PRICING DELIVERY PRIORITY AND SPEC. LEVEL OF SEMICONDUCTOR PRODUCTS

by

Kyung Jo Min and Shmuel S. Oren

Memorandum No. UCB/ERL/IGCT M90/64

26 July 1990

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Pricing Delivery Priority and Spec. Level of Semiconductor Products

Kyung Jo Min

Iowa State University

and

Shmuel S. Oren

University of California, Berkeley

ABSTRACT

We construct and analyze a new allocation mechanism for semiconductor products of a single producer via pricing of both delivery priorities and spec. levels. The investigation is aimed at improving the performance of semiconductor producers with respect to on-schedule delivery of chips. In the proposed scheme the producer offers a " product line " of priority classes under an allocation rule that supplies higher priority classes with higher spec. level chips regardless of production outcome. This product line design and allocation rule enable us to cast the producer's profit maximization problem as a non-linear programming formulation. We examine the effects of perturbation in the allocation rule on profits and the resulting insights lead to a set of conditions which can be used to determine a priori the profitability of downgrading (re-labelling and selling chips of a given spec. level as chips of lower spec. level). We also obtain a set of general conditions from which one can infer the optimality of the allocation rule employed.

1. Introduction

The semiconductor industry is generally characterized as capacity constrained. Even in recession periods, due to the diversity in their usage, there are certain products whose demand exceeds the production capacity. And in expansion periods, the production capacity rarely catches up with the potential market for the products (Leachman (1986)). In addition, due to unique manufacturing characteristics of semiconductor devices, the control and prediction of outcomes for production planning are complex and often harder than in other industries. Under such circumstances, on - schedule delivery of chips of given spec. (specification or minimum quality) ordered by customers is inherently difficult, and over-booking by the suppliers or cancellation of orders by the customers are common practices. In case of over-booking or shortage, the scarce semiconductor products are not necessarily allocated to the most valued use by the suppliers and attempts by the customers to anticipate and avert such situations (e.g. by over-ordering) make the transactions all the more chaotic. The main objective of this paper is to construct an allocation mechanism via pricing so as to alleviate the problems described above taking both the manufacturing characteristics of semiconductor products and the market conditions for the supplier and customers into consideration.

1.1 Overview of Semiconductor Products

We first present a brief description of the semiconductor manufacturing process and its characteristics (For a deeper understanding of this aspect, see Leachman (1986) or Intel (1985)).

Wafers of silicon materials are imprinted with numerous patterns of an integrated circuit where each integrated circuit is called a "die". Then each die is probed for its capabilities and defects are marked for later disposal. Next the wafers are sawed and dies are assembled into packaged devices or chips. Finally packaged devices are extensively tested to determine the level of certain critical attribute(s). For example, Figure 1.1 depicts the distribution of dies on a wafer over two critical attributes such as speed and power consumption. According to the test results, the dies are classified into bins where each bin has a pre-determined specific range of the attribute level, e.g., a bin may be for the chips with speed level between 10 MHz and 15 MHz, and any chip in this bin is said to have the spec. level of 10 MHz. Figure 1.2 illustrates this classification process. The actual amount of chips falling into each bin (or the actual distribution of the chips over the attribute levels) varies from one production run to the next due to the random factors in the production process, which may result in shortage of certain spec. levels. (Observe that in Table 1.1, Lot #9's ratio of bin 1 and bin2 is nearly 1 : 1 whereas Lot #11's ratio is nearly 10 : 1.)

Prod #	Total	Bin 1	Bin 2	Bin 3	Bin 0
9	1458	530	453	204	271
11	1356	883	86	212	175
15	1456	646	144	327	339
18	1450	620	148	528	154

(Prod # denotes the production run number and Bin 0 denotes the number of defects.)

Table 1.1 Classified Chips and Their Distribution

Any chip with higher spec. level can be re-labelled and sold as a lower spec. chip and this industry practice is called down-binning or downgrading. Downgrading has been employed as an emergency measure to fulfill delivery commitments to customers when there is a shortage of lower spec. chips. To quote from Leachman (1986), " Scrapping of output falling into low-quality bins is common. Also, it is common to assign binned output to a customer spec. whose sales price is less than the highest priced customer spec. for which that bin is suitable. ··· a packaged device final tested to one customer spec. could be assigned to fill an order for an inferior customer spec. " In this paper, we extend the usage of downgrading such that it is applicable whether there is a shortage of lower spec. chips or not. It might be counter-intuitive that lowering quality levels, thus lowering the total economic surplus from which the producer's profit can be extracted, may actually increase the profit level. However, from a monopolist's viewpoint, downgrading is an added instrument for discrimination among potential customers and as such it may enhance profits at the expense of a reduction in total surplus.

Another important economic aspect in the above manufacturing process is the determination of the spec. levels that define the bins. In general by setting spec. levels higher (lower), the quality of chips to be sold increases (decreases) and the price the producer charges also increases (decreases) while the quantity to be sold decreases (increases). Unfortunately, often times this important decision is carried out without systematic analysis of economic consequences. To quote from Riley and Sangiovanni-Vincentelli (1986), " ··· Among the deficiencies of conventional formulations implicit in this characterization is that they would presumably involve considerable effort and computer time to precisely adjust parameters based on a <u>definition of what a "good" set of specification is , which has, to a degree, been pulled out of the proverbial air</u>". In this paper, the spec. levels are optimally determined from the producer's profit maximization problem following the economic approach shown in Riley and Sangiovanni-Vincentelli (1986) and Styblinski (1985).

The producer provides the price schedule for chips, often charging different prices for different spec. levels (see e.g., Table 1.2) while customers make order decisions taking this price schedule into considerations.

Price Schedule for Priority Interrupt Controllers

Prod.I.D.	Package Type	Spec. (Speed)	Price
9519A1DC	CER DIP	3.0 MHZ	$24.95
9519A1PC	MOLD DIP	3.0 MHZ	$15.85
9519ADC	CER DIP	2.0 MHZ	$19.20
9519APC	MOLD DIP	2.0 MHZ	$13.50

Table 1.2 Producer's Price - Spec. Level Schedule

Due to inherent production capacity limitations and randomness in the production mix, delivery commitments are frequently not met. To quote

Leachman (1986), " ··· On the whole, industry performance in this respect [delivery quality] historically has been poor compared to other industries". The consequence of that poor performance is chaotic over-ordering and over-booking, which results in allocative inefficiencies. In other words, with conventional price schedules the customers' delivery service quality is not in accord with the ranking of the customers' preferences or willingness to pay. Therefore, price schedules and the corresponding allocation schemes that better account for the delivery service quality is called for.

Even though the overview explained above may be applicable to various semiconductor products, in this paper, we limit our investigation to monopolistic semiconductor products with a single critical attribute whose end use is relatively " homogeneous " (e.g., microprocessors used in personal computers). For an extension of this work to monopolistic semiconductor products with relatively " heterogeneous " end use (e.g., certain memory chips used both in personal computers and in consumer electronics), see Min and Oren (1990).

1.2 Pricing Theory

Earlier work on non-uniform pricing was done by Mirrlees (1971) on taxation problems. In Goldman, Leland and Sibley (1984), and Oren, Smith and Wilson (1983), products are defined over the purchase quantities where as in Mussa and Rosen (1978), and Oren, Smith and Wilson (1982,1984), products are defined over quality levels. For economic analysis of products defined by bundling and multi product monopolies, see Adams and Yellen (1976) and Mirman and Sibley (1980) respectively. Examples of unbundling quality attributes in the electric power context are described in Chao et al. (1986a) and in Oren and Min (1988). Earlier work on asymmetry of information and subsequent incentive compatibility problem was done by Leland and Meyer (1978). Market segmentation and "cannibalization" among segments were studied by Moorthy (1984) while allocative distortion due to market segmentation subject to incentive compatibility was studied by Cooper (1984).

Several recent papers have focused on rationing of limited supply through priority pricing. According to this approach, available supply is allocated on the basis of contracts that specify each customer's priority. Priority pricing has been shown to achieve efficiency gains in the case of non-storable commodities or goods and services with congested demands or queues due to limited supply capacities. In the seminal work by Harris and Raviv (1981), they show that priority pricing is superior as far as monopoly profits are concerned to other allocative schemes. In Chao and Wilson (1987), in the context of electric power, it is shown that priority rationing is Pareto superior to random rationing. For a further systematic treatment of supply rationing via pricing and an extensive reference list, see Wilson (1989).

