Incentive Schemes for Semiconductor Capacity Allocation

Suleyman Karabuk
S. David Wu
Lehigh University

Report No. 01T-006

INCENTIVE SCHEMES FOR SEMICONDUCTOR CAPACITY ALLOCATION

Suleyman Karabuk[1] and S. David Wu[2]

[1]Operations Research Group, Milliken & Company, Spartanburg, SC

[2]Manufacturing Logistics Institute, Department of Industrial and Systems Engineering, Lehigh University, Bethlehem, PA

Abstract

The semiconductor capacity-planning problem involves product managers (PM) who privately observe a demand distribution for their product portfolio and the headquarters (HQ) that decides on capacity allocation based on the demand information submitted by the PMs. When the total anticipated demand exceeds available capacity, PMs compete for the scarce resources and thus use their private information as a leverage to maximize their divisional performance at the expense of company profits. We show that without a properly designed allocation game, the HQ cannot implement the optimal capacity decisions. In this paper, we design a capacity allocation mechanism that extracts privately observed demand information from the product managers while implementing the optimal solution that maximizes system wide total expected profits. This mechanism is supported by an incentive scheme that requires side payments and participation charges to the players. We find the conditions under which the surplus created by coordination exceeds the bonus payments; hence, the mechanism achieves budget balance and voluntary participation simultaneously. The results provide important insight on how to deal with misaligned incentives in the context of the capacity allocation problem in semiconductor industry.

1. Introduction

During the recent years of the economical boom in the United States, the high tech industry experienced a substantial demand increase for its products. In particular, the telecommunications sector expanded at a high rate with the demand coming from Internet companies and wireless communications products. This situation leads to a capacity shortage for the semiconductor manufacturers. Our study is based on our observations at a major microelectronics manufacturer specialized in telecommunications sector.

In semiconductor industry, the manufacturing process is made of two consecutive stages: wafer manufacturing and assembly and testing. In the first part of the process, the electronic circuit of the chip is printed on a silicon wafer, and this process has a lead-time of 6-12 weeks. Then the wafers are sent to the assembly and testing facilities to be put on printed circuit boards and become end-products during a lead time of 2-4 days. The wafer manufacturing facilities are fairly expensive to build (2-4 billion dollars) and it takes about a year to complete. The demand in high tech industry is known to be cyclic and particularly sensitive to the fluctuations in economy. Therefore, it is clear that managing the existing wafer manufacturing capacity rather than opening additional costly wafer fabs is crucial to be profitable for a high tech company. Even during the times of economic downturn, companies close down access capacity and capacity rationing becomes even more important to become or stay profitable.

In big companies with multiple administrative divisions, demand management authority is delegated to product managers (PMs) from various business units, whose performance is usually evaluated by the total divisional profits they make. These managers have the most accurate estimate on the demand as they directly interact with customers. A PM who is closely working with customers may anticipate a high demand from a certain customer but may not accurately know when that demand will take off. Since the manufacturing lead-time is very long, (6-12 weeks) it is often not possible to catch up with demand if not planned beforehand. When the total anticipated demand exceeds available capacity, PMs use their private information regarding demand to leverage it to their benefit at the expense of total company profits.

A PM performs poorly when anticipated demand is not materialized; consequently, the capacity allocations are adjusted. However, this causes lost opportunities until the adjustment is implemented due to long manufacturing lead-time. Moreover, this capacity adjustment is based on recent historical data rather than the private demand information that the PM has. For instance, right after a reduction in the capacity share of a PM due to poor past performance, demand for her business unit may increase substantially which in turn causing more lost opportunities. Hence, the allocation adjustment may not yield the optimal results without acting based on the private information from the participating PMs.

At the semiconductor manufacturer we worked with, the PMs start with an initial capacity assignment made at the start of a financial period. This is essential for the business units to negotiate contracts with customers. The PMs consider the initial allocation as a guaranteed capacity available to them throughout the period and try to fill it as much as possible. The managers at the Headquarters (HQ), who are responsible for interacting with the PMs and coordinating production planning, describe the current allocation method and its shortcoming as it relates to their business as follows.

> *"...The traditional capacity allocation method is to assign a certain number of wafer starts to each business by technology group based on some reference demand view, typically a demand view that is linked to a specific financial commitment. This type of allocation creates a sense of wafer-starts ownership, and has a tendency to cause business segments to hold on to their share of wafers until the last moment when they don't really need to make the starts, or they tend to build inventory. From a global asset utilization point of view, these allocations drive underutilization by trapping pocket of capacity to segments with a low-swing of demand, where at the same time there are segments short of supply because of a high-swing of demand. Because it is necessary to have some finite lead-time on the high-swings of demand, wafers that are relinquished at the point of execution are sometimes too late to capture the upswing". - Director of Integrated Circuit Business Planning & Rationalization.*

The HQ describes their problem as finding a method to reshuffle the capacity allocations as the demand information changes relative to the initial capacity allocation. The policy of taking away initial shares to prevent the *ownership* problem and its consequences is also under consideration.

In order to accomplish an optimal allocation of capacity, it is imperative to extract the privately known demand information from the PMs. In this study, we analyze game theoretical aspects of the capacity allocation problem and develop a capacity allocation mechanism that collects private information from the PMs and implements the optimal capacity shuffling that maximizes company wide profits. The capacity allocation mechanism, while improving system wide profits, compared to the traditional system, rewards the PMs using the existing bonus system based on the divisional profits and side payments. We also analyze the impact of the way the initial shares are set, on the incentives of the PMs.

1.1. Related Literature

The coordination of marketing and manufacturing functions of a large corporation is a well-known problem in the operations management literature. This is essentially a resource allocation problem with two important features: misaligned incentives and the private information of the participating decision makers who are thus motivated to gain personal benefits at the expense of company profits.

The problem has attracted different approaches in the literature. Celikbas et al (1999) devise penalty schemes to coordinate forecasting and production planning. Mallik and Harker (1998) develop a bonus function to extract private information from marketing and manufacturing divisions. Portheus and Wang (1991) develop a transfer pricing scheme to coordinate the capacity planning and allocation. Kouvelis and Lariviere (2000) generalize the approach of Portheus and Wang to a certain class of coordination problems. Groves and Loeb (1979) dispute the effectiveness of price mechanism for coordinating divisional managers and design performance evaluation measures based on divisional profits less the impact of bundling decisions on the profits of other divisions to facilitate coordination. On the other hand, Harris and Kriebel (1982) design optimal transfer pricing schemes for allocating resources in a certain setting. The capacity allocation problem has also been analyzed in the context of supply chain management where there is no central coordinator (Cachon Lariviere (1999)-1-2).

The major difference in our case is that an initial capacity allocation is assigned before production planning takes place at which point optimal allocation may change with respect to the initial allocation. This requirement actually distorts the incentives of the participants even more and makes it harder to facilitate coordination.

Another relevant line of research is the mechanism design with private information. The focus of this research is how to coordinate a trade between independent decision makers to maximize value created by the trade with the desirable properties of: voluntary participation and no external funding to support the transaction.

