CHAPTER FOUR

PRODUCTION PLANNING AND CONTRON

4.1. Introduction

You are familiar with the word planning if you are not don‟t worry. Planning is an integral part of a manager‟s job. If uncertainties cloud the planning horizon, it can be quite difficult for a manager to plan effectively. Forecasts can help managers by reducing some of the uncertainty, thereby enabling them to develop more meaningful plans than they might otherwise. This unit deals with business forecasting and aggregate planning. It covers basic forecasting techniques, how to monitor a forecast, the necessary steps in preparing a forecast, and elements that are common to forecasts, techniques of forecasting. Moreover, it addresses the concept of aggregate planning, techniques of aggregate planning, master scheduling and controlling.

4.2. AGGREGATE PLANNING

4.2.1. Concept of aggregate planning

Aggregate plan: statement of a company‟s production rates, workforce levels, and inventory holding based estimates of customer requirements and capacity limitations.

Production plan: a manufacturing firm‟s aggregate plan, which generally focuses on production rate and inventory holdings.

It determines the quantity and timing of production for the immediate future. Aggregate planning is an intermediate term planning decision. It is the process of planning the quantity and timing of output over the intermediate time horizon (3 months to one year). Within this range, the physical facilities are assumed to be fixed for the planning period. Therefore, fluctuations in demand must be met by varying labor and inventory schedule. Aggregate planning seeks the best combination to minimize costs.

4.2.2. Typical objectives

The many functional areas in an organization that give input to the aggregate plan typically have conflicting objective for the use of the organization‟s resources. Six objectives usually are considered during development of a production or staffing plan, and conflicts among them may have to resolved:

1. Minimize cost/maximize profit: if customer demand is not affected by the plan, minimizing cost will also maximize profit.
2. Maximize customer service: improving delivery time and on time delivery may require additional workforce machine capacity or inventory resources.
3. Minimize inventory investment: inventory accommodations are expensive because the money could be used for more productive investments.
4. Minimize changes in production rate: frequent change in production rates can cause difficulties in coordinating the supplying of materials and required production line rebalancing.
5. Minimize changes in workforce levels: fluctuating workforce levels may cause lower productivity because new employees typically need time to fully productive.
6. Maximize utilization of plant and equipment: process based on a line flow strategy requires uniformly high utilization of plant and equipment.

The weight given to each one in the plan involves cost tradeoffs and consideration of non-quantifiable factors. Balancing these various objectives to arrive at an acceptable aggregate plan involves consideration of various alternatives. The two basic types of alternatives are reactive and aggressive.

For effective and efficient aggregate planning process the following points are required:

- A logical overall unit for measuring sales and output
- A forecast of demand for an intermediate planning period in these aggregate terms
- A method for determining costs
- A model that combines forecasts and costs so that scheduling decisions can be made for the planning period.

4.2.3. Aggregate Planning Strategies

The variables of the production system are labor, materials and capital. More labor effort is required to generate higher volume of output. Hence, the employment and use of overtime (OT) are the two relevant variables. Materials help to regulate output. The alternatives available to the company are inventories, back ordering or subcontracting of items.

These controllable variables constitute pure strategies by which fluctuations in demand and uncertainties in production activities can be accommodated by using the following steps:

- **Vary the size of the workforce:** Output is controlled by hiring or laying off workers in proportion to changes in demand.
- **Vary the hours worked:** Maintain the stable workforce, but permit idle time when there is a slack and permit overtime (OT) when demand is peak.
- **Vary inventory levels:** Demand fluctuations can be met by large amount of inventory.
- **Subcontract:** Upward shift in demand from low level. Constant production rates can be met by using subcontractors to provide extra capacity.
- **Change prices** or other factors to influence demand
- **Backlogs**, backorder and stocks:

4.2.4. Planning strategies

Managers often combine reactive (workforce adjustment, anticipation of inventory, workforce utilization, vacation schedules, subcontractors and Backlogs, backorder and stocks) and aggressive (complementary products, and creative pricing) alternatives in various ways to arrive at an acceptable aggregate plan.