As an alternative to the priority pricing, one might consider spot pricing. In this pricing scheme, the prices for the chips continually change under each realization of the production outcome so as to induce an efficient allocation of chips (See e.g., Vickrey (1971) for an early treatment on spot pricing and efficiency gain). Such pricing, however, is very difficult to implement in our environment. For example, customers' ordering decision is tied to complex planning and scheduling of their end products and the cost to adapt their planning and scheduling according to continual changes in prices of available chips would be substantial.

1.3 Priority and Spec. Level Pricing

In this paper, we develop an allocation mechanism for semiconductor chips via pricing where the " product line " consists of priority levels. A product class specifies a vector of spec. levels and a forecast of corresponding delivery probabilities resulting from applying the priority supply rule. We may consider these product classes as bundled products of an intrinsic attribute (spec. level) and delivery quality. Specifically, at the end of each production period, the producer distributes the quantities classified into bins with different spec. levels according to the priority supply rule which specifies the supply order for the subscribed customers. Customers' purchase decisions are based on the producer's forecast of the probability that each spec. level will be delivered in filling the order under a given priority level. These probabilities are endogenously derived along with the corresponding prices through the producer's profit maximization subject to constraints implied by production and demand data. Figure 1.3 depicts a typical allocation procedure given a production outcome as well as a typical price table.

Conventional pricing practice does not explicitly price the delivery quality. Incorporating this attribute in the price schedule and the economic determination of spec. levels are two important features of this model. Such an approach can lead to better economic efficiency in the allocation of the supply quantity among the competing orders. (i.e. higher quality products are offered to higher valued consumption units at higher prices and the magnitude of mis-allocation due to haphazard delivery policies is greatly reduced.)

In section 2, we mathematically characterize the market. Next, we construct the basic model and provide the underlying motivation. In section 3, we investigate the downgrading option and the optimality of the allocation rule we employ. We introduce, in section 4, a two part pricing for practical implementation purpose and present an illustrative example. Finally, in section 5, we make concluding remarks and comment on further research.

2. Model Constructs

2.1 Market Characterization

The market heterogeneity is characterized in terms of a customer's type index $t \in [0, 1]$, where $t = 0$ defines the lowest ranked customer type and $t = 1$ the highest. Each customer t has his utility function $U(u, t)$ per purchase unit where $u \in [0, 1]$ denotes a normalized index characterizing the spec. level of the critical attribute under consideration such as speed or power consumption. The assumption that customer's utility depends on the spec. level rather than on the true quality of the product is reasonable when customer testing for the precise quality level and utilizing subsequent results of the tests is uneconomical. In such cases, a customer must rely on the spec. level which is the (minimum) guaranteed quality stamped on the chip. We also assume that the net benefit to the customers from the detection of the downgraded chips is negligible. In other words, customer testing for the " true " spec. level and subsequent utilization of the test results is uneconomical. We also assume that income effects (in the microeconomic sense) are negligible. Throughout this paper, the term utility function is identical to " Willingness To Pay (WTP) " function in the pricing literature (see e.g., Oren, Smith and Wilson (1983)).

We assume $U(0, t) = 0$, i.e. spec. level 0 is worthless to any customer $t \in [0, 1]$. We also assume that $U(u, t)$ is twice differentiable and that $\partial U/\partial u > 0$ and $\partial U/\partial t > 0$, i.e. the value of any chip increases with spec. level and the value of any spec. level chip is higher when used for a higher ranked end use. We also assume $\partial^2 U/\partial u \partial t > 0$, which is referred to as utility monotonicity condition. (see e.g., Oren, Smith and Wilson (1983)). It implies that there is unanimous agreement among customers regarding the preference ranking of the products. Or technically, the utility functions of customers for any pair of products never cross. Such monotonicity ensures tractable market segmentation, i.e. self-selection will result in ordered blocks or groups of customers that select the same products (See e.g., Oren, Smith and Wilson (1982), Moorthy (1984) and Smith (1986)).

Cumulative demand is characterized by $D(t)$, the total potential demand quantity by customers ranked t or higher. Thus if the market segmentation is contiguous, the size of market segment i (or customer class i) is defined to be $D(t_i) - D(t_{i-1})$ and t_i, $i = 1,\cdots, M$ will be referred to as boundary customers.

The contingency supply quantity function is characterized by $S(u, \underset{\sim}{B})$, the total amount of chips produced in a given period with spec. level u or higher under contingency $\underset{\sim}{B}$. The vector $\underset{\sim}{B} = (B_1,\cdots, B_K)$ denotes a random vector whose elements correspond to factors that affect the quality distribution of production outcomes. We assume that the sample space β for $\underset{\sim}{B}$ and the corresponding joint probability distribution $\text{Prob}\{ \underset{\sim}{b} \}$ over all possible realizations $\underset{\sim}{b} \epsilon \beta$ are known to the producer. Also the producer is assumed to have complete knowledge of the customers' type distribution D(t) and the form of the utility function U(u,t), but he can not identify the particular type of a customer. We assume that the production cost is lump sum constant because, regardless of number of defects or the actual distribution of chips over the quality level, the manufacturing and testing costs have already been incurred. For notational convenience, we suppress the arguments in the utility and demand functions whenever there is no ambiguity, i.e., we denote $D(t_i)$ by D_i and $U(u_j, t_i)$ by U_{ji}. We will also omit the vector notation for the contingencies and replace $\underset{\sim}{B}$ by B and $\underset{\sim}{b}$ by b

2.2 Basic Model

In order to understand the motivation for the proposed pricing scheme and the priority supply rule, let us consider an elementary format for pricing spec. levels and delivery probabilities from which the basic model is evolved. In this format, the priority of delivery is specified with respect to each spec. level. That is, within each spec. level, a higher priority customer is guaranteed delivery before a lower priority customer.

Let us denote the spec. levels and delivery classes by u_j's and d_i's respectively. u_j's and d_i's are arranged such that lower index implies higher quality. The probability of delivery for a u_j chip at delivery class i is denoted by pr_{ij} and the corresponding price by p_{ij}. Thus in this format, a product consists of a spec. level and the corresponding delivery priority and is denoted by the pair $< u_., d_. >$. Table 2.1 below depicts a typical price table of this format.

	u_1	....	u_N
d_1	$p_{11}/\ pr_{11}$	....	$p_{1N}/\ pr_{1N}$
d_2	$p_{21}/\ pr_{21}$	....	$p_{2N}/\ pr_{2N}$
$\vdots$	$\vdots \quad \vdots$	....	$\vdots$
d_M	$p_{M1}/\ pr_{M1}$	....	$p_{MN}/\ pr_{MN}$

Table 2.1 Price Table for the Elementary Priority and Spec. Level Pricing

The delivery probabilities which can be provided as a forecast by the supplier or inferred by customers on the basis of rational expectation must take into consideration the uncertain supply mix and the distribution of customer orders over the products $< u_., d_. >$'s. The forecast of delivery probabilities in each spec. level requires an explicit allocation rule for every contingency $b \epsilon \beta$ since high spec. chips can be always downgraded to fill an order for inferior chips. In specifying such an allocation rule, the producer must address questions such as who gets the second pick at u_1 chips after the first priority order of such chips is satisfied. Suppose, for example, that there is a shortage of u_2 chips. Then, should the excess u_1 chips be given to the second priority u_1 customers (i.e., customers who selected product $< u_1, d_2>$) or downgraded as u_2 chips and given to the first priority u_2 customers (i.e., customers who selected product $< u_2, d_1>$) ? Also it is conceivable that in case of shortage, a customer ordering u_1 chips will accept u_2 chips as a substitute. In such cases, again the producer has to decide who has the priority, i.e., customers who selected $< u_1, d_1 >$ or $< u_2, d_1 >$.

It is possible to resolve such dilemmas with some arbitrary allocation rules. Unfortunately, when such assignment rules are accounted for in evaluating the expected utility functions corresponding to a particular pair of products $< u_i, d_h >$ and $< u_j, d_k>$, there may be disagreement among customers as to the preference functions may cross and the producer's market segmentation is in general intractable (see e.g., Moorthy (1984)). Figure 2.1 depicts an example where the net expected utility functions cross more than once.