The results are mostly about impossibility of achieving this objective. Myerson and Satterthwaite (1983) prove that it is not possible to coordinate a buyer and a seller who owns an indivisible object if the support of probability distribution, that represent the beliefs, overlap. However, there are few possibility results too. Cramton et al. (1987) develop a bidding mechanism that assigns a jointly owned asset to the partner who values it most. The mechanism works only if the ownership distribution satisfies certain conditions. McAfee (1991) study conditions under which participants who have privately known amounts of a good transfer it among themselves in such a way that the final solution maximizes the total of valuations. Makowski and Mezzetti (1993) analyze a trading problem for an indivisible object with two buyers and one seller and characterize the conditions under which there is an implementable solution. Makowski and Mezzetti (1994) and Williams(1999) generalize the theory and characterize the conditions under which a mechanism with the desirable properties exist.

In the next section, we model the local problems of the PMs and describe the coordination problem. In section 3, we develop an optimal capacity allocation mechanism that satisfies the desirable properties we identify. Section 4 analyzes the mechanism and draws conclusions for its characteristics under certain conditions. A case study follows the analysis in Section 5, which is followed by conclusion in Section 6.

2. The Capacity Allocation Problem

2.1. The Product Managers' Local Problem

There are multiple planning points throughout a financial period that the business units make demand planning and prepare a priority ordered list that consists of actual customer orders and forecasts. The ordering represents the relative importance of entries in the list, which is assigned based on certainty or profitability an order. At this point, PMs take into account their capacity share and may add or delete entries in the list to match their initial share. This list is then communicated to the wafer fab to be scheduled for production. The scheduler at the fab level releases wafers to satisfy the orders in the list starting from the most important entries until the capacity share of the PM is filled. Capacity allocation has a decreasing marginal profit for a business unit, because each additional unit of capacity will be used to fill in a relatively less important order in the order list. The demand is highly volatile and even the actual customer orders placed at the time of planning are subject to significant changes throughout the manufacturing lead-time. We consider an aggregation of the order list that each PM submits and describe their model based on aggregate demand.

The demand analysis at the HQ level consists of historical data and input from the PMs. The analysis consists of identifying a nominal demand that is constant over the planning period. The variance on top of the nominal is very high and it can be accurately identified by historical data. However, the nominal demand is not easy to predict and the PMs has this information more accurately as they interact with the customers frequently and they can anticipate a shift in the nominal demand.

We describe PMs' local problem by a newsvendor model. Let ξ_i represent the realized demand for PM i's products and $F_i(\xi_i, \theta_i)$ represent the demand distribution with parameter θ_i. We assume that θ_i is the mean of the distribution, which represents the nominal demand for business unit i, and that only a PM knows it with certainty. The concavity of the newsvendor model with respect to capacity allocation accurately describes the effect of scheduling wafer releases with respect to the ordered demand list.

We express the capacity as the number of wafers that can be manufactured per period. Let r_i be the average profit from one wafer allocated to PM i. We normalize the unit production cost to zero without any loss of generality. In the highly volatile high tech industry, carrying inventory is highly undesirable. For custom-made products, a customer may stop purchasing a particular version of a chip without any contractual liability. Some other products may face a decline in demand or they may even be phased out during the manufacturing lead-time causing the inventory to be worthless to the manufacturer. In a recent example, CISCO Systems wrote off $2.25 billion inventory due to economic slowdown. Their customers suddenly cancelled their orders and the company does not see any possibility of selling that inventory in the next 12 months.

" … On Tuesday, Chief Financial Officer said the company [CISCO] plans to *scrap and destroy* the majority of the inventory because most of it *can't be sold because it was custom-built* …

Cnet.com news report May 9, 2001

Therefore, we assume that the expected resale value of inventory at the end of the period is less than the production cost. Let v_i represent the potential loss from one unit of leftover wafer at the end of the period (production cost minus salvage value) in PM i's newsvendor model. We can describe total profit function for PM i's business unit, under a demand realization of ξ_i and a capacity allocation of y_i, as follows.

$$\pi_i(y_i, \xi_i) = r_i \min(y_i, \xi_i) - v_i \max(y_i - \xi_i, 0)$$

The PMs are rewarded by a bonus based on the profits they realize after sales are finalized. Let *g(.)* denote the bonus function that is implemented by the corporate. We assume that *g(.)* is a strictly increasing function of realized profits tallied at the end of the accounting period. We assume that all decision makers are risk neutral and maximize their expected utility. Under this assumption, the PMs' local problem reduces to maximizing the total expected profits of their respective business units expressed by the following function.

$$E(\pi(y_i, \xi_i)) = \Pi_i(y_i, \theta_i) = r_i y_i - (r_i + v_i) \int_0^{y_i} F_i(\xi_i, \theta_i) d\xi_i$$

The local problem of a PM is described as maximizing the total expected profits subject to total capacity constraint of $y_i \le b$, where b is the total capacity. The optimal solution to this problem is the newsvendor solution as follows.

$$\overline{y_i}(\theta_i) = \min\left\{ F_i^{-1}\left(\frac{r_i}{r_i + v_i} \mid \theta_i \right), \; b \right\}$$

2.2. The Coordination Problem

We restrict our analysis to a corporate environment where two PMs compete for scarce capacity and the HQ, acting as a central coordinator, wants to maximize the total expected profits for the corporation. We assume the net profits less bonuses paid to PMs is always increasing in total profits. Therefore, the HQ's problem is equivalent to deciding on the capacity allocation that maximizes total expected profits across the two business units. We formally define the coordination problem as follows.

Problem CA

$$\underset{y_1,y_2}{Max}\ z(\theta_1,\theta_2)=\Pi_1(y_1,\theta_1)+\Pi_2(y_2,\theta_2)$$

$$s.t.$$

$$y_1+y_2\le b$$

For notational convenience we use $-i$ to indicate the PMs other than PM i. Let $y_i^*(\theta_i,\theta_{-i})$ be the optimal capacity allocation that solves problem CA. Notice that problem CA is concave with respect to the allocation decisions due to the newsvendor structure and therefore y_i^* is unique and defined by the first order conditions. We also define: $z^*(\theta_i,\theta_{-i})=\Pi_i(y_i^*,\theta_i)+\Pi_{-i}(y_{-i}^*,\theta_{-i})$.

We need the following assumptions to facilitate further analysis.

A1. The private information θ_i are independent.

