CHIP WAR

THE FIGHT FOR THE WORLD'S MOST CRITICAL TECHNOLOGY

CHRIS MILLER

Thank you for downloading
this Simon & Schuster ebook.

Get a FREE ebook when you join our mailing list. Plus, get updates on new releases, deals, recommended reads, and more from Simon & Schuster. Click below to sign up and see terms and conditions.

CLICK HERE TO SIGN UP

Already a subscriber? Provide your email again so we can register this ebook and send you more of what you like to read. You will continue to receive exclusive offers in your inbox.

More Praise for **CHIP WAR**

"One of the most important books I've read in years—engrossing, beautifully written. Miller shows that, for all its manifest flaws and failures, the American capitalist system has repeatedly outperformed other systems and in the process has done much to bolster the security of democracy."

> —Robert Kagan, senior fellow, The Brookings Institution, columnist for *The Washington Post*, and author of *The Jungle Grows Back: America and Our Imperiled World*

"If you care about technology, or America's future prosperity, or its continuing security, this is a book you have to read."

> —Lawrence H. Summers, 71st U.S. Secretary of the Treasury and Charles W. Eliot University Professor at Harvard University

"Outstanding. Miller's history of the chip covers all angles: technological, financial, and especially political.... The go-to reference on one of the most important industries today."

> —Dan Wang, technology analyst at Gavekal Dragonomics

"The battle for supremacy in semiconductors is one of the most important stories in geopolitics, national security, and economic prosperity. But it's also been one of the least well understood. Thankfully, we now have *Chip War* to give us a clear view and sharp read on this essential subject."

>—Andrew McAfee, coauthor of *The Second Machine Age* and author of *The Geek Way* and *More from Less*

CHIP WAR

THE FIGHT FOR THE WORLD'S MOST CRITICAL TECHNOLOGY

CHRIS MILLER

To Liya

Cast of Characters

Morris Chang: Founder of Taiwan Semiconductor Manufacturing Company (TSMC), the world's most important chipmaker; previously, a senior executive at Texas Instruments.

Andy Grove: Former president and CEO of Intel during the 1980s and 1990s; notorious for his aggressive style and success in reviving Intel; author of *Only the Paranoid Survive*.

Pat Haggerty: Chairman of Texas Instruments; led the company as it specialized in building microelectronics, including for the U.S. military.

Jack Kilby: Co-inventor of the integrated circuit, in 1958; longtime Texas Instruments employee; winner of the Nobel Prize.

Jay Lathrop: Co-inventor of photolithography, the process of patterning transistors using specialized chemicals and light; formerly of Texas Instruments.

Carver Mead: Professor at the California Institute of Technology (Caltech); advisor to Fairchild Semiconductor and Intel; visionary thinker about the future of technology.

Gordon Moore: Cofounder of Fairchild Semiconductor and Intel; creator in 1965 of "Moore's Law," which predicted that the computing power on each chip would double every couple of years.

Akio Morita: Cofounder of Sony; coauthor of *The Japan That Can Say No*; represented Japanese business on the world stage during the 1970s and 1980s.

Robert Noyce: Cofounder of Fairchild Semiconductor and Intel; co-inventor of the integrated circuit in 1959; known as the "Mayor of Silicon Valley"; first leader of Sematech.

William Perry: Pentagon official from 1977–1981 and later Secretary of Defense from 1994 to 1997 who advocated using chips to produce precision-strike weapons.

Jerry Sanders: Founder and CEO of AMD; Silicon Valley's most flamboyant salesman; an aggressive critic of what he saw as unfair Japanese trade practices in the 1980s.

Charlie Sporck: Drove the offshoring of chip assembly while leading manufacturing operations at Fairchild Semiconductor;

later CEO of National Semiconductor.

Ren Zhengfei: Founder of Huawei, China's telecom and chip-design giant; his daughter Meng Wanzhou was arrested in Canada in 2018 on charges of violating U.S. law and trying to evade U.S. sanctions.

