*Proceedings of the 2011 International Conference on Industrial Engineering and Operations Management*
*Kuala Lumpur, Malaysia, January 22 – 24, 2011*

# Production Planning under Uncertainty with Multiple Customer Classes

**Tarik Aouam**
**Department of Industrial Engineering**
**AlHosn University, Abu Dhabi, United Arab Emirates, P.O. Box 38772**

## Abstract
We consider a capacitated production system with a single stage, single item, and multiple demand classes. Demand from each customer class is assumed random without knowledge of the underlying probability distribution. The producer has the flexibility to decide on the quantity of demand to be satisfied from each customer class (order selection). We also assume that congestion takes effect after a certain critical utilization level resulting in increasing lead times. In this paper we propose an original formulation of the problem as a robust optimization model based on the concept of clearing functions (CFs). An illustrative numerical example is presented to illustrate the model characteristics and the effects of congestion and variability.

## Keywords
Production Planning, Congestion, Load-Dependent Lead Times, Robust Optimization, Multiple Customer Classes

## 1. Introduction
Models of production and inventory systems have been developed since the early days of the Operations Research and Management Science field. A major concern in the area has been to formulate models that can be solved efficiently, yet these models should not be based on over simplifying assumptions.
In particular, a few models in the literature capture the dependency between demand uncertainty, utilization, and lead-times, [8, 10]. In fact, in order to plan production of an item in the face of time-varying demands, a planned lead-time needs to be estimated. Based on this planned lead time, planning decisions are made in order to determine the amount of work to be released into the facility in a given time period, which in turn determines the production stage utilization. However, queuing models have revealed that lead-time increases non-linearly as the resource utilization approaches 100%, [6]. This constitutes the circular, non-linear dependency between lead-time and utilization. Demand uncertainty is an even more complicating factor for two reasons. Demand variability affects production lead-times, and customer requirements introduce the need to hold additional inventory, requiring additional releases that further increase utilization. In this paper we deal with the circularity problem between lead time, utilization, and demand uncertainty.
Asmundsson et al. in [1] use the idea of clearing functions (CFs) ([5, 7]), which defines the expected throughput of a capacity-constrained resource as a function of the average work-in-process inventory (WIP) at the resource. They propose mathematical programming models of capacitated production systems that capture the nonlinear relationship between workload and lead-times. Orcun et al. in [9] considers the same problem as well as demand uncertainty and the need for safety stocks. They proposed a non-linear mixed integer programming formulation of the problem and introduce an iterative heuristic as a solution approach. The proposed models in this paper use the concept of CFs but take a robust optimization approach to reflect demand uncertainty, [4], resulting in tractable formulations.
Most of the work on systems with multiple demand classes lies within the inventory systems literature. In his seminal paper, [12], Topkis considers an inventory system with a single item and multiple demand classes, in a discrete time setting. He shows that the optimal ordering policy is a base-stock policy. Benjaafar et al., [2], consider an assembly system with multiple stages, multiple items, and multiple customer classes. They formulate the problem as a Markov decision process and use it to characterize the structure of the optimal policy. Inventory models of production-inventory systems use very simple models of capacity, [8, 9, 13].
The rest of this paper is organized as follows. In section 2, we present a tractable deterministic production planning model capturing the dependency between workload and lead times. In section 3, we consider aggregate demand

uncertainty and propose a robust optimization model. In section 4, we formulate the production planning problem with multiple classes and its robust counterpart. In section 5, we present our initial numerical results. We conclude in section 6.