A lexicographic allocation rule that resolves the above dilemmas and guarantees tractable segmentation (it will be shown later) is as follows: Assign each product $< u_i, d_j >$ an ordered priority. Next, regardless of production outcome, customers who selected these products are supplied according to the priority order with chips pulled out of the bins from the top down. With such allocation rule, the ordered spec. level or the delivery priority with respect to each spec. level becomes rather meaningless and the relevant aspect now is the rank order of customers. Thus, the producer can collapse the two dimensions (spec. level and delivery priority) into one and simply define one product line as supply priority where the priority class denotes the rank in which an order is met with chips out of the bins that are depleted from the top down. We now formally state this priority supply rule as follows:

Suppose that the spec. levels have been determined and all semiconductor chips produced are classified into bins with spec. $u_1, \cdots, u_N$. The priority supply rule determines how the bin contents are assigned to customers whose orders are specified in terms of priority classes $1, \cdots, M$. According to this rule, regardless of the number of chips in each bin, the producer allocates spec. u_1 chips to the first class customers. If there is a shortage of u_1 chips, they are allocated randomly to the first class and the balance of the first class orders is met with the next best chips, spec. u_2. If spec. u_2 chips are also exhausted, then the next best spec. u_3 are allocated, and so on until the demand for first class is met. Only then does the allocation to the second class customers start in the same manner, and only after the second class,the allocation for the third class, and so on. This rationing process terminates when either all the bins are exhausted or when all the demands are satisfied.

In this way, customers who select a higher priority are provided with higher (or equal) quality mix than those who select a lower priority under every contingency $b \epsilon \beta$. One of the advantages of this allocation rule lies in the fact that it is verifiable. For instance an organization can easily verify that the quality mix in a lower class order is inferior to a higher class order under any contingency. Such easy verification mechanism is helpful for establishing the producer's credibility and inducing the desired customer response to the pricing scheme.

While the priority supply rule determines ex-post the relationship between the priority class and the quality mix, customers' ex-ante purchase decision will be based on a probabilistic forecast of that relation which will specify Pr_{ij}, the delivery probability of u_j spec. chip to a customer in priority class i averaged over all possible contingencies. This forecast must take into consideration both the rationing rule (in this case the priority supply rule) and the anticipated response by customers. Such response will obviously depend on the price corresponding to each class which is controlled by the producer. A reasonable objective in setting such prices is to induce monotone market segmentation. That is, to induce the higher ranked customer (who have a higher utility for any given spec. level) to select higher priority class. There is still the question whether such ordered segmentation is maximizing profit and we will later address this issue partially in subsections 3.3. This assumption, however, is essential for analytic tractability.

We will now proceed to express the probabilities Pr_{ij}'s corresponding to the priority supply rule under the monotone segmentation hypothesis. We introduce for this purpose, variables denoting the amount of chips available, the amount of shortage/ surplus, and the amount of downgraded chips under each contingency $b \epsilon \beta$ as follows:

V_{jb} : the total amount of chips of spec. level u_j when $B = b$.

Q_{ijb} : the remaining demand in class i after using up spec. j chips when B = b .

R_{ijb} : the remaining supply of spec. j chips after supplying class i when B = b .

W_{ijb} : the amount of originally u_i chips re-labelled as u_j chips when B = b.

The relation between W_{ijb} and the supply quantity function is as follows: for all $b \in \beta$, $j = 1, \cdots, N$

$$S(u_j, b) - S(u_{j-1}, b) = W_{jjb} + W_{jj+1} + \cdots + W_{jNb} \tag{2.1}$$

$$V_{jb} = W_{1j} + W_{2j} + \cdots + W_{jjb} \tag{2.2}$$

When we rule out the downgrading option, since every chip is sorted into the highest possible spec. bin, we can simplify the above relations as follows:

$$S(u_j, b) - S(u_{j-1}, b) = V_{jb} \quad \text{for all } b \in \beta \tag{2.3}$$

The demand for each product class is given by $D(t_i) - D(t_{i-1}) = D_i - D_{i-1}$ for $i = 1, \cdots, M$ where $D_0 = 0$. Consequently, we can express the shortage and surplus relations with respect to each class and spec. level under the priority supply recursively as follows:

$$Q_{ijb} = \max[\, Q_{ij-1b} - R_{i-1jb},\ 0 \,] \tag{2.4}$$

$$R_{ijb} = \max[\, R_{i-1jb} - (\, Q_{ij-1b} - Q_{ijb} \,),\ 0 \,] \tag{2.5}$$

$j = 1, \cdots, N$ and $i = 1, \cdots, M$ and for all $b \in \beta$
where $R_{0jb} = V_{jb}$, $Q_{i0b} = D_i - D_{i-1}$, $t_0 = 1$, and $D_0 = 0$

According to the priority supply rule, the chips of any spec. level u_j allocated to a class i are randomly distributed within this class. Hence under any contingency b, the conditional probability

$$Pr_{ij|b} = (Q_{ij-1b} - Q_{ijb})/(D_i - D_{i-1}) \tag{2.6}$$

Averaging over all possible contingencies, we have :

$$Pr_{ij} = \sum_{b \in \beta} \mathrm{Prob}\{ b \} ((Q_{ij-1b} - Q_{ijb})/(D_i - D_{i-1}))$$

$$i = 1, \cdots , M \text{ and } j = 1, \cdots , N \tag{2.7}$$

Under the proposed scheme, the producer's price schedule will consist of priority prices and corresponding forecast of delivery probabilities for the various spec. levels as shown below.

class	price	u_1	u_2	....	u_N
1	p_1	Pr_{11}	Pr_{12}	....	Pr_{1N}
2	p_2	Pr_{21}	Pr_{22}	....	Pr_{2N}
:	:	:	:	....	:
M	p_M	Pr_{M1}	Pr_{M2}	....	Pr_{MN}

(p_1 and u_1 denote the highest price and spec. level)

Table 2.2 Price Table for Priority and Spec. Level Pricing

When N = 1 and M = 1, this scheme collapses to a random rationing of a homogeneous product. When N = 1 and M > 1, we have the classical priority rationing price schedule for a homogeneous product (see Wilson (1989)). And when N > 1 and M = 1, we have a random rationing scheme that supplies heterogeneous products to a single class of customers. Finally when N > 1 and M > 1, we have the non-degenerate priority pricing of heterogeneous products under the priority supply rule.

We now turn to modeling the customer's decisions. We assume that customers are expected value decision makers. Thus, each customer t's expected utility and net expected utility when he orders a unit of priority class i are given by,

$$EU_i(t) = \sum_{j=1}^{N} Pr_{ij}\, U(u_j, t) \qquad (2.8)$$

and,
$$NEU_i(t) = \sum_{j=1}^{N} Pr_{ij}\, U(u_j, t) - p_i \qquad (2.9)$$

The optimal customers' behavior or self-selection is simply to choose priority level $\hat{i}$, where $NEU_{\hat{i}}(t) = \max_i NEU_i(t)$.

Indirect market segmentation, which is the essence of this paper, is the result of the aggregate responses by the customers optimizing their net expected utility. For this reason, throughout this chapter, we assume that $EU_i > EU_{i+1}$ for any $t \in [0, 1]$. If $EU_i = EU_{i+1}$ for a customer $t \in (0, 1]$ under the priority supply rule, then this implies that an identical spec. level is supplied to all customers of priority class i and i+1 under every contingency $b \in \beta$. In such pathological cases (i.e., when two classes are targeted to receive the identical products), the customers may not respond as intended by the producer. Also it is conceivable that there may be m classes out of M priority classes ($m < M$) that no customer selects. Such cases are referred to as market segmentation with m degenerate market segments. Degenerate market segments affect neither the producer's profit nor the customers' consumer surplus in our basic model framework. Therefore, we can consider such cases as equivalent to market segmentation with M-m non-degenerate market segments.

