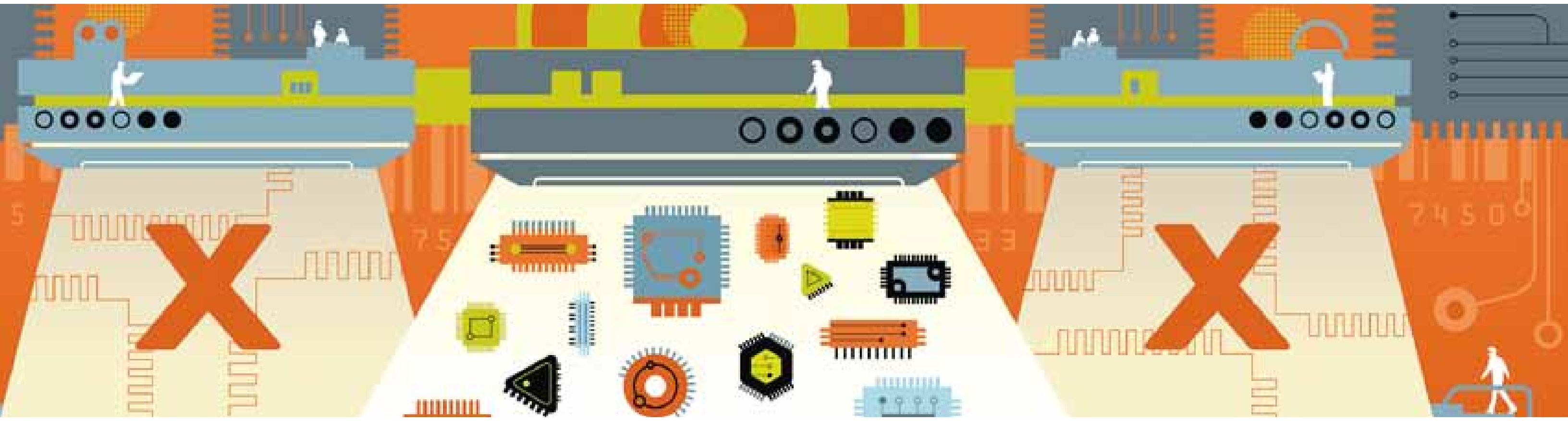


Mastering variability in complex environments

Variability adds cost to semiconductor production systems, but the ability to cope with it can also be a critical source of profit. A new approach goes beyond traditional tools to help companies control variability in their processes and make intelligent trade-offs in order to maximize return.

Harald Bauer,
Iana Kouris,
Gernot Schlögl,
Thomas Sigrist,
Jan Veira,
and Dominik Wee

Semiconductors are among the most complex products manufactured today. That fact, in combination with the pace of change in the industry and the inherent difficulties in managing dynamic production environments, leads to up-and-down performance. After all, the need to produce a portfolio of different products, each of which requires a different amount of time or resources for production, inevitably creates inefficiencies, which result in financial losses. These losses can arise either from a drop in the utilization level of expensive labor and equipment or by an increase in lead times[1] as material waits in queues. Traditional methods for the optimization of production and other systems—most famously, lean—seek to eliminate variability completely, aiming to level demand and balance flow within a system. However, in many production situations, some variability is inevitable, so manufacturers must seek additional tools.

Managing variability in a sophisticated fashion can have positive effects for semiconductor manufacturers, creating an opportunity for them to distinguish themselves from their competitors.

The particular traits that cause variability in the semiconductor industry include dramatic advancements in technology, rapidly evolving customer demands, and a growing demand for specialty products. The variability that each of these forces introduces has a strong impact on

[1] Time between production start and production end of a wafer, including processing and waiting time.

profitability. Market demand and intense product-development activities require an expanding menu of different chips in ever-smaller quantities. Yet the popularity of a given product cannot always be predicted reliably, and therefore it is possible to be blindsided by sudden spikes (or sudden drops) in demand.

Even at the best of times, then, variability presents a challenge. However, the recent global recession revealed the ways variability can make a tough situation worse. During the recession, excess capacity was so high that 5 of the top 10 semiconductor companies reported material losses as a result. Even Intel, the standout financial performer in the sector, reported a loss of $1.1 billion in its 2010 annual report—a loss attributed specifically to excess capacity. If more sophisticated management of variability could mitigate any degree of excess capacity, it could make a substantial difference during the next down cycle.