A2. For every possible realization of θ_i, all of the available capacity is allocated at the optimal solution.

$$\sum_{i=1}^{2} y_i^*(\theta_i,\theta_{-i})=b \qquad \forall\,\theta_i,\theta_{-i}$$

A3. The demand distribution function always decreases with respect to an increase in demand mean.

$$\frac{\partial F_i(\xi_i,\hat{\theta})}{\partial\theta_i}<0 \qquad \forall i,\forall\hat{\theta}$$

This assumption has important implications that we use through the analysis. In the newsvendor model, this implies that the marginal profit contribution for a constant allocation increases as θ_i increases. Consider the optimality condition for problem CA: the marginal contribution at the optimal capacity allocation is equal across the PMs. However, an increase in θ_i causes the marginal profit at y_i^* to increase, therefore invalidating the optimality of the solution. The optimal solution is then readjusted by increasing y_i^* until the marginal profits are equal across the allocation of PMs. Consequently,

$$\frac{\partial y_i^*(\theta_i,\theta_{-i})}{\partial \theta_i} > 0, \text{ and } \frac{\partial y_i^*(\theta_i,\theta_{-i})}{\partial \theta_{-i}} < 0.$$

Moreover, with A3 we have,

$$\frac{\partial z^*(\theta_i,\theta_{-i})}{\partial \theta_i} > 0.$$

That is, an increase in any one of the θ_i increases the optimal total expected profits. This is inferred with the same argument above regarding the change in optimal solution with respect to an increase in θ_i . However, this does not necessarily mean that z^* () is concave with respect to θ_i .

The optimal allocation relies on private information from the respective PMs. Without acquiring the mean of the demand distribution information from the PMs, the HQ cannot implement the optimal capacity allocation. On the other hand, the PMs are clearly motivated to exaggerate the demand mean because their allocation increases with the demand mean they report to the HQ (by A3). Therefore, it is not possible to implement the optimal allocation without any additional incentive structure. In the next section, we set up a capacity allocation game, which applies a necessary incentive structure to implement the optimal capacity allocation.

3. The Capacity Allocation Game

Our objective is to design a capacity allocation game that will enable the HQ to implement the optimal capacity allocation with the participation of the PMs within the existing bonus system in the company. We will rely on mechanism design literature from microeconomics in our analysis (e.g. see Fudenberg and Tirole 1996). The general set-up for a mechanism consists of specifying two functions $\langle s,t \rangle$: the first being the *selection rule* that maps the messages communicated by the participating players to social outcomes and the second being the *transfer scheme* that determines the incentive payment to the participants. The players simultaneously announce their messages that maximize their utility with respect to the announced mechanism <s,t> and the resulting social outcome is implemented.

For the capacity allocation problem under consideration, the HQ announces the capacity allocation rule (or function) and the bonus payments schedule as a function of reported θ_i. The players in the capacity allocation game are the PMs who simultaneously announce a θ_i value that maximize their total bonus payments. The range of θ_i values that a PM can announce (message space) is restricted by the *prior beliefs* of the other PMs and the HQ, which we assume to be shared by every participant. The prior belief function represents the information that participants have about each others' private information and it is described by a distribution function. We will refer to the private information of the PMs as their type. Let $\Phi_i(\theta_i)$ be the distribution function that represents the prior beliefs of participants about the type of PM i with support $\theta_i \in [\overline{\theta}_i, \underline{\theta}_i]$. Furthermore, let θ_i^* and $\theta_i^{'}$ denote the actual type of PM i as it is privately known to her and the announced type by PM i respectively. We restrict our attention to direct revelation mechanisms where the announced type is taken as the actual type of PM i without any loss of generality since any mechanism that interprets the announced types differently can be converted to a payment equivalent direct mechanism (*refs here ...*).

3.1 The surrogate profit function for PMs.

In order to relate the side payments to the existing bonus system we define a surrogate profit function. Let $\overline{\pi}_i(\ )$ be the surrogate profit function for PM i, to which the bonus function $g(.)$ is applied to determine the total bonuses to be paid. Let $t_i(.)$ be the function that represents side payment to PM i. Then we have:

$$\overline{\pi}_i(\theta_i^{'}, \theta_{-i}^{'}, \xi_i) = \pi_i(y_i(\theta_i^{'}, \theta_{-i}^{'}), \xi_i(\theta_i^*)) + t_i(\theta_i^{'}, \theta_{-i}^{'}) \qquad \textit{for } i = 1..2.$$

The local problem of PM i is to maximize expected surrogate profit function for any given $\theta_{-i}^{'}$ and it is stated as follows.

Problem PM

$$\underset{\theta_i^{'}}{Max}\ \overline{\Pi}_i(\theta_i^{'},\theta_{-i}^{'},\theta_i^{*}) = E_{\xi}(\overline{\pi}_i(\theta_i^{'},\theta_{-i}^{'},\xi_i))$$

The surrogate profit function pays the PMs the profits from their own business units and an additional side payment. This motivates the PMs to consider not only a coordinated solution, which is induced by the side payment, but also to keep the profits of their business units high. Therefore, the incentive structure should serve as a means of coordinating the PMs in such a way that the bonuses paid are increased (together with the expected profits) and this increase is shared by the PMs appropriately.

3.2. Properties of a desirable mechanism

We are interested in mechanisms that implement the optimal solution to problem CA. Therefore we define the function s by $y_i^*(\theta_1^{'},\theta_2^{'})$. The side payment function t should support the optimal allocation rule and should have the following properties.

Truth telling is Bayesian Nash equilibrium policy for the PMs. That is, for each participating PM, announcing her true type maximizes her surrogate profit function in expectation with regards to others' types. More formally, the mechanism <*s,t*> is required to be *Bayesian incentive compatible* hence to satisfy the following condition:

$$E_{-i}[\overline{\Pi}_i(\theta_i^{*},\theta_{-i}^{'},\theta_i^{*})] \geq E_{-i}[\overline{\Pi}_i(\theta_i^{'},\theta_{-i}^{'},\theta_i^{*})] \qquad \forall i\ \forall \theta_i^{'}\ \forall \theta_i^{*}$$

where, $E_{\theta_{-i}}$ is the expectation operator with respect to prior belief function $\Phi_{-i}()$.

We require that the incentive payment is at most based on the total realized profits including the side payments. We call this a *budget-balanced scheme* with respect to the existing bonus system. It is also needed to satisfy the assumption we made while describing problem CA. The total expected profit after bonus payments to the PMs is maximized with $y_i^*(\theta_1^{'},\theta_2^{'})$ if the side payments constitute a *budget-balanced scheme*. Otherwise, depending on the bonus system, additional profits may be offset by the additional bonuses to be paid to the PMs. The coordination should ideally allocate the capacity optimally and the increase in total profits should increase the bonuses that the PMs expect to receive. This corresponds to the following *budget balance* constraint on the side payment function.

$$\sum_{i=1}^{2} t_i(\theta_i^{'},\theta_{-i}^{'}) \leq 0 \qquad \forall \theta_i^{'},\theta_{-i}^{'}$$

Notice that the budget constraint implies that the total side payments may amount to a negative value, which indicates that the total bonus payments may be based on a value that is less than total profits. This can be justified by the PMs by a significant increase in total profits by coordination that would otherwise not be possible.

