Fabless-Foundry Partnership: Models and Analysis of Coordination Issues

Arun Chatterjee, Dadi Gudmundsson, Raman K. Nurani, Sridhar Seshadri, and J. George Shanthikumar

***Abstract*—The fabless-foundry partnership for integrated circuit (IC) manufacturing business is expected to grow from 12% in 1995 to approximately 17% (i.e., $45B) of the total IC market in 2000. The growth of this market will be even more significant for subquarter micron technologies—whose growth is driven by the multimedia industry. The customer base will extend beyond traditional fabless IC companies into vertically integrated IC manufacturers and system vendors. Given the rate of growth and the high technology profile of products, substantial investments in capital, technology, and skilled workforce have to be dedicated and managed effectively for ensuring a successful partnership. In this paper, we outline the potential coordination problems that may arise in such partnerships, and propose a framework for analyzing issues related to yield information sharing and yield improvement. Our analysis indicates that fabless-foundry contracts that are based on a fixed number of good dies, and better yield information are more profitable.**

***Index Terms*—Benchmarking, foundries, game theory, pricing, yield management.**

I. Introduction

The fabless-foundry partnership for integrated circuit (IC) manufacturing business is expected to grow from 12% in 1995 to approximately 17% (i.e., $45B) of the total IC market in 2000 [13]. The 40% annual growth rate forecast for fabless companies [1] constitutes a bright note in an other wise dampened outlook for the semiconductor industry. The forecast, based on a survey by the Fabless Semiconductor Association (FSA), also states that demand could grow at nearly 200% per annum for 0.5-μm and smaller technologies. The FSA survey points out that the high growth rate has been accompanied by an increasing number of long-term fabless foundry partnerships and in times of shortage by the denial of foundry capacity to small fabless firms. An article by Dunn [2], that was written about the same time as the survey, describes in detail how fabless companies are meeting the challenges of producing high technology products, using proprietary processes but in foundries not owned by them. We argue that, given the rate of growth and the high technology profile of products, not only large investments in capital, technology, and training of workforce have to be undertaken; but also that the interactions between fabless firms and foundries have to be managed effectively for ensuring successful partnership outcomes.

Manuscript received September 15, 1998; revised November 13, 1998.

A. Chatterjee, D. Gudmundsson, and R. K. Nurani are with KLA-Tencor Corporation, San Jose, CA 95161 USA (e-mail: arun.chatterjee@kla-tencor.com; dadi.gudmundsson@kla-tencor.com; raman.nurani@kla-tencor.com).

S. Seshadri is with the Leonard N. Stern School of Business, New York University, New York, NY 10012 USA.

J. G. Shanthikumar is with Industrial Engineering and Operations Research Department, University of California, Berkeley, CA 94720 USA.

Publisher Item Identifier S 0894-6507(99)01216-6.

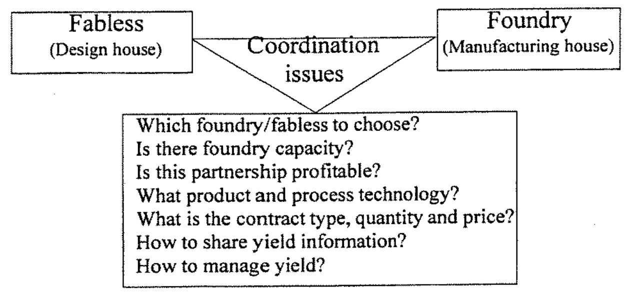


Fig. 1. Fabless-foundry partnership issues.

The outline of this paper is as follows. In Section II, we outline the potential problems of coordination that could arise in fabless-foundry partnerships. The issues are manifold but the literature on analytical models that specifically address the fabless-foundry coordination issues is limited. With that in mind, our goal is to create an analytical framework for studying the coordination issues, with special emphasis on yield management. We show that yield information sharing plays an important role in determining mutually beneficial contracts. In Section III, we present the model and results. In Section IV, we compare our results with the data from preliminary field observations. In Section V, we outline the opportunities for extending this work. For this paper, we assume that a fabless firm is a design house which does IC product design and has semiconductor process and technology know-how, but does not have the manufacturing facility. The foundry has wafer fabrication facility and manufacturing expertise but no design expertise (Fig. 1).

II. Coordination Issues

The decisions that the fabless firm and the foundry are required to coordinate upon are depicted in Fig. 1. Recent articles [1], [2], [5], [13] have identified several problems that can be traced to a lack of coordination. These problems include the following.

1) Precondition sometimes placed by foundries that the fabless firm must enter into equity participation and/or into long term arrangements.
2) High price premiums charged by foundries in times of capacity shortage.
3) Long supply lead times.

4) Proprietary process technology being exchanged for foundry capacity.
5) Fabless firms' concern about how to ramp up the yield when there are several foundry partners involved.
6) Reluctance on part of the foundries to give information to fabless firms with regard to yield and delivery performance.

We briefly discuss below the factors that not only make the coordination of decisions difficult but are also unique to this industry in this regard.

1) Product Market Risk: Fabless-foundry partnerships are meant for satisfying the demand for niche products [3]. Niche products by definition involve proprietary designs and could also involve proprietary process technologies. Leachman and Hodges [9] in their benchmarking study of wafer fabs conclude that, "While learning rates for defect density vary somewhat, rarely is a bad start overcome by subsequent rapid development. This observation underlines the critical importance of highly disciplined development activity, and of tight coupling between development and manufacturing." Recent research by Montverde [10] indicates that the need for maintaining technical dialog during the process of manufacturing niche products could act as a spur for vertical integration. However, the small volumes, short product life, and specialized skills required for designing such products act as deterrents to owning a foundry. The need for suitable contract mechanisms arises from these conflicting considerations, namely balancing the cost of maintaining a dialog between fabless firms and foundries with the risk associated with specializing in and marketing niche products. Models that capture the nature of uncertainties of this special business can reveal the extent of risk shared by the partners and whether the incentives are in proportion to the risk bearing capacities of the partners (Wilson [14]).

