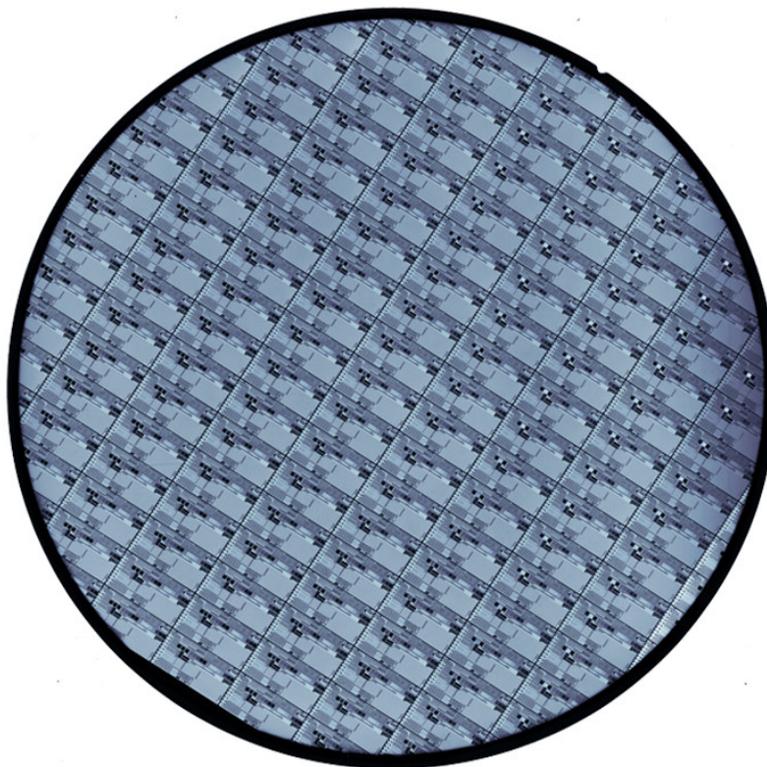


Advanced Electronics Practice

Semiconductor design and manufacturing: Achieving leading-edge capabilities

As chips get smaller and competition increases, semiconductor companies need a new strategy that considers everything from fab size to supply-chain issues.

This article was written collaboratively by the Advanced Electronics Practice. The authors include Harald Bauer, Ondrej Burkacky, Peter Kenevan, Stephanie Lingemann, Klaus Pototzky, and Bill Wiseman.



Semiconductors are the unsung heroes of the technology world, working behind the scenes to power everything from toys and smartphones to cars and thermostats. In recent years, they have enabled breakthrough technologies, including artificial intelligence and machine learning, that have transformed how we live and work. Taking the digital revolution to the next level will require even more advanced chips with greater computational power and memory capacity.

With the COVID-19 crisis disrupting supply chains and geopolitical tensions increasing, semiconductor companies have become more interested in achieving end-to-end design and manufacturing capabilities for leading-edge technology. Many governments share this interest and are attempting to support their local semiconductor markets. But new fabs and extensive R&D programs—essential for producing leading-edge technologies at high volumes—require billions in investment. A misstep in these areas, lax cost control, or lower-than-expected demand can severely decrease or even eliminate a company's return on investment. Leading-edge chip design and manufacturing also require strong capabilities in research, supply chain, talent, and intellectual-property (IP) protection, as well as the ability to navigate government policies. While semiconductor companies may excel at some of these tasks, few have top capabilities across the board.

Given the extended time frames required to build fab infrastructure and enhance workforce skills, semiconductor companies need a long-term strategy for achieving design and manufacturing excellence—one that considers construction issues, equipment costs, and the need to enhance internal capabilities. Here's a road map for moving forward.

Greater complexity and greater costs for semiconductor design and manufacturing

Over the past decade, the need for leading-edge technology leadership has transformed from

an amorphous goal to an absolute necessity at semiconductor companies. In line with Moore's law, the number of transistors on a chip roughly doubled every two years over that period, although the pace has recently slowed. While complexity has increased, structures on chips have shrunk in size.

