Moore’s Law postulates that, every two years, technological advancement of integrated circuits (ICs) doubles the number of transistors on a microchip, increasing the resulting computer processor’s computing speed and halving its cost. Scientific consensus predicts that this exponential growth will, eventually, hit a theoretical quantum mechanics barrier and falter. Nevertheless, ongoing research and production breakthroughs continue to bring forth better-than-ever computers and, to date, have not bucked this trend.

State-of-the-art integrated circuit manufacturing has focused on silicon-based semiconductor transistors that, as of 2019, have miniaturized to dimensions below five nanometers. Gallium arsenide, silicon carbide, and gallium nitride transistors, which enable specialized photovoltaic energy harvesting or light-emitting diode capabilities, round out the state-of-the-art semiconductor manufacturing space. Various lithography and crystal-growth techniques are used to fabricate semiconductor discs and pack these transistors in high-density arrangements on these wafers. As the Organisation for Economic Co-operation and Development (OECD) 2019 report highlights, these wafers are 99.9999% pure and 300 mm in diameter. Both numbers are crucial: they facilitate, respectively, reliable performance and low costs that enable everyday modern computing. These metrics stem from decades of successful materials R&D efforts.

But as demands balloon, transistors shrink, and quantum tunneling threatens to break down the performance of nanosized silicon oxide gates in ICs, continued growth demands new approaches and development strategies.

The semiconductor market
According to a September 2020 report by the Information Technology & Innovation Foundation (ITIF), over 1 trillion semiconductor processors were shipped last year; each of those contained as many as 30 billion transistors. This immense computing power enables operation of communications, electronics, and memory systems in every technology sector.

These transistors account for a vast global market. The Semiconductor Industry Association (SIA) 2020 report pegged this industry worldwide at USD$470 billion, and it will reach USD$730 billion in six years. Oxford Economics estimates that this high-technology sector contributes over USD$7 trillion to the worldwide economy. According to a September 2020 US Congressional Research Service (CRS) report, US-based manufacturers capture a 47% share of this market; 70 operating foundries (“fabs”) directly employ over 184,000 workers. However, competition from Japan, South Korea, China, Taiwan, and other countries has driven down the share of US manufactured microchips to only 12%—down from the 1990 level of 37 percent. In fact, a combination of emerging technologies, R&D demands, and supply chain disruptions may shake up this high-technology landscape.
Over the next few years, global efforts to integrate the Internet of Things into much of the existing infrastructure, along with the implementation of 5G wireless communication networks, will drive up demand for manufactured semiconductor devices with superior computing power. In the longer term, the prevalence of artificial intelligence, driverless vehicles, and quantum computing will likely eclipse those needs and require even more numerous microprocessors with overwhelmingly superior capabilities.

A January 2017 US President’s Council of Advisors on Science and Technology report highlighted impacts of potential future breakthroughs in computing calculation algorithms, carbon nanotube and DNA molecule material alternatives to silicon, and fabric-embedded integrated circuits. "Moonshot"-like transformative manufacturing technologies, such as three-dimensional (3D) printing and efficient benchtop-scale fabrication systems, may also be within reach—contingent on successes of multilateral, sustained research efforts.

Adherence to Moore’s Law comes with a hefty price tag; constant innovations in integrated circuit manufacturing require significant R&D investments. According to the SIA report, semiconductor R&D expenses grew at an annual 6.6% rate for the last 20 years. Just last year, US-based manufacturers spent almost USD$40 billion on innovation. R&D expenditures account for over 16% of sales revenues. The ITIF report states that, in order for subsequent generations of microchips to emerge, companies must affirmatively choose to reinvest generated revenues into their technology research pipelines. These mammoth capital investment hurdles allow only a select few companies—five worldwide today (Intel Corporation, Samsung, TSMC, Micron Technology, Inc. and SK Hynix), down from almost 30 two decades ago—to manufacture leading-edge semiconductors that empower computing in top-of-the-line electronics.

US Government departments and agencies, including Departments of Energy (DOE) and Defense (DoD), the National Institute of Standards and Technology, and the National Science Foundation, annually provide USD$1.5 billion of funding for basic research efforts. Recent “grand challenge”-level programs include the DoD Electronics Resurgence Initiative (funded by the Defense Advanced Research Projects Agency), National Strategic Computing Initiative, the PowerAmerica R&D Institute, and the National Quantum Initiative (funded, in part, by the DOE Office of Science).

Nevertheless, most contemporary semiconductor technology advancements, including improved device performance, lowered power consumption, and reduced size, stem from collaborative efforts of multinational research teams. The ITIF report cites achievements of academic and industry groups that included partners from the United States, China, Germany, France, Japan, Taiwan, The Netherlands, South Korea, and the UK, among others. The CRS, ITIF, and SIA reports all reach a similar consensus: international collaboration empowers the semiconductor industry’s attempts to reach its key technology goals. These include the adoption of 450-mm-wide wafers, systems-on-a-chip, lower-power integrated circuits, and groundbreaking memory and computing novel device architectures.

The international nature of semiconductor development advances hand in hand with a globally distributed supply chain. Each step of the process, including ingot manufacturing, wafer fabrication, device patterning, and ultimately microchip assembly typically takes place in a different corner of the world. The ITIF report cited that a typical integrated circuit crosses over 70 international borders on its 100-day-long manufacturing pathway.