In order to ensure voluntary participation of the PMs, the mechanism should pay off at least as much as a PM would get if she chooses not to participate. We consider this constraint in expectation with regard to the belief function of the participants. This is called the *interim individual rationality* constraint, which is ensured if the following relation holds:

$$E_{-i}[\overline{\Pi}_i(\theta_i^{*},\theta_{-i}^{'},\theta_i^{*})] \geq \Pi_i(bx_i,\theta_i^{*}) \qquad \forall \theta_i^{*}$$

where, E_{-i} is the expectation operator with respect to prior belief function $\Phi_{-i}()$ and x_i is the initial capacity share of PM i as a percentage of total capacity. The right hand side of the equation is the expected profit that PM i could have made if she did not participate in the capacity allocation game and chose to stick with her initial capacity share.

Violation of individual rationality in our context implies that by participating in the coordination a PM have to sacrifice from personal benefits, in terms of bonus payments, for the good of others. Even in an intra company environment, this is highly undesirable as a disadvantaged PMs may show less effort to increase divisional profits, which then leads to decreased system wide profits. In extreme cases, she can even seek employment elsewhere as this effectively reduces her total financial compensation.

3.3. Coordination Mechanism for Capacity Allocation

At the start of a financial period, an initial capacity share of x_i, expressed as a percentage of the total capacity, is assigned to the PMs. We assume that all available capacity is assigned initially. The basis of this assignment is the financial commitment targets of the PMs at the start of the period. However, due to demand volatility throughout the period, the demand patterns may invalidate the rationale behind the initial allocation. As the PMs observe their demand mean change, this information has to be used to update the allocation at the planning points during the financial period. The capacity allocation mechanism proposed below motivates the PMs to give up their initial share and facilitate a redistribution of capacity so as to maximize system wide expected profits. We assume that all the parameters used in the mechanism are publicly known except for the private information θ_i.

Mechanism CA

Define:

$$B = \int_{\underline{\theta}_i}^{\overline{\theta}_i} \int_{\underline{\theta}_{-i}}^{\overline{\theta}_{-i}} z^*(\theta_i, \theta_{-i}) d\Phi(\theta_i) d\Phi(\theta_{-i})$$

$$B_i(\theta_i) = \int_{\underline{\theta}_{-i}}^{\overline{\theta}_{-i}} z^*(\theta_i, \theta_{-i}) d\Phi(\theta_{-i})$$

$$C_i = \min_{\theta_i}\{B_i(\theta_i) - \Pi(x_i, \theta_i)\}$$

$$q = \frac{C_1 + C_2 - B}{2}$$

Step 1: PMs announce their private information θ_i', θ_{-i}', at the planning point, before demand is realized.

Step 2: PMs receive their share defined by the optimal capacity allocation that maximizes problem CA:

$$y_i^*(\theta_i', \theta_{-i}'), y_{-i}^*(\theta_i', \theta_{-i}').$$

Step 3: Wafer starts are released with respect to the reshuffled capacity allocation. After the manufacturing lead-time, demand is observed for both business units: ξ_i, ξ_{-i}, and divisional profits are realized and added to the surrogate profit function.

Step 4: Side payment to the surrogate profit function is made according to the following expression. The PMs get their bonuses on the final value of their surrogate profit function.

$$t_i(\theta_i^{'}, \theta_{-i}^{'}) = \pi_{-i}(y_{-i}^{*}(\theta_i^{'}, \theta_{-i}^{'}), \xi_{-i}) - C_i$$
$$+\frac{1}{2}[B + B_i(\theta_i^{'}) - B_{-i}(\theta_{-i}^{'}) - [\pi_i(y_i^{*}(\theta_i^{'}, \theta_{-i}^{'}), \xi_i) + \pi_{-i}(y_{-i}^{*}(\theta_i^{'}, \theta_{-i}^{'}), \xi_{-i})]] + \min\{0, q\}$$

Theorem 1.

Mechanism CA:

(a) implements the optimal capacity allocation,

(b) is Bayesian Incentive Compatible,

(c) supports a budget-balanced bonus structure for problem CA and satisfies interim individual rationality if and only if $q \geq 0$.

Proof: (See Appendix).

The mechanism essentially pays the realized total profits to each PM and charges a lump sum participation fee of C_i. The payment of the system wide total profits aligns the incentives of the PMs with that of the corporate, hence induces truth telling as equilibrium strategy. The charges preserve this incentive structure because they are in lump sum. The C_i is the maximum that can be charged without knowing the type of PM i and still preserving voluntary participation. According to the definition of C_i, PM i expects to make at least that much over all the possible realizations of her type in expectation with respect to the types of other PMs. Any charge above the C_i value may violate the individual rationality for PM i for certain values of her type.

The expression q measures the difference between total expected payments and the lump sum charges to the PMs. If q is positive then, the charges offset the payments, budget balance is achieved and individual rationality is ensured. However, if q is negative, then either budget balance is violated or individual rationality is jeopardized at the expense of achieving budget balance. In the former case, q represents the cost of private information in terms of additional bonuses to be paid to the PMs.

Mechanism CA is based on the existing bonus structure that is defined by the bonus function: $g()$.With the proposed mechanism the PMs are still paid based on their divisional profits, consequently they are motivated to put effort to close more sales deals and to increase the mean of their divisional demand.

4. Analysis of the Capacity Allocation Game

We want to know if we can identify cases, in which we can be ensured about the possibility of achieving budget balance and individual rationality simultaneously. In this section, we characterize the C_i values as a function of initial capacity shares and the characteristics of private information about the PMs' types and draw some insights into their interaction.

Lemma 1.

Let $\theta_i^{\min} = \{\theta_i \in [\overline{\theta}_i, \underline{\theta}_i] \mid min\{B_i(\theta_i) - \Pi_i(x_i, \theta_i)\}$. There exists $\theta_{-i}^{''} \in [\overline{\theta}_{-i}, \underline{\theta}_{-i}]$ such that the following results hold for every PM i. The value of $\theta_{-i}^{''}$ depends on the belief function and the expected profit function for PM $-i$.

Case	Condition	$\theta_i^{\min} =$	$C_i =$
(a) $x_i > 0$	(1) $y_i^*(\theta_i, \theta_{-i}^{''}) > x_i \quad \forall \theta_i \in [\overline{\theta}_i, \underline{\theta}_i]$	$\underline{\theta}_i$	$B_i(\underline{\theta}_i) - \Pi_i(x_i, \underline{\theta}_i)$

	(2) $y_i^*(\theta_i,\theta''_{-i}) < x_i \quad \forall \theta_i \in [\overline{\theta}_i, \underline{\theta}_i]$	$\overline{\theta}_i$	$B_i(\overline{\theta}_i) - \Pi_i(x_i, \overline{\theta}_i)$
	(3) else	$\{\theta_i \mid y_i^*(\theta_i,\theta''_{-i}) = x_i\}$	$\int_{\underline{\theta}_{-i}}^{\overline{\theta}_{-i}} \Pi_{-i}(y_{-i}^*(\theta_i^{\min},\theta_{-i}),\theta_{-i}) d\Phi(\theta_{-i})$
(b) $x_i = 0$		$\underline{\theta}_i$	$B_i(\underline{\theta}_i)$

Table 1. Characterization of participation charges from the PMs.