Glossary

Arm:

a UK company that licenses to chip designers use of an instruction set architecture—a set of basic rules governing how a given chip operates. The Arm architecture is dominant in mobile devices and is slowly winning market share in PCs and data centers.

Chip (also "integrated circuit" or "semiconductor"):

a small piece of semiconducting material, usually silicon, with millions or billions of microscopic transistors carved into it.

CPU:

central processing unit; a type of "general-purpose" chip that is the workhorse of computing in PCs, phones, and data centers.

DRAM:

dynamic random access memory; one of two main types of memory chip, which is used to store data temporarily.

EDA:

electronic design automation; specialized software used to design how millions or billions of transistors will be arrayed on a chip and to simulate their operation.

FinFET:

a new 3D transistor structure first implemented in the early 2010s to better control transistor operation as transistors' size shrank to nanometric-scale.

GPU:

graphics processing unit; a chip that is capable of parallel processing, making it useful for graphics and for artificial intelligence applications.

Logic chip:

a chip that processes data.

Memory chip:

a chip that remembers data.

NAND:

also called "flash," the second major type of memory chip, used for longer-term data storage.

Photolithography:

also known as "lithography"; the process of shining light or ultraviolet light through patterned masks: the light then interacts with photoresist chemicals to carve patterns on silicon wafers.

RISC–V:

an open-source architecture growing in popularity because it is free to use, unlike Arm and x86. The development of RISC-V was partially funded by the U.S. government but now is popular in China because it is not subject to U.S. export controls.

Silicon wafer:

a circular piece of ultra-pure silicon, usually eight or twelve inches in diameter, out of which chips are carved.

Transistor:

a tiny electric "switch" that turns on (creating a 1) or off (0), producing the 1s and 0s that undergird all digital computing.

x86:

an instruction set architecture that is dominant in PCs and data centers. Intel and AMD are the two main firms producing such chips.

Introduction

The destroyer USS *Mustin* slipped into the northern end of the Taiwan Strait on August 18, 2020, its five-inch gun pointed southward as it began a solo mission to sail through the Strait and reaffirm that these international waters were *not* controlled by China—at least not yet. A stiff southwestern breeze whipped across the deck as it steamed south. High clouds cast shadows on the water that seemed to stretch all the way to the great port cities of Fuzhou, Xiamen, Hong Kong, and the other harbors that dot the South China coast. To the east, the island of Taiwan rose in the distance, a broad, densely settled coastal plain giving way to tall peaks hidden in clouds. Aboard ship, a sailor wearing a navy baseball cap and a surgical mask lifted his binoculars and scanned the horizon. The waters were filled with commercial freighters shipping goods from Asia's factories to consumers around the world.

On board the USS *Mustin*, a row of sailors sat in a dark room in front of an array of brightly colored screens on which were displayed data from planes, drones, ships, and satellites tracking

movement across the Indo-Pacific. Atop the *Mustin*'s bridge, a radar array fed into the ship's computers. On deck ninety-six launch cells stood ready, each capable of firing missiles that could precisely strike planes, ships, or submarines dozens or even hundreds of miles away. During the crises of the Cold War, the U.S. military had used threats of brute nuclear force to defend Taiwan. Today, it relies on microelectronics and precision strikes.

As the USS *Mustin* sailed through the Strait, bristling with computerized weaponry, the People's Liberation Army announced a retaliatory series of live-fire exercises around Taiwan, practicing what one Beijing-controlled newspaper called a "reunification-by-force operation." But on this particular day, China's leaders worried less about the U.S. Navy and more about an obscure U.S. Commerce Department regulation called the Entity List, which limits the transfer of American technology abroad. Previously, the Entity List had primarily been used to prevent sales of military systems like missile parts or nuclear materials. Now, though, the U.S. government was dramatically tightening the rules governing computer chips, which had become ubiquitous in both military systems and consumer goods.