2) Matching Demand and Supply: It is not certain whether it is optimal for a foundry to match its capacity to the fabless demand. Given the lead time and the effort required to bring a high tech foundry on stream, it might be more profitable to allow temporary mismatches between demand and supply. Hines [5] points out that demand-supply mismatches have led to intense price fluctuations, for example, "... aggressive capacity expansion and intense competition have led to a general decline in prices for 0.35-mm foundry suppliers." Similarly, foundries are reluctant to invest in 1-mm capacity, which has led to shortages and high prices.

Related to the issue of demand-supply matching is the problem of determining the best product mix for foundries. Joseph [6] points out that the share of memory products in the mix of foundries has been small. This might be due to the fact that revenue per wafer is small for memory products and therefore not attractive to manufacture. On the other hand, small memory manufacturers might have no option but to remain fabless due to the high initial cost of setting up a foundry. Foundries have an incentive to manufacture memory products in order to benchmark their processes during the start up phase of the fab, because the process for making memory is repetitive. Foundries could derive benefit from the experience gained by manufacturing such products, thereby learn to identify process defects sooner.

3) Price, Due Dates, and Quantities: Agreements between the fabless firm and the foundry usually specify time and quantity to be the essence of the contract. Abundance of literature exists on how there can be a failure in such arrangements. For example, the profit margins to the fabless firm and the foundry could be different, which could lead to very different pricing and production quantity decisions for the fabless firm when compared to the foundry (Tirole [12]). Similarly, the incentives for meeting delivery due dates might not be identical for the partners.

4) Yield Management: It is a foregone conclusion that effective yield management is key to success in the semiconductor industry. To this end there should be significant effort on the part of both the fabless firm as well as the foundry to improve yield. The expertise and problem solving ability that are necessary to achieve high yield are still subjects of on-going research. For example, yield loss might constitute systematic and random components. The systematic yield loss could be related to both design and manufacturing whereas the random yield loss is usually related to manufacturing. Analysis of these yield components will help the fabless firm and the foundry in understanding their role toward achieving improved yield management. From the decision making perspective, the foundry has to factor in the uncertainty involved in producing high technology dies for costing such products. From the fabless firm's viewpoint, estimates of the yield and the yield ramp up curve are required to formulate a pricing strategy.

Our analysis (see Section III) indicates that it may not always be in the best interest of the foundry to provide the fabless firm with all potential yield information. We also find that given that there is uncertainty about yield, unless contracts are written correctly (to align the interests of the partners), there will be "over production" of dies. In other words, the quantity produced need not jointly maximize the expected profits of the partners.

5) Yield Improvement: The guaranteeing and the improvement of yield might be key success factors if a new design has to be brought to market quickly. Our analysis and preliminary field observations (see below) indicate that improved quality of information regarding the yield can bring about substantial benefits. This paper presents models for quantifying such benefits and studying the mechanisms required for coordinating the decisions of the fabless firm, the yield guarantor, and the foundry.

6) Scope for Yield Improvement: Our preliminary field observations indicate that there exists scope for improved yield management. In Fig. 2, we depict the probe yield for different DRAM products. The curves indicate that almost uniformly across different generations of products, the rate of increase in yield during the ramp-up phase has been 6–12% per month. In comparison, the yield improvement rates have been smaller for relatively low volume logic products as shown in Table I.

Our discussions with fabless firms reveal that not only are the improvement rates smaller during the ramp-up phase, but also that there is variation in the absolute level of achievable yield. We hypothesize that these yield differences are as much due to the misalignment of design, technology and equipment,

2

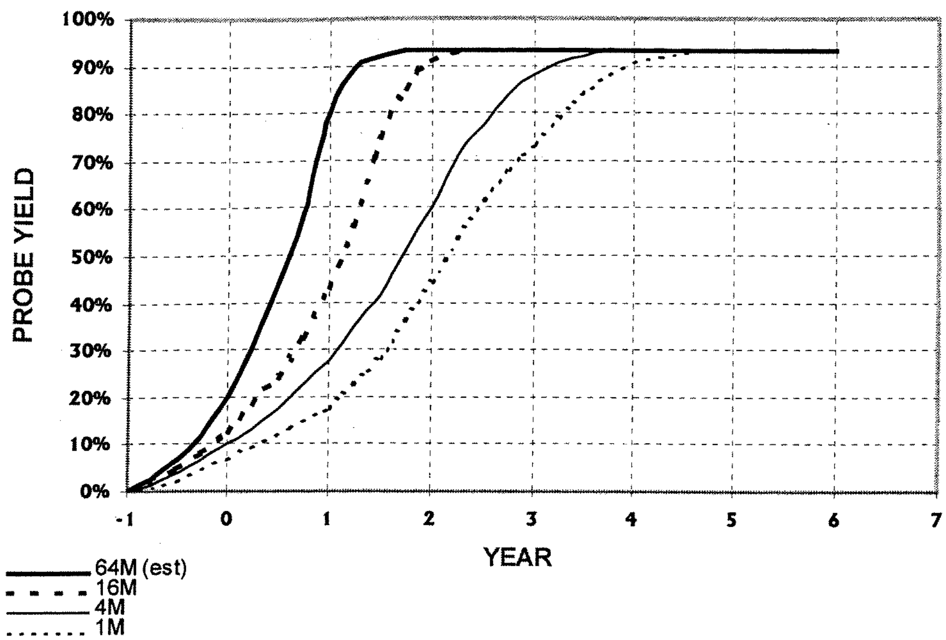


Fig. 2. Industry average probe yield curves (DRAM). Source: VLSI Research, Inc.

TABLE I
YIELD IMPROVEMENT RATES. SOURCE: (1) UNIVERSITY OF CALIFORNIA, BERKELEY. (2) KLA-TENCOR DATABASE AND [1]. (3) VLSI RESEARCH, INC., MIDTERM REPORT, 1996

Data Source	Product	Yield during ramp up		Time for ramp up Months	Improvement Rate per Month
		Begin	End		
1	CMOS Logic	0.37	0.76	54	0.01
	CMOS Logic	0.3	0.65	33	0.02
	CMOS Logic	0.18	0.63	33	0.04
	CMOS Logic	0.32	0.7	45	0.03
2	CMOS Logic	0.44	0.77	15	0.04
3	1M DRAM	0.08	0.8	40	0.06
	4M DRAM	0.1	0.8	32	0.07
	16M DRAM	0.15	0.8	18	0.1
	64M DRAM	0.2	0.8	12	0.12

as due to poor yield management practices (see Section IV).