## 2. The WIP-Based Model:

Consider a single capacitated production stage. Raw materials are released into the stage at the beginning of each time period $\in$ . The units remain in work-in-process (WIP) for *L* periods on average, when the utilization of the stage is lower than *a critical utilization level,* , after which, congestion effects increase the lead-time. Once a finished item is produced it is kept in stock. The unit costs of releases (raw materials), production, WIP holding, and finished goods holding are given by , , , and $h$ respectively. is the unit shortage penalty. The demand for period *t* is denoted by and the cumulative demand up to time *t* by $(1, )$ assumed to be deterministic in the present section. We will use a similar notation for cumulative quantities all throughout the paper. We define as the quantity released, as the quantity produced (throughput), as the WIP at end of period *t*, as the inventory level at end of period t, and as the shortage. The maximum throughput of the production stage (capacity) over one period is denoted by . The utilization at time t is given by $=$ —. We use the concept of clearing functions in order to model the nonlinear dependence between the lead- time and utilization. A clearing function (CF) is typically a concave increasing function, which yields the expected throughput of a production stage as a function of a suitable measure of average WIP, Missbauer and Uzsoy [10]. The throughput of the production stage is given by

$$= ( ), \tag{1}$$

where $=$ ——— and $(.)$ is increasing and concave with $(0) = 0$. It is also assumed that under low utilization, i.e., $\leq$ ( ), we have $( ) > -$ . For tractability, $(.)$ is approximated by the convex hull of a set of affine functions of the form, [1]:

$$( ) = \quad \dots \quad + \ , \tag{2}$$

with is a strictly decreasing series and $= 0$. CFs model the congestion effect, indeed, when utilization approaches 100% ( $\geq$ ), the average output rate will decrease because the CF is concave, resulting in increasing lead-times. In order to model the average lead time L under low utilization we use the CF form proposed in [5], based on Little's law. The throughput of the production stage can be modeled as follows,

$$= - \ . \tag{3}$$

Given that the same proportion ,− for example 25%, of the WIP is always produced (cleared from the stage), the last unit to enter the stage and hence added to the WIP should wait *L = 4* periods in the stage before it is produced. Given initial inventories and and assuming $= 0$, the WIP-based production planning model with congestion effects is given by **(CPP):**

$$\min_{\in} [ \ + \ + h \ + \ + \ ] \tag{5}$$

$$\text{s.t.} \quad = \ - \ (1, \ ) + \ (1, \ ) \qquad \forall \tag{6}$$

$$= \ + \ (1, \ ) - \ (1, \ ) + \qquad \forall \tag{7}$$

$$\leq \qquad \forall \tag{8}$$

$$\leq \ + \qquad \forall \ , \tag{9}$$

$$= 0.5 ( \ + \ ) \qquad \forall \tag{10}$$

$$, \ , \ , \ , \ , \ \geq 0 \qquad \forall \tag{11}$$

The objective function in equation (6) minimizes total cost over the planning horizon. Constraints (6) and (7) define inventory balances for each period for the WIP and for finished goods. Constraints (8) will be binding in periods under low utilization. Constraints (9) model the congestion effect through the clearing function. Constraints (10) define the average WIP over a period.

## 3. The Aggregate RO Model

In this section, we assume that demands ’s are random without knowledge of the underlying distribution,

$$= \ + \ , \tag{12}$$

where $\in [-1,1]$, and , are the mean and standard deviation of . Also, *k*>0 is the *variability factor*. Let us explore the worst case scenarios corresponding to maximum demands and minimum demands. In terms of holding cost at time *t*, the worst case corresponds to $= -1, \forall \ \leq$ . On the other hand, in terms of shortage cost the worst case is $= +1, \forall \leq$ . These two scenarios are very unlikely to happen in a real situation. In the robust optimization approach (RO), large deviations are eliminated by allocating uncertainty budgets as follows, [3]:

$$\sum \; | \; | \leq \tag{13}$$

The inventory constraint (8) can be written as, in the case of holding:

$$= \; + \; (1, \;) - \; (1, \;) + \quad {}^{*} \quad + \quad \forall \tag{14}$$

with $z_\tau^*$ is the optimal solution of the following problem,

$$\max \; \sum \tag{15}$$

$$\text{s.t.} \quad \sum \quad \leq \qquad (\;) \tag{16}$$

$$0 \leq z_\tau \leq 1 \quad \forall \tau \leq t \qquad (\;) \tag{17}$$

The quantity $\sum \; {}^{*}$ is the maximum deviation that is admissible, i.e. within the budget limit forced by constraint (16). To represent the worst case shortage, the inventory constraint is written as,

