# Capacity Loss Factors in Semiconductor Manufacturing

*by:*

***Jennifer Robinson*[1], *John Fowler*[2], *and Eileen Neacy*[3]**

*Abstract*

This paper describes characteristics and problems of the capacity planning process in semiconductor wafer fabrication facilities. Twenty-two factors that contribute to capacity loss are identified and discussed. Informaton on these loss factors was obtained through three sources: 1) a literature review; 2) an extensive survey, interview, and workshop process; and 3) a variety of queueing and simulation models.

## 1. INTRODUCTION

Every year more products, from kitchen appliances to automobiles, use sophisticated electronic technology. Computer sales have exploded. People have started spending more time on information superhighways than on real highways. Central to the advancement of all these areas is the efficient production of semiconductors. SEMATECH is a consortium founded by the US government, in cooperation with industry, to help the US to maintain a competitive position in the semiconductor manufacturing industry. JESSI is a similar cooperative effort between European partners. JESSI and SEMATECH have been cooperating in pre-competitive areas, with the expectation of advancing the state-of-the-art in manufacturing for both the US and the European Union (EU). One of the first joint efforts between SEMATECH and JESSI was a project to identify the factors contributing to capacity loss in semiconductor wafer fabrication facilities (fabs), and to improve methods of planning capacity. This paper discusses the relative importance of various semiconductor capacity loss factors, as identified by the Measurement and Improvement of Manufacturing Capacity (MIMAC) project.

[1]FabTime Inc., 325M Sharon Park Drive, #219, Menlo Park, CA 94025. www.FabTime.com
[2] Department of Industrial and Management Systems Engineering, Arizona State University, Tempe, AZ.
[3] Motorola Corporation.

 © 2003 by Jennifer K. Robinson. All rights reserved. 

# 2. PLANNING CAPACITY

### *What is Capacity Planning?*

Capacity planning, in its broadest sense, encompasses all decisions about what products a company can and should produce, and what facilities will be required to produce it. This includes long-range business decisions such as:

- Should the company enter new markets?
- Should the company build new factories, or close existing factories?
- What should be the equipment set for new factories?
- Should the process mix in a particular factory be changed?
- Should certain products be outsourced?

Capacity planning also addresses shorter-term, more strategic questions, such as:

- How many wafers can the factory produce if the mix changes?
- What additional equipment will be required if a new product is introduced into the factory?
- What impact will an operational change (for example, a setup minimization strategy) have on the throughput of a particular workstation?
- How fast can a particular product be ramped to full production?

### *Why is Capacity Planning Important?*

Planning capacity accurately is critical in today's highly competitive semiconductor industry. Equipment costs are rising, technologies are changing, and customers are demanding ever faster chips. And they want them yesterday. With some pieces of equipment costing several million dollars each, capacity planning decisions have the potential to make an immediate impact on the bottom line. Understanding capacity is also critical to maintaining profitability over time. In many cases, the demand for a fab's product is greater than the ability to meet that demand. This implies a significant penalty, in terms of lost revenue, for planning to load a factory at a level lower than its current capacity. On the other hand, significant negative consequences can stem from overloading a factory. These outcomes include long cycle times, missed delivery dates, excessive inventory, and possibly lower yields.

### *What Makes Capacity Planning Difficult?*

In addition to being a critical task, capacity planning can be a difficult one. A capacity planner must address conflicting priorities within the organization. He or she must identify trade-offs between capital costs and cycle time, for example, to aid management in making decisions. Various pressures are brought to bear to encourage higher, more profitable capacity projections, as opposed to lower, perhaps more realistic expectations. In general, underestimating capacity is safer than overestimating capacity. However, as fabs become more expensive, companies are less and less able to afford the consequences of overly conservative targets. Planning capacity tends to be a high risk enterprise.

Further complicating matters for the capacity planner is the constantly changing environment in a wafer fab. Product demands change. Process flows change. Mix

 © 2003 by Jennifer K. Robinson. All rights reserved. 

changes. Tools break down. Yield fluctuates. Start projections that looked reasonable last month are likely to be out-of-date. Furthermore, the business environment frequently changes. As a result, a single analysis typically requires evaluating many “what if” scenarios.

Finally, capacity planners often have trouble obtaining and maintaining accurate data. For newer fabs, the data simply is not available. Even for mature fabs, few companies use automated procedures for extracting capacity planning data from the shop floor control system. In many cases, industrial engineers are out on the floor with stop watches, or are calling equipment engineers, asking them what they think the downtime percentages are for a particular tool. Even when data is automatically extracted from the shop floor control system, some question how accurately the data was entered in the first place. Efforts to make the data more accurate are likely to be expensive. Problems with data appear to be nearly universal.

### *How Are People Planning Capacity Today?*

Most people in the semiconductor industry today do their capacity planning with spreadsheets. In some cases, the spreadsheets are supplemented with other techniques such as queueing or simulation models. Generally, the different models are not linked to one another electronically. This means that data must be maintained (usually manually) in multiple locations. In many cases, different people in different organizations maintain the spreadsheet and simulation models, and the data is not necessarily even consistent between them. This, of course, compounds the data problems described above.

A wide variety of complexity exists in current capacity planning models. Some include only a few simple formulas. Others have elaborate macros for easing data entry and conducting “what if” analyses. Some include only overall line yields, while others detail yield loss and rework probabilities by step. In general, however, most semiconductor capacity planning spreadsheets require similar inputs and outputs, and perform the same basic types of calculations.

Typically, capacity planning calculations are performed by treating each group of identical tools in isolation. For each tool group, the capacity starts out at 100% (usually 24 hours per day, seven days per week). This capacity is then downrated for capacity loss factors such as breakdowns, preventive maintenance, engineering time, and setups. Each loss factor is usually expressed as a percentage, and subtracted from 100%. The loss factors are usually based on the projections or experience of people who work with the tool. Often, however, loss factors are a source of negotiation between different levels of the organization. For example, the fab manager might push for no more than 20% equipment downtime in the model (including random failures and preventive maintenance), even as the equipment engineer protests that the tool will operate with closer to 30% downtime.

After the loss factors are subtracted from the 100% capacity, the amount remaining is usually multiplied by a contingency factor ranging from 75% to 90%. This results in a planned idle time percentage on the tools of 25% down to 10%. The contingency factor summarizes management’s intuition that equipment cannot be operated with no idle time in the unreliable environment of a wafer fab. Some fabs refer to the contingency factor as a “cycle time factor,” because tools with lower utilization will usually have lower cycle times. Others call it a “variability factor,” or “catch-up capacity.”

 © 2003 by Jennifer K. Robinson. All rights reserved. 

Typically, the same contingency factor is used for all tool groups, although some fabs use tighter contingency factors for more expensive bottleneck tools.

Once the loss and contingency factors are accounted for, the amount available for production remains. Process flow data is then used to determine the machine time required at each tool group for a particular mix. These processing requirements are then used to find either the number of tools of each type needed to produce a target output, or the amount of wafers that can be produced by each tool group. The tool group with the lowest capacity limits the capacity of the factory.

Most people feel that this method of planning capacity is 'fairly accurate.' It has the advantages of being fast, easy to interpret, and easy to use. However, people often see room for improvement. In particular, the method does not capture any dynamic behavior such as interactions between tool groups. It also does not provide estimates of work-in-process and cycle time. Some people iteratively use a simulation model with a spreadsheet model. The spreadsheet model provides a lower bound on the required toolset for a given product mix. Tools are then added to the simulation model until projected cycle times become acceptable. People using this procedure tend to perceive their results as more accurate than do people using spreadsheets alone. For additional discussion of capacity planning, see Neacy *et. al.* [54], Spence and Welter [65], or Karmarkar *et. al.* [39].

# 3. MEASURING PERFORMANCE

### *How Do People Measure Factory Performance?*

The performance of semiconductor fabs is constantly being evaluated. Factory throughput, moves, work-in-process (WIP), cycle time, yield, wafer cost, machine utilization, on time delivery, machine availability, overall equipment effectiveness, and linearity of shipments are just some of the measures used. The relative importance of factory performance measures depends on the product market and the company philosophy. High volume, single product factories tend to focus more on equipment utilization, while custom producers rely heavily on cycle time. However, a survey of over 100 people from various US and European factories found that cycle time, on time delivery, wafer cost, line yield, and number of good die per wafer were most often cited among the most important performance metrics [54]. Among the earliest surveys distributed, cycle time was named most frequently. Among the later surveys, on time delivery dominated. The importance of cycle times in today's competitive environment was apparent from these results.

