McKinsey & Company

McKinsey on Semiconductors

Creating value, pursuing innovation, and optimizing operations

Number 8, November 2021

McKinsey on Semiconductors is written by experts and practitioners in McKinsey & Company's Semiconductors Practice along with other McKinsey colleagues.

To send comments or request copies, email us: McKinsey_on_ Semiconductors@ McKinsey.com.

Cover image:
© leezsnow/Getty Images

Table of contents

Introduction

When chip shortages first shut down automotive production lines in 2021, the semiconductor industry found itself in an unaccustomed spotlight. Suddenly everyone was talking about the tiny chips that enabled so many different car functions, from interior lighting to seat control to blind-spot detection. When some high-tech and consumer-electronics companies began to experience chip shortages or voiced concerns about supply chains, the attention intensified. It's now clear to all: we are living in a semiconductor world.

Semiconductor companies have increased throughput to meet surging demand, with revenue growth of about 9 percent expected in 2021—up from the approximately 5 percent seen prepandemic in 2019. And some governments are upping their investment in semiconductor technology to lessen the impact of global supply-chain disruptions. But the complexities of semiconductor production, combined with the long timelines for building new fabs and uncertain demand patterns, are presenting many challenges.

This issue of *McKinsey on Semiconductors* reflects on the unexpected developments of the past two years and looks ahead to the postpandemic future. In the first article, "Value creation: How can the semiconductor industry keep outperforming?," we analyze sector performance, including the accelerated growth seen in recent years. The article then examines future demand drivers, including self-driving cars, the Internet of Things, artificial intelligence, blockchain technology, and the coming shift to the 5G connectivity standard. One major finding: as semiconductor companies pursue new opportunities, they will benefit from tapping into M&A and expanding partnerships.

Continuing with this theme, "Semiconductor design and manufacturing: Achieving leading-edge capabilities" discusses how to maximize value. The article offers a practical look at core competencies, including the best way to scale and cluster fabs, strategies for improving the research and supply-chain functions, and a new approach to talent recruitment. Next, in "Scaling AI in the sector that enables it: Lessons for semiconductor-device makers," we make the case for pursuing innovative technologies that could take value to new heights.

The next three articles focus on the automotive sector, where customer segments will evolve significantly in the coming years. The first, "Coping with the auto-semiconductor shortage: Strategies for success," examines the root causes of the shortage, including complications related to the COVID-19 pandemic, and then offer solutions that automotive OEMs might consider when dealing with the imbalance of supply and demand over both the short and long term.

In "Automotive semiconductors for the autonomous age," the authors describe how semiconductor companies can enable some of the most sophisticated and essential vehicle systems in a growing market. The focus on autonomous vehicles (AVs) has already altered demand patterns for automotive semiconductors, with sales of specialty silicon—chips tailored to specific applications—growing steadily. Semiconductor companies, OEMs, and tier-one suppliers must understand how the market is evolving to keep production lines running and revenues strong.

Sophisticated AVs require more software content, and the cyberrisk of connected cars has become a critical issue because security researchers have identified technical vulnerabilities that could allow hackers to control vehicle functions. In "Cybersecurity in automotive: Mastering the challenge," the authors explore a vital question: How can OEMs secure hardware and software while meeting regulatory requirements and satisfying customer expectations?

Although the automotive sector drives much of the business for semiconductor companies, smaller, emerging segments also deserve attention. “The next wave of innovation in photonics” presents some of the opportunities being created as more end products integrate lasers with sensors and optics. The authors describe the market for these components and then discuss next steps for companies that want to make their mark in photonics.

The issue concludes with perspectives from two industry leaders: “Navigating through change: An interview with NXP Semiconductors’ Kurt Sievers” and “Renesas’s Hidetoshi Shibata on leadership through difficult times.” Both Sievers and Shibata reflect on the challenges of the past year, as well as their hopes for their companies. We also get their take on opportunities in specific sectors, including automotive.

We hope that these articles provide food for thought as you consider your next steps, and we welcome questions and comments about them, especially those that relate to your own experience.

Ondrej Burkacky
Senior partner,
Munich

Abhijit Mahindroo
Partner,
Silicon Valley

Value creation: How can the semiconductor industry keep outperforming?

The industry has made vast profitability gains in recent years. Can it seize the moment to accelerate value creation even further?

by Ondrej Burkacky, Marc de Jong, Ankit Mittal, and Nakul Verma

The massive global disruptions caused by the COVID-19 pandemic gave the already rapid adoption of digital technologies an extraordinary boost, as electronic devices proved vital to connecting people and businesses during lockdowns. As demand from both consumer and business customers soared in 2020, shareholders in semiconductor companies saw high double-digit returns, despite supply-chain issues and growing divergence in global trade. The shortages across the value chain that resulted from the surge in demand have been accompanied by growing consolidation as chip makers rushed to gain the benefits of scale.

With accelerated digitization likely to continue in the post-COVID-19 world, semiconductor companies might benefit from developing strategies that address the shifts in the competitive landscape. To understand the value-creation trends in the industry, we analyzed its performance over the past two decades, looking at 11 segments and three key global regions. Our findings suggest that semiconductor players might want to consider focusing on gaining leadership in profitable segments by leveraging M&A and partnerships, building agility, and pursuing new technologies and innovations. Burgeoning investments in self-driving cars, the Internet of Things, and artificial intelligence, along with the coming shift to the 5G connectivity standard, present opportunities for further growth and specialization. Semiconductor players that make wise strategic choices now may gain lasting industry leadership.

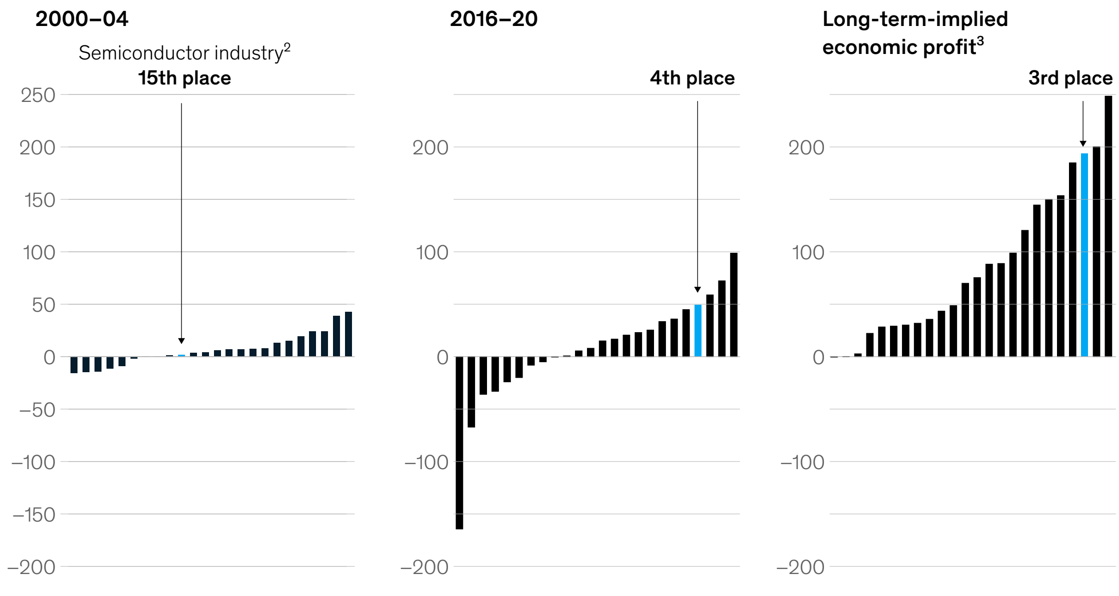
Widespread gains: Shifting pools across two decades

Since the start of the new millennium, the semiconductor industry has gone through two distinct phases. In the early 2000s, profit margins were low, and most companies generated returns below the cost of capital. Profitability improved during the past decade, however, spurred by soaring demand for microchips in most industries, the rapid growth of the technology sector, and increased cloud usage, as well as ongoing consolidation in many subsegments.

Exhibit 1

In recent years, the semiconductor industry's economic profitability has improved relative to others, and this trend is expected to continue.

Average economic profit (EP), by industry, $ billions (n = 2,644[1] companies in 24 industries)

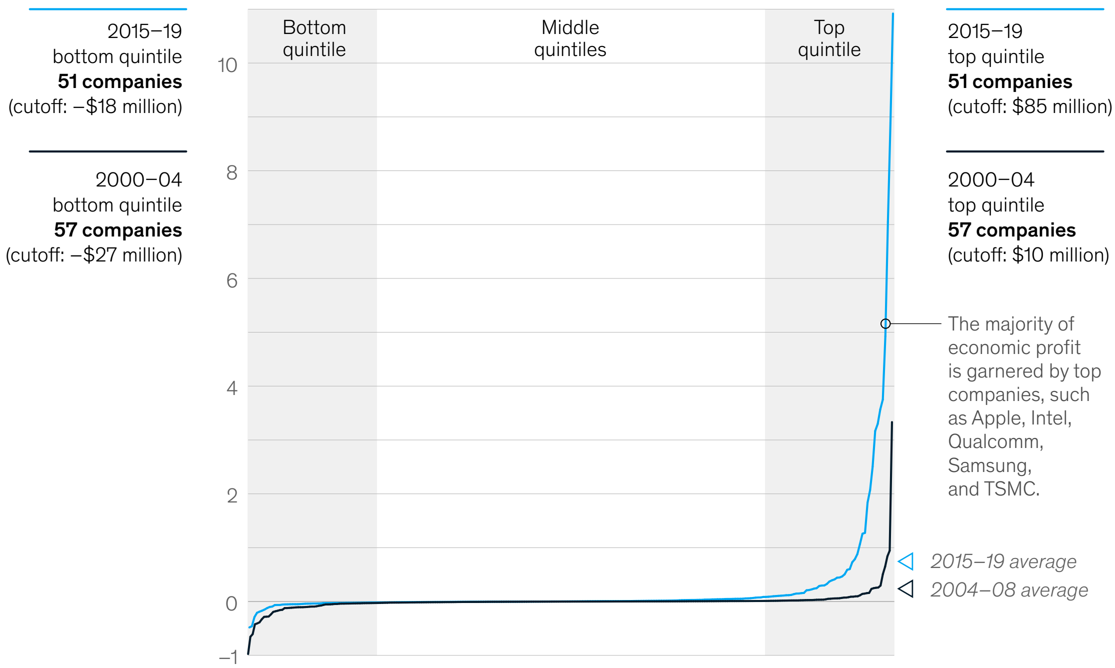


[1]Top publicly listed companies by revenues.
[2]Semiconductor industry's position versus other industries (based on top pure-play companies only).
[3]Long-term-implied economic profit is based on Sept 13, 2021, market valuations.
Source: McKinsey Strategy Practice and Corporate Performance Analytics

Exhibit 2

Over the past two decades, leading companies have captured a greater share of the semiconductor industry's profits.

Average annual economic profit of semiconductor companies, by quintile,[1] $ billions



[1]Total sample of ~380 companies, with data available from 283 companies (2000–04) and from 254 companies (2015–19).
Source: Company reports; Capital IQ; Corporate Performance Analytics by McKinsey

One consequence is that the semiconductor industry's profit pool (defined by aggregated economic profitability) improved significantly relative to other industries (Exhibit 1). The industry's power curve of economic profit,[1] which covers approximately 2,600 top companies around the world in 24 industries, shows that semiconductors rose from 14th place during the period from 2000 to 2004 to fourth place from 2016 to 2020. Chip makers' aggregate annual economic profit, which totaled $3.5 billion for the first period, increased dramatically to $49.3 billion during the second period. Average profitability peaked in 2017 and 2018, but pricing pressure in memory chips caused a significant decline in 2019. Late 2020 saw a rebound in profitability, however. Interestingly, the semiconductor industry's relative ranking versus other industries is expected to continue improving and reach third place over the long run. (Note: to understand the expected shift in profit pools, we used current market capitalization to calculate market-implied long term economic profit for each industry.)

The strong are growing stronger

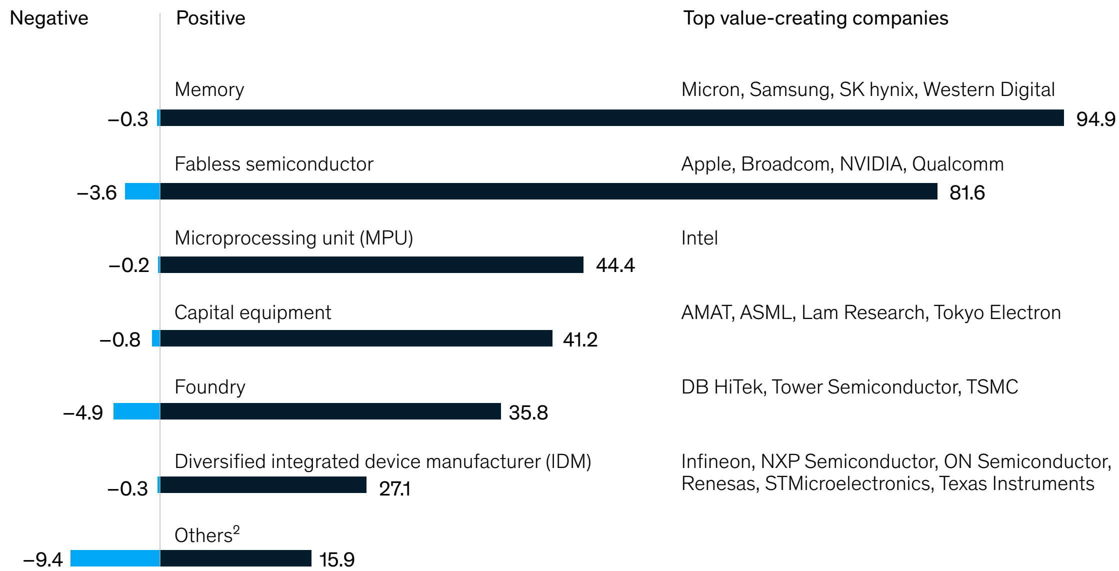
Although the semiconductor industry's economic profit has substantially increased, companies and industry segments vary significantly because value pools have shifted over time, and the strongest players have increased their lead over competitors. The industry power curve has steepened sharply at the top during the past five years: the highest quintile of companies captured a majority of the economic profit from 2015 to 2019 (Exhibit 2). The gap between the leaders and the laggards is widening as the strongest players take advantage

[1] Economic profit is the amount left over after subtracting the cost of capital from net operating profit. We chose to focus on economic profit because this metric comprehensively captures both profitability and the opportunity cost of the capital deployed.

Exhibit 3

From 2015 to 2019, the semiconductor industry's memory and fabless segments generated the greatest value.

Cumulative economic profit[1] value creation, 2015–19, by segment, $ billions



[1]Economic profit is calculated as NOPLAT – (capital charge, where capital charge is invested capital, including goodwill at previous year-end × WACC [weighted average cost of capital]); based on sample of ~380 companies.
[2]Includes EDA, IP, analog IDM, IDM others, and packaging and assembly.
Source: Company reports; Capital IQ; Corporate Performance Analytics by McKinsey

of their scale and diversified customer base to entrench their dominant positions. The difference in average economic profit between the tenth and 90th percentiles—approximately 140 percent from 2000 to 2004—swelled to a staggering 400 percent in the period from 2015 to 2019.

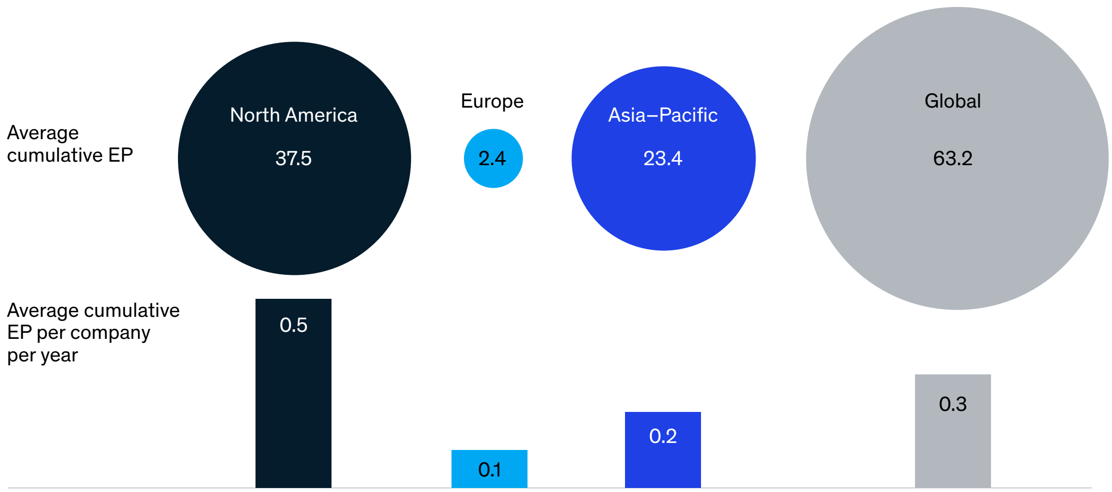
Ebbs and flows in value are also evident in the relative performance of industry categories. For instance, a look at company performance shows that Intel captured almost all the economic profit in the early 2000s. When looking at product categories, analysis shows that five segments generated the most value: memory, microprocessor units (MPU), fabless, capital equipment, and foundry. From 2015 to 2019, these five captured more than 60 percent of the industry's $335 billion in cumulative economic profit (Exhibit 3). Memory manufacturers benefitted from a surge in demand for electronic devices and rising prices up to 2018, when oversupply and pricing declines started to depress returns. Fabless was the second-best performer during this period, and Apple is estimated to have earned roughly a quarter of the total economic profit in that category.

The trends driving these patterns of performance are likely to persist. The industry continues to move toward the fabless production model as companies seek to leverage the benefits of leading-edge technology while sharing the necessary investments. Apple's M1 chip (for notebooks, low-end desktops, the Mac Mini, and tablets) exemplifies this move into in-house chip design, which leverages foundries to manufacture products. Even companies with well-established in-house manufacturing facilities, such as Intel, are considering partial outsourcing to chip foundries to benefit from greater production flexibility and cost reductions.

Exhibit 4

All regions contributed to the semiconductor industry's global value pool.

Average annual economic profit (EP)[1] of semiconductor companies, 2015–19, $ billions



[1]Economic profit is calculated as NOPLAT – (capital charge, where capital charge is invested capital, including goodwill at previous year-end × WACC); based on sample of ~380 global companies (with data available from 254 companies 2015–19).
Source: Company reports; S&P Capital IQ; Corporate Performance Analytics by McKinsey

The distribution of the semiconductor industry's economic profit also varies significantly among regions (Exhibit 4). North America, home to some of the largest fabless players (such as Apple, NVIDIA, and Qualcomm), accounted for approximately 60 percent of the global value pool during the 2015–19 period. Europe accounted for 4 percent of the industry's total economic profit, which accrued primarily to capital-equipment companies. Asia, still the hub for contract chip manufacturing, accounted for the remaining 36 percent of the value the sector generated.

What capital markets expect of the industry

Capital markets have rewarded the industry's surging profitability: semiconductor companies delivered an annual average of 25 percent in total returns to shareholders (TRS) from the end of 2015 to the end of 2019 (Exhibit 5). Last year, shareholders saw even higher returns, averaging 50 percent per annum, as consumers and businesses upped their purchases of digital equipment of all kinds. Investors expect this trend to continue.

As shareholders look to capitalize on high valuations, scale has become more important, and the industry has consolidated rapidly. Its M&A activity has reached the second-highest pace in history: $118 billion worth of deals in 2020. Our analysis of capital-market sentiment suggests that investors count on continued strong growth in the semiconductor industry. More than half of its current enterprise value is based on earnings-growth expectations, which are also reflected in current valuations: investors expect long-term growth of 7 to 8 percent, assuming the recent margin trajectory.

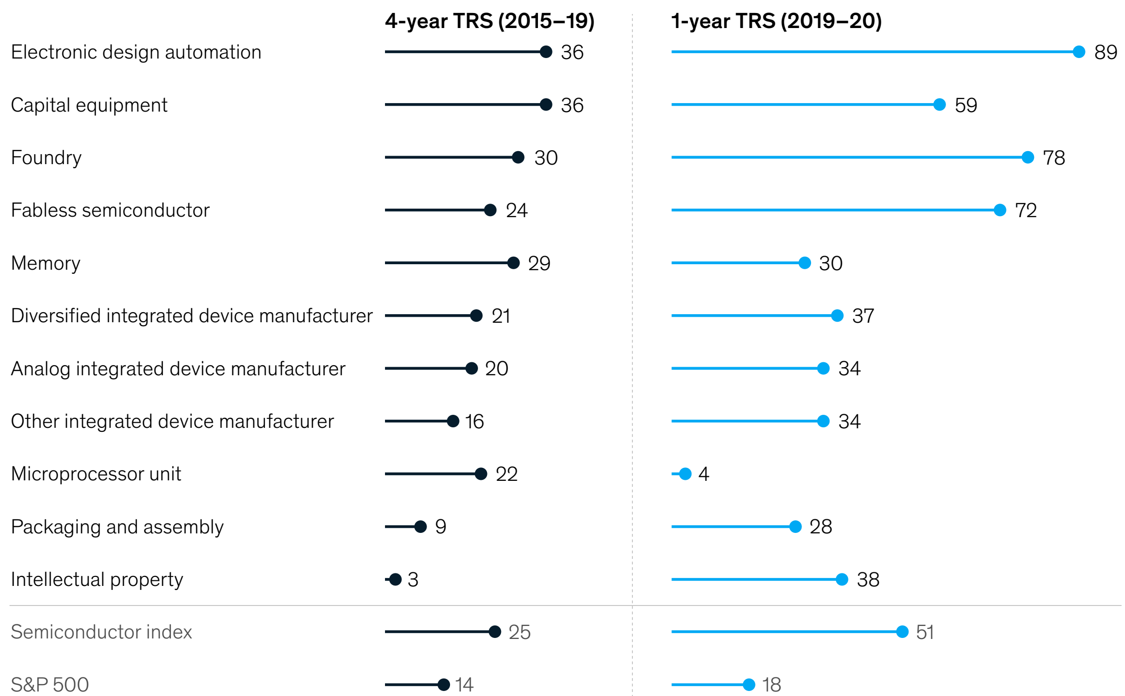
Preparing for the next phase of growth

To meet investors' expectations for continued high growth amid a shifting industry landscape, chip makers can take inspiration from the leading

Exhibit 5

The semiconductor companies' total returns to shareholders have increased significantly in recent years.

Total returns to shareholders (TRS), %



[1]Based on a sample of ~380 semiconductor companies, excluding conglomerates: Apple, Fuji Electric, Fujitsu, Nikon, Samsung, Sony, and Toshiba.
Source: Company reports; S&P; Corporate Performance Analytics by McKinsey

players and the tactics that they followed to create the majority of the industry's economic profit. Securing leadership in profitable segments has been their key success factor. By continued capital investments or R&D, the leading players have developed distinctive offerings that have further extended their leadership position. Notably, recent history shows that companies find it hard to catch up with rivals that lead in well-defined technology areas. Together, these findings suggest that semiconductor companies should explore three avenues when developing strategies to thrive in the postpandemic future.

Rethink collaboration with other players

Partner across the value chain to expand the customer base. Industries increasingly require application-specific solutions, such as those that automakers embed in self-driving vehicles. Many such requirements come from players that previously did not design their own integrated circuits. While semiconductor companies must ensure that their order volumes are sufficient to justify the rising R&D cost of leading-edge custom chip design, efforts to tap into bespoke requirements of customers down the value chain could provide pathways into high-growth industry niches.

