per unit of product p at assembly station a

q^{h}_{m}: Total quantity of modules m required to assemble all products demanded by the dealers over the planning horizon

q^{r}_{srm}: Quantity of resource r required to produce one unit of module m at supplier s

q^s_m: Minimum ordering lot size imposed by the supplier of module m

r_{pzt}: Actualized revenue from delivering a unit of product p to dealers in zone z at time t

r^a_{wap}: Number of workers of type w required at assembly station a when assembling product p

$r^c_{wapp'}$: Number of workers of type w required at assembly station a when changing over from product p to product p'

s_p: Space occupied by a unit of product p in a transport vehicle

s^v: Space availability in a transport vehicle

t^{ss}: First time period at which a module may be ordered from a supplier ($t^{ss} \leq 0$)

t^{sw}: First time period at which any change can be made to the work assignment of any station of the production center (generally ≤ 0)

t^e_{srb}: End time of bucket b

t^s_{srb}: Start time of bucket b

v^t_z: Travel time for a round-trip to zone z by a delivery vehicle

Variables.

A_{pt}: Binary variable stating whether or not a unit of product p is to be finished at time t

$C_{pp't}$: Binary variable stating whether or not a changeover from product p to p' starts at time t, as perceived at the end of the production center

C^c: Nonnegative real variable summing the actualized total changeover cost

C^e: Nonnegative real variable summing the actualized total cost for empty production slots, not finishing products at each time in the planning horizon

C^i: Nonnegative real variable summing the actualized total product inventory cost

C^o: Nonnegative real variable summing the actualized total line opening cost

C^s: Nonnegative real variable summing the actualized total supply cost

C^{s-}: Nonnegative real variable summing the actualized total cost for deviation from minimal safety stock targets for products

C^{sf}: Nonnegative real variable summing the actualized total speculative product stock cost at the end of the planning horizon

C^v: Nonnegative real variable summing the actualized total vehicle transport cost

C^w: Nonnegative real variable summing the actualized total personnel cost

$\boldsymbol{D_{pzt}}$: Nonnegative real variable computing the cumulative dealer network delivery of product p at time t

$\boldsymbol{D^{+}_{pzt}}$: Nonnegative real variable computing the over-delivery of product p to dealer network at time t

$\boldsymbol{D^{-}_{pzt}}$: Nonnegative real variable computing the under-delivery of product p to dealer network at time t

$\boldsymbol{D^{t}_{pzt}}$: Nonnegative real variable computing the punctual delivery of product p to dealer network at time t

$\boldsymbol{E_t}$: Binary variable equal to one only when no product is to be finished at time t

$\boldsymbol{E^{T}_{u}}$: Nonnegative real variable computing the number of unused assembly time slots (not producing products) belonging to the cost segment u, during the entire planning horizon

$\boldsymbol{I_{pt}}$: Nonnegative real variable computing the distribution center inventory of product p at time t

$\boldsymbol{I^{m}_{mt}}$: Nonnegative real variable computing the inventory of modules m at time t

$\boldsymbol{I^{s+}_{pt}}$: Nonnegative real variable computing the positive deviation from the safety stock target for a product p at time t

$\boldsymbol{I^{s-}_{pt}}$: Nonnegative real variable computing the negative deviation from the safety stock target for a product p at time t

$\boldsymbol{I^{sf}_{pn}}$: Nonnegative real variable computing the speculative stock of product p at the end of the planning horizon, belonging to cost segment n

$\boldsymbol{L_{srb}}$: Nonnegative real variable computing the load on critical resource r of capacitated supplier s during a planning time bucket b

$\boldsymbol{L^{+}_{srb}}$: Nonnegative real variable computing the above average loading on critical resource r of capacitated supplier s during a planning time bucket b

$\boldsymbol{L^{-}_{srb}}$: Nonnegative real variable computing the under average loading on critical resource r of capacitated supplier s during a planning time bucket b

$\boldsymbol{N_{wt}}$: Nonnegative integer variable computing the number of workers of type w active in the production line at time t

$\boldsymbol{N^{+}_{wt}}$: Nonnegative integer variable computing the number of workers of type w added in the production line at time t

$\boldsymbol{N^{-}_{wt}}$: Nonnegative integer variable computing the number of workers of type w removed from the production line at time t

$\boldsymbol{O^{f}_{t}}$: Binary variable stating whether the production center is open at time period t, as perceived from the end of the center

$\boldsymbol{O^{s}_{mt}}$: Binary variable stating whether or not an order of modules m is transmitted to its supplier at time t

$\boldsymbol{P_{pt}}$: Nonnegative real variable cumulating production of product p up to time t

$\boldsymbol{Q^s_{mt}}$: Nonnegative real variable computing the quantity of modules m ordered from its supplier at time t

$\boldsymbol{R_{wt}}$: Nonnegative real variable computing the number of workers of type w required at time t

$\boldsymbol{R^d}$: Nonnegative real variable computing the actualized total revenue generated from deliveries to dealer network

$\boldsymbol{R^s_{mt}}$: Nonnegative real variable computing the cumulative total number of modules m required from its supplier up to time t

$\boldsymbol{U_u}$: Binary variable stating whether or not there is a greater-than-zero number of unused assembly time slots during the planning horizon corresponding to cost segment u

$\boldsymbol{V_{zt}}$: Nonnegative integer variable computing the number of transport vehicles departing to zone z at time t

2.2 Objective function

Production planning greatly influences the flow of revenues and costs through the planning horizon. From the revenue side, in the studied cases, sales are registered once a product is delivered to a dealer. Dealers order a variety of products, with distinct margins associated to each product. Therefore the production sequence affects the availability of products for delivery, which affects the deliveries to dealers, which affects the revenue stream. Also, especially in the agile and client-centric demand and supply chain of Figure 12.2, time spent on changeovers in the assembly center reduces the potential for producing products required by customers, thus having an impact on overall sales. This influence of production planning on revenue leads to using a maximizing objective function as stated in equation (12.1).

Objective function.

$$\text{Maximize } R^d - (C^o + C^c + C^w + C^i + C^v + C^s + C^{s-} + C^{sf} + C^e) \quad (12.1)$$

The above statement of the objective function is purposefully limited to identifying the main aggregate revenue and cost variables. Detailed specification of each cost variable is to be addressed in a modular fashion in the next sections. However it is important to state that all variables in the objective function are actualized, taking into consideration the present value of future costs and revenues. All cost and revenue parameters are also allowed to be time dependent.

The costs in the objective function include, in their presentation order in (12.1):

(1) Center opening cost;
(2) Product changeover cost;
(3) Personnel cost;
(4) Product inventory cost;
(5) Transport-to-dealer cost;
(6) Supply cost;
(7) Safety stock target deviation cost;
(8) Speculative product stock cost;
(9) Empty production slot cost.