In most factories, evaluating performance is a separate activity from planning capacity. Part of the reason for this is the long lead time required for buying new pieces of equipment and bringing them up to speed. In the rapidly changing environment of the fab, by the time an equipment set is up and running, the product mix or processing requirements are likely to differ from those that drove the original equipment purchases. This makes comparisons difficult. Another problem is that traditional capacity planning tools (mostly spreadsheets) do not directly estimate cycle times. In many cases the capacity planner, without access to cycle time estimates, designs a particular tool set. Production is then responsible for using that tool set to meet some pre-specified cycle time target.

 © 2003 by Jennifer K. Robinson. All rights reserved. 

The above approach does not take into account the physical relationship between cycle time and start rate for a given factory (discussed in detail in Section 4). This can set production up for failure, by requiring a cycle time target that cannot be met with the given tool set and start rates. Because of the negative consequences of this sequence of events, an increasing number of factories are using simulation to decide whether a planned capacity corresponds with "reasonable" cycle times. Few companies have formalized this type of analysis, however, and the cycle time component of the analysis is frequently omitted due to time and budget constraints.

For additional information on performance measures in semiconductor manufacturing, see Baudin *et. al.* [3], the Competitive Semiconductor Manufacturing Study Report [10], Burman *et. al.* [9], or Hicks [29]. For examples of simulation used to estimate factory performance, see Berlow *et. al.* [5], Miller [49], New *et. al.* [55], or Hood [30]. For examples of queueing models used to estimate fab performance, see Connors *et. al.* [11], Whitt [74], or Inoue and Yoneda [34].

# 4. UNDERSTANDING CAPACITY AND IMPROVING PLANNING

### *Can Current Capacity Planning Methods Be Improved?*

The primary goal of the Measurement and Improvement of Manufacturing Capacity (MIMAC) project was to understand the factors that contribute to capacity loss. Another goal was to improve methods for planning capacity. To tackle these questions, a project team was assembled that consisted of engineers from JESSI, SEMATECH, and JESSI and SEMATECH member companies, as well as researchers from several universities. Most members of the team did not work directly in a manufacturing environment. To better understand the capacity planning process, and to ensure that the results obtained through the study would be applicable, the team commenced with an extensive survey and interview process. They sent written surveys to fab managers, operations managers, capacity planners, production controllers, and shift supervisors from all of the SEMATECH and JESSI member companies. They also conducted on-site interviews at several companies, to ask more in-depth questions than could be covered with the written form. The surveys asked about capacity planning, performance measures, and the impact of specific loss factors on fab capacity. Specifically, participants were asked to rank a list of 22 factors in terms of their effect on capacity. The 22 factors originally listed are shown in Figure 4.1. The results of the survey are discussed in detail in [54].


© 2003 by Jennifer K. Robinson. All rights reserved.

| Tool Dedication | Batching Policy |
|---|---|
| Unscheduled Maintenance | Dispatch/Sequencing |
| End-of-Shift Effect | Factory Shutdown |
| Hot Lots / Engineering Lots | Inspection |
| Operator Cross-Training | Lot Size |
| Operator Availability | Order Release |
| Preventive Maintenance | Product Mix |
| Reentrant Flow | Rework |
| Setup | Shift Plans |
| Time Constraints Between Steps | Lack of Tool Redundancy |
| WIP Control Strategy | Yield |

*Figure 4.1 Capacity Loss Factors in Semiconductor Manufacturing*

In parallel with the survey effort, the MIMAC team collected an extensive bibliography of research concerned with capacity planning in semiconductor manufacturing, and with the 22 capacity loss factors [23]. Following the survey and literature review, team members studied each of the factors in isolation, using small simulation and queueing models. These experiments were known as the MIMAC local effect experiments.

The team then conducted an extensive set of designed simulation experiments to investigate the factory-level impact of several of the factors. They used a set of factory-level datasets, assembled by SEMATECH as part of a separate effort. European data was added and validated under the MIMAC project. The purpose of collecting the datasets was to aid academics and suppliers in developing new models and tools for industry. The datasets contain actual manufacturing data from both ASIC and logic wafer fabrication facilities, organized into a standard format. They include no real product names, company names or other nomenclature that could serve to identify the source of the data. Each dataset contains the minimum information necessary to model a factory, including: product routings and processing times; rework routings; equipment availability; operator availability; and product starts. For a more detailed description of the datasets, refer to Fowler, Leachman, and Feigin [19]. The methodology used in the designed experiments, and the results obtained, are described in [1] and [21].

The focus of Section 5 of this paper is the impact of the factors, as evaluated through the survey process, the literature review, the local effect experiments, and the factory-level designed experiments.

### *Can Capacity Planning Be More Closely Linked with Factory Performance?*

Cycle time emerged as the number one performance metric for MIMAC survey and interview respondents. Partly in response to this, the MIMAC team investigated the relationship between capacity, variability, and manufacturing cycle time. It is commonly known that work-in-process (WIP) levels and cycle times increase as factory output rate increases. At the output rate that drives the bottleneck of the factory to 100% utilization, the cycle times and WIP become infinite. The rate at which the cycle times and WIP

 © 2003 by Jennifer K. Robinson. All rights reserved. 

increase depends upon the amount of variability in the system. An example of cycle time vs. start rate curves for a two-product logic factory under different amounts of variability is shown in Figure 4.2. The relationship between capacity and cycle time has been well documented by such authors as Bitran and Tirupati [6], Dayhoff and Atherton [13], Fordyce and Sullivan [18], Fromm [24], Hopp *et. al.* [32], and Najmi [50].

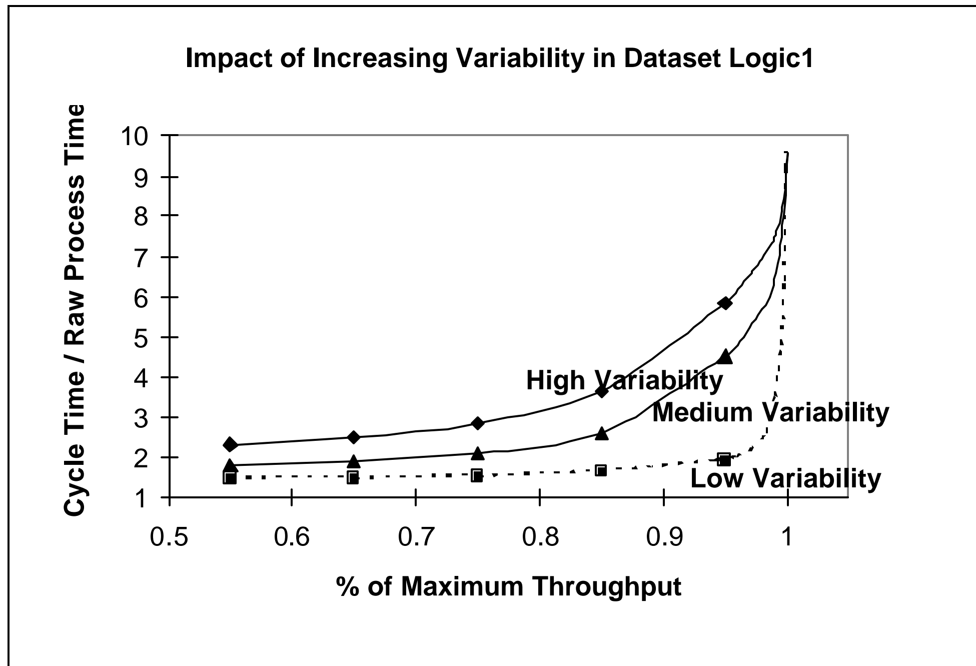


*Figure 4.2 The relationship between cycle time, throughput, and variability in a two product logic factory, one of the SEMATECH testbed datasets.*

Traditional capacity planning methods have focused on finding the maximum capacity that drives the bottleneck to 100% utilization, despite the fact that infinite cycle times are infeasible in practice. Leonovich [45] proposed optimizing WIP, cycle time, and output rate by defining the relationship between WIP and output rate, and then selecting the output rate that corresponded to some WIP target. Spence and Welter [65] used the cycle time versus throughput trade-off curve to define the operational capacity of a given factory. Martin [47], in an independent effort, imposed a cycle time requirement on a manufacturing line to improve capacity planning.