Develop a programmatic M&A strategy. Amid ongoing industry consolidation, semiconductor companies might want to consider developing a programmatic M&A strategy—a serial approach to smaller acquisitions, along a specific theme—that would guide their approach to acquisitions aimed at branching into adjacent areas or adding capabilities essential for future growth. The current scarcity of targets requires potential acquirers to investigate and execute mergers quickly. Chip makers might also want to consider making major deals that could open up important new markets. For instance, in a deal now under review by regulators, graphics-card maker NVIDIA has proposed acquiring Arm, an intellectual-property specialist for mobile applications, from SoftBank. If approved, this acquisition would allow NVIDIA to gain access to the broader computing market.

Stay agile and responsive to a more volatile world
Acknowledge that supply chains are changing. Semiconductor players can gain an advantage by increasing the resilience of their supply chains as global trade diversifies, particularly for sophisticated technologies. Several large chip manufacturers are already exploring diversifying production so that they can rely on more than one vendor or supplier at a time. These moves are motivated in part by new government subsidies designed to support capabilities for manufacturing advanced chips.

Strengthen pricing and allocation strategies in response to supply shortages. Especially in the automotive and industrial sectors, chip shortages could become the new normal, so semiconductor companies would benefit from thinking carefully about allocating inventory and about fair pricing strategies. These companies can also explore the potential of inviting customers to co-invest in the development of custom chips, which would help buyers to reduce the risk of supply shortages while assuring manufacturers of real demand for new designs. Chip makers could also engage with the broader industry to explore solutions to the continuing shortages.

Drive adoption of new technologies and innovations
Use advanced analytics to reimagine time to yield and yield ramp-ups. Chip makers could partner with equipment manufacturers to apply advanced analytics to speed up the yield learning curve. For example, modeling made possible by advanced combinatorial learning could replace the physical testing of chips and thus reduce both the cost of introducing them and their time to market.

Leverage innovation within and beyond Moore's law. Innovation in line with Moore's law (which predicts the shrinking of chip structures) will certainly continue, but additional advances through system-on-a-chip architectures using "chiplets" may be possible as well. Manufacturers could also explore innovations beyond Moore, such as compound semiconductors that use silicon carbide and gallium nitride, which could provide performance benefits superior to those of traditional silicon.

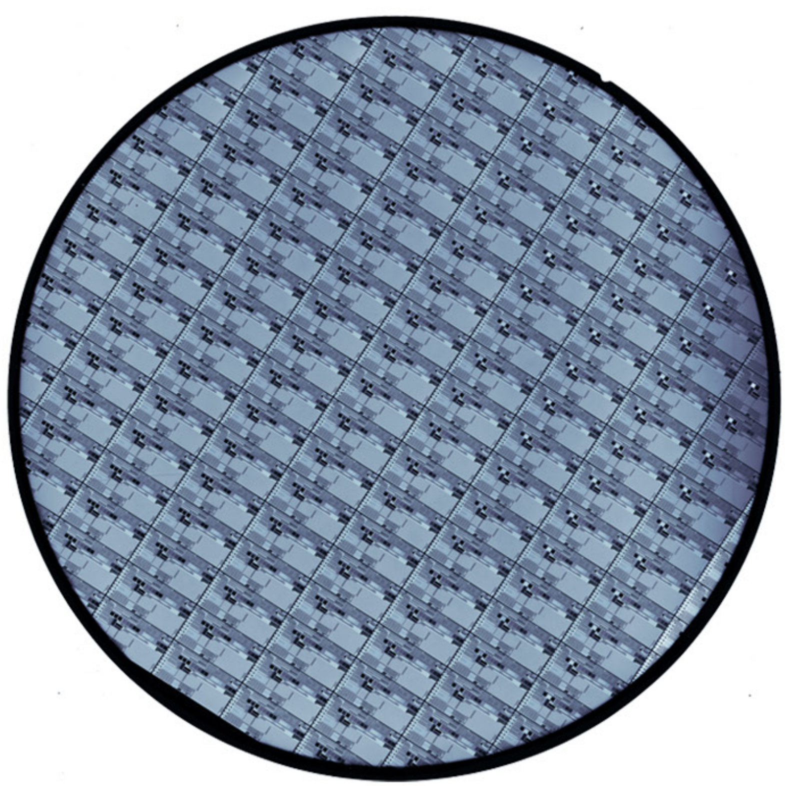
Following a period of rapid growth, leaders of semiconductor businesses should prepare for a world where increasingly challenging demand–supply matching, geopolitical issues, and a need for specialized products will make novel demands on the industry. To meet shareholders' expectations for continuing high returns, semiconductor players could expand their partnerships and look for industry-wide solutions to product shortages and supply-chain challenges.

Ondrej Burkacky is a senior partner in McKinsey's Munich office; **Marc de Jong** is a senior partner in the Amsterdam office; and **Ankit Mittal** is a knowledge expert in the Gurugram office, of which **Nakul Verma** is an alumnus.

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 is a collaborative effort by Harald Bauer, Ondrej Burkacky, Peter Kenevan, Stephanie Lingemann, Klaus Pototzky, and Bill Wiseman, representing views from McKinsey's Advanced Electronics Practice.



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

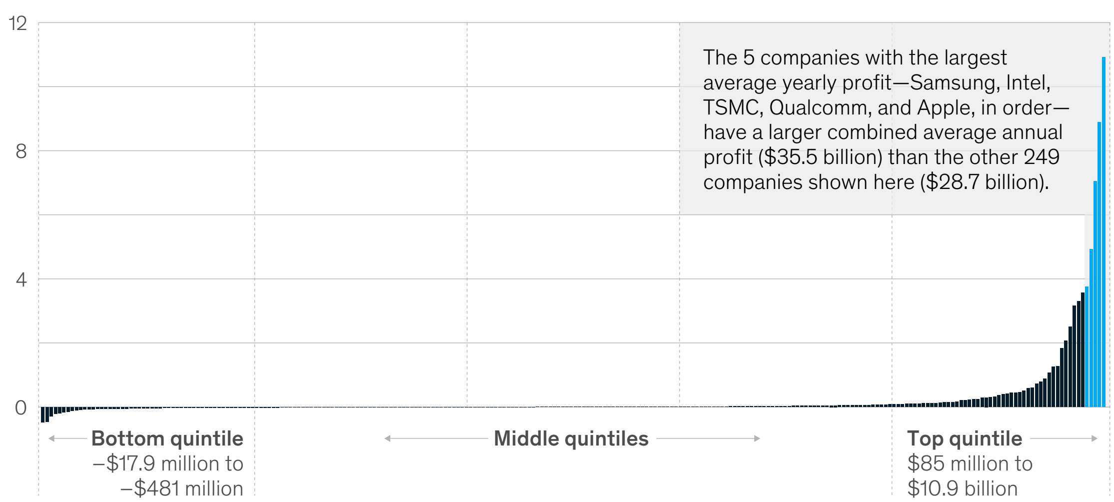
The smallest structures in chips, which are five nm and three nm in size, confer important benefits within select market segments and application areas, including data centers, 5G smartphones, edge computing, and machine learning. Semiconductor customers in these segments will pay a premium for leading-edge 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, $ billions



[1]Total 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

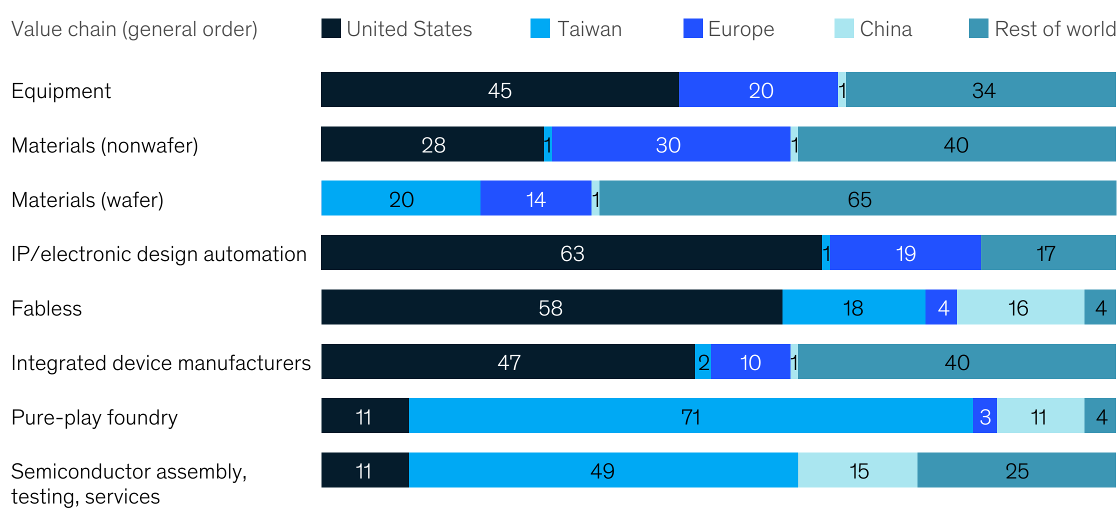
Leadership within manufacturing and R&D is not simply a tool for winning the market; it is also essential to cost control. That benefit is now more important than ever because expenses have soared over the past ten years.

First, consider fab construction. The cost of building and equipping a facility with five 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 ten 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 five nm chip costs about $540 million for everything from validation to IP qualification. That is well above the $175 million required to design a ten nm chip and the $300 million required for a seven nm chip. We expect that R&D costs will continue to escalate, especially for leading-edge products.

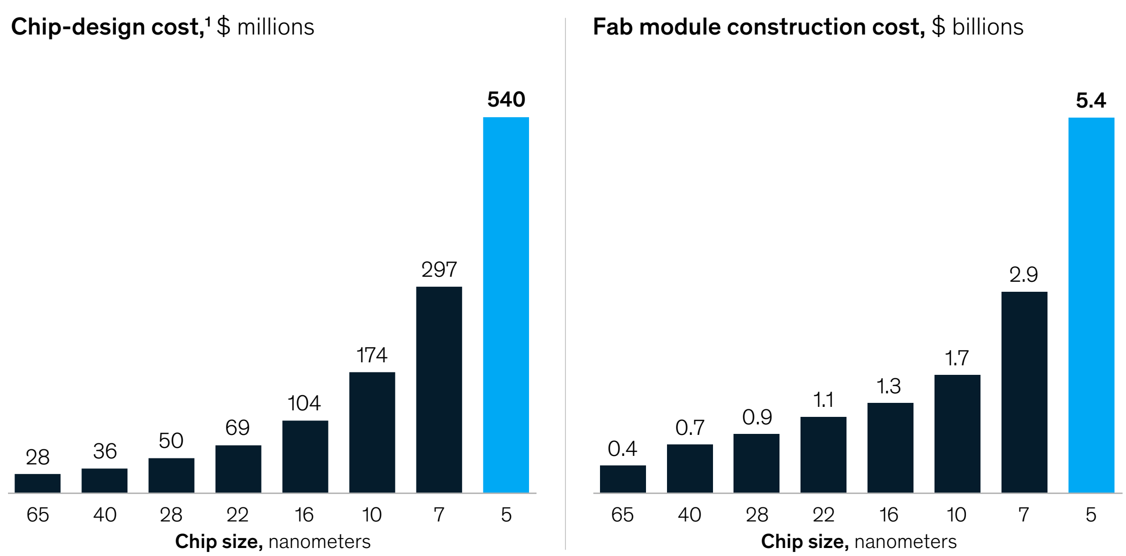
Although semiconductor companies must devote billions to new fabs, they will not see an immediate return on their investment. It takes about 12 to 24 months to build a shell of a fab and install the required tools, plus another 12 to 18 months to ramp up to full capacity. And if demand falls beyond projections,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



[1]Major 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

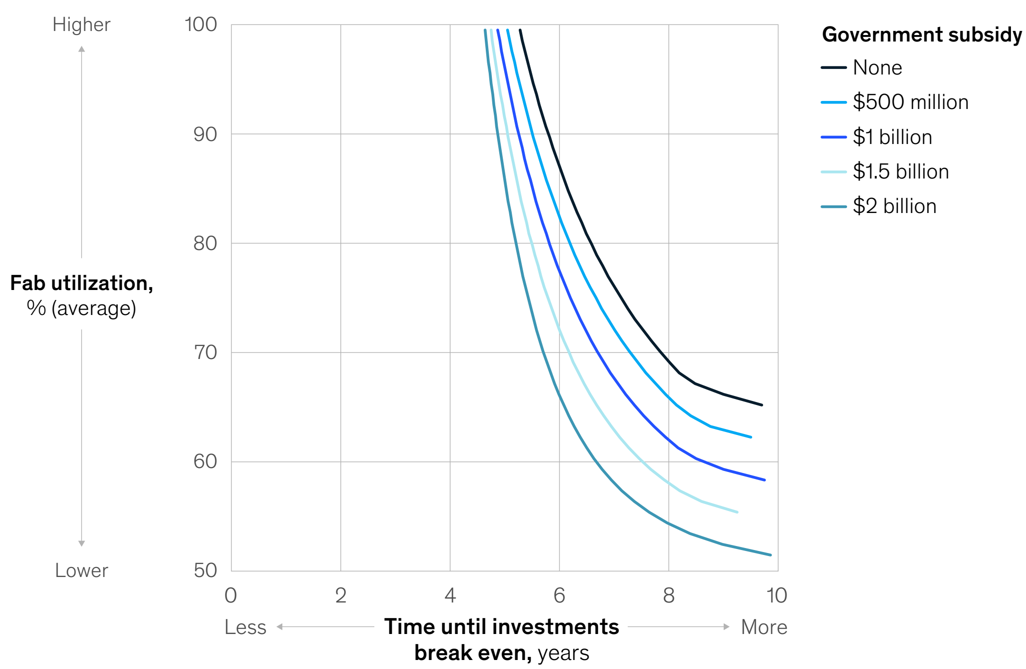
In a best-case scenario, with high utilization, our model shows that semiconductor companies can expect to see a positive cash flow from their new fabs within about five years, even with little to no government support (Exhibit 4). If the fabs are utilized below capacity, however, this time frame could 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 semiconductor company 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

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

also consider the impact of any government subsidies, since that could significantly shorten 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 semiconductor company'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.

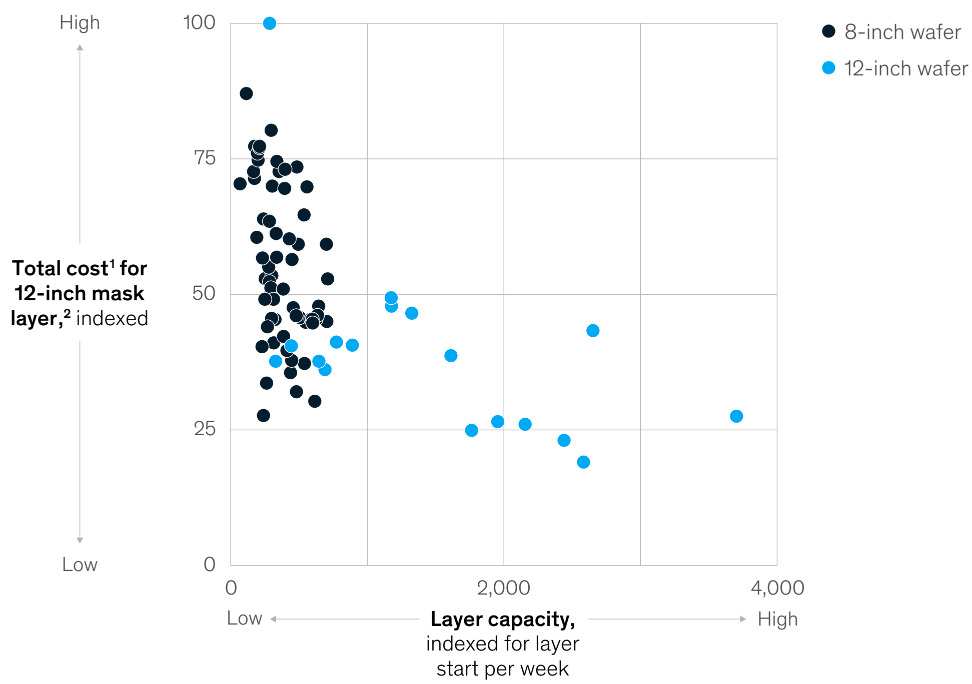
Scaling can help semiconductor players standardize processes and reduce issues related to time zone or language differences. Large fabs, or groups of fabs, will also have greater leverage when negotiating with local suppliers, since they will be purchasing greater volumes. From a knowledge perspective, 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



[1]Includes both operating expenditures and capital expenditures.
[2]For 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 semiconductor company 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, semiconductor companies 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, semiconductor companies 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 example, 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, semiconductor companies 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.

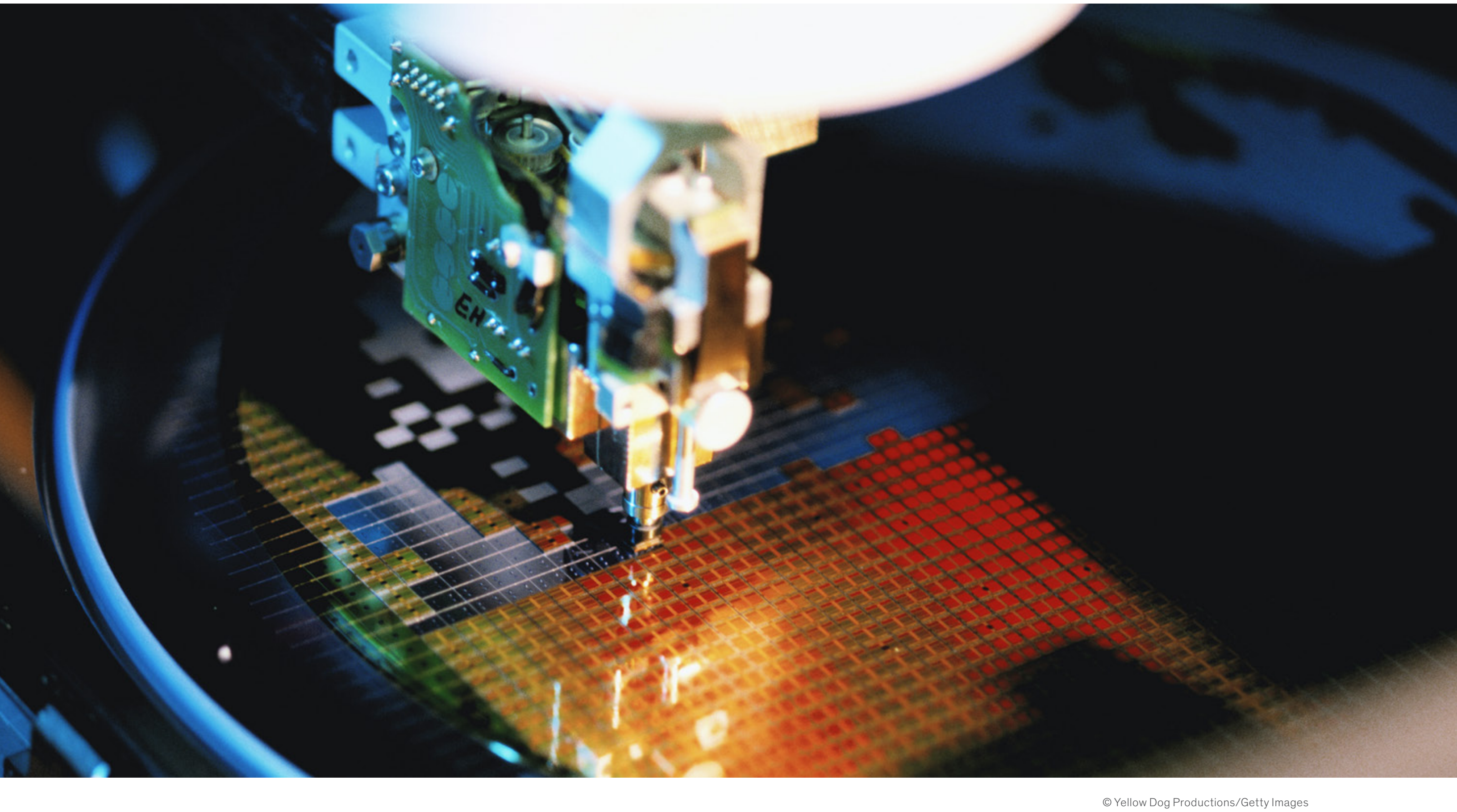
Harald Bauer is a senior partner in McKinsey's Frankfurt office; **Ondrej Burkacky** is a partner in the Munich office, where **Stephanie Lingemann** is an associate partner and **Klaus Pototzky** is a consultant; **Peter Kenevan** is a senior partner in the Tokyo office; and **Bill Wiseman** is a senior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

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

Scaling AI in the sector that enables it: Lessons for semiconductor-device makers

Artificial intelligence has significant value-creation potential in the semiconductor industry. How can semiconductor companies deploy AI at scale and capture this value?

by Sebastian Göke, Kevin Staight, and Rutger Vrijen



Artificial intelligence/machine learning (AI/ML) has the potential to generate huge business value for semiconductor companies at every step of their operations, from research and chip design to production through sales. But our recent survey of semiconductor-device makers shows that only about 30 percent of respondents stated that they are already generating value through AI/ML. Notably, these companies have made significant investments in AI/ML talent, as well as the data infrastructure, technology, and other enablers, and have already fully scaled up their initial use cases. The other respondents—about 70 percent—are still in the pilot phase with AI/ML and their progress has stalled.

We believe that the application of AI/ML will dramatically accelerate in the semiconductor industry over the next few years. Taking steps to scale up now will allow companies to capture the full benefits of these technologies.

This article focuses on device makers, including integrated device manufacturers (IDMs), fabless players, foundries, and semiconductor assembly and test services, or SATS (for more information on our research, see sidebar, "Our methodology"). In a future article, we will look more closely at the implications for equipment players.

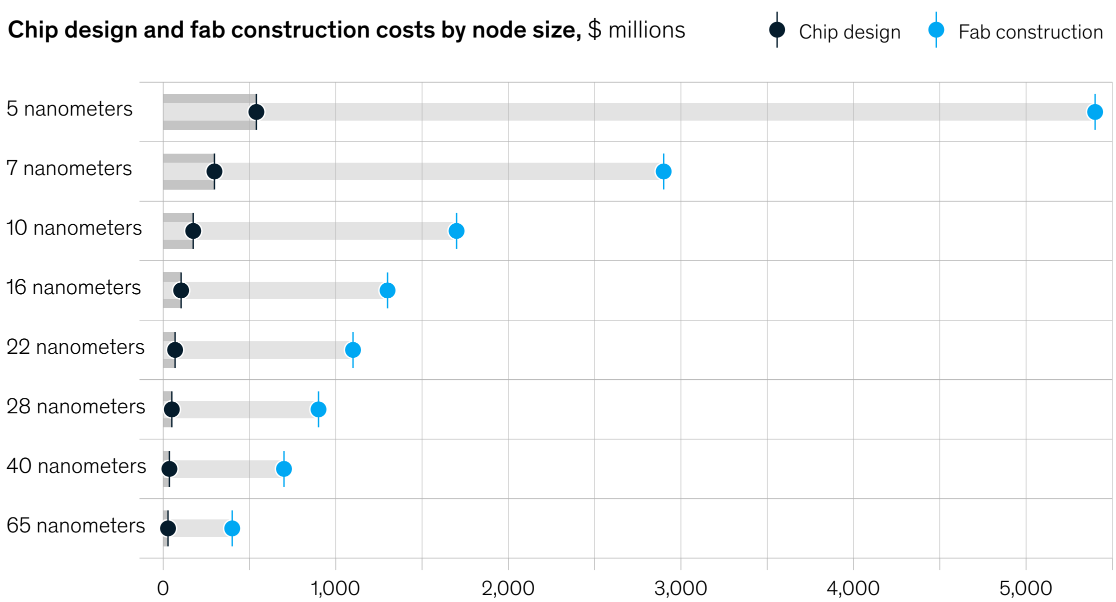
AI's role in tackling the challenges ahead

Because of their high capital requirements, semiconductor companies operate in a winner-takes-most or winner-takes-all environment. Consequently, they have persistently attempted to shorten product life cycles and aggressively pursue innovation to introduce products more quickly and stay competitive. But the stakes are getting increasingly high. With each new technology node, expenses rise because research and design investments, as well as capital expenditures for production equipment, increase drastically as structures get smaller. For example, research and design costs for the development of a chip increased from about $28 million at the 65 nanometer (nm) node to about $540 million at the leading-edge five nm node (Exhibit 1). Meanwhile, fab construction costs for the same nodes increased from $400 million to $5.4 billion.