The operations and transportation shutdowns of the COVID-19 pandemic caused major disruptions in this multipronged supply chain. According to the November 2020 Cybersecurity and Infrastructure Security Agency (CISA) report, recent events underscored a long-held belief among industry professionals that supply chain diversification into multiple regions is necessary to minimize risks associated with single-location sources. Reliance on lean inventories and poor understanding of operations of subsidiary subcontractors and materials and component suppliers caught semiconductor manufacturers by surprise and exacerbated their production disruptions. From mining to manufacturing to distribution and sales, the pandemic strained every choke point in the supply chain. Although the CISA report recommends larger inventories and dual sourcing from a broader geographical range of manufacturers, it acknowledges economic realities that may pose insurmountable hurdles for this risk mitigation process.

Supply chain disruptions pose a looming challenge to the US economy. Domestic manufacturers, which maintain operations in eight US states, depend on a viable and accessible worldwide market in order to generate the required revenues that sustain profitable operations. In addition to upfront capital costs, ITIF highlights nonmarket-based competition from governments, such as China, that may subsidize costs and unfairly disadvantage enterprises that operate by rules of unrestrained market economies. To maintain the viability of US semiconductor manufacturing and promote onshoring of manufacturing operations, federal and state governments have, in the past, offered tax incentives, grants, low-interest loans, and free land for facilities.

In June of 2020, the US Congress introduced the Creating Helpful Incentives to Produce Semiconductors (CHIPS) for America Act (H.R. 7178; S. 3933). This legislation authorizes US government agencies to fund USD$7 billion of semiconductor R&D. The bill further creates several manufacturing institutes for public–private partnerships and
advanced prototyping and manufacturing development. Furthermore, it provides tax incentives for domestic capital manufacturing costs and establishes a USD$10 billion trust fund that will match state and local incentives for semiconductor manufacturers. A similar bill, the American Foundries Act of 2020, was introduced in July 2020 (S. 4130) in the US Senate. Although Congress has not acted on either bill, the Fiscal Year 2021 National Defense Authorization Act (NDAA) incorporates key provisions from the CHIPS for America Act. The SIA report estimates that these incentives will increase the number of domestic operating foundries by 27 percent. Perhaps most importantly, pending legislation aims to establish an overarching, long-term national semiconductor strategy. Major global players in the integrated circuit market have a clear strategy—except for the United States.

Counterfeit microchips
Renewed policy interest in semiconductors is fueled not only by economics but also by national security interests. Diminishing market shares for domestic US manufacturers, coupled with COVID-19 shutdowns, create supply chain voids that can be exploited by counterfeit microchip manufacturers. This is a deep-rooted problem; the US Department of Commerce Bureau of Industry and Security 2010 report found over 3300 counterfeit integrated circuits in the DoD supply chains in 2005, and over 8000 of these inauthentic products in 2008.

According to a 2015 Defense Technical Information Research & Analysis article, dummy microchips with fabricated labels, scavenged components from scrapped and defective electronics, and unauthorized, low-quality components are supplied as cost-saving measures. More nefariously, if purposefully compromised components are installed in national security platforms, they can leak sensitive information to adversaries or become reprogrammed to malfunction in critical components.

University of Maryland professor Sandor Boyson, who extensively studies national security implications of supply chain breakdowns, says, “The threat to the global IT supply chain is ceaseless. The threat actors we have seen recently range from the crude to the sophisticated. IT companies need to be hypervigilant and establish formal supply chain-wide risk programs, with mitigation playbooks that include a range of actions such as tougher supplier selection and monitoring standards, quarantining code and hardware from overseas suppliers before acceptance, and round-the-clock video surveillance of critical production processes.”

The US Government has launched multiple initiatives to counter this threat of counterfeits in its supply chain. The DoD Trusted Foundry Program ensures that the integrated circuits in the DoD supply pipeline are authentic and reliable. However, this 16-year-old program is still very small and accounts for less than 2% of the integrated circuits acquired by the Department of Defense. Section 224 of the enacted Fiscal Year 2020 NDAA aims to transition the program away from the sole-source foundry approach and more into implementation of strict standards and guarantees on the part of vetted suppliers. DoD is also pursuing a similar Microelectronics Innovation for National Security and Economic Competitiveness program. On the basic research side, DARPA’s Supply Chain Hardware Integrity for Electronics Defense (SHIELD) program, as well as DNA-embedded inks and optical scanning funded by the Defense Logistics Agency and the Army Research Office, are bringing forth approaches to rapidly authenticate microchips.

Emerging legislation, such as the CHIPS for America Act, recognizes the international nature of the semiconductor market. One of the bill’s provisions calls for a USD$750 million trust fund to be allocated for a consortium of democratic foreign governments and is contingent on agreements that bolster transparency of supply chains, coordinated technology development efforts, and harmonized policy agreements. Numerous pending legislative proposals, as well as recent government directives and provisions in the enacted Fiscal Year 2019 NDAA, aim to balance market forces and global cooperation with the need to protect against intellectual property theft and product tampering.

And while subnanometer-sized transistors pose a formidable technology challenge for researchers in foundries, governments around the world—including the United States—grapple with shifting market forces and instability in their supply chains. Comprehensive and timely policies that address these challenges will shape the semiconductor and computing landscape in the decades ahead.

Boris Dyatkin