23June 2026

Musha Research Commentary

Strategy Bulletin (No. 402)

The AI Price Revolution: The Emergence of an Astonishingly High-Profit Economic Zone

～The AI Market Is Far From a Bubble, We're Still Just at the Beginning～

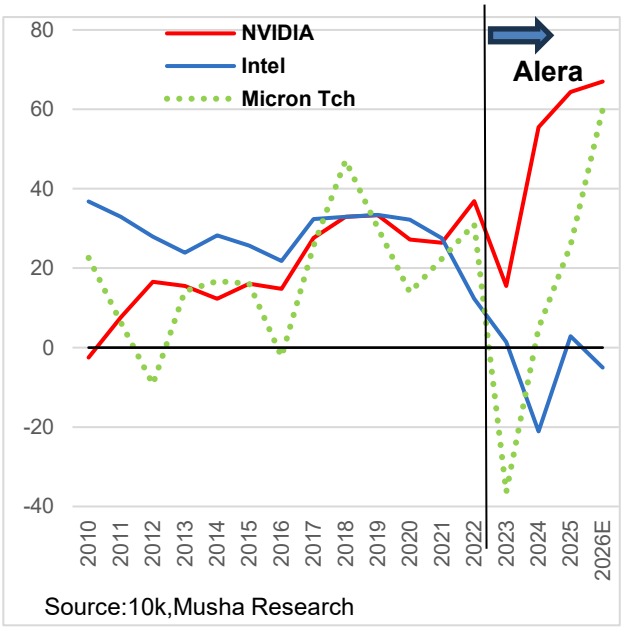
(1) The Current Upheaval: An Anomaly or a Historic Event?

Events that, by conventional wisdom, should not be happening are occurring with increasing frequency. Is this upheaval an anomaly or a passing phenomenon, or is it a structural or historic shift? The first example is the NVIDIA phenomenon. Until ten years ago, NVIDIA was viewed as a second-tier semiconductor company specializing in GPUs for gaming. Yet it has surged not only to the top of the semiconductor industry but has become the world's largest company by market capitalization. NVIDIA has completely rewritten the conventional wisdom of the semiconductor industry: 1) Cyclicality has disappeared from the semiconductor industry; 2) The norm has shifted from "price declines are the norm" to "price increases are the norm"; and 3) It has solidified its dominance through software rather than hardware. Since selling prices are rising in a semiconductor market where costs are constantly falling, profits are surging. Operating profit margins have soared from around 15% a decade ago to 60–70%. By investing these excess profits into building a software and AI ecosystem, NVIDIA has secured its position as a central player in the development of AI infrastructure. What is surprising is that it is not just NVIDIA; a steady stream of companies is emerging that are achieving dramatic profit growth due to rising prices. The trend of rising prices is spreading across a wide range of products, from semiconductor memory to electronic components and cables, as seen in companies such as Samsung, SK Hynix, Micron Technology, and Kioxia.

Musha Research Co., Ltd.
President
Ryoji Musha
musha@musha.co.jp
http://www.musha.co.jp

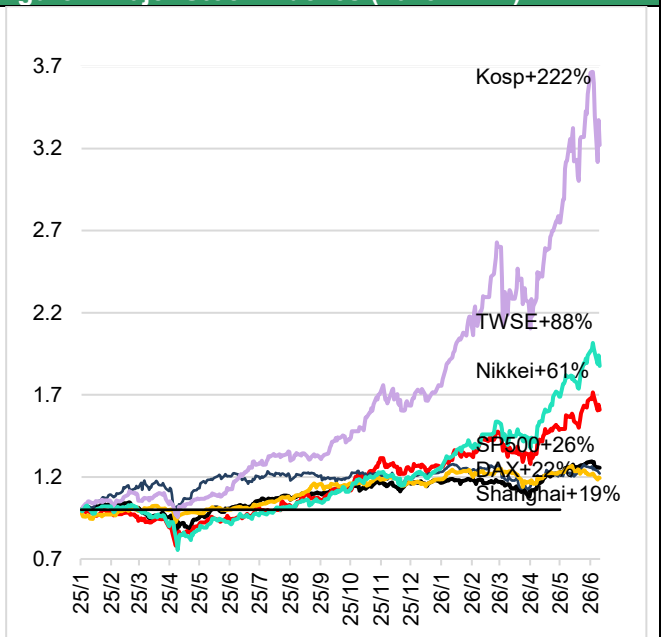
Tel ＋81-3- 3740 2143
2-16-7-3912, Konan,
Minato-ku, Tokyo
108-0075, Japan

Figure1:OP Margin of Semiconductor Companies .