Exhibit 1

Costs for chip design and fab construction have soared as chips become increasingly complex.

Chip design and fab construction costs by node size, $ millions



Source: IBS; McKinsey analysis

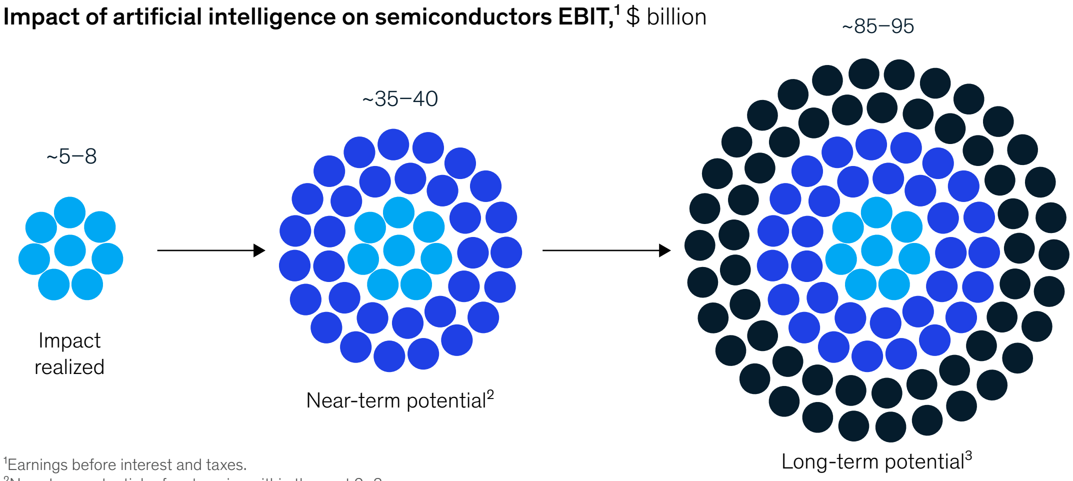
As companies attempt to increase productivity within research, chip design, and manufacturing, while simultaneously accelerating time to market, AI/ML is becoming an increasingly important tool along the whole value chain.

Our research shows that AI/ML now contributes between $5 billion and $8 billion annually to earnings before interest and taxes at semiconductor companies (Exhibit 2). This is impressive, but it reflects only about 10 percent of AI/ML's full potential within the industry. Within the next two to three years, AI/ML could potentially generate between $35 billion and $40 billion in value annually. Over a longer time frame—gains achieved four or more years in the future—this

Exhibit 2

Artificial intelligence could generate $85 billion to $95 billion for semiconductor companies over the long term.

Impact of artificial intelligence on semiconductors EBIT,[1] $ billion



[1]Earnings before interest and taxes.
[2]Near-term potential refers to gains within the next 2–3 years.
[3]Long-term potential refers to gains achieved 4 years or more in the future.

Our methodology

To determine the impact of artificial intelligence (AI) on the semiconductor industry, we first identify state-of-the-art advanced-analytics use cases. We then complement these with use cases from other advanced industries, if applicable, to build a comprehensive map. Use-case domains are defined as areas of AI impact that involve several different use cases. To determine the extent of the overall impact, including the repercussions for revenue, we evaluate every use case within the context of the semiconductor industry, looking at detailed criteria, such as the maturity of the use case, the estimated timeline for value creation, and the relevant impact on the baseline (capital expenditures, cost of goods sold, spending on research and chip design, and costs for selling, general and administrative). Internal and external experts have validated the research.[1]

[1] When evaluating impact, the main focus is on cost reduction because total revenue largely depends on external factors. We also consider revenue gains resulting from increased capture of market share.

figure could rise to between $85 billion to $95 billion per year. That amount is equivalent to about 20 percent of the industry's current annual revenue of $500 billion and almost equal to its 2019 capital expenditures of $110 billion.[1] While a significant portion of this value will inevitably be passed on to customers, the competitive advantage of capturing it, particularly for early movers, will be impossible to ignore.

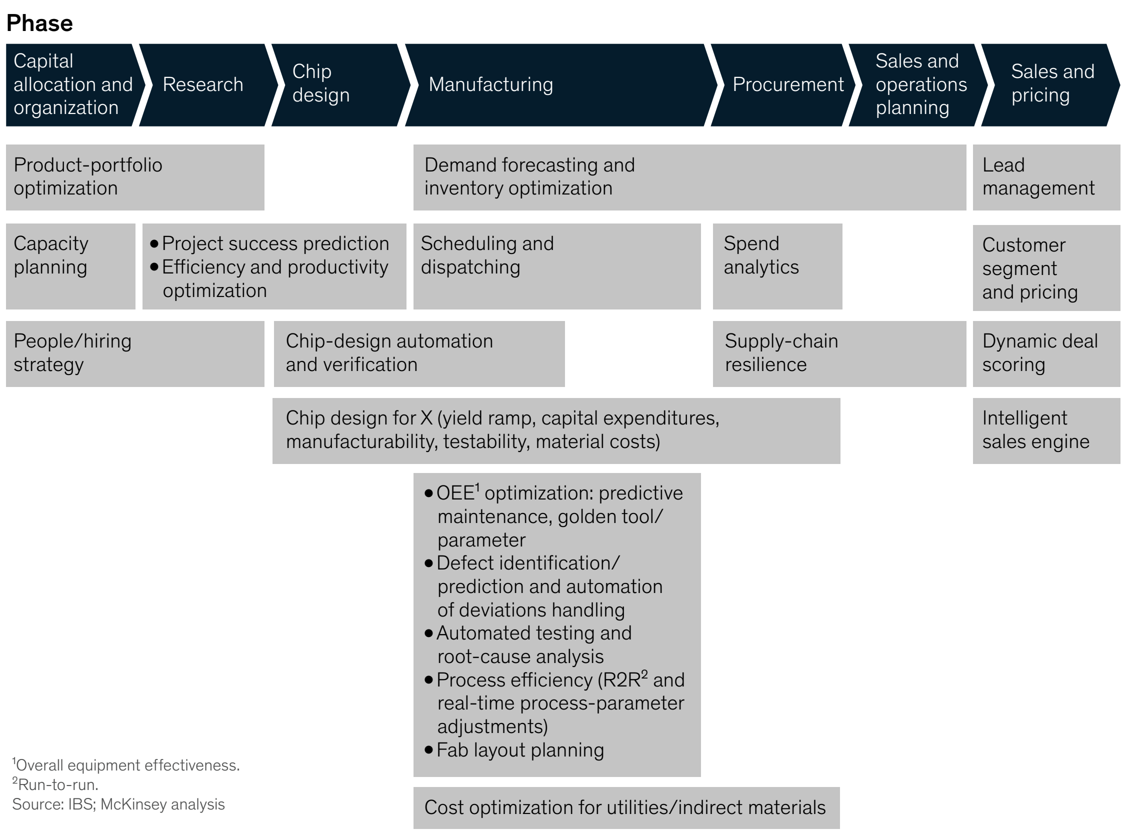
AI/ML use cases in the semiconductor industry

Our comprehensive map of AI/ML use-case domains—areas that contain multiple specific use cases—spans the entire value chain for semiconductor-device makers (Exhibit 3). A use-case domain can also extend across several value-chain activities. For example, the demand-forecasting and inventory-optimization domain is relevant to manufacturing, procurement, and sales and operations planning.

Industry-wide, manufacturing will accrue the most value from AI/ML (Exhibit 4). This is not a surprise given the capital expenditures, operating expenditures, and material costs involved in semiconductor fabrication. The greatest relative spend reduction will occur in research and design, primarily resulting from the automation of chip design and verification. We will investigate the main use cases in the next section.

Exhibit 3

A comprehensive heat map of use cases allows individual companies to focus and set priorities.



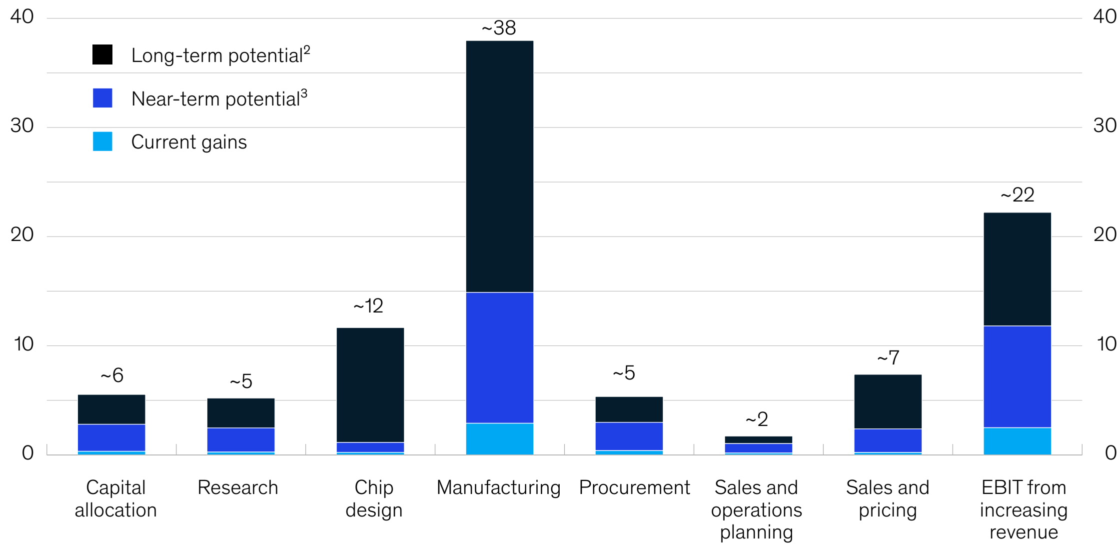
[1]Overall equipment effectiveness.
[2]Run-to-run.
Source: IBS; McKinsey analysis

[1] Industry revenues are for IDM, fabless, pure-play foundries, and SATS.

Exhibit 4

Artificial intelligence will deliver the most value by reducing manufacturing costs, but the largest relative impact will be in R&D.

Impact on earnings EBIT,[1] by semiconductor key activities, $ billions



Resulting overall potential of artificial intelligence	**28–32%** reduction in research and design costs	**13–17%** reduction of cost of goods sold, including depreciation	**10–14%** reduction in sales, general, and administrative spending

[1]Earnings before interest and taxes.
[2]Long-term potential accessible on timescales longer than 4 years.
[3]Near-term potential accessible within next 2–3 years.
Source: IBS; McKinsey analysis

AI/ML use cases in manufacturing
Manufacturing is the semiconductor industry's largest cost driver, and AI/ML use cases will deliver the most value—about 40 percent of the total—here. They can reduce costs, improve yields, or increase a fab's throughput. Over the long term, we estimate that they will decrease manufacturing costs (both cost of goods sold and depreciation) by up to 17 percent. Consider a few examples.

Adjustment of tool parameters. When defining steps in process recipes, semiconductor companies typically specify one constant time frame for each one. But the time frame required for some individual wafers may show statistical or systematic fluctuations, so a process could keep running after it has produced the desired outcome (for instance, a particular etch depth). That can increase timelines and waste or even damage the chip.

To achieve greater accuracy, semiconductor companies can use live tool-sensor data, metrology readings, and tool-sensor readings from previous process steps, allowing machine-learning models to capture nonlinear relationships between process time and outcomes, such as etch depth. The data collected might include electric currents in the etching process, light intensity in lithography, and temperatures in baking. With these models, optimal process times can be implemented on a per-wafer or per-batch basis to shorten processing time, improve yield, or both, thus decreasing cost of goods sold (COGS) and increasing throughput.

Visual inspection of wafers. This step, which helps ensure quality by detecting defects early in the front-end and back-end production processes, is frequently conducted during production—for example, using cameras, microscopes, or scanning-electron microscopes. Those images are still commonly evaluated manually by operators for potential defects, however, leaving them subject to error and backlogs and driving up costs.

Modern wafer-inspection systems, made possible by advances in deep learning for computer vision, can be trained to detect and classify defects on wafers automatically, with an accuracy on par with or better than human inspectors. Specialized hardware, such as tensor-processing units, and cloud offerings enable automated training of computer-vision algorithms. This, in turn, allows for faster piloting, real-time inference, and scalable deployment.

With this approach, companies can obtain early insights on potential process or tool deviations, allowing them to detect problems earlier and improve yields, all while reducing costs.

AI/ML use cases in research and chip design

AI/ML use cases can help semiconductor companies optimize their portfolios and improve efficiency during the research and chip-design phase. By eliminating defects and out-of-tolerance process steps, companies can avoid time-consuming iterations, accelerate yield ramp-up, and decrease the costs required to maintain yield. They may also automate the time-consuming processes related to physical-layout design and the verification process.

Although we are not yet at the point where AI/ML acceleration can be applied to all designs and to all stages of chip design, we do not see a fundamental reason why it cannot penetrate further over time. Therefore, AI/ML may eventually reduce the current R&D cost base by as much as 28 to 32 percent, which is even higher than the gains expected from manufacturing.

Automated yield learning in integrated circuit design. If there are missteps during integrated circuit (IC) design, semiconductor companies have to undertake multiple costly and complicated iterations based on feedback from manufacturing.

Semiconductor companies may avoid this problem by deploying ML algorithms to identify patterns in component failures, predict likely failures in new designs, and propose optimal layouts to improve yield. During the process, IC designs are broken down into key components with the support of AI-based analytics. The algorithms then compare these component structures with existing designs to identify problematic locations within the layout of individual microchips and improve the design. Thus, AI- and ML-aided design can significantly reduce COGS, increase terminal yields, and decrease time to market for new products. It can also decrease the effort required to maintain the terminal yield.

Other areas. All other functions, including planning, procurement, sales, and pricing, will benefit from AI/ML use cases. Often, these use cases are not specific to the semiconductor industry and are partially established in other industries, thus allowing implementation to occur more rapidly. Overall, applying AI/ML use cases to additional functions could yield up to $20 billion in annual value.

Six critical enablers for successful AI/ML implementation at scale

To assist semiconductor companies with AI/ML transformations and deploy use cases at scale, we focus on six enablers that are part of the McKinsey playbook for digital and analytic transformations: the creation of a strategic road map, talent strategy, agile delivery, technology, data, and adoption and scaling (Exhibit 5).

Exhibit 5

Our research indicates that six enablers are critical for successful implementation of artificial intelligence at scale.

6 enablers of artificial-intelligence (AI) implementation

Strategic road map	Talent strategy	Agile delivery	Technology	Data	Adoption and scaling
• Make AI a top priority • Prioritize value and time to value • Don't reinvent the wheel	• Centralize AI capabilities for critical mass • Assemble various team capabilities/profiles • Add business expertise in central AI team • Add data-science expertise in local teams	• Conduct development sprints and encourage rapid learning	• Connect the fab • Leverage the cloud • Build out the edge in fabs	• Unlock your data and systems • Remember that not all data are relevant • Establish rigid data governance	• Design use cases for scale from the beginning • Demand use of best-known methods across sites • Ensure smooth integration of use cases into digital workflows

Creation of a strategic road map

Above all, scaling AI/ML efforts must be a strategic priority for companies. The initial effort, which involves coordinating data, agreeing on priority use cases, and encouraging collaboration among the right business, data science, and engineering talent, is too great to be successful as a bottom-up project.

Ideally, the AI/ML effort will be linked to clear business targets, giving business units and business functions a joint interest in making the transformation successful. For example, companies could identify cost savings for predictive maintenance and provide resources for the appropriate AI/ML use case. The resulting savings would help the function that sponsored the use case and provided the appropriate resources, allowing it to achieve its business targets. Such gains will give functions a strong incentive to support AI/ML implementation. Setting clear business targets will also help companies measure the benefits of each use case over time.

In line with their defined targets, companies should identify specific business domains and value levers that will be their focus. They can then select relevant use cases that allow them to apply these levers.

When prioritizing use cases in the strategic road map, companies should emphasize their total value, feasibility, and time to value. As their experience and capabilities grow, they can undertake additional use cases that are more difficult to implement or take longer to achieve. As they determine the value of potential use cases, companies should examine levers that often get overlooked, such as the competitive advantages associated with decreased time to market and higher quality. Such details will allow them to size and prioritize initiatives accurately.

After setting their priorities, semiconductor companies must allocate sufficient resources to their AI/ML initiatives and investigate supportive partnerships with third parties that have complementary skills, rather than trying to reinvent the wheel themselves. Some larger players may have the spending power required to develop most capabilities in-house, as well as sufficient data from their large installed tool fleet to train AI/ML models, allowing them to retain full control over all associated intellectual property. Given the required resources, smaller players might find it beneficial to leverage commercially available solutions where available, or to partner with others to develop or share algorithms, or to create joint data-sharing platforms that increase the amount of information available to train models. Examples of potential partners include other semiconductor-device makers, companies involved in electronic design automation, hyperscale cloud providers, or equipment OEMs.

Talent strategy

Most companies that successfully implement AI/ML create a centralized organization, such as a center of excellence (COE), that focuses on such initiatives. This group serves as a clear home for the new talent required and is responsible for defining common standards and building a central repository for best practices and knowledge. Some of the leading semiconductor companies have already made significant investments in AI/ML COEs that include hundreds of engineers.

When hiring technology staff for the central team, semiconductor companies should carefully balance the role composition to ensure that it has the right capabilities to move from pilot to full scale-up of a use case. For example, data scientists and data engineers are required for piloting an AI/ML use case, but ML engineers, infrastructure architects, or full-stack developers are needed to drive the scale-up. Typically, semiconductor companies do not have employees with these profiles and must recruit them externally.

The centralized AI/ML function cannot be isolated from the business and functions in which it will deploy use cases. To build connections, people with business/operations domain expertise, such as R&D designers, process engineers, and equipment engineers, should be included in the AI/ML function. These team members have a critical role in identifying AI/ML use cases and also act as ambassadors for AI/ML solutions within the organization.

Likewise, successful companies will ensure that local sites—fabs, functions, or both—add data-science expertise to their AI/ML teams. The employees trained to become “data citizens” can work jointly with specialist roles from the AI/ML COE to lead use-case selection and support implementation in cross-functional teams.

Agile delivery

To avoid a situation where AI/ML use cases become stuck in a “proof-of-concept” spiral with limited use or scale, teams should focus on achieving business value, with a heavy emphasis on iterative improvement.

An agile approach, which is central to software development, can help semiconductor companies attain this focus. Although AI/ML development involves intense discovery and exploration, semiconductor companies should receive continuous feedback from people who use insights from their models. Many agile teams have found success by leveraging the vertical-sliver approach, which involves creating an end-to-end analytic pipeline that includes data ingestion, modeling, recommendation development, and deployment to users—typically business owners or engineers who work on the fab floor—in the first or second sprint. The vertical-sliver approach may be counter to many established practices since semiconductor companies typically only make changes within manufacturing engineering when they are completely certain that the shift will deliver perfect results.

From an operational perspective, agile teams are beneficial because they reduce dependencies on people outside the team. Typically, it is difficult to avoid such dependencies since there are often organizational divisions among data owners, AI/ML experts, and IT infrastructure. But agile AI/ML teams are cross-functional and encompass all needed expertise for the use case even if some members are only included for a limited number of sprints. Agile teams can also leverage self-serve resources, such as access to data and infrastructure.

The shift to agile AI/ML delivery should occur as soon as possible and will be more likely to gain traction if top leaders lend their support and companies attempt to change mindsets as well as processes.

Technology

Within the fabs, successful companies establish a connectivity layer for real-time access to relevant data sources, including production and measurement tools, auxiliaries, facilities, and others. Tool OEMs can help ensure this connectivity, which is particularly essential for manufacturing use cases. We will explore the role of tool makers in enabling AI/ML in a second article.

Semiconductor companies also require a common data-integration layer. This layer first combines the data before deploying the analytics engines and use cases in a development environment. For best results, semiconductor companies must find ways to combine data and use cases from different tool vendors to limit complexity and prevent multiple Internet of Things stacks in parallel silos.

Successful companies will leverage both edge and cloud computing to support their AI/ML use cases. Since some tools generate tremendous amounts of data, edge-computing capabilities—deploying the AI/ML use case within or close to the tools—are often required for real-time applications. Cloud solutions provide economies of scale and enable links among different fabs, increasing the pool of training data for use cases. (Semiconductor companies are historically cautious around data security, however, so they may limit deployment of sensitive data to on-premises solutions.)

Data

Semiconductor companies have several hundred tools in each fab, some of which generate terabytes of data, and it would be impossible to examine every piece of information. To ensure maximum effectiveness and efficiency, players must prioritize data that might enable multiple use cases since this will have a much greater impact than a single initiative.

Even if players limit the amount of information analyzed, their AI/ML initiatives will still require extensive time and resources, such as sufficient numbers of data engineers on AI/ML teams. Strict data-governance policies are required to ensure that existing data and newly generated data are immediately ready for use, consistently high in quality, and trustworthy. Successful companies typically have a dedicated data-governance team to ensure data consistency as well as the quality of new and existing data.

Adoption and scaling

Semiconductor companies should stringently focus on the scalability of prioritized use cases, beginning in the design phase. Experts from multiple sites or fabs must be included early on to ensure that use cases can later be deployed across locations. Some semiconductor companies are creating focus groups within the fab landscape to plan for scale-up. For specific domains, they pick a fab to serve as the lead site, and it then identifies use cases, collects requirements from the other fabs, creates the implementation plan, and ensures transfer of knowledge. As noted earlier, semiconductor companies will need to prioritize use cases for deployment based on their value after full scale-up.

Second, semiconductor companies should ensure that the entire organization follows standards and best-known methods (BKMs) when developing and scaling up use cases. Codifying and enforcing the use of BKMs across the organization can ensure that solutions are sustained and improved over time, allowing machine learning to gain maximum scale across sites. Typically, the central AI/ML team oversees this critical task.

Finally, semiconductor companies must seamlessly integrate use cases into an end user's digitized workflows to ensure adoption. Many companies overlook this step, but this oversight has major consequences. In our survey, nearly half of semiconductor-device makers stated that lack of integration was the second-biggest problem in scaling AI/ML use cases. If organizations form tight links between the AI/ML function and the business side, it will be significantly easier to take the user's perspective when initially designing the use case.

The semiconductor industry is at a turning point, and companies that don't devote significant resources to AI/ML strategies could be left behind. Although semiconductor companies may take different approaches, depending on business model, experience with AI/ML, and strategic priorities, the goal is the same: to take productivity and innovation to new levels.