$$= \; + \; (1, \;) - \; (1, \;) - \quad {}^{*} \quad + \quad \forall \tag{18}$$

Now, because the linear programming problem (22 - 24) is feasible and bounded, by strong duality the optimal objective (total deviation) of this problem is equal to the optimal objective of its dual. By replacing $\sum \; {}^{*}$ in equations (21) and (22) by the dual objective function and by defining $= \max\{h \;, \;\}$, we obtain the following robust counterpart of the CPP problem, which represent the robust production planning model and is given by **(RO-CPP):**

$$\min \sum_{\in} [\; + \; + \; + \;] \tag{19}$$

$$. . \quad = \; - \; (1, \;) + \; (1, \;) \qquad \forall \tag{20}$$

$$\geq h \,(\; + \; (1, \;) - \; (1, \;) + \; + \sum \;) \quad \forall \tag{21}$$

$$\geq - \; (\; + \; (1, \;) - \; (1, \;) - \; - \sum \;) \quad \forall \tag{22}$$

$$+ \; \geq \qquad \forall \;, \tag{23}$$

$$(8) - (11)$$

$$, \;, \; \geq 0 \qquad \forall \tag{24}$$

Bertsimas and Thiele in [3] suggest that evolves as $\sqrt{\; + 1}$. The uncertainty budgets are increasing in time since uncertainty increases with the number of future time periods considered. Also, the uncertainty budgets cannot increase by more than 1 at each time period, i.e., $0 \leq \; - \; \leq 1, \forall$ . This means that the increase should not exceed the number of new parameters added at each time period. Based on these assumptions, we suggest the use of $= \gamma\sqrt{\; + 1}$, where $0 \leq \gamma \leq \frac{}{\sqrt{}}$ in order for the assumptions to hold at each time period. γ is referred to as the *budget factor*.

## 4. The RO Model with Multiple Customer Classes

In this section we assume that there are several customer classes and the producer has the flexibility to decide at the beginning of the planning horizon on the quantity of demand to be satisfied from each class.

### 4.1 The Nominal Model

Let us assume that there are *N* customer classes and that the demand of a customer class *n* in period *t* equals to . The producer has the flexibility to decide on the quantity to be committed to. $\geq 0$ denotes the quantity of demand that the producer has already committed to in the past and has the obligation to meet; this quantity is assumed to be always less than the demand. The reservation price, i.e., the price a class is willing to pay for the product is denoted by . The nominal production planning problem with congestion and multiple demand classes can be formulated as follows, (**CPPMC):**

$$\min_{\in} \; + \; + h \; + \; + \; - \sum_{\in} \tag{25}$$

$$\text{s.t.} \quad = \; - \; (1, \;) + \; (1, \;) \qquad \forall \tag{26}$$

$$= \; + \; (1, \;) - \sum_{\in} \; (1, \;) \qquad \forall \tag{27}$$

$$\leq \qquad \forall \;, \tag{28}$$

$$\leq \qquad \forall \;, \tag{29}$$

$(8) - (11)$

The objective function (25) is equivalent to maximizing net profit. Constraints (28) and (29) are restrictions on the quantities to produce.

### 4.2 The RO Model

Now assume that the demand for each customer class is random and is represented by

$$= \quad + \quad , \tag{34}$$

with $\in [-1,1]$, and $\quad , \quad$ are the mean and standard deviation of . Only constraints (29) in problem CPPMC are affected by demand uncertainty. A robust, yet very conservative approach is to consider the worst case scenario (in terms of constraint violation) corresponding to $= -1$, i.e.,

$$\leq \quad - \qquad \forall \ , \tag{35}$$

The aim is to formulate a robust model that can provide solutions that reflect the producer's level of conservatism. Notice that constraints (29) imply that (*but are not equivalent to*),