7) Rating Vendor Performance: We are faced with the possibility that given the nature of uncertainties, fabless firms (unless they are well established) may have difficulty in convincing their customers about delivery reliability and product quality. To alleviate such problems, data regarding the range of contracts entered into, the contract performance, and the prospects for improved joint decision making could be made available for buyers in this industry. A pioneering effort in this direction was carried out by Leachman and Hodges [9] for wafer fabrication. In our preliminary field study we observed that fabless firms were enthusiastic about obtaining benchmark studies of foundries. We believe that such studies help in avoiding incorrect choice of foundry partners, and that yield management is currently viewed as being secondary when it comes to benchmarking.

As described above, there are several issues to consider when designing a fabless-foundry contract. Modeling all these issues simultaneously is a complex problem. Instead, in the next section, we model and analyze the role of yield information sharing for determining the order quantity and the pricing of a fabless-foundry contract. The aim of our analysis is to provide insights and to establish the need for further research in this direction.

III. MODELS AND ANALYSIS

In this section we develop simple models for understanding the role played by uncertain yield in determining the contract price and quantity.

A. Pricing

In this subsection we focus on the price charged by the foundry. We assume the following.

1) Fabless firm is a price taker, and that it can sell every good die at a price p.
2) We consider a single period model, in which the fabless firm contracts for production, obtains product, pays the foundry, and sells only the good output (dies).
3) The foundry can produce a die at a unit cost of c. This cost could differ depending on the utilization of the foundry.
4) The foundry agrees to a contract and executes the order.
5) If the foundry produces Q dies, then the number of good dies will be given by the random quantity, $Y_T(Q)$, where the subscript T stands for the true yield. The dies that are not good have zero salvage value.
6) Let $E[\cdot]$ denote the expected value. We shall assume that $E[Y_T(Q)]$ as well as the estimates of the expected number of good die are increasing and concave in Q.
7) We shall assume that the true expected yield of good dies is known to the foundry. The fabless firm need not be aware of the true expected yield.

3

8) Cooperative Outcome: We shall work in the cooperative two person game theoretic framework. This assumption is valid as the contract between the two parties can be enforced if necessary (see Kohli and Park [7], Kreps [8]).
9) We shall assume that the optimal production quantity is always finite.

1) Price Model #1—Fabless Firm Buys Wafers and Tests Dies: In this model the fabless firm contracts for a quantity of Q dies, pays a price of c_T per die to the foundry, locates the good dies and sells what it is good. We shall use the Nash fixed threat bargaining model as well as the Raiffa–Kalai–Smorodinsky solution to the bargaining problem (Friedman [4]).

Proposition 1 (Complete Information on Yield): Given that a quantity Q is contracted for production, and if the true value of expected yield is known to both parties, then:

1) the fabless firm will pay a price of $c_T = pE[Y_T(Q)]/2Q + c/2$ per die. The profit functions of both parties will be $pE[Y_T(Q)]/2 - cQ/2$;
2) the foundry does not have an incentive to put in its best effort in attaining or exceeding the promised die-yield (defect density).

Proof:

1) The Nash solution for the contract price is determined by $\max_{c_T} \ ((pE[Y_T(Q)] - c_TQ)(c_TQ - cQ))$. The Raiffa–Kalai–Smorodinsky (RKS) solution to the bargaining problem is given by connecting the threat point, i.e., the point if no contract ensues, (0, 0) to the ideal point at which both parties get the maximum payoff, i.e., $((pE[Y_T(Q)]-cQ), (pE[Y_T(Q)]-cQ))$ and setting the "solution" to be the point of intersection of this line with the Pareto frontier. The Pareto frontier is simply the line joining the points $(0, (pE[Y_T(Q)]-cQ))$ and $((pE[Y_T(Q)] - cQ), 0)$. We leave it to the reader to verify that both the solutions will split the difference between the maximum and minimum price of output, i.e., c and $pE[Y_T(Q)]/Q$.
2) The foundry gets paid as per the number of dies produced, and thus has no incentive to exercise due diligence in manufacturing the dies (assuming there is no reputation effect). ■

Proposition 2 (One Sided Information on Yield): Assume only the foundry has information about the true yield, and that the fabless firm accepts the value of yield provided by the foundry (subject if necessary to some ceiling on expected yield which the fabless firm believes in). Then the foundry will attempt to provide an overestimate of the expected yield.

Proof: The expected yield of good dies is assumed to be increasing and concave in Q. If the foundry overestimates the expected yield say by a constant factor $f > 1$, then it stands to benefit in two ways.

1) By equating marginal revenue and marginal cost the quantity contracted will not be smaller, say $Q_{\text{LIE}} \geq Q_{\text{TRUTH}}$.
2) As per Proposition 1, the revenue to the foundry will be given by $(pfE[Y_T(Q_{\text{LIE}})] - cQ_{\text{LIE}})/2$. But,

$$(pfE[Y_T(Q_{\text{LIE}})] - cQ_{\text{LIE}})/2 \\ \leq ((pE[Y_T(Q_{\text{TRUTH}})] - cQ_{\text{TRUTH}})/2$$

where the inequality follows from the assumption of concavity and because Q_{LIE} is the optimal value for Q, when the yield is increased by a constant factor f. Therefore, the foundry makes as much or greater profit on the first Q_{TRUTH} units by suggesting a higher yield. On the units produced beyond Q_{TRUTH} (i.e., Q_{TRUTH} to Q_{LIE}), the foundry does not lose money due to the uniqueness of the optimum and the assumption of concavity.

Remark: In contracts as described in Model #1, the foundry attains the greater share of profits, specifies a higher yield, and delivers a lower yield. This hypothesis is being tested using field data.

2) Price Model #2—Foundry Carries out the Testing: In this model, the fabless firm pays only for the good dies at the contract price, c_T. The foundry is given the (fixed) price per good die, c_T, and the foundry decides upon the quantity to be produced. The foundry (costlessly) screens the dies produced, identifies the good dies, and then delivers the dies to the fabless firm.