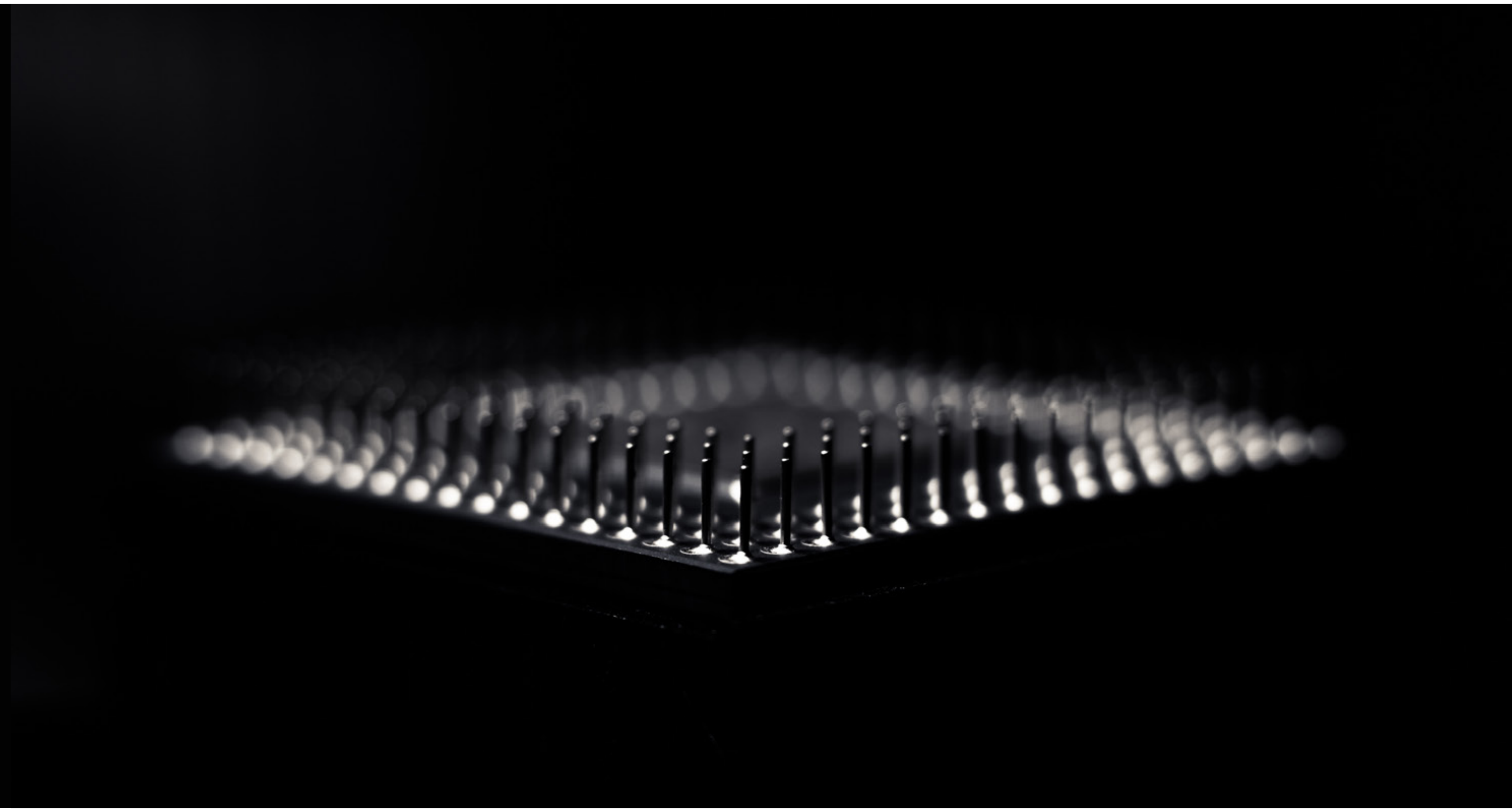
Sebastian Göke is a consultant in McKinsey's Berlin office; **Kevin Staight,** an analytics expert at QuantumBlack, a McKinsey company, is based in the Boston office; and **Rutger Vrijen** is a partner in the Silicon Valley office.

The authors wish to thank Sven Borghardt, Stefan Burghardt, Eike Ebel, Julia Kompalla, Maximilian Schulz, and Filippo Scopel for their contributions to this article.

Coping with the auto-semiconductor shortage: Strategies for success

Just as cars and trucks go digital, a scarcity of semiconductors is causing billions of dollars in lost revenue for the automotive industry. Here's why it's happening and how to move forward.

by Ondrej Burkacky, Stephanie Lingemann, and Klaus Pototzky



The automotive industry is running out of chips. The global semiconductor shortage that began in the first quarter of 2021 has halted assembly lines around the world, as the long lead time for the tiny silicon chips has slowed production of everything from smartphones and home appliances to driver-assistance systems. Major car makers, including a US-based OEM, have already announced significant rollbacks in their production, lowering expected revenue for 2021 by billions of dollars.

That challenge in the auto industry is the latest in a series of them that began in the early months of the COVID-19 pandemic, when auto sales plummeted as much as 80 percent in Europe, 70 percent in China, and nearly 50 percent in the United States. The lack of demand for new cars shuttered auto factories and sent home millions of workers, while orders for semiconductors—used in myriad ways, including in fuel-pressure sensors, digital speedometers, and navigation displays—dropped off precipitously.

The effects of the semiconductor shortage have extended beyond the auto sector, with other industrial players struggling to secure chips. This highlights the fragility of those supply chains, which largely rely on Asia as a hub of semiconductor manufacturing. Many automakers are now operating in crisis mode, and few expect a rapid resolution. Auto manufacturers and chip makers alike will need to work together to tackle the imbalance in demand.

This article addresses both how the shortage happened and what remedies for it exist.

How the shortage happened

No single incident or disruption caused the semiconductor shortage. Instead, a confluence of events contributed to the situation the auto industry now faces.

Struggles during the COVID-19 crisis

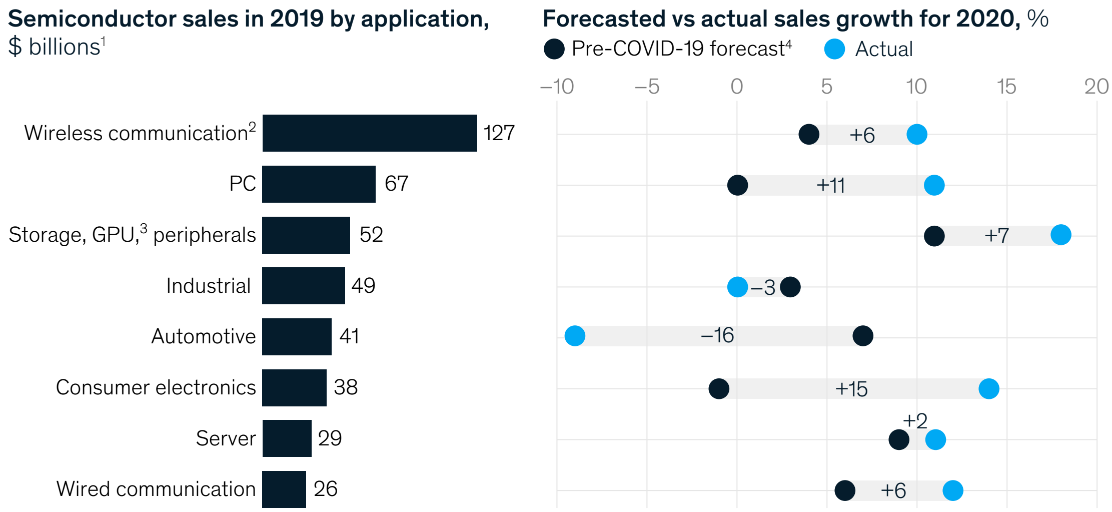
In the first half of 2020, the auto industry faced a substantial drop in demand. Moreover, while new-vehicle sales grew in the second half of the year, the highly ambiguous sales outlook at the time meant that automakers didn't meaningfully increase their semiconductor orders. At the same time, driven by the shift to remote work and the associated greater need for connectivity, consumer demand significantly rose for personal computers, servers, and equipment for wired communications, all of which heavily depend on semiconductors. That meant that even as the auto industry drastically cut chip orders, other sectors faced an increased need.

Our analysis of IHS Markit data reveals that the actual demand for semiconductors in the auto industry in 2020 trailed a prepandemic estimate by around 15 percentage points (Exhibit 1). Over the same period, most other segments (with the exception of the industrial sector) experienced rapid

The effects of the semiconductor shortage have extended beyond the auto sector, with other industrial players struggling to secure chips.

Exhibit 1

Automotive semiconductor sales lagged in 2020, but growth in most other segments is expected to exceed pre-COVID-19 estimates.



[1]Products include memory, microcomponents, logic, analog, discrete, optoelectronic, and sensors/actuators.
[2]Includes Chinese inventory effect; growth rate without inventory expected to be –4 to –8%.
[3]Graphics processing unit.
[4]As of December 2019. The estimates for 2020 were calculated using a 2019 baseline, and percentages have been rounded.
Source: IHS Markit; Strategy Analytics; expert interviews

expansion, resulting in an average increase from 5 to 9 percent in semiconductor sales beyond the forecast growth. Because of that, when the auto sector's demand recovered faster than anticipated in the second half of 2020, the semiconductor industry had already shifted production to meet demand for other applications.

Lack of new capacity

The semiconductor industry has matured in recent years through consolidation and the achievement of greater scale. Its capacity has expanded modestly but steadily—by around 4 percent annually, in line with sales (Exhibit 2). In parallel, semiconductor utilization has been consistently high (at or above 80 percent) in the past decade. In fact, utilization in 2020 was close to 90 percent, which many industry leaders regard as full utilization, since exceeding that level often results in disproportionately longer lead times. Therefore, while the semiconductor industry has increased its production capacity by nearly 180 percent since 2000, its total capacity is nearly exhausted at the current high utilization rate.

Geopolitical tensions

Because of geopolitical tensions, some consumer-electronics makers have considerably increased their chip-inventory levels to get through a period of limited access to semiconductor manufacturing. We estimate that such stockpiling caused a surge in semiconductor demand of 5 to 10 percent in the wireless space—the equivalent of one-third of auto-market chip sales.

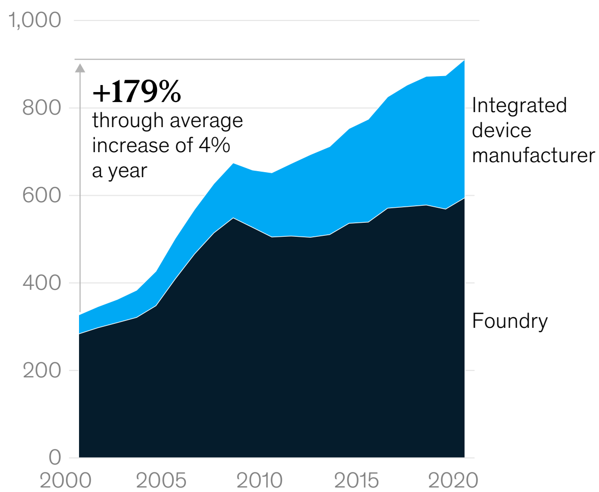
Contract terms

The typical contracts for sourcing parts in the auto industry differ significantly from other industries, which are more often governed by long-term binding agreements (so-called take-or-pay deals) and provide semiconductor suppliers with purchase orders that go well beyond six to 12 months. Amid an auto supply chain that is

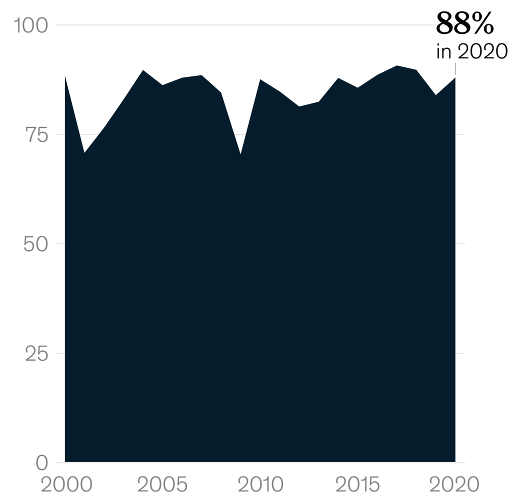
Exhibit 2

Over the past two decades, semiconductor capacity increased by an average of 4 percent a year, while utilization has remained high.

Semiconductor production capacity, million square inches per month



Semiconductor production capacity utilization, %



Source: IHS-Omdia; McKinsey analysis

complex and often heavily outsourced, the chip-sourcing commitment cycle for the auto industry, however, tends to be shorter term—especially with respect to binding purchase commitments on the order of a few weeks to a few months. While the auto industry has had a good reputation for stable demand in the past, semiconductor manufacturers are now committed to more conventional, longer-term contracts from other fast-acting industries.

Limited stock

Just-in-time manufacturing practices, which can minimize waste and increase efficiency by keeping on-hand inventory low, are widely leveraged in the auto supply chain. In normal times, the reduction of inventory is financially beneficial; however, in the event of an unexpected shortage, the practice causes immediate disruption of the entire supply chain. Since many players didn't expect the chip shortage in 2020 and 2021, they likely had very limited stock available to weather the crisis.

5G rollout and overlapping chip demand

Industry demand for semiconductors varies by node size. Chips in the smaller size ranges, the most advanced of which are seven and 14 nanometers or smaller, are often used in leading-edge technology applications but aren't required by many automakers. Our analysis reveals several knock-on effects of large-scale technology adoption that the auto industry must consider. For example, an expansive 5G rollout requires a large number of radio-frequency semiconductors manufactured at the same, larger node sizes as auto chips. The same is true for power-electronic chips needed to boot up servers and PCs (Exhibit 3). That amount of overlap means that as the rollout of 5G occurs over the next few years, automakers might see a continuing scarcity of chips.

Exhibit 3

A high amount of overlap exists between chips used for current technologies and those used for the auto industry, particularly for larger node size.

Overlap between trend and automotive nodes

	Electrification	5G	IoT edge computing
Leading edge Less than or equal to 28 nanometers (nm)	**Not applicable**	**Low overlap:** • Logic • Field programmable gate array • Application-specific integrated circuit	**Low overlap:** • Main processing unit • Memory
Trailing edge Greater than 28 nm	**High overlap:** • Discretes • Power management • Power supply units	**Medium overlap:** • Radio-frequency switches • Duplexer • Antenna	**Medium overlap:** • Sensor • Microcontroller • Analog (communication)

Source: McKinsey analysis

Prospects for recovery

The global semiconductor shortage isn't likely to resolve in the short term, because of factors such as the complexity of semiconductor manufacturing and the increasingly sophisticated chips needed in auto design. Because of that, we offer OEMs some shorter- and longer-term strategies to consider as businesses deal with the imbalance in semiconductor supply and demand.

Short-term strategies

In the short term, we don't see any indication that the current supply and demand imbalance for semiconductors will resolve. That's because typical lead times for semiconductor production can exceed four months for the products that are already well established in a manufacturing line (Exhibit 4). Increasing capacity by moving a product to another manufacturing site usually adds another six months (even in existing plants). Switching to a different manufacturer (for example, changing foundries) typically adds another year or more because the chip's design requires alterations to match the specific manufacturing processes of the new manufacturing partner. Additionally, chips can contain manufacturer-specific intellectual property that may require alternations or licensing. Also, alternative suppliers in the auto industry must go through a lengthy and complex qualification process.

Our analysis suggests that chip capacity won't catch up with demand in the short term for the auto industry. That is primarily because of the continued increases in volume and sophistication levels of the chips needed to power new technologies, such as advanced driver-assistance systems and autonomous driving.

Leading companies have taken a variety of measures to deal with the current situation. Many have established dedicated war rooms that combine their supply and demand intelligence to create greater transparency. For instance, automatically generated dashboards combine data from multiple sources on many segments, such as a company's supply chain and a semiconductor player's commitments. The use of analytics to match supply with demand helps reduce a cumbersome and error-prone manual effort.

Exhibit 4

Lead times for semiconductor production can exceed four months, while switching to a new manufacturer takes a year or longer.

Semiconductor development and production timelines

	Capacity buildup		Product development		Production
	New fab buildup	**Fab ramp**	**Chip design**[1]	**Yield and volume ramp-up**	**Production (cycle time)**
Approximate typical duration, months	12–18	6–18	12–36[2]	6 or more	4 or more
Influencing factors	•Clean-room building •Facilitation (HVAC, gas, electric, etc)	•Tool lead times •Hookup and qualification	•Product complexity	•Product complexity •Fab utilization	•Product complexity •Physical and chemical processes •Fab utilization
Other considerations		Required for expansion of existing plant		Required when transferring existing product between fabs	

Fab-ramp and chip-design stages may run in parallel

[1]Chip design can be driven independent of fab manufacturing capacity.
[2]Eg, ~12-month product lifecycle for mobile phones; 24–36-month development time for automotive microcontroller units.
Source: McKinsey analysis

The goal is to provide clear input for internal communication and for communication to suppliers and customers. Companies usually view that as a no-regrets move.

Beyond that, many automakers and tier-one suppliers continue to collect and analyze more sophisticated intelligence on the semiconductor value chain and chip-manufacturing locations. To make more informed decisions, company leaders continually reassess the competitive landscape by weighing the technological applications for prioritization at the level of the individual chips required. Several tier-one and semiconductor players have complained about the lack of transparency regarding real demand levels (driven partially by the recent crisis-mode practice of overordering to secure a basic level of supply) and prioritization among individual components.

In our experience, a joint discussion involving an OEM, its tier-one suppliers, and semiconductor suppliers can help align the goals of all participants. Offering extra payments to expedite the production of wafers when capacity amounts to less than 5 percent of the production volume can also help. Other options involve replacing back-ordered components with similar but more feature-rich units (for example, swapping in chips with more memory) and using consumer-grade chip sets that receive additional quality tests.

Solving the long-term problem

In the longer term, the auto industry will need to rethink the way it structures contracts for semiconductor-related sourcing. As a good place to start, OEMs and tier-one players could make up-front volume commitments more binding (for instance, by moving to 12 months on a production- or technology-corridor[1] level and six months on a chip-set level). A more balanced risk-sharing plan aligned along the value chain could also help drive adoption rates.

In addition, companies might have to reconsider, at least in part, the current practice of just-in-time delivery and low stock levels along their value chains. There is also a need to align with the current push by various governments for more regional sourcing, since many government leaders are concerned about the fragility of supply chains and the prospect of depending on single suppliers and distant countries for vital needs. A McKinsey Global Institute analysis found that supply-chain shocks affecting global production occur just under every four years, on average, with companies losing 42 percent of one year's earnings every ten years.

In addition, auto players can revise their strategies for sourcing various chips. While a sourcing decision for a specific type of chip might appear on paper as less expensive than other options, the assessment might change when factoring in the cost of complexity in areas such as sourcing resilience and the software life cycle. In the qualification process, companies might need to reconsider some parameter constraints (for example, temperature range) to create the right balance of broader sourcing opportunities and product reliability.

Companies could also consider making selective investments in supply-chain resilience, with a clear-eyed view of their dependency on selected components and supply uncertainties. Such investments could range from spending on dual-source manufacturing qualification jointly with semiconductor suppliers to adjusting pricing levels with supply guarantees to bundling volumes to achieve greater negotiation power.

The current chip shortage is disrupting auto businesses across the value chain as OEMs and their suppliers rush to procure reliable chip sources. As auto players ponder their next moves and semiconductor manufacturers struggle to keep up with demand, both industries need to align their short- and long-term strategies to weather the supply-chain disruption as successfully as possible.

[1] "Technology corridor" refers to the use of a specific technology common for multiple chips.

Ondrej Burkacky is a partner in McKinsey's Munich office, where **Stephanie Lingemann** is an associate partner and **Klaus Pototzky** is a consultant.

McKinsey & Company

GSA
Where Leaders Meet

Automotive semiconductors for the autonomous age

The rise of autonomous vehicles is shifting demand for automotive chips and prompting OEMs to consider in-house design. How could this reconfigure value chains across industries?

by Ondrej Burkacky, Johannes Deichmann, Larissa Rott, and Alexander von Falkenhausen

Established OEMs and start-ups have invested $106 billion in autonomous-driving capabilities since 2010. Much of this funding has gone toward enhancing advanced driver-assistance systems (ADAS), which handle braking, object detection, and other critical vehicle functions. Many OEMs are also envisioning a day when fully self-driving cars will move from pilots and become mainstream, although development timelines are constantly changing and approval dates are uncertain.

The strong focus on autonomous vehicles (AVs) has already altered demand patterns for automotive semiconductors, with sales of specialty silicon—chips tailored to specific applications—growing strongly. These customized chips are only available from a few semiconductor companies, and some OEMs are now designing them in-house to reduce development timelines and gain more control. With demand for specialty silicon continuing to grow, other OEMs could take the same route.

To navigate the changes ahead and become market leaders, both automotive and semiconductor companies must understand the impact of new and future AV technologies on chip demand. Tier-one suppliers must also reassess their capabilities and offerings to ensure that their products remain relevant. With autonomous chips—those used to enable AV functions—expected to generate $29 billion in revenue by 2030, the stakes are high.

A changing market for automotive semiconductors

While the COVID-19 pandemic sent consumer purchases into a free fall—car sales were down 47 percent in the United States and 80 percent in Europe in April 2020—some countries are rebounding strongly across sectors. With demand rising for everything from smartphones to appliances to new cars, semiconductor orders are up, and fabs are struggling to increase their output.

After recent economic gains in some countries, the automotive sector is once again an important source of revenue for semiconductor companies. Most growth in this area stems from the shift to ADAS, since these systems must instantly process data when responding to unexpected changes, such as a sudden stop in traffic. Such capabilities require multiple interconnections within the vehicle and high-performance chips. In consequence, they have a more centralized electrical and electronic (E/E) architecture compared with traditional vehicles, as well as more sensors and compute-electronics content.

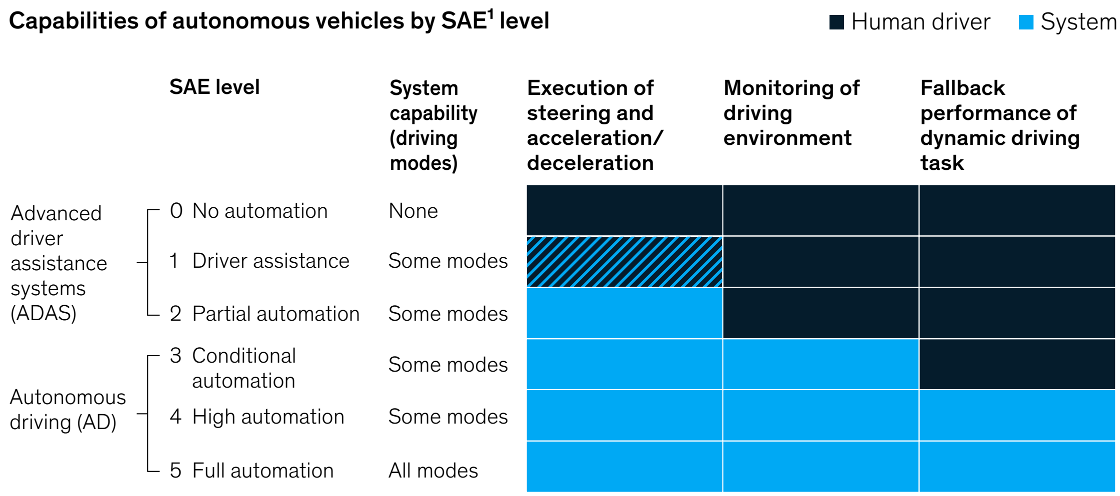
Not all AVs are alike, however, and the number and type of semiconductors required will depend largely on their level of automation. One common AV-classification system, which was defined by the Society of Automotive Engineers (SAE), divides AVs into six categories ranging from 0 (no automation) to 5 (full automation with self-driving capabilities) (Exhibit 1). Note that level 2 is divided into two parts: entry level (vehicles with some autonomous features, such as braking, but still require drivers to keep their hands on the wheel at all times) and advanced level (vehicles with more extensive autonomous features that allow drivers to take their hands from the wheel at times).

Vehicles in levels 0 through 2 (entry) achieve adequate performance with standard chips, but those in levels 2 (advanced) through 5 are expected to require a growing share of specialty silicon. Such chips are more efficient, enable rapid performance increases within vehicle systems, and allow the execution of complex software functionalities and analytics, such as those that enable sensor fusion of cameras, laser, LiDAR, and other devices. But many OEMs now have difficulty obtaining the silicon that exactly matches their needs, which interferes with their ambitious AV-development programs. What's more, there are few solutions available that fit to the targeted software stack of OEMs.

Faced with these constraints, some OEMs have already begun designing chips in-house. This route comes with several benefits, including optimized performance for specific algorithms and shorter development timelines for continuous

Exhibit 1

Autonomous vehicles have been divided into levels based on their capabilities.