Cost variables (1) and (2) encompass the most basic costs which an enterprise is conscious of when planning production. Their computation is modelled in Section 2.3 dealing with operation and changeover in the assembly center. Third, the personnel cost variables are cumulating costs associated to manpower requirements and variations. They are modelled in Section 2.4 dealing with personnel. Product inventory and transport-to-dealer cost variables (4) and (5) are addressed in Section 2.5 dealing with the dealer network. The sixth cost variable corresponds to the supply cost and is modelled in Section 2.6 dealing with the supply network. The last three cost variables (7) to (9) are modeling costs associated with dealing with the dynamic uncertainty in a rolling horizon mode, which is dealt with in Section 2.7.

2.3 Dealing with operation and changeover in the assembly center

The core operational decision variables in planning production in the assembly center are the A_{pt} and $C_{pp't}$ variables. The former variables state whether or not product p is to be assigned to production time slot t, finishing the product in the last assembly station at time t. The latter variables state whether or not a changeover from product p to product p' is to be initiated in time slot t. From these are derived the following sets of constraints defining the core operations and costs of the assembly center.

Operations and changeover constraints.

$$P_{pt} = P_{p,t-1} + A_{pt}, \quad \forall p \in P; \forall t \in T^a \tag{12.2}$$

$$P_{pt^l} \geq o_p, \quad \forall p \in P \tag{12.3}$$

$$\sum_p A_{pt} + E_t = 1, \quad \forall t \in T^a \tag{12.4}$$

$$\begin{aligned}(1 - C_{pp't}) \geq \sum_p A_{pt'}, \quad & \forall t' \in [t, t-1+e_{pp'}]; \forall t \in T; \\ & \forall (p \in P, p' \in P) \mid p' \in C_p\end{aligned} \tag{12.5}$$

$$\sum_{p' \neq p} C_{pp't} \leq A_{p,t-1}, \quad \forall p \in P, \forall t \in T^a \tag{12.6}$$

$$A_{p't} \leq A_{p'(t-1)} + \sum_{p} C_{pp'(t-e_{pp'})}, \quad \forall p' \in P; \forall t \in T^a \tag{12.7}$$

$$\sum_{p'} C_{pp't} = 1, \quad p = 0; t = 1 \tag{12.8}$$

$$O_t^f = \sum_{p \in P} A_{pt} + \sum_{p \in P} \sum_{\substack{p' \in P \\ e_{pp'} > 0}} \sum_{t' \in [t, t-1+e_{pp'}]} C_{pp't'}, \quad \forall t \in T^a \tag{12.9}$$

$$C^o = \sum_{t \in T^a} c_t^o O_t^f \tag{12.10}$$

$$C^c = \sum_{p \in P} \sum_{\substack{p' \in P \\ p' \neq p}} \sum_{t \in T^a} c_{pp't}^c C_{pp't} \tag{12.11}$$

Constraint set (12.2) computes the cumulative production of each product p at each time t. Constraint set (12.3) imposes that the cumulative production of each product p at the end of the planning horizon be at least as high as the total number of orders for product p, ensuring that all orders are planned to be fulfilled. Constraint set (12.4) limits each production time slot to be assigned at most one product unit. It also identifies the time slots during which no product units are produced, the center being either idle or changing over from one product to another. Constraint set (12.5) imposes that no product be assembled during the changeover time from a product p to another p’ starting at time t. For example, in Figure 12.3, this imposes the three-time-slots changeover time from product B to product M in each assembly station. Note that this changeover time may be dependent on the pair of from-to products. For example, changing over from B to M may take 3 time slots, but changing from B to S may take 20 time slots. Constraint set (12.6) restricts a changeover from p to p' to be allowed at time t only if product p is allocated for production in time $t-1$. Conversely, constraint set (12.7) restricts production of product p' in time t to be allowed only if product p' was already produced in the previous time slot or if a changeover has been performed from some product p to product p' in the previous time slots, according to the specified changeover duration. Constraint set (12.8) initializes the production by deciding which product is to be produced first, requiring beforehand an initial setup. It uses product zero as surrogate for stating the initial changeover time requirements. Constraint set (12.9) determines whether or not each time slot is open or not, active either in producing a product unit or changing over from one product to another.

Constraints (12.10) and (12.11) respectively cumulate the costs for opening production time slots and for the inter-product changeovers.

The operational and changeover constraint sets can be generalized to deal with multiple parallel assembly centers, each with specific capabilities in terms of which products it can assemble. This generalization is beyond the scope of the chapter and left as an exercise to the interested reader.

2.4 Dealing with personnel

Assembly centers such as those modelled in this chapter may readily employ multiple hundreds of people. These represent important costs. The production planning decisions influence the need for personnel, and therefore the personnel costs. Opening time slots implies having an adequate number of persons to operate the center during those slots. Furthermore, it is often the case that each product requires a specific number of workers of each type in each assembly station. For example, a small and simple product may require less people in the assembly center than a large complex product. Two products may have very similar personnel requirements at each station, except a few where they differ dramatically due to their specifications.

In many enterprises, personnel cost is not dealt with explicitly when developing the master assembly plan. In such cases the plan is forwarded to the human resources center which has the responsibility of providing and assigning the right set of people to the center. The production plan is thus optimized without considering the personnel costs, and the personnel costs are afterward minimized given the production plan constraints. Adjustments may be made to the production plan to deal with infeasibilities.

Below are presented the constraint set allowing to consciously integrate personnel into the production planning optimization.

Personnel constraints.

$$n_{wt}^{\min} \leq N_{wt} \leq n_{wt}^{\max}, \quad \forall w \in W; \forall t \in T^w \tag{12.12}$$

$$N_{w,t-1} + N_{wt}^{+} - N_{wt}^{-} = N_{wt}, \quad \forall w \in W; \forall t \in T^m \tag{12.13}$$

$$N_{w,t-1} = N_{wt}, \quad \forall w \in W; \forall t \in \{T^w - T^m\} \tag{12.14}$$

$$R_{wt} \leq N_{wt}, \quad \forall w \in W; \forall t \in T^w \tag{12.15}$$

$$R_{wt} \geq \left[\sum_{a \in A} \sum_{p \in P} r^{a}_{wap} A_{p,t+l_a} + \sum_{a \in A} \sum_{p \in P} \sum_{p' \in C_{pa}} r^{c}_{wapp'} \sum_{s=t+l_a}^{t+l_a-1+e_{pp'}} C_{pp's} \right], \quad \forall w \in W; \forall t \in T^{w} \qquad (12.16)$$

$$C^{w} = \sum_{t \in T^{a}} \sum_{w \in W} c^{w}_{wt} N_{wt} + \sum_{t \in T^{m}} \sum_{w \in W} (c^{w+}_{wt} N^{+}_{wt} + c^{w-}_{wt} N^{-}_{wt}). \qquad (12.17)$$

The key decision variables related to personnel are the N_{wt} and R_{wt} variables which state the number of workers of type w to be respectively working and required in the assembly center in time t. For each time t, constraint set (12.12) bounds this number to be lower than the available pool of workers of the specific type and to be higher than the union-negotiated and/or strategically-planned lower limit on the number of workers of this type. Constraint set (12.13) computes the increase and decrease in workforce of each type occurring at each time period. Constraint set (12.14) forces workforce increases or decreases to be occurring only at allowable times. It does so by forcing the workforce to be the same as in the precedent time slot whenever workforce changes are not allowed in the time slot. This set is included for presentation clarity. However it leads to variable set reduction prior to solving the problem.

