

ClearBridge

A Franklin Templeton Company

Can AI Capex Extend the Semiconductor Cycle?

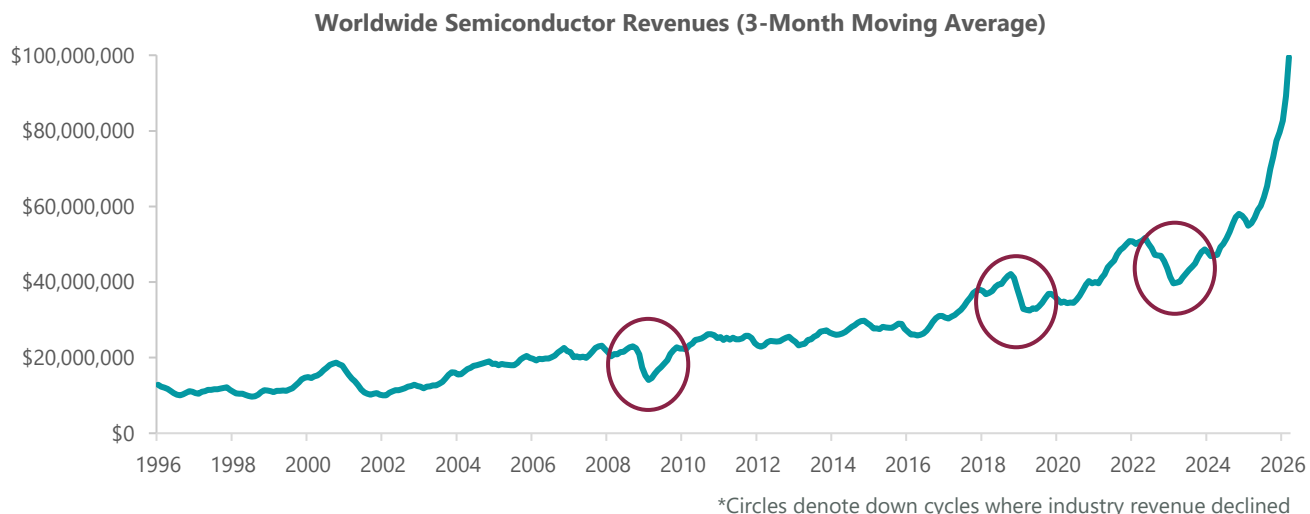
June 1, 2026

Key Takeaways

- ▶ Specialized memory chips and traditional CPUs have joined GPUs as key enablers of AI workloads and agents. In addition to a greater range of chips supporting AI development, several factors could cause the current cycle to last longer than expected.
- ▶ Durable demand for chips to run inference workloads has expanded the supplier base to include companies developing application-specific integrated circuits as well as hyperscalers producing silicon in-house.
- ▶ We see the availability of server power components as the next gating factor in building out AI capacity, highlighting the importance of analog companies manufacturing power management chips.
- ▶ Memory remains a key source of upside, with tight high-bandwidth memory and DRAM supply likely to underpin stronger pricing and earnings through 2027.

Accelerating capital spending on AI buildouts by mega cap hyperscalers and emerging AI model developers continues to surpass expectations. The positive trajectory of capex has supported equity performance across the semiconductor industry, with memory players seeing particularly strong gains. Just as skepticism has emerged over the potential return on investment from an unprecedented period of capex, investors have also begun to raise concerns over the duration of the current semiconductor cycle (Exhibit 1).

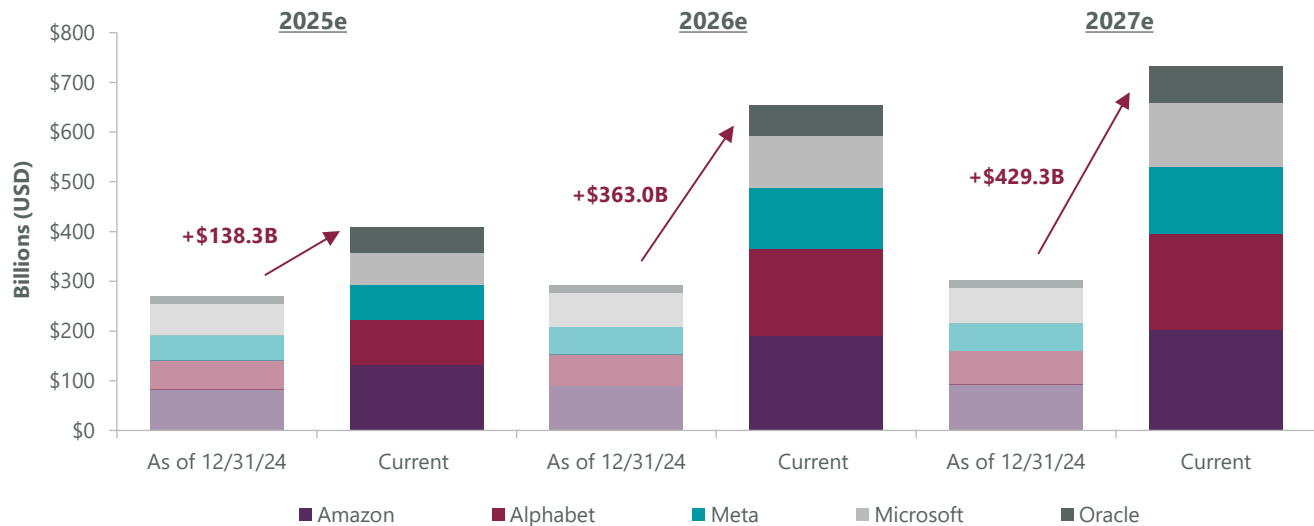
Exhibit 1: Semiconductor Revenue Highly Cyclical



Data as of March 31, 2026. Source: WSTS.