The target was Huawei, China's tech giant, which sells smartphones, telecom equipment, cloud computing services, and other advanced technologies. The U.S. feared that Huawei's products were now priced so attractively, partly owing to Chinese government subsidies, that they'd shortly form the backbone of

next-generation telecom networks. America's dominance of the world's tech infrastructure would be undermined. China's geopolitical clout would grow. To counter this threat, the U.S. barred Huawei from buying advanced computer chips made with U.S. technology.

Soon, the company's global expansion ground to a halt. Entire product lines became impossible to produce. Revenue slumped. A corporate giant faced technological asphyxiation. Huawei discovered that, like all other Chinese companies, it was fatally dependent on foreigners to make the chips upon which all modern electronics depend.

The United States *still* has a stranglehold on the silicon chips that gave Silicon Valley its name, though its position has weakened dangerously. China now spends more money each year importing chips than it spends on oil. These semiconductors are plugged into all manner of devices, from smartphones to refrigerators, that China consumes at home or exports worldwide. Armchair strategists theorize about China's "Malacca Dilemma"— a reference to the main shipping channel between the Pacific and Indian Oceans—and the country's ability to access supplies of oil and other commodities amid a crisis. Beijing, however, is more worried about a blockade measured in bytes rather than barrels. China is devoting its best minds and billions of dollars to developing its own semiconductor technology in a bid to free itself from America's chip choke.

If Beijing succeeds, it will remake the global economy and reset the balance of military power. World War II was decided by steel and aluminum, and followed shortly thereafter by the Cold War, which was defined by atomic weapons. The rivalry between the United States and China may well be determined by computing power. Strategists in Beijing and Washington now realize that all advanced tech—from machine learning to missile systems, from automated vehicles to armed drones—requires cutting-edge chips, known more formally as semiconductors or integrated circuits. A tiny number of companies control their production.

We rarely think about chips, yet they've created the modern world. The fate of nations has turned on their ability to harness computing power. Globalization as we know it wouldn't exist without the trade in semiconductors and the electronic products they make possible. America's military primacy stems largely from its ability to apply chips to military uses. Asia's tremendous rise over the past half century has been built on a foundation of silicon as its growing economies have come to specialize in fabricating chips and assembling the computers and smartphones that these integrated circuits make possible.

At the core of computing is the need for many millions of 1s and 0s. The entire digital universe consists of these two numbers. Every button on your iPhone, every email, photograph, and YouTube video—all of these are coded, ultimately, in vast strings of 1s and 0s. But these numbers don't actually exist. They're

expressions of electrical currents, which are either on (1) or off (0). A chip is a grid of millions or billions of *transistors*, tiny electrical switches that flip on and off to process these digits, to remember them, and to convert real world sensations like images, sound, and radio waves into millions and millions of 1s and 0s.

As the USS *Mustin* sailed southward, factories and assembly facilities on both sides of the Strait were churning out components for the iPhone 12, which was only two months away from its October 2020 launch. Around a quarter of the chip industry's revenue comes from phones; much of the price of a new phone pays for the semiconductors inside. For the past decade, each generation of iPhone has been powered by one of the world's most advanced processor chips. In total, it takes over a dozen semiconductors to make a smartphone work, with different chips managing the battery, Bluetooth, Wi-Fi, cellular network connections, audio, the camera, and more.

Apple makes precisely *none* of these chips. It buys most off-the-shelf: memory chips from Japan's Kioxia, radio frequency chips from California's Skyworks, audio chips from Cirrus Logic, based in Austin, Texas. Apple designs in-house the ultra-complex processors that run an iPhone's operating system. But the Cupertino, California, colossus can't manufacture these chips. Nor can any company in the United States, Europe, Japan, or China. Today, Apple's most advanced processors—which are arguably the world's most advanced semiconductors—can only be produced by

a single company in a single building, the most expensive factory in human history, which on the morning of August 18, 2020, was only a couple dozen miles off the USS *Mustin*'s starboard bow.