Chase strategy: a strategy that matches demand during the planning horizon by varying either the workforce level or the output rate. When a chase strategy uses the first method, varying the workforce level to match demand, it relies on just one reactive alternative-workforce variation. This chase strategy has the advantage of no inventory investment, overtime, or under-time. However, it has some drawback, including the expense of continually adjusting workforce levels, the potential alienation of the workforce, and the loss of productivity and quality because of constant change in the workforce.

The second chase strategy, varying the output rate to match demand, opens up additional reactive alternatives beyond changing the workforce level. Sometimes called the utilization strategy, the\ extent and timing of the workforce's utilization is changed through overtime, under- time and vacation are taken. Subcontracting, including temporary help during the peak season, is another way of matching demand.

Level strategy: a strategy that maintains a constant workforce level or constant output rate during the planning horizon. When a level strategy used the first method, marinating a constant workforce level, it might consist of not hiring or laying off workers (except at the beginning of the planning horizon), building up anticipation inventories to absorb seasonal demand fluctuations, using under-time in slack periods and overtime up to contracted limits for peak periods.

When level strategy uses the second method, maintaining a constant output rate, it allows hiring and layoffs in addition to other alternatives of first level strategy. The key to identifying a level strategy is whether the workforce or output rate is constant.

Mixed strategy: Strategies that consider and implements a full range of receive alternatives and goes beyond a "Pure" chase or level strategy. Whether management chooses a pure strategy or some mix the strategy should reflect the organizations environment and planning objectives.

4.2.5. The planning process

Determining demand requirements: the first step in planning process is to determine the demand requirements for each period of planning horizon. The planner can derive future requirements for finished goods from backlogs (for make to order operations) or from forecasts for product families made to stock (for make to stock operations).

Identifying alternatives, constraints, and costs: the second step is to identify the alternatives, constraints, and costs for the plan.

Constraints represent the physical limitations or managerial policies associated with the aggregate plan. Typically, many plans can satisfy specific set of constraints. The planner usually considers several types of costs when preparing aggregate plans:

- Regular time costs (wages health insurance, dental care, social security, and retirement funds pay for vacation, holidays, and certain other types of absence).
- Overtime costs:
- Hiring and layoff costs
- Inventory holding costs
- Backorder and stock out costs

Preparing an acceptable plan: developing an acceptable plan is an iterative process i.e. plan may need to go through several revisions and adjustments. A prospective, or tentative, plan is developed to start. The plan must then be checked against constraints and evaluated in terms of strategic objectives.

Implementing and updating the plan: the final step is implementing and updating the final plan. Implementation requires the comment of manager in all functional areas. The planning committee may recommend changes in the plan during implementation or updating to balance conflicting objectives better. Acceptance of the plan does not necessarily mean that everyone is in total agreement, but it does imply that everyone will work to achieve it.

4.2.6. Aggregate planning with mathematical method

The large cost involved with aggregate plans is a motivation to seek the best possible plan. Several mathematical methods can help with this search process.

Transportation method of production planning: the use of transportation method to solve production planning problems, assuming that a demand forecast is available for each period, along with workforce level plan for regular time. Capacity limits on over time and subcontractor production also are need for each period. The other assumption is all costs are linearly related to the amount of goods produced –that is, that a change in the amount of goods produced creates proportionate change in costs. With these assumptions, the transportation method yields the optimal mixed strategy production plan for the planning horizon.

Production planning with backorder:

We start with a table called tableau of the workforce level, capacity limits demand forecast quantities beginning inventory level, and costs of each planning horizon.

Each raw in the tableau represents an alternative for supplying output. The columns represent the periods that the plan must cover, plus the unused and total capabilities available. The box in the upper right hand side corner of each cell shows the cost of producing a unit in one period and in some cases, carrying the unit in inventory for sale in future period.