Proposition 3 (Complete Information on Yield): Given that a quantity Q is contracted for production, and if the true value of expected yield is known to both parties, then

1) the fabless firm will pay a price of $c_T = (p/2 + cQ/(2E[Y_T(Q)])$ per good die. The profit functions of both parties will be $(pE[Y_T(Q)]/2 - cQ/2)$;
2) unlike in Proposition 1, the foundry has an incentive to put in effort for attaining the promised yield.

Proof: Similar to Proposition 1.

Proposition 4 (One Sided Information on Yield): Assume only the foundry has information about the true yield, and the fabless firm accepts the value of yield provided by the foundry (subject if necessary to some floor on the expected yield which the fabless firm believes in). Then the foundry will attempt to provide an underestimate of the expected yield.

Proof: The quantity decision is now in the hands of the foundry. (In the previous model, the fabless firm decided the quantity to purchase.) By underestimating the yield, the foundry stands to gain by obtaining a higher price (the unit price equals $p/2 + cQ/2E[Y_T(Q)]$). Once this "higher" price has been agreed upon, the foundry optimizes with respect to the quantity to be produced. In this case, the quantity produced will be again larger (!) than the optimum level. As argued earlier, the foundry stands to gain on output until Q_{TRUTH}, and not to lose on the extra quantity produced because of the uniqueness of the optimum and the assumption of concavity. ?

Remarks:

1) We can set up similar hypotheses (as were given after proposition 2) for field testing in the case of Model #2.
2) It is indeed intriguing that there will be overproduction in both models whenever only the foundry knows the true parameters of the yield.

4

Proposition 5: The following sequence of statements by the fabless firm will elicit the truth about the true yield.

1) Give me an estimate of your yield.
2) I shall then decide whether to follow Models 1 or 2.

Remark: Propositions 1–4 continue to hold when the demand is stochastic and the yield is increasing and concave in a stochastic sense (see [11, ch. 6]).

B. Determining the Order Quantity

In this subsection we allow demand, denoted as D, to be stochastic. We assume that the yield and demand are independent random variables. We assume in addition that the fabless firm can sell the minimum of the quantity of good dies and the (random) demand. Any unsold quantity is assumed to be scrapped (at zero salvage value) and unmet demand is considered to be lost. There is no loss of goodwill due to lost sales. We shall assume that the expected marginal yield from producing an extra die 1) does not depend on the yield of dies already produced and that 2) the expected marginal yield is either the same or larger with increasing Q. Let $my(Q)$ be the marginal yield from producing the Qth die. Then

$$my(Q) = E[Y_T(Q) - Y_T(Q-1)|Y_T(Q-1)]$$

and

$$E[Y_T(Q)] = my(1) + my(2) + \cdots + my(Q).$$

Let $Y_T^{-1}(Q_g)$ stand for the inverse yield function, i.e., the number of dies that have to be produced in order to obtain Q_g good dies. Then

$$Y_T^{-1}(Q_g) = inf_Q\{Y_T(Q) = Q_g\}.$$

Proposition 6: The fabless firm and the foundry stand to gain by contracting for a fixed number of good dies, Q_g, when compared to contracting for a fixed quantity of dies, Q.

Proof: When the contract is for a fixed quantity of production, i.e., Q, then the expected total profit $\pi_{\mathrm{PROD}}(Q)$ is given by (from Propositions 1–4):

$$\pi_{\mathrm{PROD}}(Q) = pE[\min\{Y_T(Q), D\}] - cQ. \quad (1)$$

If the contract were for a fixed number of *good* dies, Q_g, instead, then the total expected profit will be

$$\pi_{\mathrm{GOOD}}(Q_g) = pE[\min\{Q_g, D\}] - cE[Y_T^{-1}(Q_g)]. \quad (2)$$

Let $z = E[Y_T(Q)]$. Then by concavity of the min $\{\}$ function,

$$E[\min\{Y_T(Q), D\}] \leq (E[\min\{E[Y_T(Q)], D\}]. \quad (3)$$

Define $\lceil z \rceil$ and $\lfloor z \rfloor$ to be the smallest and the largest integers greater than or equal to and smaller than or equal to z. Let $I\{A\}$ stand for the indicator function of the event A. Notice that when $Y_T^{-1}[\lceil z \rceil]$ dies have been made a total of $\lceil z \rceil$ good dies have been produced. Consider the case when $Y_T^{-1}[\lceil z \rceil] \leq Q$. In this case, because the expected marginal yield is independent of the yield of previous dies and because by our assumption that $my(Q)$ is nondecreasing in Q

$$\begin{aligned} E[(\lceil z \rceil + (Q - Y_T^{-1}[\lceil z \rceil])my(Q))I\{Y_T^{-1}[\lceil z \rceil] \leq Q\}] \\ \geq E[Y_T(Q)I\{Y_T^{-1}[\lceil z \rceil] \leq Q\}]. \end{aligned}$$

In other words, instead of stopping when we have produced $\lceil z \rceil$ good dies we proceed to produce the extra dies $(Q - Y_T^{-1}[\lceil z \rceil])$. These extra dies give an expected marginal yield that is not greater than $my(Q)$. Thus, we end up producing more than the yield from Q dies whenever $Y_T^{-1}[\lceil z \rceil] \leq Q$. Because $z = E[Y_T(Q)]$, we may add $E[Y_T(Q)I\{Y_T^{-1}[\lceil z \rceil] \geq Q\}]$ to both sides of the above inequality to obtain

$$\begin{aligned} E[(\lceil z \rceil + (Q - Y_T^{-1}[\lceil z \rceil])my(Q))I\{Y_T^{-1}[\lceil z \rceil] \leq Q\}] \\ + E[Y_T(Q)I\{Y_T^{-1}[\lceil z \rceil] > Q\}] \geq z. \end{aligned}$$

This inequality upon taking the term involving $\lceil z \rceil$ from the left hand side of the inequality to the right-hand side of the inequality and upon simplification yields

$$\begin{aligned} E[(Q - Y_T^{-1}[\lceil z \rceil])I\{Y_T^{-1}[\lceil z \rceil] \leq Q\}]my(Q) \\ + E[Y_T(Q)I\{Y_T^{-1}[\lceil z \rceil] > Q\}] \\ \geq \lceil z \rceil \Pr\{Y_T^{-1}[\lceil z \rceil] > Q\} - (\lceil z \rceil - z). \end{aligned} \quad (4)$$