A new approach has been developed to help semiconductor companies manage this uncertainty. This approach allows companies to openly ask a question that pertains to many if not all production systems: what is the quantitative relationship among variability, lead time, and utilization? By quantifying the impact of changes in lead time and utilization on variability using a new operating-curve methodology, semiconductor players can manage that variability—specifically, the variability that is left after more traditional tools such as lean have been taken as far as they can go. The ideal proceeding, now practiced by some leading semiconductor companies, is to identify the "sweet spot" in the management of their production systems—a spot determined by as much reduction of variability as possible, along with effective management of the remaining variability.

Use well-known tools

As noted, fast technological progress and rapidly changing customer demand make it hard for semiconductor manufacturers to control variability. However, significant opportunity exists to reduce variability in the production system.

The first step is to apply classical tools, such as lean manufacturing. Lean seeks to maximize the throughput of a production line by identifying and eliminating waste from the steps that cause bottlenecks in the process, with typical improvements of 20 to 30 percent. In semiconductor manufacturing, three sources of waste are prevalent: loss of availability (scheduled and unscheduled downtime), loss of utilization (changeover time, idle time, and loss of speed), and loss of quality (rework and scrap). By quantifying these three sources of waste, lean can identify "hidden" capacity in both leading-edge and lagging fabs, with different kinds of improvement levers applied. A focus on automation and handling inside the tools is most valuable for leading-edge fabs. Lagging fabs typically benefit most from the application of operator-efficiency levers.

In pursuit of a competitive advantage, some leading semiconductor firms are now moving beyond these traditional tools to apply new levers to optimize high-variability production environments. For example, a sophisticated approach to maintenance management can reduce the variability caused by equipment failures while minimizing the impact on production speed of planned maintenance events. One semiconductor fab reduced lead time per wafer by 20 percent by splitting a small number of relatively long maintenance shutdowns into a larger number of short ones (Exhibit 1). Even though the new strategy meant production lines were shut down for 15 percent longer

Exhibit 1

Splitting maintenance events can increase speed, despite reducing overall equipment availability.

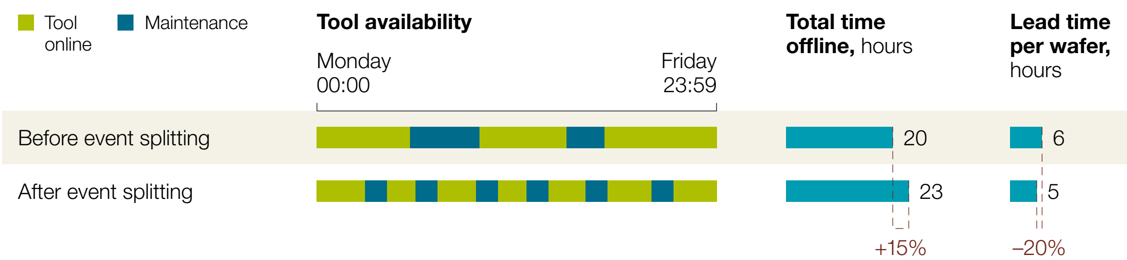
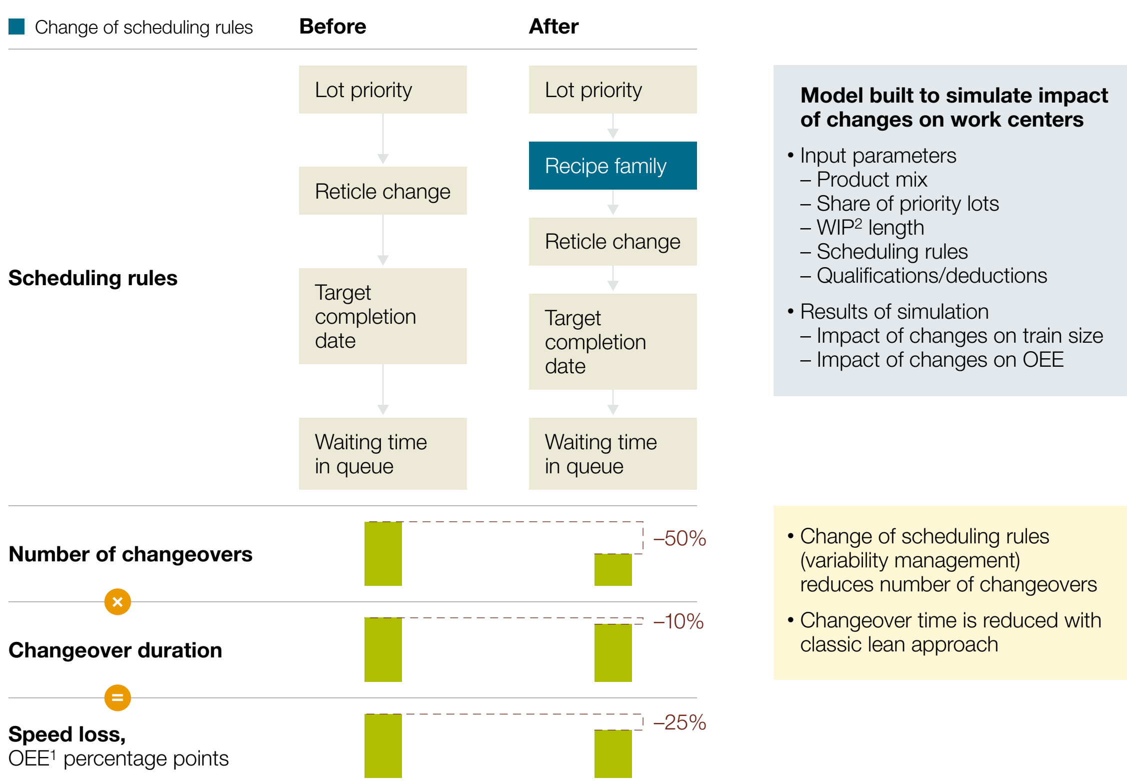