Use the following procedure to develop an acceptable aggregate plan

Step1. Select a workforce adjustment plan (Rt values) using a chase strategy, level strategy, or a mixed strategy. Identify the capacity constraints on overtime (Ot values) and on subcontracting (St values). Usually a period‟s overtime capacity is a percent of its regular-time capacity. Also identify the no-hand anticipation inventory (Io values) currently available before the start of the period.

Step2. Input the cost parameters (h, r, c, s, u and b) for the different reactive alternatives

Step3. Forecast the demand for each future period, and insert the forecasts as the values in the tableau‟s last row. The last period‟s requirement should be increased to account for any desired inventory at the end of the planning horizon. The unused capacity cell in the last raw equals the total demand requirements in the last raw, minus the total capacity in tableau‟s last column.

Step 4. Solve the transportation problem just formulated with a computer routine to find the optimal solution (based on the workforce adjustment plan). The sum of all entries in a row equals the total capacity for that raw, and the sum of all entries in a column must equal the requirement in the last column.

Step5. Return to step 1 and try other staffing plans until you find the solution that best balances cost and qualitative consideration.

Linear programming for production planning

Linear programming models for production planning seek the optimal production plan for a linear objective function and a set of linear constraints; that is, there can be no cross products or powers of decision variables or other types of nonlinear terms in the problem formulation. Liner programming models can be used to determine optimal inventory level, backorders, subcontractor quantities, production quantities overtime production, hires and layoffs. The main drawbacks are that all relationships between variables must be linear and that the optimal values of the decision variables may be fractional.

4.3. SCHEDULING

4.3.1. Concepts of scheduling

Now you can understand the concept of scheduling. Scheduling is the processes of determining the starting and completion times to jobs. It is the determination of when labor, equipment and facilities are needed to produce a product or provide a service. Scheduling is a time table for: performing activities, using resources, or allocating facilities. Generally, a schedule specifies the timing and sequence of production. Schedule must be realistic; that is they must be capable of being achieved within the capacity limitations of the manufacturing facilities. Scheduling should be clearly differentiated from aggregate planning. The purpose of scheduling is to ensure that available capacity is efficiently and effectively used to achieve the organization's objectives. The purpose of aggregate planning is to determine the resources (labor, equipment, space etc.) that should be acquired for scheduling. Often several jobs might be processed at one or more work stations. Typically a variety of tasks can be performed at each work station which make effective scheduling a must rather than an alternative. Sequencing & loading should be considered during scheduling activity.

Sequencing: sequencing is concerned with developing an exact order (or sequence) of job processing. It is the determination of the order in which jobs are processed. One of the oldest sequencing methods is the Gant chart. Gant chart is a bar chart that shows a job‟s progress graphically or compares actual against planed performance.

***Loading*:** is the assignment of work to specific resources/ machines. It is simply the process of assigning work to individual workers or machine. Loading can be: finite loading, infinite loading, back ward loading and forewarned loading.

- *Finite loading*: refers to loading activities with regard to capacity. Tasks are never loaded beyond capacity.
- *Infinite loading*: loading activity without regard to capacity.

- *Back ward loading* begins with the due date for each job and loads the processing time requirements against each work centre by proceeding back ward in time. The purpose of back ward loading is to calculate the capacity required in each work centre for each time period.
- *Forewarned loading* begins with the present date and loads jobs forwards in time. The processing time is accumulated against each work centre, assuming infinite capacity. The purpose of forewarned loading is to determine the approximate completion date of each job and the capacity required in each time period.

4.3.2. Objectives of scheduling

The objectives of scheduling include minimizing: job lateness, response time, completion time, overtime, idle time, work in process inventory, and maximizing machine and labor utilization.

Scheduling is common for manufacturing as well as service sectors.

4.4. Scheduling in manufacturing

Scheduling in manufacturing is the process of assigning priorities to manufacturing orders and allocating workloads to a specific work centers. Scheduling is challenging if the task variety is high. This is a case particularly for the job shop scheduling. In the following discussion, we will concentrate on scheduling issues for job shop production.