However, once again because the expected marginal yield is independent of the yield of previous dies and because by our assumption that it is nondecreasing in Q

$$\begin{aligned} E[Y_T(Q)I\{Y_T^{-1}[\lceil z \rceil] > Q\}] \\ + E[(Y_T^{-1}[\lceil z \rceil] - Q)I\{Y_T^{-1}[\lceil z \rceil] > Q\}]my(Q) \\ \leq \lceil z \rceil \Pr\{Y_T^{-1}[\lceil z \rceil] > Q\}. \end{aligned} \quad (5)$$

In other words, when $Y_T^{-1}[\lceil z \rceil] > Q$, we produce beyond Q until we obtain $\lceil z \rceil$ good dies. The additional dies produced, namely, $(Y_T^{-1}[\lceil z \rceil] - Q)$, have an expected marginal yield that is greater than $my(Q)$, resulting in (5).

Subtracting (5) from (4) we obtain

$$\begin{aligned} E[(Q - Y_T^{-1}[\lceil z \rceil])I\{Y_T^{-1}[\lceil z \rceil] \leq Q\}]my(Q) \\ - E[(Y_T^{-1}[\lceil z \rceil] - Q)I\{Y_T^{-1}[\lceil z \rceil] > Q\}]my(Q) \\ \geq -(\lceil z \rceil - z) \end{aligned}$$

or

$$Q \geq E[Y_T^{-1}[\lceil z \rceil]] - (\lceil z \rceil - z)/my(Q). \quad (6a)$$

A similar argument reveals that

$$Q \geq E[Y_T^{-1}[\lfloor z \rfloor]] + (z - \lfloor z \rfloor)/my(Q). \quad (6b)$$

Combining the inequality (3) and equation (1) we get

$$\begin{aligned} \pi_{\mathrm{PROD}}(Q) &= pE[\min\{Y_T(Q), D\}] - cQ \\ &\leq pE[\min\{z, D\}] - cQ. \end{aligned}$$

We now claim that $pE[\min\{z, D\}] - cQ$ is at least equal to or smaller than one of

$$pE[\min\{\lceil z \rceil, D\}] - cE[Y_T^{-1}[\lceil z \rceil]] \quad \text{and}$$
$$pE[\min\{\lfloor z \rfloor, D\}] - cE[Y_T^{-1}[\lfloor z \rfloor]].$$

Assume to the contrary that it is not. Then using (6a)

$$\begin{aligned} 0 &< pE[\min\{z, D\}] - cQ - (pE[\min\{\lceil z \rceil, D\}] \\ &\quad - cE[Y_T^{-1}[\lceil z \rceil]]) \\ &\leq -(\lceil z \rceil - z)p \Pr\{D > \lfloor z \rfloor\} + (\lceil z \rceil - z)c/my(Q) \\ &= -(\lceil z \rceil - z)(p \Pr\{D > \lfloor z \rfloor\} - c/my(Q)). \end{aligned}$$

5

Using (6b)

$$\begin{aligned} 0 < & pE[\min\{z, D\}] - cQ - (pE[\min\{\lfloor z \rfloor, D\}] \\ & - cE[Y_T^{-1}[\lfloor z \rfloor]]) \\ & \leq (z - \lfloor z \rfloor)p \Pr\{D > \lfloor z \rfloor\} - (z - \lfloor z \rfloor)c/my(Q) \\ & = (z - \lfloor z \rfloor)(p \Pr\{D > \lfloor z \rfloor\} - c/my(Q)). \end{aligned}$$

Notice that both $(\lceil z \rceil - z)$ and $(z - \lfloor z \rfloor)$ are nonnegative quantities. Therefore, both these inequalities *can not hold simultaneously*. Therefore producing either (or both) $\lfloor z \rfloor$ or $\lceil z \rceil$ good dies should yield equal or greater profit compared to producing Q dies. ■

Remarks:

1) The results of Proposition 5 need not hold when the demand is stochastic. This is because the foundry is aware of the fact that when the conditions of Proposition 6 hold a contract for good dies gives higher expected profit. The conditions of Proposition 6 are more likely to hold true for new products because the yield can be expected to improve as the foundry and the fabless firm learn to produce the new dies. These products might be the ones for which sharing information about the yield is very important. Therefore, for the reasons described in this paper and reasons such as the fabless firm's or the foundry's reluctance to share information about new products, the problem of yield management becomes more pronounced for such products. We are carrying out a field study to verify these predictions.
2) The proposition also reveals [see (2)] that the expected profit decreases with increased volatility of demand and yield.
3) An aspect that has not been captured by these models is the possibility that during the production of the dies the foundry experiences problems or finds the production to be easier than assessed earlier. In both cases, such experience could be used to revise the order quantity dynamically. We expect that dynamic reaction to the yield will be made easier if the contract were to be based on good dies. We might also expect that the marginal cost of production is more closely matched to the marginal revenue when the order quantity is changed dynamically.

C. Yield Prediction and Improvement

In this section we determine 1) the relationship between demand and the magnitude of the benefit obtained from improving yield, and 2) when additional information regarding yield is beneficial. We shall assume that the contract is for good dies.

Proposition 7: We are given two products (1 and 2) with the same cost of production and with the same yield functions. Assume that the yield can be improved such that the marginal expected cost of producing the nth good die becomes smaller for all values of n.

Case (i): The two products have the same mean demand but product 2 has a more volatile demand compared to product 1, i.e., its demand is larger in the usual convex order compared

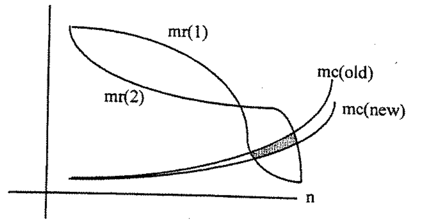


Fig. 3. Proof of case (ii).

to the demand for product 1—see Shaked and Shanthikumar [11] for the definition of convex ordering. Then improving the yield of the product whose present order quantity is higher gives greater benefit.