First, we focus on the no-initial-share policy, case (b) in lemma 1. We have the following strong conclusion stated by the next theorem.

Theorem 2.

Individual rationality and budget balance always holds with zero initial share: $x_i = 0 \ \forall i$. Moreover, $q > z^*(\underline{\theta}_i, \underline{\theta}_{-i})$.

Proof (See Appendix).

This result proves that discontinuing the initial capacity assignment policy overcomes the inefficiency that may be caused by private information under any business environment that can be described by the model. This is an important conclusion in that it provides a trade off between disadvantaging the PMs by putting them in uncertainty during their dealings with the customers and implementing an incentive compatible, individually rational optimal capacity allocation. Depending on the business environment this may be a viable policy.

Unfortunately, the positive-initial-share case is more complicated. The three regions for the initial capacity share, as shown in Table 1, have important practical implications. If x_i satisfies (a.1), then this implies that PM i will always have, in expectation with regards to the types of the other PM, more capacity after the capacity allocation game is played. That is, she expects to be a capacity buyer from the other PM, because we assume that initial shares sum up to the total capacity. Similarly, if x_i satisfies (a.2), then this implies that PM i will be a capacity seller in expectation. Any initial capacity allocation that leads to leaving one of the PMs to be an expected buyer or an expected seller is undesirable at the corporate level because of the delicate politics among the PMs which we cannot capture with our game theoretical model. Therefore, we will focus on the initial capacity allocations that fall into case (a.3) for all PMs. The next proposition restricts the feasible solution space for the initial allocation even more.

Lemma 2.

Assume case (a.3) in Table 1.

(a) If for PM i $\theta_i^{\min} = \overline{\theta}_i$, then $\theta_{-i}^{\min} < \overline{\theta}_{-i}$ $\forall i$.

(b) If for PM i $\theta_i^{\min} = \underline{\theta}_i$, then $\theta_{-i}^{\min} = \underline{\theta}_{-i}$ $\forall i$, if and only if $\sum_{i=1}^{2} x_i < b$.

The possibility of achieving both individual rationality and budget balance is not conclusive with the positive initial share case. It depends on the belief functions and the cost structure of the local newsvendor problem as well as how the initial shares are distributed. However, if we relax the requirement to distribute the total capacity as initial shares, then we find a compromise that ensures all the desirable properties.

Theorem 3.

Individual rationality and budget balance always hold if the initial shares are set as:

$$x_i = y_i^*(\underline{\theta}_i, \theta_{-i}^{"}) \qquad \forall i,$$

where, $\theta_{-i}^{"}$ is defined as in Lemma 1.

Proof (See Appendix).

The theorem states that if the initial shares are set in such a way that all the PMs are going to end up with the same capacity share, in expectation with regards to the other PMs' type, at their lowest type, then individual rationality and budget balance holds simultaneously. However, by lemma 2, we know that this can be achieved if and only if the initial shares are less than total capacity. The initial shares described by theorem 3 represent the least acceptable quantity from PMs perspective as initial capacity assignment. The unassigned capacity will be in the ownership of the HQ and it will be totally distributed with the mechanism CA. Therefore, this policy does not interfere with the politics between the PMs. Deferring the assignment of part of the capacity can be viewed as a bargaining tool for the HQ against the private information of the PMs.

5. Case Study

In this section, we describe a case we have observed at the company and apply the theory we have developed in the previous sections.

We consider two PMs from different business units competing for the total capacity. We assume that the PMs have identical newsvendor models and that their local problem differs only in the demand mean. This implies that the unit profit from one wafer and the cost for one unsold wafer are the same for both PMs, and that the demand distribution has the same variance at the same mean value. We observed that this assumption seemed to hold at the company. The wafer manufacturing costs are very similar because they are manufactured at same fabs and the only difference is the setup made for different circuits that are printed on the wafer. The company tries to attain a constant average profit rate for all technologies; therefore, on the average the business units have similar profit margins. At the end-product level the profit rate may be different but at the wafer level the average is quiet similar. We also observed that individual orders can accurately be represented by a normal distribution. With the help of the central limit theorem, we further assume that the total demand is also normally distributed.

A significant portion of the demand comes from big customers in the form of custom-made chips for their products. Once a production run is completed for a customer, that particular batch of wafers cannot be used to satisfy demand for another customer. We model this case by a normally distributed demand with constant coefficient of variation (γ). Therefore, the demand distribution takes the following form, where θ_i is the mean and (γ) is the coefficient of variation of the demand distribution: $F_i(\xi_i \mid \theta_i, \gamma\theta_i)$. As the expected volume of orders increases, so does the variance of the distribution. We assume that mean is the private information and the coefficient of variation is publicly known in the corporate.

The demand distribution defined this way also satisfies Assumption 3 in section 2.2. Under these assumptions optimal allocation can be represented by a simple proportional allocation rule (Cachon and Lariviere, 1999-2) as defined by:

$$y_i^*(\theta_i) = \min\left\{\overline{y_i}(\theta_i),\ b\frac{\theta_i}{\theta_1 + \theta_2}\right\} \qquad i = 1,2$$

for any coefficient of variation (γ). When the PMs get allocations in proportion to their expected demand, the corporate wide total expected profits are achieved.

We represent the belief functions by beta distribution, which is well suited to describe a random process in the absence of relevant data (Law and Kelton, 1991). The nominal demand, which is the private information observed by the PMs, is not readily extractable from historical data that is available to all the PMs and the HQ. The short product cycle-time and the dynamic market for the output of the industry limit the use of past performance data for planning purposes. However, historical or market research data can be used to determine the support of the belief function distribution. An important characteristic of demand in the industry is that the nominal demand can frequently ramp up or ramp down depending market conditions and the life cycle of the product. There are even odd cases when demand is expected to ramp up shortly before ramping down and eventually resulting in phasing out of the product. We assume that a ramp up (down) can be represented by an average constant nominal demand through the planning period.

If very little is known about the demand behavior, which is often the case for newly introduced products, a uniform distribution can accurately describe the beliefs. A probability density that is skewed right assigns higher probability to higher nominal demand values within its range; therefore, such a distribution indicates an anticipation of a ramping up of nominal demand. Similarly, a left skewed probability density represents anticipation for ramp down in nominal demand. The two parameters of the beta distribution can be set to obtain a variety of shapes for the probability density including the uniform distribution. Therefore, it provides a uniform framework to describe the beliefs of the participants in the capacity allocation game.

The data values used in this example is listed in table 2 below.

Data	Value	Data	Value
r	50	θ_1	[1000,2000]
v	25	θ_2	[1500,2500]
b	2550	$\phi_1(\theta_1), \phi_2(\theta_2)$	Beta(1,1) : uniform – no information Beta(1,3) : skewed right – ramp up expected Beta(3,1) : skewed left – ramp down expected
γ	0.05		

Table 2. Data used in the example.