The monotonicity condition on utility functions is a standard assumption in the literature dealing with product line pricing (see e.g., Oren, Smith and Wilson (1982), Moorthy (1984) and Smith (1986)). In our model, mathematical representation of the monotonicity condition is as follows:
For priority class $i = 1, \cdots, M$

$$(EU_i(t_1) - EU_{i+1}(t_1)) - (EU_i(t_2) - EU_{i+1}(t_2)) > 0 \quad \text{for } t_1 > t_2 > 0.$$

The above condition states that the valuation difference between the products increases in t (see Figure 2.2). In the following proposition we show formally that the expected utility functions over product classes resulting from the priority supply rule satisfy the above monotonicity condition. Again, for notational convenience, we suppress the argument in expected utility functions whenever such expression does not create ambiguity. That is, $EU_i(t_j)$ is denoted by EU_{ij} for all i and j defined.

<u>Proposition 2.1</u>

Let us assume that customers' utility functions satisfy the assumptions of section 2.1.1. That is, $\partial U/\partial t > 0$, $\partial U/\partial u > 0$, and $\partial^2 U/\partial u \partial t > 0$. Then, the expected utility functions of customers ordering the priority classes under the priority supply rule satisfy the monotonicity condition shown below : for priority class $i = 1, \cdots, M$, the quantity

$$(EU_{i1} - EU_{i+11}) - (EU_{i2} - EU_{i+12}) > 0 \tag{2.10}$$

for $t_1 > t_2$.

<u>Proof</u>: We define $EU_i^b(t)$ to be the conditional expected utility of customer t who selects priority class i given B = b, and a_{ijb} to be the amount of chip j delivered to priority class i given B = b under the priority supply rule.

We first prove that

$$(EU_{i1}^b - EU_{i+11}^b) - (EU_{i2}^b - EU_{i+12}^b) \geq 0 \tag{2.11}$$

Collecting the first term of each parenthesis,

$$\begin{aligned} & EU_{i1}^b - EU_{i2}^b \\ &= (a_{i1b}/(D_i - D_{i-1}))\ (U_{11} - U_{12}) + (a_{i2b}/(D_i - D_{i-1}))\ (U_{21} - U_{22}) \\ &\quad + \cdots + \ (a_{iNb}/(D_i - D_{i-1}))\ (U_{N1} - U_{N2}) \end{aligned} \tag{2.12}$$

Collecting the second terms of each parenthesis,

$$\begin{aligned} & EU_{i+11}^b - EU_{i+12}^b \\ &= (a_{i+11b}/(D_{i+1} - D_i))\ (U_{11} - U_{12}) + (a_{i+12b}/(D_{i+1} - D_i))\ (U_{21} - U_{22}) \\ &\quad + \cdots + \ (a_{i+1Nb}/(D_{i+1} - D_i))\ (U_{N1} - U_{N2}) \end{aligned} \tag{2.13}$$

By conditioning on whether the sum of conditional delivery probability of each spec. level for class i, $\sum_{j}^{N} (a_{ijb}/(D_i - D_{i-1}))$ is strictly less than 1 or is equal to 1, it can be verified that quantity (2.12) is bigger than or equal to quantity (2.13).

Since the expected utility of priority class i, EU_i is strictly greater than the expected utility of priority class i+1, EU_{i+1}, the quantity (2.12) minus (2.13) is strictly positive for at least one contingency b. Thus,

$$\sum_{b\epsilon\beta} Pr\{b\} \, [(EU^b_{i1} - EU^b_{i+11}) - (EU^b_{i2} - EU^b_{i+12})]$$
$$= (EU_{i1} - EU_{i+11}) - (EU_{i2} - EU_{i+12}) > 0.$$

$\square$

Now that we have met the sufficiency condition for monotone market segmentation, we can represent the market segmentation in terms of the following boundary customers relations.

For $i = 1, \cdots, M-1$

$$\sum_{j=1}^{N} Pr_{ij} \, U(u_j, t_i) - p_i = \sum_{j=1}^{N} Pr_{i+1j} \, U(u_j, t_i) - p_{i+1} \tag{2.14}$$

and,

$$\sum_{j=1}^{N} Pr_{Mj} \, U(u_j, t_M) - p_M = 0 \tag{2.15}$$

The above relations state that each of the boundary customers $t_1, \cdots, t_{M-1}$ is indifferent between the two neighboring market segments and the last boundary customer t_M is indifferent between selecting priority level M or withdrawing from the market.

Conventional formulae for the profit, consumer surplus and total surplus to be used in the numerical example in section 4 are provided as follows (see e.g., Oren and Min (1988)).

$$\text{Profit } \pi = \sum_{i=1}^{M} p_i \, (D_i - D_{i-1}) \tag{2.16}$$

Consumer Surplus:

$$CS = \sum_{i=1}^{M} \int_{t_{i-1}}^{t_i} \left(\sum_{j=1}^{N} Pr_{ij} \, U(u_j, t) \right) dD(t) - \pi \tag{2.17}$$

Total Surplus:

$$TS = \sum_{i=1}^{M} \int_{t_{i-1}}^{t_i} \left(\sum_{j=1}^{N} Pr_{ij} \, U(u_j, t) \right) dD(t) \tag{2.18}$$

The entire formulation for the producer's profit maximization problem is shown in Appendix 1.

3. Market Segmentation Analysis

From the boundary customers relations developed in section 2, we find that expected utility levels obtained from the priority classes as well as the corresponding prices are interdependent. In fact we can interpret priority class i as the upgraded version of priority class i+1, i = 1, ⋯, M-1 where priority class M is the "base" product. From this perspective, we define the premium price $\bar{p}_i$ for class i to be the incremental price paid for upgrading priority class i+1 to priority class i. We can then express the price p_i paid to purchase priority class i as the sum of incremental upgrade prices, i.e.,

$$p_i = \bar{p}_i + \bar{p}_{i+1} + \cdots + \bar{p}_M \tag{3.1}$$

Now from the boundary customers relations,

$$\bar{p}_i = EU_{ii} - EU_{i+1i} \qquad i = 1, \cdots, M-1$$

$$\bar{p}_M = EU_{MM} \tag{3.2}$$

and, profit $\pi = \bar{p}_1 \, D_1 + \bar{p}_2 \, D_2 + \cdots + \bar{p}_M \, D_M$ (3.3)

To further analyze the impact of chip distribution in each priority class on profits, we define the following.

a_{ijb} : the amount of chip j delivered to priority class i after downgrading, under the priority supply rule, when B = b

c_{ijb} : the per unit contribution of a_{ij} to the profit

We now derive the relation between the profit and quantities a_{ijb}'s in the following propostion..

Proposition 3.1

Given non-degenerate market segmentation under the priority supply rule, the profit is linear in a_{ijb}'s and the corresponding per unit contribution to the profit is given by

$$c_{ijb} = \Pr\{b\} \; ((-U_{ji-1} D_{i-1} + U_{ji} D_i)/ (D_i - D_{i-1})) \; .$$

Proof : From equations (3.2) and (3.3), the profit is given by

$$\pi = \sum_{b\in\beta} \Pr\{b\} \; [(EU^b_{11} - EU^b_{21}) D_1 + (EU^b_{22} - EU^b_{32}) D_2 + \cdots + EU^b_{MM} D_M] \qquad (3.4)$$

while the conditional expected utility of boundary customer j for priority class i is given by, $EU^b_{ij} = \sum_k a_{ikb} U_{kj}/(D_i - D_{i-1})$ (3.5)

By factoring a_{ijb} out, it can be verified that:

$$\pi = \sum_{a_{ijb}} [\Pr\{b\} \; (-U_{ji-1} D_{i-1} + U_{ji} D_i) \; / \; (D_i - D_{i-1}) \;] \; a_{ijb} \qquad (3.6)$$

which proves this proposition. □

Since the per unit contribution to the profit under any contingency is proportional to the probability of that contingency, it is convenient to define the conditional per unit contribution to the profit c_{ij} which is independent of any contingency. We denote the conditional per unit contribution by c_{ij}, where $c_{ij} = c_{ijb} / \Pr\{b\}$. We note that the c_{ij} depends on the market segment sizes and utility functions. The conditional per unit contribution c_{ij} may or may not be positive. This is the motivation to study " single exchange " perturbations of the priority supply rule. By single exchange, we mean that one unit of given spec. chips from a priority class is exchanged with another chip from another priority class. We show that such minor perturbation in the priority supply rule may or may not yield positive change in profit. This in turn indicates the profitability of downgrading for some cases. We now present the following corollary concerning the single exchange perturbation. The proof can be easily obtained from Proposition 3.1.