Case (ii): The demand for product 2 is larger than the demand for product 1 (in the usual stochastic order). In this case improving the yield of product 2 gives higher benefit.

Proof: Let p_n be the probability that the demand for dies equals n.

Case (i): Let the demand for the two products be denoted as D_i, $i = 1, 2$. The marginal expected revenue for product i from the $(n+1)$st die is given by

$$\begin{aligned} & p\left(\sum_{j=1}^{n+1} jp_j + (n+1)\Pr\{D_i > n+1\}\right) \\ & \quad - p\left(\sum_{j=1}^{n} jp_j + n\Pr\{D_i > n\}\right) \\ & = p\Pr\{D_i \geq n+1\}. \end{aligned}$$

By our assumption (see Shaked and Shanthikumar [11]), the graph of $\Pr\{D_1 \geq n+1\}$ crosses the graph of $\Pr\{D_2 \geq n+1\}$ once and from above. The proof is complete by referring to Fig. 3. In Fig. 3, $mr(1)$ and $mr(2)$ refer to the marginal revenue from the two products. The marginal costs before and after yield improvement are denoted by mc(old) and mc(new). We depict the case in which the original optimal order quantity is larger for product 2. In this case additional profit that is equal to the shaded area between the two marginal costs curves is obtained by undertaking the improvement of the yield of the second product. The proof when the order quantity is larger for the first product (instead of being larger for the second product) is similar. *Case (ii)*: The proof is similar to the proof of case (i).

We now determine sufficient conditions under which additional information provided by a third party regarding the yield will be beneficial. We assume that the marginal expected revenue from the nth good die is given by $mr(n)$. We assume that the contract is for good dies. We denote the true marginal expected cost of producing the nth die is denoted by $mc(n)$. The estimate of the marginal cost without the additional information provided by the third party is assumed to be given by

$$\widehat{mc}_1(n) = mc(n)(1 + \varepsilon(1)) \tag{7}$$

where the quantity $\varepsilon(1)$ is a random error term. We assume that the estimate of the marginal expected cost with the additional

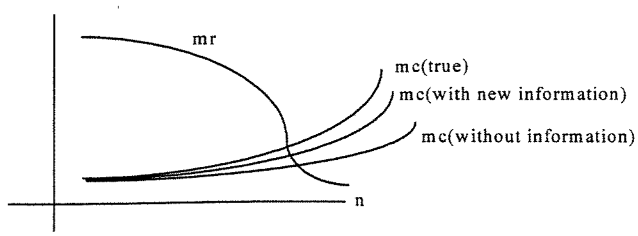


Fig. 4. Value of information.

information provided by the third party is given by

$$\widehat{mc}_2(n) = mc(n)(1 + \varepsilon(2)). \tag{8}$$

Proposition 8: The new information provided by the third party will give higher expected profit if either of the following conditions hold.

1) If the error in predicting the yield is smaller by a constant positive multiple, $\varepsilon(2) = f \cdot \varepsilon(1)$, $0 < f < 1$.
2) If the errors are independent of the quantity produced, their distribution is symmetric around zero, and the absolute value of $\varepsilon(2)$ is smaller in probability compared to the absolute value of $\varepsilon(1)$.

Proof: The proof is similar to the proof of case (i) of proposition 7. The main point about these conditions is that the marginal expected cost estimate is closer to the true marginal expected cost when the third party provides the additional information. This is depicted below in Fig. 4. ■

In the next proposition we analyze how a third party can be compensated for improving the yield. We assume that the third party is of sufficient reputation, that information given to the third party will be kept confidential, and that the fabless firm and the foundry share data regarding yield, costs, and profit margins. To keep the exposition simple we assume that the quantity produced can take real values. The result of Proposition 9 given below is easily modified to the case when the quantity can take only integer values. We assume that the third party exerts an effort, a, to improve the yield; where a can take values in the interval [0, 1). The cost of this effort is $e(a)$ to the third party. The function $e(a)$ is assumed to be increasing and convex in a. Without loss of generality we assume that the inverse yield function decreases linearly in a, i.e. the number of dies required to produce Q_g good dies is given by,

$$Q = Y_T^{-1}(Q_g)(1 - a) + aQ_g.$$

We assume that a is determined at the beginning of the yield improvement program by the third party and does not change dynamically over time. Assume that the foundry wishes to produce a quantity Q_g good dies. This quantity is arrived at after mutual consultations between the three parties. The main issue is whether the effort exerted by the third party will be optimal.

Proposition 9: If the third party is compensated an amount equal to $K + c(E[Y_T^{-1}(Q_g)] - Q)$, where K is a constant, then the effort exerted will be optimal.

Proof: The expected profit to the third party is given by

$$\begin{aligned} K &+ E[c(E[Y_T^{-1}(Q_g)] - Q)] - e(a) \\ &= K + cE[Y_T^{-1}(Q_g)] - c((1 - a)E[Y_T^{-1}(Q_g)] \\ &\quad + aQ_g) - e(a) \\ &= K + caE[Y_T^{-1}(Q_g)] - caQ_g - e(a) \end{aligned} \tag{9}$$

where we have substituted $Y_T^{-1}(Q_g)(1 - a) + aQ_g$ for Q to obtain the first equality.

It is clear that the third party will attempt to maximize profit and thus set $cE[Y_T^{-1}(Q_g)] - cQ_g - e'(a) = 0$. The total expected profit for all three parties combined for a given value of Q_g is given by [from (2)]

$$\begin{aligned} \pi_{\mathrm{GOOD}}(Q_g) = {} & pE[\min\{Q_g, D\}] - c((1 - a)E[Y_T^{-1}(Q_g)] \\ & + aQ_g) - e(a). \end{aligned} \tag{10}$$

Notice that by our assumptions that the expected profit is a concave function of the effort. Differentiating this profit with respect to a gives $cE[Y_T^{-1}(Q_g)] - cQ_g - e'(a) = 0$. ■

Remark: The third party will enter into this contract only if it is profitable. On the other hand, the foundry and the fabless firm might be unwilling to give away all the savings from cost reduction to the third party. Therefore, the value of K might require negotiation. There will always be a feasible contract if $[cE[Y_T^{-1}(Q_g)] - cQ_g - e'(a)]_{a=0} > 0$, i.e. if the marginal cost of the effort required to increase the current yield is *less* than the total cost of producing Q dies. This suggests that in a majority of cases in which there is scope for yield improvement, the use of a third party to benchmark and improve yield will be found to be not only feasible but also beneficial.