The data we used captures the relationship between data items as we have observed at the company. We have set the total capacity such that at their lowest types both PMs get the optimal allocation that solves their local problem (i.e. $y_i^* = \overline{y}_i, \forall i$).

In this experiment, we compute the C_1, C_2 and B values under a variety of initial capacity allocations and belief functions. The results are reported in Tables 3-5. In the company, every PM gets at least 25% of the capacity as initial share, therefore we covered initial allocations from 25% to 75% for both PMs. We used Maple version 6.0 to carry out the computations. In order to reduce computational requirements to match to our hardware capabilities, we discretized the beta distribution for the belief functions. We divided the support of the probability distribution to 10 equal intervals and took the middle point of the interval as the value of the random variable and the probability density of the interval as the probability. We conducted pilot runs to find the number of intervals so that the accuracy loss resulting from discretization is negligible. The C_i values are found by total enumeration over the discretized values of θ_i.

First, we look at the uniformly distributed belief function case. As seen from Table 3 below, there are no initial allocation settings where q is nonnegative except the (0.75, 0.25) case. However, initial shares from (0.6, 0.4) through (0.75, 0.25) essentially falls into case (a.2) and (a.1) for PM 1 and 2 respectively and are therefore unacceptable. Particularly, at the initial share configuration of (0.75, 0.25) PM 1 is allocated a capacity that is more than her optimal newsvendor solution ($\overline{y_i}(\theta_i)$) for most of her possible types. Such initial configurations actually defeat the purpose of assigning an initial allocation.

Another point Table 3 shows is that q is very small compared to B. Therefore, budget shortage actually leads to a negligible amount of extra bonuses to be paid to the PMs. Alternatively, the violation of individual rationality can be acceptably low. This situation can be explained by looking at the effect of increasing θ_i on the optimal total expected profit function $z^*()$. As θ_i increases so does $z^*()$. However, the rate of increase drops at a high rate at high values of θ_i as described by the derivative below (and Lariviere, (1999-2) and by the envelop theorem):

$$\frac{\partial z^*(\theta_i,\theta_{-i})}{\partial \theta_i} = (r+v)\frac{1}{\theta_i^2}F(\frac{b}{\theta_i+\theta_{-i}},1).$$

This implies that $z^*()$ shows little sensitivity to θ_i and therefore the value of information in this environment is relatively low.

x_1,x_2	$\theta_1^{\min},\theta_2^{\min}$	C_1, C_2	C_1+C_2	q
0.25,0.75	*1050*$^-$, *2450*$^+$	95308, 31875	127183	-274
0.30,0.70	*1050*$^-$, *2450*$^+$	88933, 38250	127183	-274
0.35,0.65	*1050*$^-$, 2150	82559, 44624	127184	-273
0.40,0.60	1150, 1750	76430, 50993	127424	-33
0.45,0.55	1350, *1550*$^-$	70123, 57122	127245	-212
0.50,0.50	1650, *1550*$^-$	63749, 63433	127183	-274
0.55,0.45	*1950*$^+$, *1550*$^-$	57375, 69808	127183	-274
0.60,0.40	*1950*$^+$, *1550*$^-$	51000, 76183	127183	-274
0.65,0.35	*1950*$^+$, *1550*$^-$	44627, 82558	127186	-271

0.70,0.30	*1950*$^{+}$, *1550*$^{-}$	38386, 88933	127319	-138
0.75,0.25	*1950*$^{+}$, *1550*$^{-}$	33599, 95308	128907	1450

Table 3.Computational results for uniformly distributed belief fn. (B=127457).
($^{+,\ -}$ indicates the upper and lower support for the parameter respectively).

Next, we look at the effects of different belief functions as shown in Table 4 below. It is clear that coordination becomes easier as higher level of demand mean is anticipated. A higher demand means an increase in the total system profits, which leads to an increase in the surrogate profit function for the PMs. Thus, they become more willing to participate to get the befits from increasing total profits. Especially with the ramp-up case the q value, although negative, is negligibly small. Another observation is that the initial allocation value of (0.4, 0.6) induced the highest q under all the belief functions. This happens to be very close to the optimal allocation that is based on the expected value of belief functions: $y_i^*(E[\theta_i], E[\theta_{-i}])$. The expected value based optimal allocations are (0.41, 0.59), (0.42, 0.58) and (0.43, 0.57) for ramp-down, uniform and ramp up respectively. It seems that, a good rule of thump is to set initial shares based on expected value of beliefs.

x_1,x_2	q (Ramp-down)	q (Uniform)	q (Ramp-up)
0.25,0.75	-437	-274	-2
0.30,0.70	-437	-274	-2
0.35,0.65	-436	-273	-1
0.40,0.60	-17	-33	-1
0.45,0.55	-377	-212	-1
0.50,0.50	-437	-274	-2
0.55,0.45	-437	-274	-2
0.60,0.40	-437	-274	-2
0.65,0.35	-434	-271	1
0.70,0.30	-301	-138	134
0.75,0.25	1287	1450	1722

Table 4. Comparison for different belief functions.

We also look at the initial shares that ensure individual rationality and budget balance as characterized by theorem 3. As shown in Table 5, the total allocation varies between 80- 90% of the total capacity. Similar to what we have observed before, as high demand levels are anticipated, the total that needs to be allocated decreases with an increase in the q value, hence the budget surplus.

Belief	x_1,x_2	Total allocated	q
Ramp down	0.35, 0.55	0.90	11803
Uniform	0.35, 0.55	0.90	12224
Ramp up	0.30, 0.50	0.80	25486

Table 5. The initial capacity allocation that satisfies conditions of theorem 3.

6. Conclusion

In this study, we have analyzed a capacity allocation problem at a major US semiconductor manufacturer in a game theoretical setting. We showed that the incentives of the PMs who are competing for scarce capacity are not properly aligned with the company wide interests under the current bonus structure. This is due to the private information that the PMs have regarding the demand mean for their business units. Another factor that contributed to the problem is the initial capacity share that is assigned to PMs before the planning period.

We have developed a capacity allocation mechanism in the form of side payments as a function of the reports of PMs about their private information. Under the proposed allocation mechanism, the PMs are induced to announce their true private information knowing that they will forego their initial share and that system wide optimal allocation is going to be implemented by the HQ. The bonuses are then paid after the profits in all business units are realized and observed. We investigate the situations when we can attain voluntary participation of the PMs and budget balanced bonus payments by the HQ. Our results characterize the conditions under which these properties can be attained. Our main conclusion is that the inefficiency generated by the PMs having private information can be counteracted by reducing the total of initial assignment and keeping a fraction of the capacity for competition at the time of production planning. The example in our case study indicated that the initial assignment that supports a desirable coordination could be as high as 90% of the total capacity.