Fabricating and miniaturizing semiconductors has been the greatest engineering challenge of our time. Today, no firm fabricates chips with more precision than the Taiwan Semiconductor Manufacturing Company, better known as TSMC. In 2020, as the world lurched between lockdowns driven by a virus whose diameter measured around one hundred nanometers —billionths of a meter—TSMC's most advanced facility, Fab 18, was carving microscopic mazes of tiny transistors, etching shapes smaller than half the size of a coronavirus, a hundredth the size of a mitochondria. TSMC replicated this process at a scale previously unparalleled in human history. Apple sold over 100 million iPhone 12s, each powered by an A14 processor chip with 11.8 billion tiny transistors carved into its silicon. In a matter of months, in other words, for just one of the dozen chips in an iPhone, TSMC's Fab 18 fabricated well over 1 quintillion transistors—that is, a number with eighteen zeros behind it. Last year, the chip industry produced more transistors than the combined quantity of all goods produced by all other companies, in all other industries, in all human history. Nothing else comes close.

It was only sixty years ago that the number of transistors on a cutting-edge chip wasn't 11.8 billion, but 4. In 1961, south of San

Francisco, a small firm called Fairchild Semiconductor announced a new product called the Micrologic, a silicon chip with four transistors embedded in it. Soon the company devised ways to put a dozen transistors on a chip, then a hundred. Fairchild cofounder Gordon Moore noticed in 1965 that the number of components that could be fit on each chip was doubling annually as engineers learned to fabricate ever smaller transistors. This prediction—that the computing power of chips would grow exponentially—came to be called "Moore's Law" and led Moore to predict the invention of devices that in 1965 seemed impossibly futuristic, like an "electronic wristwatch," "home computers," and even "personal portable communications equipment." Looking forward from 1965, Moore predicted a decade of exponential growth—but this staggering rate of progress has continued for over half a century. In 1970, the second company Moore founded, Intel, unveiled a memory chip that could remember 1,024 pieces of information ("bits"). It cost around $20, roughly two cents per bit. Today, $20 can buy a thumb drive that can remember well over a billion bits.

When we think of Silicon Valley today, our minds conjure social networks and software companies rather than the material after which the valley was named. Yet the internet, the cloud, social media, and the entire digital world only exist because engineers have learned to control the most minute movement of electrons as they race across slabs of silicon. "Big tech" wouldn't

exist if the cost of processing and remembering 1s and 0s hadn't fallen by a billionfold in the past half century.

This incredible ascent is partly thanks to brilliant scientists and Nobel Prize–winning physicists. But not every invention creates a successful startup, and not every startup sparks a new industry that transforms the world. Semiconductors spread across society because companies devised new techniques to manufacture them by the millions, because hard-charging managers relentlessly drove down their cost, and because creative entrepreneurs imagined new ways to use them. The making of Moore's Law is as much a story of manufacturing experts, supply chain specialists, and marketing managers as it is about physicists or electrical engineers.

The towns to the south of San Francisco—which weren't called Silicon Valley until the 1970s—were the epicenter of this revolution because they combined scientific expertise, manufacturing know-how, and visionary business thinking. California had plenty of engineers trained in aviation or radio industries who'd graduated from Stanford or Berkeley, each of which was flush with defense dollars as the U.S. military sought to solidify its technological advantage. California's culture mattered just as much as any economic structure, however. The people who left America's East Coast, Europe, and Asia to build the chip industry often cited a sense of boundless opportunity in their decision to move to Silicon Valley. For the world's smartest

engineers and most creative entrepreneurs, there was simply no more exciting place to be.

Once the chip industry took shape, it proved impossible to dislodge from Silicon Valley. Today's semiconductor supply chain requires components from many cities and countries, but almost every chip made still has a Silicon Valley connection or is produced with tools designed and built in California. America's vast reserve of scientific expertise, nurtured by government research funding and strengthened by the ability to poach the best scientists from other countries, has provided the core knowledge driving technological advances forward. The country's network of venture capital firms and its stock markets have provided the startup capital new firms need to grow—and have ruthlessly forced out failing companies. Meanwhile, the world's largest consumer market in the U.S. has driven the growth that's funded decades of R&D on new types of chips.