IV. Discussion

We interviewed ten fabless companies with regard to the propositions described in the previous section. Their responses tend to agree with many of the propositions stated in Section III. For instance, fabless companies that contract on "wafers" (all dies) said that "Yield obtained from the foundries was lower than yield promised," which agrees with Proposition 2. Fabless companies that mainly contract on "good dies" responded by saying "Yield obtained from the foundries was higher than yield promised," in agreement with Proposition 4. During our interviews and a preliminary survey we also observed that some fabless-foundry partnerships are moving away from contracts based on wafers to contracts based on good dies. This agrees with Proposition 6.

We observed that the yield risk is higher when a fabless firm designs its product using leading edge technology. We hypothesize that the alignment of Design know-how, Equipment parameters and process Technology parameters is very critical to accelerate the yield learning process. As mentioned in Section II, the alignment between Design, Equipment, and Technology (DET) may be difficult to attain in the case of fabless firms. Fig. 5 lists the steps involved in DET. During the preliminary survey we encountered evidence that it is possible to obtain the DET alignment quite rapidly when a third party provides the expertise and benchmark information.

7

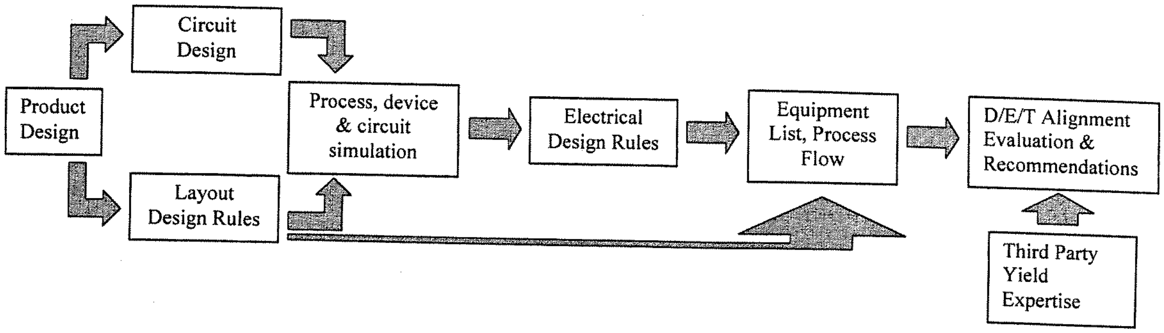


Fig. 5. DET alignment verification flowchart.

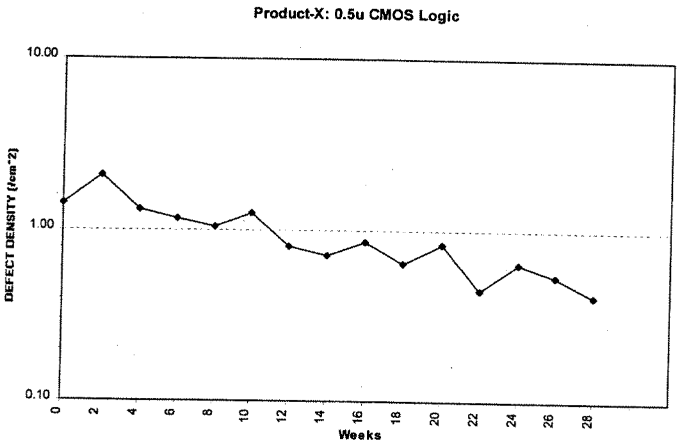


Fig. 6. Example of yield improvement attained when third party consults on DET alignment.

Fig. 6 shows such an example for a fabless firm's product. In this example, the yield improvement rate with expert intervention was over 4% per month (compare with the learning rates Table I). In the first phase (time = 1–8), the processing parameters were optimized to align Technology and Equipment. In the second phase (time = 9–15), the layout design was optimized to align Technology and Design.

Every fabless firm we spoke to agreed on the importance of having benchmark data. This agrees with Proposition 9 and the remark following the proposition. At the firm level the benefits from benchmarking include: being able to make informed choice of foundry partner, obtaining alignment among Design, Equipment, and Technology, obtaining accurate yield information, being able to determine the optimal contracting quantity, able to accurately assess the scope for yield improvement, and other benefits including being able to make better pricing decisions and determine reliable delivery schedules. The benefits are not just restricted to the firm level but could very well extend to the entire industry due to improved matching of capacity to demand, capability studies becoming available to prospective customers and investors, and due to improvement in yield and productivity.

In summary, we noticed that the nature of products, technological complexity, constantly changing product volume mix, and the highly competitive environment in the fabless-foundry industry necessitate a sophisticated yield prediction and management system. Given the encouraging results from our models and preliminary field study, we plan to continue our research with a more exhaustive field study.

V. Conclusion and Extensions

Given the substantial share fabless-foundry manufacturing business is of the total IC market, we are motivated to identify and study the coordination issues in such ventures. Game theoretic models were used to study pricing, order quantity decisions, and yield prediction and improvement in fabless-foundry partnerships We presented two price models, model 1 assumed a contract that is based on the fabless firm purchasing wafers, while model 2 assumed a contract based on good dies only. Propositions 1–4 showed how both models, combined with either full or one-sided knowledge of true foundry yield, led to nonoptimal results from a cooperative standpoint. Since the two price models deviated in opposite directions, proposition five states that the fabless firm should not indicate to the foundry which price model will be used, forcing the foundry to reveal the true yield. Proposition 6 asserts that it is better to contract for "good dies," as opposed to wafers. It also points out that stochastic demand can impact joint decision making since the foundry knows that price model 2 is better.