Source:10k,Musha Research

Figure2: Majer Stock Indexes (2025.1.1=1)

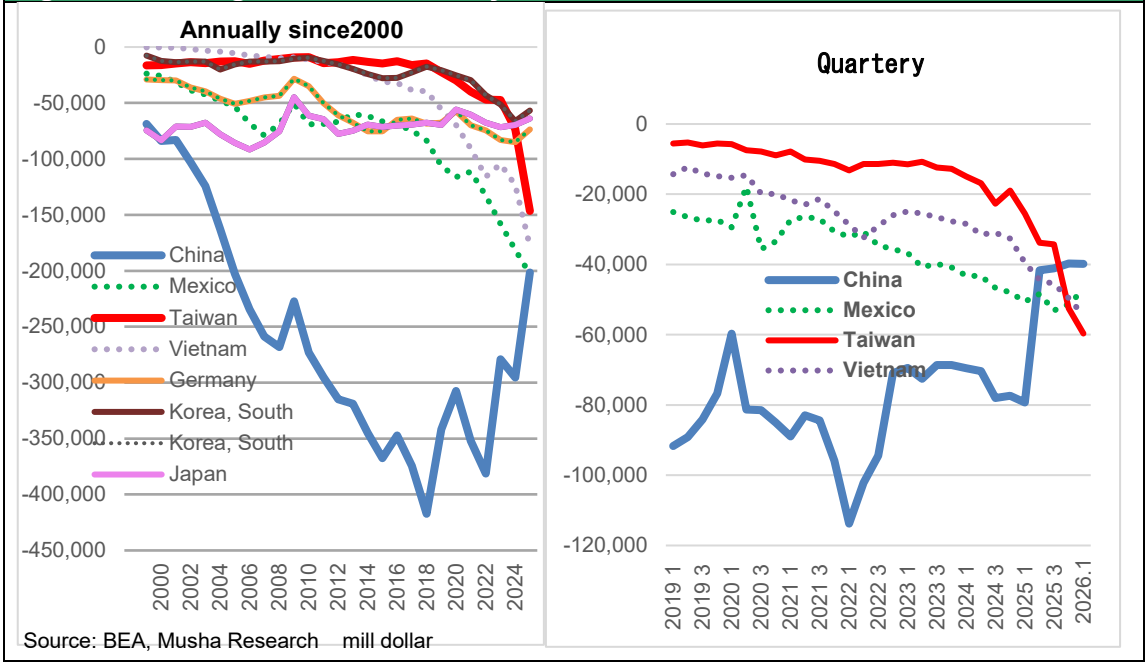


Source: Bloomberg, Musha Research

The second significant shift is the change in the stock market, specifically the exceptional performance of AI-related stocks. Globally, stock price increases in South Korea and Taiwan stand out. In Japan, two of the top three companies by market capitalization are AI-related (SBG and Kioxia), and SoftBank has emerged as Japan's largest company.

Third, these changes have begun to manifest in trade patterns. The United States' largest trade deficit partner has shifted from China to Taiwan. From January to March 2026, the U.S. trade deficit by country was: 1) Taiwan, $59.7 billion; 2) Vietnam, $54.5 billion; 3) Mexico, $47.2 billion; and 4) China, $39.8 billion (or $24.7 billion if Hong Kong is included). Taiwan, which is the sole supplier of U.S. AI-related hardware, has surpassed China—a country that has long boasted industrial clusters based on overwhelmingly low unit labor costs (ULC).