[1] Society of Automotive Engineers.
Source: Society of Automotive Engineers

feature improvements. In-house design also allows OEMs to define the software stack and gives them greater control over chip design so they can create customized solutions that could differentiate their AVs (for instance, by allowing earlier time to market or providing greater availability of features). Some tier-one suppliers are also claiming ownership of certain vehicle systems, such as software stacks. With such divisions, the value chain is becoming increasingly disaggregated, and the division of expertise means that companies increasingly look for partners that complement their skills and products.

Given these developments, players in the industry commonly ask several questions: What model will most OEMs favor when it comes to chip procurement? How do we need to evolve to remain strong in the changing landscape? And what is our future role in the value chain and technology stack?

Navigating the changing landscape for automotive semiconductors

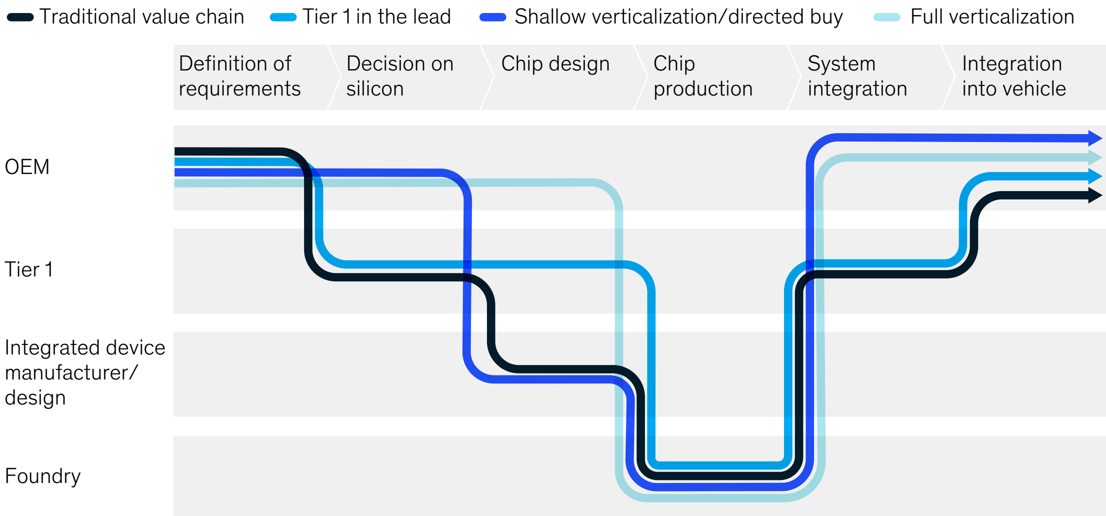
Analysis suggests that most OEMs will fit into one of four models when it comes to procuring semiconductors (Exhibit 2). In the first model, which is most common today, OEMs work directly with tier-one suppliers. In addition to defining chip requirements, the suppliers contract integrated device manufacturers (IDMs) to design or select the chips required. After the chips are produced by the IDM or a foundry, the supplier builds them into the system. In the second model, which is a variation of the first model, a tier-one supplier takes the lead for both defining requirements and designing the chips. Since IDM services are not required in the second model, the tier-one supplier would directly approach a foundry for production.

The third model, which is emerging as OEMs become more involved with design, involves what could be termed shallow verticalization. Under this

Exhibit 2

OEMs will likely follow one of four models to acquire or produce semiconductors.

Paths for acquiring and producing semiconductors



model, OEMs define the chip requirements and directly approach IDMs and design services. They then commission production at a foundry. Some OEMs like this model because it requires limited in-house talent and shifts the burden for ensuring quality to the IDM or manufacturer. On the downside, the shallow-verticalization model increases material costs and gives OEMs little opportunity to incorporate their own specifications and create specialty chips. The lack of customization may make it difficult to differentiate their products from the pack, especially if competitors use specialty chips that deliver better computational efficiency.

A McKinsey survey of more than 100 leading automotive and semiconductor specialists showed that 68 percent of respondents believed that OEMs would favor the shallow-verticalization approach to navigating the value chain. This path makes sense for many companies, since use-case requirements are similar across OEMs and chip-development costs are higher when smaller volumes are produced.

The last model, full verticalization, is just emerging. This model gives OEMs the most independence, since they define requirements and oversee chip design before commissioning production at a foundry. There are two potential strategies for full verticalization:

— *The independent route,* in which OEMs develop their own silicon, architecture, and chip design—all while keeping material costs relatively low. By working independently, OEMs can create specialized chips that differentiate their products from competitors. Of course, independence comes with greater risks, since a single OEM bears all costs and has complete responsibility for meeting timelines and ensuring quality. OEMs may also encounter some delays, at least initially, since many lack employees with strong expertise in chip design and architecture.

— *The formation of a cross-OEM consortium,* in which multiple companies collaborate on chip design and architecture. While individual companies might be able to incorporate some of their own specifications into the chips, compromises are inevitable and there are limited opportunities for differentiation. In some cases, it may be difficult to reach an agreement about joint requirements and priorities or to align on processes. On the plus side, OEMs share both development costs and risks, reducing the burden on individual companies. Consortium members can also pool their staff, which may reduce the competition for top talent.

Opportunities still abound for semiconductor companies

The increased involvement of OEMs in chip design could cut into profits at semiconductor companies. That said, the growth of AVs could significantly increase the market for automotive chips and could help compensate for much lost business. Analysis suggests that revenues for autonomous chips—one important subcategory—are expected to rise to about $29 billion per year by 2030, representing about $350 per vehicle (Exhibit 3). That's up from $11 billion in 2019.

Most demand is expected to come from vehicles with autonomy levels of 2 or higher. In 2019, these vehicles only accounted for about 40 percent of revenues for automotive chips. By 2030, they will account for 85 percent of demand.

Shifting chip demand

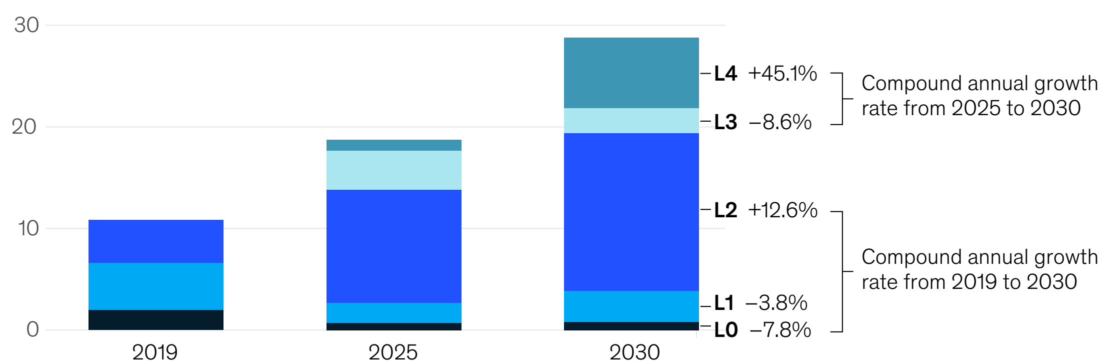
High-performance central-compute chips, such as domain control units (DCUs) and sensors, could see the fastest growth. Revenues from these products are expected to rise by about 12 percent annually from 2025 through 2030, when they will account for almost one-third of revenues. Chips for decentral electronic control units (ECUs) and sensors are only expected to see annual growth of about 6 percent from 2019 through 2030.

Looking more deeply at high-performance chips, analysis revealed that ECUs and DCUs are likely to account for a greater proportion of autonomous-chip revenues than sensors, with their share expected to increase from about 55 percent in 2019 to around 70 percent in 2030, driven primarily by the trend of increasing central calculation of autonomous-driving (AD) operations without significant preprocessing in

Exhibit 3

The autonomous-chip market is expected to nearly triple by 2030.

Autonomous semiconductor market, $ billions



Source: IHS Markit; McKinsey Center for Future Mobility the sensors. The most dramatic changes may occur between 2025 and 2030, when ECU and DCU values could nearly double because of faster technology adoption and reduced R&D costs. Over the same period, the value derived from sensors is likely to remain relatively flat because price decreases, such as those for LiDAR, will offset any increases in sales.

Regional variations

An examination of regional differences suggested that the Chinese market may be poised for strong growth. In 2019, China accounted for under 30 percent of global autonomous semiconductor revenues—about $3 billion. With an expected annual growth rate of about 12 percent, the Chinese semiconductor market is expected to see its share rise to nearly 40 percent by 2030. The main driver of this strong growth is high ADAS and AD adoption in China, supported by a receptive public, high demand, and a strong regulatory push. This increase, combined with overall market growth, could bring annual revenues in China to about $11 billion. Markets in the rest of the world (ROW) are expected to see more moderate revenue growth of about 8 percent annually through 2030.

The outlook for semiconductor manufacturers

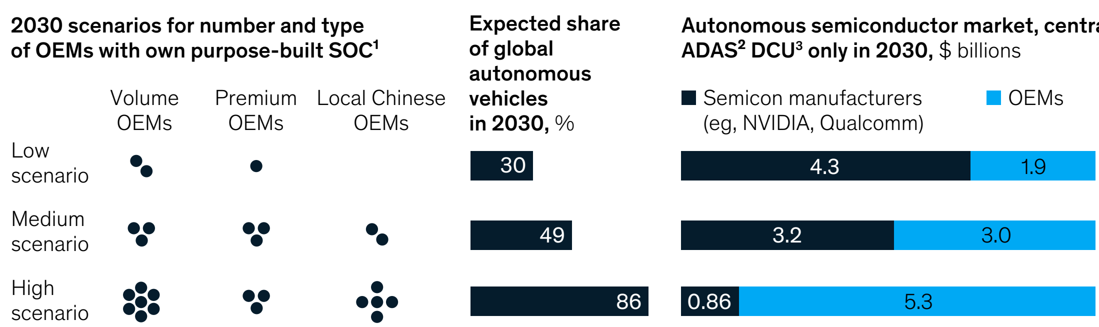
Semiconductor companies will need to understand how the possible shifts could affect their market share. It is difficult to quantify the exact impact, however, since it will largely depend on the extent to which OEMs begin to design their own chips. The shift to in-house design of central DCUs could have the greatest impact, and some major OEMs have already announced their plans to do so.

With so much uncertainty ahead, it may be helpful to consider three scenarios about the shift toward in-house design at OEMs. Out of the about $6 billion market for central DCUs in 2030, OEMs would account for approximately one-third of the value, or about $2 billion, if there is a low shift in this direction. That figure would rise to about 50 percent ($3 billion) in the medium scenario and 80 percent ($5 billion) in the high scenario (Exhibit 4). Even better, those shifts can lead to large revenue increases for the semiconductor companies if they have appropriate strategies to assess the changes and react accordingly. Those figures all represent lost revenue for semiconductor IDMs or fabless companies.

A more detailed breakdown of the Chinese market showed that volume OEMs, premium OEMs, and local Chinese OEMs could all potentially gain value, although the revenue impact varies by scenario. For instance, the low scenario assumes that two volume OEMs and one premium OEM will move to in-house product in 2030. The scenario does not assume that any local Chinese OEMs will make the shift, so we

Exhibit 4

Three scenarios can estimate the value of domain-control-unit chips produced by OEMs in 2030.



[1]System on a chip.
[2]Advanced driver-assistance system.
[3]Domain control unit.

did not allot them any value. The high and medium scenarios do assume that Chinese OEMs will make the shift, along with premium and volume OEMs.[1]

A strategic response

Semiconductor companies could benefit from new strategies. A few no-regret moves could help in all instances:

— *Taking a holistic approach to chip design.* Semiconductor companies could benefit from combining their software and semiconductor knowledge and talent. Consider a chip that enables vehicles to perceive people walking in the street. To build a chip with the right capabilities, it might be helpful to have software and analytics experts involved from the outset or performance may fall short.

— *Onboarding talent for chip design, E/E architectures, and software.* Semiconductor companies could benefit from aggressively seeking the best talent. To attract software specialists who are in great demand, they may need to adapt their company culture by adopting the agile principles common to software development. Semiconductor companies may also need to evaluate the skill sets of current employees, mainly around chip design and intellectual property (IP) block integration, and determine if any gaps exist.

— *Ensuring a thorough understanding of new requirements for specialized silicon.* OEMs that collaborate directly with semiconductor players on chip design could get important insights that are not available from tier-one suppliers. For semiconductor companies, this close collaboration could result in a deeper understanding of customer requirements and a better end product.

— *Define separate regional strategies.* E/E and software architectures, as well as the requirements for specialized silicon, may be different in certain regions, compared with the rest of the world, because of country-specific regulations, technology requirements, and customer preferences. For instance, more than 90 percent of respondents in a McKinsey survey of automotive and semiconductor executives expect that regional variations in the software stack will require different silicon in China. Most respondents—about 60 percent—also said that the best solution for serving multiple countries involved building one joint architecture that can be customized to fit market-specific requirements.

— *Building strong partnerships to ensure excellence throughout the value chain.* In addition to partnerships with OEMs, semiconductor companies could form strong relationships with other companies along the value chain, including specialty-software players, integrated-device manufacturers, and tier-one suppliers. Such collaboration with complementary silicon products or IP companies could further improve end products, since each player can contribute specialized knowledge and different expertise.

Strategic imperatives for automotive OEMs and tier-one suppliers

While the upcoming changes to the AD chip market might hit semiconductor companies first, companies in other industries would be wise to rethink their strategic position and invest in essential capabilities to understand and navigate the changing environment.

[1] In the medium scenario, we assume that an additional two global-premium OEMs and two Chinese OEMs will move to in-house chip design, on top of the two volume OEMs and one premium OEM from the low scenario. The high scenario assumes an additional five volume OEMs and two Chinese OEMs will shift design in-house, on top of the companies that made this move in the medium scenario.

Automotive OEMs face a multitude of challenges because of the rising complexity of the on-board software and electronics architecture required to fulfill autonomous driving, connected vehicles, electrification of the powertrain, and shared mobility (ACES) requirements. While most OEMs decided to ramp up their in-house capabilities for software development and integration, they are just beginning to create strategies for semiconductors. A few no-regret moves could assist their efforts:

— developing a strong understanding of end-to-end architecture, from software applications to semiconductors, to optimize systems

— ramping up recruitment and retention for semiconductor talent; regardless of their design decisions, they need staff that can understand and judge concepts

— ensuring that they focus their efforts if they plan to design semiconductors, which will help to avoid fragmentation of talent and allow them to double down where it really matters (for instance, concentrating on only one application area in ADAS)

— identifying strategic partners to help implement their semiconductor strategies and accelerate their efforts

Tier-one suppliers are also encountering many challenges as OEMs move down the technology stack by insourcing more parts of software development and undertaking semiconductor design. Simultaneously, semiconductor players are becoming more aggressive in linking proprietary platform software with their chips. To remain strong, tier-one suppliers could benefit from a clear semiconductor strategy, especially in the R&D-intensive areas of ADAS and AD. First, they could develop flexible offerings to avoid being locked out of the market. For instance, ADAS applications could be configured to to run on any chip. Tier-one suppliers could also consider investing in semiconductor capabilities to complement their electronics and systems-integration knowledge. Finally, tier-one suppliers could actively pursue partnerships with OEMs, semiconductor companies, and tech players within their ecosystem to avoid becoming the odd one out.

The pursuit of the best semiconductors within the automotive sector could become more severe as ADAS systems, AD, and vehicle connectivity advance. Semiconductor companies can gain insight into shifting chip demand and then develop the capabilities and products needed to remain strong. Likewise, OEMs and tier-one suppliers will benefit from taking a hard look at their capabilities and current strategies, especially in relation to semiconductors. In the new landscape, collaboration among semiconductor companies, automotive OEMs, and tier-one suppliers along the value chain may be critical for success. Those companies that monitor market trends and proactively take action may be in the best position to succeed.

This work is independent, reflects the views of the authors (McKinsey & Company and GSA), and has not been commissioned by any business, government, or other institution.

Ondrej Burkacky is a senior partner in McKinsey's Munich office, **Johannes Deichmann** is a partner in the Stuttgart office, and **Larissa Rott** and **Alexander von Falkenhausen** are consultants in the Berlin office.

Cybersecurity in automotive: Mastering the challenge

With the software content of cars increasing, what do automotive players need to know about cybersecurity?

This article is a collaborative effort by Ondrej Burkacky, Johannes Deichmann, Benjamin Klein, Klaus Pototzky, and Gundbert Scherf, representing views from the McKinsey Center for Future Mobility.

© vladru/Getty Images

The four ACES disruptions—autonomous driving, connected cars, electric vehicles, and shared mobility—have dominated the agenda of automotive industry leaders in recent years. These innovations, built on the digitization of in-car systems, the extension of car IT systems into the back end, and the propagation of software, turn modern cars into information clearinghouses while also making them tempting targets for cyberattacks.For this report, we worked with the Global Semiconductor Alliance to explore the consequences of this shift. We focused on providing a perspective on three key questions for the automotive industry:

— What are the specific trends and drivers of cybersecurity in the automotive industry, and why will this represent a paradigm shift for the industry?

— How are these drivers going to affect the automotive industry's long-established value chains?

— How can players inside and outside the industry prepare and position themselves for the upcoming market developments and anticipated segment growth?

Cybersecurity is becoming a new dimension of quality for automobiles

Over the last few years, the cyberrisk of connected cars has become clear with security researchers revealing various technical vulnerabilities. In these cases, the attackers disclosed their findings to OEMs to help them fix the issues before malicious attackers caused harm.

Currently, only narrow standards and guidelines exist for specific technical procedures for securing hardware and software in vehicles, such as standards for hardware encryption or secure communication among electronic control units (ECUs). That will soon change, however. The World Forum for Harmonization of Vehicle Regulations (WP.29), under the UN Economic Commission for Europe (UNECE) is planning to release new regulations on cybersecurity and over-the-air software updates. These will present cybersecurity as nonnegotiable for securing market access and type approval across UNECE WP.29 member countries (Exhibit 1).

While the UNECE WP.29 regulations on cybersecurity and software updates set a regulatory framework and minimum requirements for automotive players along the value chain, they do not include detailed implementation guidance for translating the requirements into concrete operational practices. However, the new International Standardization Organization (ISO)/Society of Automotive Engineers (SAE) 21434 standard, "Road vehicles – cybersecurity engineering," and ISO 24089, "Road vehicles – Software update engineering," lay out clear organizational, procedural, and technical requirements throughout the vehicle life cycle, from development to production to after-sales in terms of cybersecurity and software-updates.

These standards will allow the industry to implement common cybersecurity practices specific to vehicle development and manufacturing. They will also allow an assessment of adherence to the practices and attestation by third parties, which can be used between industry players to demonstrate adherence to the standards, for example, in contracts between OEMs and suppliers.

Securing hardware and software in modern vehicles will require new skills and talent

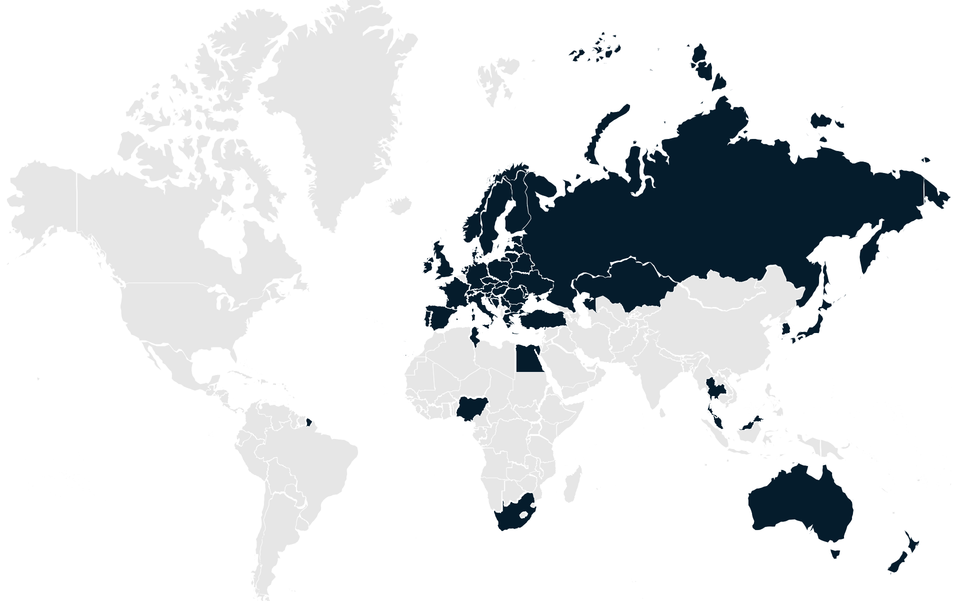
To secure hardware and software while meeting regulatory requirements and customer expectations, current automotive employees will need new skills and ways of working throughout the entire development cycle, including the phases involving specification, design, development, integration, and testing (Exhibit 2). Employees in other areas, such as procurement, project management, dealerships, and customer communications, will also need upskilling related to cybersecurity.

Exhibit 1

Cars in more than 60 countries will be affected under the new World Forum for Harmonization of Vehicle Regulations framework on cybersecurity and software updates.

World Forum for Harmonization of Vehicle Regulations (WP.29) under the UN Economic Commission for Europe (UNECE)

■ Countries party to the 1958 agreement[1] (as of Dec 2018)



[1] "Agreement concerning the Adoption of Harmonized Technical United Nations Regulations for Wheeled Vehicles, Equipment and Parts which can be Fitted and/or be used on Wheeled Vehicles and the Conditions for Reciprocal Recognition of Approvals Granted on the Basis of these United Nations Regulations" (original version adopted in Geneva on March 20, 1958).

Source: UNECE ECE/TRANS/WP.29/343/Rev.27 – Status of the Agreement, of the annexed Regulations and of the amendments thereto – Revision 27

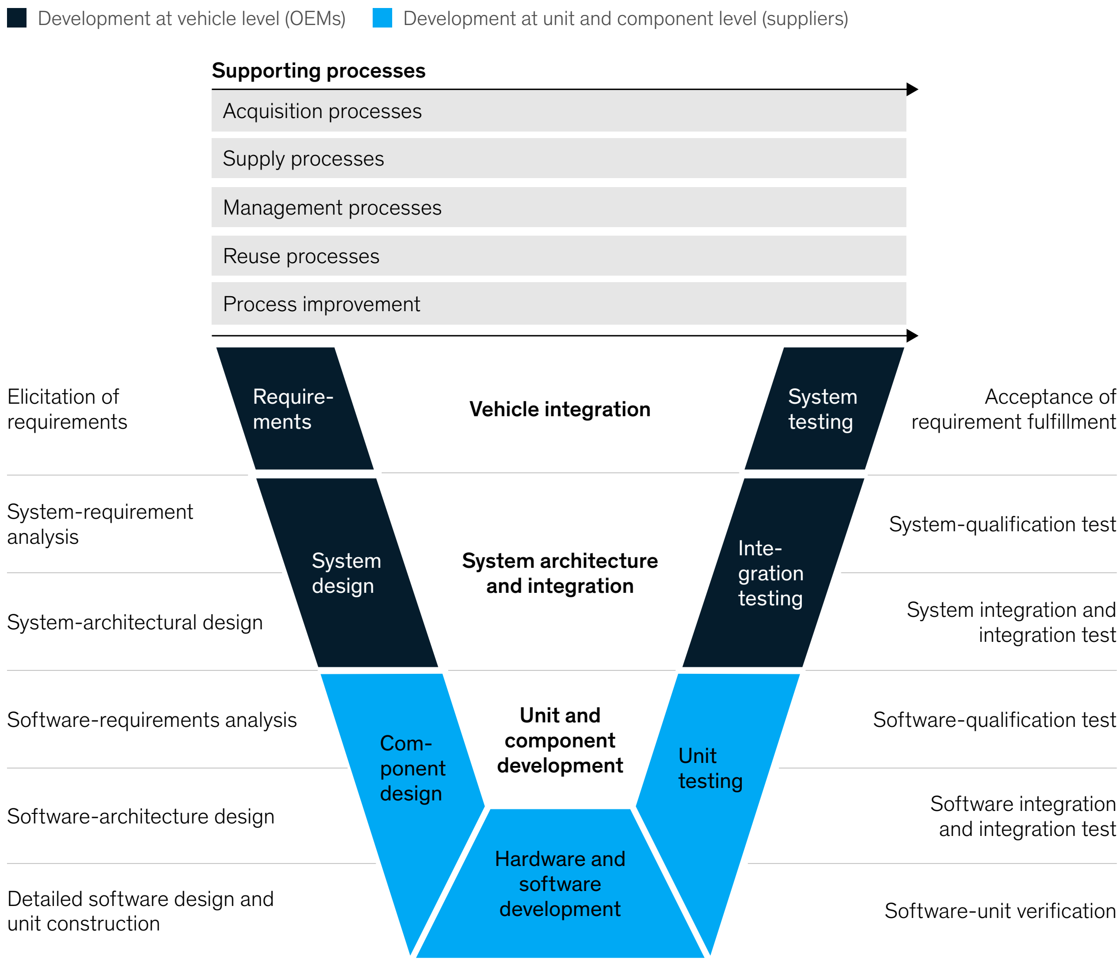
The boundaries and names shown on this map do not imply official endorsement or acceptance by McKinsey & Company.