The MIMAC project team defined the cycle-time constrained capacity of a factory as the maximum output rate that a system could achieve for a given output mix, under a constraint on the average cycle time. They computed this capacity by drawing the characteristic curve of cycle time versus output rate for the factory, and finding the output rate that corresponded to the desired cycle time constraint. An example is shown in Figure 4.3. Under this definition, cycle times are usually expressed as multiples of the weighted average raw processing time (RPT). So, for example, a cycle time constraint of twice the average raw processing time would be referred to as a 2X RPT constraint. The output rate corresponding to this constraint would then be called the 2X capacity. For a more theoretical discussion of this method, refer to [1]. Cycle time constrained capacity was used as the performance measure for most of the MIMAC studied.

 © 2003 by Jennifer K. Robinson. All rights reserved. 

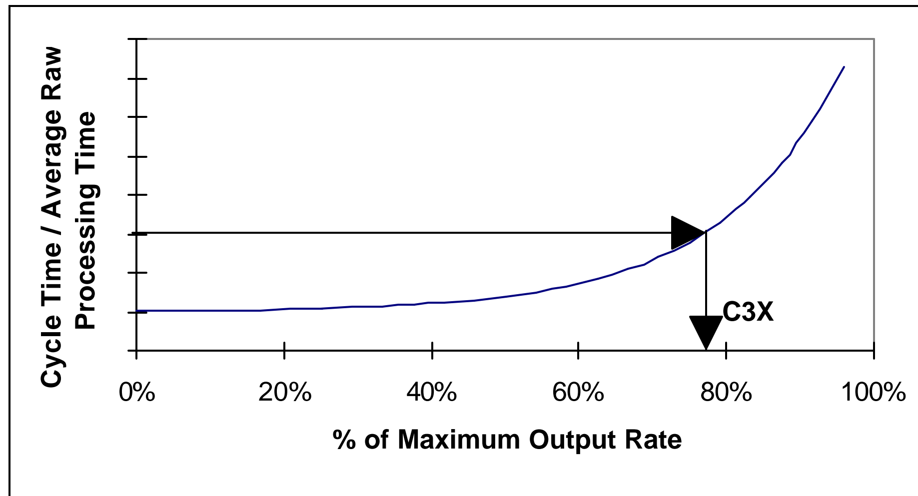


*Figure 4.3 The 3X cycle time constrained capacity for a system*

Cycle time constrained capacity is related to the contingency factor used in many spreadsheet capacity planning models. Imposing a contingency factor of 80%, for example, is equivalent to drawing a vertical line on Figure 4.3 that intersects the x-axis at 0.8. The primary difference in the two methods is that with cycle time constrained capacity, the cycle time target drives the maximum system capacity. With the traditional method, the contingency selected drives the resulting cycle time. The only way to impose a specific cycle time goal is by trial and error. The MIMAC team has proposed using the definition of cycle time constrained capacity as a means of including cycle time goals in the capacity planning process. To find the cycle time constrained capacity of a given toolset, capacity planners can use a simulation or queueing model to generate the characteristic curve. They can then read across and down to find the capacity that will allow them to meet a given cycle time target. This closes the loop between planning capacity and measuring performance, and allows factories to set more realistic targets.

# 5. RESULTS CONCERNING LOSS FACTORS

The MIMAC team, discovering several strongly overlapping areas of analysis, combined several pairs of factors into single items. Unscheduled maintenance and preventive maintenance were combined into a category called equipment downtime. Inspection was grouped with yield, end of shift effect with shift plans, and WIP control strategy with order release policy. The team found, through their surveys and experiments, that five of the remaining 18 factors tend to cause the majority of capacity loss in wafer fabs. These are equipment downtime, yield loss, setup, dispatch rule, and batching policy. The first three are measurable loss factors that are generally included in capacity planning spreadsheets. The last two are control policies that can have a strong influence on the other loss factors, particularly in a cycle time constrained environment. Each is described in detail in Section 5.1. Section 5.2 discusses several other, relatively controllable, loss factors. Other factors that are more inherent in semiconductor manufacturing are reviewed in Section 5.3.

 © 2003 by Jennifer K. Robinson. All rights reserved. 

## 5.1 The five biggest loss factors

### *Equipment Downtime*

Unscheduled downtime was ranked by MIMAC survey respondents, on average, as the most significant capacity loss in wafer fabs. Simulation experiments conducted using four factory-level datasets from real fabs confirmed that equipment failures contributed to significant capacity loss for these factories. See [1] for details. This result, by itself, is not surprising. Consider a single tool. If the tool is down for 20% of the time, its capacity to produce is reduced by that 20%. This is equivalent to moving the asymptote of the characteristic curve for the tool from 100% to 80%, and shrinking the rest of the curve accordingly. Changing the curve in this manner is equivalent to using a 20% loss factor for equipment downtime in a spreadsheet-based capacity planning model. Downtime is virtually always included in spreadsheet-based capacity planning models, although the downtime percentages for individual tools can be sources of considerable negotiation within an organization.

What is less commonly considered in planning is that equipment downtime can also dramatically change the shape of the characteristic curve. Both random failures and preventive maintenance events increase the variability of the processing times experienced by lots arriving at machines. That is, if a lot arrives at a down machine, the time spent waiting for the machine to come back up becomes effectively part of the lot's processing time. Because variability in processing times increases lot cycle times, downtime moves the characteristic curve upward. This in turn decreases the cycle time constrained capacity of the fab. The magnitude of this effect is influenced by the amount of variability in the downtimes. The MIMAC studies showed that more variable downtime distributions (e.g. exponential) resulted in lower cycle time constrained capacities than less variable distributions (such as triangular or even constant distributions). Similarly, for the same downtime percentage, shorter, more frequent failures result in higher cycle time constrained capacities than do longer, less frequent failures. This is because the long repair times introduce more variability into the system. The implications of this for fab managers suggest (where possible) scheduling more frequent, shorter maintenance events, rather than taking a machine down for a very long period of time.

For other research on equipment downtime, see Baum and O'Donnell [4], Gurnani and Akella [27], Law [42], or Hopp and Wu [33].

### *Yield Loss*

Yield loss did not rank very high as a capacity loss factor in the MIMAC surveys. Apparently, most capacity planners do not consider yield loss a variable that they can control. In some cases, line yield is treated as a translation factor between total wafer starts and wafer outs. That is, a capacity is computed assuming no yield loss, and the final number is multiplied by the line yield to obtain the actual capacity. This is a conservative method, but results in underutilized tools near the end of the line. In other cases, individual step yields are accounted for in computing the processing requirements for each tool. This can be difficult to do, however, particularly in the presence of changing product mix. Often, to simplify computations, yield loss is assumed to occur linearly, with an equal probability of scrap at each step. Of course, when the actual yield loss is not linear, this method can lead to problems. Line yield and number of good die per wafer are both recognized as important performance measures for the fab, but they are not necessarily accounted for accurately in capacity planning.

 © 2003 by Jennifer K. Robinson. All rights reserved. 

The MIMAC experiments found that any yield loss after the bottleneck tool for a factory results in a corresponding loss in factory capacity. This is because time spent on the bottleneck tool processing the scrapped wafers cannot be recovered. Yield loss before the bottleneck increases variability in lot arrivals, and thus increases cycle times, and decreases cycle time constrained capacity. In the MIMAC designed experiments, yield loss was assessed by comparing the overall capacity of a fab to the overall capacity of the same fab with perfect yield and the same output mix. As an example, yield losses caused an average 10% reduction in 2X cycle time constrained capacity for a two-product logic fab with overall line yields of 84% and 87%. The unconstrained capacity loss observed was 7.5%.

For additional research on yield, see Tang [67], Lee and Yano [44], or Cunningham [12]. For a discussion on the relationship between yield and cycle time, see Wein [71]. For research on inspection in an environment with reentrant flows, see Narahari and Khan [51].

### *Setup*

Setup ranked in the top half of capacity loss factors in the MIMAC surveys. Most capacity planners include a specific loss factor for setups at several tool groups (e.g. implanters and steppers). The loss factor is usually based on experience with the tool. Sometimes it is based on simulation results. The actual amount of setup observed at a tool is a function of tool utilization, setup durations, product mix, equipment dedication, and dispatch rule. Loss factors based on historical data are often inaccurate, because product mix and equipment dedication strategies are subject to change. Part of the MIMAC research focused on deriving formulas to approximate the amount of setup at a workstation. The formulas account for product mix, setup lengths, number of tools per tool group, start rate, and operating policies. They can be included in capacity planning spreadsheets, to replace more subjective loss factors. A paper describing the setup approximations is in progress. For additional information, contact Jennifer Robinson or John Fowler.