Figure 3: US Largest Trade Deficit by Countries



Source: BEA, Musha Research mill dollar

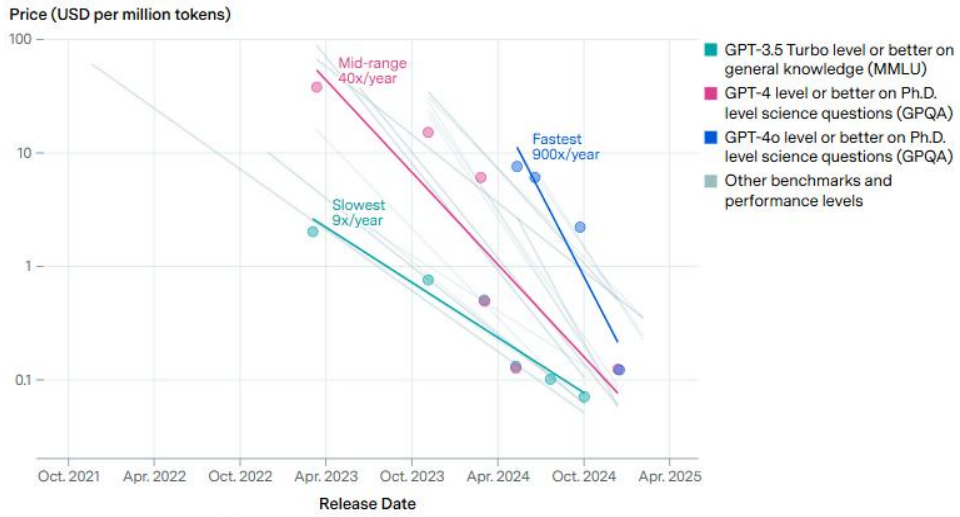
(2) The Ongoing Dramatic Productivity Revolution Driven by AI

At the root of this shift are dramatic productivity gains and cost reductions in AI. The astonishing advancement of the information society over the past 50 years, including the convergence of communications and computing and the emergence of the Internet—was driven almost entirely by Moore's Law (the exponential growth of semiconductor technology and the exponential decline in costs). This represents an astonishing technological evolution in which the density of transistors on a chip doubles (and costs are roughly halved) every 18 to 24 months. At a time when concerns were growing that Moore's Law was entering a mature phase and slowing down, AI emerged. Technological advances and cost reductions driven by AI are demonstrating performance improvements that far exceed those of Moore's Law.

Technological progress in AI is generally referred to as the "scaling law." It is said that the scale of input determines the progress of AI. These inputs consist of three elements: (1) the volume of training data, (2) the size of the model (i.e., the number of parameters), and (3) computational resources (the computing power of a data center or the number of NVIDIA AI chips installed). The theory holds that as each of these is scaled up, AI technological progress will continue. The speed of this progress is astonishing.
Epoch AI, a U.S.-based research organization, reports that the computational cost required for AI to generate text and images—measured per token (the smallest unit of question and response)—continues to decline at a rate of one-ninth per year for the smallest models and one-nine-hundredth per year for the largest models. Since the cost reduction under Moore's Law is 30% per year (approximately 1/1.4), this means that AI costs are falling at a rate 6 to 600 times faster.

While it is difficult to verify how accurate and sustainable these estimates are, at this point, the overwhelming disparity between Moore's Law and the scaling laws of AI is hard to deny. In fact, Anthropic—a key player in AI development—is calling on AI research institutions and others to consider slowing down the pace of development. The company argues that AI systems are advancing so rapidly that they may soon be able to self-improve without human intervention, posing a serious risk to society. (WSJ, April 26)

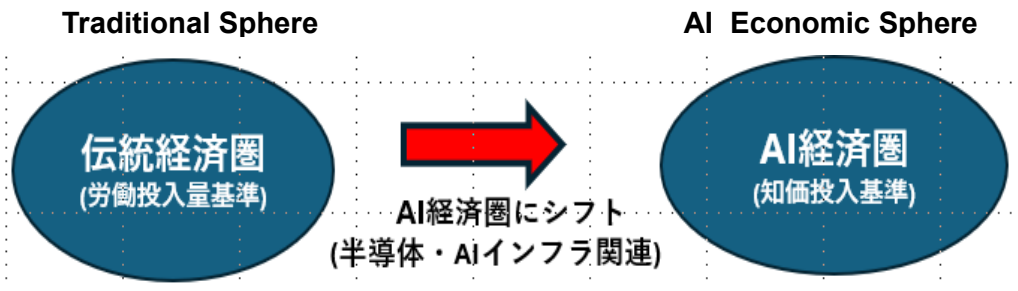
Figure4: AI Price per tokens（WSJ9/8/25）



(3) The AI-Driven Price Revolution Is Underway… Deflationary Pressure and Price Spikes