$$(1, \ ) - \quad (1, \ ) + \sum \qquad \leq \ 0 \qquad \forall \ , \tag{35}$$

For each constraint corresponding to a class *n* and a period *t*, a budget of uncertainty parameter, $\Gamma \ \in [0, \ ]$, is introduced. Through this parameter the producer can adjust the robustness of the solution against the level of conservatism, which may vary from one class to another. The worst case scenario in terms of the chances to violate the constraint corresponds to $= +1, \forall \ \leq$ . However, it is unlikely in practice that during all the periods $\leq$ , demands will be equal to their maximum. In fact this situation would be extremely conservative and would correspond to Soyster's model, [11]. The approach of Bertsimas and Sim, [4], is followed in this work. It is assumed that up to $\lfloor \Gamma \ \rfloor$ periods demand will be at their maximum and during one of the periods demand will change by $(\Gamma \ - \lfloor \Gamma \ \rfloor)$ . Constraints (35) can now be written as,

$$(1, \ ) - \quad (1, \ ) + \sum \quad ^{*} \qquad \leq \ 0 \qquad \forall \ , \tag{36}$$

with $z^{*}_{\tau}$ is the optimal solution of the following problem,

$$\max \ \sum \tag{37}$$

$$\text{s.t.} \ \ \sum \quad \leq \qquad ( \quad ) \tag{38}$$

$$0 \leq \quad \leq 1 \quad \forall \tau \leq t \qquad ( \quad ) \tag{39}$$

The quantity $\sum \quad ^{*}$ is the maximum deviation that is admissible and represents the protection from violation of the constraint. The linear program (37 - 39) is feasible and bounded we conclude by strong duality that the optimal objective of this problem is equal to the optimal objective of its dual $^{*} \quad + \sum \quad ^{*}$ . Now that a protection is provided against demand uncertainty, demands can be replaced by their mean in constraints (29). The robust production planning problem with congestion and multiple demand classes can be formulated as follows,

**(RO-CPPMC):**

$$\min_{\in} \qquad + \qquad + h \quad + \qquad -_{\in} \tag{40}$$

$$\text{s.t.} \ (26), (27), (29)$$

$$\leq \qquad \forall \ , \tag{41}$$

$$(1, \ ) - \quad (1, \ ) + \qquad + \sum \qquad \leq \ 0 \qquad \forall \ , \tag{42}$$

$$+ \qquad \geq \qquad \forall \ , \ , \ \leq \tag{43}$$

$$(8) - (11)$$

$$, \qquad \geq 0 \qquad \forall \ , \ , \tag{44}$$

## 5. Numerical Results

This section presents an illustrative example. The optimization models have been implemented using GAMS and solved using CPLEX 11.0. The customer demand is forecasted over a horizon of three months each consisting of four working weeks (12 weeks horizon). In this example we assume that demand during months 1, 2, and 3 are equal to 80, 100, and 200 respectively. Each demand period has a standard deviation of 20% of the mean. Notice that product demands are heavily skewed towards the later periods, which is the case in practice since it is observed that often up to 60% of customer orders are shipped in the last 3 weeks of a quarter. Because of limited production capacity, production takes place ahead of time during early periods. The choice of values for and can be arbitrary, but the values we use are those recommended by Graves (1988), $= \; . \; \bar{} = 254$ and $= 1.645 * 0.2 * 127 * \sqrt{2} \cong 59$. We assume that the capacity is $= 200$ (almost 160% of the average demand). We assume that $= \$20$, $= \$100$, $h = 0.2$ , $= 0.1$ , $= 0.6$ . Based on [7], we assume the following functional form of the clearing function, $( \; ) = \frac{}{}$, where $\mathrm{K} = \; . \; \; .(1 - \; )$, =0.8, and $L$ = 2 weeks. We consider four customer classes distinguished by their reservation price (high/low) and their budget of uncertainty (high/low). Specifically, $= \; = \$200$, $= \; = \$180$, $= \; = 1.2\ \gamma\sqrt{\; + 1}$, and $= \; = 0.8\ \gamma\sqrt{\; + 1}$. where $0 \leq \gamma \leq \frac{}{\sqrt{}}$ is the budget factor. We also assume that $= 1.2 * -$ .

