Leveraging the Learning Process in Manufacturing

by

Timothy H. Ingle

B.S., Civil Engineering, University of Tennessee

Submitted to the Sloan School of Management and the
Department of Mechanical Engineering in Partial Fulfillment of the Requirements for the Degrees of

Master of Science in Management

and

Master of Science in Mechanical Engineering

at the
Massachusetts Institute of Technology
June 1997

©1997 Massachusetts Institute of Technology
All rights reserved

Signature of Author

Sloan School of Management
May 16, 1997

Certified by

J. M. Utterback
Sloan School of Management
Thesis Supervisor

Certified by

A. C. Thornton
Department of Mechanical Engineering
Thesis Supervisor

Accepted by

Lawrence S. Abein
Director of Master’s Program, Sloan School of Management

Accepted by

Ain N. Sonin
Committee on Graduate Students, Mechanical Engineering

JUL 01 1997
ARCHIVES
Massachusetts Institute of Technology

Abstract

Leveraging the Learning Process in Manufacturing

by

Timothy H. Ingle

Submitted to the Sloan School of Management and the Department of Mechanical Engineering on May 16, 1997 in partial fulfillment of the requirements for the Degrees of Master of Science in Management and Master of Science in Mechanical Engineering

Yield performance of a semiconductor fab is a major factor in its competitiveness. One component of overall yield, die yield, is studied here in detail. This thesis proposes that die yield improvement over time is the result of a carefully devised process of solving problems and instituting solutions. The process is referred to generally as a learning process. The learning process of a fab producing memory devices is analyzed and specific recommendations are made.

The learning process in semiconductor manufacturing is described in detail. This description includes both learning technologies and how they are integrated by the fab organization. Data from benchmarking efforts are presented to compare the subject fab with its competitors within the firm and across the semiconductor industry. Next, a simulation model is proposed to evaluate the impact of proposed changes to the learning process on the fab's die yield. Results of the simulation are presented.

Included is a discussion of the relationship between learning and competitiveness. If competitiveness is measured in terms of the trajectory of performance per unit cost over a period of time, then the learning process will play an important role in determining competitiveness. While the most competitive firms in the semiconductor industry today are vertically integrated, dis-integrated alliances of firms employing novel process architectures are beginning to emerge. A survey of the literature is presented to analyze this threat.

The thesis concludes with recommendations for the fab that was the subject of this study. Benchmarking data and results simulation model will be integrated to support these recommendations. Also, potential directions for evolution of learning systems in semiconductor manufacturing will be discussed.

Thesis Supervisors: James M. Utterback
Professor of Engineering and Management

Anna C. Thornton
Assistant Professor of Mechanical Engineering
Acknowledgments

I would like to gratefully acknowledge Anna Thornton and Jim Utterback. Without their guidance and patience, this document would be much less valuable. My gratitude is also due to Randy Williams, whose encouragement and challenging feedback led me to pursue this work. I would like to thank my family and friends who encouraged me throughout the past two years. Finally, I would like to thank the faculty, sponsors, and fellows of the Leaders for Manufacturing Program for their support of this work and for the fantastic personal growth the program has offered me.
# Table of Contents

Abstract 3
Table of Contents 7
Table of Figures 9

## Chapter One — Introduction 11

- Value of the learning curve 12
- Learning by doing 14
  - Types of learning 15
- Models of learning processes 15
- Learning in semiconductor manufacturing 16

## Chapter Two — Learning in Semiconductor Manufacturing 19

- Semiconductor manufacturing 20
- The learning framework 23
  - Collecting Data 24
  - Converting data into information 24
  - Using information to direct the acquisition of process knowledge 25
- Improving the process 25
- Benchmarking of Die Yield Performance in Semiconductor Manufacturing 27
- Conclusion 28

## Chapter Three — Modeling the Learning Process 29

- Scope of the model 29
- Structure 30
- Modeling in-line data collection 31
- Applying the yield ramp model 36
  - Implementation of in-line inspection system 37

---

7
Production process cycle time reduction 37
New manufacturing tool 38
Reducing time required to implement improvements. 38

Observations 38
Process cycle time reduction 38
Using in-line monitors 39
Introducing new manufacturing tools 40
Reducing time required to implement changes 41

Conclusions 42

Chapter Four — Manufacturing Learning and Competitiveness 45

The semiconductor industry 45
Semiconductor technology supply chain 46
Growth of the dis-integrated supply chain 47
Evolution of the semiconductor fabrication process 49
Research into supply chain dis-integration and competitiveness 50
Implications on learning and competitiveness 52

Conclusions 54

Chapter Five — Summary and Conclusions 55

Recommendations for the memory fab 55
Evolving the architecture of the learning process 56
Extending this analysis to other industries 58

References 59
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Improvement of Process Performance in Semiconductor Manufacturing</td>
<td>12</td>
</tr>
<tr>
<td>1.2</td>
<td>Relationship between fab capacity, line yield, and die yield</td>
<td>13</td>
</tr>
<tr>
<td>2.1</td>
<td>Example of yield ramps (Stapper, 1995)</td>
<td>20</td>
</tr>
<tr>
<td>2.2</td>
<td>Semiconductor Wafer Fabrication Process</td>
<td>21</td>
</tr>
<tr>
<td>2.3</td>
<td>Example of Organizational Interaction in the Learning Process (Weber 1996)</td>
<td>22</td>
</tr>
<tr>
<td>2.4</td>
<td>Examples of Learning Loops</td>
<td>23</td>
</tr>
<tr>
<td>3.1</td>
<td>Model Process Flow</td>
<td>30</td>
</tr>
<tr>
<td>3.2</td>
<td>Comparison of Actual Die Yield to Die Yield Simulation</td>
<td>36</td>
</tr>
<tr>
<td>3.3</td>
<td>Die yield improvement due to cycle time reduction</td>
<td>39</td>
</tr>
<tr>
<td>3.4</td>
<td>Die yield improvement from in-line inspection</td>
<td>40</td>
</tr>
<tr>
<td>3.5</td>
<td>Die yield impact of a new manufacturing tool</td>
<td>41</td>
</tr>
<tr>
<td>3.6</td>
<td>Die yield impact of reducing time to implement process changes</td>
<td>42</td>
</tr>
<tr>
<td>4.1</td>
<td>Integrated Supply Chain</td>
<td>46</td>
</tr>
<tr>
<td>4.2</td>
<td>Dis-integrated Supply Chain</td>
<td>47</td>
</tr>
<tr>
<td>4.3</td>
<td>from Bower and Christensen (1994)</td>
<td>52</td>
</tr>
<tr>
<td>5.1</td>
<td>Feedforward Yield Methodology</td>
<td>57</td>
</tr>
</tbody>
</table>
Chapter One — Introduction

In management generally and manufacturing specifically, the learning curve, improvements in cost, performance or quality over time, is ubiquitous. Many firms count on the learning curve effect to establish profitability. Frey (1994) admonishes managers to understand "the strategic and tactical use of the manufacturing learning curve." Pindyck and Rubinfeld (1994) state that "[t]he learning curve is crucial for a firm that wants to predict the cost of producing a new product" and "learning curve effects can be important for a firm deciding whether it is profitable to enter an industry."

Some of these experts treat improvements from the learning curve effect as simply the result of production. Said another way, simply by doing a firm will learn. These experts do attempt to explain neither why learning by doing happens nor why differences in learning rates across companies or industries. Other researchers have recognized that the learning curve effect is the result of conscious attempt within a firm to identify and solve problems that limit productivity (Adler and Clark, 1994; vonHippel and Tyre, 1995; LaTourette, 1996). The critical importance of learning to competitive strategy have been demonstrated by Christensen (1994), Leonard-Barton (1991), and Gruber (1994). Some managers have recognized the competitive importance of the learning curve effect, and they have instituted organizations and procedures to manage the learning process.

The impressive learning curve effect in the semiconductor industry is an example of improvements in performance resulting from managed learning. Figure 1.1 (Brynjolfson, 1995) shows the improvement in the number of transistors per chip for both microprocessors and dynamic random access memory (DRAM). While the capacity of DRAMs has increased, the price per unit of capacity has fallen. This cost/performance improvement has required producing smaller and smaller features on the chip, from several micrometers in the 1970's to today's processes which can produce features less than 0.25 micrometers. Intense competition has driven firms to increase the cost performance of their factories in shorter and shorter periods of time. Therefore, to stay competitive, firms must acquire the knowledge needed to produce ever more demanding products ever more rapidly.
Figure 1.1—Improvement of Process Performance in Semiconductor Manufacturing

This thesis attempts to link the learning process in semiconductor manufacturing with the performance factors that define competitiveness in this industry. The fundamental message of this thesis is that learning is active, not passive, process that requires a set of capabilities fostered by management. The ultimate goal for this thesis is to better understand what makes a learning processes effective and give managers as well as engineers the tools to design effective processes for enhancing learning as a strategic capability. These tools were used to evaluate the learning system within a fab producing memory devices. The techniques used and results obtained are presented.