Astounding productivity gains are triggering a price revolution. The first phase of this revolution is a deflationary one. Intellectual tasks—such as research, design, translation, writing, and analysis—that were traditionally performed by humans are being rapidly automated or semi-automated by AI. Since AI has virtually negligible variable costs, the marginal costs for companies will eventually become minimal. Eventually, AI will replace the majority of intellectual labor. This will trigger a chain reaction of price declines—including lower product design and development costs (for both software and hardware), reduced corporate fixed costs (labor and outsourcing expenses), and lower service unit prices (for translation, design, consulting, etc.)—leading to general deflation and large-scale job cuts. Therefore, proactive economic policies that alleviate unemployment and reduce deflationary pressures by stimulating demand will be essential.

Figure5: AI Price Revolution and Polarrization of Economy



Source: Musha Research

At the same time, however, AI is causing rapid price increases in certain localized areas. In AI-driven economic sectors with exceptionally high productivity, prices for goods such as semiconductors—where demand has surged—are skyrocketing, leading to shortages and drastic shifts in opportunity costs. If AI-specific memory generates far greater profits than gaming memory—even within the semiconductor sector, this price surge will result in significant lost opportunities in traditional

出所: 武者リサーチ

semiconductor markets. The price surge is not limited to semiconductors and computing resources (such as GPUs); it is spreading to data centers, AI-skilled personnel, and other areas.

This polarization of prices will trigger a polarization of economic spheres. A split is likely to occur between traditional sectors where prices are falling and AI-related sectors where prices are rising, leading to a process of repricing. While long-term trends are difficult to predict, for the time being, market differentiation is likely to continue between the traditional economic sphere—where deflationary pressures are intensifying—and the AI economic sphere, where price increases have become the norm.

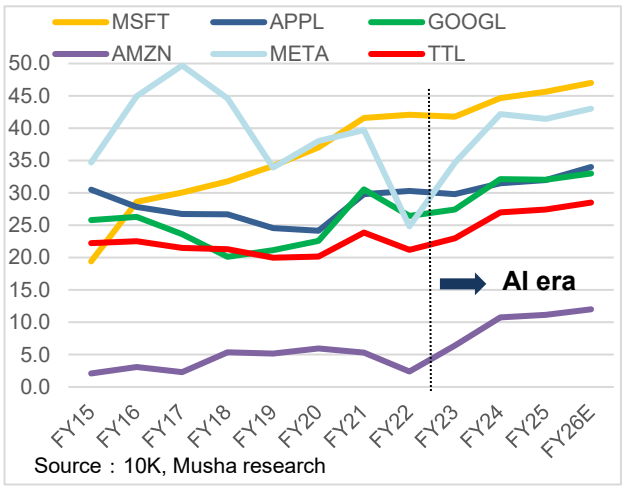
(4) A Bipolar Economic Landscape: The AI Economic Sphere Displaces the Traditional Economic Sphere

The economy is divided into traditional economic sphere A, which relies on labor input, and AI economic sphere B, which features significantly higher productivity. Even for the same product, economic utility increases when it is deployed in Economy B.

The sharp rise in profit margins among semiconductor manufacturers (the "NVIDIA phenomenon") can also be attributed to the shift from the traditional economy to the AI economy. In other words, even for the same product, becoming part of the AI economy triggered a profit margin revolution. This phenomenon is now spreading across the board.