In addition to upskilling employees, OEMs and other companies along the value chain must establish stricter cyberrisk-management processes and compliance documentation. The decision to modify systems or adopt new ones often depends on a company's organizational structure and maturity. Companies may also need to adjust roles, responsibilities, and formal processes for assessing and managing cyberrisks to vehicles.

In the new environment, OEMs will need to respond immediately to security incidents, including those in which companies discover a new or potential vulnerability, or in which their vehicles are attacked by malicious hackers. This will require organizational, procedural, and technical capabilities for detecting and addressing cybersecurity events. Providing security patches throughout the full vehicle life cycle will also be essential for safe vehicle operation. Vehicles are often driven for ten years or longer, requiring regular updates over a very long period. This makes them more akin to aircraft or vessels, which see software updates provided over longer periods than those for consumer products, such as PCs, smartphones, tablets, and smart appliances.

Exhibit 2

OEMs and suppliers will need to integrate cybersecurity measures throughout the development, which will require new talent and skills.



Source: Automotive Software Performance Improvement and Capability dEtermination (ASPICE) framework; McKinsey analysis

Automotive cybersecurity is expected to nearly double in the coming decade

We have broken down the automotive cybersecurity market into three elements: cybersecurity hardware, cybersecurity-related software-development efforts, and cybersecurity processes and solutions. Based on external expert interviews, McKinsey analyses, and predictive modeling, we estimate that the total cybersecurity market will increase from $4.9 billion in 2020 to $9.7 billion in 2030, corresponding to annual growth of more than 7 percent (Exhibit 3).

To capture value in this growing cybersecurity market, players along the value chain are following different strategies. We expect to see a significant amount of change in the following areas in particular:

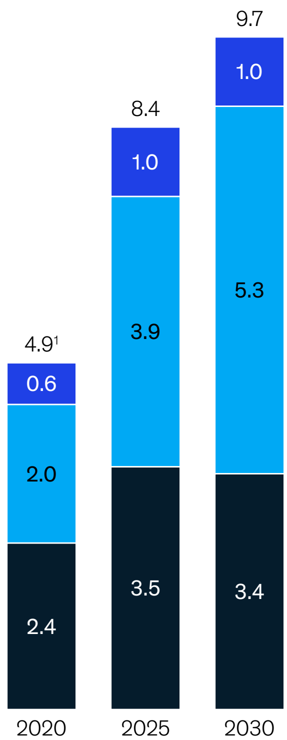
— OEMs are pursuing vertical integration (for instance, by building their own cybersecurity components or even software stacks).

— Suppliers are pushing their way up and down the value chain, such as by offering specialized cybersecurity-consulting services.

Exhibit 3

The cybersecurity market will grow significantly for automotive in the coming years.

Market size, $ billions



Submarket	**Compound annual growth rate, 2020–30,** %	
Total	7	
Cybersecurity hardware components	6	Dedicated security components for encryption and key storage (eHSM[2] and TPM[3])
Cybersecurity-related software-development efforts	10	Implementation of cybersecurity components (eg, encryption functionality) and requirements in functional domains Integration and testing of cybersecurity components and additional effort due to cybersecurity requirements in functional domains
Cybersecurity processes and solutions	4	Efforts to implement new regulatory requirements Software traceability (inventory and compatibility management, and impact assessment) Risk management and incident response Vehicle monitoring using security operations centers

[1]Figures may not sum to 100%, because of rounding.
[2]Embedded hardware security module.
[3]Trusted Platform Module.
Source: Analysis based on data from Ondrej Burkacky, Johannes Deichmann, and Jan Paul Stein, *Automotive software and electronics 2030: Mapping the sector's future landscape*, July 2019, McKinsey.com

- Start-ups are entering the market with innovative solutions, including specialized threat-detection applications or vehicle security operations centers (SOCs) as a service.

- IT and operational-technology (OT) companies are expanding into the adjacent automotive-cybersecurity market (for instance, by offering back-end solutions or cybersecurity components).

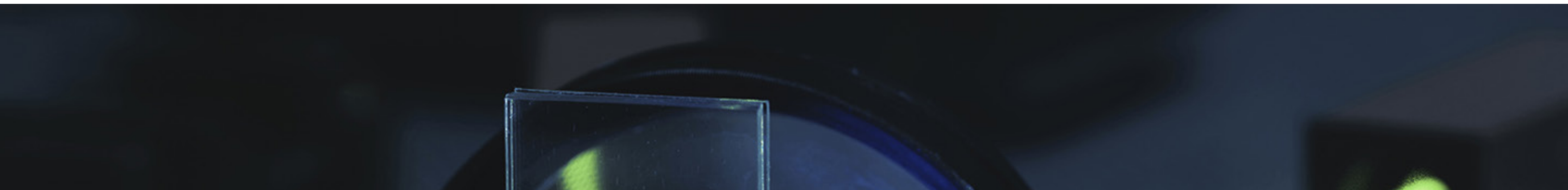
- Semiconductor companies are pushing their way up the value chain through various measures, such as by providing software that is optimized for their chips.

Ondrej Burkacky is a partner in McKinsey's Munich office, where **Klaus Pototzky** is a consultant; **Johannes Deichmann** is a partner in the Stuttgart office; and **Benjamin Klein** is a specialist in the Berlin office, where **Gundbert Scherf** is a partner.

The next wave of innovation in photonics

More end products are integrating lasers with sensors and optics, opening new opportunities for photonics manufacturers.

This article is a collaborative effort by Gaurav Batra, Ryan Fletcher, Kairat Kasymaliev, Abhijit Mahindroo, and Nick Santhanam, representing views from McKinsey's Advanced Electronics Practice.



Einstein laid the foundations for laser technology in his groundbreaking article "The quantum theory of radiation," published in 1917. After years of development, the first widely commercialized lasers hit the market in the 1960s, when they were used for applications ranging from science to surgery. Since those early days, the unique ability of lasers to create a narrow, focused beam of light has enabled many other use cases, including barcode scanning, DNA sequencing, and semiconductor-chip manufacturing. In one of the most novel applications, the NASA rover *Curiosity* used laser-enabled equipment to blast rocks on Mars, allowing scientists to analyze chemicals in the resulting vapors.

Although the laser market has steadily increased since the 1970s, innovation and revenue growth have slowed over the past decade. Many low-cost companies have entered the market as the core technology has matured. That put pressure on the average sales price for lasers used in high-volume end products, including those related to telecom transmission, marking and engraving, and biosensing. But the sector may now be on the cusp of a new age of innovation in which lasers are increasingly combined with optics and sensors to enable even more sophisticated applications. These integrated devices, many of which are still in development in a number of industries, could not only put the laser market back on a high-growth trajectory, but also become the main source of value.

To help photonics-industry stakeholders evaluate the opportunities ahead, we assessed recent developments across laser end-markets. We then explored the optics and sensor sectors in detail, focusing on the unique capabilities that such technologies can provide when combined with lasers. Industry stakeholders—including owners, operators, and board members—have recognized these advantages and are quickly moving to broaden the technological capabilities of their companies through mergers, acquisitions, and strategic partnerships. Investors too are taking heed.

An evolving and exciting market

Although laser technology has continuously matured since its inception, two eras of innovation stand out. Through the 1970s and 1980s, researchers made important discoveries in core laser physics that advanced the technology, although many applications were limited to scientific, laboratory, and R&D settings. Over the most recent three decades, laser devices truly moved from the lab to the commercial sphere as they were refined to improve performance, robustness, and reliability. Many new laser applications, such as surgery, lithography, and welding, emerged at this time, enabling breakthroughs in industries ranging from healthcare to electronics to industrial manufacturing. These innovations helped the laser-device market achieve a value of $17 billion by 2020.

Despite the industry's technological advances and strong revenues, some recent indicators raise concerns. Take the pace of innovation as measured by the number of patents registered. From 2001 through 2010, researchers filed more than 29,000 US applications for laser-related patents—up more than twofold from the previous decade (Exhibit 1). For the years from 2011 through 2020, however, only about 24,000 applications were filed. This drop was an aberration in an industry where patent filings have traditionally doubled each decade.

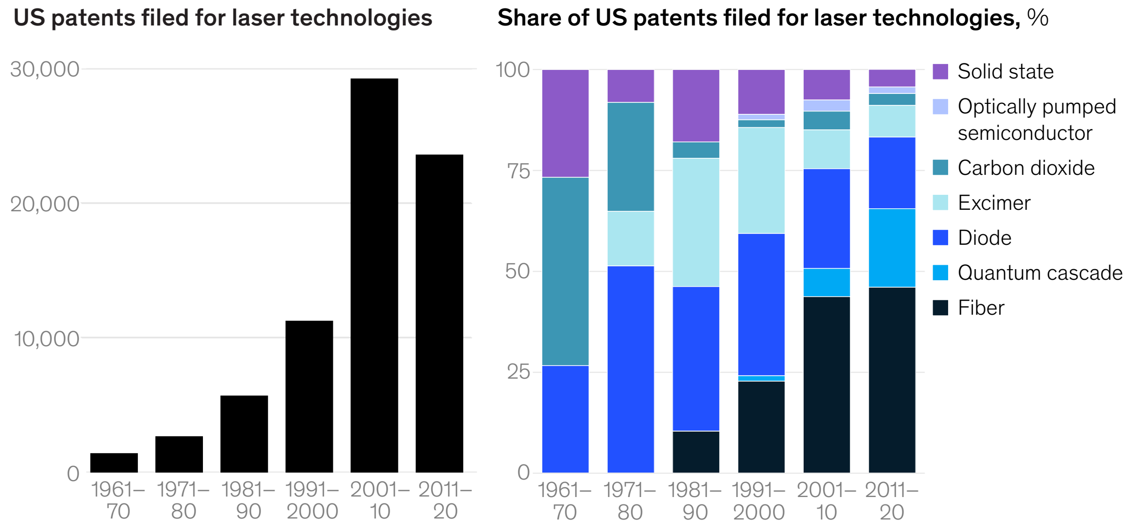
In tandem with the fall in the number of patent applications, the focus of technology is shifting for some of the most important laser technologies—fiber, diode, solid state, carbon dioxide (CO_2), excimer, and quantum cascade. (These categories are briefly described in the sidebar "An overview of laser technology.")

Fiber, the largest laser category, now accounts for more than 45 percent of all patents filed. Fiber has gained share, in part, because it can focus the beam size down to the micro level and generates more power in a smaller package than CO_2 does. It thus provides optimal speed and precision for cutting metal and welding, among other applications. Fiber has also enabled new medical applications, especially for dermatology procedures.

Although quantum-cascade lasers have encountered significant development challenges over the past 20 years, their efficiency and wavelength range have opened up new opportunities

Exhibit 1

The number of new patents is dropping, and the focus of technology continues to shift.



Source: US patent-registration data

in medical diagnostics, optical communications, and industrial-process monitoring. Recently, this category has grown more rapidly than fiber. It accounted for 19 percent of patents over the past decade, up from 7 percent from 2001 to 2010.

Overall, the number of laser patent applications is declining because these devices tend to have staying power once they gain a foothold within an industry. Innovative laser technologies traditionally require decades of R&D and hundreds of millions in funding before they are market ready, so companies are not likely to search for alternatives once they find a workable solution for an application. That in turn means that the category share for all core laser technologies is not expected to shift substantially over the next few years (Exhibit 2). Diode, fiber, solid-state, and CO_2 technologies, which now account for 90 percent of laser revenues, will continue to dominate the market. Fiber technology is projected to see the most growth, primarily because of its simple design and cost advantage over other laser types.

As in the past, most industries will rely on more than one type of laser, since their applications have diverse needs. For example, industrial companies may use fiber lasers to cut metal but rely on CO_2 lasers for plastics, glass, and wood, given different emission wavelengths and performance requirements by material type.

Even with the number of patent applications down, the laser-device market should have relatively strong 10 percent growth through 2025, and is expected to reach a value of about $28 billion (Exhibit 3).

An overview of laser technology

Lasers can use solids, liquids, or gases as a gain medium (a source of optical amplification) to create the desired beam of coherent light. Such beams are composed of photons—particles representing the smallest discrete amount, or quantum, of electromagnetic radiation—that have the same frequency and waveform. This uniformity prevents the beam from spreading and diffusing. Gas lasers use CO_2 or other gases as their gain medium and typically provide more uniform emission, with less loss, than solid-state or liquid lasers do. Examples of laser categories include the following:

- Quantum-cascade lasers emit light in the mid- to far-infrared portion of the electromagnetic spectrum. Often used in military-sensing applications, they have promise for optical communications, medical diagnostics, and industrial-process control. Their high cost is still prohibitive in many cases, however.

- Excimer lasers emit very concentrated light in the ultraviolet region of the spectrum. They are far costlier than fiber but can produce high power at lower wavelengths.

- CO_2 lasers, the highest-power continuous-wave gas lasers now available, are still some of the most widely used devices. They can be less expensive than other types of lasers, but maintenance costs outweigh this advantage in high-use applications, such as materials processing.

- As their name suggests, solid-state lasers use a solid gain medium. They have recently been losing market share to fiber and diode lasers because they are often less robust and efficient, as well as more costly to maintain.

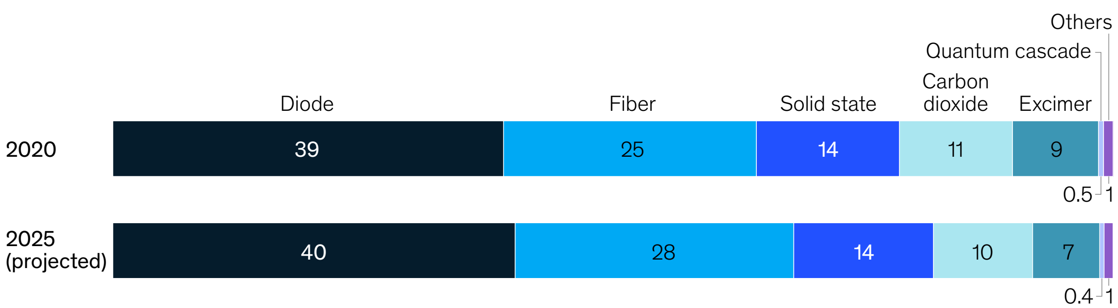
- Fiber lasers rely on optical fibers doped with rare-earth elements, such as erbium, neodymium, or ytterbium. Their use has grown steadily because of their simple, robust design and cost advantages versus other categories.

- Diode lasers generate radiation through semiconductors composed of alloys of aluminum or gallium rather than silicon. They are seeing steady growth as their power levels increase and range of wavelengths expands. Their cost is also falling.

Exhibit 2

The use of core technologies is expected to remain relatively stable through 2025.

Laser devices' share of revenue by technology, $ billions[1]



Note: Figures may not sum to 100%, because of rounding.
[1]Complete laser device, including emitter, primary optics, power supply, basic driving electronics, and thermal management (if required).
Source: McKinsey analysis

The aerospace-and-defense sector is well positioned to achieve the highest growth per year (24 percent), given increased use of high-performance, high-cost lasers for sensing, tracking, and countermeasures. A few applications, such as data storage and printing, are likely to decline as next-generation technologies shift away from laser.

The industries seeing the strongest growth increasingly depend on applications that combine lasers with advances in other technologies:

— optics, including a range of active and passive materials that can direct, filter, or change certain portions of light

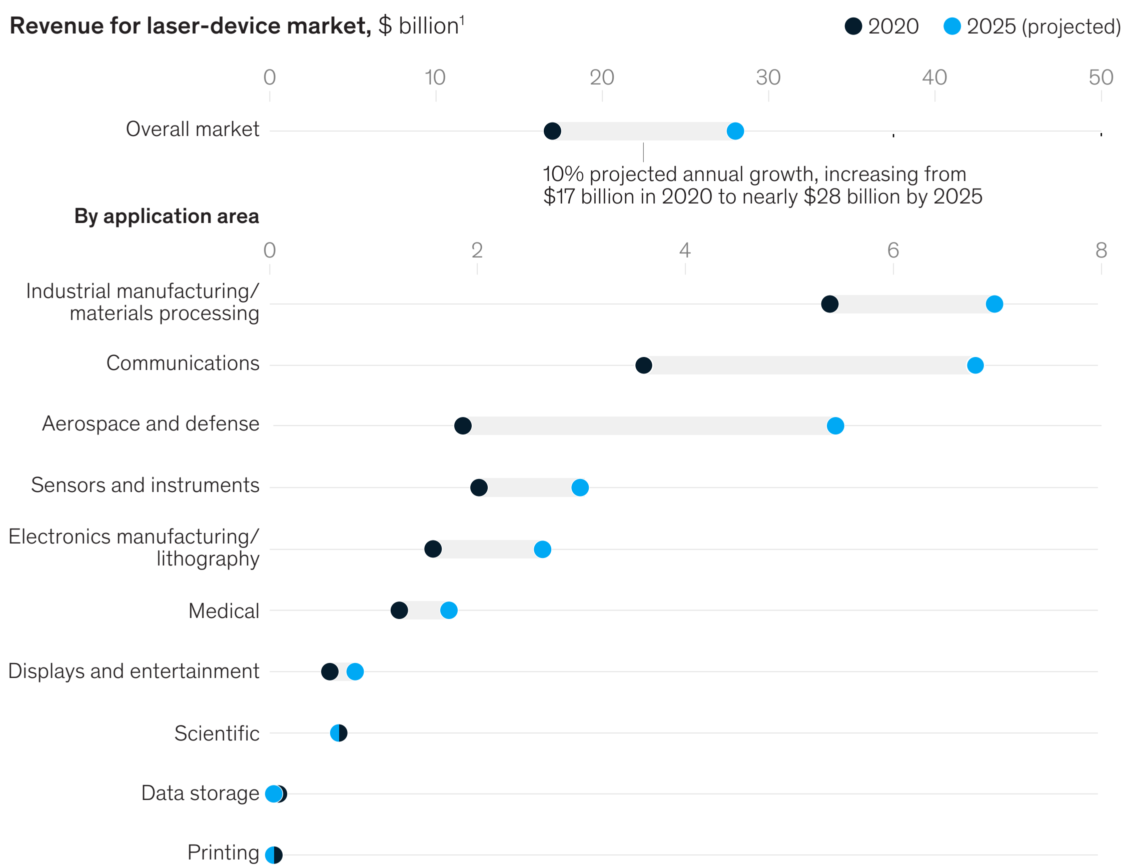
— photonic sensors (which detect precise emissions of light or energy within the photonic spectrum), including some UV and IR wavelengths, which are processed into information about the environment or application in which the sensor operates

Carefully coordinating optic, laser, and sensor technologies—particularly for power, wavelength, and optical design—is critical for their success.

In addition to expanding the number of potential applications, optics and sensors can also take laser performance to a new level. For example, integrated devices are already critical to optical-coherence tomography, a noninvasive procedure for taking 2-D and 3-D images of retinal tissue. To determine the full potential of integrated laser-based systems, we first examined the precision optics and photonic-sensor sectors, looking at core technologies, recent growth, and go-forward adoptions. We found that both markets are now thriving and that the uptick in

Exhibit 3

The laser-device market is expected to grow by 10 percent annually through 2025.



[1]Complete laser device, including emitter, primary optics, power supply, basic driving electronics, and thermal management (if required).
Source: McKinsey analysis

integrated laser devices could increase their value even further.

The optics market

With an overall value of $33 billion, the optics market includes components that can enhance the precision and performance of laser-based systems, such as:

— filters, usually consisting of coated glass or plastic, that selectively transmit specific wavelengths of light while blocking or reflecting others

— lenses, classified by the type and degree of their curvature, that focus or disperse light

— mirrors that selectively reflect light, typically to steer or fold beams

— beamsplitters, which separate light (typically by wavelength or direction), and are often used in devices that include both lasers and sensors

— prisms, typically machined from glass, that disperse light into its components by wavelength

— adaptive optics, which typically integrates multiple optical components and adjusts their properties through mechanical or electrical articulation and control

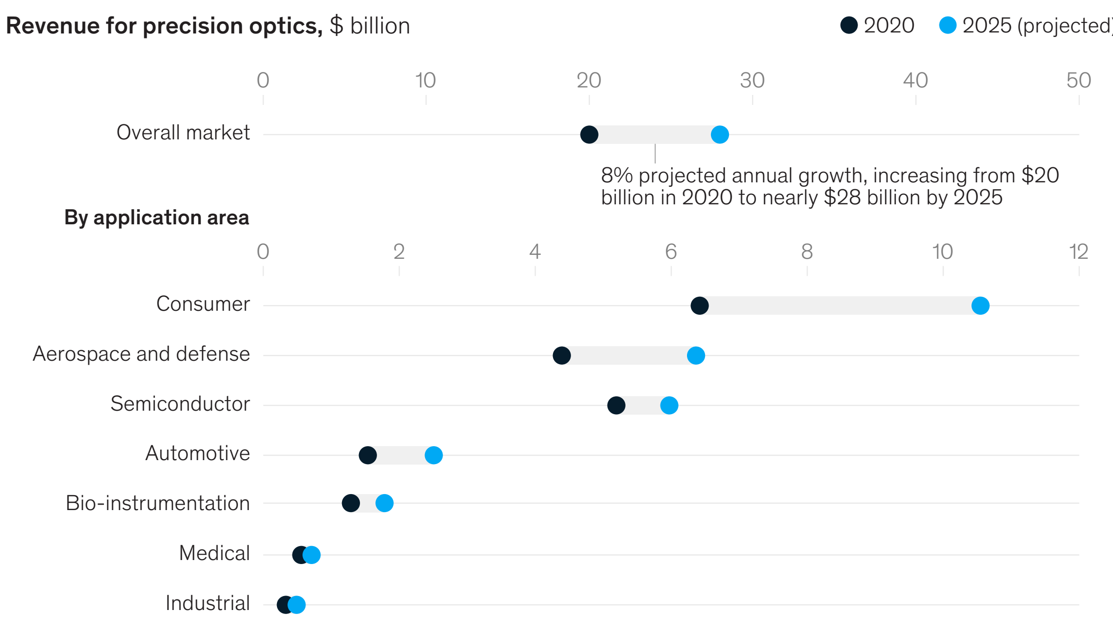
Precision optics, valued at $20 billion, represents about two-thirds of the value of the total optical-components market, and strong growth of 8 percent is expected through 2025 (Exhibit 4). Consumer applications, such as biosensing, security, and portable device LiDAR, are likely to drive most of the demand. The automotive, semiconductor, and space

Exhibit 4

Most segments of precision optics are growing significantly.

Revenue for optical components by type in 2020, %

Revenue for precision optics, $ billion



[1]Precision optics is defined by tighter manufacturing tolerances (eg, thickness, diameter, centering, curvature, surface, and coating uniformity) 2–10x higher than those of standard optics, at a minimum.
Source: McKinsey analysis

sectors will also account for a large proportion of precision-optics revenues.