Value of the learning curve.

There are many factors that define competitiveness in semiconductor manufacturing. In the microprocessor segment, for example, the instruction set architecture of Intel dominates the market with over 90% market share among firms producing chips to this standard. Firms with very large scale, like Samsung, gain a competitive advantage in DRAM production, where product designs is not as important a factor in market share. In all segments of the business,
however, manufacturing cost is a major factor. DRAM makers compete on cost, and microprocessor manufacturers need profits to plow back into design of subsequent product generations.

![Relationship between fab capacity, line yield, and die yield](image)

**Figure 1.2--Relationship between fab capacity, line yield, and die yield**

One of the most important ways to decrease costs of semiconductor manufacturing is to improve yield. The financial impact of yield in semiconductor manufacturing is demonstrated in Equation 1.1. Capacity and variable costs of running a fabrication facility, or fab, is a function of the number of wafers that can be processed in a given period of time. However, only a fraction of the wafers that begin the process complete the process. This fraction is known at the line yield. Similarly, only a fraction of die, or chips, on a wafer function properly. This fraction is known as die yield. The relationship among capacity, line yield, and die yield is depicted in Figure 1.2. Using a simple activity based cost accounting system, the unit cost is given by:

**Equation 1.1**

\[
Unit \ Cost = \left( \frac{Wafer \ Cost}{LY \times DY \times DPW} \right)
\]

where:

- \(LY\) = line yield, or fraction of wafers completing process
- \(DY\) = die yield, or fraction of good dice per wafer
- \(DPW\) = total die per wafer
For the fab, there are at least two learning curves, or changes in yield over time: (1) the learning curve for line yield, and (2) the learning curve for die yield. The nature of the difference between these two learning curves will be discussed later. For now, though, one can readily see that as the organization learns and both types of yield improve, unit costs go down. If the firm can increase die yield while selling all of the chips it produces, a common scenario in parts this industry, the benefit of learning not only reduces unit cost, but increases revenue overall.

Beyond the simple financial benefits of increased die yield lie strategic benefits that can be even more staggering. Improved performance can stimulate market demand, and, if the trajectory of performance improvement is faster than its competitors, can lead to market dominance by a firm (Bower and Christensen, 1995; Gruber, 1994). If learning can be effectively transferred to new increments of capacity, whether within a given factory or across a number of factories, that learning is leveraged to make this new capacity more productive. Transfer of learning across subsequent generations of products and processes can help to accelerate their respective learning rates (Weber 1996).

**Learning by doing.**

The original notion of a learning curve was introduced by Wright (1936). In his study he noted that as aircraft were produced, workers and engineers learned how to better produce them and unit costs fell over time. He defined the learning curve as an empirical relationship between a unit cost and cumulative units produced. Unit cost was a function of the number of units produced raised to an empirically determined factor. Many functional forms of the empirical relationship between unit cost and units produced have been proposed (Badiru, 1992), and several proxies of cumulative production have been proposed. Some of the proxies included are time (Asher, 1956) and cumulative investment (Sheshinski, 1967). Only recently, however, have researchers studied why such relationships exists and how it takes happens. Several have studied the productivity improvements of direct labor due to learning (Levy, 1965). Increased efficiency in manufacturing due to improvements in product and process engineering are also important, however. Unfortunately, Adler and Clark (1994) note that both managers and researchers often assume that the learning curve for machine processes, such as an automated assembly line, is slow compared with labor intensive processes.
Types of learning

As von Hippel and Tyre (1996) explain, engineers and tool operators do not learn by repetition the way a manual worker does. Rather, they learn through problem solving and, in turn, making adjustments to product designs, tools, and procedures. The literature cited above uses many different terms to distinguish these modes of learning, but here two are used:

- behavioral learning—increased productivity of direct labor due to repetition
- cognitive learning—increased productivity of tools, processes, and organizations due to problem solving

To better understand how these terms apply, return to the semiconductor manufacturing situation above. Line yield losses typically are due to mishandling (dropping wafers) or misprocessing of wafers (using the wrong tool recipe) by the technician. Mishandling losses decrease as technicians' handling skills improve, and misprocessing losses decrease through either better procedures or better adherence to the procedures by the technician. Because the primary mode of improvement is through repetition, line yield might be a measure of behavioral learning.

The mechanisms for a majority of die yield improvements are quite different, however. Most die losses are due to contamination of circuits by tools or the environment, or failure of the process to build circuit features to specifications. Either way, no amount of repetition without problem solving can improve die losses for these reasons. The organization must develop solutions to these problems and implement them to improve die yield. Die yield, then, both represents accumulated learning and affects costs as shown in Equation 1.1. This accumulated learning is partially responsible for the performance improvements seen in Figure 1.1.

Models of learning processes

In recent years there has been some research into the learning-by-doing process. Adler and Clark (1991) describe a learning process model that draws a distinction between behavioral and cognitive learning described in chapter 1. Their empirical analysis, using firm-level data such as cumulative engineering hours and training time, shows that using models based on cognitive learning factors yield better results than those relying on behavioral factors such cumulative volume. Bohn (1992, 1994) has done significant mathematical modeling of the learning process and from this work identified eight stages of knowledge. His work suggests that there are
optimal methods for managing the learning process based on the stage of knowledge. For example, if engineers can use statistical methods to calculate process capability, their optimum quality strategy is to use methodologies like statistical process control. Simply sorting between good and bad die would be an inefficient use of accumulated learning, but knowledge is not sophisticated enough to adopt a feed-forward quality methodology.

**Learning in semiconductor manufacturing**

Methods for managing the learning process for competitive advantage will be explored in three sections:

Describe the learning process in manufacturing.

A generalized version of the learning process in a semiconductor fab will be described in detail. It includes a sketch of the semiconductor manufacturing process, and it highlights the basic technical issues associated with die yield losses. The collection of manufacturing information relevant to die yield is described, and the process for analyzing that data to drive problem solving is explored. It will also describe how the architecture of the learning process can be redesigned to achieve higher die yield performance, and it will present some benchmarking data that suggest how to sustain that performance.

Simulate the learning curve effect.

To analyze cost performance and assess the impact of different learning processes on die yield performance is valuable to fab management. Using the key variables identified in the discussion of the learning process, a method of simulating the die yield improvement process will be proposed. The model will seek to simulate the effect of the configuration, or architecture, of the learning process on productivity for a factory producing memory devices. Most of all, the model is designed to be simple enough in order that it is relatively easy to collect input data and the function of model is clearly understood by the team responsible for the learning process.

Discuss the relationship between learning and competitive advantage.
Chapter four will review research to define the relationship between learning and competitive advantage. In the first section, the emergence of disintegrated supply chains in semiconductor manufacturing will be described, and its potential as a threat to integrated manufacturers will be explored. In a disintegrated supply chain, individual firms are free to concentrate on ways they can contribute to improved cost performance, through either design or manufacturing. Good design firms can then partner with fast learning fabrication firms to achieve competitive advantage. This chapter will continue by drawing on research in supply chain dynamics and analysis of improvements in cost performance over time.

Conclusions

This thesis will conclude with recommendations for the memory device fab that was the focus of this study. It will continue by summarizing the findings of previous chapters and what they imply about the potential of innovation of the learning process in manufacturing. It will also discuss what the findings suggest about how current improvements in cost and performance in semiconductor manufacturing might be sustained.
Chapter Two — Learning in Semiconductor Manufacturing

As described in the introduction, accelerating cost performance creates and strengthens competitive advantage. Using learning processes to accelerate die yield performance over time is an important part securing this advantage. This chapter how the learning process is developed to sustain die yield improvements. Semiconductor manufacturers actively manage the learning process, usually through organizations and techniques crafted to find die yield problems and direct process engineers to eliminate them. This chapter will describe the learning process of the semiconductor industry in detail. This is done to highlight particularly valuable practices in the semiconductor learning process and to lay the foundation for the modeling in chapter four.

One goal of the learning process in semiconductor manufacturing is to ramp the yield of the manufacturing process. Yield ramp, for this thesis, refers to the rate of change of die yield as a function of time (see figure 3.1). This is critical to profitability as faster yield improvement directly leads to higher productivity (see equation 1.1). Early in the lifecycle of a process, die yields tend to be low. Engineers identify the faults in die that do not function properly, determine the root cause of those faults, and eliminate them. Here, this activity is referred to as a learning loop. The term learning is used because the process change is retained. The faster learning loops are completed, the faster yield can be ramped to profitable levels.