Corollary 3.1

Given non-degenerate market segmentation under the priority supply rule, suppose that we perturb the priority supply rule by exchanging one unit of u_i chip from class h with one unit of u_j chip from class k under contingency $b \epsilon \beta$ and adjust prices accordingly so as to preserve the original market segmentation. We further assume that this perturbation is sufficiently small so that the monotonicity condition of section 2 on the expected utility function over priority classes is preserved. Then $\delta(\pi)$, the change in the profit, is given by:

$$\delta(\pi) = \Pr\{b\} \; [\; (U_{ih-1} - U_{jh-1}) \, D_{h-1} + (-U_{ih} + U_{jh}) \, D_h] \; / \; (D_h - D_{h-1}) \\ + \Pr\{b\} \; [\; (-U_{ik-1} + U_{jk-1}) \, D_{k-1} + (U_{ik} - U_{jk}) \, D_k] \; / \; (D_k - D_{k-1}) \quad (3.7)$$

or equivalently,

$$\delta(\pi) = -c_{hib} + c_{hjb} + c_{kib} - c_{kjb} \quad (3.8)$$

Figure 3.1 illustrates the result of Corollary 3.1 for the case of M = 3. For simplicity the expected utility functions are assumed to be linear. We show the effect of exchanging one unit of u_i chip ($u_i > u_j$) from market segment 3 with one unit of u_j chip from market segment 2 while maintaining the same market segmentation. Such an exchange increases the expected utility of priority class 2 while it decreases the expected utility of priority class 3. Consequently the producer is able to increase the premium price $\bar{p}_2$. However, in order to prevent customer crossovers, the premium prices $\bar{p}_1$ and $\bar{p}_3$ must be reduced. The change in the prices and the profit can be seen in Figure 3.1. (prices are represented by the arrows and the changes in profit are represented by the shaded rectangles). We note that the total change in profit can be either positive or negative.

Now consider a case where one unit of u_i chip is downgraded as u_j chip in class k when B = b. This transaction can be viewed as a degenerate single exchange where u_i chip is exchanged with u_j chip from an external imaginary source. Then the incremental profit $\delta(\pi)$ is given by :

$$\delta(\pi) = \Pr\{b\} \; [\; (U_{ik-1} - U_{jk-1})\, D_{k-1} + (-U_{ik} + U_{jk})\, D_k \;]/(D_k - D_{k-1}) \quad (3.9)$$

which is the difference between the per unit contribution of a_{kjb} and a_{kib}.

Again the resulting $\delta(\pi)$ can be either positive or negative. Figure 3.2 depicts the net effect when a u_i chip is downgraded as a u_j chip in the second market segment. Here the example is drawn in such a way that for the producer, the extra profit [A] is much smaller than the loss [B]. Thus, in this case, downgrading decreases profit. This situation, however, is reversed in Figure 3.3 which shows larger [A] relative to [B] resulting in a increased profit.

3.2 Analysis of Downgrading

In this subsection we first present an example of downgrading. Next, we develop sufficiency conditions under which profit maximization under the priority supply rule will be achieved without downgrading. Such investigation is of importance because whenever these sufficiency conditions are met, the producer may rule out downgrading as a strategic option. This in turn will eliminate the possibility of "customer by-pass" (i.e., reinspection and re-classification of spec. levels) even when price differentials among spec. levels make such by-pass economically attractive. It will also simplify the formulation by reducing the number of variables and thus facilitate numerical solutions.

Example 3.1: Downgrading

In order to focus on the downgrading aspect, we will assume that the distribution of the relevant quality attribute is a spike at $u = 1$. That is, the entire production has uniform quality and the production quality is sufficient to meet the demand with uniform quality level $u = 1$. The question we want to address is whether the supplier could increase his profits by labelling some of the chips produced as a lower quality spec. level, which will then allow him to offer two priority levels at two different prices and segment the market. As for the demand data, let $D(t) = 1 - t^2$ and
$U(u, t) = (10u + 1)^t - 1$ where $t \in [0, 1]$ and $u \in [0, 1]$. To examine the effect of downgrading in this example we first solve the profit maximizing problem without downgrading option using relation (2.3) instead of (2.1) and (2.2). Then we solve the profit maximization problem with the downgrading option using the formulation shown in Appendix 1. Without the downgrading option, the optimal price table is as follows.

class	price	$u_1 = 1$
1	$p_1 = 4.55$	1

Table 3.1 Price Table for Example 3.1 with M = 1

At the optimum, we obtain profit $\pi = 2.226$, demand quantity $D_1 = 0.489$ and boundary customer $t_1 = 0.715$. The delivery schedule is as follows:

u_1 chips are allocated to all customers of priority class 1. The total chips delivered equals to the total demand quantity D_1.

With the down grading option, assuming two priority classes and solving for the prices and spec mix, we obtain the optimal price table as follows:

class	price	$u_1 = 1$	$u_2 = .36$
1	$p_1 = 4.617$	1	0
2	$p_2 = 1.297$	0	1

Table 3.2 Price Table for Example 3.1 with M = 2

The corresponding optimal profit is $\pi = 2.249$, demand quantity $D_1 = 0.436$ and $D_2 = 0.667$, boundary customer $t_1 = 0.751$ and $t_2 = 0.676$. The delivery schedule is as follows:

u_1 chips are allocated to all customers in priority class 1 while u_2 chips are allocated to all customers in priority class 2. The total amount of u_1 chips delivered to priority class 1 is 0.436 while the total amount of u_2 chips delivered to priority class 2 is 0.107.

We observe that the optimal strategy for a monopoly supplier is to downgrade some of the chips produced and offer two priority classes. By doing so, the supplier is able to achieve a higher profit without cannibalizing the higher quality market segment. These results show that downgrading can improve market penetration and increase profits.

To examine the relations between the conditional per unit contribution c_{ij}'s and profit changes in Example 3.1, we compute the numerical values of c_{11}, c_{12}, c_{21}, and c_{22} by applying Proposition 3.1. The calculation yields 5.056, 2.144, -0.003, .4186 respectively. With these values, assuming that the same market shares are maintained, we observe that downgrading one unit of u_1 as u_2 in the first segment is not profitable because the resulting change in the profit

$\delta(\pi) = -c_{11} + c_{12} < 0$ (by applying Corollary 3.1)

and indeed the numerical solution shows that no u_2 chips (i.e. downgraded chips) should be delivered to first class customers.

The above calculation suggests that profitability of downgrading may be inferred from the values of c_{ij}'s. This observation is formalized in the Proposition 3.2 below. The proof of Proposition 3.2 (as well as Proposition 3.3 in subsection 3.3) relies on the following lemma and corollary concerning the priority supply rule. The proofs can be obtained by simply applying the definition of the priority supply rule given in section 2.

Lemma 3.1

Given non-degenerate market segmentation, suppose chips from each bin are allocated to priority classes according to an allocation rule A. Then the property P described below holds if and only if the allocation rule A is the priority supply rule.

P: No single exchange or unilateral transfer that upgrades the product mix of the higher priority class can be made between any two priority classes k and l ($k < l$) for all $b \in \beta$.

Corollary 3.2

Under the identical assumptions as in lemma 3.1, property P holds if and only if property T described below holds.

T: No single exchange or unilateral transfer that upgrades the product mix of the higher priority class can be made between any two neighboring priority classes k and k+1, $k = 1, \cdots, M-1$.

We now present Proposition 3.2 as follows.

Proposition 3.2

Given non-degenerate market segmentation, assume that at the optimal solution with downgrading option under the priority supply rule, the per unit contributions c_{ij}'s are as follows.

$$c_{ij} \geq c_{ij+1} \text{ for } j = 1, \cdots, N-1 \text{ and } i = 1, \cdots, M .$$

Then, downgrading will not yield additional profit.