Only a few companies are capable of designing and manufacturing the most advanced chips with node sizes of 14 nanometers (nm) and below because of the skills and large investment required for design, R&D, scaling, and other activities. Meanwhile, demand for these chips is soaring. In some major market segments, including artificial intelligence and machine learning, chips under 14 nm are critical because they combine strong performance with lower power consumption (see sidebar "Major markets for sophisticated chips").

A tough market for all but a few

The semiconductor industry's record of steady technological improvement has created a winner-take-all dynamic that makes leading-edge capabilities vital within several segments. If a company's product or service is even slightly better than a competitor's, it typically captures an outsize portion—or even the vast majority—of industry revenue. This phenomenon is apparent along the entire value chain, from equipment production to chip manufacture. Companies that want to challenge the winner may find it difficult to catch up, since the leading players are often several years ahead in technology development.

The implications of the winner-take-all culture became apparent when we examined the economic profit generated by 254 semiconductor companies from 2015 through 2019 (Exhibit 1). Our analysis covered all types of industry players, from equipment and material providers to those that create specialized chips for different end markets. We found that many companies experienced a strong increase in economic profit as they recovered from the 2008 financial crisis, although their margins were similar to past levels.

Among our sample, a few players stood out for their exceptionally high economic profit. Most of these top companies focused on leading-edge technology and continued to strive for ever-smaller and more efficient semiconductors. Typically, the top companies focused on one product segment or a single step in the value chain, since intense effort is required to achieve and retain R&D and manufacturing leadership.

A particular business—or even a group of businesses within a region—may become a hub of expertise (Exhibit 2). Consider a few examples of company- and market-specific strengths:

- Intel dominates the market for desktop and laptop CPUs.
- Qualcomm is the leader in the smartphone system-on-a-chip market.
- TSMC in Taiwan is the top manufacturer for chips at ten nm or below.

Major markets for sophisticated chips

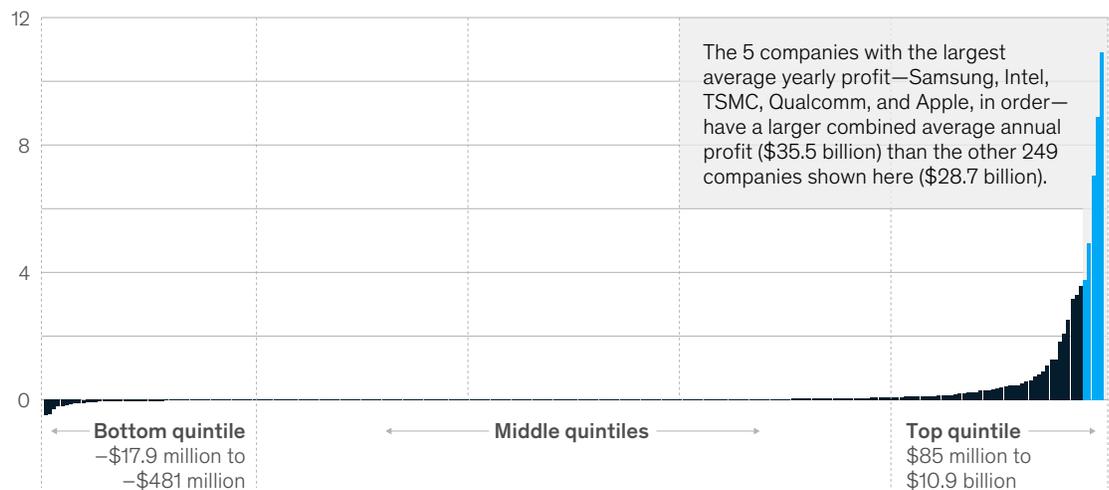
The smallest structures in chips, which are 5 nm and 3 nm in size, confer important benefits within select market segments and application areas, including data centers, 5G smartphones, edge computing, and machine learning. Semiconductor customers in these segments will pay a premium for leading-edge chips, since they need a combination of strong performance and low power requirements to win their markets. The resulting sales boost will more than compensate for higher chip costs.