This chapter begins with a short tutorial on semiconductor manufacturing to put learning activities modeled in this thesis into the context of the manufacturing process. An explanation of the metrology used to diagnose the health of the manufacturing process will be included. Information generated by this metrology will be followed through the learning process, and each activity will be explained. Next, a particular improvement in the basic learning process, in-line defect inspection, will be introduced. This chapter includes by suggesting how the yield improvement process has been improved by some factories to be more effective.

This new type of yield system was considered by a fab producing memory devices. To support the fab's decision making process, a benchmarking study was used to determine some factors that have been used successfully by other semiconductor firms to drive yield ramps. That benchmarking data is described as well as some of the implications for the potential success of in-line defect inspection for this fab.
Fig. 12. Yield learning for DRAM's.

Figure 2.1—Example of yield ramps (Stapper, 1995)

**Semiconductor manufacturing**

Semiconductor manufacturing essentially consists of two activities: building transistors or other electrical component on a silicon substrate and wiring those components together into a circuit. These activities are usually referred to as the front-end and back-end, respectively. The structure is built in series of layers, each of which is produced in a cycle of thin film deposition, patterning, and etch or implant (see Figure 2.2). A series of these cycles resulting in a particular structure is called a module. Once the modules are connected within the structure, the result is a very large scale integration, or VLSI, device.

The tools used in this process subject the wafer to very high temperatures, very low pressures, and complex chemistries. The products themselves may consist of millions of transistors. This combination of high complexity in both the product and process requires many engineers dedicated to each product and process tool in the factory. The very high degree of interaction
between product and process requires intricate coordination among engineers. The organization used to achieve this coordination is critical to accomplishing a rapid yield ramp. Yield engineers must effectively direct information on fault analysis to process engineers, who in turn work with integration engineers to assure that constraints due to multiple processes being run on multiple tools are not violated. Figure 2.3 depicts the organizational relationships between process engineers for functional areas, such as thin films, lithography and etch, and integration engineers across several generations of processes (G1, G2, and G3). Weber (1996) discusses the relationship between organizations and learning in semiconductor manufacturing.

Semiconductor Wafer Fabrication Process

Figure 2.2—Semiconductor Wafer Fabrication Process

Once the devices have been built they are each electrically tested for proper function and sorted into good and failing die. Failure can occur for two main reasons. First, the actual structure of the device may not conform to specifications. For instance, the gate of a transistor may be too thin and leak charge when turned on. This type of failure is known as parametric. However, some perfectly constructed die may fail. This is usually due to a defect, such as a particle, that
corrupts the die during processing and interfere with normal electrical function. This type of failure is usually described as defect driven. It is important to note that a particular defect may or may not result in a detectable fault.

![Diagram](image)

**Figure 3.1: Organizational complexity of full VLSI process experiments. Each solid line depicts a necessary communication channel between an integration engineer and an engineer of a constituent process technology.**

**Figure 2.3—Example of Organizational Interaction in the Learning Process (Weber 1996)**

Early in the lifecycle of a process, a majority of die failures are parametric in nature. Sometimes this is due to a structure, such as a transistor gate, not functioning as designed. Other times, process tooling is unable to produce structures within specifications on a reliable basis. As engineers work to correct these faults, many are eliminated. By the time production volumes are increased to full scale production, most faults are caused by defects. This study focuses on the process for detecting and eliminating defect-driven faults. The reader should be aware, however, that the process for eliminating parametric faults in fundamentally the same, and it is an important part of the learning curve effect early in the product lifecycle.
The learning framework

Semiconductor manufacturers learning by doing, but how does this happen? Adler and Clark examine this process, but, because they use firm level data, it would be different to analyze tactical decisions based on their model. Bohn's work strongly suggests what stage, or level, of knowledge should be pursued by a firm, but, like Adler and Clark, he does not provide methods to analyze specific learning processes that could be used. There is, however, a learning framework in their studies that functions like a feedback loop of problem identification, problem solving, and knowledge acquisition. The same basic framework is used here to analyze specific options for learning (see figure 2.4). Data on particle counts and defect characterization are collected during the manufacturing process using specially designed tools, and it is converted into information by yield analysis engineers. That information is used to instruct product and process and integration engineers in process improvement. They identify root cause, eliminate the problem, and embody new knowledge in a procedure or tool modification, for example. This process is described in more detail in the paragraphs that follow.

Yield Improvement Methodologies

End-of-line Yield Analysis

In-line Yield Analysis

Figure 2.4—Examples of Learning Loops
Collecting Data

Many semiconductor manufacturers collect voluminous amounts of data during the manufacturing process. Data collected can include tool settings, readings from process sensors, and measurement of critical dimensions of circuit features. This data is stored in a complex database that will allow engineers to conduct sophisticated data analysis.

For defect-driven yield analysis, an important type of data collected during the process regards the number and type of defects present. Defects are typically foreign objects, such as specks of dust, that can interfere with the proper function of the circuit. Tools will scan the surface of the wafer to find defects, which can be on the order of 0.1 micrometers in diameter. Technicians will classify the defect according to its shape, and its size will be noted. This data can be used in a control chart to monitor the state of process control. If the data is archived, it can be used for future analysis as well.

Data is also collected after die are sorted. First, data is collected in where die fail and what particular type of electrical failure, or fault, was detected. Wafers are then sampled for a procedure that determines the location of faults and strips back each process layer until the fault is isolated and characterized. A typical fault may be “called out” as a “flake at layer 5”. This data is compiled for many bad die on many wafers into a pareto chart according to the total die impact of each fault type. From this chart, teams of engineers prioritize their defect reduction activities. This method of using data from the sort and fault isolation processes is known as end-of-line analysis.

Converting data into information

Next, yield analysis engineers determine the root cause for the fault. For particles, this may include determining the composition of the particle and finding its source. Typical sources of particles may be skin from operators or residue from process tools. Finding this type of information in the mass of data collected requires sophisticated analytical tools. Particle composition may be determined using mass spectroscopy techniques, and the results are used to hypothesize the source of the particle. If the particle composition is similar to the by-products of a chemical vapor deposition step, then engineers may focus on residues in the CVD tool. Regression and correlation analyses may be used to find differences in performance of various
tools. A yield engineer may find a statistically significant difference in die yield between one CVD tool and the other three.

This analysis requires some amount of time. This time may be defined by the number of microscopes or mass spectrometers available for engineers to conduct analyses. Another common problem simply is not having the engineering staff available to identify these problems as quickly as might be ideal.

**Using information to direct the acquisition of process knowledge**

Once the root cause has been identified, the learning process resembles the Plan-Do-Check-Act cycle common in the Total Quality Management literature (Shiba et. al, 1993). Process engineers, working with information from yield analysis engineers, hypothesize about how to eliminate the root cause of a defect driven fault. Next, they plan and conduct experiments to test the hypothesis. Experiments can consist of two types. Short-cycle experiments involve only a few process steps. In the example of the layer 5 fault cited above, the experiment will only span the process steps associated with layer 5. Usually, though, permanent changes will not be made until full-cycle experiments have been completed. If the hypothesis is proven (the check of the PDCA cycle), the fab makes a permanent change in its processing using established change control procedures to institutionalize the change.

**Improving the process**

Using the current learning process, faults are not identified until the end of the manufacturing process, or end-of-line. This means that a large part of the time required for a learning loop, from data collection through experimentation to process change, is consumed by manufacturing steps that have nothing to do with those that caused the fault. If many faults are generated during front-end processes, engineers wait for the back-end processes to be complete before they begin root cause and PDCA activities?

Some manufacturers have short circuited the learning loop by leveraging data from in-line defect monitors (Leachman, 1996). These monitors scan the surface of the wafer to detect excursions in the process defect density. Wafers pass through the monitors following specified process steps. As long as the defect counts fall within control limits, processing continues as normal.
Returning to the layer 5 flake scenario above, an engineer can recall the in-line monitor data to see how many of what type particle was found after the layer where the flake was found. If there is a high correlation between the defect count found by the monitor and the faults found during the end-of-line failure analysis, end-of-line yield losses due to faults at that layer can be predicted. While there can be some interaction between a particles deposited at one layer and those close to it, it is uncommon.

This method of yield analysis relies on an existing base of knowledge. Engineers cannot tell technicians how to classify defects before they have seen them. Some amount of knowledge must exist before this system can work. For example, engineers must know what kind of defects to look for, and they must know which types of defects have high probabilities to cause faults. Most factories will begin to "ramp" a new process by expediting a few lots through the process to get them to end-of-line. Faults can then be isolated and characterized, and set of defect calls is established.