Other countries have found it impossible to keep up on their own but have succeeded when they've deeply integrated themselves into Silicon Valley's supply chains. Europe has isolated islands of semiconductor expertise, notably in producing the machine tools needed to make chips and in designing chip architectures. Asian governments, in Taiwan, South Korea, and Japan, have elbowed their way into the chip industry by subsidizing firms, funding training programs, keeping their exchange rates undervalued, and imposing tariffs on imported

chips. This strategy has yielded certain capabilities that no other countries can replicate—but they've achieved what they have in partnership with Silicon Valley, continuing to rely fundamentally on U.S. tools, software, and customers. Meanwhile, America's most successful chip firms have built supply chains that stretch across the world, driving down costs and producing the expertise that has made Moore's Law possible.

Today, thanks to Moore's Law, semiconductors are embedded in every device that requires computing power—and in the age of the Internet of Things, this means pretty much *every* device. Even hundred-year-old products like automobiles now often include a thousand dollars worth of chips. Most of the world's GDP is produced with devices that rely on semiconductors. For a product that didn't exist seventy-five years ago, this is an extraordinary ascent.

As the USS *Mustin* steamed southward in August 2020, the world was just beginning to reckon with our reliance on semiconductors—and our dependence on Taiwan, which fabricates the chips that produce a third of the new computing power we use each year. Taiwan's TSMC builds almost all the world's most advanced processor chips. When COVID slammed into the world in 2020, it disrupted the chip industry, too. Some factories were temporarily shuttered. Purchases of chips for autos slumped. Demand for PC and data center chips spiked higher, as much of the world prepared to work from home. Then, over 2021,

a series of accidents—a fire in a Japanese semiconductor facility; ice storms in Texas, a center of U.S. chipmaking; and a new round of COVID lockdowns in Malaysia, where many chips are assembled and tested—intensified these disruptions. Suddenly, many industries far from Silicon Valley faced debilitating chip shortages. Big carmakers from Toyota to General Motors had to shut factories for weeks because they couldn't acquire the semiconductors they needed. Shortages of even the simplest chips caused factory closures on the opposite side of the world. It seemed like a perfect image of globalization gone wrong.

Political leaders in the U.S., Europe, and Japan hadn't thought much about semiconductors in decades. Like the rest of us, they thought "tech" meant search engines or social media, not silicon wafers. When Joe Biden and Angela Merkel asked why their country's car factories were shuttered, the answer was shrouded behind semiconductor supply chains of bewildering complexity. A typical chip might be designed with blueprints from the Japanese-owned, UK-based company called Arm, by a team of engineers in California and Israel, using design software from the United States. When a design is complete, it's sent to a facility in Taiwan, which buys ultra-pure silicon wafers and specialized gases from Japan. The design is carved into silicon using some of the world's most precise machinery, which can etch, deposit, and measure layers of materials a few atoms thick. These tools are produced primarily by five companies, one Dutch, one Japanese,

and three Californian, without which advanced chips are basically impossible to make. Then the chip is packaged and tested, often in Southeast Asia, before being sent to China for assembly into a phone or computer.

If any one of the steps in the semiconductor production process is interrupted, the world's supply of new computing power is imperiled. In the age of AI, it's often said that data is the new oil. Yet the real limitation we face isn't the availability of data but of processing power. There's a finite number of semiconductors that can store and process data. Producing them is mind-bogglingly complex and horrendously expensive. Unlike oil, which can be bought from many countries, our production of computing power depends fundamentally on a series of choke points: tools, chemicals, and software that often are produced by a handful of companies—and sometimes only by one. No other facet of the economy is so dependent on so few firms. Chips from Taiwan provide 37 percent of the world's new computing power each year. Two Korean companies produce 44 percent of the world's memory chips. The Dutch company ASML builds 100 percent of the world's extreme ultraviolet lithography machines, without which cutting-edge chips are simply impossible to make. OPEC's 40 percent share of world oil production looks unimpressive by comparison.