Most people use some sort of setup reduction strategy at tools with significant setups. Setup avoidance policies are common. Under these policies, when a machine finishes processing a lot, the operator checks the queue for a lot that will not require a setup. Only when there are no such lots in the queue is the machine set up for a different product type. The factory-level experiments conducted under MIMAC found that setup avoidance policies significantly mitigate the negative capacity effect of setups (see [1] and also [61]). However, even under setup avoidance policies, setups lead to a measurable capacity loss. This is because time spent in setup is time lost to processing. The situation is particularly critical in semiconductor fabs because the tools with large setups are among the most expensive in the factory. Setups, like failures, increase the variability of lot processing times, and hence restrict cycle time constrained capacity.

For additional references on setup, see Dietrich *et. al.* [15] or Zhou and Egbelu [76].

### *Batching Policy*

Batch processing in semiconductor manufacturing usually refers to situations where more than one lot may be processed at one time on a given machine (such as a furnace operation). In a system which is not cycle time constrained, batching does not necessarily lead to capacity loss at the batch machine itself because it is usually possible to operate the machine with full batches. Capacity is more likely be lost (even

 © 2003 by Jennifer K. Robinson. All rights reserved. 

with no cycle time constraint) for a serial-batch-serial system. This type of loss occurs because the time spent forming the batch may lead to idle time at the second serial machine, time which cannot always be recovered. Imposing a cycle time constraint on any system may require making smaller batches. This effectively reduces the capacity of batch machines, and can result in batch machines limiting overall capacity.

The most common control policies used for batch machines are threshold policies. Under a threshold policy, a batch is started when the machine becomes idle and there are at least $L$ lots of some product available for processing. Two common thresholds are $L = 1$, known as the greedy policy, and $L$ = maximum batch size, called a full batch policy. In both cases, lots can typically only be processed together if they share certain characteristics, expressed by having a common batch I.D. Batching policy does not affect the maximum capacity of a tool. However, batching policy can have a significant effect on lot cycle times, and hence on cycle-time constrained capacity. At high traffic intensities, greedy and full batch policies result in similar cycle times, because most of the batches processed are full. However, at low traffic a full batch policy can result in increased cycle times, because operators must wait to form full batches. The MIMAC factory level experiments found that for ASIC fabs, where there were many different batch I.D.s, forcing full batches caused very large losses in cycle time constrained capacity.

For additional research on batching policy, see de Haut de Sigy [14], Gurnani *et. al.* [28], Fowler *et. al.* [20], Glassey and Weng [26], Weng and Leachman [73], Robinson *et. al.* [62], Tran-Gia and Schoemig [68], or Dobson and Nambimadom [16].

***Dispatch Policy***

Dispatch policies, by themselves, are not measurable capacity loss factors. However, they can strongly influence the capacity loss due to other factors such as setups and batching. They can also affect the amount of variability in a system, and thus change the shape of the characteristic curve. Experiments conducted using four factory-level datasets found that setup and batching dispatch policies can dramatically influence cycle time constrained capacity. Failure to use a setup avoidance policy at tools with large setups results in significantly decreased capacity. Similarly, forcing full batches at all tools degrades cycle time constrained capacity, particularly when many different products are present. Other dispatch rules, such as first-in-first-out (FIFO) and shortest processing time (SPT) have a much less significant effect on cycle time constrained capacity (and no effect on unconstrained capacity).

For additional references on dispatching, see Blackstone *et. al.* [7], O'Neil [57], Lu *et. al.* [46], Panwalkar and Iskander [58], Wein [72], or Johri [36]. For detailed results on the impact of dispatch rule on the testbed datasets, refer to Robinson *et. al.* [61].

## 5.2 Other loss factors that are somewhat controllable

Fab personnel have some level of control over several of the other loss factors studied under MIMAC. These include fab-level factors such as the number of hot lots in the fab, the timing of fab shutdowns and the order release policy followed. The number of operators available, the level of operator cross training observed, and the shift plan followed are also relatively controllable. At the tool level, in addition to the dispatch rules discussed in Section 5.1, fab personnel often decide about dedication strategies to follow. These factors are each discussed below.

 © 2003 by Jennifer K. Robinson. All rights reserved. 

### *Hot Lots/Engineering Lots*

Hot lots are lots that have a higher priority than other lots for sequencing at the various machines. This high priority allows hot lots to be run through the system quickly, typically at a cost to the cycle time of regular lots. Hot lots increase the cycle time of regular lots by:

- Causing regular lots to wait until the hot lots are finished;
- Forcing additional setups;
- Forcing the processing of small batches at furnaces; and
- Increasing variability.

Studies have shown that although hot lots increase the cycle times of regular lots, because the hot lots themselves have such low cycle times, overall average fab cycle times remain approximately the same. See Bonvik [8], Trybula [69], or Ehteshami *et. al.* [17] for details. However, hot lots do significantly increase the variance of lot cycle times, and hence degrade overall fab efficiency. Experiments performed under the MIMAC project confirmed these results. Hot lots did not appear as a significant factor in the overall designed experiments, which focused on cycle times, but not on cycle time variability. However, separate experiments showed that hot lots increased cycle time variability, particularly in the presence of large setups.

### *Order Release/WIP Limits*

It is a well-known result (Little's Law) that for a given start rate the amount of inventory is proportional to the average cycle time. Therefore, some production planning techniques focus on limiting the work-in-process (WIP) for a factory (or an equipment group) as a way to control cycle times. This can be accomplished by monitoring the amount of WIP and then controlling the release of lots into the line (or equipment group). Kanban systems are an example of this philosophy. WIP limits were imposed at the factory level (i.e. CONWIP) for the MIMAC project. In CONWIP systems, the total amount of WIP in the system is held to some specified constant (see [31] for details). This is a simplification of kanban systems.

An experiment was conducted under MIMAC to evaluate the impact of CONWIP limits on a two product logic factory. Simulations were then performed for various CONWIP limits, where the output variables recorded were the number of outs per time period as well as the observed cycle times. The observed throughput numbers under the CONWIP limits were then used to determine an equivalent start rate for each product (under each WIP limit) that would yield the same throughput. Simulations were then run using these equivalent start rates, and the resulting cycle times recorded.

The results showed an interaction with minimum batch size. Because most of the batch machines in the factory studied required more than one lot to begin processing, very low levels of WIP resulted in machines sitting idle while waiting for enough lots to arrive. This increased cycle times. At higher levels of WIP, the cycle times increased in proportion to increasing WIP. Interestingly, the cycle times were consistently higher for systems operating under a CONWIP policy than for those with equivalent throughput where the lots were started at constant intervals. This may have occurred because while the CONWIP systems had no variability in the amount of WIP in the system, they did have variability in the time between lot releases (a lot was released only when another lot completed processing). The other systems had no variability in lot interarrival times, and hence had lower cycle times.

 © 2003 by Jennifer K. Robinson. All rights reserved. 

WIP limits did not appear as a significant capacity loss effect in the MIMAC survey results. Because imposing a WIP limit has a similar effect to imposing a cycle time constraint, WIP limits were not included in the MIMAC designed experiments. However, considerable work has been done on the study of order release and WIP limits. For additional research on order release in semiconductor manufacturing, see Glassey and Resende [25], Philipoom and Fry [59], Roderick *et. al.* [63], or Ragatz and Mabert [60]. For release on WIP control, see Hopp and Spearman [31], Krajewski *et. al.* [41], or Mitra and Mitrani [48].

### *Alternative Tools (Equipment Dedication and Non-Identical Tools)*

Alternative tools are tools that are the same or similar in function that are assigned to specific process recipes. The term “equipment dedication” is commonly used to describe alternative tools. Three reasons to have alternative tools in a semiconductor factory are: equipment location, contamination, and equipment capability. In dedication for equipment location, a piece of equipment may be selected to run specific process recipes because of its location in the factory. Dedication for contamination occurs when there is concern that processing two different recipes on the same machine will lead to contamination. This type of dedication also reduces the setups that would be required to clean a single machine between process recipes. Dedication for equipment capability occurs when a fab has non-identical tools that are expected to perform similar functions. This situation may arise when the fab buys a newer version of a tool, but still keeps the older version for extra capacity. Also, non-identical tools may arise due to cost considerations. The fab might buy a very expensive piece of equipment that can perform all of the needed developing operations, but also buy a cheaper piece of equipment that can handle the less critical layers.