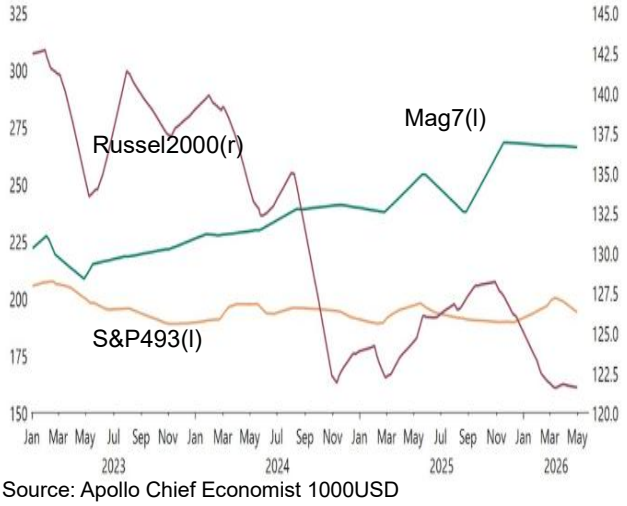
Looking back, a pricing revolution (i.e., price increases resulting from the shift to the AI economy, leading to improved margins) had also occurred within GAFAM.
GAFAM's profit margins have clearly risen; the average operating profit margin for the five companies has increased from 21.2% in 2022 to 28–29% this year. In 2022, GAFAM saw smartphone demand peak, causing revenue growth to slow significantly, and it was thought that the sector had entered a period of maturity. Each company simultaneously embarked on large-scale restructuring. However, as shown in Figure 6, each company's profit margins have risen sharply since bottoming out in 2022. It is believed that the shift toward AI—which began with the announcement of ChatGPT at the end of 2022—has boosted these companies' margins. While AI adoption should have significantly reduced costs, prices did not fall; rather, they continued to rise. This is also reflected in the changes in revenue per employee reported by Apollo Chief Economist Torsten Slok (Chart 7). While the "Magnificent Seven"—companies rapidly advancing their AI shift—have seen a significant increase, the other SP 493 companies show a downward trend. The decline is even steeper for the smaller companies in the Russell 2000.

Figur6: OP Margine of GAFAM



Source：10K, Musha research

Figure7: Per capita Rev. (Mag7, S&P493, Russel2000)



Source: Apollo Chief Economist 1000USD

The AI price revolution signifies a shift in the source of value from labor to knowledge (or "divine intelligence"), and could even be considered a historic revolution on par with the Agricultural and Industrial Revolutions. This may mark a moment when the conventional wisdom of existing economics—which views labor input as the source of value—is fundamentally rewritten.

(5) The AI Revolution Is Still in Its Early, Dawn Stage

Broadly speaking, the AI ecosystem consists of the following three layers, plus data accumulation. The first layer is the infrastructure layer, which serves as the foundation for running AI. It consists of computing resources—such as semiconductors, data centers, and power and cooling facilities—as well as the cloud platform layer, which aggregates data and transforms the infrastructure into a service. (e.g., AWS, Microsoft Azure, Google Cloud) and AI development platforms (training and inference environments), which provide the foundation for making AI "ready for use."

The second layer is the model layer, which can be described as the "brain" of AI. Based on large language models (LLMs), companies such as OpenAI, Anthropic, and Google DeepMind are developing core models. The third layer is the application layer, which serves as the interface with users. This includes ChatGPT, Copilot, Google Gemini, as well as search, AI image generation, business automation AI, AI for healthcare, finance, and education, and physical AI used in industry. These constitute the edge portion of the AI ecosystem and serve as the gateway through which AI is implemented in society. Among the GAFAM companies, Apple has made virtually no massive investments in infrastructure; this is believed to be based on a strategy of specializing in edge AI, leveraging its dominance in smartphones and PCs—which serve as AI gateways.

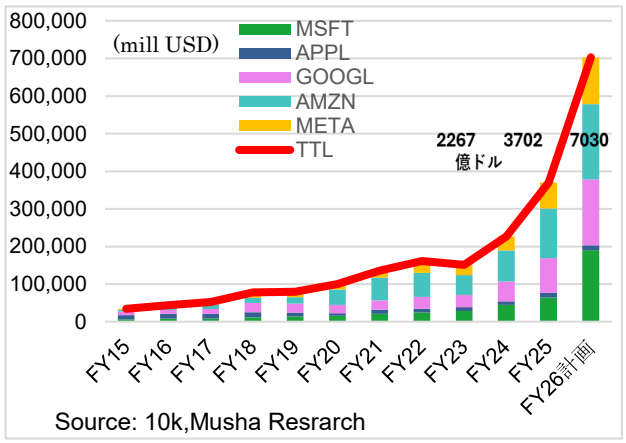
As a fourth component, vast amounts of data circulate at ultra-high speeds within the ecosystem. This data is the driving force of the AI industry—a resource equivalent to "oil" in the traditional economy—and can be viewed as a cross-layered resource that underpins the entire ecosystem.