Proof : Consider a product strategy S which includes downgraded chips when B = b and assume that the profit corresponding to this strategy is strictly higher than the maximum profit achievable without downgrading. Let us now replace all the downgraded labels in strategy S with the original labels. Given any two neighboring classes k and k+1, if there is any chip with higher quality in class k+1 than the lowest quality chip given to class k, we can make a single exchange between these two classes to upgrade priority class k. Such an exchange will not decrease profit due to the following reasoning.:

Suppose a u_i chip of class k+1 (which may have been downgraded as u_j chip) is exchanged with a u_l chip of class k (which may have been downgraded as u_m chip). Then, from Proposition 3.2,the conditional profit difference $\hat{\delta}(\pi) = \delta(\pi) / \Pr\{b\}$, is as follows:

$$\hat{\delta}(\pi) = c_{ki} - c_{k+1j} + c_{k+1l} - c_{km} \qquad (3.10)$$

1) Since $u_i > u_m$ (the exchange upgrades the product mix quality of a higher priority class at the expense of a lower priority class), from the supposition that $c_{ij} \geq c_{ij+1}$ for all i and j, we have $c_{ki} - c_{km} \geq 0$.

2) Since $u_l \geq u_j$ (if $u_l < u_j$, then it would violate the priority supply rule assumption), from the supposition that $c_{ij} \geq c_{ij+1}$ for all i and j, we have $c_{k+1l} - c_{k+1j} \geq 0$.

It follows that $\hat{\delta}(\pi)$ of equation (3.10) is non-negative. Thus, the corresponding $\delta(\pi) \geq 0$.

Single exchanges of the kind described above may be repeated until such exchange can no longer be made. Thus, from Lemma 3.1 and Corollary 3.2, the resulting allocation is identical to that under the priority supply rule.

These exchanges will obviously alter the expected utilities of the various priority classes. However, the resulting expected utilities satisfy the monotonicity condition of section 2 and we can adjust the prices corresponding to these classes so as to maintain the same market segmentation as in the original product strategy S. The effect of such an adjustment on the profit is already accounted for in evaluating the cumulative effect of all the profit perturbation $\delta(\pi)$'s. The construction above thus restored all the labels so that no chip is downgraded without reducing profit. This, however, yields a contradiction since we assumed that the profit obtained with strategy S is strictly higher than the maximum profit achievable without downgrading.
□

Corollary 3.3

If the utility function is of the form U(u, t) = g(u) h(t) (this is referred to as product form) then the maximum profit is achievable without downgrading.

The above corollary can be easily proved by showing that when the utility function is of product form then the sufficient conditions of Proposition 3.2 are always satisfied (See Min (1989)).

3.3 Optimality of the Priority Supply Rule

Much of the discussion so far was based on the assumption that the allocation of the chips under every contingency follows the priority supply rule. This supply rule, which provides higher spec. level chips to higher ranked customers, is the most desirable allocation rule given our product line format from an economic efficiency point of view. The question still remains, however, whether the priority supply rule is also optimal from a profit maximization point of view. Perhaps there exist other allocation rules that will allow the supplier to increase his profits. In addressing this question, we will restrict ourselves only to contingency allocation rules which will preserve the monotone market segmentation. As indicated before, such monotonicity is essential to keep the problem tractable. In subsection 3.3, the definitions of a_{ijb} and c_{ijb} are extended so that they are applicable to any contingency supply rule consistent with the monotone market segmentation criteria as well as the priority supply rule. If the market segmentation is monotone, the premium prices and profit relations of (3.2) and (3.3) still hold. Therefore, this relaxation does not affect the formulae for c_{ij}'s or the profit change due to a single exchange given by Proposition 3.1 and Corollary 3.1 respectively.

First we illustrate the question we are addressing via an example.

Example 3.2 The Optimality of the Priority Supply Rule

Let $D(t) = 1 - t$ and $U(u, t) = ut$ where $t \epsilon [0, 1]$ and $u \epsilon [0, 1]$. For convenience both supply and demand are normalized to be within the unit interval. Also let the number of priority classes $M = 2$ and the number of spec. levels $N = 2$. In order to focus on the contingency allocation rules, we will make the following simplifying assumptions.

1) u_1, u_2, t_1, t_2 are pre-specified so that $u_1 = 0.9$, $u_2 = 0.6$, $t_1 = 0.7$, and $t_2 = 0.5$.

2) There is only one contingency b. (i.e., the production distribution over the classified attribute is identical for all the production runs.) and $S(u_1, b) = 0.3$ while $S(u_2, b) = 0.5$.

3) Downgrading is not allowed.

The question we want to address is whether the producer can increase his profit by using a contingency supply rule that is consistent with the monotone market segmentation criteria, but is different from the priority supply rule. In this example, we compare the prices and profit under the priority supply rule with those obtained under a supply rule T defined as below.

Under the priority supply rule all of u_1 chips are allocated to the first class while all of u_2 chips are allocated to the second class. By contrast, under supply rule T, from the u_1 chips produced, 0.2 are allocated to the first priority class and the remaining 0.1 to the second priority class. And from the u_2 chips produced, 0.1 are allocated to the first class and the remaining 0.1 to the second class.

Since u_1, u_2, t_1 and t_2 are pre-determined, prices can be uniquely derived from the boundary customers relations of (2.14) and (2.15).

The resulting prices and corresponding profit under the priority supply rule are given by, $p_1 = 0.51$, $p_2 = 0.3$, and $\pi = 0.213$. And the resulting prices and corresponding profit under the supply rule T are given by, $p_1 = 0.41$, $p_2 = 0.375$, and $\pi = 0.198$.

To examine the relation between c_{ij}'s and the supply rules, we compute the numerical values of c_{11}, c_{12}, c_{21}, and c_{22} by applying Proposition 3.1 and obtain the respective values 0.63, 0.42, 0.18, and 0.12. Since the conditional per unit contribution c_{ij}'s depend on the utility function, demand function, spec. levels and the boundary customers and in this example all of them are pre-determined, the c_{ij}'s values under the priority supply rule and those under the supply rule T are the same. Suppose that we alter the supply rule T so that all 0.3 of u_1 chips produced are allocated to the first class while all 0.2 of u_2 chips produced are allocated to the second class. (Note that the altered allocation rule is identical to the priority supply rule.) Then, by Corollary 3.1, the net change in the profit is given by,

$$\begin{aligned}\delta(\pi) &= 0.1 \; (c_{11} - c_{12} - c_{21} + c_{22}) \\ &= 0.015 > 0,\end{aligned}$$

The above calculation suggests that the relative profitability of a particular contingency supply rule (as compared to the priority supply rule) may be inferred from c_{ij} values. This observation is formalized in the following proposition.

<u>Proposition 3.3</u>

Let the supply rule R be an allocation rule which differs from the priority supply rule. Assume that when prices and spec. levels are optimized under allocation rule R, then the resulting price table leads to non-degenerate monotone market segmentation. Suppose that at the optimal solution, the c_{ij} values satisfy the following conditions:

$$c_{ki} - c_{kj} - c_{k+1i} + c_{k+1j} \geq 0 \quad \text{for all k, i, and j } (u_i > u_j) \tag{3.11}$$

$$\text{and,} \quad c_{ki} \geq c_{k+1i} \quad \text{for all k and i} \tag{3.12}$$

Then, the optimal profit under supply rule R can not be higher than the optimal profit under the priority supply rule.

<u>Proof</u> : Suppose the optimal profit under the supply rule R is strictly higher than that under the priority supply rule. Then,

(a) If we can make a single exchange of a u_i chip from class k+1 with a u_j chip from class k when B = b, then by (3.11) the conditional net profit change is

$$\hat{\delta}(\pi) = c_{ki} - c_{kj} - c_{k+1i} + c_{k+1j} \geq 0$$

(b) If we can make a unilateral transfer of a u_i chip from class k+1 to class k when B = b, then by (3.12)

$$\hat{\delta}(\pi) = c_{ki} - c_{k+1i} \geq 0$$

Single exchanges or unilateral transfers as described above may be repeated until such exchanges or transfers can no longer be made. Then, by Lemma 3.1 and Corollary 3.2, the altered allocation rule is the priority supply rule. By employing similar argument as in the proof of Proposition 3.2, we can produce a contradiction to the supposition that the optimal profit under R exceeds the optimal profit under the priority supply rule.

□

Corollary 3.4

If the utility function is of product form (i.e. $U(u, t) = g(u)\ h(t)$) then the maximum profit achieved with the the priority supply rule is at least as high as that achievable with any other contingent allocation that preserves monotone market segmentation.