In other market segments, most companies have little incentive to move to smaller structures, since this shift would require major changes in their design and production processes. What’s more, their customers often have more important priorities than low power requirements. For instance, a company that primarily develops safety critical automotive applications will focus on chips with superior reliability rather than highest performance.

Exhibit 1

The semiconductor industry is ‘winner takes all,’ with a limited set of top performers.

Average yearly profit of semiconductor companies,¹ 2015–19, \$ billion



¹Total sample of ~380 companies (with data available for 254 companies from 2015–19). Source: S&P, Corporate Performance Analytics by McKinsey

- ASML, a Dutch company, produces most lithography equipment, especially leading-edge products.
- Samsung in Korea leads the memory market.
- NVIDIA in the United States dominates the market for graphics cards.
- Almost all specialty chemicals used in semiconductor manufacture come from Japan.
- Japanese and South Korean companies produce the majority of wafers.

While specialization confers competitive advantages, it also means that semiconductor companies and related businesses are highly

interdependent. Today, no local market or company has all the capabilities required for end-to-end semiconductor design and manufacturing. If there is a major disruption to the supply chain, similar to what has happened this year because of the COVID-19 crisis, production bottlenecks could occur and certain chips could be in short supply.

Steep investment and delayed returns

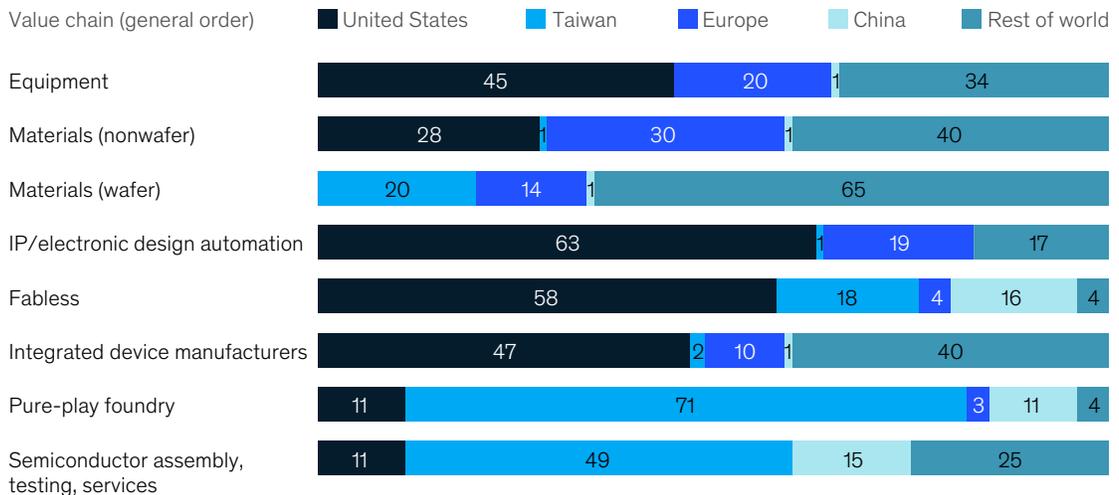
Leadership within manufacturing and R&D is not simply a tool for winning the market; it is also essential to cost control. That benefit is now more important than ever because expenses have soared over the past ten years.

First, consider fab construction. The cost of building and equipping a facility with 5 nm production lines now runs about \$5.4 billion—more than

Exhibit 2

No region has end-to-end capabilities for semiconductor design and manufacturing.