Constraint set (12.15) insures that the number of workers of each type in each time slot is always greater or equal to the required number to realize the production at each assembly station. Constraint set (12.16) computes the number of workers of each type required at each assembly station at time t given the production and changeover decisions.

Constraint (12.17) cumulates the total personnel cost, combining for each time slot the cost of working employees, the cost of adding employees and the cost of removing employees. These latter costs may be expensive when numerous variations of staffing level occur. Constraint (12.17) could be made even more rigorous by adding the cost of moving personnel around in the center to deal with varying personnel requirements at each station from one time slot to the next, whenever this cost becomes significant and influenced by the production plan. This is left as an exercise to the interested reader.

2.5 Dealing with the dealer network

Demand and supply chains such as those studied in this chapter involve thousands of dealers geographically spread throughout large regions. At production planning it is generally too cumbersome to explicitly deal with each dealer. So enterprises make compromises.

Most mass production oriented enterprises completely discard them from their modeling. In fact the dealers only appear in aggregate form in constraint set (12.3) which makes sure that somehow during the production season all orders from all dealers are produced. Dealing with the assignment of production to dealers is left as an aftermath decision to be dealt with by distribution managers.

Even within the framework of the demand and supply chain of Figure 12.1, there is potential and reward for production planning to explicitly integrate the dealer network in its optimization modeling. In a demand and supply chain that is customer-centric, agile and/or offering personalization, explicitly dealing with the dealer network is a requisite. Below is presented a set of constraints which allows modeling the issues relevant to production planning that are related to the dealer network.

By the way, this is where the notion of inventory is introduced. If the dealer network is not explicitly modelled, then the production planner either assumes that finished products will be shipped efficiently to dealers in a prompt manner after their availability for delivery or that the inventory-transport decisions will be subordinated to the production plan without significant loss of optimality. Dealing with product inventory involves defining the set of variables I_{pt} stating the inventory of product p at time t.

Dealer network constraints.

$$I_{pt} = P_{pt} - \sum_{z \in Z} D_{pzt}, \qquad \forall p \in P; \forall t \in T^a \tag{12.18}$$

$$P_{p(t-l^d)} \geq \sum_{z \in Z} D_{pzt}, \qquad \forall p \in P; \forall t \in T^a \tag{12.19}$$

$$D_{pzt} - d_{pzt} = D^{+}_{pzt} - D^{-}_{pzt}, \qquad \forall p \in P; \forall z \in Z; \forall t \in T^a \tag{12.20}$$

$$D^{t}_{pzt} = D_{pzt} - D_{pz,t-1}, \qquad \forall p \in P; \forall z \in Z; \forall t \in T^a \tag{12.21}$$

$$s^v V_{zt} \geq \left[\sum_{p \in P} \sum_{t \in T^a} s_p D^{t}_{pzt} \right], \qquad \forall z \in Z; \forall t \in T^a \tag{12.22}$$

$$\sum_{z \in Z} \sum_{t' \in [t+1-v^z_t, t]} V_{zt'} \leq f^v, \qquad \forall t \in T^a \tag{12.23}$$

$$R^d = \sum_{p \in P} \sum_{t \in T^a} \sum_{z \in Z} r_{pzt} D^{t}_{pz,t-v^t_z/2} \tag{12.24}$$

$$C^i = \sum_{t \in T^a} \sum_{p \in P} c^{ip}_{pt} I_{pt} \tag{12.25}$$

$$C^v = \sum_{t \in T^a} \sum_{z \in Z} c^{v}_{zt} V_{zt} \tag{12.26}$$

In order to model the dealer network, it is clustered in a set of dealer zones grouping nearby dealers. Furthermore, the delivery timing preferences of each dealer are pooled to generate the set of parameter d_{pzt} stating the preferred cumulative deliveries of product p in zone z at time t. Two key sets of variables allow modeling the dealer related decisions, these are D_{pzt} and V_{zt}. The D_{pzt} variables compute the cumulative delivery of product p to the dealers in zone z up to time t. The V_{zt} variables stipulate the number of transport vehicles departing to zone z at time t, so as to deliver products to dealers in that zone z.

Constraint set (12.18) computes the inventory of product p in time t as the difference between the cumulative production of product p up time t and the sum over all regions of the cumulative deliveries of product p to these regions up to time t. Constraint set (12.19) insures that sufficient production of product p is realized prior to its delivery to dealers, with enough lead time to permit transit from the factory to the distribution center and preparation for delivery. Constraint set (12.20) computes the positive and negative differences between the preferred and achieved cumulative deliveries of product p to dealers in zone z at time t.

The punctual deliveries of product p to zone z shipped at time t are determined through constraint set (12.21). Cumulating these punctual product-to-zone deliveries at time t for each zone, constraint set (12.22) determines the corresponding number of transport vehicles departing to the zone, given the space requirements of each product in the vehicle and the spatial capacity of the vehicles. Constraint set (12.23) limits the number of vehicles simultaneously travelling to never exceed the fleet size. These constraint sets can easily be generalized to concurrently deal with volume and weight capacities, distinct types of transport vehicles, and combinations of internal fleet and external transporters.

The total actualized revenues are computed through constraint set (12.24). It assumes revenues to be registered at delivery time to dealers, approximated to mid round trip to the dealer zone. Other revenue actualization can be similarly modelled to reflect specific situations. Constraints (12.25) and (12.26) respectively compute the total actualized inventory and transport costs.

2.6 Dealing with the supplier network

By their intrinsic nature, supply networks studied in this chapter involve a large number of supplied parts, components and/or modules, supplied by many suppliers and subcontractors. Indeed it is common to deal with several thousands of items by hundreds of organizations throughout the world.