4.4.1. Job shop scheduling

For job shop production scheduling decision can be quite complex. What makes Scheduling so difficult in a job shop is the variety of jobs (customer orders) that are processed, each with distinctive routing & processing requirement. In a pure job shop, there are several jobs to be processed, each of which may have different routing among department or machines in the shop. In designing a scheduling and control system for a high variety of activities, sequencing and prioritizing should be emphasized.

4.4.2. Priority rules for allocating jobs to machine- sequencing

As discussed above, sequencing is prioritizing jobs that have been assigned to limited resources. Sequencing is simple if work centers are lightly loaded and need the same processing time. But if work centers are heavily loaded there will be longer waiting time and idle time. In this case to minimize the waiting and idle time, we must prioritize the tasks by using priority rules. Priority rules are the criteria by which the sequence of jobs is determined. The process of determine which job is started first on a particular machine or work center is known as priority sequencing. Some of the more common priority rules for sequencing jobs are:

- First come, first serve (FCFS): orders are run in the order that they arrive in the department i.e. the oldest first rule.
- Shortest processing time (SPT): run the job with the shortest completion time first i.e. shortest operating time first.
- Earliest due date first (due date): run the job with earliest due date first. Thus, a job that is due tomorrow has a higher priority than the job that is due next week or next month.
- Critical ratio (CR): this is calculated as the difference between the due date and the current date divided by the work remaining. Orders with the smallest CR are run first. In the CR rule, jobs are sequenced from lowest CR to highest CR. Those with a CR less than one are considered behind schedule and need to be expedited. And CR greater than one implies that the job is ahead schedule and can be de-expedite.

$$CR = \frac{\text{due date – today's date}}{\text{Total shop time remaining}} = \frac{\text{time remaining}}{\text{lead time remaining}}$$

- Last come first serve (LCFS): this rule occurs frequently by default. As orders arrive they are placed on the top of the stack and the operator usually picks up the order on top to run first.
- Random order – whim: the supervisors or the operators usually select which ever job they feel like running.
- Queue ratio (QR): this is calculated as the slack time remaining in the schedule divided by the planned remaining queue time. Orders with the smallest QR are run first.
- Slack time remaining (STR): This is calculated as the difference between the times remaining before the due date minus the processing time remaining. Obviously, job with negative slacks are behind schedule; those jobs might require expediting to get caught up. Under the minimum slack rule, jobs are sequenced based on their slack: those with the

most slack receive the lowest priority and those with the least slack receive the highest priority i.e. Orders with the shortest STR are run first.

- Slack time remaining per operation(STR/OP):

$$\text{STR/OP} = \frac{\text{time remaining before due date} - \text{remaining processing time}}{\text{Number of remaining operations}}$$

4.4.3. Performance measures of job shop scheduling

From the operations manager‟s perspectives, identifying performance measures to be used in selecting a schedule is important. If the overall goals of the organizations are to be achieved the schedule should reflect managerially acceptable performance measure. The following describes the most common performance measure used in operation schedule.

✓ *Job flow time*: refers to time a job spends in the shop. It is the sum of the moving time between operations, waiting time for machines or work orders, process time (including set up), and delays resulting from machine breakdowns, component unavailability and the like. The objective here is to minimize the average flow time.

Flow time = waiting time + processing time

Average flow time = (sum of total flow time) ÷(no of job processed)

✓ *Make span time*: refers to time to process a set of jobs. It is the total amount of time required to complete a group of job. The general objective is to minimize the make span time. (Generally, make span time =Σprocessing time of each job)

✓ *Tardiness (past due)*: refers to the amount by which completion time exceeds the due date of a job. If a job is completed before its due date, tardiness is zero. The objective is to minimize the number of tardy jobs. Tardiness= Σ past due, & (Average *tardiness* = Σtimes past due ÷total no. of jobs)