Exhibit 2

Clever batch sequencing reduced the number of changeovers—and lean methods made them faster.



[1]Overall equipment effectiveness.
[2]Work in process.

in total, the shorter stops kept material flowing more quickly through the system.

Optimizing manufacturing lines within a production facility helps, too. Companies can often dedicate particular lines to certain types of products, minimizing variability in individual lines and allowing each line to be optimized to suit specific product requirements. Some lines may be optimized for high-volume production, while others concentrate on specialty products.

Advanced batching and scheduling systems also help to reduce variability in production. By grouping different products with similar "recipes," for example, companies can minimize the number and effect of changeovers. Exhibit 2 shows how a semiconductor line used a combination of better sequencing and traditional lean tools to reduce both the number and the length of changeovers between batches, leading to a 25 percent reduction in speed losses.

Another important challenge in many environments is batch prioritization. Some jobs must be completed more quickly than others, either because customer requirements call for extra speed through the process or because waiting at one point will result in low utilization further downstream. Unfortunately, adjusting equipment to handle priority batches often incurs extra losses and delays. By establishing optimal rules for batch priorities, companies can ensure that priority batches are handled efficiently, without an excessive cost or time penalty.

Manage the rest of the variability

While not quite universal, attempts to minimize variability in production systems are widespread in most manufacturing companies. The methods described above can be successful in reducing variability to a large extent—but they cannot eliminate it completely. It is also the case that some variability-reduction mechanisms will prove not to be worth the cost. As a consequence, all semiconductor companies will reach a point at which variability has been reduced as far as is possible, or as far as is economically viable. What then?

The next challenge for a fab is to decide how it should operate its systems with the variability that remains. At this point, we recommend the introduction of an operating curve to calculate the effect of variability on a system. This curve is based on a theoretical approach known as queuing theory. Developed in academia in the 1970s, queuing theory has proved applicable in industrial applications over the last 10 to 15 years. Every production process has its own operating curve, which can be plotted using data on output, variability, and speed. The curve shows how different combinations of utilization (along the x axis) and lead time (along the y axis) generate different levels of variability (Exhibit 3).

A "perfect" production system with no variability at all, a hypothetical car plant perhaps, has an operating curve like the dotted line in the exhibit. Products flow smoothly through the process, and increasing volumes have no effect on lead time, so the equipment can achieve its maximum possible utilization.