After this initial learning period, some faults can be predicted using the in-line monitor set. Not only can root cause analysis can begin sooner, but engineers do not have to wait for end-of-line data to confirm the results of their experimentation. This process would be like taking a few pieces of dimensional variation and correlating it to a measure of quality of a finished assembly. With that data, the process improvement part can begin at the component level rather than the level of the finished assembly. The shorter learning loop allows engineers to solve problems faster and put less inventory at risk for failure.

This learning methodology presented here is simulated using the model presented in chapter 3. That model is used to simulate the effects of the change from using end-of-line data to in-line data to drive the learning process. Those changes are compared with other options for increasing the yield ramp such as reducing manufacturing throughput time. While this learning methodology is effective at identifying and solving problems that limit the die yield of a process, it does have its limits. While equipment vendors are developing new defect identification and information systems to expedite the time required for learning loops based on this methodology, it is still a feedback process that reacts to faults. Suggestions for a potentially better approach are made in chapter 5.
Benchmarking of Die Yield Performance in Semiconductor Manufacturing

While the basic process for learning presented above is common in the semiconductor industry, how does one explain differences in learning rates across the semiconductor industry described by Leachman (1996)? A semiconductor manufacturing company requested benchmarking of die yield and learning systems within the firm and across the industry. Particularly, the firm was interested in whether or not the more advanced die yield system described above might be of financial benefit to this factory. In answering this question, it was decided to rely on a combination of benchmarking data supported by a simulation of how die yield in that factory might respond to the more advanced system.

While the basic yield methodology described above is used pervasively in the semiconductor industry, Leachman (1996) has identified variations in learning systems that significantly impact line yield. He divides these differentiating factors into three categories: information systems, organizational practices, and process technology. Information systems that have substantial die yield benefit include statistical process control, automatic data logging, and yield correlation analysis, among others. The in-line yield management system would incorporate each of these.

Organizational factors Leachman singles out include integration across engineering and manufacturing groups, operator and technician improvement teams, and close coordination with process development factories. While these factors were present to some extent within the factory, the new system did nothing to change current practice. Internal benchmarking did reveal new resource demands in connection with data logging. Other factories had integrated their manufacturing operators with yield management technicians to address this problem.

Process technology factors that Leachman highlights include process flow re-design, product re-design, and machine modifications. There is substantial literature that suggests that these process technology factors would be closely linked with organizational factors. Iansiti (1995) studied the development of microprocessors for the mainframe computer industry. He observed that organizations that effectively climbed the performance curve were those that excelled in technology integration. This integration was accomplished by planning process innovation according to the relationship between technical options and existing capabilities of the organization. Weber (1996) provides a detailed study of integration of knowledge across
organizations in order to drive experience curves. Again, refer to figure 3.3 for an example of flow of knowledge through the semiconductor organization he studied.

This work implies that the plan for implementing an in-line inspection system should include defining an organization to address the technology involved and how it will impact the capabilities of the factory's technical staff. New technologies include data management systems and yield correlation software. Fab management assigned an engineer to bring the new software, which was developed in the firm, into this fab. A more experienced engineer was assigned to work with fab's information systems department to make the necessary changes to the data logging and retrieval system. At the time of this study, however, making changes to sampling rates and new demands on both in-line and end-of-line defect metrology had not yet been planned. This would impact both manufacturing as well as engineering. This type of integration would be critical to the success of the project.

**Conclusion**

In this chapter, the process of learning by doing in the semiconductor industry was explained in some detail. The collection of manufacturing information was described, and the use of that information for problem solving and knowledge acquisition was explained. Even though this basic process is present across the industry, differences in learning rates are present. Benchmarking data on yield learning systems was presented. In chapter five, this information will be combined with information in subsequent chapters will be integrated to form recommendations for the fab that was the subject of this study.
Chapter Three — Modeling the Learning Process

In chapter two, the yield improvement process at a semiconductor factory was reduced to its constituent activities. In the production process, these activities include collecting process data and testing finished die for functionality. Aside from the production process engineers isolate failures, identify root causes, and experiment to eliminate those root causes. In this chapter, that process is reconstituted in a simulation model. The objective of the simulation is to explore the dynamics between variables from the manufacturing system and those from yield improvement.

Chow, et al., (1991) suggests that simulation of new technology can be critical to effective implementation. Using the simulation presented here, a manager can better understand how changes to certain variables effect the performance of the yield improvement system over time. Data used to drive this model was readily available without extensive validation, and the results were directly related to the variables that define the fab's performance. Finally, this kind of model helps to provide information for the decision support system for factory management.

In this chapter, the general principles of the model are explained. First, the key variables in the simulation model will be introduced, and the assumptions behind the model will be explained. Next, the function of the simulation model on those variables will be described. Finally, some generalized results from analysis of data for the memory fab will be presented. Detailed information on mathematic equations and non-dimensionalized data from simulations runs are presented in the appendix.

Scope of the model

Die yield performance in a semiconductor fab is an aggregation of numerous learning curves. There is a learning curve for each of several hundred tools, each product layer, and for the interaction between product and process. Disruptions in knowledge occur as new processes are introduced, and there is turnover among the technical staff. There are unforeseen events, such as power outages, that result in yield losses. The overall factory yield performance depends on the yield from each of these activities or variables.

A single model accounting for the myriad of variables associated with each of these curves would be unfeasible. This simulation is intended to capture and model learning at the process
technology level. As such, it aggregates tool knowledge for all tools used in that process. It also aggregates the problem solving skill of the operating technician and process engineers, and thus reflects the state of organizational learning at the factory level. The model does not seek to model disruptions of process knowledge due to staff turnover, nor does it attempt to model the impact of excursions.

**Structure**

The structure of the simulation model is taken from the process described in chapter 2. As shown in Figure 3.1, batches of product flow from one stage of the production process to the next. Along the way, engineers collect data on the production process and use that information to eliminate the root cause of yield problems. Each of the steps in this process take time just as stages in the production process do. Together, these activities form a learning loop. By changing the amount of time required to complete a learning loop, one can manipulate the rate of yield improvement. For instance, if the process cycle time is reduced, the overall time for a learning loop shortens accordingly, and yields improve.

![Figure 3.1—Model Process Flow](image-url)
LaTourette (1995) used this concept to develop a yield learning model. He modifies a classical learning curve model by replacing units or production with learning cycles. Next, he replaces time required for production, or cost, with defect density. His model is described as follows:

\[ DD = DD_0 \left( \frac{t_{lc}}{T} \right)^{\frac{\log(b)}{\log(2)}} \]

where,
- \( DD_n \) = Fault density in time period \( n \) \((n = \frac{t_{lc}}{T})\)
- \( DD_0 \) = Initial fault density
- \( t_{lc} \) = Time required for learning loop
  - = process throughput time + problem solving time
- \( b \) = learning factor empirically derived for the fab

The key variable here is \( t_{lc} \), or average time to complete a learning loop. LaTourette defines it to be the sum of variables such as process cycle time, fault isolation time, and time required to implement solutions. If process cycle time is reduced, the defect density at any time period \( T \) will be lower.

**Modeling in-line data collection**

Breaking the coupling between process cycle time and learning cycle time that is assumed in LaTourette’s model poses difficulties. If there are many stages, tracking average learning cycle times \( t_{lc} \) and learning factors \( b \) becomes quite complicated. Within some fabs, experimentation capacity is limited to a certain percentage of total capacity. Because of this limitation, experimentation capacity is allocated according to the magnitude of the yield problem. If defect density for stage one is twice that of stage four, accordingly more experiments will be dedicated to stage one. Defect density in stage one will decrease faster than stage four, therefore, their learning factors, as defined by LaTourette, are different.

To address these weaknesses, the concepts from LaTourette are employed in a queuing model that accounts for shifts in problem solving priorities over several production stages. The queuing simulation is divided into two portions: the production process where faults are generated and detected, and the problem solving process where root cause is established and eliminated. Once a root cause is eliminated, the production processes is altered, closing the feedback loop. Each
of the steps in these processes takes time. This model queues die through production and information about faults through the problem solving process.

The model presented here is like the LaTourette (1996) model in two important ways:

1. Each problem solving loop results, on average, in some die yield improvement.
2. Decreasing the time required to complete a learning loop accelerates die yield improvement.

Some of the simplifications of the model, such as process times being deterministic, have already been discussed. Two other important simplifications and their implications are discussed here.