Currently, the benefits of the AI economy are concentrated in hardware—specifically, infrastructure construction (semiconductors and data center assets) where large-scale investment is taking place—which is driving the semiconductor boom. However, the profit structure of the AI economy is expected to expand to include AI models (core models, applications), infrastructure, and edge models (user interfaces).

(6) AI Investment Surges to 100 Trillion Yen by GAFAM Companies, but It's Not a Bubble

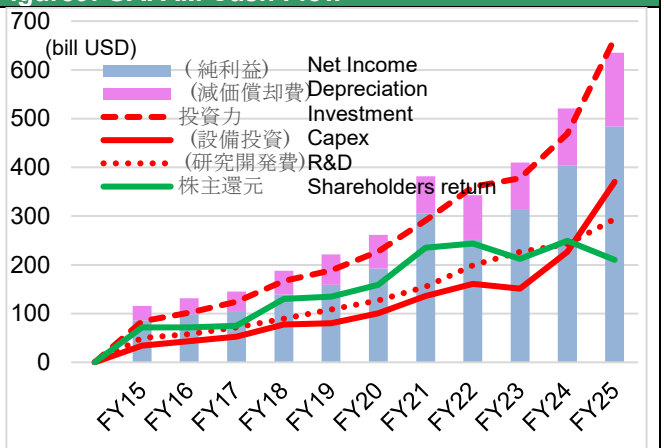
Capital expenditures by GAFAM are surging. From a previous year-over-year growth rate of 20–30%, they are projected to reach $226.7 billion in 2024 (a 50% increase), $370.2 billion in 2025 (a 63% increase), and $703–735 billion in 2026 according to company forecasts (nearly doubling). This is due to a dramatic surge in AI-related investments, particularly in data centers.