We considered two competing PMs in our analysis. Our next step is to generalize the results to any number of PMs. In our analysis, we assume that the belief functions are common knowledge among the PMs and the HQ. Consequently, we also require that incentive compatibility and individual rationality hold in expectation. However, our assumptions may not hold true in some cases. For example, PMs who are operating in different markets may not posses a belief function regarding the demand of the others. In order to relax the common knowledge assumption we have to impose tighter restrictions on the incentive compatibility and individual rationality. Specifically, these conditions have to hold without regard to the type of the participating PMs. Then the analysis will differ in that the participation charges to the PMs will have to be less in order to accommodate the relaxed assumptions. Resolving this issue is reserved for future research.

APPENDIX

Proof: Theorem 1.

(T1.a) By definition of $y_i^*(\theta_i', \theta_{-i}')$ in section 2.2.

(T1.b) Consider the payments to the surrogate profit function under the mechanism $<y^*, t>$.

(*i*) Each PM receives the total realized profits.

(*ii*) We define the *total participation charge* in expectation with regards to demand for a PM as follows.

$$h_i(\theta_i', \theta_{-i}') = C_i - \frac{1}{2}[B + B_i(\theta_i') - B_{-i}(\theta_{-i}') - z^*(\theta_i', \theta_{-i}')] - \min\{0, q\}$$

The expected participation charge with respect to the type of the other PMs is:

$$E_{\theta_{-i}}[h_i(\theta_i', \theta_{-i}')] = C_i - \frac{1}{2}[B + B_i(\theta_i') - B - B_i(\theta_i')] - \min\{0, q\} = H_i \qquad \forall \theta_i'$$

It is clear that the participation charge is *lump sum* in expectation with regards to the type of the other PMs.

From (*i*) and (*ii*), <*y**, *t*> is a Groves mechanism in expectation and by the equivalence theorem, <*y**, *t*> is *Bayesian incentive compatible* (Makowski and Mezzetti 1994).

(T1.c) The mechanism pays the realized profits to the respective business units plus the extra payments defined by *t*. We show that the total extra payments sum up to be less than or equal to zero.

$$\sum_{i=1}^{2} t_i(\theta_i^{'}, \theta_{-i}^{'}) = B - [C_i + C_{-i}] + \min\{0,\ [C_i + C_{-i}] - B\}$$

Consider the right hand side of the equation above. If total expected pay, *B*, is larger than the total expected charge, (C_1+C_2), then the third term on the right hand side deducts the difference from the total payments to the surrogate function and makes the equation evaluate to zero. On the other hand, if there is a surplus in the expected payments, then total payments evaluate to a negative value without any adjustment and the surplus is captured by the designer in terms of less bonus payment.

(T1.d) By participating in the capacity allocation game, PM *i* is paid $B_i(\theta_i)$ in expectation with regards to the type of the other PM. On the other hand, she gives up what her initial share would pay her, which is $\Pi(x_i, \theta_i)$. The PMs are expected utility maximizers, therefore the payments are considered as expected values with regards to demand. Consequently, without knowing PM *i*'s type, the HQ can at most charge a participation fee of C_i. The PM *i* expects to make at least C_i or more compared to not participating, at any realization of her type θ_i. Consider the lump sum participation charge H_i defined in proof of (1.b). H_i equals C_i if and only if $q \geq 0$, otherwise $H_i > C_i$ and the interim individual rationality of PM *i* is violated.

Proof: Lemma 1.

(L1.a) Consider $x_i > 0$.

(*i*) By the envelop theorem we have,

$$\frac{\partial C_i(\theta_i)}{\partial \theta_i} = \int_{\underline{\theta}_{-i}}^{\bar{\theta}_{-i}} \left(-(r_i + v_i) \int_0^{y_i^*(\theta_i, \theta_{-i})} \frac{\partial F_i(\xi_i, \theta_i)}{\partial \theta_i} d\xi_i + (r_i + v_i) \int_0^{x_i} \frac{\partial F_i(\xi_i, \theta_i)}{\partial \theta_i} d\xi_i \right) d\Phi(\theta_{-i}) = \Delta .$$

By assumption 3, it is clear that Δ increases as θ_i increases.

(*ii*) By the first mean value theorem for integrals, there exists $\theta_{-i}^{"} \in [\bar{\theta}_{-i}, \underline{\theta}_{-i}]$ such that $B_i(\theta_i) = z^*(\theta_i, \theta_{-i}^{"})$ (e.g. see Trench 1978). Therefore, we can rewrite Δ as,

$$\Delta = -(r_i + v_i) \int_0^{y_i^*(\theta_i, \theta_{-i}^{"})} \frac{\partial F_i(\xi_i, \theta_i)}{\partial \theta_i} d\xi_i + (r_i + v_i) \int_0^{x_i} \frac{\partial F_i(\xi_i, \theta_i)}{\partial \theta_i} d\xi_i$$

According to the mean value theorem $\theta_{-i}^{"}$ may change value with respect to θ_i. However, we know from (*i*) that Δ is an increasing function of θ_i. Therefore, Δ increases regardless of the value of $\theta_{-i}^{"}$. The term Δ is negative as long as $y_i^*(\theta_i, \theta_{-i}^{"}) < x_i$, hence C_i is decreasing up to the θ_i value until $y_i^*(\theta_i, \theta_{-i}^{"}) = x_i$ where it evaluates to zero. Beyond that point, as θ_i is increased and $y_i^*(\theta_i, \theta_{-i}^{"}) > x_i$ satisfied. Consequently, Δ always takes a positive value and C_i increases.

Derivation of C_i values for cases (a.1) and (a.2) is straightforward. For case (a.3): by the mean value theorem for integrals we have,

$C_i = \Pi_i(y_i^*(\theta_i^{\min}, \theta_{-i}^{"}), \theta_i^{\min}) + \Pi_{-i}(y_{-i}^*(\theta_i^{\min}, \theta_{-i}^{"}), \theta_{-i}^{"}) - \Pi_i(x_i, \theta_i^{\min})$.

Since $y_i^*(\theta_i^{\min}, \theta_{-i}^{"}) = x_i$, it follows that

$$C_i = \Pi_{-i}(y_{-i}^*(\theta_i^{\min}, \theta_{-i}^{"}), \theta_{-i}^{"}) = \int_{\underline{\theta}_{-i}}^{\bar{\theta}_{-i}} \Pi_{-i}(y_{-i}^*(\theta_i^{\min}, \theta_{-i}), \theta_{-i}) d\Phi(\theta_{-i}).$$

(L1.b) Consider $x_i = 0$.

This is a special case of the previous case where the optimal allocation is always greater than initial share, which is zero. Therefore $\theta_i^{\min} = \underline{\theta}_i$ and the second term in C_i evaluates to zero.

Proof: Lemma 2

(a) The *if* part implies that $y_i^*(\overline{\theta}_i,\theta_{-i}^{"}) = x_i$. From assumption 2, $y_{-i}^*(\overline{\theta}_i,\theta_{-i}^{"}) = b - x_i$. Suppose PM $-i$ has $\theta_{-i}^{\min} = \overline{\theta}_{-i}$, then $y_{-i}^*(\theta_i^{"},\overline{\theta}_{-i}) = b - x_i$. However, this cannot hold true since $\overline{\theta}_{-i} > \theta_{-i}^{"}$ and $\overline{\theta}_i > \theta_i^{"}$, and by assumption 3. In fact, $y_{-i}^*(\theta_i^{"},\overline{\theta}_{-i}) > b - x_i$. Since y_{-i}^* is increasing with respect to $\theta_{-i}^{\min}$, the *then* part should be satisfied.