Share of 2018 sales based on company headquarters, %



Source: Gartner; IHS; Strategy Analytics; McKinsey

three times the \$1.7 billion required for a fab with 10 nm production lines (Exhibit 3). Most of the increase is related to the greater precision required to manufacture small structures. For instance, lithography now uses extreme ultraviolet light, increasing costs for tools.

Now consider R&D. As chips get smaller, R&D becomes more challenging as researchers deal with quantum effects, minor structural variations, and other factors that can complicate development. Designing a 5 nm chip costs about \$540 million for everything from validation to IP qualification. That is well above the \$175 million required to design a 10 nm chip and the \$300 million required for a 7 nm chip. We expect that R&D costs will continue to escalate, especially for leading-edge products.

Although semiconductor companies must devote billions to new fabs, they will not see an immediate return on their investment. It takes about 12 to 24 months to build a shell of a fab and install the required tools, plus another 12 to 18 months to ramp up to full capacity. And if demand falls beyond projections, or if costs exceed expectations, the anticipated returns could be much lower than expected.

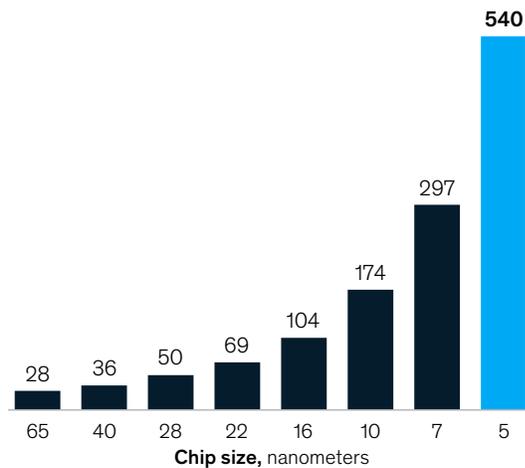
Strategies for building leading-edge fabs while controlling costs

To increase clarity about risks and potential returns, we created a simplified model to predict when semiconductor companies will see profits from a new fab. Rather than providing precise cost insights, the model is designed to give some general estimates

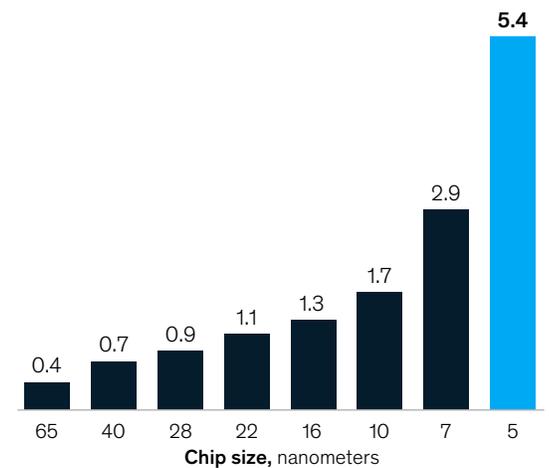
Exhibit 3

R&D for chips and fab module construction costs are soaring.

Chip-design cost,¹ \$ million



Fab module construction cost, \$ billion



¹Major components include IP qualification, architecture, verification, physical, software, prototyping, and validation. Source: IBS; McKinsey

about timelines for receiving a return on investment. The model focuses on the following factors:

- the amount of capital expenditures (capex) invested, adjusted for depreciation time, cost of capital, and other factors
- revenue, based on the average selling price for products manufactured at the fab, year over year
- operational expenditures, including utilities, labor, and materials, since this will influence how quickly a company can pay off its initial capex
- the amount of government support received— one of the most critical factors in determining when companies will realize a profit

In a best-case scenario, with high utilization, our model shows that semiconductor companies can expect to see a positive cash flow from their new