To account for the multistage nature of the production process and the prioritization of problem solving within the process, this model functions as a deterministic queuing network with “m” stages. Each stage agglomerates the time required to process wafers through the tools within the jth stage, or the stage throughput time, “nj”. A deterministic model was chosen because the total time through the stage is the important variable, not queue sizes. For the process studied, the expected throughput time for a stage was established using historical data. For a process being planned, standard queuing network models can be used to generate the expected throughput time data needed for this model.

Due to their very nature, excursions, or yield losses due to special causes, cannot be predicted. This class of yield losses includes unexpected equipment failure, operator error, and acts of God. This model only considers the baseline, or chronic, die yield. Experience shows that, over a long period of time, excursions will occur, and factory output will suffer. Minimizing the risk of excursions is better analyzed using methods such as Bean (1997).

Disruptions to organizational knowledge could be examined using this type of model. A knowledge disruption occurs when existing knowledge becomes less effective or irrelevant due to a technological change. An example of this type of change would be the introduction of a new manufacturing tool with a lower defect density. While the tool potentially may be much cleaner, initially the engineering staff might understand less about how to achieve lower defect densities. Given some problem solving experience, process engineers, hopefully, will return the tool to higher yield levels. This effect has been modeled empirically by Hatch (1994). His study relied on data from across the semiconductor industry, however, and he used only dummy variables to
determine the magnitude of the effect. While there is not enough data to use a similar approach for this study, this kind of disruption could be modeled at the user's discretion. For instance, if another fab had introduced a particular tool in the factory environment, their experience could be factored into the model.

Capacity at each stage of this model is assumed to be infinite. The state of in-process inventory at any stage \( m \) is described mathematically as follows:

\[
Q_{j,t} = \sum_{i=1}^{t} (P_{j-1,i} - P_{j,i})
\]

where: \( Q_{j,t} \) = number of wafers at stage \( j \) at time \( t \) 
\( P_{j,i} \) = number of wafers process through stage \( j \) at time \( t \)

Each stage takes \( n \) amount of time, so:

\[
Q_{j,t} = \sum_{i=t-n}^{t} (P_{j-1,i})
\]

At each stage, faults are added to the wafer according to the fault density of each stage, \( \text{FD}_{j,t} \). Therefore, at any point within the process, the fault density on a wafer is equal to the sum of the fault densities of each previous stage. As wafers complete the process, they are sampled for fault isolation, where the layers of failing die are stripped back to expose and characterize faults. The process of isolating faults is like any other process with its own process time and queue.

Once faults are isolated and the root cause of that fault is found, experiments can be performed to solve the problem of how the eliminate the root cause. Because most fabs run at full capacity, there is a cost associated with running experimental wafers, especially if salable die cannot be salvaged from them. Therefore, experimentation capacity is fixed as part of the overall factory production plan. Experimentation capacity is allocated to different parts of the process according to a fault pareto. In the model, experiments are allocated proportionally. That is, if 40% of faults occur in stage 2, 40% of experimentation capacity for the given time period will be
allocated to reducing the fault density in stage 2. This is a good approximation of practice in the fab where this model was constructed.

Experiments are put into the queue in the stage in which it is targeted. To continue the above example, a batch that has completed stage 1 in the given time period is used to conduct the experiment. This assumes that process changes in stage 2 will not interact with the processes in stage 1. This assumption is consistent with current practice for reducing particle-driven faults. When results from the experiment are available depends on whether in-line monitors at the end of the stage can be used to effectively predict die yield. Effectiveness is measured as the percent faults that can be identified using in-line defect data. If the current in-line data effectiveness is greater than a standard set by engineers, that data is available at the end of the stage. Otherwise the results are not available until the end of the sort process.

Once experiments are completed, the model assumes that fault density will decrease by a fixed percentage. Certainly, the experimental die yield response may not rise above process noise, and it may even be negative. On the average, however, they do yield positive results. Using an average also helps to avoid accounting for complex and unique iterations of experiments that cannot be predicted.

Die yield improvement due to completed experiment is modeled as follows:

\[
FD_{j,t} = FD_{j,t-1}(1 - FD_{j,u} - r_t)
\]

where: \( FD_{j,t} \) = fault density at stage j for time t
\( FD_{j,u} \) = ultimate fault density achievable
\( r_t \) = percentage reduction in fault density per completed experiment

The terms \( r_t \) and \( FD_{j,u} \) are determined using historical data. If the process under consideration is already in production, \( r_t \) can be determined by fitting the yield curve from this model to actual data. Otherwise, historical data could be used. Studies of several process generations in this fab yielded values for \( r_t \) consistent enough to be used as a predictive value, although a sensitivity analysis would be warranted.
This can be compared to LaTourette's learning curve by taking the term \((T/KLC)\) from equation \(x\) as the number of completed experiments, \(E\), and rewriting that equation as:

\[
FD_n = FD_0 E^{\log(b)/\log(2)}
\]

Equation 3.5

Using equation \(y\), one can calculate the fault density from this model after \(E\) experiments as follows:

\[
FD_{j,t} = FD_{j,0} (1 - FD_{j,u} - r_j)^E
\]

Equation 3.6

Comparing equations 4.1 and 4.6, one can see that:

\[
(1 - FD_{j,u} - r_j)^E = E^{\log(b)/\log(2)}
\]

Equation 3.7

Results of experiments are not immediately implemented. Change control procedures in a factory may require experiments to confirm the benefit of a change and management review before a change can be made. If the change requires some kind of equipment procurement, there may be lead times to consider as well. For the examples shown in the appendix, the expected implementation time was based on historical data. Once the implementation process is completed, new values for \(FD_{j,t}\) take effect in the production process.

Fault density is converted to process die yield using a yield model. A common yield model is the Murphy model (Murphy, 1964) which relates the two as follows:

\[
DY = \left(\frac{1 - A*FD}{e^{-A*FD}}\right)^2
\]

where:

- \(DY\) = die yield
- \(A\) = die area
- \(FD\) = fault density
Other yield models, either open or proprietary, may be used as is suitable.

**Applying the yield ramp model**

The semiconductor fab that was the focus of this study was producing memory devices using an end-of-line yield methodology described in chapter 2. Fab management was aware of accelerated yield ramps at other fabs using an in-line method, and wanted to know if the same techniques, applied in this factory, would provide similar results. Benchmarking data was gathered as described in chapter 2. The benchmarking data helped to gauge costs of the new system, but management was not confident that higher yield ramps these fabs had could be used to estimate the benefit of in-line inspection. The model was used to fulfill this role.

![Yield Time Series](image)

**Figure 3.2—Comparison of Actual Die Yield to Die Yield Simulation**

Before the model could be used in this role, there had to be confidence that it could be used to predict yield within the factory. In figure 3.2, the yield ramp of the simulation is compared with the actual yield over a given period of time. This data has been disguised to protect proprietary information. Die yield from the last four time periods was withheld from the fitting of the model. The model reflects the slow start of the ramp as the new process is brought into the factory. As fault data become available and experiments are completed, die yield begin to increase. The magnitude of yield improvements decreases over time because of effective prioritization. The factory engineering staff was comfortable that this simulation at least
suggested how output might change with changes in the yield system, even if it could not predict the yield at a given period of time. To reiterate, the goal of the model is to estimate the impact of system changes, not to predict die yield.

The benchmarking exercise and study of the current yield process suggested several scenarios for improving die yield performance. Four of those scenarios are presented here: in-line inspection system, process cycle time reduction, tool upgrade, and streamlining the implementation of process improvements.

**Implementation of in-line inspection system**

Implementing an in-line inspection system means augmenting data collection to correlate in-line defects with end-of-line yield. As discussed in chapter 2, this can enable engineers to determine root cause and prioritize problems while the wafer is in-process. The overall time required to close the learning loop is reduced, thus accelerating the yield ramp.

Implementing this system requires investment in metrology, information systems, and the time of process engineers. Relative to the very high capital costs of the semiconductor industry, this investment is small. Significant engineering resources are required, however, to optimize metrology tools to detect important process defects. Therefore, management needs to know if this investment is justified by the increase in output that results.

**Production process cycle time reduction**

Reducing the production process cycle time is one way to decrease the time required for learning loops. Both production and engineering wafers travel through each process stage faster. Therefore, process faults can be diagnosed and process improvements developed faster.

Cycle time reduction must be weighed against the corresponding decrease in throughput. Other studies (Perrin, 1997) show that, for a given yield, overall throughput is economically more valuable than the working capital invested in in-process inventories. Higher yield ramps from lower cycle times must be weighed against the reduced number of wafers produced in a given period of time.
New manufacturing tool

Engineers are tempted to expect that yields will improve in a step function when old tools are replaced with new tools. Tools are closely evaluated to determine what the expected yield improvement will be. This improvement can be compared with reduction in cycle time, etc., to decide how to allocate resources.