The global network of companies that annually produces a trillion chips at nanometer scale is a triumph of efficiency. It's

also a staggering vulnerability. The disruptions of the pandemic provide just a glimpse of what a single well-placed earthquake could do to the global economy. Taiwan sits atop a fault line that as recently as 1999 produced an earthquake measuring 7.3 on the Richter scale. Thankfully, this only knocked chip production offline for a couple of days. But it's only a matter of time before a stronger quake strikes Taiwan. A devastating quake could also hit Japan, an earthquake-prone country that produces 17 percent of the world's chips, or Silicon Valley, which today produces few chips but builds crucial chipmaking machinery in facilities sitting atop the San Andreas Fault.

Yet the seismic shift that most imperils semiconductor supply today isn't the crash of tectonic plates but the clash of great powers. As China and the United States struggle for supremacy, both Washington and Beijing are fixated on controlling the future of computing—and, to a frightening degree, that future is dependent on a small island that Beijing considers a renegade province and America has committed to defend by force.

The interconnections between the chip industries in the U.S., China, and Taiwan are dizzyingly complex. There's no better illustration of this than the individual who founded TSMC, a company that until 2020 counted America's Apple and China's Huawei as its two biggest customers. Morris Chang was born in mainland China; grew up in World War II–era Hong Kong; was educated at Harvard, MIT, and Stanford; helped build America's

early chip industry while working for Texas Instruments in Dallas; held a top secret U.S. security clearance to develop electronics for the American military; and made Taiwan the epicenter of world semiconductor manufacturing. Some foreign policy strategists in Beijing and Washington dream of decoupling the two countries' tech sectors, but the ultra-efficient international network of chip designers, chemical suppliers, and machine-tool makers that people like Chang helped build can't be easily unwound.

Unless, of course, something explodes. Beijing has pointedly refused to rule out the prospect that it might invade Taiwan to "reunify" it with the mainland. But it wouldn't take anything as dramatic as an amphibious assault to send semiconductor-induced shock waves careening through the global economy. Even a partial blockade by Chinese forces would trigger devastating disruptions. A single missile strike on TSMC's most advanced chip fabrication facility could easily cause hundreds of billions of dollars of damage once delays to the production of phones, data centers, autos, telecom networks, and other technology are added up.

Holding the global economy hostage to one of the world's most dangerous political disputes might seem like an error of historic proportions. However, the concentration of advanced chip manufacturing in Taiwan, South Korea, and elsewhere in East Asia isn't an accident. A series of deliberate decisions by government officials and corporate executives created the far-flung supply

chains we rely on today. Asia's vast pool of cheap labor attracted chipmakers looking for low-cost factory workers. The region's governments and corporations used offshored chip assembly facilities to learn about, and eventually domesticate, more advanced technologies. Washington's foreign policy strategists embraced complex semiconductor supply chains as a tool to bind Asia to an American-led world. Capitalism's inexorable demand for economic efficiency drove a constant push for cost cuts and corporate consolidation. The steady tempo of technological innovation that underwrote Moore's Law required ever more complex materials, machinery, and processes that could only be supplied or funded via global markets. And our gargantuan demand for computing power only continues to grow.

Drawing on research in historical archives on three continents, from Taipei to Moscow, and over a hundred interviews with scientists, engineers, CEOs, and government officials, this book contends that semiconductors have defined the world we live in, determining the shape of international politics, the structure of the world economy, and the balance of military power. Yet this most modern of devices has a complex and contested history. Its development has been shaped not only by corporations and consumers but also by ambitious governments and the imperatives of war. To understand how our world came to be defined by quintillions of transistors and a tiny number of

irreplaceable companies, we must begin by looking back to the origins of the silicon age.