To understand the increasing impact of precision optics on the performance of laser-based devices, consider flow cytometry. In this process, a laser is projected through a biological sample to evaluate the physical and chemical characteristics of individual cells or particles, including those in blood. Flow-cytometry systems use bandpass and dichroic filters to restrict the wavelengths of light that pass to detectors, allowing scientists to identify specific cells or particles within each sample (Exhibit 5).

In recent years, researchers have improved flow-cytometry filters to increase their accuracy and precision and to make it possible to identify multiple components within a single sample simultaneously. These upgrades have pushed the limits of design and manufacturing. The importance of leading-edge bandpass and dichroic filters is reflected in the fact that they can represent, on average, 10 to 20 percent of overall flow-cytometer system costs.

Other innovators have enhanced flow cytometry by replacing traditional optics, including mirrors and filters, with a dispersive optic spectrometer. In addition to improving the accuracy of these devices, such innovations have significantly accelerated the throughput of samples.

The sensor market

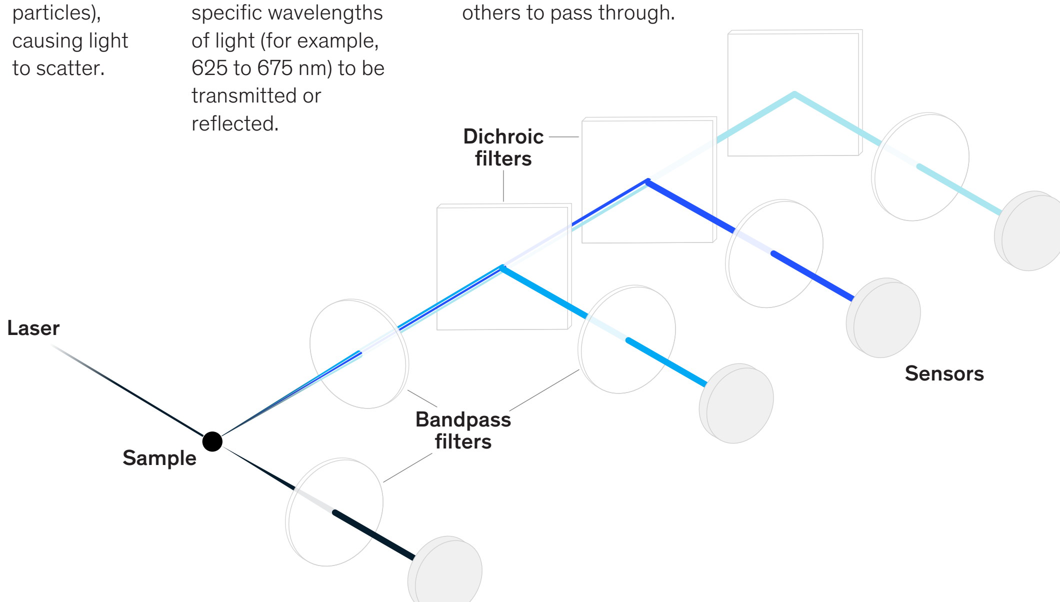
Photonic sensors represent a $29 billion market—about 16 percent of the broader $180 billion sensor market. This segment is projected to see

Exhibit 5

The performance of flow cytometers relies on the precision and consistency of optical filters.

How filters are used in a flow cytometer to determine the characteristics of a cell or particle[1]

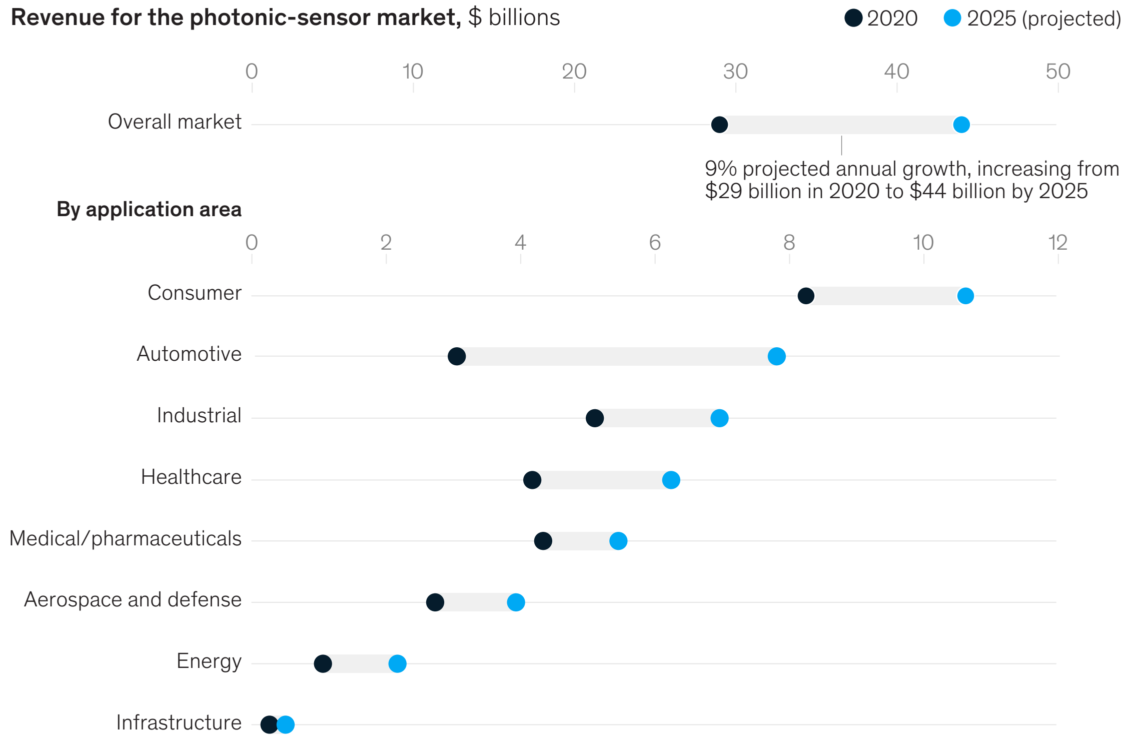
1. A laser is directed at a sample (a collection of cells or particles), causing light to scatter.
2. The scattered light passes through optical filters, typically bandpass and dichroic. These cause only specific wavelengths of light (for example, 625 to 675 nm) to be transmitted or reflected.
3. Dichroic filters are arranged at an angle to reflect specific wavelengths toward a sensor while allowing others to pass through.
4. The filtered light is received by sensors, each of which detects the presence or absence of different parameters, corresponding to specific wavelengths of light, within the sample.



[1]Specific requirements vary depending on the parameters of a given detection system.
Source: McKinsey analysis

Exhibit 6

Photonic sensors are expected to experience 9 percent annual growth through 2025.



Source: McKinsey analysis

strong growth of 9 percent annually through 2025, reaching $44 billion in revenues that year (Exhibit 6). A few application areas stand out:

— In automotive, demand is expected to grow by 21 percent a year because advanced driver-assistance and autonomous-driving systems require sensors with high precision and resolution.

— Annual growth is expected to reach 8 percent for aerospace and defense as automated applications, the expanded use of aerial LiDAR, and new remote-sensing tools drive demand.

— The energy sector could grow by 15 percent a year as fiber-optic sensor technology is incorporated into monitoring and measurement applications, some of which can help reduce waste and pollution.

— Within infrastructure, sensor demand will rise by an estimated 14 percent as innovators develop more integrated laser devices to measure the physical characteristics of buildings, including strain and vibrations.

Photonic-sensor technologies include silicon photodiodes, which are broadly used in applications where a large quantity of detectors are required. For instance, silicon photomultipliers are employed in LiDAR (which uses light in the form of pulsed lasers to measure distance) and time-of-flight use cases (which involve determining distance or depth between the source and another object). Similarly, charge-coupled-device (CCD) sensors and complementary metal-oxide semiconductor sensors, both of which use silicon photodiodes, have broad use cases in spectroscopy, machine vision, and defense applications.

As another example, industrial cutting lasers used in manufacturing are beginning to gain new capabilities through the integration of precision optics and sensors. Initially, machine operators set parameters and the laser completed the cut

As OEMs increasingly turn to photonics systems to address customer needs, the lines between component suppliers, subsystem providers, and device integrators will probably continue to blur.

exactly as ordered, with no midprocess adjustments. More recent devices include sensors that detect parameters, such as surface finish, density, depth of cut, and thermal stress on materials. Such devices not only provide for real-time adjustments but also contain precision optics, often beam-splitting filters, to enable both laser cutting and laser measurement in the same optical path (Exhibit 7).

Next steps for companies in the photonics sector

As industry stakeholders pursue opportunities within photonics and integrated devices, M&A merits increased attention. Despite the recent wave of deals, the laser-device industry remains fragmented, with many small players with less than $250 million in revenues focusing on specialized niches. This fragmentation suggests that operators, board members, and investors may find many opportunities for synergistic combinations or partnerships.

Some laser manufacturers and end customers are already pursuing such deals to facilitate the creation of devices that integrate precision-optics, sensors, and lasers. For instance, a major supplier of lithography systems recently acquired a precision-optics company to gain additional capabilities for extreme-ultraviolet and deep-ultraviolet products. Another leading industrial-applications company acquired minority stakes in some laser-technology firms to boost its capabilities in materials-processing applications. It also acquired a company that manufactures many of the photonics components and products used in sensors for autonomous driving, smartphones, and digital data transmission.

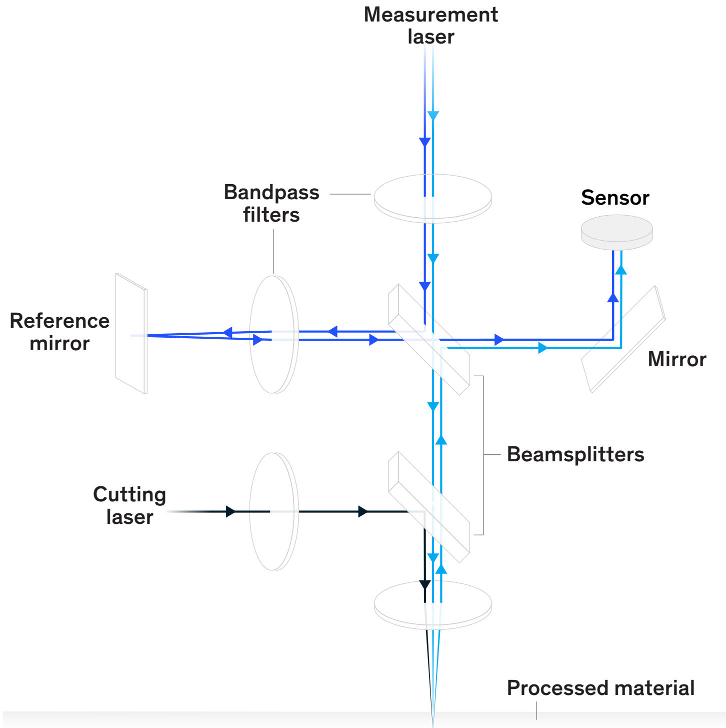
As integration among lasers, sensors, and optics becomes increasingly important for creating value in next-generation systems, operators and board members must rethink their product strategies and reposition themselves along the value chain. For instance, the need for effective integration and real-time monitoring is likely to increase the importance of software in this traditionally hardware-driven industry. New serviceability requirements, such as remote diagnostics, adjustments, and calibration, could also create additional opportunities to provide services over the life of each system. And as OEMs across market segments increasingly turn to photonics systems to address customer needs, the lines between component suppliers, subsystem providers, and device integrators will probably continue to blur.

Like any high-tech sector, photonics must innovate to survive. Although the speed of innovation in

Exhibit 7

New quality-monitoring systems use sensors to improve precision in laser cutting.

Paths of cutting and measurement lasers in an industrial-cutting machine



1. The cutting laser passes through bandpass filters and a beamsplitter to reach the target material.
2. Additional bandpass filters and beamsplitters enable the beam from a measurement laser to reach the target material along the same axis, as well as a beam that hits a reference mirror. Light from both eventually reaches a sensor, where it is compared and analyzed to detect material properties, depth, and quality of cut.
3. Information processed from the sensor is used to adjust the cutting laser's power, position, and speed in real time.

Source: McKinsey analysis

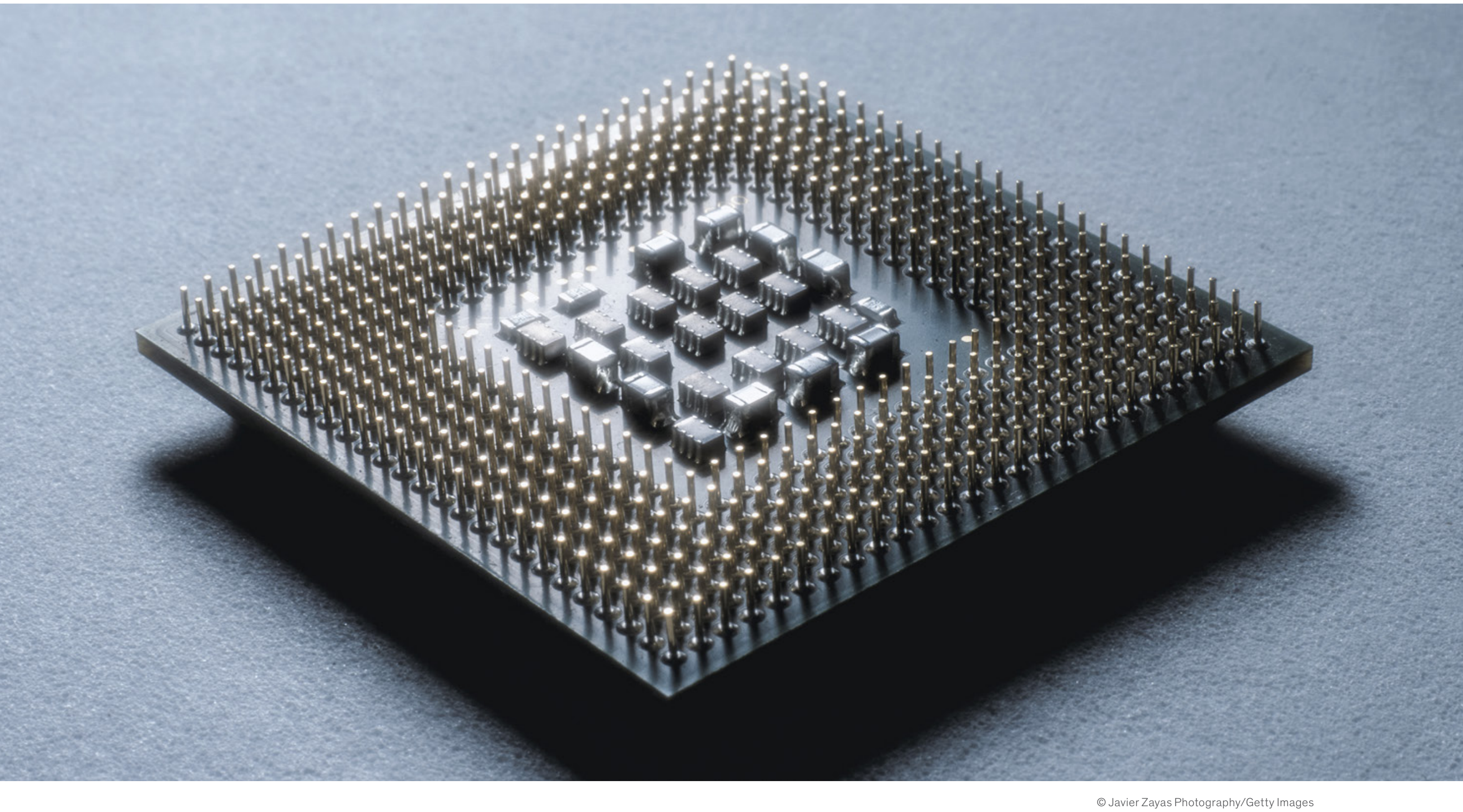
laser technology has been dropping, the creation of integrated devices combining lasers, sensors, and optics could usher in a new age of opportunity. Companies that develop such devices now could have a first-mover advantage, since end customers are likely to seek strategic partnerships to explore new applications and build product offerings. The move to integrated devices might require new capabilities, but opportunities abound for rapidly sourcing them within the fragmented industry landscape.

Gaurav Batra is a partner in McKinsey's Washington, DC, office; **Ryan Fletcher** is an associate partner in the Southern California office, where **Abhijit Mahindroo** is a partner; **Kairat Kasymaliev** is a consultant in the New York office; and **Nick Santhanam** is a senior partner in the Silicon Valley office.

The authors wish to thank Barbara Castro, Brendon Earle, Paulo Guimaraes, Jo Kakarwada, Adam Matula, Jwalit Patel, Mark Sawaya, Emily Shao, Dan Trunzo, and Scott Whitehead for their contributions to this article.

Navigating through change: An interview with NXP Semiconductors' Kurt Sievers

A new CEO reflects on the past year and looks at what's ahead for semiconductors.



© Javier Zayas Photography/Getty Images

Kurt Sievers was nominated to be president and CEO of NXP Semiconductors in early March of 2020, days before the COVID-19 pandemic began shutting down economies worldwide. McKinsey's Abhijit Mahindroo, Sven Smit, and Anupama Suryanarayanan recently spoke with Sievers about the challenges presented by the COVID-19 crisis, NXP's evolving strategy, and the future of the semiconductor industry.

McKinsey: Your first days have been spent shepherding NXP through the COVID-19 crisis. What have you learned about leadership during these times?

Kurt Sievers: It's important to remember how a well-considered succession plan can help a company. I am a very lucky member of an organization with a very thoughtful plan. NXP began preparing the plan at least two years before I took over in March, which made it comparatively easy for me to take the lead, despite the pandemic and the suddenly difficult environment. For my own succession one day, I will do my best to replicate this plan.

Second, human relationships clearly matter, and I was fortunate to have relationships with all essential stakeholders when I was announced as CEO, because I wasn't new to the company. I wasn't new to investors, customers, or employees. You learn the importance of human relationships in those moments when you have to rely on other people.

I made very timely decisions on work-from-home regimens, keeping visitors from our sites, and similar topics during the pandemic, so we took action more quickly than many governments or other institutions. NXP's actions are not simply a product of my wisdom; I just had the courage to listen to my advisers and move quickly.

The game isn't over yet, but we've been successful in keeping the infection numbers in the company very low. In addition to protecting employees' safety and health, this has allowed us to maintain business continuity. We didn't have to stop R&D projects or close factories because of the timeliness of our decisions. Sometimes I think it's better to make a quick decision, even if you don't have all the context yet.

Finally—and this is more of an experience than a learning—it's amazing how far you get if you manage to keep your people engaged. It's amazing what an organization can deliver with such people on staff, even if they are working remotely and separated from each other.

McKinsey: What will be different about the post-COVID-19 world?

Kurt Sievers: The use of digital tools will not go backward after COVID-19 is controlled, since we've all learned to use them for the better. I don't believe that all companies will maintain a work-from-home routine for all employees, though, and certainly not in our industry. Like many other industries, the semiconductor sector involves complex innovation, and that's not a one-person job. It's something you have to do together, and I think inventiveness and the power of innovation are suffering because people are not together in a room. That's why I'm not so sure that corporate life will change that much. I also believe that the focus will return to global trade issues and climate change—two things that are more sustained challenges than the pandemic.

McKinsey: Let's shift gears from the macro-environment to NXP. NXP has been a leader in the semi industry—one of the first players in transistors and integrated circuits, one of the first integrated device manufacturers, one of the first companies to undertake large-scale private-equity plays in semi, and one of the first to follow an asset-light model, among other portfolio-shaping moves. Looking forward, what is the path ahead for NXP?

Kurt Sievers: We do have a history of firsts, and a lot of our tech leadership is our heritage from Philips and Motorola, given our acquisition of Freescale.

I think in ten-year steps. My vision of the future is based on the fact that semiconductor growth is driven by very few absolute killer applications, and these come and go. Between, say, 2000 and 2010, the focus was on computers and laptops; between 2010 and 2020, smartphones, tablets, and cloud computing were the big deal. In the next ten years, cloud computing will continue, but it will be very strongly complemented by edge computing and edge applications. That's not just the processing itself, but everything that comes with it.

Over the past five years, we have assembled a portfolio that aligns with the needs of edge applications. That's why we acquired Freescale. We needed the processing portfolio to improve our capabilities. We also needed connectivity, which is why we acquired wireless-connectivity assets from Marvell almost a year ago. We do have very good capabilities in low power, cybersecurity, and functional safety. We are building muscle in artificial intelligence. Those are the elements that are really needed to build complete edge applications.

I really think we are fortunate. It's partially a result of how we've been building the portfolio over the past years, but it also feels like the time is ripe for finding an opportunity out there.

McKinsey: NXP has undertaken some major portfolio realignment. Can you talk about the company's evolution?

Kurt Sievers: We deliberately went away from mobile communications on the modem side and from products for digital consumers, including TVs, optical storage, and even audio offerings. Those were good businesses from a market perspective, but we felt that sustainable growth with reasonable profitability was impossible. We instead focused on building automotive, even though the sector wasn't a given in those days. We also built a position in industrial, which was a well-kept secret but is now becoming more popular, and we're now a major player there.

We do still have a focused play in mobile—for example, contactless payments, transit ticketing, door opening in hotels. That all has to do with security, ultra-wideband technology, and the use of communications. Although I wouldn't say we're a mobile company, we are experiencing strong growth in this area because we use the mobile ecosystem

'In the next ten years, cloud computing will continue, but it will be very strongly complemented by edge computing and edge applications. That's not just the processing itself, but everything that comes with it.'

to enable our technology, which ends up in smartwatches and different applications. We still have our communications infrastructure, and we are clearly a leader in the radio-power side of things. We're already investing in 6G, by the way, but that's more of a focused play.

Overall, you could say we really have only two large segments—automotive and industrial—and a few very well-defined focused plays, which is certainly very different from what it used to be.

McKinsey: Let's go deeper into some of these portfolio elements now. Automotive electronics are proliferating like never before, fundamentally changing the driving experience. How do you envision the future of automotive—in terms of the electronics that will be deployed—and the role that NXP will play in the transformation?

Kurt Sievers: The future is clearly about advanced driver-assistance systems [ADAS] and electrification. And that will not just dominate for the next two to four years; I believe it's going to keep us busy for easily ten years. The growth of ADAS looks straightforward on the surface, but it has very deep implications for the architecture of the car and the whole value chain. I still believe that autonomous driving is feasible, although this development is further out. When that occurs, companies will need business models that are designed to facilitate transportation, such as fleet services, rather than ones that focus on meeting the needs of individual car owners.

Electronics will be the backbone of a car, and OEMs will need to have strong electronics and software to survive. I cannot stress enough the importance of software, since I think it will largely determine who will be in the value chain and what value is created at different levels. OEMs will increasingly work directly with semiconductor and software companies to drive innovation.

That shift is obviously significant. Tesla already follows that model, and it's part of the company's DNA. The rest of the industry is still trying to catch up and understand how to make the shift. Obviously, it's more difficult when you have legacy systems and processes.

This is a prime time for semiconductor companies to help with innovations, including those that improve safety and reduce emissions. The challenge, of course, is that the volume of new cars is relatively low. Normally, we deal with industries that have very high scaling factors for individual products but that show little variation in terms of models. That lets semiconductor companies get their products into many devices. However, in automotive, there are only 85 million cars produced in a good year; for 2020, it will probably be somewhere in the neighborhood of 74 million due to the impact of COVID-19. Unless you have an application that is very common in cars, volumes will be low. You will need either incredible market share or very high levels of standardization across the industry. Otherwise, you will never achieve the scale required to recoup the investment you made to develop sophisticated systems.