(b) The argument is same as (a) for the stated types.

Proof: Theorem 2.

Let $\overline{C}(\theta_1,\theta_2) = z^*(\underline{\theta}_1,\theta_2) + z^*(\theta_1,\underline{\theta}_2)$, $\overline{B}(\theta_1,\theta_2) = z^*(\theta_1,\theta_2)$. Hence, from Lemma 1 and Theorem 1 we have,

$$C_1 + C_2 = \int_{\underline{\theta}_1}^{\overline{\theta}_1}\int_{\underline{\theta}_2}^{\overline{\theta}_2}[\overline{C}(\theta_1,\theta_2)]d\Phi(\theta_1)d\Phi(\theta_2)$$

$$B = \int_{\underline{\theta}_1}^{\overline{\theta}_1}\int_{\underline{\theta}_2}^{\overline{\theta}_2}\overline{B}(\theta_1,\theta_2)d\Phi(\theta_1)d\Phi(\theta_2)$$

If we can show that $\overline{C}(\theta_1,\theta_2) \geq \overline{B}(\theta_1,\theta_2)$, then this implies that $C_1 + C_2 \geq B$.

(*i*) At the lowest values of types $(\underline{\theta}_1,\underline{\theta}_2)$ we have,

$\overline{C}(\theta_1,\theta_2) = 2z^*(\underline{\theta}_1,\underline{\theta}_2) \geq \overline{B}(\theta_1,\theta_2) = z^*(\underline{\theta}_1,\underline{\theta}_2)$.

(*ii*) At any (θ_1,θ_2)

$$\frac{\partial\overline{C}(\theta_i,\theta_{-i})}{\partial\theta_i} = -(r_i + v_i)\int_0^{y_i^*(\theta_i,\underline{\theta}_{-i})}\frac{\partial F_i(\xi_i,\theta_i)}{\partial\theta_i}d\xi_i \text{ and } \frac{\partial\overline{B}(\theta_i,\theta_{-i})}{\partial\theta_i} = -(r_i + v_i)\int_0^{y_i^*(\theta_i,\theta_{-i})}\frac{\partial F_i(\xi_i,\theta_i)}{\partial\theta_i}d\xi_i .$$

With assumption 3 we can conclude that at any (θ_1,θ_2), $\overline{C}(\theta_1,\theta_2)$ increases at a higher rate than $\overline{B}(\theta_1,\theta_2)$. With (*i*), this completes the proof.

Proof: Theorem 3.

By Theorem 1 and Lemma 1 we have,

$$C_1 + C_2 = \int_{\underline{\theta}_1}^{\bar{\theta}_1} \int_{\underline{\theta}_2}^{\bar{\theta}_2} [\Pi_1(y_1^*(\theta_1, \underline{\theta}_2), \theta_1) + \Pi_2(y_2^*(\underline{\theta}_1, \theta_2), \theta_2)] d\Phi(\theta_1) d\Phi(\theta_2)$$

$$B = \int_{\underline{\theta}_1}^{\bar{\theta}_1} \int_{\underline{\theta}_2}^{\bar{\theta}_2} [\Pi_1(y_1^*(\theta_1, \theta_2), \theta_1) + \Pi_2(y_2^*(\theta_1, \theta_2), \theta_2)] d\Phi(\theta_1) d\Phi(\theta_2).$$

By assumption 3, At any (θ_1, θ_2) the expression inside the integral in $C_1 + C_2$ is greater that the one inside B. Therefore, $C_1 + C_2 > B$.

References

Cachon, G.P., and Lariviere, M.A., 1999-1, "An equilibrium analysis of linear, proportional and uniform allocation of scarce capacity", *IIE Transactions*, 31, pp. 835-849.

Cachon, G.P., and Lariviere, M.A., 1999-2, "Capacity choice and allocation: strategic behavior and supply chain performance", *Management Science*, Vol. 45, No. 8, pp. 1091, 1108.

Celikbas, M., Shanthikumar., J.G. and Swaminathan, J.M., 1999, "Coordinating production quantities and demand forecasts through penalty schemes", *IIE Transactions*, 31, pp. 851-864.

Cramton, P., and Gibbons, R., and Klemperer, P., 1987, "Dissolving a partnership efficiently", *Econometrica*, Vol. 55, No. 3. pp. 615-632.

Fudenberg, T., Tirole, J., 1991, Game Theory, MIT Press, Cambridge, MA.

Groves, T., and Loeb, M., 1979, "Incentives in a divisionalized firm", *Management Science*, Vol. 25, No. 3.

Harris, M., Kriebel, C.H., and Raviv, A., 1982, "Asymmetric information, incentives and intrafirm resource allocation", *Management Science*, Vol. 25, No. 6.

Kouvelis, P., and Lariviere, M.A., 2000, "Decentralizing cross functional decisions: coordination through internal markets", *Management Science*, Vol. 46, No. 8, pp. 1049-1058.

Lariviere, M.A., 1998, "Supply Chain Contracting and Coordination with Stochastic Demand", *Quantitative Models for Supply Chain Management*, S. Tayur, M. Magazine and R. Gaeshan, (editors) Kluwer Academic Publishers, Norwell, MA.

Law A., and Kelton, W.D., 1992, *Simulation Modeling and Analysis*, New York, McGraw-Hill.

Makowski, L., and Mezzetti, C., 1993, "The possibility of efficient mechanisms for trading an invisible object", *Journal of Economic Theory*, 59, pp. 451-465.

Makowski, L., and Mezzetti, C., 1994, "Bayesian and weakly robust first best mechanisms: characterizations", *Journal of Economic Theory*, 64, pp. 500-519.

Mallik, S. and Harker, P. T. 1998, "Coordinating Supply Chains with Competition: Capacity Allocation in Semiconductor Manufacturing", *Working Paper, Fishman-Davidson Center, The Wharton School.*

McAfee, R.P., 1991, "Efficient Allocation with Continuous Quantities", *Journal of Economic Theory,* 53, pp.51-74.

Myerson, R. and Satterthwaite, M., 1983, "Efficient mechanisms for bilateral trading", *J. Econ. Theory*, 28, pp. 265-281.

Porteus, E.L., and Whang S., 1991, "On manufacturing/marketing incentives", *Management Science*, Vol. 37, No. 9.

Trench, W. F., 1978, Advanced Calculus, Harper & Row Publishers: New York, Hagerstown, San Francisco, London.

Williams, S.R., 1999, "A Characterization of Efficient, Bayesian Incentive compatible mechanisms", *Economic Theory*, 14, pp. 155-180.