Figure8: GAFAM Capex



Source: 10k,Musha Resrarch

Figure9: GAFAM Cash Flow



Source: 10K, Musha Research

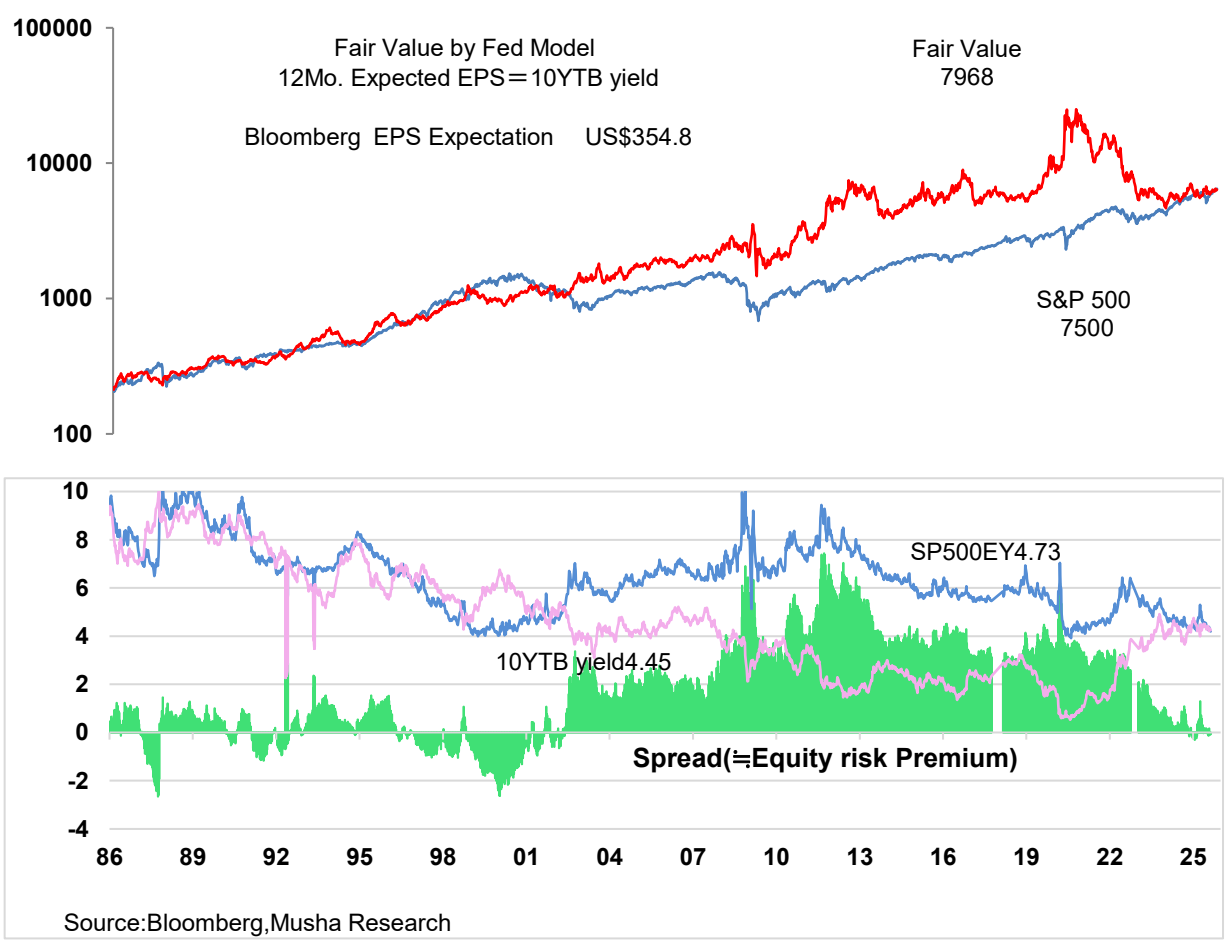
However, even with these massive investment figures, they remain within the scope of cash flow, demonstrating just how much the profitability of each GAFAM company has improved
In addition, massive investment plans—such as the "Stargate" project led by Oracle and SoftBank, among others—are underway outside of GAFAM, driving up the U.S. economic growth rate. The U.S. GDP growth rate for 2026 is projected to be around 2.0%. Within this context, capital expenditures—

driven primarily by AI-related investments—are surging, and AI and data center investments are expected to reach a scale of more than 2% of GDP. Since this figure is set to double (even if a significant portion is financed through imports), AI investment alone is estimated to be boosting the GDP growth rate by 0.5–1%. If we factor in productivity gains in traditional sectors and new services resulting from AI implementation, it can be inferred that the actual contribution to the economy is even greater.

China, however, lags significantly behind in terms of large-scale investment, estimated to be about one-tenth that of the United States. There is a large gap in profitability between Chinese internet platform companies (Tencent, Baidu, Alibaba, etc.) and GAFAM. Because China lags behind in terms of capital and hardware, it has no choice but to compete using "low-cost, small-scale models" based on algorithmic innovations like DeepSeek. However, to reach the ultimate frontier (AI in the realm of the divine), scaling enabled by the U.S.'s overwhelming infrastructure (massive investment) remains indispensable, and in the long term, U.S. dominance will remain unshaken.

What makes this high level of investment possible is not only high profitability but also high stock prices. While some have pointed to a U.S. stock market bubble in light of the high stock prices, this is likely an exaggeration. Figure 9 shows the spread between the 10-year Treasury yield and the stock earnings yield—the most basic valuation metric. Despite rising interest rates, this spread remains near zero, indicating that the market is within fair-value territory. Furthermore, the P/E ratio for the "Mag7" is just under 26x; considering a 20% pace of earnings growth, this is by no means overvalued.