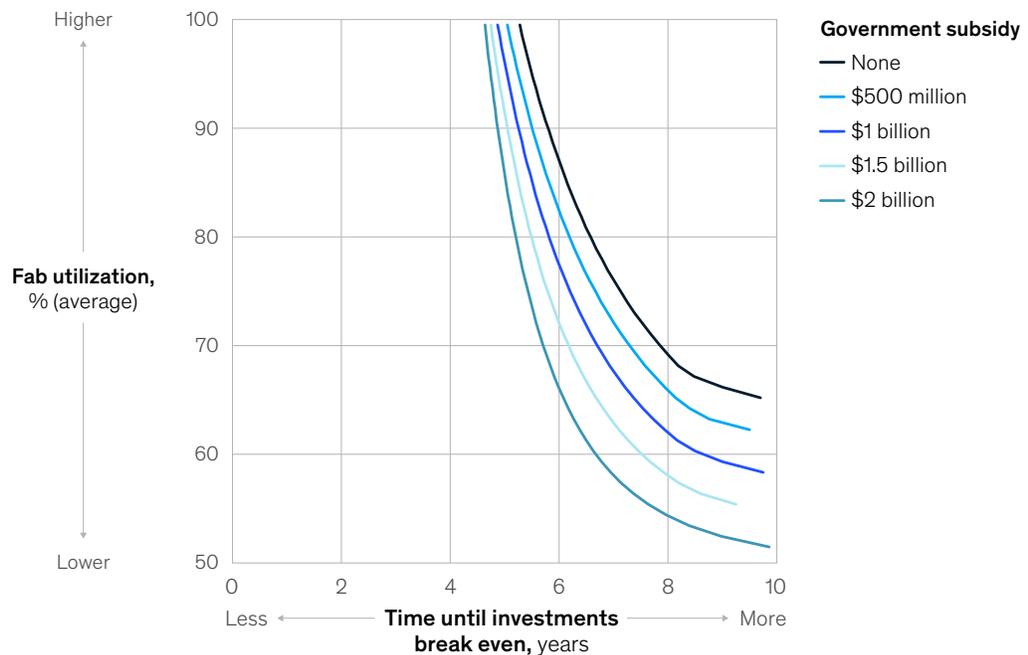
fabs within about five years, even with little to no government support (Exhibit 4). If the fabs are utilized below capacity, however, this time frame could increase significantly. In extreme cases, where utilization drops below about 55 to 65 percent of full capacity and remains at that level, companies could be in a dire situation. Even in a best-case scenario, where a semico receives high government subsidies, it would take almost ten years to see a return on investment when operating at 55 percent capacity.

With so much at stake, semiconductor companies must thoroughly investigate product demand, including possible long-term shifts, before beginning fab construction. If they decide to proceed, they should also attempt to reduce costs per wafer through two strategies: fab scaling and industry clustering. As noted earlier, they should also consider the impact of any government subsidies, since that could significantly shorten

Exhibit 4

Fabs that are used below capacity take significantly longer to reach the breakeven point.

Analysis based on model for leading-edge fab producing 5 nanometer products



Source: McKinsey analysis

the time frame to the breakeven point, making investments more appealing.

Fab scaling

Today, semiconductor manufacturers will often build large fabs with extensive production lines. Only a few years ago, a fab with 20,000 wafer starts per week (wspw) was a big deal. Now leading fabs routinely scale production to deliver 100,000 wspw.

Beyond increasing output, large fabs confer many other advantages. At the building stage, they have lower costs per square meter. A fab's shell accounts for a relatively small proportion of the total investment; the real cost drivers at new fabs are cleanroom technologies and manufacturing equipment.

After the construction phase, large-scale operations can reduce overhead and increase labor productivity, decreasing the cost per wafer (Exhibit 5). For instance, companies will be able to centralize some production functions, such as industrial engineering, as well as some support functions, such as human resources and accounting. Companies operating two smaller fabs would have duplicative functions at each site and higher costs. Labor productivity will also increase because direct manufacturing staff, such as members of the maintenance and engineering teams, will have less downtime. If one area or production line is slow, their services will likely be needed elsewhere.