Equipment dedication is not generally considered a huge capacity loss factor. However, it does make capacity planning much more difficult. Tools dedicated because of equipment location or contamination can usually be treated as separate tool groups for the purposes of planning. However, they are frequently used to back one another up in the event of a serious equipment failure. This is difficult to account for in static (an even in many dynamic) models. Non-identical tools are even more difficult to account for, because they typically require different processing times to perform the same process steps. Assumptions must usually be made concerning what percentage of operations will be performed on the slower versus the faster piece of equipment.

Equipment dedication was found under MIMAC to reduce the impact of setups, but to also increase the effect of lack of tool redundancy (discussed in Section 5.3). It is difficult to speak in general of the capacity loss effect due to alternative tools, because the effect in a particular case is very dependent on the overall toolset and product mix. To fully understand the implications of alternative tools in a given situation will generally require the use of dynamic models. Johri [35], Leachman and Carmon [43], and Rohan [64] both discuss the problem in greater detail.

### *Shift Plans/End of Shift Effect*

Two different effects may occur in a wafer fab when a shift ends. First, disturbances may be observed even when the next shift has the same staffing levels. This effect is difficult to model, because it stems from the behavior of operators. For example, operators may be reluctant to start a product run that will not finish before they leave, leading to lost capacity. Even more impact is seen when the next shift has different staffing levels. The end of shift effects result in non-linearity of product shipments.

 © 2003 by Jennifer K. Robinson. All rights reserved. 

Shift effects result from the operation of the fab on a given day. Today, nearly all fabs are operated 24 hours a day in a three-shift mode, seven days a week. Sometimes not all shifts have the same staffing, but this is not normally considered a huge problem. Under MIMAC studies was made of the capacity loss due to shorter working time per day, and the increased WIP due to different staffing levels. Both were found to have highly predictable, linear responses. This factor was not studied in great detail, because it was not believed (by researchers or MIMAC survey respondents) to have a very significant effect.

### *Factory Shutdown*

Many factories are shut down for some time over the course of the year (e.g., for the week between Christmas and New Year's Day, or for two weeks in the summer). With this type of shutdown, no operators remain in the factory, and the equipment is turned off. While this results in cost savings, the tradeoffs in lost capacity are not always clear.

For MIMAC the impact of scheduled factory shutdowns was studied using simulation. The models used included twelve hour windows to prepare equipment prior to shutdown, and a similar amount of non-productive time after each shutdown to re-stabilize the fab environment. Before the equipment preparation, wafers were stockpiled at "safe points" in the process. These points were selected so that the wafers were not vulnerable to moisture damages during the shutdown. In stockpiling these wafers at "safe" points, production was halted at designated process steps at predetermined times.

Overall, this analysis found that the capacity loss caused by the stockpiling at "safe" points and by the actual factory shutdown was predictable. These results confirmed interview and survey results that factory shutdown had a predictable impact on capacity. A key point is that stockpiling the wafers at "safe" points had the same impact regardless of the length of the factory shutdown. Capacity is lost when wafers are held up at safe points and again when production begins as the bubbles of WIP at safe points are processed. This stockpiling loss is nearly constant for all shutdown lengths. It is also a relatively small loss. For this analysis, one model that contained only the stockpiling process (no shutdown involved) resulted in only small increases in mean cycle time that were not greater than the normal "noise" level of the factory.

Capacity loss is somewhat affected by the location and number of "safe" points in the process flow. If "safe" points occur before a certain tool set, then the tool set will have a WIP bubble to process after the shutdown. This capacity loss is affected by the proximity of the "safe" points to the factory bottleneck. Furthermore, the smaller the number of "safe" points, the larger those WIP bubbles will be. The changes in the process flow, product mix, tool set, and equipment throughput will also affect the capacity loss during stockpiling. While these capacity loss factors are complex, their effects can be readily estimated through discrete-event simulation. Factory shutdown was not included in the MIMAC designed experiments, because the overall loss did not appear significant in today's manufacturing environment.

### *Operator availability*

Operator availability refers to the number of operators assigned to a particular equipment group, and their break schedules. Operator availability is not generally included in capacity planning, although some companies include a loss factor to account

 © 2003 by Jennifer K. Robinson. All rights reserved. 

for “operator fatigue” or operator absence. The impact that operator availability has on capacity of a workcell depends on the following parameters:

- The size of the workcell (redundancy of tools).
- The amount of time operators are needed for processing (relative to the processing time).
- The number of operators per machine in the workcell.
- The cycle time constraint.

Experiments have shown that the first parameter, size of the workcell, influences capacity loss. For example, preliminary studies showed that 4 machine workcells are more sensitive to reductions in operator availability than 8 machine systems. This is consistent with intuitions about granularity of resources. The second and third parameters can be captured in the following ratio:

(operators per machine)/(% of operator use per process step)

A system with a value of 1.0 for the above ratio has exactly the amount of operators that would be specified by a pure static analysis. For example, a 4 machine system, with 2 operators, where an operator is needed for 50% of the processing time would have a ratio of 1.0. A ratio greater than 1 indicates the presence of “extra” operators, while a ratio of less than 1 indicates that operators constrain the system. The final parameter, cycle time constraint, clearly influences any operator availability experiments. Even with plenty of operators, a cycle time constraint may reduce the capacity of a workcell.

A series of small simulation experiments was conducted to assess the impact of number of operators on a system with multiple machines. The model had a single product, processed through one of four, eight, or 16 identical machines, where an operator was required for either 25% or 50% of the processing time. The number of operators in each model was chosen to yield operator availability ratios of 2.0, 1.5, and 1.0. The results of the experiment showed that having exactly “enough” operators according to a static calculation led to a reduction in capacity when compared to the system with “extra” operators. The difference was most pronounced in the smaller workcell, and amounted to up to a 15% difference in cycle time constrained capacity. This occurred because of the variability in the system, and the fact that operators were sometimes needed at more than one machine at a time.

While operator availability can have a significant impact at the workcell level when not enough operators are present, operator availability did not appear as significant in the MIMAC factory-level experiments. The reason for this was that in all of the datasets used, operators, as modeled, were not a constraint. However, human behavior is one of the most difficult aspects of a facility to model accurately, and this does not imply that operators do not have a large impact on real factories. For a discussion of the issues in modeling operator behavior, see Spier and Kempf [66].

### *Operator cross training*

Level of operator cross-training refers to the ability of operators to handle more than one type of equipment. Full cross-training means that every operator can handle every tool, so that one large pool of operators is possible. No cross-training means that every operator handles a single type of tool. When operators are not cross-trained, a single operator may be responsible for more than one piece of equipment. When both

 © 2003 by Jennifer K. Robinson. All rights reserved. 

machines require the operator at the same time, capacity is lost because the operator can only be in one place at a time. Cross training mitigates this effect.

For MIMAC the level of operator cross-training was studied using simulation models ranging from four workstation models to full factory models. Cross-training was found to increase capacity (particularly cycle time constrained capacity) when operators were highly utilized. Cross-training did not show up as a significant effect in the full factory models, however, because operators were not highly utilized in those datasets. In general, cross-training can be helpful if it increases the pool of operators who can manage a given machine. If, however, cross-training is used to reduce the number of operators in the factory, it will probably have a negative effect on cycle time constrained capacity. For additional results on operator cross-training, see Ward *et. al.* [70] or O'Ferrell [56].

## 5.3 Other loss factors that are inherent in semiconductor manufacturing

Several other capacity loss factors are inherent in the nature of semiconductor manufacturing, or result from strategic manufacturing decisions. The former include reentrant flow, rework, and time bound sequences. The latter include lack of tool redundancy, mix, and lot size. Lack of tool redundancy is driven by factory size and equipment procurement decision. Mix is driven by demand. Lot size is driven by highly strategic influences. Changing lot size would require equipment redesign, and is almost an industry-level decision. Each factor is discussed below.

### *Reentrant Flow*

Reentrant flow refers to the situation where lots return to the same equipment group more than once in the course of processing. This is common in semiconductor manufacturing, where lots repeat a similar sequence of operations for each layer added. Reentrant flow by itself does not lead to capacity loss, since the processing time for each visit to the equipment group can be accounted for in capacity planning. However, when there are setups between different operations on the same machine, or when different operations cannot be batched together, capacity of the system may be reduced. Dedication of a machine to a particular operation can reduce the impact of reentrant flow, but is often infeasible for systems without a high level of redundancy. Reentrant flow may be quantified somewhat by determining the percentage of flow that is reentrant at each station. In effect, reentrant flow is a mix problem at the machine level. Reentrant flow is something inherent in semiconductor manufacturing, which capacity planners have little control over. Reentrant flow was ranked of low importance by MIMAC survey respondents. Therefore, it was not studied further in the project. For additional information on reentrant flow, see Bai and Gershwin [2], Kumar [40], or Narahari and Khan [52].