### 5.1 Robust Model

By solving the RO-CPP with and without the congestion constraints one can compare the effect of congestion on the cost and the various planning decisions. In fact, the congestion leads to an increase in the total cost of 5%. Furthermore, Figure 1 shows the effect of congestion on the various planning decisions. As one can see, the congestion increases the WIP level in the productions stage, which limits the releases and the production output. As a consequence, we see that the level of finished goods inventory is lower.

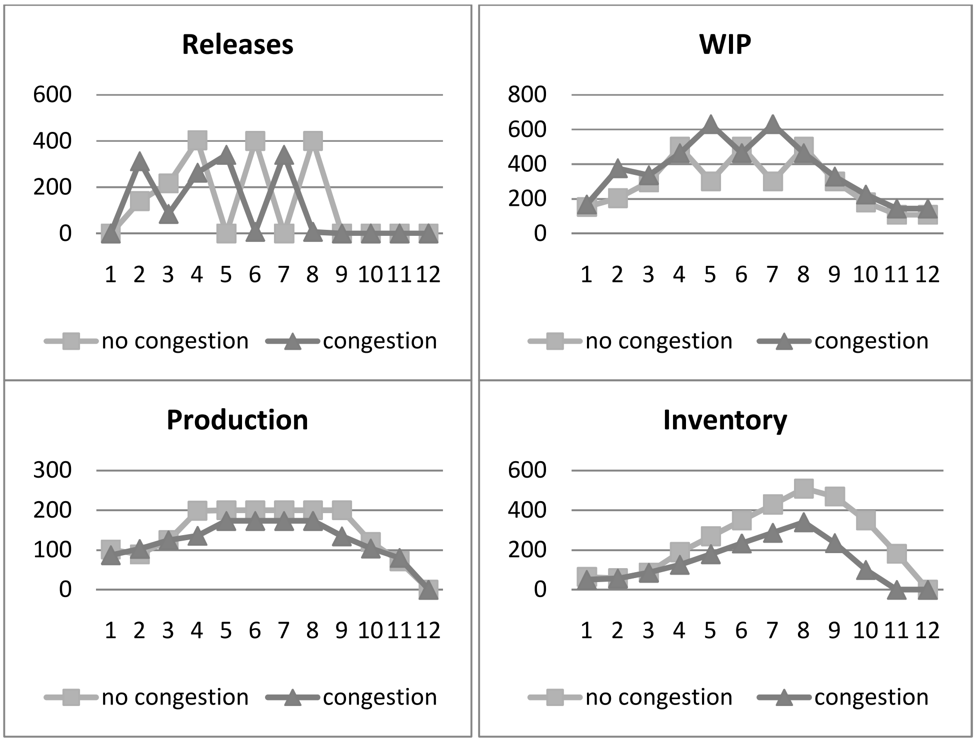


Figure 1: The effect of Congestion on Production Decisions

### 5.2 Robust Model with Customer Classes

We would like to demonstrate the effect of the budget of uncertainty, , and the variability factor, , on the fulfillment rate defined as $\frac{\Sigma}{\Sigma}$ for each customer class *n*. Three levels of each parameter are considered and the respective fulfillment rates are reported in Table 1. These results were obtained from solving model RO-CPPMC with and without congestion constraints. The results lead to the following conclusions:

- an increase in the variability factor or in the uncertainty budget decreases the fulfillment rate,
- congestion decreases the fulfillment rate,
- an increase in the reservation price increases the fulfillment rate; compare n1 to n3 or n4 to n2.

- an increase in the level of budget of uncertainty (producer more conservative) decreases the fulfillment rate; compare n1 to n4 or n2 to n3.