Everybody having been involved in assembly factories readily recognizes that a lot of the operational problems leading to difficulties in delivering productively, fast and reliably are related to missing or incorrect supplies. In fact, when supplies are always on time at the right assembly stations, piloting the assembly center becomes much easier. Furthermore, even though everyone dreams of agile nonconstraining suppliers delivering perfect products just in time with short notice, there are nearly always suppliers with significant and unreliable delivery times, imposing minimal supply lots, and subject to limited supply capacity. These may have significant impact on the feasibility and profitability of production plans. Yet in most cases, supply planning is performed subject to a predetermined production plan and forecasts, according to a material requirement planning (MRP) logic. Below are presented sets of constraints allowing for integration of supply planning of influential inputs and suppliers to production planning optimization.

Dealing with the supply network.

$$R^s_{mt} = R^s_{m,t-1} + \sum_{a \in A} \sum_{p \in P_{ma}} q_{mpa} A_{p,(t+l_a+l^m_{ma})}, \qquad \forall m \in M; \forall t \in T^a \tag{12.27}$$

$$\sum_{\substack{t' \in T^s \\ t' \leq t - l^s_m}} Q^s_{mt'} = R^s_{mt} + I^m_{mt}, \quad \forall m \in M; \forall t \in T^s \tag{12.28}$$

$$Q^s_{mt} \leq q^h_m (1 - O^s_{mt}), \quad \forall m \in M; \forall t \in T^s \tag{12.29}$$

$$Q^s_{mt} \geq q^s_m O^s_{mt}, \quad \forall m \in M; \forall t \in T^s \tag{12.30}$$

$$L_{srb} = \sum_{m \in M_{sr}} q^r_{srm} \left(R^S_{mt^s_{srb}} - R^S_{mt^e_{srb}} \right), \qquad \forall s \in S^c; \forall r \in R_s; \forall b \in B_{sr} \tag{12.31}$$

$$L_{srb} \leq l^r_{srb}, \quad \forall s \in S^c; \forall r \in R_s; \forall b \in B_{sr} \tag{12.32}$$

$$L_{srb} - l^{ra}_{srb} = L^+_{srb} - L^-_{srb}, \quad \forall s \in S^c; \forall r \in R_s; \forall b \in B_{sr} \tag{12.33}$$

$$C^s = \sum_{t \in T^s} \sum_{m \in M} (c^{om}_{mt} O^s_{mt} + c^{qm}_{mt} Q^s_{mt} + c^{im}_{mt} I^m_{mt} + c^{sl}_{srb} L_{srb}) + \sum_{s \in S^c} \sum_{r \in R_s} \sum_{b \in B_{sr}} (c^{sl+}_{srb} L^+_{srb} + c^{sl-}_{srb} L^-_{srb}). \tag{12.34}$$

The key variables allowing to deal explicitly with supply are the R^s_{mt}, Q^s_{mt} and I^m_{mt} variables, respectively deciding the supply requirements, supply quantity ordered and the current inventory of module (part, component, etc.) m at time t. The link between supply and production is set by parameters q_{mpa} stating the quantity of modules m required per unit of product p treated at assembly station a of the assembly center.

Constraint set (12.27) transposes the assembly plan decisions into their supply implications. Explicitly, they update the cumulative supply requirements for each module m at time t by adding to the previous cumulative requirements at time $t-1$ the material requirements generated by the products assigned to each assembly station at time t. They allow for setting a required time lag between the delivery of a module m from its supplier and its use in assembly station a for internal logistic considerations and protection against supplier delivery time unreliability.

Constraint set (12.28) balances on one side the cumulative quantity of ordered modules m planned to have been delivered at time t, taking delivery lead time in consideration, and on the other side the combination of current module inventory and cumulative consumption of these modules due to production requirements up to time t.

Constraint set (12.29) makes sure that an order of modules m is transmitted at time t to the supplier for actually allowing the ordered quantity of module m to be greater than zero. This then allows constraint set (12.30) to impose supplier specified minimum order quantities whenever an order of modules m is transmitted to its supplier at time t.

Constraint sets (12.31) to (12.33) allow supporting collaborations with critical suppliers allowing the enterprise to know and exploit in its planning their key resource constraints so as to insure supply feasibility and minimize their joint supply costs. For example, if a supplier of a specialized module is known to have a production capacity of 10 modules per day, then the enterprise can integrate this knowledge in the assembly plan and therefore avoid both keeping unnecessary product inventory and avoiding supply disruptions associated with infeasibilities due to limited capacity at the supplier site. Similarly, if it is important for a supplier to smooth its loading on a critical resource, then this can be taken into consideration through the planning optimization. The constraint sets permit at the extreme to synchronize the constraints to the assembly center takt time, but allow for aggregating the resource constraints limitations in terms of a time bucket (shift, day,. week, etc.) specified for each critical resource r of critical supplier s. These time buckets can be set to variable durations stated in terms of shifts, days, weeks, and so on through the use of parameters t^s_{srb} and t^e_{srb} setting the start and end times of bucket b for resource r of supplier s.

Constraint set (12.31) transposes the supply requirements into loads on resource r of supplier s during time bucket b. Constraint set (12.32) limits the load on resource r not to exceed its capacity during time bucket b. Constraint set (12.33) permits to model supplier resource smoothing by computing punctual positive and negative deviations from the ideally smoothed average loading of resource r of supplier s during bucket b.

Constraint (12.34) adds up all supply related costs. These include the order costs, the purchase costs, the inventory costs, the critical resource usage costs and the critical resource smoothing costs.

The costs are linearly modelled through constraint (12.34). Variants can be developed to allow, for example, the module purchase cost to be a stepwise function of the quantity ordered exhibiting economies of scale. Similarly, constraint sets (12.28) to (12.34) assume a single supplier per item supplied as is currently the most common case for significant supplied items. The model can be readily upgraded to deal with multiple suppliers per item. This is again left as an exercise for the interested reader.

2.7 Dealing with future dealer demand uncertainty

The mass production oriented demand and supply chain of Figure 12.1 does not allow dealers to reorder after signing their original pre-season order. In such a context, all products to be assembled during the production season are known a priori. On the contrary, in the agile customer-centric demand and supply chain of Figure 12.2, the dealers may order any time they want prior to and during the selling season. This means that at planning time, the enterprise has in its hands a set of orders from dealers and a set of forecasts for future demands. These forecasts can be collected and synthesized from forecasts provided by individual dealers and/or can be generated by the enterprise based on sales history and a variety of predictive indicators.

When continuous ordering is allowed, it does not make sense for the planning to specify production assignments of products to time slots in the assembly centers way ahead in the future, far beyond the range of actual orders, since the future demand is unknown and forecasts are bound to be imperfect. Therefore the production planning horizon is generally much shorter than up to the end of the entire selling season, often in the order of days or weeks. Furthermore, contrary to the more stable and deterministic case of Figure 12.1 where the production planning horizon covers the entire production season and the production plan is only to be re-optimized due to major events in the supply chain such as supply problems, in the agile and customer-centric context of Figure 12.2 the production plan is to be re-optimized in a very frequent rhythm, often daily, in order to adjust to events in the demand chain such as new orders and adjusted forecasts as well as events in the supply chain.