### *Rework*

Rework refers to the case where a lot fails to pass inspection, and is sent back to repeat earlier processing. Two types of rework can be distinguished:

- immediate rework of a process step
- a rework route that requires several rework operations before the lot can return to its original path

 © 2003 by Jennifer K. Robinson. All rights reserved. 

Obviously, capacity is lost when rework occurs. Not only is the workload of individual tools increased, but the variability of process times and interarrival times also increases. Therefore, particularly in a cycle time constrained environment, more just the rework percentage is lost.

Parameters relevant to rework include rework percentage, which is the percentage of lots which have to be reworked after an operation, together with the rework path the lot has to follow. Usually, fabs measure the number of runs a piece of equipment makes for rework lots. The rework percentage is the number of rework runs divided by the total number of runs. This percentage is sometimes included in capacity planning spreadsheets as a loss factor.

Rework was included in the MIMAC factory-level experiments, but did not appear as a significant effect. The reason was probably that none of the datasets used included significant rework percentages. Rework was not ranked as a very significant loss factor in the MIMAC surveys, although it was ranked as more serious by people from pilot or research and development lines. For more information on rework, see Zargar and Ehteshami [75].

### *Time Bound Sequences*

Time bound sequences (TBS) occur when a particular operation must follow another operation within a set time period (possibly with intervening operations). If the time expires before the lot reaches the final operation, the lot must be either reworked or scrapped. The most effective use of the capacity of a TBS is never to have scrap or rework, which implies that the maximum capacity of the sequence is the maximum capacity of its (local) bottleneck, subject to the time constraint. Because of the time constraint, the real maximum batch size which can pass through the system may be limited. TBS thus interact with the batching factor. TBS also have high interaction potential with breakdowns (since there is less time available to recover from a failure).

Preliminary studies on TBS indicated that they had only a small impact on capacity of the factory overall. Because of this small impact, and because they were ranked unimportant by MIMAC survey respondents, time bound sequences were not studied further in the project.

### *Lack of Tool Redundancy*

Redundant tools are multiple (generally identical) tools that can process a given step. In a traditional fab layout, these similar tools are placed together in tool groups. A lot that is at a step with redundant tools can be processed on any of the available tools. Typically, as reported on the MIMAC surveys, lack of redundant tools is a problem for smaller fabs, or fabs that are just coming on line. Large-scale production fabs have very few one-of-a-kind tools, and hence do not have a major problem with this loss factor.

The presence of redundant tools leads to decreased cycle times, particularly when the equipment is subject to failures. The main reason for this is that with more tools in the tool group, the possibility of a lot finding all of the tools in the group down for maintenance at the same time decreases. Consider that if a one-of-a-kind tool is down 20% of the time, lots have a 20% chance of finding the machine down. If, however, a second machine with the same downtime characteristics is added, the probability of both of them being down is $0.2*0.2 = 0.04$ (4%). This effect can be readily illustrated using queueing or simulation models, and is commonly understood by people in fab management.

 © 2003 by Jennifer K. Robinson. All rights reserved. 

Lack of redundancy was not tested as part of the MIMAC designed experiments, because this would have required fundamental changes to the four datasets (changes in toolset). However, a separate experiment using just one of the datasets showed that by doubling the number of tools in each tool group, and the number of operators in each operator group, it was possible to more than double the cycle time constrained capacity. Above a certain size, however, no additional benefits of tool redundancy were apparent.

### *Mix*

Product and process mix are often cited as significant capacity loss factors. Increasing product mix results in increased setups, longer waiting time for batches of the same batch I.D., and increased likelihood of misprocessing. Capacity planning is typically done for an assumed product mix. It is difficult to assess the impact of deviations from the original mix on actual fab capacity, because it is difficult to quantify the deviations meaningfully. Some studies have shown (see [53]) that mix itself (for reasonably similar flows) does not lead to capacity loss if the fab is designed for that mix. However, when a fab deviates from its designed mix, serious capacity loss can result. Mix was not included in the MIMAC factory level experiments, because changing mix required essential changes in the factory datasets. Improving methods for planning capacity in the presence of mix is, however, an area of increasing interest among manufacturers, and is recommended as an area for further research. For additional discussions of mix, see Karmarkar and Kekre [38] or Neacy *et. al.* [53].

### *Lot Size*

Lot size refers to the number of wafers in a production lot. This is typically the same as the number of wafers in a transfer lot, though transfer lots and production lots can differ under some circumstances. Some tools process single wafers, others single lots, and still others batches (groups of lots). A typical lot size in semiconductor manufacturing is 25 wafers. However, lot sizes ranging from 12 to 48 wafers are sometimes used.

The effect of lot size on system capacity depends heavily on interactions with other factors, including load and unload times, yield loss, and maximum batch sizes at batch tools. Lot size was not included in the MIMAC designed experiments, because it would have required significant process changes to the datasets. However, a separate experiment using one of the datasets was performed to study varying the nominal lot size between 12 and 48 wafers. The aim of this single-effect study was to get some indication of the order of magnitude change in capacity resulting in an arbitrary change of lot size, with no further changes to the dataset. Because single-wafer lots were expected in particular to require process changes, they were not investigated in this study. Instead, single-wafer processing was left as a potential topic of future research.

A slight capacity loss was observed for the cycle time constrained cases in going from 48-wafer lots to 24-wafer lots. A more dramatic (nearly 50%) capacity loss was observed in all cases in going from 48-wafer lots to 12-wafer lots. Several factors were at play here. There was a reduction in the cycle time for the smaller lot sizes due to wafers not having to wait as long after completing single wafer operations. However, the smaller lot sizes resulted in increased setups, and increased processing on tools with per lot processing times. Also, wafers were sometimes required to wait longer to form full batches on machines with minimum batch sizes (for example, several machines had 96-wafer minimum batch sizes).

 © 2003 by Jennifer K. Robinson. All rights reserved. 

A more realistic study of the impact of smaller lot sizes would require information from fab personnel on setup times, load and unload times, and per lot processing times for candidate lot sizes. Control policies and batch machine threshold values might also need to be adjusted. For additional work on lot size, see Karmarkar [37].

# 6. CONCLUSIONS AND DIRECTIONS FOR FURTHER RESEARCH

This paper has described, at a high level, some of the difficulties inherent in planning capacity for semiconductor fabs. Both difficulties in the planning process and specific capacity loss factors, have been discussed. Some specific conclusions and recommendations for further research are listed below.

- The methodology derived under MIMAC for measuring cycle time constrained capacity may be helpful in reducing the disconnect between planning capacity and measuring performance.
- Cycle time constrained capacity may eliminate the need to use arbitrary contingency factors to account for variability.
- The biggest leverage opportunity in increasing fab capacity seems to lie in improving equipment reliability. Further research on the specific causes and effects of equipment downtime would be welcomed by people in the semiconductor industry.
- Another area in which improvements are needed is that of planning capacity in the presence of mix changes.
- The trend in capacity planning is towards using some type of stochastic model with the static (spreadsheet) model. Current tools are not sufficiently integrated to satisfy the needs of the semiconductor industry, and are a recommended area for further research.
- Data (in terms of availability and accuracy) remains one of the biggest sore spots for people in capacity planning.

## Acknowledgments

The authors are grateful to the many people who participated in the MIMAC surveys, interviews, and workshops, and who supported the project. There are far too many to name them individually.

 © 2003 by Jennifer K. Robinson. All rights reserved. 

## References


[1] F. Chance, J.K. Robinson, J.W. Fowler, O. Gihr, B. Rodriguez, and L.W. Schruben, “A Design of Experiments Methodology for Semiconductor Wafer Fab Capacity Planning,” submitted for publication, 1995.

[2] S.X. Bai and S.B. Gershwin, “Scheduling Manufacturing Systems with Work-in-Process Inventory Control: Reentrant Systems,” Report No. LMP-91-036, Laboratory for Manufacturing and Productivity, Massachusetts Institute of Technology, 1991.