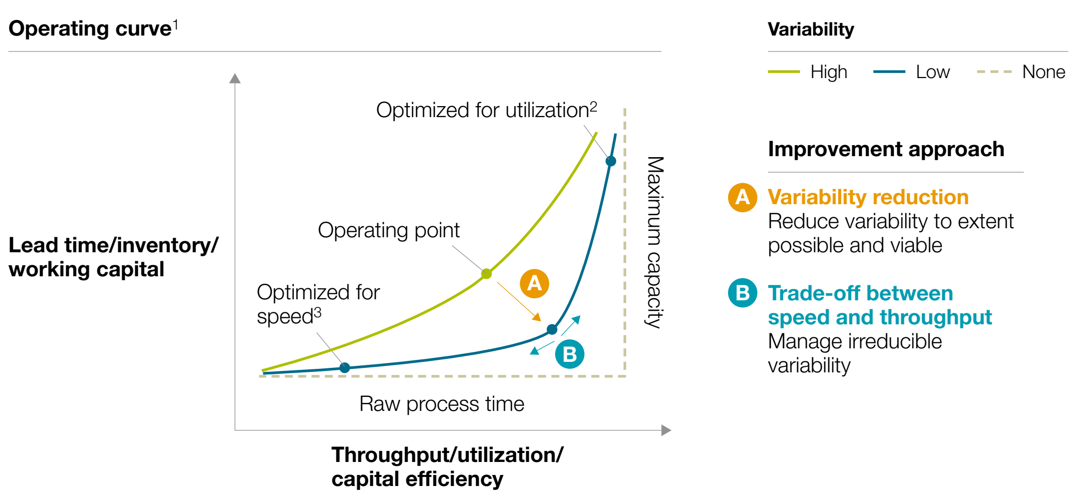
When variability is introduced, the operating curve looks more like the solid lines in the exhibit. The more the production assets are utilized, the longer the overall lead time becomes, as work in process increases in order to get the best out of the equipment on the line. To optimize systems with variability in them, their owners can use two levers: they can try to reduce variability wherever possible, as discussed above, moving the curve lower (that is, to the right). Then they can examine the curve for the optimal operating point, the place where utilization is maximized, without producing unacceptably long lead times. This is the most profitable point of operation.

In essence, the operating curve quantifies the trade-off between speed and utilization that every system must face in the presence of variability. Our experience suggests that building a reliable curve requires several months' worth of production data. This amount of data is at once statistically significant and sufficient for clients to see the financial impact of decisions that affect speed and utilization. It also allows clients to compare the impact of reducing variability using various levers with the cost of doing so.

Most semiconductor plants produce a combination of high-volume "commodity" products and lower-volume runs for specialty products. These plants are hugely expensive to build and run, with the result that utilization has traditionally been a priority for plant operators at the expense of speed. When one semiconductor manufacturer decided to take a new look at its operations, it realized that reducing variability and finding a new point along the operating curve would have a positive impact on its profits. First, the company took all the economically viable steps it could to reduce variability losses in its production system. Next, it made the trade-off described above and decided to reduce

Exhibit 3

Different combinations of lead time and utilization generate different levels of variability.



[1]Bending of the curve is determined by the system's variability.
[2]High work in process (WIP); high utilization; low speed.
[3]Low WIP; low utilization; high speed.

overall utilization in order to improve speed. The result, after an 18-month improvement effort, was a 20 percent decrease in manufacturing costs.

• • •

The best semiconductor companies are applying these two fundamental levers in an iterative way, working with their customers to control the demand for variability, relentlessly improving their production systems to reduce its impact, and adjusting their operating blend to make the best of the remaining variability. Soon, we expect to see variability management added to the arsenal of standard manufacturing methodologies in all companies: lean focuses on reducing variability, and the approach presented here will reduce the costs associated with the variability that lean cannot eliminate. ○

Harald Bauer is a principal in the Frankfurt office, where **Iana Kouris** is a consultant. **Gernot Schlögl** is an associate principal in the Santiago office. **Thomas Sigrist** is an alumnus of the Zurich office. **Jan Veira** is a consultant in the Berlin office. **Dominik Wee** is a consultant in the Munich office.