Again, the above corollary is proved by showing that the sufficient conditions stated in proposition 3.3 are always satisfied when the utility function is of product form (See Min (1989)).

4. Two Part Pricing and an Illustrative Example

We turn our attention to the problem of how the producer actually charges the customers. Up until now the prices of the priority classes are defined to be the total charges to be collected from each customer according to the priority class he selects. But in practice, if the total charge is collected after the delivery (ex-post), customers may be tempted to place multiple orders with the intention to cancel some orders depending on the production outcome. On the other hand, if the total charge is collected before the delivery, it might be objectionable to most customers.

A reasonable approach to resolve this dilemma is to decompose the price of each priority class into two parts. That is, the price consists of a nonrefundable priority charge paid in advance to secure delivery of a given priority, and a delivery charge ex post for the actually delivered chips. To discourage multiple orders, the delivery charge paid by customers upon delivery should depend on the actual chips delivered, i.e., the same chips delivered under different priorities cost the same. The format of such two part pricing is analogous to two part tariffs discussed in the context of capacity pricing (installation charge and usage charges of Oren, Smith and Wilson (1985)) or in electric power pricing (demand charge and energy charge of Chao, Oren, Smith and Wilson (1986b)).

The customers are expected value decision makers as we have assumed throughout this paper. Thus, a customer's priority selection depends on the expected total price he pays, which is given by

$$p_i = \hat{p}_i + \sum_j Pr_{ij} \tilde{p}_j \tag{4.1}$$

where $\hat{p}_i$ is the priority charge for class i while $\tilde{p}_j$ is the delivery spec. charge for spec. u_j

The decomposition of these prices into priority and delivery prices may not be unique unless the probabilities Pr_{ij} have some special structure. When such ambiguity arises, additional criteria may be imposed. For example, in the electric power industry, the demand charge (ex-ante) corresponds to the cost of fixed capital while the energy (delivery) charge attempts to reflect the marginal cost of producing electric power.

In the following example, we elucidate important features of our model.

The production data of a certain kind of chip called product 10c from an anonymous semiconductor manufacturer (Leachman (1988)) are as follows.

Prod #	Total	Bin 1	Bin 2	Bin 3	Reject
1	616	162	353	101	0
2	618	115	423	78	2
3	617	137	378	101	1
4	618	111	416	91	0

(Prod # denotes the production run number)

Table 5.4 Actual Production Data between Oct. - Dec. 1987

In the period of October - December of 1987, there were four production runs for this particular chips. After testing, each chip was classified into one of the four bins according to the test results. The bins are numbered so that bin i chips have higher spec. level than bin i+1 chips for all i. Since chips of this particular kind are graded according to their overall performance on the tests, the corresponding spec. levels do not have physical units, e.g., nano-second. In this example, we assume that the spec. levels of bin 1, 2, and 3 have been set at $u_1 = 0.65$, $u_2 = 0.35$, and $u_3 = 0.05$ respectively. We fit the available data to a quadratic functional form shown below.

$$S(u, B) = 1 - (1-B)\, u^2 - B\, u$$

Using least square error fit for each production run and denoting the realizations of B by b_1, b_2, b_3, and b_4, we obtain the resulting estimates:

$b_1 = 0.7691$, $b_2 = 0.8545$, $b_3 = 0.857$, $b_4 = 0.9044$

The demand data are assumed to be as follows: $D(t) = 1 - t$ and $U(u, t) = ut$ where $t \epsilon [0, 1]$ and $u \epsilon [0, 1]$. We also assume that the number of priority classes M = 2 and the number of spec. levels N = 3. We solve this problem according to the formulation shown in Appendix 1. The resulting optimal solution is given by Table 4.2 and 4.3 below.

class	price	$u_1 = .89$	$u_2 = .778$	$u_3 = .585$
1	$p_1 = .51$	.508	.484	.008
2	$p_2 = .319$	0	.019	.967

Pr_{ij}

Table 4.2 Price Table of Example 4.1

B value	class 1	class 2
b_1	$a_{11} = .133$	$a_{21} = 0$
	$a_{12} = .114$	$a_{22} = .016$
	$a_{13} = 0$	$a_{23} = .187$
b_2	$a_{11} = .125$	$a_{21} = 0$
	$a_{12} = .122$	$a_{22} = .001$
	$a_{13} = 0$	$a_{23} = .202$
b_3	$a_{11} = .124$	$a_{21} = 0$
	$a_{12} = .123$	$a_{22} = 0$
	$a_{13} = 0$	$a_{23} = .203$
b_4	$a_{11} = .12$	$a_{21} = 0$
	$a_{12} = .119$	$a_{22} = 0$
	$a_{13} = .008$	$a_{23} = .191$

(a_{ij} denotes the amount of u_j chips delivered to class i)

Table 4.3 Contingency Delivery Table of Example 4.1

From Tables 4.2 and 4.3, we observe the following:

1) The corresponding profit, consumer surplus and total surplus defined by equations (2.21), (2.22) and (2.23) are given by

$\pi = .191$, CS = .0652, TS = .2562 respectively.

2) Since the utility function is of product form, without any calculation, we deduce that at the optimal profit under the priority supply rule is higher than or equal to the optimal profit under any contingency supply rule which results in monotone market segmentation (by Proposition 3.4). Furthermore, again due to the product form utility function, we deduce that at the optimal solution, downgrading is unprofitable (by Proposition 3.5). Therefore, a priori, we can simplify the problem by replacing the production - downgrading relations (A.4) and (A.5) with relation (2.3) shown in section 2.

The price table (Table 4.2) lists only the expected total price for each priority. Using the numerical values obtained above, we have from equation (4.1) the following system of linear equations

$$p_1 = .51 \;\; = \hat{p}_1 + .508\ \tilde{p}_1 + .484\ \tilde{p}_2 + .008\ \tilde{p}_3$$
$$p_2 = .319 = \hat{p}_2 \qquad\qquad + .019\ \tilde{p}_2 + .967\ \tilde{p}_3$$

The solution for the above equations is not unique since the system of equations is under determined (2 equations and 5 unknowns). A set of prices that satisfy the above relations are given by:

the delivery charges, $\tilde{p}_1 = 0.459$, $\tilde{p}_2 = 0.401$, and $\tilde{p}_3 = 0.322$

and, the priority charges, $\hat{p}_1 = 0.08$ and $\hat{p}_2 = 0$

In order to compare relative welfare levels of producer and customers in this particular example, we now compare the optimal profit, consumer surplus and total surplus with those obtained under the case where only one delivery class is offered at one price and the different spec. level chips (N = 3) are allocated to the customers through random rationing as described in subsection 2.2. The latter benchmark case can be formulated as an optimization problem using relations (A.1) to (A.8) restricted to M = 1. The optimal solution (prices, spec. levels and market segmentation) obtained under such a delivery scheme is summarized in Table 4.4 below.

class	p_1	u_1	u_2	u_3	t_1	D_1
1	.454	.876	.751	.623	.6	.4

Table 4.4 Price Table and Market Share under the Random Rationing

The corresponding profit, consumer surplus, and total surplus are given by, $\pi = 0.181$, CS = 0.0609, and TS = 0.2419

As shown above, the profit, consumer surplus and total surplus corresponding to the two priority classes are higher than those corresponding to a single delivery class. Therefore in this particular example, both customers and the producer are better off with the priority scheme, i.e., the priority scheme is Pareto Superior.

In this example we employed simple least square error fit. However, there exists an extensive literature dealing with the theory of density estimation, which provides more sophisticated methods than the one employed above. For examples, one can use the kernel estimation method or the maximum penalized likelihood estimation method (see e.g., Silverman (1986)).

5. Conclusion

In this study, we explored a new way of pricing and allocating semiconductor products. It is shown how a product line defined in terms of spec. levels and delivery probabilities with appropriate pricing results in monotone market segmentation where higher spec. level chips are allocated to higher valued consumption. Such pricing and allocation mechanism can greatly reduce the mis-allocation of chips due to the prevailing haphazard delivery policies in the capacity constrained semiconductor industry. Exploiting the fact that the products' composition (spec. levels and the corresponding delivery probabilities) is endogenously derived from the supply and demand data, we were able to derive the effect on profit level of a single exchange perturbation. The market segmentation analysis based on the single exchange perturbation in turn provides fundamental insights on important issues such as profitability of downgrading and optimality of the priority supply rule. We derived sufficiency conditions which guarantee that profit maximization under the priority supply rule is achieved without downgrading. We also derive sufficiency conditions under which the priority supply rule maximizes profit among a broad class of contingency supply rules.