[3] M. Baudin, V. Mehrotra, B. Tullis, D. Yeaman, and R.A. Hughes, “From Spreadsheets to Simulations: A Comparison of Analysis Methods for IC Manufacturing Performance,” Proceedings of the 4th Annual International Semiconductor Manufacturing Science Symposium, 1992.

[4] S. S. Baum and C. M. O’Donnell, “An Approach to Modeling Labor and Machine Down Time in Semiconductor Fabrication,” Proceedings of the 1991 Winter Simulation Conference, B. L. Nelson, W. D. Kelton, G. M. Clark (egs.), 448-454, 1991.

[5] S.M. Berlow, S.J. Hood, and C.Y. Wong, “On Improving Semiconductor Line Cycle Time without Losing Throughput,” Proceedings of the IBM International Manufacturing Productivity Symposium, IBM East Fishkill, New York, October 12-15, 1993.

[6] G.R. Bitran and D. Tirupati, "Tradeoff Curves, Targeting and Balancing in Manufacturing Queueing Networks," *Operations Research*, Vol. 37, No. 4, 547-564, 1989.

[7] J.H. Blackstone, Jr., D.T. Phillips, and G.L. Hogg, "A State-of-the-Art Survey of Dispatching Rules for Manufacturing Job Shop Operations," *International Journal of Production Research*, Vol. 20, No. 1, 27-45, 1982.

[8] Asbjoern M. Bonvik, "Estimating the Lead Time Distribution of Priority Lots in a Semiconductor Factory," Working Paper from Operations Research Center Massachusetts Institue of Technology, 1-26, 1994.

[9] D.Y. Burman, F.J. Gurrola-Gal, A. Nozari, S. Sathaye, and J.P. Sitarik, “Performance Analysis Techniques for IC Manufacturing Lines,” *AT&T Technical Journal*, Vol. 65, No. 4, 46-57, 1986.

[10] The Competitive Semiconductor Manufacturing Program, “The Competitive Semiconductor Manufacturing Survey: First Report on Results of the Main Phase,” Report CSM-02, Engineering Systems Research Center, University of California, Berkeley, April 2, 1993.

[11] D. Connors, G. Feigin, and D. Yao, “A Queueing Network Model for Semiconductor Manufacturing,” 1993. Submitted to *IEEE Transactions on Semiconductor Manufacturing*.

[12] J.A. Cunningham, “The Use and Evaluation of Yield Models in Integrated Circuit Manufacturing,” *IEEE Transactions on Semiconductor Manufacturing*, Vol. 3, No. 2, 60-71, 1990.


 © 2003 by Jennifer K. Robinson. All rights reserved. 

[13] J.E. Dayhoff and R.W. Atherton, "Signature Analysis of Dispatch Schemes in Wafer Fabrication," *IEEE Transactions on Components, Hybrids, and Manufacturing Technology*, Vol. CHMT-9, No. 4, 518-525, 1987.

[14] R.J. de Haut de Sigy, “Loading Control Policy for a Batch Machine,” Report No. LMP 90-001, Laboratory for Manufacturing and Productivity, Massachusetts Institute of Technology, 1990.

[15] B.L. Dietrich, J. Lee, and Y.S. Lee, "Order Selection on a Single Machine with High Set-Up Costs," *Annals of Operations Research*, Vol. 43, 379-396, 1993.

[16] G. Dobson and R.S. Nambimadom, “The Batch Loading and Scheduling Problem,” Working Paper No. QM 92-03, William E. Simon Graduate School of Business Administration, University of Rochester, Rochester, NY, 1992.

[17] B. Ehteshami, R.G. Petrakian, and P.M. Shabe, “Trade-Offs in Cycle Time Management: Hot Lots,” *IEEE Transactions on Semiconductor Manufacturing*, Vol. 5, No. 2, 101-105, 1992.

[18] K. Fordyce and G. Sullivan, "Cycle Time Versus Machine Utilization: Moving Along the Curve vs. Shifting the Curve," IBM Technology Report No. TR 21.1440, 1991.

[19] J.W. Fowler, R.C. Leachman, and G. Feigin, “Semiconductor Manufacturing Testbed: Testbed Datasets.” Available from John Fowler.

[20] J.W. Fowler, D.T. Phillips, and G.L. Hogg, “Real-Time Control of Multiproduct Bulk-Service Semiconductor Manufacturing Processes,” *IEEE Transactions on Semiconductor Manufacturing*, Vol. 5, No. 2, 158-163, 1992.

[21] J.W. Fowler and J.K. Robinson, “Measurement and Improvement of Manufacturing Capacity (MIMAC) Designed Experiment Report,” SEMATECH Technology Transfer #95062860A-TR, 1995.

[22] J.W. Fowler and J.K. Robinson, “Measurement and Improvement of Manufacturing Capacity (MIMAC) Final Report,” SEMATECH Technology Transfer #95062861A-TR, 1995.

[23] J.W. Fowler and J.K. Robinson, “Measurement and Improvement of Manufacturing Capacity (MIMAC) Project Bibliography,” SEMATECH Technology Transfer #94062424A-XFR, 1994.

[24] Fromm, Hans, “Some Remarks on Cycle Time, Variability, Zero Inventories, and Costs in Microelectronics Manufacturing Lines,” IBM Technical Report TR 28.167, 1992.

[25] C. R. Glassey and M. G. Resende, “Closed-Loop Job Release Control for VLSI Circuit Manufacturing,” *IEEE Transactions on Semiconductor Manufacturing*, Vol. 1, No. 1, 36-46, 1988. See also *Operations Research Letters*, Vol. 7, No. 5, 213-217, 1988.

[26] C.R. Glassey and W.W. Weng, “Dynamic Batching Heuristic for Simultaneous Processing,” *IEEE Transactions on Semiconductor Manufacturing*, Vol. 4, No. 2, 77-82, 1991.

[27] H. Gurnani and R. Akella, “Issues in Capacity Planning Under Different Repair Strategies,” Proceedings of the 1990 Japan-USA Symposium on Flexible Automation, 1113-1118, 1990.

 © 2003 by Jennifer K. Robinson. All rights reserved. 

[28] H. Gurnani, R. Anupindi, and R. Akella, "Control of Batch Processing Systems in Semiconductor Wafer Fabrication Facilities," Graduate School of Industrial Administration, Carnegie Mellon University, *IEEE Transactions on Semiconductor Manufacturing*, vol. 5, no. 4, 319-328, 1992. Also available in the Proceedings of the 1991 IEEE International Conference on Robotics and Automation, Sacramento, CA.

[29] D.T. Hicks, " Time-Based Competition: Evaluating Time-To-market as a Future Factory Manufacturing Concept," SEMATECH Technology Transfer #95012673A-XFR, 1995.

[30] S.J. Hood, "Detail vs. Simplifying Assumptions for Simulating Semiconductor Manufacturing Lines," Proceedings of the Ninth IEEE International Electronics Manufacturing Technology Symposium, 103-108, 1990.

[31] W.J. Hopp and M.L. Spearman, "Throughput of a Constant Work in Process Manufacturing Line Subject to Failures," *International Journal of Production Research*, Vol. 29, No. 3, 635-655, 1991.

[32] W.J. Hopp, M.L. Spearman, and D.L. Woodruff, "Practical Strategies for Lead Time Reduction," *Manufacturing Review*, Vol. 3, No. 2, 78-84, 1990.

[33] W.J. Hopp and S-C Wu, "Machine Maintenance with Multiple Maintenance Actions," *IIE Transactions*, Vol. 22, No. 3, 226-233, 1990.

[34] G. Inoue and K. Yoneda, "VLSI Production Analysis Using Queueing Network Model," Proceedings of the Fifth Symposium on Automated Integrated Circuits Manufacturing, Hollywood, FL, 33-44, 1989.

[35] P. K. Johri, "Overlapping Machine Groups in Semiconductor Wafer Fabrication," *European Journal of Operational Research*, Vol. 74, 509-518, 1994.

[36] P.K. Johri, "Practical Issues in Scheduling and Dispatch," *Journal of Manufacturing Systems*, Vol. 12, No. 6, 474-485, 1993.

[37] U.S. Karmarkar, "Lot Sizes, Lead Times and In-Process Inventories," *Management Science*, Vol. 33, No. 3, 409-418, 1987.

[38] U. Karmarkar and S. Kekre, "Manufacturing Configurations, Capacity and Mix Decisions Considering Operational Costs," *Journal of Manufacturing Systems*, Vol. 6, No. 4, 315-324, 1987.