Table 1. Fulfillment Rate for Customer Classes

| | | **No Congestion** | | | | **Congestion** | | | |
|---|---|---|---|---|---|---|---|---|---|
| **K** | | n1 | n2 | n3 | n4 | n1 | n2 | n3 | n4 |
| 1 | 0 | 1 | 1 | 0.7 | 1 | 1 | 0.72 | 0.54 | 0.96 |
| 1 | 1.2 | 0.88 | 0.84 | 0.88 | 0.91 | 0.88 | 0.62 | 0.67 | 0.91 |
| 1 | 2.4 | 0.82 | 0.86 | 0.79 | 0.86 | 0.82 | 0.63 | 0.6 | 0.86 |
| 2 | 0 | 1 | 1 | 0.7 | 1 | 1 | 0.72 | 0.51 | 1 |
| 2 | 1.2 | 0.76 | 0.82 | 0.76 | 0.82 | 0.76 | 0.59 | 0.61 | 0.82 |
| 2 | 2.4 | 0.63 | 0.71 | 0.63 | 0.71 | 0.63 | 0.56 | 0.44 | 0.71 |
| 3 | 0 | 1 | 1 | 0.7 | 1 | 1 | 0.72 | 0.54 | 0.96 |
| 3 | 1.2 | 0.64 | 0.73 | 0.64 | 0.73 | 0.64 | 0.53 | 0.55 | 0.73 |
| 3 | 2.4 | 0.45 | 0.57 | 0.45 | 0.57 | 0.45 | 0.49 | 0.32 | 0.57 |

## 6. Conclusion

This work considers the single item production planning problem for a single capacitated production stage with demand uncertainty and multiple demand classes. Robust WIP-based production planning models are formulated capturing the circularity problem between demand uncertainty, resource utilization, and lead-time. An illustrative numerical example shows how congestion can affect greatly production decisions. It is also established that the fulfillment rate is greatly affected by variability. This work can be extended in two different ways. This work should be completed by a thorough simulation study to evaluate the actual performance of the RO approach. Also, there are several potential avenues for future research. One can of course consider the extension to multiple products. A more challenging extension would be the introduction of setup costs.

## References


1. Asmundsson, J. M., Rardin, R. L., Turkseven, C. H., and Uzsoy, R., 2009, “Production planning models with resources subject to congestion.” Naval Research Logistics, 56, 142-157.
2. Benjaafar, S., ElHafsi, M., C.-Y. Lee, W. Zhou, 2010, “Optimal Control of an Assembly System with Multiple Stages and Multiple Demand Classes”. Operations Research (in press).
3. Bertsimas, D. and Thiele, A., 2006, “A robust optimization approach to inventory theory.” Operations Research, 54, 150-158.
4. Bertsimas, D., M. Sim. 2004. The price of robustness. Operations Research, 52, 35–53.
5. Graves, S. C., 1985, “A tactical planning model for a job shop.” Operations Research, 34, 522-583.
6. Hopp, W. J. and Spearman, M. L., 2001, Factory physics: Foundations of manufacturing management. Irwin/McGraw-Hill, Boston.
7. Karmarkar, U. S., 1989, “Capacity loading and release planning with work-in-progress and lead-times”. Journal of Manufacturing and Operations Management, 2, 105-123.
8. Missbauer, H. and Uzsoy, R., “Optimization models for production planning.” To appear in Handbook of Production Planning, K. G. Kempf, P. Keskinocak, R. Uzsoy, Kluwer International Series in Operations Research and Management Science, Kluwer Academic Publishers (in press).
9. Orcun, S., Uzsoy, R., and Kempf, K. G., 2009, “An integrated production planning model with load-dependent lead-times and safety stocks.” Computers and Chemical Engineering, 33, 2159-2163.
10. Pahl, J., S. Voß, D. L. Woodruff, 2007, “Production planning with load dependent lead times: an update of research”, Annals of Operations Research, 153, 297–345
11. Soyster, A. L. 1973. “Convex programming with set-inclusive constraints and applications to inexact linear programming.” Operations Research, 21, 1154–1157.
12. Topkis, D., 1968, “Optimal Ordering and Rationing Policies in a Nonstationary Dynamic Inventory Model with n Demand Classes,” Management Science, 15, 160-176.
13. Zipkin, P. H., 1977, Foundations of inventory management. Irwin, Burr Ridge, IL.