Throughout this study, the customers were assumed to be expected value decision makers. However, modification of this assumption to incorporate risk aversion is certainly conceivable. Also we have assumed that the economic net benefit from detecting downgraded chips is negligible. Nonetheless, there is implicit limit on the magnitude of downgrading. If the price differentials are large enough, the customers might re-inspect the chips and reclassify their spec. levels independent of the spec. levels of the producer. This may alter the optimal pricing strategy of the producer.

In this study, we only considered the case of a single critical attribute. This in itself is not a severe restriction because published price schedules (see e.g., AMD (1986)) indicate that the majority of chips are priced over a single quality attribute. However, occasionally we also find chips priced over two critical quality attributes such as speed and power consumption. Therefore, it would be valuable in practice to devise models for at least two quality attributes. Finally, the semiconductor chips in our study are assumed to be monopolistic products. However, for some semiconductor products, there may be (functional) substitutes. For these products, the competition aspect must be represented in the pricing strategy and supply rules. We think that a game theoretic approach appropriate for such an environment will provide additional insights to the complex and difficult issues of pricing and allocation of semiconductor products.

Appendix 1 The formulation for the Producer's Profit Maximization Problem

$$\max \sum_{i=1}^{M} p_i \, (D_i - D_{i-1}) \tag{A.1}$$

s.t. $1=t_0 \geq t_1 \geq \cdots \geq t_M$ $\quad$ $1=u_0 \geq u_1 \geq \cdots \geq u_N$

$p_1 \geq p_2 \geq \cdots \geq p_M$ $\quad$ all variables ≥ 0

Boundary Customers Relations:

$$\sum_{j=1}^{N} Pr_{ij} \, U(u_j, t_i) - p_i = \sum_{j=1}^{N} Pr_{i+1j} \, U(u_j, t_i) - p_{i+1} \qquad i = 1,\ldots,M-1 \tag{A.2}$$

$$\sum_{j=1}^{N} Pr_{Mj} \, U(u_j, t_M) - p_M = 0 \tag{A.3}$$

Production – Downgrading Relations:

for all $b \epsilon \beta$, $j = 1, \cdots, N$

$$S(u_j, b) - S(u_{j-1}, b) = W_{jjb} + W_{jj+1} + \cdots + W_{jNb} \quad (A.4)$$

$$V_{jb} = W_{1j} + W_{2j} + \cdots + W_{jjb} \quad (A.5)$$

Priority Supply Relations:

$$Q_{ijb} - (Q_{ij-1b} - R_{i-1jb}) \geq 0 \quad (A.6a)$$

$$Q_{ijb} (Q_{ijb} - (Q_{ij-1b} - R_{i-1jb})) \leq 0 \quad (A.6b)$$

$$R_{ijb} - (R_{i-1jb} - (Q_{ij-1b} - Q_{ijb})) \geq 0 \quad (A.7a)$$

$$R_{ijb} (R_{ijb} - (R_{i-1jb} - (Q_{ij-1b} - Q_{ijb}))) \leq 0 \quad (A.7b)$$

$j = 1, \cdots, N$ and $i = 1, \cdots, M$ and for all $b \epsilon \beta$

where $R_{0jb} = V_{jb}$, $Q_{i0b} = D_i - D_{i-1}$, and $D_0 = 0$.

Delivery Probability Relations:

$$Pr_{ij} = \sum_{b \epsilon \beta} Prob\{ b \} ((Q_{ij-1b} - Q_{ijb})/(D_i - D_{i-1}))$$

$$i = 1, \cdots, M \text{ and } j = 1, \cdots, N \quad (A.8)$$

References

Adams, W. and J. Yellen (1976), " Commodity Bundling and the Burden of Monopoly," *Quarterly Journal of Economics*, 90, 475-498.

Advanced Micro Devices (1986), *Price Schedule/ OEM Price Book*, 1-51.

Chao, H., S. Oren, S. Smith and R. Wilson (1986a), " Unbundling the Quality Attributes of Electric Power; Models of Alternative Market Structures," *Report UER-169*, *University-wide Energy Research Group*, University of California at Berkeley.

_______, _______, _______, _______, (1986b), " Multilevel Demand Subscription Pricing for Electric Power," *Energy Economics*, (October), 199-217.

Cooper, R. (1984), " On allocative distortions in problems of self-selection," *Rand Journal of Economics*, 15-4, 568-577.

Goldman, M., H. Leland and D. Sibley (1984), " Optimal Non-Uniform Pricing," *Review of Economic Studies*, LI, 305-319.

Harris, M. and A. Raviv (1981), " A Theory of Monopoly Pricing Schemes with Demand Uncertainty," *The American Economic Review*, 347-365.

Intel (1985), *Component Quality/ Reliability Handbook.*
Leland, H. and R. Meyer (1976), " Monopoly pricing structures with imperfect discrimination," *The Bell Journal of Economics*, 7-2, 449-462.

Leachman, R. (1986), " Preliminary Design and Development of a Corporate-Level Production Planning System for the Semiconductor Industry," *Report no. ORC* 86-11, *Operations Research Center*, University of California at Berkeley.

__________ (1988), private communications.

Maskin, E. and J. Riley (1984), " Monopoly with incomplete information," *Rand Journal of Economics*, 15, 171-196.

Mirman, L. and D. Sibley (1980), " Optimal Nonlinear Prices for Multiproduct Monopolies," *The Bell Journal of Economics*, 11-2, 659-670.

Mirrlees, J. (1971), " An exploration in the Theory of Optimal Income Taxation," *Review of Economic Studies*, 38, 175-187.

Min K. J. (1989) "Pricing Delivery Priority and Spec. Level of Custom Semiconductor Chips", Ph.D Dissertation, Department of Industrial Engineering and Operations Research, University of California, Berkeley.

Min K. J. and S. S. Oren (1990), " Extending Priority and Spec. Level Pricing of Semiconductor Products to Multiple Customer Industries," Technical Report, Center for Interdisciplinary Research in Coordination, University of California, Berkeley.

Moorthy, K. S. (1984), " Market Segmentation, Self-Selection, and Product Line Design," *Marketing Science*, 3-4, 288-30.

Mussa, M. and S. Rosen (1978), " Monopoly and Product Quality," *Journal of Economic Theory*, 18-2, 301-317.

Oren, S. and K. J. Min (1988), " Priority Pricing for Electric Power with Upper Limits on the Number of Interruptions," *Report UER*-215, *University-wide Energy Research Group*, University of California, Berkeley.

Oren, S., S. Smith and R. Wilson (1982), " Linear Tariffs with Quality Discrimination," *The Bell Journal of Economics*, 13 (Autumn), 455-471.

__________, __________, ________ (1983), " Competitive Nonlinear Tariffs," *Journal of Economic Theory*, 29-1, 49-71.

__________, __________, ________, (1984), " Pricing a Product Line," *Journal of Business.*

__________, __________, ________, (1985), " Capacity Pricing," *Econometrica*, 53 (May), 545-567.

Riley, D. and A. Sangiovanni-Vincentelli (1986), " Models for a New Profit-Based Methodology for Statistical Design of Integrated Circuits," *IEEE Transactions on Computer-Aided Design*, CAD-5, 131-169.

Smith, S. (1986), " New Product Pricing in Quality Sensitive Markets," *Marketing Science*, 5-1, 70-87.

Styblinski, M. (1985), " A new formulation of yield optimization problem: maximum profit approach," *Proceedings 1983 IEEE international Symposium on Circuits Systems, Kyoto, Japan*, 1273-1276.

Vickrey, W (1971), " Responsive Pricing of Public Utility Services," *Bell Journal of Economics*, 2, 337-346.

Wilson, R. (1989), " Efficient and Competitive Rationing ", *Econometrica*, 57-1, 1-40, (1989).