✓ *Work in process inventory*: any job in waiting line, moving from one operation to the next, being delayed for some reason, being processed and the like are work in process. This measure can be expressed in units, number of jobs, and birr value for the entire system and so on. (Average WIP inventory =Σjob flow times÷makes pan times)

✓ *Total inventory in the system*: is the sum of the schedule receipt and on hand inventories. (Average total inventory = Σ jobs in the system÷makes pan time)

✓ *Utilization* = total processing time ÷total flow time

✓ *Average number of jobs in the system* $= \dfrac{\Sigma(\text{no. of jobs})\ (\text{production time of the job})}{\text{Makes pan time}}$

4.4.4. Job shop scheduling techniques and priority rules.

There are a number of job scheduling techniques. However, we will take two most widely practiced job scheduling techniques: scheduling „n" jobs – one machine problems and scheduling „n" jobs on two machines.

i. scheduling n jobs on one machine(n/1cases).

Let us compare some of the priority rules in a static scheduling situation involving some jobs and one machine. The theoretical difficulty of this type of problem increases as more machines are considered.

Example 1 Mesfin industrial engineering has engaged in assembling different model cars. Assume five customers submitted their orders for different model of cars: the first customer order model type A, the seconded model type B, the third model type C, the fourth for model type D and the fifth order model type E. Further assume that all order requires the use of only one machine which forces the company to decide on the processing sequence for the five orders. The following is the time (in days) required to complete the job on in the machine.

Job (in order of arrival)	*processing time (in days)*	*No .of days till due or due date (days hence)*
Model A	3	5
Model B	4	6
Model C	2	7
Model D	6	9
Model E	1	2

Required: based on FCFS rule, determine makes pan time, total flow time, average flow time, average tardiness (average job lateness), average WIP inventory, average total inventory, utilization & no. of jobs in the system.

Solution The FCFS rule results in the following computation.

Job order	Begin work	processing_Time(in days)	job flow time	scheduled customers pick up time	actual customer pick up time	days early	days past due
A	0	3	3	5	5	2	-
B	3	4	7	6	7	-	1
C	7	2	9	7	9	-	2
D	9	6	15	9	15	-	6
E	15	1	16	2	16	-	14

Make span time = 3+4+2+6+1= 16

Total flow time= 3+7+9+15+16=50

Average total flow time= 50□5 = 10 days

Average tardiness= 0+1+2+6+14□5 = 4.6 days

Average WIP inventory= 50□ 16 = 3.125

Average total inventory = (5+7+9+15+16) □16= 3.25

Utilization = 16 □ 50 = 0.32= 32%

No. of jobs in the system = (3X5) + (4X4) + (2X3) + (6X2) + (1X1) □ 16 = 3.125

Example 2: ABC factory is engaged in gentle men garment business. The order receiving unit has received the orders from its customers for coats, gowns, shirts, trousers, and jackets respectively. There is only one job centre that all kinds of jobs are performed. The following is the time (in days) required to complete the job in the work centre.

Job order	*Processing time(in days)*	*No of days till due*
Coats(C)	7	8
Gown (G)	2	3
Shirt(S)	5	7
Trouser (T)	3	9
Jacket(J)	6	6

Required:

i. Using CR priority rule, schedule the jobs.

ii. Evaluate the in terms of performance measure i.e. make span time, average day early, average job lateness, average WIP inventory, average total inventory utilization and no. of jobs in the system.

Order	No of days till due (a)	processing time (b)	CR= (a÷ b)
C	8	7	1.14
G	3	2	1.5
S	7	5	1.4
T	9	3	3
J	6	6	1

From the highest critical to lowest critical ratio, the job can be scheduled as: J, C, S, G, T in the order of 1 st , 2nd , 3rd , 4th , 5th respectively.

Job order	Begin work	processing Time(in days)	job flow time	scheduled customers pick up time	actual customer pick up time	days early	days past due
J	0	6	6	6	6	-	-
C	6	7	13	8	13	-	5
S	13	5	18	7	18	-	11
G	18	2	20	3	20	-	17
T	20	3	23	9	23	-	14

Now , you can determine make span time, total flow time, average flow time, average tardiness (average job lateness), average WIP inventory, average total inventory, utilization & no. of jobs in the system as follows.