The good thing is that trends like advanced driver-assistance systems and electrification drive strong growth in semiconductor content in cars. One example is radar, which is in the ADAS domain, where we have already reached a $500 million run rate. That's obviously still embryonic because there are many, many cars that don't even have radar yet. We will see a triple acceleration of this because all cars will eventually get radar; many will get multiple radars, such as front-facing radars and side radars for blind-spot detection. The value per radar system is going up because the performance requirements are continually increasing for accuracy and other features. NXP has a large market share for radar.

That brings up another point: automotive requirements are becoming incredibly complex, and that will be an issue. In the past, semiconductor companies focused on simple microcontrollers and analog products within automotive. Quality was important, but the products were relatively unsophisticated, and the required investment was small. The business case for simple products was decent because you could rely on long-

‘Small industrial companies are typically not knowledgeable about electronics and have a great need for solutions. . . . Industrial solutions may also be more sticky than automotive ones because nobody wants to change the design in that market.’

standing demand. Now, the equation is tilting because the product volumes are small but the complexity per application is skyrocketing. I don’t think that the industry has determined how it will deal with this challenge.

McKinsey: You called industrial applications the industry’s best-kept secret. What are some of the challenges related to the Industrial Internet of Things, and how is NXP gearing up for those?

Kurt Sievers: Industrial is totally different from automotive. The market includes Siemens, Schneider, and a few other large companies, but there’s also a very large number of relatively small companies, so the name of the game is distribution. You need to be exposed in the distribution channel to be able to serve the market.

Small industrial companies are typically not knowledgeable about electronics and have a great need for solutions. In the best case, they will need a reference design from us with a very well-developed software-development kit that allows them to get a product out in two or three months. Industrial solutions may also be more sticky than automotive ones because nobody wants to change the design in that market.

NXP has the required solution capabilities, including security, connectivity, and our MCU/processor capabilities. Connectivity was a missing piece. That’s why we acquired wireless-connectivity assets from Marvell. The fact that most industrial applications are now cloud connected makes things even more interesting, and we know how to enable this.

Small industrial companies need someone like us to facilitate the relationship with the large cloud players.

McKinsey: Let’s discuss your focused plays, including your role in mobile payments. After powering semiconductor growth over the past decade, the mobile sector has matured in terms of growth. What disruptions or innovations do you foresee for this sector?

Kurt Sievers: As I mentioned, we’re not a mobile player per se but have built a position in mobile wallet over the past six to eight years. Mobile-payment applications usually have three anchors: hardware-based cybersecurity, near-field-communication [NFC] radio, and a great deal of security software. NXP’s role in mobile-wallet software may be a well-kept secret, but we are definitely not just a hardware company. If we did not offer a significant portion of the software required, we would never have been able to help smaller companies that have lagged behind the large players that went into mobile wallet earlier.

We are interested in expanding our mobile-payment technology to other use cases, such as transit ticketing, and we have made a big move into ultra-wideband radio, where NFC is one of the enablers of the mobile wallet. The UWB [ultra-wideband] chip, which is just a different radio, will be the enabler of completely different use cases. Just like we use mobile payment to get rid of coins and cash, we could use ultra-wideband technology to get rid of mechanical keys for autos and houses if it is embedded in smartwatches, phones, or other devices. The principle is the same as with mobile payments—you need security at the personal level. NXP is already deeply embedded in IoT [Internet of Things], and we are working with lock companies to create electronic door locks. Next year, many prominent car companies will allow you to open vehicles with your phone based on NXP's ultra-wideband technology.

McKinsey: Shifting gears, recently there have been many big-ticket M&A deals announced within the semiconductor industry. If they proceed, 2021 could see another wave of consolidation. What does this trend portend for NXP and the broader industry?

Kurt Sievers: In general, I think the semi industry is not yet at the end of the required consolidation. At the same time, I do believe that geopolitical issues may complicate M&A.

I believe many companies are becoming involved in deals because of their very high valuation. It's just that they have new currency called equity, so they are leveraging that—nothing unusual. We have our eyes wide open and know what others are doing, but we're not going to be misled by this and just fall into a fever of doing a deal. We know where we want to go.

McKinsey: Let's expand on the geopolitical issues you mentioned. In 2020, we also saw unprecedented geopolitical developments that had material

Kurt Sievers

Education
Earned a Master of Science degree in physics and IT from the University of Augsburg

Career highlights
NXP Semiconductors
(2020–present)
President and CEO

(2018–20)
President

(2009–18)
Member of executive management team

(1995–2009)
Progressed through a series of sales and marketing, product definition and development, strategy, and general-management-leadership positions across a broad number of market segments

implications for the semiconductor industry. How do you see this playing out in 2021 and beyond?

Kurt Sievers: Many trade tensions are really about the ownership of technology. That has been true historically and will continue to be true. Economic leadership will depend so much on technology over the next ten, 20, or 30 years.

An industry like semiconductors requires global scale, and that's why geopolitical issues are becoming such an inhibiting factor. In the worst case, they force companies to burn a lot of money and decelerate the speed of innovations that could have helped the world manage problems such as climate change and poverty. That's the tragic negative implication from all of this.

McKinsey: What advice would you give to people who are just starting their careers in the semiconductor industry?

Kurt Sievers: My advice is totally independent of semiconductors. Ideally, you should have the courage to take the time to find out what you really like to do. That is not a naive statement. Of course, you will have a couple of mornings when you get up and don't feel particularly enthusiastic about your day, but you should be in an area that plays to your strengths, ignites your passion, and ideally is more than a profession for you. That's my most important advice—to have the courage to do this.

Learning from others is a big deal. Lessons that come from a book are not the most important; it's what life teaches you in professional and nonprofessional circumstances, and that really has to do with the individuals with whom you interact. There's a little bit of luck involved in finding good leaders to serve as role models; it's not always easy. I've been fortunate to work with many different good leaders. From each person, I could pick and choose what appealed to me.

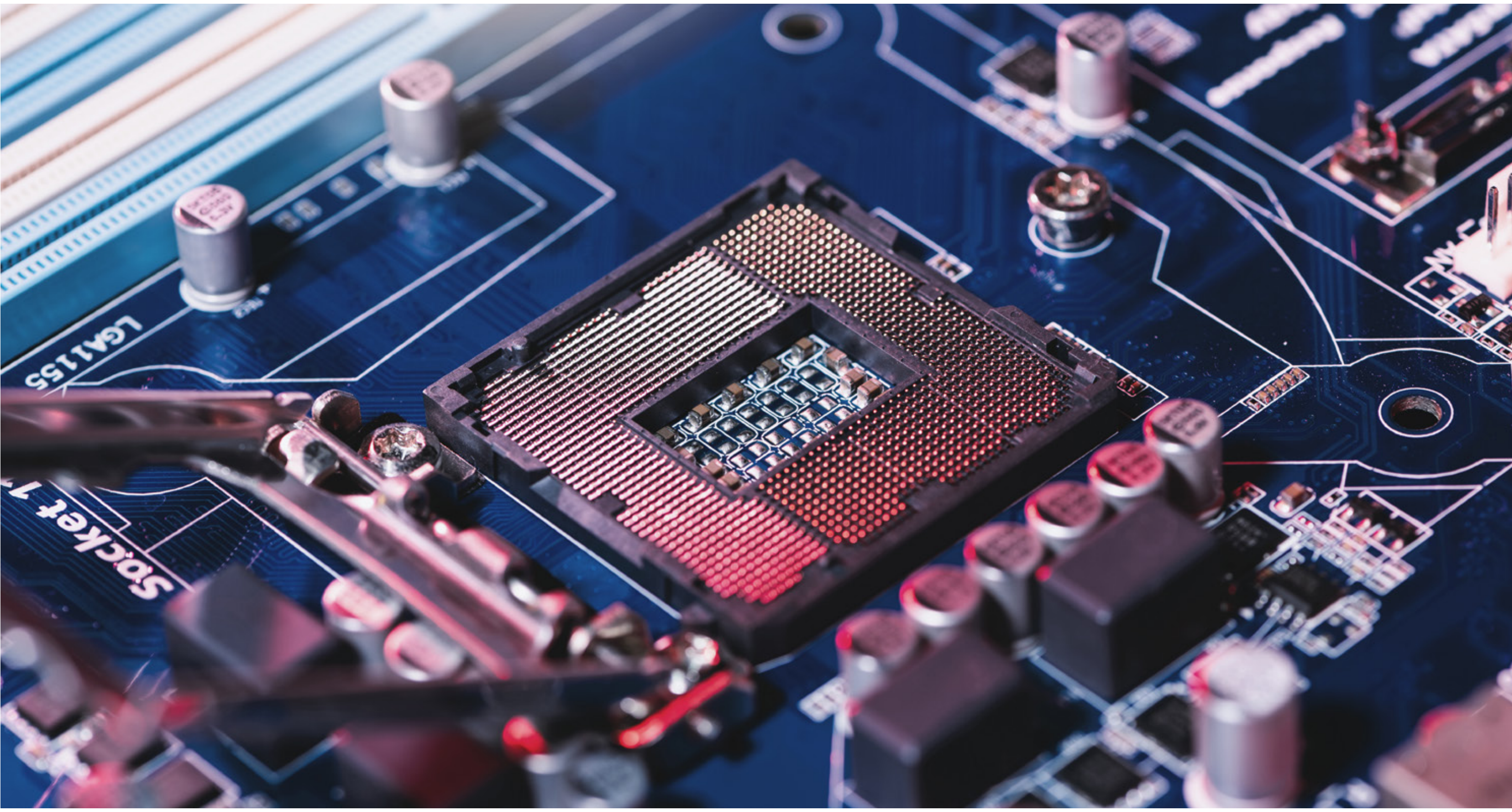
The semiconductor industry is small; you meet the same people over and over again. Learning from mentors and growing through character is essential for keeping you in the game—hopefully, ahead of it—and sustaining you. If you do not have character, you may have a lot of quick wins but not sustainable success.

And, lastly, if you want to lead people, you need people who actually love to follow you. They will do it because they like you. That doesn't mean you are everybody's favorite, but people understand you because you're authentic and you feel trustworthy. It boils down to building your character and being real.

Kurt Sievers is the president and CEO of NXP Semiconductors. This interview was conducted by **Abhijit Mahindroo**, a partner in McKinsey's Southern California office; **Sven Smit**, a senior partner in the Amsterdam office; and **Anupama Suryanarayanan**, a consultant in the Silicon Valley office.

Renesas's Hidetoshi Shibata on leadership through difficult times

A CEO discusses past and future strategy in the evolving semiconductor sector.



Tokyo-based Hidetoshi Shibata is licensed to drive the Shinkansen bullet trains that are famous worldwide for their speed and punctuality. Some might say this was good preparation for his current role leading Renesas Electronics Corporation, the Japanese semiconductor company with annual revenues of about $7 billion, in an era of nonstop change for the semiconductor industry. McKinsey's André Andonian, Abhijit Mahindroo, and Anupama Suryanarayanan recently spoke with Shibata-san, who was appointed president and CEO of Renesas in July 2019, about his aspirations for the company and the outlook for the semiconductor industry.

McKinsey: What was it like to drive a Shinkansen bullet train?

Hidetoshi Shibata: Back then, I may have been one of the fastest men on earth. Now I sometimes think I might need to move even faster.

McKinsey: Well, you've taken some quick actions since taking over Renesas. What are your hopes for the company, and how would you like it to evolve?

Hidetoshi Shibata: I constantly ask myself the same question. What type of growth, or what type of transformation, would I like to see? I choose to be pragmatic and realistic, so what matters most is the relative growth—relative to both our past and to our industry peers. First and foremost, I would like to take Renesas back to a robust growth path. There are multiple ways to do this, and we have a compelling product suite and technology offerings that can help our customers increase their business. But, candidly, what has been challenging at Renesas is the organizational capability required to get those products and solutions to those who desire them most, in the timeliest and most price-competitive manner. In my first six months at Renesas, I focused on reshaping our growth strategy. Now I'm completely focused on execution.

McKinsey: It was an ambitious decision to acquire Intersil and IDT [Integrated Device Technology]. Can you talk a little bit more about the rationale for those deals? In hindsight, what worked well in terms of integration, and what, if anything, would you have done differently?

Hidetoshi Shibata: Coming from the private-equity and investment sphere, I believed we needed to expand our portfolio to include many more analog and mixed-signal offerings, especially in areas like power management, RF [radio frequency], and connectivity. So that was one reason for the deals.

I also carefully designed the sequence of the acquisitions to make sure they'd be successful. If you look at Intersil, it is primarily a power-management supplier. If you look at our SoCs [systems on a chip] and microcontrollers, there's a power-management integrated circuit [PMIC] attached to each one. We used to use PMICs from other competitors on the reference board. Internalizing Intersil capabilities and their product portfolio allowed us to offer a comprehensive solution.

At the same time, I was really careful about the cultural fit. Like it or not, we are what we are. Integrating with Intersil, which has traditionally been a little more aggressive organization, could have come with significant challenges. But I wanted to enable cultural and behavioral changes from this cross-pollination and combination of different cultures.

To my pleasant surprise, younger colleagues in Renesas, who had begun their careers with the company, told me that they enjoyed having supervisors from different countries. They thought it was refreshing and exciting, and they liked the way the company was changing. That said, the cultural transformation was eventually not as dramatic as I'd expected, probably because the two companies had many cultural similarities from the beginning. Overall, however, the integration went very smoothly.

Then came IDT. Its leadership is even more aggressive in financial management and pursuing growth. So now we are in the middle of more exciting cultural and behavioral changes, particularly in areas outside auto.

McKinsey: Renesas is the result of several integrations—NEC and the chip operations of Mitsubishi and Hitachi. You also have a diverse investor base that includes Nissan and Toyota, as well as the Innovation Network Corporation of Japan. What challenges and opportunities arise from the mix of cultures? How can you use your base as an advantage?

Hidetoshi Shibata: There are two interesting observations related to Renesas being a mixture of three cultures. As you can imagine, while Hitachi, Mitsubishi, and NEC have different cultures and behavior patterns, they also have many similarities. In such a situation, people can become myopic and focus on the gaps or differences. Also, culturally, because all of the companies involved were Japanese, the organizations were used to making decisions by committee, rather than assigning individual responsibility. In a case like that, often no one accepts the responsibility for bad decisions.

Another negative was having a patchwork of business processes and systems because there was no central force to drive the entire organization in one direction. The introduction of IDT and Intersil has been a great help in that respect, since it helped open the eyes of a lot of colleagues who had spent their entire careers with Renesas. They saw the necessity of adopting simpler, industry-standard practices—for example, in IT systems and business processes—to accommodate further acquisitions. Otherwise, we'd have to engage in step-by-step, case-by-case discussions every time we welcomed a new partner.

Going forward, I would like to focus on the diversity that we now embody as a result of those mergers and acquisitions. Even within Japan, I believe this will set a solid foundation for increasing diversity even further in terms of technical backgrounds, cultural backgrounds, religions, and races. Hopefully, I can make Renesas a truly cosmopolitan organization down the road.

McKinsey: Now let's talk about some specific sectors. Automobiles are increasingly connected and need more and more electronics. How has your strategy and approach to the automotive sector evolved?

Hidetoshi Shibata: Electronics or semiconductor content for automobiles will likely continue to rise, and that puts us in a very good position in the semiconductor space as a whole. Within automotive, it looks like the progress is occurring in two distinct ways. First, compute capabilities are increasing

'We're moving away from mastering individual devices. Instead, we're investing more in software-development tools and ecosystem enrichment, such as connectivity to the cloud.'

consistently, and that enables us to streamline the spaghetti-like electronics architectures in autos. The higher compute power is also enabling more advanced safety functions, and more enjoyable connectivity and infotainment functionalities.

A second growth angle involves simple electrification. This is about the electrification of not just the power train but also simple things, such as power seats and power windows. Historically, we've tried to provide as many discrete high-capability semiconductor devices as possible, whether they were microcontrollers, insulated-gate bipolar transistors [IGBTs], or bipolar complementary metal-oxide semiconductors [BiCMOS]. We were trying to master those products individually and provide them to industry-leading large customers. But now we are trying to take a bimodal approach. For simple electrification, thanks partly to the acquisitions of IDT and Intersil, we can now provide many more ready-to-use solutions. In some cases, we provide reference designs or proofs of concept, but others are truly systems that can be used out of the box. We are putting a lot more effort into making those system-level solutions available to our customers.

Likewise, we're moving away from mastering individual devices. Instead, we're investing more in software-development tools and ecosystem enrichment, such as connectivity to the cloud

Mr. Hidetoshi Shibata

Education

Earned an MBA from Harvard Business School and a Bachelor of Engineering degree from the University of Tokyo

Career highlights

Renesas Electronics Corporation
(2019–present)
Representative Director, President, and CEO

(2013–19)
Executive Vice President and CFO

Innovation Network Corporation of Japan (INCJ)
(2009–13)
Executive Managing Director

Merrill Lynch
(2007–09)
Managing Director, Global Private Equity

MKS Partners—a private equity firm based in Japan
(2001–07)
Partner

Central Japan Railway Company
(1995–2001)
Project manager for a large-scale business-process reengineering project; earned a license to drive Shinkansen bullet trains

services of Amazon, Microsoft, and others. That strategy is helping us gain traction, primarily in markets like China and India.

We're also seeking growth by creating scalable and more flexible digital-product offerings that allow customers to design a car's entire electronics architecture more easily. Historically, we offered rudimentary 16-bit microcontrollers, 32-bit microcontrollers, and ARM-based SoCs—all with completely different architectures, which made it difficult for customers to scale their applications across our product line, or port applications from one product to another.

We're now in the middle of making our product offerings more scalable. With the new product suite, customers don't have to rewrite their applications as they transition across different controllers. They can also make simple tweaks between the central electronic control unit and what some customers call zone controllers, which control specific sensors and actuators. We're making it easier for customers to make changes from one generation of cars to the next, or from one trim to another within the same generation.

McKinsey: Now a question on 5G and the evolution of wireless technology: What is Renesas's role in enabling these breakthrough technologies, especially in Japan?

Hidetoshi Shibata: Regardless of the type of connectivity, I believe our role is to provide more easy-to-use ingredients to help more customers. Historically, we focused on leading telecom-equipment providers. Now we're also serving newer, smaller companies with fast-moving, innovative solutions. We realized that we couldn't address their needs on our own, so we're also trying to strengthen, deepen, and expand our partnerships with targeted distribution partners.

So we're developing what we call "winning combination" solutions to enable customers to focus on developing their application software rather than spending time and resources on hardware configuration. We are working to provide combinations of ingredients that are proven to work together seamlessly. Along with those, we are also working to provide the relevant development tools and software and connectivity to standard cloud services. So customers will be able to open a box and immediately test their application software. If we're successful, we will be able to provide more of our technologies—like massive multiple input, multiple output [MIMO]; RF; or antenna technologies—to a more diverse customer base, and enable new applications beyond conventional cellular communications, like connected vehicles, autonomous factories, and remote healthcare.

That said, what's often missing for breakthrough technologies, at least in Japan, is demand traction. If we can drive early traction in demand for innovative technologies, it becomes a lot easier for private-sector companies like us to develop and provide the solutions to enable them.

McKinsey: Like 5G, Industry 4.0—the Fourth Industrial Revolution—is a big topic. How is Renesas planning to participate in this opportunity?

Hidetoshi Shibata: It is very similar to what we discussed regarding connectivity in 5G and beyond, with one key difference: a focus on AI [artificial intelligence], and what we call endpoint AI in particular. The endpoint represents the true point of action, as opposed to the edge of the cloud. Near-zero latency is critically important in those endpoints. So that is a key focus of our solutions. For example, we provide dynamically configurable hardware architectures that let customers enjoy the benefit of hardware-processing speeds while also providing more flexibility to change the functionality or algorithms.

We are also enriching our translator capabilities to effectively port AI models trained on industry-standard networks to our hardware. Toward that end, we are working closely with advanced endpoint-compiler software partners to enable customers to test capabilities purely in the cloud before porting them over to our boards, and then, finally, to implement these seamlessly on our chips.

‘I think the semiconductor industry will be more and more separated into two worlds: a US-centric world and a China-centric world.’

McKinsey: Do you think recent geopolitical issues will affect the semiconductor industry?

Hidetoshi Shibata: There will be both headwinds and tailwinds. I think the industry will be more and more separated into two worlds: a US-centric world and a China-centric world. Therefore, despite the continued robust growth of the industry as a whole, the size of the addressable market will become smaller for many companies. In my opinion, that will ignite additional consolidation and drive more efficiencies. It might also create a little more volatility in pricing, inventory management, demand management, and other areas, which could make things difficult for the industry. However, I'm hopeful that the industry will overcome these challenges. Further, with consolidation and the changing nature of competition, there might even be a better outlook for investors, particularly those who prioritize stable cash flows.

McKinsey: The year 2020 will be remembered for the global pandemic. What was your greatest challenge this year, and what will the post-COVID-19 world look like?

Hidetoshi Shibata: I have to admit that I was horrified, as a brand-new CEO, to see demand plummet catastrophically, especially in the auto space. But we were successful in managing our spending and production in a timely manner until the market demand bounced back. As a result, we expect 2020 to see the highest operating profit in the history of Renesas.

Back in April and May of 2020, when I was facing the most stress I'd ever faced in my job, I received a lot of compassion, empathy, and encouragement from employees worldwide. A significant number volunteered to swap out a portion of their cash pay for company stock. The trust they exhibited was very encouraging, and I realized that I could mold this organization into a more animated, motivated, high-energy group of people.

I've also realized that in both our private and professional lives, we are able to live without physically interacting with each other. This does come with a lot of limitations and can be frustrating. Even so, people are figuring out ways to connect and collaborate. This could pave the way for an efficient, hybrid model for conducting business and pursuing innovations throughout the world. For example, you don't have to go to Shenzhen or Silicon Valley to initiate a very innovative conversation. You can instantly start those electronically. I find that amazing.

McKinsey: What is your leadership philosophy?

Hidetoshi Shibata: That might be one of the most important questions we've discussed. I'm trying to take a completely different approach from many people. Typically, in a well-run, professional-services organization, you tend to come across very effective "servant leadership." I'm trying to embody servant leadership at Renesas, using my private-equity experiences as a guide.

As needed, I may provide inspiration to employees, or help them with data points and facts, or suggest some alternatives to ensure that everything runs smoothly and efficiently. I think this allows employees to achieve their full potential and growth aspirations, both personally and for the company. Employees can collectively share the experiences of delivering results and meeting and beating targets that they themselves set.

I don't know to what extent this approach is effective and practicable in an industrial manufacturing company with 20,000 employees. I am experimenting now and may change or tweak my philosophy as I make progress.

I also enjoy comparing notes with my peer CEOs, most of whom run larger-scale organizations. I've connected with some of them at McKinsey's T-30 conference, and I've found it interesting to hear what other people are doing.

McKinsey: Is there any advice you would give young professionals beginning a career in the semiconductor industry?

Hidetoshi Shibata: This is not specific to semiconductors, but I would advise them to be as simple and straightforward as possible about their passion or motivation. They don't have to be focused on groundbreaking or world-changing ideas, although that is always an option. Instead, they might concentrate on more incremental growth. Being true to their hearts will help people in many respects. It will encourage them to expand their horizons, raise the bar, and experiment with something completely new, such as meeting with new people in new geographies. Being simple and straightforward is always the best approach in good times and bad, be it in the semiconductor industry or elsewhere.

Hidetoshi Shibata is the representative director, president, and CEO of Renesas Electronics Corporation. This interview was conducted by **André Andonian,** a senior partner in McKinsey's Seoul office, **Abhijit Mahindroo,** a partner in the Southern California office, and **Anupama Suryanarayanan,** a consultant in the Silicon Valley office.

This McKinsey Practice Publication meets the Forest Stewardship Council® (FSC®) chain-of-custody standards. The paper used in this publication is certified as being produced in an environmentally responsible, socially beneficial, and economically viable way.

Printed in the United States of America

October 2021
Designed by Global Editorial Services

McKinsey.com