There are multiple ways to model demand uncertainty. In this chapter, it is treated by differentiating four uses for products assembled in a period:

(1) Fulfillment of an actual order from a dealer in a zone;
(2) Expected fulfillment of probable orders within the planning horizon, based on average expected demand per time period for each product;
(3) Amplification of a safety stock specified for allowing fast response to forthcoming yet unknown dealer orders;
(4) Amplification of a speculative stock to deal with future demand beyond the planning horizon.

As an illustrative example, assume that the production planning horizon is set to two weeks and that the current time is set in the early stages of the selling season, before the selling peaks. The first usage covers all officially registered orders from dealers. This may correspond to 25% of the assembly center capacity during the two-week horizon. The second usage corresponds to the expected average consumption of products, forecast to be coming from not yet registered dealer orders during the next two weeks. For example, this may correspond to an average of two units of product 46 per day. The combination of all these forecasts may occupy 40% of the center capacity. This leaves 35% remaining capacity to deal with the third and fourth usages. Relative to the third, given the forecast uncertainty and the required level of service, the enterprise may set a target safety stock to be maintained at all times during the planning horizon. For example, it may set a target safety stock of 10 units of product 46 in the first week and 12 units during the second week. Given the current level of stock of product 46, equal to eight units, then this implies raising it by two units in the first week and two more units in the second week. Adjusting the safety stocks may, for example, require 10% of the center capacity. This leaves 25% remaining capacity for usage four, which is to build anticipatory stock to be used after the planning horizon, to prepare for the forthcoming peaks. If the forthcoming peaks are not so high and that future capacity is expected to be able to handle them, then the enterprise may decide not to produce anything more during the current two-week horizon, as planned at the current time. This decision may be altered in the coming days if new increased forecasts become available. To the contrary, the enterprise may profit from the currently available capacity to build inventory to sustain high future peaks beyond future capacity to handle by itself. The task is then to decide what quantity of each product to assemble for anticipatory stock. For example, the enterprise may decide to increase its production of product 46 by 20 units. The above options leave a lot of decisions to be made which are subjects of the production planning optimization.

The first usage is already modelled through the previous sets of equations, those related to the operational constraints and those related to the dealer network. In order to avoid having to express new sets of constraints, the third usage uses for the in-horizon forecast orders the same variable and constraint sets as the registered orders. The only difference lies in setting the parameters. For example, the revenue per unit can be weighted to express the confidence level in the forecast. A posteriori analysis of the solution is to show that some planned transports to dealer zones are to deal with registered orders while others are mostly potential transports delivering expected forecast orders. The shorter the planning horizon, the less important is to be the set of forecast orders to be added to the set of registered orders.

The constraint sets (12.35) to (12.43) below describe how the second and fourth usages are to be dealt with.

Future dealer demand uncertainty constraints.

$$I_{pt} = i^{s}_{pt} + I^{s+}_{pt} - I^{s-}_{pt}, \qquad \forall p \in P; \forall t \in T^{a} \tag{12.35}$$

$$C^{s-} = \sum_{t \in T^{a}} \sum_{p \in P} c^{s-}_{pt} I^{s-}_{pt} \tag{12.36}$$

$$I^{s+}_{pt^{l}} = \sum_{n \in N^{sf}_{p}} I^{sf}_{pn}, \qquad \forall p \in P \tag{12.37}$$

$$I^{sf}_{pn} \leq i^{sf}_{pn}, \qquad \forall p \in P; \forall n \in N^{sf}_{p} \tag{12.38}$$

$$C^{sf} = \sum_{p \in P} \sum_{n \in N^{sf}_{p}} c^{sf}_{pn} I^{sf}_{pn} \tag{12.39}$$

$$\sum_{t \in T^{a}} E_{t} = \sum_{u \in U} E^{T}_{u} \tag{12.40}$$

$$E^{T}_{u} \leq e^{T}_{u} U_{u}, \qquad \forall u \in U \tag{12.41}$$

$$U_{u} \leq E^{T}_{u-1} / e^{T}_{u-1}, \qquad \forall (u \neq 1) \in U \tag{12.42}$$

$$C^{e} = \sum_{t \in T^{a}} c^{e}_{t} E_{t} + \sum_{u \in U} c^{T}_{u} E^{T}_{u}. \tag{12.43}$$

Constraint sets (12.35) and (12.36) deal with the goal of maintaining target safety stocks for each product p at each time t. Set (12.35) simply contrasts the current inventory level of product p with its target and computes the positive and negative deviations from target. Constraint set (12.36) computes the cost associated with lower than targeted safety stocks. Larger stocks are not disruptive from a safety stock perspective and therefore are not penalized from this perspective. Setting the expected marginal cost for lower than targeted safety stock requires an a priori statistical and economical analysis.

The build up of inventory for dealing with anticipated demand beyond the planning horizon results in an inventory of the product at the last time slot in the planning horizon that is above the level required for safety stock purposes. Therefore the level reached by variable $I^{s+}_{pt^l}$ corresponds to the anticipatory inventory of product p. Constraint sets (12.37) to (12.42) serve the purpose of correctly balancing the compromise between using current production capacity for creating such anticipatory stocks and not using this current production capacity, thus avoiding production and inventory costs.

Consider the first unit of product 46 stocked for anticipation. How much does it cost to stock it? A deterministic answer is impossible since its usage depends on future demands which have not yet materialized. However it is easy to differentiate between two levels of stocked quantities. The first level includes units which are practically certain to be ordered by dealers before the end of the selling season. For example, if the future demand for product 46 for the entire season beyond the planning horizon is forecasted to behave according to a Normal distribution with an average of 300 and a standard deviation of 30, then the 200 first products stocked in anticipatory mode are almost certain to be ordered by dealers. At this first level, the only uncertainty lies in the timing of the eventual order. Based on the time phased forecasts, it is possible to compute the expected duration-of-stay of each additional unit in anticipatory stock.

The second level includes the stocked units that are beyond the practical certainty of eventual demand. At this level, there is a significant probability that a stocked unit may never be ordered by any dealer during the selling season. In the above example for product 46, the 300th unit stocked in anticipation has a 50% chance of never being ordered by a dealer during the selling season. If not sold, it would either have to be dismantled, sold to a bargain market, or kept in stock until the next selling season, to be sold at discounted price in competition with next year's products. Therefore the cost for stocking it must incorporate both the duration-of-stay factor and the probability-of-demand factor. Again, through statistical and economical analysis, the marginal cost of an additional unit in anticipatory stock may be computed at this second level. For practical purposes, in the product 46 example, the second level ends at about 400 units. Beyond that it makes practically no sense to produce any further product unit.