Make span time =6+7+5+2+3=23 days

Total flow time = 6+13+ 18+ 20+23= 80 days

Average flow time = 80 □5 = 16 days

Average tardiness = (5+ 11+ 17+14) □5= 9.4 days

Average WIP inventory = 80 □23= 3.478

Average inventory= (6+13+18+20+23) □23= 3.478

Utilization = 23□80 0.2875= 28.75%

No. of jobs in the system= (6x5) + (7x4) + (5x3) + (2x2) + (3x1) □23= 3.478

Exercise : Based on the above data, determine make span time, total flow time, average flow time, average tardiness (average job lateness), average WIP inventory, average total inventory, utilization & no. of jobs in the system using:

i. FCFS ii. LCFS iii. EDD, iv. SPT

4.5. Scheduling in services

Scheduling is also important in service organizations. For example nurses must be scheduled in hospital, and truck must be scheduled for deliveries for furniture distributors. One important distinction between manufacturing and services that affects scheduling is that service operations cannot create inventories to buffer demand uncertainties. A second distinction is that in service operations demand often is less predictable; customers may decide on the spur of the moment that they need a hamburger, a hair cut or a plumbing repair. Thus capacity, often in the form of employees is crucial for service providers. In this section we discuss various ways in which schedule systems can facilitate the capacity management of service providers.

4.5.1. Scheduling customer demand

One way to manage capacity is to schedule customers for arrival times and definite periods of service time. With this approach, capacity remains fixed and demand is leveled to provide timely service and utilize capacity. Three methods are commonly used: appointment, reservation and backlog.

Appointment: an appointment system assigns specific times for services to customers. The advantages of this method are timely customer service and high utilization of servers. Doctors, dentists, lawyers, and automobile repair shops are examples of service providers that use appointment systems. Doctors can use the system to schedule parts of their days to visits hospital patients and lawyers can set aside time to prepare cases. If timely service is to be provided, however, care must be taken to tailor the length of appointment to individual customer needs rather than merely scheduling customers at equal time interval.

Reservation: reservation system, although quite similar to appointment systems, are used when the customer actually occupies or uses facilities associated with the service. For example, customer reserve hotel rooms, automobiles, airline seats, and concert seats. The major advantage of reservation system is the lead time they give service managers to plan the efficient use of facilities. Often, reservation requires require some form of down payment to reduce the problem of no-shows.

Backlog: a less precise way to schedule customers is to allow backlogs to develop; that is customers never know exactly when service will commence. They present service request to an order taker, who adds it to the waiting line of orders already in the system. TV repair shops, restaurants, banks, grocery stores, and barber shops are examples of the many types of business that uses this system. Various priority rules can be used to determine which order to process next. The usual rule is first come, first served, but if the order involves rework on a previous order, it may get a higher priority.

4.5.2. Scheduling the worker:

Another way to manage capacity with a schedule system is to specify the on duty and off duty periods for each employee over a certain time period as in assigning postal clerks, nurses, pilots, attendants, or police officers to specific work days and shifts. This approach is used when customers demand quick response and total demand can be forecasted with reasonable accuracy. In this instant capacity is adjusted to meet the expected loads on the service system. The work force capacity available each day must meet or exceed daily work force requirements. If it does not, the scheduler must try to rearrange days off until the requirements are met. If no such scheduling can be found, management might have to change the staffing plan and authorize more employees, over time hours or large backlogs. Managers usually use rotating schedule than fixed schedule to assign workers on their duty. In rotating schedule rotate employees through a series of work days or hours. Thus, over a period of time, each person has the same opportunity to have weekends and holidays off and to work days, as well as evening and nights. A rotating schedule gives each employee the next employee's schedule the following week. In contrast, a fixed schedule calls for each employee to work the same days and hours each week.