[39] U.S. Karmarkar, S. Kekre, and S. Kekre, "Capacity Analysis of A Manufacturing Cell," *Journal of Manufacturing Systems*, Vol. 6, No. 3, 165-175, 1987.

[40] P. Kumar, "Re-Entrant Lines," *Queueing Systems: Theory and Applications: Special Issue on Queueing Networks*, Vol. 13, 87-110, 1993.

[41] L.J. Krajewski, B.E. King, L.P. Ritzman, and D.S. Wong, "Kanban, MRP, and Shaping the Manufacturing Environment," *Management Science*, Vol. 33, No. 1, 39-57, 1987.

[42] Averill M. Law, "Models of Random Machine Downtimes for Simulation," Proceedings of the 1990 Winter Simulation Conference, 314-317, 1990.

[43] R.C. Leachman and T.F. Carmon, "On Capacity Modeling for Production Planning with Alternative Machine Types," *IIE Transactions*, Vol. 24, No. 4, 62-72, 1992.


 © 2003 by Jennifer K. Robinson. All rights reserved. 

[44] H.L. Lee and C.A. Yano, "Production Control in Multistage Systems with Variable Yield Losses," *Operations Research*, Vol. 36, No. 2, 269-278, 1988.

[45] G. Leonovich, "An Approach for Optimizing WIP/Cycle Time/Output in a Semiconductor Fabricator," 1994 IEEE/CPMT Int'l Electronics Manufacturing Technology Symposium.

[46] S.C.H. Lu, D. Ramaswamy, and P.R. Kumar, "Efficient Scheduling Policies to Reduce Mean and Variance of Cycle-Time in Semiconductor Manufacturing Plants," *IEEE Transactions on Semiconductor Manufacturing*, Vol. 7, No. 3, 1994, 374-380.

[47] D.P. Martin, "Key Factors in Designing a Manufacturing Line to Maintain Tool Utilization and Minimize Turnaround Time," IBM Technology Products, Essex Junction, VT, 1994.

[48] D. Mitra and I. Mitrani, "Analysis of a Kanban Discipline for Cell Coordination in Production Lines," *Management Science*, Vol. 36, No. 12, 1548-1566, 1990.

[49] D.J. Miller, "Simulation of a Semiconductor Manufacturing Line," *Communications of the ACM*, Vol. 33, No. 10, 98-108, 1990.

[50] A. Najmi, "Management of Cycle Time in Semiconductor Wafer Fabrication," Engineering Systems Research Center, Report No. 93-3, the University of California at Berkeley, 1993.

[51] Y. Narahari, and L. M. Khan, "Modeling Re-Entrant Manufacturing Systems With Inspections," Proceedings of the 1995 IEEE International Conference on Robotics and Automation. Part 2 (of 3), Nagoya, Jpn, 1738-1743, 1995.

[52] Y. Narahari and L.M. Khan, "Performance Analysis of Scheduling Policies in Re-Entrant Manufacturing Manufacturing Systems," *Computers Ops Res*, Vol. 23, No. 1, 1996, 37-51.

[53] E. Neacy, N. Abt, S. Brown, M. McDavid, J. Robinson, S. Srodes, and T. Stanley, "Cost Analysis for a Multiple Product / Multiple Process Factory: Application of SEMATECH's Future Factory Design Methodology," Proceedings of the Advanced Semiconductor Manufacturing Conference, Boston, MA, 1993.

[54] E. Neacy, S. Brown, R. McKiddie, "Measurement and Improvement of Manufacturing Capacity (MIMAC) Survey and Interview Results," SEMATECH Technology Transfer #94052374A-XFR, 1994.

[55] S.J. New, A.G. Lockett, and R.J. Boaden, "Using Simulation in Capacity Planning," *Journal of the Operational Research Society*, Vol. 42, No. 4, 271-279, 1991.

[56] D.S. O'Ferrell, "Manufacturing Modeling and Optimization," Proceedings of the IEEE/SEMI Advanced Semiconductor Manufacturing Conference, 334-339, 1995.

[57] P. O'Neil, "Performance Evaluation of Lot Dispatching and Scheduling Algorithms Through Discrete Event Simulation," Proceedings of the 3rd International Semiconductor Manufacturing Science Symposium, Burlington, CA, 21-24, 1991.

[58] S.S. Panwalkar and W. Iskander, "A Survey of Scheduling Rules," *Operations Research*, Vol. 25, No. 1, 45-61, 1977.


 © 2003 by Jennifer K. Robinson. All rights reserved. 

[59] P.R. Philipoom and T.D. Fry, "Capacity-Based Order Review/Release Strategies to Improve Manufacturing Performance," *International Journal of Production Research*, Vol. 30, No. 11, 2559-2572, 1992.

[60] G.L. Ragatz and V.A. Mabert, "An Evaluation of Order Release Mechanisms in a Job-Shop Environment," *Decision Sciences*, Vol. 19, 167-180, 1988.

[61] J.K. Robinson, F. Chance, and J.W. Fowler, "The Impact of Lot Dispatch Strategy on Semiconductor Wafer Fab Capacity," submitted for publication, 1995.

[62] J.K. Robinson, J.W. Fowler, and J.F. Bard, "The Use of Upstream and Downstream Information in Scheduling Semiconductor Batch Operations." *International Journal of Production Research*, Vol. 33, No. 7, 1849-1870, 1995.

[63] L.M. Roderick, D.T. Phillips, and G.L. Hogg, "Comparison of Order Release Strategies in Production Control Systems," *International Journal of Production Research*, Vol. 30, No. 3, 611-626, 1992.

[64] D. Rohan, "Resource Sharing in Capacity Analysis," Proceedings of the 1992 IEEE/SEMI Advanced Manufacturing Conference, 39-42, 1992.

[65] A.M. Spence and D.J. Welter, "Capacity Planning of a Photolithography Work Cell in a Wafer Manufacturing Line," Proceedings of the IEEE International Conference on Robotics and Automation, Raleigh, NC, 702-708, 1987.

[66] J. Spier and K. Kempf, "Simulation of Emergent Behavior in Manufacturing Systems," Proceedings of the IEEE/SEMI Advanced Semiconductor Manufacturing Conference, 90-94, 1995.

[67] C.S. Tang, "Designing an Optimal Production System with Inspection," *European Journal of Operational Research*, Vol. 52, No. 1, 45-54, 1991.

[68] P. Tran-Gia and A. Schoemig, "Discrete-time Analysis of Batch Servers with Bounded Idle Time," University of Wuerzburg, Wurzburg, Germany, submitted for publication, 1996.

[69] W.J. Trybula, "Hot Jobs, Bane or Boon," Proceedings of the 1993 IEEE/CHMT International Electronics Manufacturing Technology Symposium, Santa Clara, CA, 317-322, 1993. Also available as SEMATECH Technology Transfer No. 93041617A-ER.

[70] P.T. Ward, P.D. Berger, J.G. Miller, and S.R. Rosenthal, "Manufacturing Process Technology and Support Staff Composition: An Empirical View of Industry Evidence," *Production and Operations Management*, Vol. 1, No. 1, 5-21, 1992.

[71] L.M. Wein, "On the Relationship Between Yield and Cycle Time in Semiconductor Wafer Fabrication," *IEEE Transactions on Semiconductor Manufacturing*, Vol. 5, No. 2, 156-158, 1992.

[72] L.M. Wein, "Scheduling Semiconductor Wafer Fabrication," *IEEE Transactions on Semiconductor Manufacturing*, Vol. 1, No. 3, 115-126, 1988.

[73] W.W. Weng and R.C. Leachman, "An Improved Methodology for Real-Time Production Decisions at Batch-Process Work Stations," *IEEE Transactions on Semiconductor Manufacturing*.

[74] W. Whitt, "The Queueing Network Analyzer," *The Bell System Technical Journal*, Vol. 62, No. 9, 2779-2814, 1983.

 © 2003 by Jennifer K. Robinson. All rights reserved. 

[75] A.M. Zargar and B. Ehteshami, “Tradeoffs in Cycle Time Management: Reworked Bonus Lots,” Proceedings of the Summer Computer Simulation Conference, 1039-1043, 1991.

[76] C. Zhou and P.J. Egbelu, “Scheduling in a Manufacturing Shop with Sequence-Dependent Setups,” *Robotics and Computer-Integrated Manufacturing*, Vol. 5, No. 1, 73-81, 1989.

 © 2003 by Jennifer K. Robinson. All rights reserved. 