Close examination of the learning process suggests that the step function model may not be true. There could be a yield disruption as process engineers become familiar with the new tool. Once the expected yield improvement is achieved, there may be the opportunity for a higher rate continuous improvement. Unfortunately, others’ experiences with this tool were not available for such an analysis.

Reducing time required to implement improvements.

The process for approving changes to the manufacturing process are carefully controlled through a set of procedures that include pilot-scale manufacturing and one or more levels of management review. This serves to carefully check the results of process engineers and, if necessary, coordinate changes across different organizations. The cost/benefit of streamlining the change control process is evaluated using this model.

Observations

Because of proprietary information concerns, results of the actual die yield analysis cannot be revealed. However, the general implications of the results can be indicative of managing the learning within this factory. For each of the scenarios tested, percentage change in die output from current practice were computed. These results are presented in table 3.1 below.

Process cycle time reduction

Of the four scenarios tested for this factory, the model suggests that process cycle time reduction had the greatest impact on die yield ramp (see figure 3.3). In a typical semiconductor manufacturing process, over half of the overall cycle time is consumed by transit from one process to the next and by waiting in buffers. If that waiting time is reduced significantly, information can be collected much faster. Of course, there can be an economically significant decrease in throughput (Perrin, 1997) because the system becomes more susceptible to the
breakdown of tools and variability of service times within the process. Early in a process lifecycle, though, the yield benefit may outweigh loss of throughput.

Figure 3.3—Die yield improvement due to cycle time reduction

Process cycle time has a very large impact on yield because the effect is magnified by the experimentation process. Information may flow through a particular process numerous times. One way semiconductor manufacturers mitigate this problem is by giving priority to experimental batches. This practice may be counterproductive if it confounds the results of the experiments.

**Using in-line monitors**

The value of in-line monitors for continuous die yield improvement varies with the process cycle time. One can imagine that when cycle times are very long, data from in-line monitors will be more valuable. Because of the very high demand and margins for logic chips, cycle times are very long to maximize throughput. In-line monitors can be very valuable in this case. If cycle times are short, though, the value of in-line monitors is diminished. For the simulation shown in figure 3.4, the output effect was much smaller than might be observed for products with longer process cycle times.
Figure 3.4—Die yield improvement from in-line inspection

The value of in-line monitors can also be a function of the distribution of faults through the process. Consider an example where the bulk of faults are added to the die during the early part (or front end) of the process. Even if the cycle time is short, the time required for the die to get to end-of-line fault analysis can be quite long. Short circuiting the feedback loop using in-line monitors is quite useful in this case.

Introducing new manufacturing tools

Semiconductor manufactures take great care in assembling and altering their process equipment sets. After all, these can be multi-million dollar investments. Process engineers understand what improvement in defect density they expect to see for the given investment (see figure 3.5). If the simulation would provide a standard basis of comparison between cycle time reduction and tool acquisition, managers would be more comfortable making decisions about tradeoffs. In the example analysis presented in the appendix, a new tool was found to be of about as much benefit as using in-line monitors for problem solving. For some tools, the investment in a new tool can be an order of magnitude higher than the investment in inspection metrology.
Figure 3.5—Die yield impact of a new manufacturing tool

**Reducing time required to implement changes**

In the example scenario analyzed, implementation time was reduced by 25%. While this reduction is the same as the amount of process cycle time reduction, the corresponding benefit is much lower (see figure 3.6). This is because change control happens once per learning loop, while the manufacturing process re-enters the learning loop through at least one iteration of experimentation.
Figure 3.6—Die yield impact of reducing time to implement process changes

**Conclusions**

The die yield simulation model presented above was able to estimate differences in die yield over a period of time. Results, disguised to protect proprietary data, are presented in the table 3.1 below.

**Table 3.1—Summary of simulation results**

<table>
<thead>
<tr>
<th>Improvement Scenario</th>
<th>Normalized impact on output, new tool=1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Process cycle time reduction</td>
<td>6</td>
</tr>
<tr>
<td>New manufacturing tools</td>
<td>1</td>
</tr>
<tr>
<td>Reduce implementation time</td>
<td>1</td>
</tr>
<tr>
<td>In-line inspection system</td>
<td>1</td>
</tr>
</tbody>
</table>

These results would suggest that each of the proposed scenarios would improve the output of the factory. The largest of these increases comes from cycle time reduction, in which case output would increase six times as much as from other scenarios. However, as mentioned, there may be a corresponding loss in output due to overall functioning of the manufacturing system. A full scale simulation of the queuing system in the factory would be needed to determine the net change in output. Such a simulation was outside of the scope of this project. The fab used for this simulation effort has very low cycle time, and according to benchmarking data (Leachman,
1996) this fab would rate among industry leaders. It is questionable how much more cycle time reduction could be reasonably accomplished. Decisions about new tools are made regularly in a semiconductor fab, and the costs and benefits are well characterized. In some cases, using an in-line inspection methodology would be a good choice for improving die yield, as well. Successful implementation of an in-line inspection system could have a positive impact of about the same magnitude of installing better manufacturing tools, depending of course on how much benefit is realized from the tool being considered. This comparison was helpful within the factory that was the object of this study. Being able to compare the impact of the in-line inspection system with a tool upgrade can be an important factor in the fab's decision support system.

It is important to note here that this simulation assumes a successful implementation of the new technology. The organizational factors presented in chapter two cannot be captured in this model, however, they deserve serious consideration. In chapter five, the simulation results will be combined with the benchmarking of chapter two to make final recommendations.
Chapter Four — Manufacturing Learning and Competitiveness

This chapter discusses the relationship between learning as a competitive factor and trends in the semiconductor industry. The first part of this chapter describes the state of industry and the dominant structure of the supply chain. Next, it will explore some of the trends among newer entrants in the semiconductor industry. Specifically, trends in virtual fab supply chains will be described, and the implications of these trends on learning and cost performance will be hypothesized in light of the literature.

This chapter will proceed as follows:

- Integrated firms supplying the computer market dominate the semiconductor industry.

- The very high capital costs of integration has driven semiconductor firms serving smaller markets to pursue alliances of dis-integrated firms. Single wafer fabrication process flows run on standardized tools are another way to reduce capital costs.

- Single wafer process flows may have an advantage in process learning, and standardization of the interfaces in the semiconductor supply chain will allow competitors to adopt this type of innovation more quickly than establish firms.

- Established firms can take steps to assure they do not miss adopting a technology that can radically change the competitiveness of their industry.

The semiconductor industry

Currently, the semiconductor industry is dominated by a few large players, such as Intel in the microprocessor segment and Fujitsu in the dynamic random access memory (DRAM) segment. These companies are highly integrated, with operations spanning circuit design, circuit fabrication, and chip packaging. Within these firms there is tight coupling between the design of product family generations and process technology generations.

Intel, for instance, has competitors in the microprocessor market, such as Advanced Micro Devices and Cyrix. These companies have developed very competitive designs, but they have not been produced in high volumes at low prices necessary to take significant market share away
from Intel. Capital intensity, large scale and highly-productive processes have emerged to be significant barriers to entry against new firms. In the past several years, however, these barriers are being threatened by new alliances that can replicate some of the competitive strengths of today’s industry powerhouses. This creates the potential that firms with valuable device designs could partner with firms who have established fast ramping manufacturing processes to pose a significant threat to established firms. As the literature will show below, it is important for dominant firms to recognize and assess the potential of this threat.

Semiconductor technology supply chain

Fundamentally, the dominant supply chain for the semiconductor industry is typified by Intel (see figure 4.1). For logic processors, Intel designs chips based on its instruction set architecture. It also designs the packaging that houses the chip and provides the interface with the computer architecture. Intel designs the process used to manufacture the chips, as well. While it purchases equipment from vendors, integrated firms often significantly modify that equipment for its given tasks using their base of proprietary knowledge.

Figure 4.1-Integrated Supply Chain

This helps to create and maintain the competitive advantage Intel gains from producing the chips themselves. Integrated firms package chips using the same philosophy of running proprietary processes on standard industry equipment.
The capital required for a fabrication and packaging facilities is on the order of $1.5 billion to $2 billion. Recouping this investment requires very high utilization and high throughput. To achieve these throughputs, the large semiconductor manufacturers use batch sizes of around twenty five wafers. Usually, tooling is built to accommodate such batch sizes, but, as will be discussed later, there are exceptions to this.

Along with equipment and manufacturing processes, semiconductor firms have developed methods for learning about the process to drive process improvements. A typical system will be presented in chapter three. One should note here, however, that the learning system encompasses the range of the supply chain from product design through production to packaging.