These lower overhead costs, combined with greater profits from the larger outputs, will increase a semico's profits.

Scaling will also help semiconductor companies increase overall equipment efficiency (OEE)—the amount of productive manufacturing time. Demand peaks and drops will level out, since a single fab will be serving multiple customers, and utilization will increase for high throughput tools. What's more, semiconductor companies will be able to churn out more products and offer volume discounts to customers who bundle their orders or need large quantities.

If companies cannot expand existing sites or do not want to build a large facility, they could locate multiple fabs within a single region to obtain benefits of scale. They will still be able to centralize many functions and shift employees between facilities, as well as among production lines, depending on demand.

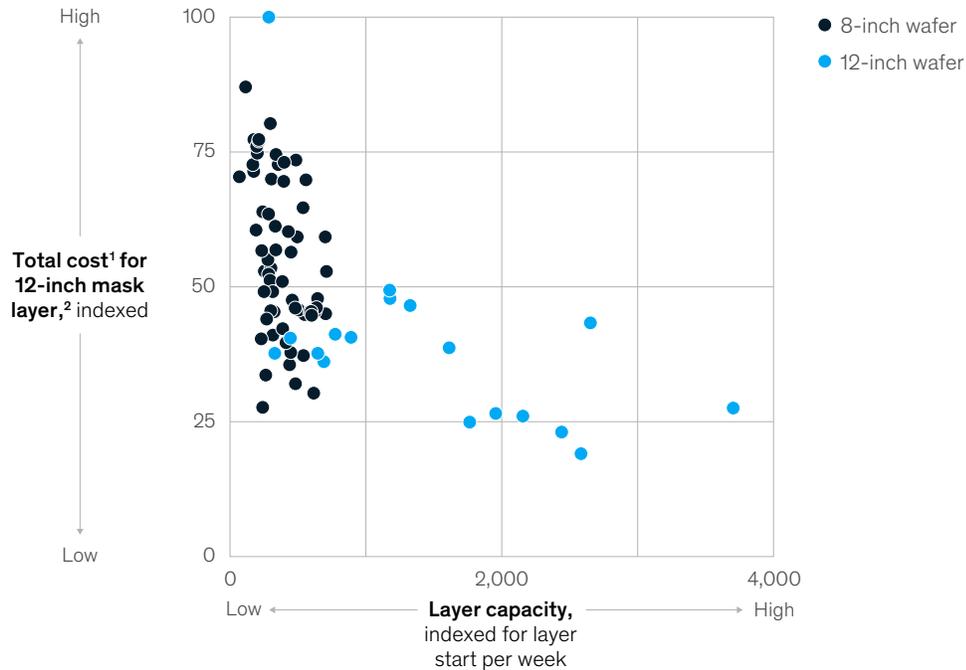
Scaling can help semiconductor players standardize processes and reduce issues related to time zone or language differences. Large fabs, or groups of fabs, will also have greater leverage when negotiating with local suppliers, since they will be purchasing greater volumes. From a knowledge perspective, scaling will help companies identify and share best practices gleaned from extensive production experience.

If a company's product or service is even slightly better than a competitor's, it typically captures an outsize portion of industry revenue.

Exhibit 5

Large-scale operations can significantly reduce costs at fabs.

Total costs for 12-inch mask layer based on capacity



¹Includes both operating expenditures and capital expenditures.
²For equivalent size comparison, the cost of an 8-inch mask layer is multiplied by 2.25.
 Source: McKinsey analysis

Fab clustering

When building a fab, location is one of the most important considerations. Whenever possible, semiconductor players should try to place new facilities in industry clusters—places where several semiconductor companies, or related businesses, have established a strong presence within a relatively small area. These clusters create an environment that encourages collaboration and promotes synergies among companies, even traditional rivals. Cluster participants can achieve a much higher level of performance and greater international visibility than a single company operating independently.