Given the above logic, the enterprise is to prepare a priori an anticipatory stock cost function for each product and to generate a piecewise linear approximation of this convex cost function. Constraint sets (12.37) to (12.39) implement this approximation. For each product, the speculative end-of-planning-horizon inventory is split in n_p^f linear cost segments. Each has a fixed unit cost c_{pn}^{sf} and a maximum allowed inventory i_{pn}^{sf}. The first segment has the lowest unit cost while the nth segment has the highest unit cost. Each segment n of every product p has its inventory variable I_{pn}^{sf} stating the current level of anticipatory inventory at time t. Constraint set (12.37) insures that the sum of the current anticipatory inventories associated with each segment for a product p equals the total anticipatory stock for that product, as expressed by I_{ptl}^{s+}. Constraint set (12.38) simply bounds the segment-specific anticipatory stocks not to exceed the specified maximum for that segment. Constraint set (12.39) cumulates the anticipatory stock costs for all segments of all products.

In the above constraint sets, stocking a product for anticipatory use beyond the production planning horizon has been considered to be a cost. There must be a balancing cost promoting the build up of such stock, otherwise the optimal solution to the problem will never construct such stock. This balancing cost is associated with lost capacity when nothing is produced during a time slot in the assembly center. Consider for example the first time slot in the planning horizon. It is clear that if nothing is planned to be produced during this time, then nothing will ever be produced during this time, and potential capacity will be lost forever.

How does one compute an expected value for this capacity? The answer is similar in nature as what has been explained for the expected anticipatory stock cost. Assume that when summing the forecasts for all products in all dealer zones, the global demand remaining beyond the production planning horizon is forecast to behave according to a Normal distribution with an average of 20,000 units and a standard deviation of 1,000. Now assume that beyond the current production planning horizon, there remains only 15,000 potential time slots in the assembly center, with no possibility to increase that number. Then it is clear that with almost certainty there is a remaining demand of at least 17,000 units. Thus there is lack of 2,000 time slots. If time slots are not found for producing them, then the result will be a loss of at least 2,000 sales and their associated margins. Assuming that the current planning horizon covers ten eight-hour days with a 3-minute takt time, there are 1,600 time slots during the current planning horizon. This means that each one of them not used for producing a product is practically certain to result in a lost sale and associated lost margin.

It is easy to set up another example depicting the other extreme situation where expected subsequent demand is to be so small that producing anticipatory stock cannot make any business sense, as is often the case late in the selling season.

In between these extreme situations it is possible through Monte Carlo simulations and statistical and economical analysis to generate an expected concave cost function for each additional assembly slot left empty during the planning horizon, not being used to assemble a product. The enterprise must then develop a piecewise linear approximation of this cost function. Constraint (12.40) computes the total number of empty time slots in the planning horizon and then redistributes this sum over all cost segments corresponding to the piecewise linear approximation of the concave cost function. Constraint set (12.41) simply states whether or not a cost segment u is used, with a greater than zero membership. Constraint set (12.42) insures the validity of the linear relaxation of the concave cost function by insuring that a lower numbered cost segment is fully used prior to allowing the opening of the next cost segment.

Constraint (12.43) adds up all the unused capacity costs over all cost segments. It also adds up a sum of unused capacity costs specific to each time slot. The reasoning for these costs is to factor in the fact that when using a rolling horizon with frequent re-optimizations, then among all time slots of a given planning horizon, the first slots are more costly to leave empty than the latest slots. The first time slots, if not planned to be used for production, have a probability of one of never being used, their potential capacity gone forever. So in a planning horizon, if a time slot is to be planned to be unused, it is preferable for that time slot to be among the last slots rather than the first slots. By computing an expected unused marginal cost differentiation among time slots, the enterprise is in a position to set cost parameters c_t^e for each time slot. The contribution of these costs is then summed in constraint (12.43), added whenever a time slot t is not planned to be used for production.

3. Production planning optimization in alternative demand and supply chains

This section examines the production planning optimization modeling of assembly centers in gradually more collaborative, customer-centric, agile and personalized demand and supply chains. It describes the implications of these transformations as well as of production planning optimization knowledge and technology.

3.1 Pushy and rigid demand and supply chain

The rigid demand and supply chain of Figure 12.1, in its pushy variant which does not take care of delivery timing preferences of dealers, still generically faces the entire production planning problem described in section two, except the portions dealing with the dealer network and with dealer demand uncertainty.

The assembly center(s) of such chains are generally designed from a mass production paradigm. They are efficient at assembling a product once setup for it. However the changeover from one product to the next can require highly significant time and generate significant costs. It is common to let a product run for days prior to switching to the next. In fact a product is often run only a few times, if not a single time, during the production season. The operations and changeover constraint sets (12.2) to (12.11) are therefore at the core of the production planning problem.

In order to size the complexity of the problem, assume that only these constraint sets are considered by the planning optimization, and that the enterprise imposes a single run per product. Then the problem reduces to the Traveling Salesman Problem, well known to be NP-complete, where a city becomes a product, the travelling distance between pairs of cities becomes the inter-product changeover cost (including the cost of lost time slots in the assembly center during the changeover), the traveling salesman becomes the assembly center, and the objective of touring all cities in minimal travelled distance becomes the objective of touring all products in minimal overall changeover cost. In the contexts studied in this chapter, it is common to deal with hundreds of products. Furthermore, the single run per product is an extreme solution which can be examined but is not to be a priori imposed except in very precise conditions guaranteeing its optimality. Such conditions involve a complete dominance of changeover costs and an absolute ignorance of dealer delivery preferences.

The sheer complexity and size of the overall problem helps understand why in most cases, the planning decisions result from a decomposition of the problem, from relying on decision rules and heuristics, and from heavy reliance on human intelligence.

Problem decomposition generically assumes away many constraint sets when generating the master assembly plan, which becomes an input to the other sub problems, according to a hierarchical planning strategy. Supply and personnel planning become subordinates of the master assembly plan resulting from the production planning optimization. There is heavy reliance on hierarchical decomposition in practice. Such decomposition makes sense when there is indeed a dominance of the operational and changeover variables in terms of feasibility and cost optimization. However the personnel and supply feasibility issues and costs are often significant.

As described in the introduction to personnel constraints sets (12.12) to (12.17) and supply constraints sets (12.27) to (12.30) and (12.34), there are intricate relationships between these and the operational and changeover constraints sets (12.2) to (12.11). An example stemming from the fact that such enterprises are often among the biggest employers in their region, is the signature of social or union contracts guaranteeing that the assembly center is to maintain as stable a workforce as possible, with a fixed bottom level and penalties for not achieving its engagements. This imposes constraints which affect the feasibility and profitability of production plans. Lack of in-depth knowledge of these relationships and their impact, coupled with the lack of adequate planning optimization technology to support the effective integrated treatment of the integrated problem formulation, are the key factors inhibiting their consciously integrated treatment.