**Growth of the dis-integrated supply chain**

While microprocessors and DRAMs are produced in very large volumes, there are many applications of semiconductor processes. Those include everything from digital signal processors to home blood pressure kits. Manufacturers of these devices can find capital investment and production costs too high to risk on an unproven product. Also, semiconductor manufacturing processes are evolving to smaller and smaller geometries, and thus lower costs, as rapidly as the products themselves. Therefore, long-term competitiveness would require constantly updating process equipment sets.

![Dis-integrated Supply Chain Diagram](image)

Figure 4.2-Dis-integrated Supply Chain
Facing this very difficult supply problem, a particular device company might consider how to approximate a semiconductor manufacturing supply chain. A device company typically has strong design resources, and may or may not dictate the architecture of the instructions on which the device will operate. To have the device produced, the firm must communicate product design requirements with the foundry. For the integrated semiconductor firm, design rules and informal relationships within the firm provide this interface. This is a simpler proposition when there are a relatively small number of products to be made. For a flexible foundry, these interfaces are likely to be complicated by numerous device designs and many different companies. Once the device has been fabricated, the device must be packaged and tested. Again, there needs to be some type of communication among the parties to assure that product specifications are met.

Semiconductor manufacturing facilities for these industry segments have been provided by foundries such as Taiwan Semiconductor Manufacturing Company. TSMC maintains process capability and scale comparable with industry leaders. Therefore a device company can design its own circuits and contract with TSMC for production. This contract not only provides capacity, but it provides process expertise and a consistently competitive process capability. For instance, a new TSMC fab under construction in the United States will have the capability to produce features with dimensions as small as 0.18 microns.

After several years of production experience, the firms that cooperate to form these supply chains have employed different approaches to the interface challenge. Early on, the firms tended to develop partnerships across engineering teams of the firms. Design engineers at the device company would work with process engineers at the foundry to move the device into the production phase. Recently, foundries have worked with circuit design software companies to develop a standardized set of libraries to make device designs more portable from the desks of the device firm's design staff to the tools of the foundry. This helps to standardize the interface between design and manufacturing, and makes the two steps in production more interchangeable. TSMC has entered into partnership agreements with device packaging firms to complete the supply chain.

To compete, the virtual fab arrangement described above must also develop a learning system to at least match the rate of improvement, or trajectory, in cost and performance of their integrated
counters. The partnerships between the design firms and fabrication firms described above provides the organizational foundation for a learning process. The virtual fab must then design the optimal learning system. Assessing the actual cost/performance trajectory is described below.

**Evolution of the semiconductor fabrication process**

As described above, dominant firms of the semiconductor industry process wafers in batches. In the past several years, there has been significant research into transitions from batch processing to single wafer flow. Equipment manufacturers now produce "cluster tools" that consist of three of more modular processing chambers linked by a wafer handling system. The interface of the chambers is standardized such that, as processes evolve, it can be changed at much lower costs.

Semiconductor foundries, such as Taiwan Semiconductor Manufacturing Company, might include single wafer flow as part of their competitive strategy. Langlois (1992) suggests that this approach is well suited to flexible, low-volume production. Since research in single-wafer flow began, dominant firms have transitioned from 6-inch to 8-inch and now to 12-inch wafer processing. All of these process generations continue to be based on batches, however. While cluster tools are common in these fabs, this is driven more by process physics than production economics. This is because the demanding environments of semiconductor manufacturing are easier to maintain in small chambers versus larger ones. Still, the dominant firms often use cluster tools to process wafers within a batch in parallel rather than in series, as single-wafer flow would suggest.

Making the transition from batch flow to single wafer flow for high volumes would certainly be consistent with the advice of some who have studied manufacturing in detail. Womack (1996) advocates single piece flow as part of a system of lean production, or production with very little waste. He cites anecdotes from several industries, including automobiles and aerospace. Bergendahl (1990) suggests that single piece flow in semiconductor manufacturing will accelerate the process of cognitive learning described in this thesis. By processing wafers one at a time, problems are detected on a single wafer, and could potentially be eliminated on subsequent wafers. In batch processing, all of the wafers within a batch would presumably be affected by the problem.

49
Utterback (1994) has examined the dynamics of innovation in numerous industries. He divides products into two general types: assembled and non-assembled. He concludes from this research that discontinuous change in process architecture, such as the transition from batch to single-wafer flow, has much more impact on non-assembled products than assembled ones. So are VLSI circuits assembled or non-assembled? Langlois draws a parallel between semiconductor manufacturing and chemical manufacturing, implying that VLSI circuits are non-assembled. One could argue, however, that the several hundred processing steps for a semiconductor device better fits Utterback's description of an assembled product.

**Research into supply chain dis-integration and competitiveness**

Christensen (1994) has explored cycles of vertical integration in the disk drive industry. Early in the industry, it was dominated a few vertically integrated firms such as IBM. Through the 1970's, however, disk packs for the drives became modular, and the sub-components came to be supplied by independent firms. The demand for high volumes of computers at low costs drove standardization of more components. By the late 1980's, the leaders of the disk drive industry were disintegrated firms that, Christensen says, had three advantages: faster product design cycles, rapid incorporation of new component technology, and innovative system designs.

By the late 1980's through the early 1990's, the trend reversed. High disk drive densities now called for new technologies such as control circuitry and thin film recording heads. Incorporating these technologies into the drive architecture drove designs to be more integral. Christensen presents evidence that integral designs, designs where interfaces between components are tightly coupled, drive firms toward vertical integration. Secondly, he argues, economies of scale were such that drive assembly was highly competitive with thin profits. Some firms were driven to integrate backwards into components where scale was more of a factor in profitability.

Applying Christensen's findings to the semiconductor manufacturing industry requires drawing parallels between product architecture of disk drives and process architecture of VLSI circuit production. The recording heads and actuator motors of the disk drive are paralleled by the tooling of the production process and how inventory is managed within the process. If these components of the supply chain can be modularized, then industry can move away from vertical integration. These components can be modularized by standardizing the interface between them.
The software firms who provide design tools to smooth the transition of a device design into a semiconductor foundry would be an example of modularization. Also, process innovation within the individual firms that is independent of the overall supply chain can begin.

The extent to which supply chain disintegration can continue is limited. Semiconductor devices are by nature integral, and accelerating cost/performance trends are reinforcing product integrality. High processing speeds require that the circuits for different parts of a microprocessor be closely located. Also, substantial interaction among layers and modules of the device requires close integration of the fabrication process. The emerging virtual semiconductor firms have struck upon available opportunities, the separation of design, fabrication and assembly. Device firms can move from one foundry to the next chasing improvements in process performance, and they can choose the most appropriate packaging technology for their application. Within the standardized interfaces of this supply chain architecture, firms are free to innovate and find more efficient combinations of design, fabrication, and packaging to maximize cost/performance trajectories.

Langlois (1992) suggests another possible advantage of this system. He studied development of standards for semiconductor manufacturing equipment and contends that such standardization serves as a partial substitute for internal integration of production and long-term relationships between equipment suppliers and the firm. Extending his observations to the overall supply chain for a semiconductor device suggests that dis-integrated device manufacturers can achieve economies of scope through standardization of the interfaces between design, fabrication, and packaging. This helps to make the dis-integrated supply chains competitive with integrated competitors. The work of Langlois and Christensen would also suggests that this approach can create external economies of scope by allowing for increased innovation within the stages, as resources do not need to be committed to integrating the links in the supply chain.

While dis-integrated supply chains allow more competition within firms, by itself it only facilitates competition along existing cost/performance curves. Foundries can pursue continuous improvement along with existing integrated firms. Leapfrogging the competition requires a leap in process technology that is now facilitated by a shift in integration trends within an industry.
**Implications on learning and competitiveness**

The research cited above suggests that evolving dis-integrated semiconductor alliances can have the opportunity to develop lower cost production technologies based on concepts such as single wafer flow. Should dominant firms be concerned with this? These firms achieved their dominance through intense competition to supply the growing computer market, and they earn high margins doing so. The trajectory of improvement in price and performance is growing faster than the ability of software companies to use it. Some argue that there is no rational reason to seek a radical departure from this proven system of production.

![How to Assess Disruptive Technologies](image)

**Figure 4.3—from Bower and Christensen (1994)**

Bower and Christensen (1994) describe numerous industries, from disk drives to retailers, where the same cycle of success kept dominant firms from catching the next wave of innovation in their industry. They assert that firms are consistently looking up-market, and defining a cost/performance trajectory consistent with the needs of their current customers. New entrants with new but potentially competitive technologies enter the market, usually by finding new market segments. The needs of existing customers are met by the dominant firms. In some cases, the cost/performance trajectory of the new technologies is higher than that of existing
firms, and eventually existing technologies are overtaken. This kind of effect is depicted in Figure 4.3. The cost/performance position for the disruptive technology begins well below that of the established technology. The disruptive technology, however, is on a higher cost/performance trajectory than established technology and eventually overtakes it.