Industry clusters develop in places where several semiconductor companies, or related businesses, have established a strong presence. Within the

United States, for instance, clusters exist in Silicon Valley, Phoenix, upstate New York, and Austin for chip manufacture and equipment. Other clusters include those in China, Germany, Japan, Singapore, South Korea, and Taiwan. Clusters are normally led by one or more world-class manufacturers and a nearby research institute—a constellation that attracts small-to-medium enterprises and eventually creates an ecosystem to support research and manufacturing. Semiconductor companies located in industry clusters will also have access to a larger talent pool, since potential employees with the right skills gravitate to the area and local academic institutions may have ties to local companies that facilitate student recruitment. The risk of supply-chain disruptions is also lower if companies are receiving components or other inputs from nearby vendors.

Industry clusters can help participants reduce costs because they create the potential to share utilities and logistics costs. For instance, several companies may be able to use the same warehouses or consolidate their deliveries into a single shipment. The nearby presence of suppliers may also help if a semico needs assistance, such as quick help from a vendor's manufacturing technician.

As local governments see industry clusters grow and become more important to the regional economy, they may be more likely to subsidize fab construction, invest in leading-edge R&D, or provide tax incentives that reduce operational costs. This support, above all, may motivate semiconductor leaders to become part of an industry cluster.

Creating an environment for successful semiconductor design and manufacturing

Beyond a solid strategy for scaling and clustering fabs, semiconductor companies must develop long-term initiatives for improving their research, supply chain, and talent functions. Simultaneously, they must learn how to navigate ever-changing government policies and ensure IP protection.

Promoting leading-edge research

Semiconductor companies understand that their success hinges on the strength of their R&D programs, especially the ability to increase chip performance while shrinking structure size. But the global economic uncertainty arising from the COVID-19 crisis could hit industry revenues hard, reducing the internal funding available for innovation. External investors may also retreat, given the long development timelines for many leading-edge semiconductor innovations. Consider extreme ultraviolet technology, which received its first research funding around 1995. Now, 25 years later, it is finally being deployed in production. Semiconductor research also has a high element of risk, since companies often investigate unproven technologies and processes. For instance, semiconductor companies might expand their business to quantum computing, but it is difficult to predict when demand for this technology might accelerate (see sidebar “Quantum computing and the semiconductor industry”).

Quantum computing and the semiconductor industry

Some high-tech leaders consider quantum computing to be the next great technological advance, and industry insiders are closely following developments in this area. A few have even questioned whether the new technology will eventually replace today's computers for all functions. But based on the current state of quantum computing, we believe that it will complement current computers, rather than replace them—at least over the short and medium term. For instance, quantum computing could be incorporated into devices or applications to decrease the processing time for encryption and specific data processing tasks. With other functions, traditional computers will remain the default option because quantum computers do not provide additional benefits.

As companies continue to invest in quantum computing, this technology could eventually produce game-changing advances in processing power and speed for multiple functions. Simultaneously, quantum computing could decrease in cost. If semiconductor companies closely follow developments in this segment and are prepared to wait for a return on investment, they could eventually find many opportunities to capture value.

In this environment, government support for research—always important—may become even more essential to success. Companies may have to step up their efforts to secure public funding and select fab locations with this in mind. As discussed earlier, companies that are part of an industry or company-specific cluster may have an easier time obtaining public funding because of their importance to the local economy.

Companies may also gain an R&D edge by closely tracking research activities from academia and start-ups, including publications about the use of new materials and technologies. This vigilance will help ensure that semiconductor companies do not overlook potential market movements.

Industry clusters can help participants reduce costs because they create the potential to share utilities and logistics costs.

Finally, semiconductor companies should enhance R&D by taking a comprehensive, cross-portfolio view of their activities. They should have a clear logic for prioritizing R&D projects and adhere to it diligently. If semiconductor companies consistently monitor R&D spending, reprioritize resources, and ensure that funding goes to areas that could see high future demand, they will maximize their returns.