The semiconductor industry has historically proven volatile through waves of technology innovation where chip demand has increased faster than available supply, yet we believe the unique and expansive nature of generative AI could elongate and reduce the severity of the traditional boom/bust cycle (Exhibit 2).

Exhibit 2: Hyperscaler AI Capex Keeps Surpassing Expectations



Source: The term “consensus” for the Hyperscaler capex estimates refers to the average of capex estimates made by sell-side analysts. Data as of March 31, 2026.
Source: FactSet.

Overall chip demand continues to inflect much faster than supply as the generative AI revolution has transformed what the installed enterprise technology base needs to be. Adding supply, especially for high-end applications, can be an especially lengthy process. For example, an extreme ultraviolet lithography machine that etches intricate circuit patterns required to produce leading-edge chips can take six to 12 months to manufacture and ship to a customer. Indeed, supply is constrained by both physics as well as the production capacity of semiconductor capital equipment makers like ASML and leading foundries such as Taiwan Semiconductor Manufacturing, especially when building the foundation for an all-new type of compute.

While graphics processing units (GPUs) have been the workhorse of initial generative AI development, specialized high-bandwidth memory (HBM) chips as well as traditional central processing units (CPUs) are playing increasing roles as more companies develop cloud capabilities to host AI workloads and offer AI agents. In addition to a greater range of chips supporting AI development, we believe three factors could cause the current cycle to last longer than expected.

The first is the durability of inference silicon demand. In addition to market leader Nvidia, companies such as Broadcom have seen strong uptake for their application specific integrated circuits (ASICs) designed for inference functions — the outputs produced by commercial large language models. More recently, Advanced Micro Devices, Taiwan’s MediaTek and newly public AI chip designer Cerebras have reported growing demand from model developers for silicon to support not only processing but also networking and connectivity. Meanwhile, to better control supply in a constrained market for chips to support AI workloads and to customize silicon for their own needs, hyperscalers Google and Amazon are also producing semiconductors in-house (TPU and Trainium, respectively).

The second trend supporting the semiconductor cycle is the increasing power required to efficiently and effectively operate server racks in data centers. We see the availability of server power components as the next gating factor in building out further AI capacity. Analog semiconductor makers such as Texas Instruments and Monolithic Power produce power management chips that deliver and regulate the precise voltage requirements for servers based on workload intensity. Tightening capacity is causing customers to pay higher prices for

analog power chips — from 5% to 20% higher — to ensure available supply. We believe the diversity of customers in this market, a function of more chip makers adding value across the AI supply chain, should help smooth out previous volatility among analog stocks. Notably, the AI tailwind from power management is occurring in conjunction with a broader cyclical recovery in the analog segment.

Memory Still a Long Way from Equilibrium

The third driver of visible semiconductor demand is the memory market. Surging demand for DRAM has created a super cycle that we see lasting through 2026 and into 2027 as model developers scoop up HBM, which are a specialized form of DRAM. These products stack DRAM chips together to enable parallel processing. The HBM market is consolidated, with Micron Technology, Samsung Electronics and SK Hynix the primary suppliers. DRAM prices are expected to increase for the rest of 2026 and flow directly into supplier earnings. We believe consensus forecasts still underestimate the pricing power these companies maintain.

Not nearly enough DRAM supply has been built in the current cycle. This is due to HBM consuming three to four times more silicon wafers than traditional DRAM. In addition, these top three HBM suppliers currently do not have enough clean room space to build for the type of demand we see. The six to nine months it takes from equipment installation to product distribution is one driver of an extended memory up cycle. What makes this upcycle different than previous ones is the accelerated growth of HBM, which has higher margins than other forms of memory — those margins, based on fixed pricing set prior to the supply crunch, are currently below commodity DRAM margins due to the fluid nature of commodity pricing that is being squeezed by severe supply constraints — and is sold under longer-term agreements. HBM supply does not go as far as commodity DRAM either due to the need for more wafers to perform its processes, which has also caused the margins on commodity DRAM to improve dramatically as demand exceeds slower supply adds across memory types. The inherently wafer-intensive HBM demand surge has lowered the risk of an oversupply situation when demand cools.

In a period where generative AI compute capacity is so constrained, primarily due to supply shortages, and capex continues to move up and to the right, the tide has the potential to lift all boats in AI-related semiconductor segments. We expect leading edge capacity to remain tight well into 2027, and possibly into 2028. As supply and demand eventually reach equilibrium and the current constraints are released, however, we expect the best operators in compute (GPU and CPU) and analog to stand out as durable growth winners, with the possibility of the memory cycle remaining stronger for longer due to the unique characteristics and increasing revenue mix of HBM.

About the Author



Dalya Hahn

- Vice President, Research Analyst for Semiconductors and Hardware
- 8 years of investment industry experience
- Joined the firm in 2025
- B.A. in neuroscience from Princeton University

ClearBridge Investments

One Madison Avenue, New York, NY 10010 | 800 691 6960 | [ClearBridge.com](https://www.clearbridge.com)

Past performance is no guarantee of future results. Copyright © 2026 ClearBridge Investments. All opinions and data included in this document are as of the publication date and are subject to change. The opinions and views expressed herein are of the author(s) and may differ from other managers, or the firm as a whole, and are not intended to be a forecast of future events, a guarantee of future results or investment advice. This information should not be used as the sole basis to make any investment decision. The statistics have been obtained from sources believed to be reliable, but the accuracy and completeness of this information cannot be guaranteed. Neither ClearBridge Investments, LLC nor its information providers are responsible for any damages or losses arising from any use of this information.