3.2 Collaborative yet rigid demand and supply chain

When considering the more dealer-collaborative variant of the rigid demand and supply chain of Figure 12.1, then the problem statement needs to take explicit consideration of constraint sets (12.18) to (12.26). On the one hand, it transforms the problem from a cost minimization perspective to a profit maximization perspective since the timing and constitution of shipments to dealer zones have direct impact on the revenue stream. On the other hand, it forces to understand the implications of not satisfying dealer preferences on the profitability of the enterprise. These involve the following:

- Each dealer aims to have early on in his possession the set of products maximizing his expected showcasing effect, thus maximizing his customer attraction potential and his revenue expectations.
- Each dealer aims not to have product over stock in order to avoid both product financing and storage costs and minimizing security risks.
- Each dealer, when feeling badly treated by the enterprise, has higher probability to lose brand loyalty and switch to sell products of competitors, bringing with them a significant percentage of their customers, thus having long term impact on sales by the enterprise.
- Each dealer is a business unit that is a flagship of the enterprise in the region he deserves. The final customers interact with the enterprise

through the dealer. The relative prosperity of the dealer generally reflects on the perception of the enterprise by the final customer, thus affecting purchasing probability.

In practice, integration of dealer network considerations in the production planning optimization, when attempted, is generally limited to insuring a first release of top products for early showcasing effect and dealing with the most important dealers (or dealer groups) which have significant weight on the enterprise sales. Again lack of in-depth knowledge, resulting on misperceptions relative to the potential gains for the committed efforts, as well as lack of planning optimization technology, restrains enterprises to exploit more collaboration with dealers.

Even in the pushy variant where products are shipped to dealers as soon as produced, subject to localized vehicle transport optimization, enterprises should understand the impact of delivery schedule to dealers on the revenue stream. They can profit from integration of constraints (12.18) to (12.26), setting all delivery preferences as soon as possible and the deviation costs to zero. This would help them optimize revenues and minimize transport costs.

Collaboration can be established on both the demand and the supply sides (Montreuil et al., 2000). Exchange of status information with suppliers and subcontractors, coupled with the establishment and modeling of their key constraints and smoothing objectives, lead to generate constraint sets (12.31) to (12.33) and to refine cost constraint (12.34). This collaboration allows suppliers to not protect themselves with extensive lead times and high minimal quantities, knowing that the enterprise takes explicit care of its constraints, costs and objectives in its production planning optimization. By systematically developing the collaboration with all critical suppliers and subcontractors, the production plan has the potential to reach levels of profitability not attainable without collaboration.

3.3 Customer-centric, agile and collaborative demand and supply chain

In between the demand and supply chains of Figures 12.1 and 12.2 lies the potential for increasing the agility and customer-centricity of a collaborative demand and supply chain organized such as in Figure 12.1, yet offering continuous ordering throughout the selling season.

In the context studied in this chapter, extreme agility involves:

- No significant changeover times and costs;
- Highly scalable operations allowing up and down swing in production level with negligible costs and constraints;

- Highly polyvalent workforce with negligible constraints and costs on manpower level modification;
- Insignificant supply lead time and constraints from suppliers and subcontractors;
- Non constraining transportation capabilities.

When extreme agility is achieved, most of the constraints melt away and production can be operated with minimal planning, mostly in a sense and respond model, producing just-in-time in pull mode, keeping minimal safety stocks to allow instant delivery when required. In such a case the operations are in a perfect position to enable the enterprise to be customer-centric, delighting both dealers and customers.

The problem lies in the fact that in most cases some degree of agility is reached, yet many constraints remain. Therefore all constraint sets (12.2) to (12.34) may be potentially needed, but applied to a smaller number of entities (e.g., lower number of modules and suppliers) and with more limited impact (e.g., smaller number of worker types, each capable of performing a wider scope of tasks).

Customer centricity involves putting the goal of delighting customers at the forefront of the business preoccupations, meeting or exceeding their expectations in terms of product offering and availability, delivery speed and price, anytime prior or during the main selling season. The most direct implication is the importance of allowing dealers to order throughout the selling season to respond to customer demand, rendering impossible the termination of the production season prior to the main selling season. High availability of all products requires the maintnance of a safety stock and to be rapidly able to replenish stock in light of consumed demand by dealers. It also may involve the building of anticipatory stocks ahead of time in order to maintain high availability and fast delivery throughout the season peaks.

Such a demand and supply chain has to dynamically update its production plan, perhaps every day depending on the market dynamics. It also has to recognize explicitly that it deals with uncertain demand. These facts lead it to have to face such constraints as those in sets (12.35) to (12.43). In such a setting, the production planning horizon may, for example, be set to a few weeks. During these weeks, the production plan is precisely optimized, setting the production assignment to each time slot in the assembly center. The production plan optimization attempts to best decide on what products to produce at each time or yet when not to produce anything or start a changeover to another product. As stated in Section 2.7, it must choose between fulfilling actual and forecast orders during the planning horizon, amplifying the safety stock of selected products, or building anticipatory stocks to deal with future demand beyond the production planning horizon.

3.4 Personalized, customer-centric, collaborative and agile demand and supply chain

The demand and supply chain outlined in Figure 12.2 adds personalization to the characteristics described in Section 3.3 above. This means that instead of imposing a fixed product mix to the customers, often composed of a few hundred products, the enterprise lets customers personalize products according to their needs and tastes. Section 1 explicitly states the personalization offer in the illustrative example. According to personalization framework introduced by Montreuil and Poulin (2004), it combines popularizing, accessorizing and parametering. The popularizing option aims to offer a limited set of popular products to be highly available on the shelves so as to satisfy the needs of customers wanting their product *right now.* The accessorizing option offers a series of ready-to-personalize products coupled with a variety of accessory modules allowing each customer to personalize his product. The parametering option offers personalized products defined by the customers through parameter and option settings.

In order to deliver the personalization offer, the demand and supply chain of Figure 12.2 is transformed by the addition of fulfillment centers and the specialization of the main assembly center. The assembly center is to produce popular products, ready-to-personalize products and parametered products. The fulfillment centers assemble accessorized products from ready-to-personalize products and sets of accessories. Both types of centers are agile and highly polyvalent personnel, requiring no changeover time when switching from a product to the next, yet they both have limited capacities. Both types collaborate with both their demand side and supply chain partners.

From a production planning perspective, the fulfillment centers cannot produce an accessorized product without a customer having actually ordered that product. Therefore their planning horizon corresponds to the time required to go through the order booking. The key decisions involve the sequencing of products to be accessorized, mostly to best compromise between overall manpower smoothing, supply availability and cost from suppliers and from the assembly center, and delivery promises and transport to dealers.