Increasing supply-chain resilience

Semiconductor companies obtain components and equipment from suppliers around the world, with a few vendors dominating the market for certain products and services. These dynamics leave semiconductor companies vulnerable to supply-chain disruptions. If a key vendor cannot ship products or meet delivery deadlines, production lines could grind to a halt. More widespread upheaval, such as the global lockdowns in response to the COVID-19 pandemic, have even greater repercussions.

Adding to the supply-chain risk, semiconductor demand is volatile. To accommodate unexpected shortages, companies need flexible and resilient supply chains that can quickly adjust. Sourcing critical parts from multiple vendors may increase supply-chain resilience; if that is not possible, semicos could stockpile essential components and materials to guard against future disruptions.

While such measures may keep production moving, they also raise costs. At present, while the COVID-19 crisis is creating massive upheaval, such outlays may be justified. Over the longer term, however, semicos must increase resiliency by closely monitoring the supply chain, identifying weak spots, and building

a strategy to address them. For instance, they could identify and upskill local vendors to serve as a second supply source for critical resources. The exact solution will vary by business, since each company has different needs and vendor networks.

Gaining access to global talent

Semiconductor companies that are not part of an industry cluster may struggle to recruit employees, as the local talent pool might not offer enough people with the right expertise and specialized knowledge. If companies cannot join a cluster, they could recruit more foreign talent, provided that all laws and regulations are met. Some governments may assist with this effort—for instance, by relaxing visa requirements—to stimulate the local semiconductor market. Alternatively, companies can look into building R&D centers in new locations, such as Eastern Europe or India. Although some areas might lack talent for chip design, they offer a rich pool of software developers. With software gaining importance for the semiconductor industry, semicos may get a competitive advantage from combining their existing chip-design expertise with new software capabilities.

Navigating government policies and protecting intellectual property

Governments support their local semiconductor industries in multiple ways: setting regulations that favor growth, providing financial support for innovation, and creating a favorable environment to attract talent. Ideally, governments would maintain that support consistently, but the past few years have seen frequent policy shifts in many locations. Some of these have been detrimental to the industry, such as immigration regulations that limit the entry

of skilled workers. But governments have also recently expressed greater interest in supporting their local semiconductor industries, with some offering new subsidies. Semiconductor companies that monitor policy changes and rapidly adjust can help their businesses maintain healthy growth and capture new opportunities.

Semiconductor companies should likewise monitor the patent environment to protect their IP. The need for this protection is greater than ever, since many new players in other industries, including automotive companies, hyperscalers, smartphone companies, and start-ups, have started to develop and patent their own chips. Without a reliable IP policy, semiconductor companies may have difficulty sustaining innovation, since the payback period is long.

Building leading-edge chips is a challenging and cost-intensive process that requires a huge investment in R&D and even greater expenditures for fab construction. To attain excellence in semiconductor design and manufacturing while still keeping costs in check, companies need a strategic plan for R&D investments and for potentially building new fabs—perhaps the first they have ever owned. They must also consider opportunities related to government incentives and subsidies, as well as the beneficial network effects they could obtain as members of a semiconductor cluster. These actions, when combined with other improvement initiatives in supply chain, IP, and other critical areas, will help semiconductor companies achieve manufacturing excellence in an increasingly competitive landscape.

Harald Bauer is a senior partner in McKinsey's Frankfurt office; **Ondrej Burkacky** is a partner in the Munich office, where **Stephanie Lingemann** is an associate partner and **Klaus Pototzky** is a consultant; **Peter Kenevan** is a senior partner in the Tokyo office; and **Bill Wiseman** is a senior partner in the Seattle office.

The authors wish to thank Stefan Burghardt, Larissa Rott, and Klaus Seywald for their contributions to this article.

Designed by McKinsey Global Publishing
Copyright © 2020 McKinsey & Company. All rights reserved.