Bower and Christensen suggest that there is a method to identify and cultivate disruptive technologies. Dominant firms in the semiconductor manufacturing industry might apply this method as follows:

**Determine whether the technology is disruptive or sustaining.** A change from batch flow to single-wafer flow would certainly be disruptive. Currently, most process engineers are responsible for a single process technology, such as chemical vapor deposition. If cluster tools are used to produce layer using single-wafer flow, process engineers may be organized around layers or modules of the product rather than individual processes. Inventory management systems would be significantly different, as well.

**Define the strategic significance of the disruptive technology.** The key here is determining tomorrow’s needs. If the dominance of established firms is threatened and margins begin to fall, cost/performance trajectories may need to be higher than today to remain competitive. If the personal computer market begins to saturate, what customers will drive new growth in semiconductors, and what will their performance needs be? If single-wafer flow offers an accelerated trajectory satisfying these scenarios, then it may indeed be strategically significant.

**Locate the initial market for the disruptive technology.** This may be a bit tricky, but Bower and Christensen suggest that adopting a second-mover strategy may be a good one. If TSCM is using single-wafer flow, who are their customers and what are their needs. Dominant firms would have to act quickly, however. Research suggests that there may be a limited window of opportunity for adopting innovation (Christensen, 1996).

**Place responsibility for building a disruptive technology business in an independent organization.** If the profit margins of the new markets where this technology is applied
are lower than existing markets, success can look like failure. An independent organization is free to define success apart from the existing business. Keeping that organization independent allows the firm to kill off the existing condition if market conditions allow. If the firm does not, competitors inevitably will.

Conclusions

From the research cited above, one can better summarize the threat posed by newer entrants in semiconductor manufacturing. First, the virtual supply chain frees firms to match superior device designs with superior manufacturing processes to maximize the threat to incumbent firms. If the interfaces of the virtual supply chain are sufficient well developed to allow individual firms to concentrate on their core competency, the firms realize positive externalities. Finally, the transition to single wafer flow is one method that manufacturing firms could use to develop learning systems superior to their dominant rivals.

Chapter three of this thesis will describe the learning processes that underlie the current performance trajectory in semiconductor manufacturing, and chapter four describe methods for analyzing changes to the learning process. Chapter five draw conclusions about the future of the semiconductor industry by combining results from chapter four and industry trends described above.
Chapter Five — Summary and Conclusions

**Recommendations for the memory fab**

The benchmarking data that was presented in chapter two suggests that the memory device fab that was the object of this study should pursue an in-line inspection methodology to achieve world-class die yield. The results of the simulation model from chapter three suggests that twenty-five percent decrease in cycle time would have greatly improve die yield ramp, and output would increase by six times the benefit from other scenarios. However, this fab is already at world-class cycle times, and lower cycle times further could expose the fab to lower levels of output. The in-line inspection methodology would have a relatively modest increase in output of about equal to other potential options, such as the output increase expected for acquisition of a new manufacturing tool, although the capital investment required for in-line inspection could be far less than that required for a new tool.

The research on organizing for learning presented in chapter two suggests that implementation of the new system will be critical issue. The management of this fab should do the following to assure effective implementation:

- Identify the systemic impact of this technological change on the capabilities of the fab’s technical staff. The capabilities required will include both detailed knowledge of new technologies, such as analytical software and information systems, as well as detailed knowledge of the current yield learning system.

- Identify changes in organizational relationships necessary to make the new system work. To dedicate the additional resources required for integrating defect data and fault data, for example, other fabs have formed teams of manufacturing technicians and yield technicians.

This fab has spent substantial resources in the development of this new yield system. It should complement that investment by dedicating engineering hours to the integration of the new technology with the existing system for yield improvements. Otherwise, the gains projected by the simulation model may not be realized.
For the longer run, the discussion of chapter four would suggest that the firm might consider the potential disruptiveness of single wafer flow at semiconductor foundries. Single wafer flow would be very disruptive to the existing process architecture of integrated firms. The strategic significance should explored by this firm and potential markets such as digital signal processors should be identified. Finally, if the cost/performance trajectory of single-wafer flow appears to be higher than that of batch processing, the firm should move this technology outside of its existing organizations to pursue these markets.

**Evolving the architecture of the learning process**

Technologies are evolving to allow for more rapid feedback from in-line inspection systems. The typical technology for detecting defects in-line measures wafers after particles have been deposited. Sensor systems are being developed to measure the generation of potentially contaminating particles in-situ. These particles can be detected before they are deposited. Using this technology, engineers can estimate defect densities before they are actually measured at an inspection station down the line. They could also be used to better determine when the tool should be taken off-line for maintenance. The model described in chapter three could be modified to simulate such a system by treating each tool with in-situ defect monitors as an individual stage within the simulation model. The benefit of this system could then be quantified and compared with the current system.

Even with a learning process based on in-situ defect data, it is still feedback. Defects that exist on the wafer are statistically correlated to electrical faults. The deposition of a defect signals the potential for a fault, and its root cause can be eliminated sooner. The makers of metrology tools have developed new systems of metrology and information system to shorten the time required for the feedback loop (Singh 1997). While this approach is certainly valid, it offers only diminishing returns. Making additional reductions in the feedback loop requires more investment in sensors, information systems, and analysis tools. A feedback loop can only be reduced to some finite amount above zero.

Bohn (1994) suggests that as the state of process knowledge matures, the goal should be to develop models that allow for feed-forward. A feed-forward methodology is compared schematically with a feedback methodology in figure 5.1. A feed-forward paradigm would require detailed knowledge of what variables effect the deposition of defects that eventually
result in faults. Knowledge of this level of sophistication does exist to some extent. For instance, in-situ defect monitors could be used to anticipate the need for bring tools down for maintenance.

Extending this concept, engineers can learn that, for a particular tool, the rate at which defects are depositied will depend on the incoming wafer state, quality of raw materials, and the time since the latest tool maintenance. One can then determine, using process data, how many faults one can expect from a given process. There has been work on finding sensitivity of a circuit to defects with respect to structural parameters (such as layer thickness, line widths, line spacing, etc.). If one knows the relationship of tool and wafer states to defects at a particular level, that information can be used to tune the process to build structures that are less sensitive to that particular defect. For the example in figure 5.1, information from tool j could be used to set process parameters to minimize defects at a tool b steps down the line.

Extending Yield Improvement Methodologies

Feedforward Methodology

![Diagram of Feedforward Yield Methodology]

Figure 5.1—Feedforward Yield Methodology

Again, the pump must be primed with knowledge that accumulates from the older methodologies. It also requires a level of process knowledge that is unusual in the industry. If such a process could be instituted, however, a mature process may no longer need in-line metrology. The process would tune itself through the feed-forward mechanism, and throughput times would fall because metrology steps would become a diminishing requirement.
Extending this analysis to other industries

The previous chapters have shown that processes designed to accelerate cost/performance curves in semiconductor manufacturing have a direct impact on firm competitiveness. Accelerating cost/performance curves requires, among other activities, redesigning the learning process to take advantage of process knowledge and technological capabilities. New process designs can be simulated to assess the financial and competitive value of various process architectures. Simulation results are also useful to managers as part of their decision support process.

While this model has been described in the context of semiconductor manufacturing, the same techniques can be transferred to other industries. First, a firm should determine its own cost/performance trajectory and importance of the firm's learning system relative to other cost/performance factors. Then, determine competitive threats as described in chapter 4. Identify the constituent activities of the learning process, and develop a simple model that can be used to test difference process architectures.
References


THESIS PROCESSING SLIP

FIXED FIELD:  ill. __________________ name __________

index __________ biblio __________

► COPIES: Archives Aero Dewey Eng Hum

Lindgren Music Rotch Science

TITLE VARIES: □

NAME VARIES: ☒ Hardin middle name

IMPRINT: (COPYRIGHT)

► COLLATION: 61 p

► ADD. DEGREE: M.S.  ► DEPT.: M.E.

SUPERVISORS: __________________

__________________________

__________________________

__________________________

__________________________

__________________________

NOTES:

__________________________

________ ______
cat'r. date:

 ► DEPT: MgT [Page: 3131 + 202]

 ► YEAR: 1997  ► DEGREE: M.S.

 ► NAME: INGLE, Timothy H.