Relative to the fulfillment centers, the production planning optimization problem has the following characteristics:

- Operational constraint sets (12.2) to (12.4), (12.9) and (12.10) are enforced, with no consideration for changeovers. All changeover constraints are not to be enforced.
- Personnel constraints (12.12) to (12.17) are imposed, yet with a lower number of worker types.
- Dealer network constraints (12.18) to (12.26) are still necessary, yet the proximity to market of the fulfillment centers shortens the trip times to dealer zones. Also, dealer zones for fulfillment centers are to be much smaller than in the case of Figure 12.1.
- Supply network constraints (12.27) – (12.34) are to be imposed, maximally exploiting collaboration with partners, on one side with accessory suppliers and on the other side with the assembly center.
- Even though demand is uncertain and planning is to be dynamically updated through a rolling horizon, constraint sets (12.35) to (12.39) can be discarded since there is no possibility to stock not yet ordered products. Constraint sets (12.40) to (12.43) may be imposed to ensure that leaving an empty production slot is to be cost adequately through the production planning optimization.

In this personalized setting, the clients of the assembly center have become the fulfillment centers rather than the dealers themselves. Even though it has to wait for an actual order from a fulfillment center to produce a parametered product, it can produce popular products as well as ready-to-personalize products prior to getting actual orders for them from fulfillment centers. Therefore its planning horizon is to be longer than the planning horizon of fulfillment centers. The production planning optimization has the following characteristics:

- Operational and personnel constraints (12.2) to (12.4), (12.9), (12.10) and (12.12) to (12.17) are to be imposed as in the fulfillment centers.
- The dealer network constraints (12.18) to (12.26) become the fulfillment center network constraints. The zones correspond to a single fulfillment center.
- Supply network constraints (12.27) – (12.34) are to be imposed, maximally exploiting collaboration with suppliers.
- Constraint sets (12.35) – (12.43) are to be imposed as described in Section 2.7, replacing dealers by fulfillment centers.

For both the fulfillment centers and the assembly center, collaboration with very agile suppliers and subcontractors can permit to take them out of the production planning optimization model since they are not constraining. Fast and accurate information transfer with them is then sufficient for them to supply the centers with the required modules in time for their assembly into the products being manufactured. Similarly, in the case of suppliers having very limited influence on the production plan optimality, the supply relationship can be decoupled, operated in pull mode using a dynamically updated kanban system insuring sufficient stock to avoid shortages.

4. Conclusion

First the chapter has introduced a spectrum of demand and supply chain alternatives for high value consumer products, varying in terms of collaboration, customer centricity, agility and personalization. Second it has introduced a comprehensive production planning optimization enabling to adequately model these alternatives. Third it has provided a thorough analysis of production planning optimization modeling as a demand and supply chain is transformed to incorporate more collaboration, customer centricity, agility and personalization. It has also discussed the impact of production planning optimization knowledge and technology.

It is shown that some of the transformations increase the complexity of the model to be solved while some others decrease this complexity.

- Collaboration increases model complexity, yet this added complexity allows to achieve better global optimization.
- Agility generally decreases model complexity by removing constraints and imposing less restrictive parameters. It generally leads to improved global optimization.
- Customer centricity has a double effect. On one hand it decreases model size by switching from rare optimization using long planning horizon to frequent optimization using a shorter planning horizon. On the other hand it increases model complexity in order to deal adequately with the inherent market uncertainty involved in attempting to delight customers through the dealers.
- Personalization generically increases modeling complexity both due to the explosive product scope and the required structural transformation of the demand and supply chain. However some centers end up with more simple models due to the fact that they end up producing strictly to order.
- Lack of production planning optimization knowledge generally results in lower model complexity, and mostly in model inadequacy, through the ignorance or inadequate representation of constraints and cost factors. Entire panes of modeling can end up ignored or delegated to being imposed production planning decisions. It results in potential for lower global optimization.
- Lack of available adequate production planning optimization technology results in problem decomposition in order to avoid having to deal

with the overall complexity of the global problem. This inherently leads to global sub-optimization.

The chapter opens many avenues for further research. Below are listed a few promising avenues for the research community:

- The introduced comprehensive production planning model for assembly centers in demand and supply chains is not yet solved in an integrated manner, either to global optimization for small cases or heuristically for larger cases.
- The impact of networked collaborative and decomposition approaches to dynamically address production planning of assembly centers in demand and supply chains is not empirically studied.
- The impact of agility, customer centricity and personalization on production planning of assembly centers in demand and supply chains is not empirically studied.
- Simulation technologies are needed to enable the efficient realization of empirical studies as stated above, coupling optimization and stochastic system dynamics, and enabling adequate representation of all stakeholders in the demand and supply chain, from customers and dealers to production planners and suppliers.
- The current model should be expanded to integrate engineering issues related to new product introduction, yearly product evolution and on going engineering changes.
- The type of research reported in this chapter should be performed for other important demand and supply chain contexts.

The chapter brings to light important insights for the professional community:

- Production planning of assembly centers in demand and supply chains is a complex optimization problem having major financial and feasibility impacts. It lies at the core of a center contribution to the enterprise performance and should be addressed accordingly.
- Transformations in the demand and supply chain have significant impacts on the production planning problem formulation. They alter constraints, cost and operational parameters, the degrees of freedom, and the essence of the objective function. Overall they affect its complexity, its size and its profitability potential.
- Emphasis should be put on adequate training of managers and planners, making sure that they have the knowledge and in-depth understanding required to adequately address production planning in such contexts. This is *not* a minimal impact problem to be solved by untrained, unprepared, under equipped administrative personnel.

- Current and proposed production planning optimization technology should be carefully audited as to the degree with which it supports an adequate comprehensive modeling and solution. The technology should fit the needs. Gaps may have significant impacts on global optimization and enterprise performance.

Acknowledgements This research has been supported by the Canada Research Chair in Enterprise Engineering. Special thanks to Caroline Cloutier and Marc Poulin for valuable editorial help.

References

Hoover, W.E., Jr., Eloranta, E., Holmström J., and Huttunen, K. (2001). *Managing the Demand-Supply Chain: Value Innovations for Customer Satisfaction.* Wiley.

Montreuil B., Frayret J.-M., and D'Amours, S. (2000). A strategic framework for networked manufacturing. *Computers for Industry*, 42:299 - 317.

Montreuil, B. and Poulin M. (2004). Demand and supply network design scope for personalised manufacturing. Forthcoming in *International Journal of Production Planning and Control.*

Poulin, M., Montreuil, B. and Martel, A. (2004), Implications of personalization offers on demand and supply network design: A case from the golf club industry. Forthcoming in*European Journal of Operational Research.*