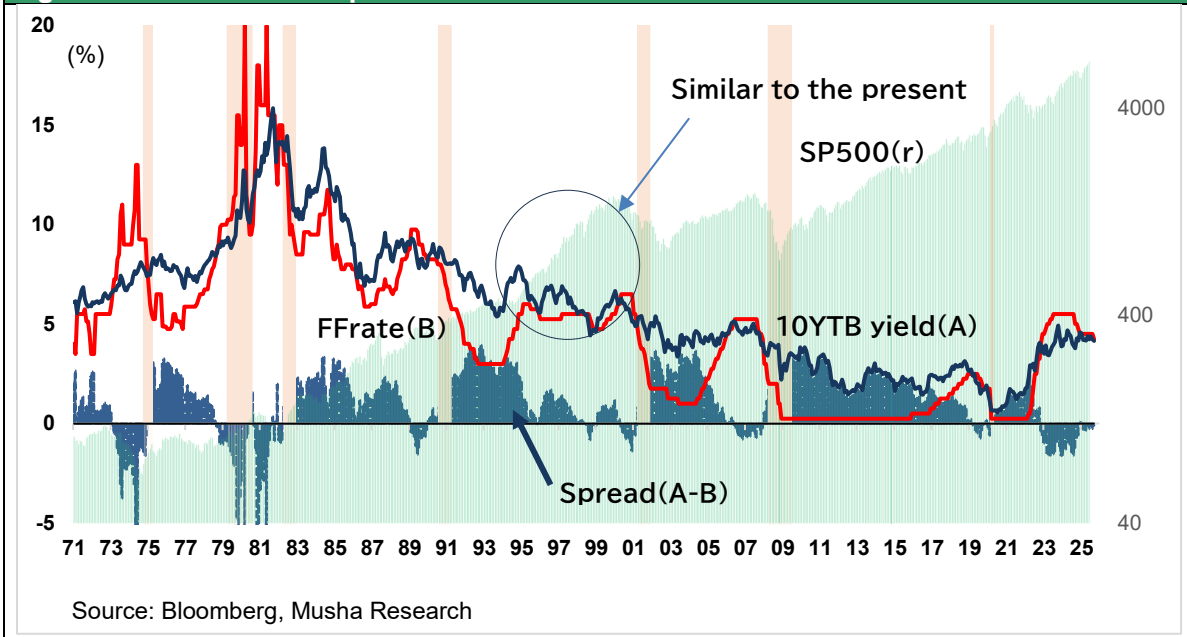
Fegure10: SP500 Fair Value Model (Fed Model) (latest 2026 May 30)



The current environment resembles that of 1995–96. While there is a possibility that the market could eventually surge into a bubble, that is not the case at present. The financial environment is similar to that period in the following respects: 1) investment is booming due to the technological revolution, and people's "animal spirits" are running high; 2) interest rates remain elevated and the yield curve is flattening; and 3) stock prices are rising. In late 1996, Chairman Greenspan cautioned against the stock

market boom, calling it "irrational exuberance," and maintained high short-term interest rates. However, long-term interest rates actually declined, flattening the yield curve, and stock prices continued to rise for more than three years thereafter. Comparing the dot-com bubble of 2000 with the current situation, the current market is far less overvalued in terms of valuations. Furthermore, the magnitude of the technological revolution is on a completely different scale between the Internet and the AI revolution. Even if a bubble does eventually form, it would be fair to say that we are still only at the very beginning.

Figure11: Interest rates Spread and S&P500



Source: Bloomberg, Musha Research

If one is concerned about a bubble bursting, it is essential to analyze monetary policy. The dot-com bubble was triggered by then-Federal Reserve Chairman Greenspan's persistent interest rate hikes. In contrast, interest rate hikes are currently being restrained. Inflation driven by crude oil prices will eventually subside. The current Trump administration's economic policies—including 1) deregulation, 2) channeling massive global capital into AI investment, and 3) promoting the creation of effective demand through expansionary fiscal and monetary policies—exhibit a strong bias toward accelerating stock prices.

If circumstances were to arise that would justify the bubble theory, it would likely be due to changes in underlying conditions such as: 1) a significant breakdown in AI scaling laws; 2) a halt in investment due to energy and power constraints; 3) a sharp rise in interest rates (leading to a contraction in risk assets); 4) regulatory-imposed restrictions on AI infrastructure investment; or 5) geopolitical supply chain disruptions. In other words, this would occur if the analyses in sections (2) through (5) of this report were to lose their validity; however, the likelihood of such an outcome is deemed low.