The remaining propositions addressed yield prediction and improvement. In Proposition 7, two almost identical products are differentiated using their demand distributions so that decision makers can target yield improvement effort to the product that gives higher benefit. Proposition 8 provides sufficient conditions under which yield information provided by a third party will be beneficial. In proposition 9 a compensation scheme for the services of a third party is presented and shown to be optimal in the cooperative framework.

The models and analysis presented in this paper provide interesting insights and provide a need for further research in this direction. Some of the potential future work include, extending the models to a multiple period and multiple product setting, analyzing the scenarios with multiple fabless companies and foundries, and conducting extensive field survey. Also, we plan to include the possibility of renegotiation/bargaining in the contracts based on yield information that becomes available over time.

Acknowledgment

The authors thank KLA-Tencor management for providing the opportunity to conduct this research. They are grateful to P. Franklin, Senior VP of Operations, S3 Incorporated, S. Reddy, Executive VP of Operations, MMC Networks, and

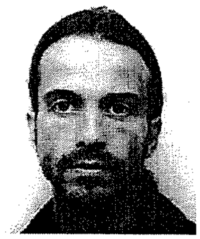
J. McColum, Founder, Actel Corporation, and several other anonymous fabless firms and foundries for providing helpful insights. We also thank the associate editor and the referees for their suggestions on the earlier version of this paper.

References

[1] "Fabless growth to beat industry," *Electron. News Mag.*, Apr. 22, 1996.
[2] D. Dunn, "How fabless forces battle supply-demand shifts," *Electron. Buyers News*, Apr. 15, 1996.
[3] "Multimedia fuels hot fabless growth," *Electron. News*, Apr. 29, 1996.
[4] J. W. Friedman, *Game Theory with Applications to Economics.* New York: Oxford Univ. Press, 1986.
[5] J. F. Hines, "A review of foundry wafer prices," *Fabless Forum*, vol. 4, no. 4, Dec. 1997, Fabless Semiconductor Association.
[6] J. Joseph, "Does it make sense for a memory supplier to be Fabless," *Fabless Forum*, vol. 4, no. 4, Dec. 1997, published by the Fabless Semiconductor Association.
[7] R. Kohli and H. Park, "A cooperative game theory model of quantity discounts," *Manag. Sci.*, vol. 35, no. 6, pp. 693–707, 1989.
[8] D. M. Kreps, *A Course in Microeconomic Theory.* Princeton, NJ: Princeton Univ. Press, 1990.
[9] R. C. Leachman and D. A. Hodges, "Benchmarking semiconductor manufacturing," *IEEE Trans. Semiconduct. Manufact.*, vol. 9, pp. 158–169, May 1996.
[10] K. Montverde, "Technical dialog as an incentive for vertical integration in the semiconductor industry," *Manag. Sci.*, vol. 41, pp. 1624–1638, 1995.
[11] M. Shaked and J. G. Shanthikumar, *Stochastic Orders and Their Applications.* San Diego, CA: Academic, 1994.
[12] J. Tirole, *The Theory of Industrial Organization.* Cambridge, MA: MIT Press, 1990.
[13] F. C. Tseng, "Foundry technologies," in *Proc. IEDM Conf.*, San Francisco, CA, 1996.
[14] R. Wilson, "The theory of syndicates," *Econometrica*, vol. 36, pp. 119–132, 1968.

Arun Chatterjee received the M.S. degree in materials science from the University of Washington, Seattle.

From 1975 to 1996, he was with Motorola, Signetics, Fairchild, Advanced Micro Devices, Data General and Cirrus Logic in developing and managing semiconductor process technologies (device development and process integration) for logic/microprocessor products. At Cirrus Logic, he was the Senior Manager of strategic wafer fab operations and Technical Coordinator to oversee the joint venture manufacturing partnership between IBM and Cirrus Logic. In 1996, he joined KLA-Tencor, San Jose, CA, and is currently the Director of integration solutions development for yield management consulting organization. His current interests include emerging technologies, interconnect solutions, technology integration and transfer, and fabless-foundry partnership development.



Dadi Gudmundsson received the B.S. degree in industrial engineering from the University of Arkansas, Fayetteville, in 1995, and the M.S. degree in industrial engineering and operations research from the University of California, Berkeley, in 1997. He is currently pursuing the Ph.D. degree at UC-Berkeley.

He currently works as a Researcher for KLA-Tencor, San Jose, CA. His current work focuses on sample planning and process control for random defect and metrology inspection in the semiconductor industry. Primary interests include stochastic modeling and simulation of industrial systems. While pursuing the M.S. degree, he worked with the Alpha Laboratory, where he contributed to IEEE's ICRA and IROS conferences on robotics, and worked as a Consultant to Hewlett-Packard and Adept.

Raman K. Nurani, for a photograph and biography, see this issue, p. 2.

Sridhar Seshadri, for a photograph and biography, see this issue, p. 2.

J. George Shanthikumar received the B.S. degree in mechanical engineering from the University of Sri Lanka in 1972, and the M.A.Sc. and Ph.D. degrees in industrial engineering from the University of Toronto, Toronto, Ont., Canada, in 1977 and 1979, respectively.

He is Professor of industrial engineering and operations research, College of Engineering and manufacturing and information Technology, Walter A. Haas School of Business, University of California, Berkeley. His research interests are in production and service systems modeling and analysis, queueing theory, reliability, scheduling, stochastic processes, simulation and supply chain management. He is author or coauthor of over 250 technical papers on these topics. He is coauthor, with J. A. Buzacott, of *Stochastic Models of Manufacturing Systems*, and a coauthor, with M. Shaked, of *Stochastic Orders and Their Applications*. He is or has been a member of the editorial boards of the *IIE Transactions on Design Manufacturing,* the *International Journal of Flexible Manufacturing Systems*, the *Journal of Discrete Event DYNAMIC Systems*, the *Journal of Production and Operations Management, Operations Research, Operations Research Letters, OPSEARCH, Probability in the Engineering and Informations Sciences*, and *Queueing Systems: Theory and Applications*.

Dr. Shanthikumar received the E. O. E. Pereira Gold Medal as outstanding student graduating from the College of Engineering, University of Sri Lanka. He was granted the Canadian Commonwealth Scholarship from 1975 to 1979 for his M.A.Sc. and Ph.D. studies at the University of Toronto.